HAL
open science

Analyse morphologique et logarithmique de grandes banques d'images

Guillaume Noyel

▶ **To cite this version:**

Guillaume Noyel. Analyse morphologique et logarithmique de grandes banques d'images. Vision par ordinateur et reconnaissance de formes [cs.CV]. Université de Reims Champagne-Ardenne, 2021. tel-03343079

HAL Id: tel-03343079

https://theses.hal.science/tel-03343079

Submitted on 13 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANALYSE MORPHOLOGIQUE ET LOGARITHMIQUE DE GRANDES BANQUES D'IMAGES

MÉMOIRE D'HABILITATION

présenté à
l'Université de Reims Champagne-Ardenne
par

Guillaume NOYEL

le 25 juin 2021

pour l'obtention de

L'HABILITATION
à
DIRIGER DES RECHERCHES

Garant d'Habilitation : Nicolas PASSAT

Jury :

Isabelle	BLOCH	*Rapporteur*
Thierry	FOURNEL	*Rapporteur*
Hugues	TALBOT	*Rapporteur*
Dominique	JEULIN	*Examinateur*
Michel	JOURLIN	*Examinateur*
Laurent	KODJIKIAN	*Examinateur*
Nicolas	PASSAT	*Examinateur*
Chris	ROBERTSON	*Examinateur*

Résumé

Avec l'utilisation de plus en plus massive de la photographie numérique, de grandes banques de données d'images ont été constituées durant ces dernières décennies tant dans l'industrie que pour la médecine. Cependant, les images ont généralement été acquises avec différentes conditions d'éclairement, poses de l'appareil photographique et types d'appareil. Ceci cause des difficultés de comparaison entre les images d'une même scène. J'ai proposé des méthodes pour l'analyse de ces grandes banques de données qui répondent à cette problématique. Elles ont été appliquées dans le domaine des pneumatiques pour le contrôle de leur aspect visuel et l'étude de leur performances ainsi que dans celui de la santé pour le diagnostic et le suivi de la rétinopathie diabétique. Ces méthodes appartiennent aux cadres de travail de la Morphologie Mathématique et du Logarithmic Image Processing (LIP). Grâce à des collaborations académiques, j'ai développé des approches de segmentation morphologiques et de classification par apprentissage statistique. Elles ont été utilisées pour l'analyse d'images texturées et tridimensionnelles de la surface des pneumatiques, à des fins de recherche de défauts. Elles ont aussi servi à l'analyse automatique des images du fond d'œil chez les patients atteints de diabète. Du fait de la grande quantité de données à analyser, j'ai amélioré des algorithmes tant au niveau de leur complexité qu'au niveau de leur programmation efficace. Le programme peut être exécuté en parallèle sur plusieurs cœurs de processeurs et écrit de manière vectorielle (i.e. plusieurs nombres traités à la fois au lieu d'un seul). Grâce au modèle LIP, j'ai proposé des méthodes d'amélioration des contrastes couleur des images rétiniennes. J'ai étudié les propriétés des métriques fonctionnelles d'Asplund définies soit avec la loi LIP-additive, soit avec la loi LIP-multiplicative. La première métrique, qui est LIP-additive, est robuste aux variations d'éclairement causées par un changement du temps d'exposition de la caméra ou de l'intensité de la source. La seconde, qui est LIP-multiplicative, est robuste aux variations d'éclairement dues à un changement d'opacité de l'objet capturé. Ces métriques sont utiles à la reconnaissance des formes, grâce aux cartes de distances entre une fonction de référence et une image. J'ai établi le lien entre ces cartes de distances d'Asplund et la Morphologie Mathématique. Puis, j'ai créé le nouveau cadre de la Morphologie Mathématique Logarithmique qui consiste à définir les opérations fondamentales d'érosion et de dilatation avec la loi LIP-additive. Ceci leur donne la propriété intéressante d'être adaptatives aux variations d'éclairement causées par un changement du temps d'exposition de la caméra. D'autres opérateurs, robustes à ces variations, ont été définis dans ce nouveau cadre de travail. Celui-ci et le modèle LIP ont permis l'introduction de méthodes de segmentation robustes aux variations d'éclairement. Des techniques de percolations ont aussi été étudiées avec des contrastes logarithmiques couleur. En outre, j'ai conçu et réalisé un système d'acquisitions tridimensionnelles et multimodales par stéréovision dans le domaine de la lumière visible et dans celui des rayons X pour l'imagerie des pneumatiques. La chaîne complète comprenant l'acquisition stéréoscopique, le calibrage des systèmes d'acquisition et la reconstruction 3D a été étudiée et validée. De plus, afin de comparer les images de coupes radiales de pneumatiques avec leurs épures (i.e. leurs plans 2D), une méthode de recalage déformable à partir de B-splines cubiques a été testée avec succès. Pour l'analyse longitudinale (i.e. temporelle) des images du fond d'œil, j'ai introduit un nouveau modèle de superposition composé d'une homographie affine (i.e. une rotation, une translation et une mise à l'échelle anisotrope) et d'une ou deux corrections de distorsions radiales selon le nombre de caméras employées lors de l'acquisition. Comparée à plusieurs méthodes de l'état de l'art, sur des séries de paires d'images provenant de banques de données de santé publique acquises avec une année d'intervalle, l'approche proposée donne de meilleurs résultats. Ces travaux de recherche constituent ma contribution à l'analyse des grandes banques d'images dans l'industrie ou en médecine.

Mots clés : traitement massif d'image, grandes banques d'images, morphologie mathématique, logarithmic image processing, apprentissage statistique, apprentissage profond, intelligence artificielle, analyse de texture, contrôle industriel, accélération d'algorithmes, parallélisation, calcul vectoriel, amélioration des contrastes, couleur, métriques fonctionnelles d'Asplund, morphologie mathématique logarithmique, robustesse aux variations d'éclairement, images à faibles contrastes, homogénéité de région, contrastes perceptuels, segmentation, percolation, stéréovision, reconstruction 3D, acquisition multimodale, calibrage, radiographie, recalage d'image, recalage déformable d'image, B-splines, homographie affine, distorsion radiale, industrie, rétinopathie diabétique, segmentation vasculaire, rétine, fond d'œil, santé publique, imagerie médicale.

Abstract

With the massive use of digital photography, large databases of images have been built during the last decades, both for industry and medicine. However, the images were generally captured with different lighting conditions, poses of the camera and types of camera. This causes difficulties when comparing images of the same scene. I have proposed methods to analyse these large databases which address these issues. They have been applied to the field of tyres for the control of their visual aspect and the study of their performances as well as to the field of healthcare for the diagnosis and the follow-up of diabetic retinopathy. These methods fall into the framework of Mathematical Morphology and Logarithmic Image processing (LIP). In particular, thanks to partnerships with academia, I have developed approaches of morphological segmentation and classification by machine learning. They have been used for the analysis of images of tyre surfaces (which can be textured or in three dimensions) in order to look for defaults. They have also served for the automatic analysis of eye fundus images from patients with diabetes. Due to the large amount of data to analyse, I have speeded up algorithms with regards to their complexity as well as their efficient programming. The program can be run in parallel over several processor cores or written in a vector way (i.e., several numbers processed at the same time in place of a single one). Thanks to the LIP model, I have proposed methods to improve colour contrast in retinal images. I have studied the properties of these functional metrics which have been defined either with the LIP-additive law, or with the LIP-multiplicative law. The first metric, which is LIP-additive, is robust to lighting variations caused by a change in the exposure time of the camera or in the source intensity. The second one, which is LIP-multiplicative, is robust to lighting variations due to changes of opacity of the captured object. These metrics are useful for pattern matching, thanks to distance maps between a reference function and an image. I have established the link between these maps of Asplund distances and Mathematical Morphology. Then, I have created the new framework of Logarithmic Mathematical Morphology which is based on the fundamental operations of erosion and dilation with the LIP-additive law. This gives them the interesting property of being adaptive to lighting variations caused by a change of camera exposure-time. Other operators, robust to these variations, have been defined in this novel framework. The latter and the LIP model have allowed the introduction of segmentation methods robust to lighting variations. Percolation techniques have also been studied with logarithmic colour contrasts. In addition, I have conceived and made a three-dimensional and multimodal acquisition system by stereovision, in the domain of the visible light and in the one of X-rays for tyre imaging. The complete chain has been studied and validated, including stereoscopic acquisition, calibration of the acquisition system and 3D reconstruction. In addition, in order to compare images of radial cuts of tyres with their 2D plan, a deformable registration method based on cubic B-splines has been successfully tested. For the longitudinal (i.e., temporal) analysis of eye fundus images, I have introduced a superimposition model made of an affine homography (i.e., a rotation, a translation and an anisotropic scaling) and one or two corrections of radial distortions depending on the number of cameras used for the acquisition. Compared to other state-of-the-art methods, in series of image pairs from public health databases captured with an interval of a year, the proposed approach gives better results. These research works form my contribution to the analysis of large image databases in industry or in medicine.

Keywords: massive image processing, large image databases, mathematical morphology, logarithmic image processing, machine learning, deep learning, artificial intelligence, texture analysis, industrial control, speed up of algorithms, parallelisation, vector processing, contrast enhancement,

colour, functional Asplund metrics, logarithmic mathematical morphology, robustness to lighting variations, low contrast images, region homogeneity, perceptual contrasts, segmentation, percolation, stereovision, 3D reconstruction, multimodal acquisition, calibration, radiography, image registration, deformable image registration, B-splines, affine homography, radial distortion, industry, diabetic retinopathy, vessel segmentation, retina, eye fundus, public health, medical imaging.

Remerciements

L'écriture de ce mémoire d'Habilitation à Diriger des Recherches m'a offert la possibilité de faire un point sur mes travaux douze années après ma thèse. Ceux-ci n'auraient pu être accomplis sans l'aide et la collaboration de plusieurs personnes à qui je tiens, tout particulièrement, à signifier ma gratitude.

Le Professeur Nicolas Passat, de l'Université de Reims Champagne-Ardenne, qui a été le garant de ce travail. Les nombreux échanges que nous avons eus et ses relectures très attentives ont permis la mise en perspective de mon travail de recherche et la rédaction de ce mémoire.

Les rapporteurs du jury, les Professeurs Isabelle Bloch, de Télécom Paris, Thierry Fournel, de l'Université Jean Monnet à Saint-Etienne, et Hugues Talbot, de CentraleSupélec, pour avoir accepté d'examiner ce mémoire et mes travaux de recherche.

Le Professeur Michel Jourlin, de l'Université Jean Monnet, avec qui j'ai travaillé à de nombreuses reprises autour du *Logarithmic Image Processing* et qui toujours été de bons conseils.

Le Professeur Dominique Jeulin, du Centre de Morphologie Mathématique de l'Ecole des Mines de Paris, avec qui j'ai collaboré lorsque je travaillais au Centre de Recherche de Michelin et que j'ai revu dans plusieurs colloques.

Le Professeur Chris Robertson, du département de Mathématiques et Statistiques de *University of Strathclyde*, à Glasgow en Ecosse au Royaume-Uni, qui m'a accueilli comme chercheur visiteur dans son équipe et avec qui j'ai pu travailler sur plusieurs projets.

Le Professeur Laurent Kodjikian, ophtalmologiste aux Hospices Civils de Lyon et à l'Université Claude Bernard Lyon 1, qui m'a apporté son soutien et donné accès aux images du fond d'œil de son service. Au sein de celui-ci, je tiens en particulier à remercier le Docteur Christine Vartin pour son avis et ses relectures des images du fond d'œil.

Le Professeur Peter Boyle, de l'*International Prevention Research Institute* (IPRI), qui m'a permis de monter l'axe de recherche en traitement d'image dans son institut en tant que *Research Director*.

Le Professeur David Littlejohn de *University of Strathclyde*, *Executive Dean of Science and Associate Principal* qui m'a permis de devenir chercheur visiteur dans son université.

Le Professeur David Owens, diabétologue à Swansea University au Pays de Galles, Royaume-Uni, qui m'a expliqué les enjeux de la rétinopathie diabétique et donné accès au centre national de dépistage : le *Diabetic Eye Screening Wales* de Cardiff. Parmi les membres de cette équipe, je tiens à remercier le Docteur Rebecca Thomas, Simon Iles, Gavin Bhakta et Andrew Crowder.

Le Professeur Alberto d'Onofrio, mon collègue à l'*International Prevention Research Institute* et à *University of Strathclyde*, avec qui j'ai organisé le montage de plusieurs projets européens et avec qui j'ai apprécié de travailler.

Le Docteur Patrice Hauret, du Centre de Recherche de Michelin, avec qui j'ai échangé sur les projets de recherche industriels.

Les stagiaires et le jeune chercheur que j'ai encadrés.

Confidentialité

Certaines images et sections de ce mémoire n'ont pas été rendues publiques pour des questions de confidentialité. Les images confidentielles apparaissent comme un rectangle gris avec le mot « Image ». Les sections confidentielles ne sont pas montrées.

Table des matières

1 Introduction **1**
1.1 Ma carrière scientifique 1
1.2 Ma démarche de recherche : des applications à la théorie 2
1.3 Objectif et plan de ce mémoire 3

2 Contexte : l'analyse d'images pour l'industrie ou la médecine **5**
2.1 Recherche d'anomalies dans des banques d'images pour le contrôle industriel . . 5
2.2 L'analyse d'image pour l'étude des performances des pneumatiques 8
2.3 L'analyse de banques d'images pour le diagnostic de la rétinopathie diabétique . 15
2.4 Conclusion du chapitre 19

3 Segmentation morphologique et apprentissage statistique dans les grandes banques d'images **21**
3.1 Fondements de la Morphologie Mathématique 21
3.1.1 Morphologie Mathématique pour les ensembles 21
3.1.2 Morphologie Mathématique pour les fonctions 23
3.1.3 Filtrage morphologique 26
3.1.4 Géodésie 27
3.1.5 Ligne de Partage des Eaux 29
3.2 Segmentation morphologique d'objets en relief à la surface des pneumatiques . . 30
3.2.1 Détection des motifs sur la partie externe des pneumatiques 30
3.2.2 Détection des motifs sur la partie interne des pneumatiques 36
3.3 L'analyse de texture pour la recherche d'anomalies en contrôle industriel 42
3.3.1 Méthode initiale d'analyse de texture 42
3.3.2 Méthode améliorée d'analyse de texture 46
3.4 L'accélération d'algorithmes 52
3.4.1 Segmentation rapide d'anomalies pour le contrôle industriel 52
3.4.2 Algorithmes rapides en Morphologie Mathématique 61
3.5 Classification d'images du fond d'œil 69
3.5.1 Evaluation de la qualité des images 69
3.5.2 Classification automatique des images du fond d'œil 75
3.6 Conclusion du chapitre 86

4 Traitement d'Image Logarithmique et métriques d'Asplund pour l'analyse des images à faible contraste **89**
4.1 Traitement d'Image Logarithmique 89
4.1.1 Traitement d'Image Logarithmique pour les images à niveaux de gris . . 89
4.1.2 Traitement d'Image Logarithmique pour les images couleur 92
4.2 Amélioration des contrastes 96
4.2.1 Application aux images du fond d'œil 96
4.2.2 Amélioration du contraste global des images à niveaux de gris 98
4.2.3 Amélioration du contraste global des images couleur 98
4.2.4 Amélioration du contraste local des images à niveaux de gris 101
4.2.5 Amélioration du contraste local des images couleur 101

4.3 Les métriques d'Asplund . . . 102
4.3.1 Travaux connexes . . . 102
4.3.2 Les métriques d'Asplund à niveaux de gris . . . 103
4.3.3 Analyse de forme avec la métrique d'Asplund LIP-multiplicative . . . 105
4.3.4 Analyse de forme avec la métrique LIP-additive de type Asplund . . . 111
4.3.5 Expériences et résultats . . . 115
4.3.6 Lien entre les deux métriques d'Asplund fonctionnelles LIP-multiplicative et LIP-additive . . . 121
4.3.7 Conclusion sur les métriques d'Asplund fonctionnelles . . . 122
4.3.8 Les métriques d'Asplund pour les images couleur et multivariées . . . 123
4.3.9 Application aux images rayons X des pneumatiques . . . 132
4.4 Morphologie Mathématique Logarithmique . . . 136
4.4.1 Motivation . . . 136
4.4.2 Travaux connexes : opérateurs de détection robustes aux variations d'éclairement . . . 137
4.4.3 Les opérateurs de Morphologie Mathématique Logarithmique . . . 139
4.4.4 Opérateurs robustes aux variations d'éclairement . . . 144
4.4.5 Illustrations . . . 151
4.4.6 Perspectives . . . 152
4.5 Segmentation robuste aux variations d'éclairement . . . 153
4.5.1 Segmentation des micro-anévrismes et des vaisseaux du fond d'œil . . . 153
4.5.2 Critères d'homogénéité de région . . . 163
4.6 Percolation . . . 172
4.6.1 Percolation robuste aux variations d'intensité lumineuse . . . 172
4.6.2 Percolation pour l'analyse des images de pneumatique . . . 181
4.7 Conclusion du chapitre . . . 185

5 Stéréovision pour l'analyse des performances des pneumatiques 187
5.1 Système de stéréovision par photographie . . . 187
5.1.1 Modèle de transformation . . . 187
5.1.2 Estimation des paramètres du modèle . . . 190
5.1.3 Calibrage d'un système de stéréovision par appareil photographique . . . 193
5.1.4 Reconstruction 3D . . . 195
5.2 Système de stéréovision par radiographie . . . 200
5.3 Conclusion du chapitre . . . 200

6 Recalage d'images pour les pneumatiques et pour le fond d'œil 201
6.1 Recalage d'images pour les coupes de pneumatiques . . . 201
6.2 Recalage d'images du fond d'œil dans les grandes banques de données de santé publique . . . 201
6.2.1 Ajustement d'un nouveau modèle de recalage sur les points SIFT . . . 201
6.2.2 Ajustement d'un modèle de recalage sur les vaisseaux . . . 222
6.3 Conclusion du chapitre . . . 228

7 Conclusion et perspectives de recherche 231
7.1 Conclusion . . . 231
7.2 Perspectives . . . 233
7.2.1 Méthodes morphologiques et logarithmiques . . . 233
7.2.2 Morphologie Mathématique et apprentissage statistique . . . 235

Annexe : Enseignement et encadrement de projets de recherche 237
1 Enseignement . . . 237

2 Encadrements de projets de recherche, bourses et autres activités de recherche . . 238
2.1 Projets de recherche ayant donné lieu à des transferts 238
2.2 Financements pour les projets 239
2.3 Encadrement de jeunes chercheurs et d'étudiants 240
2.4 Membre du comité de sélection du prix de la meilleure thèse lors d'une conférence internationale 241
2.5 Prix 241
2.6 Relecteur dans des revues à comité de lecture 241
2.7 Relecteur pour des conférences internationales 241
2.8 Membre du comité de programme de conférences internationales 242
2.9 Président de session dans des conférences 242
2.10 Membre de comités techniques 242
2.11 Membre de sociétés scientifiques 242
2.12 Développement de logiciels 242
3 Responsabilités collectives et activités administratives 244

Liste de publications **245**
1 Revues internationales à comité de lecture 245
2 Chapitre d'ouvrage 245
3 Brevets 246
4 Conférences avec actes 247
5 Conférences avec résumé 248
6 Conférences sans actes 249
7 Thèse 249
8 Conférences nationales 249

Bibliographie **251**

Chapitre 1

Introduction

Ce mémoire qui présente mes travaux de recherche depuis mon doctorat a été rédigé sous la direction du Professeur Nicolas Passat de l'Université de Reims Champagne-Ardenne. Du fait de l'enregistrement systématique de nombreuses photographies numériques depuis au moins deux décennies dans de grandes banques de données tant dans le domaine industriel que dans celui de la médecine, il est désormais possible de les analyser de manière massive afin de pouvoir les comparer. Cependant, comme les images n'ont souvent pas été acquises avec le même appareil photographique, ni avec les mêmes conditions d'éclairement et la même pose, une telle comparaison ne peut être réalisée directement. En effet, deux images de la même scène peuvent présenter différents contrastes, orientations, résolutions ou même des distorsions différentes créées par les lentilles des caméras. Ces problèmes ne sont souvent pas clairement pris en compte lors de l'analyse des images. Mon parcours de chercheur m'ayant conduit à analyser ces grandes banques, tant dans l'industrie que dans le monde académique, j'ai dû développer des méthodes qui s'adaptent aux changements entre les images tout en permettant un traitement suffisamment rapide des données. Pour ces raisons, j'ai proposé des méthodes de traitement dans le cadre bien défini de la Morphologie Mathématique [275], [277] et dans celui du Traitement d'Image Logarithmique [143], [144] qui est bien adapté aux images acquises sous faible éclairement. En outre, j'ai étudié plusieurs méthodes de classification par apprentissage statistique : (i) pour l'analyse des textures présentes à la surface des pneumatiques ou (ii) en santé publique, dans les banques d'images de la rétine utilisées pour le suivi de la rétinopathie diabétique. J'ai étudié et créé de nouveaux modèles de recalage d'images afin de faciliter leur comparaison. En outre, j'ai réalisé un système de reconstruction 3D par stéréovision dans le domaine de la lumière visible et dans celui des rayons X pour l'étude des pneumatiques. L'introduction est organisée comme il s'ensuit. Après un rappel de mon parcours, j'expliquerai ma démarche de recherche puis j'exposerai l'objectif et le plan de ce mémoire.

1.1 Ma carrière scientifique

Depuis de nombreuses années, mon parcours s'est orienté autour de la recherche en analyse et traitement d'images. De 2002 à 2005, grâce aux cours du Professeur Michel Jourlin, donnés à CPE-Lyon, j'ai découvert les domaines particulièrement intéressants que sont la Morphologie Mathématique [113], [176], [177], [201], [275], [277], [284], le Traitement d'Image Logarithmique (en anglais *Logarithmic Image Processing* – LIP) [143], [144], [150] et leurs nombreuses applications pratiques. En 2004-2005, lors de mon Master Recherche à l'Université Jean Monnet de Saint-Etienne, le Professeur Dominique Jeulin m'a fait connaître l'étude et la modélisation des structures aléatoires par des méthodes morphologiques. Durant mes stages d'élève ingénieur, j'ai pu découvrir le monde de la recherche en image : (i) en 2004, au *Utah Center for Advanced Imaging Research* (UCAIR) du Département de Radiologie de *University of Utah* à Salt Lake City, Etats-Unis, sous la direction du Professeur Larry Zeng et (ii) en 2005, au Laboratoire de Systèmes Périphériques de l'Ecole Polytechnique Fédérale de Lausanne en Suisse, sous la direction du Professeur Roger Hersch.

En 2005, le Professeur Dominique Jeulin et le Dr. Jesus Angulo m'ont recruté pour préparer un doctorat à l'Ecole des Mines de Paris au Centre de Morphologie Mathématique. Ma thèse s'intitulait : « Filtrage, réduction de dimension, classification et segmentation morphologique hyper-

spectrale » [212]. Elle concernait non seulement les images hyperspectrales satellitaires mais aussi les séquences d'IRM dynamique (DCE-MRI – *Dynamic Contrast Enhanced Magnetic Resonance Imaging*).

En 2008, après ma thèse, j'ai choisi d'avoir une expérience industrielle significative comme ingénieur-chercheur au Centre Mondial de Recherche de Michelin, regroupant 3500 chercheurs, à Clermont-Ferrand ; ceci afin de comprendre en détail les problèmes applicatifs posés au traitement des images dans ce domaine. J'ai mené deux principaux projets de recherche. Le premier portait sur l'inspection automatique par analyse d'image des pneumatiques manufacturés. La conception et le développement d'un premier démonstrateur m'ont permis de détecter les objets en relief à la surface des pneumatiques. J'ai ensuite organisé une collaboration de recherche avec l'Ecole des Mines de Paris incluant plusieurs chercheurs de l'équipe du Professeur Dominique Jeulin et du Dr. Michel Bilodeau. L'objectif était la détection sur la surface interne des pneumatiques : des anomalies dans la texture et des anomalies tridimensionnelle (3D) de manière rapide. Le deuxième projet portait sur la compréhension des performances des pneumatiques, grâce à l'analyse des images de leur surface et de leurs composants internes. J'ai ainsi fait de la stéréovision multimodalités dans les domaines du visible et des rayons X. J'ai encadré plusieurs stagiaires ingénieurs sur la stéréovision, la détection de formes dans des radiographies grâce aux métriques fonctionnelles d'Asplund et le recalage d'une image de coupe de pneumatique avec son épure (i.e. son plan) numérique. J'ai été amené à collaborer avec le Professeur Michel Jourlin de l'Université Jean Monnet à Saint-Etienne et avec des Petites et Moyennes Entreprises (PME) innovantes dans le domaine de l'image.

En 2014, j'ai rejoint l'*International Prevention Research Institute* (IPRI), à Lyon, en tant que *Research Director*. L'IPRI était dirigé par le Professeur Peter Boyle, épidémiologiste de renommée mondiale. Dans cet institut, j'ai créé l'axe de recherche sur l'analyse d'images médicales et, en particulier, l'analyse des images du fond d'œil chez les patients atteints de diabète. J'ai ainsi été amené à collaborer avec le Professeur David Owens, diabétologue, à *Swansea University* au Pays de Galles, Royaume-Uni, puis avec le Professeur Laurent Kodjikian, ophtalmologiste à l'Hôpital de la Croix Rousse et à l'Université Claude Bernard, Lyon 1. J'ai notamment obtenu deux financements de la Banque Publique d'Investissement France et un autre du programme Horizon 2020 de l'Union Européenne pour le projet Eye-Light [1]. En 2017, je suis aussi devenu *Chercheur Visiteur* à *University of Strathclyde* à Glasgow en Ecosse au Royaume-Uni dans le Département de Mathématiques et Statistiques dirigé par le Professeur Chris Robertson. Cette collaboration m'a permis d'orienter ma recherche en traitement d'image vers les questions de santé publique. J'ai ainsi créé une méthode d'amélioration des images du fond d'œil afin de faciliter le diagnostic de la rétinopathie diabétique. Une étude clinique pour valider la méthode a aussi été préparée en collaboration avec le Professeur Robertson et l'équipe du Professeur Kodjikian.

1.2 Ma démarche de recherche : des applications à la théorie

Après ma thèse, j'ai volontairement choisi d'avoir une expérience significative de recherche dans les domaines industriels et médicaux avant de revenir à des fonctions plus académiques. Mes expériences, dans l'industrie chez Michelin et en lien avec le monde médical à l'IPRI et à *University of Strathclyde*, m'ont confronté aux applications pratiques de ces domaines. Dans le monde industriel ou médical, le chercheur se trouve de prime abord face à une question concrète posée par l'industriel ou le médecin. Celle-ci peut être destinée à améliorer un processus ou à mieux comprendre un phénomène médical. Il est alors nécessaire de bien interpréter la question concrète pour formuler une problématique de recherche permettant de la résoudre ou d'y apporter une contribution significative. La problématique ainsi formulée va permettre d'accroître le domaine de recherche.

1. `https://cordis.europa.eu/project/id/717108/fr`

Par exemple, chez Michelin, on souhaitait assister par ordinateur une tâche manuelle de contrôle des pneumatiques grâce à de l'analyse d'image. Le contrôle manuel portait sur plusieurs dizaines de critères de qualité. Dans un tel système, on cherche à avoir un taux de faux négatifs aux critères presque nul et un taux de faux positifs pas trop élevé. Le problème se complexifie lorsque l'on a des milliers de pneumatiques à contrôler, dans des gammes différentes et avec reliefs variés qui doivent être présents à leur surface. Nous avons ainsi été amenés à faire des acquisitions d'intensité d'un laser sur le pneumatique et des acquisitions tridimensionnelles de leur surface. Il a ensuite été nécessaire de segmenter les marquages obligatoirement présents à la surface des pneumatiques pour les différencier des reliefs anormaux. Sur la surface interne des pneumatiques, certains défauts apparaissaient comme des anomalies dans une texture. Il a donc été nécessaire de mettre en place une technique d'analyse de texture par apprentissage statistique.

Dans le domaine médical, lorsque l'on modifie une image du fond d'œil, il est important de mettre en évidence les lésions recherchées sans enlever de l'information, comme par exemple atténuer d'autres lésions. Le problème devient complexe lorsque des milliers d'images provenant de banques de santé publique doivent être analysées. On est alors amené à développer des outils de comparaison ou de segmentation qui soient robustes aux variations d'éclairement. J'ai ainsi proposé une méthode d'amélioration des contrastes du fond d'œil afin de rendre le diagnostic de la rétinopathie diabétique plus facile et plus rapide. J'ai aussi recalé les images de la rétine numérisées à environ un an d'intervalle. Dans ce but, une approche robuste aux changements de caméras, aux mouvements de l'œil et aux variations d'éclairement a été introduite. J'ai en outre développé une méthode d'extraction des vaisseaux robustes à ces variations. Ce parcours dans le monde industriel et médical, m'a permis : (i) d'acquérir une compréhension précise des problèmes applicatifs ; puis (ii) de formuler des problématiques théoriques sur ces applications et enfin (iii) de développer la théorie adéquate pour résoudre les problèmes applicatifs. Ma démarche de recherche a donc consisté à aller des applications à la théorie.

1.3 Objectif et plan de ce mémoire

L'objectif de ce mémoire est d'exposer (i) les applications et problématiques de recherche auxquelles j'ai été confronté dans mon parcours et (ii) la théorie développée pour répondre à ces problématiques. En particulier, dans le chapitre 2, le contexte de mon travail sera présenté. Dans le chapitre 3, je décrirai les méthodes de segmentation morphologique, d'analyse de texture et de classification par apprentissage que j'ai développées moi-même ou grâce à des collaborations entre l'industrie et le monde académique. Des applications industrielles et médicales les illustreront. Dans le chapitre 4, nous verrons les méthodes de Traitement d'Image Logarithmique et les métriques d'Asplund fonctionnelles pour analyser des images à faibles contrastes notamment sur le fond d'œil. Je présenterai le lien entre les métriques d'Asplund et la Morphologie Mathématique et j'introduirai le cadre de la Morphologie Mathématique Logarithmique. Dans le chapitre 5, je présenterai les méthodes de stéréovision développées pour l'analyse des pneumatiques. Dans le chapitre 6, j'exposerai des méthodes de recalage d'images pour le pneumatique et pour l'analyse temporelle de la rétinopathie diabétique dans les banques d'images de santé publique. Les annexes décriront mes activités d'enseignement, l'encadrement de mes projets de recherche, mes activités administratives et responsabilités collectives ainsi que ma liste de publications. Ce mémoire expose de manière détaillée mes activités de recherche depuis ma thèse car, du fait de mon parcours à la fois industriel et académique, je souhaitais montrer précisément ma contribution scientifique dans le domaine du traitement d'image.

Chapitre 2

Contexte : l'analyse d'images pour l'industrie ou la médecine

Depuis ma thèse, mon travail de recherche dans l'industrie ou pour la médecine a été fortement marqué par l'analyse de problèmes pratiques. Dans ce cas, on dispose d'images qui ont été préalablement acquises dans un cadre donné et que l'on doit analyser. Dans les grandes banques d'images existantes, il est souvent nécessaire de se contenter des images telles qu'elles sont.

Une question courante dans ces grandes banques est la suivante : comment comparer ces images entre elles ou à un ensemble de référence contenant des anomalies ou lésions à rechercher ? Cet ensemble de référence est généralement plus ou moins représentatif des objets recherchés et les images ont souvent été acquises dans des conditions différentes d'éclairement, d'appareil photographique, etc. Cette comparaison d'images pose donc des problèmes de variations de contrastes, de recalage et de comparaison à un ensemble de référence.

Nous allons illustrer ces problématiques par des exemples.

2.1 Recherche d'anomalies dans des banques d'images pour le contrôle industriel

Entre 2008 et 2010, au Centre de Recherche de Michelin, j'ai travaillé sur un projet de recherche pour le contrôle des pneumatiques par inspection visuelle. Ce contrôle s'effectuait de manière manuelle et visuelle par des opérateurs spécialement formés qui contrôlaient scrupuleusement plusieurs dizaines de critères de qualité à une cadence donnée pour chaque pneumatique. Les pneumatiques étant composés de très nombreux constituants (Fig. 2.1) tels que de la gomme mais aussi des tringles et des nappes en fils métalliques ou nylons, et comme tous les pneus sont contrôlés, cela représente un travail conséquent.

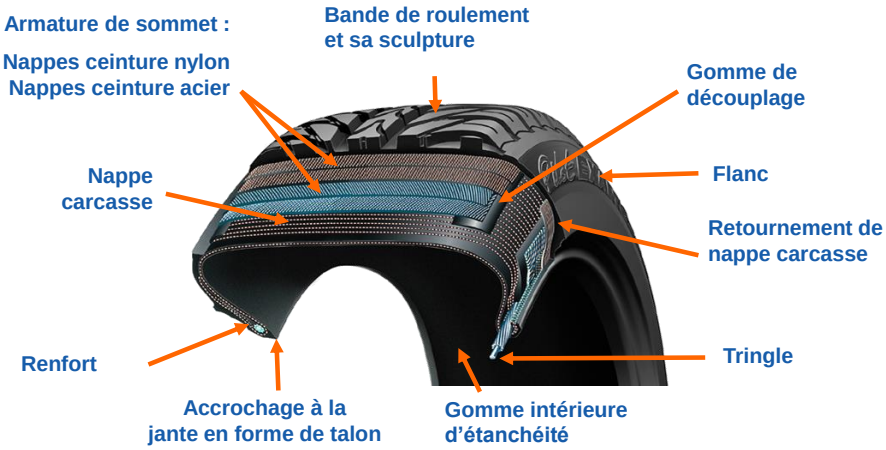


FIG. 2.1 – L'architecture et les composants d'un pneumatique.

L'objectif était de déterminer si ce contrôle manuel pouvait être assisté par une analyse d'image automatique. Pour cette étude, une image en relief tridimensionnel de la surface des pneumatiques, dite « 3D », ainsi qu'une image d'intensité lumineuse d'un laser étaient disponibles [194]. Ces images étaient acquises par la projection d'une ligne laser sur le pneumatique et sa numérisation par une caméra. Lors du calibrage préalable du système, la correspondance entre la déformation de la ligne laser et une hauteur de référence était établie ; ceci permettait d'obtenir une image des altitudes de la surface. Un tel système s'avérait très utile pour détecter les reliefs sur un objet faiblement contrasté comme le pneumatique. Les figures 2.2 et 2.3 illustrent des images de la surface 3D du pneumatique. La figure 2.4 montre un exemple d'image d'intensité de laser.

Notre but était de détecter des anomalies présentes dans ces images. Cependant, il existait plusieurs difficultés.

i) Certaines anomalies avaient une très faible taille par rapport à celle des images, selon un rapport de $1/10^5$.

ii) Des anomalies avaient un signature 3D alors que d'autres avaient uniquement une signature texturée dans l'image d'intensité lumineuse.

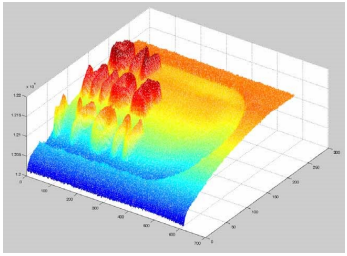
iii) Il existait aussi des reliefs 3D qui étaient tout à fait normaux.

iv) Les images étaient de très grandes dimensions (supérieures à 100 millions de pixels).

v) L'acquisition du pneumatique était faite en plusieurs sous-images qui n'étaient pas directement superposables.

Sur la partie externe des pneumatiques, il fallait trouver des anomalies 3D [215] alors que la surface présente de nombreux reliefs normaux (marquages réglementaires, logo, etc.) avec une signature 3D. Les reliefs 3D changent aussi en fonction des gammes de pneumatiques. La figure 2.2 montre un exemple des marquages en relief à détecter sur la surface du flanc.

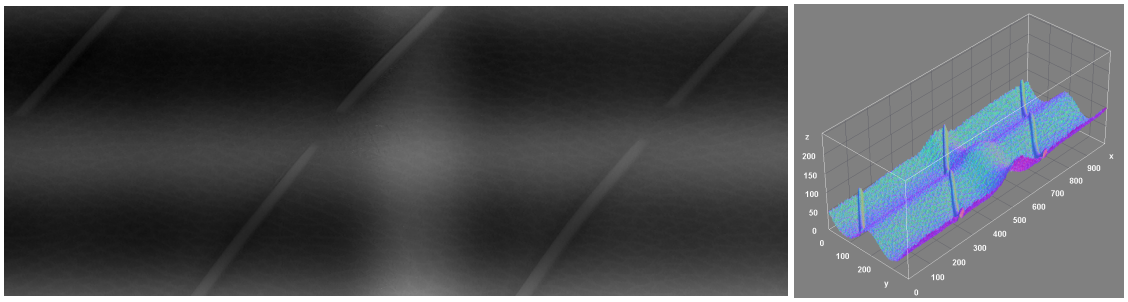
(a) Image 3D de surface.



(b) Surface topographique.

FIG. 2.2 – (a) Image de la surface tridimensionnelle externe du flanc du pneumatique. Les niveaux de gris correspondent à une altitude. (b) Surface topographique d'une partie de (a). Source [215].

Sur la partie interne des pneumatiques (i.e. la gomme intérieure), il fallait détecter des anomalies texturées [213], [225] ou en relief [214], [226] alors que cette partie est elle-même très texturée et présente des reliefs normaux. La figure 2.4 montre un exemple d'anomalies dans la texture sur la gomme intérieure. La figure 2.3 illustre des formes diagonales en relief, appelées « stries », qui doivent être présentes sur certains pneumatiques. La figure 2.5 présente plusieurs anomalies du même type qui apparaissent sur la texture de la partie interne.



(a) Image 3D de surface. (b) Surface topographique.

FIG. 2.3 – (a) Image tridimensionnelle de la surface interne du pneumatique (i.e. la gomme intérieure). Les niveaux de gris correspondent à une altitude. (b) Surface topographique de (a). Les formes en diagonales correspondent à des stries qui doivent être présentes dans certains pneumatiques. Source [214].

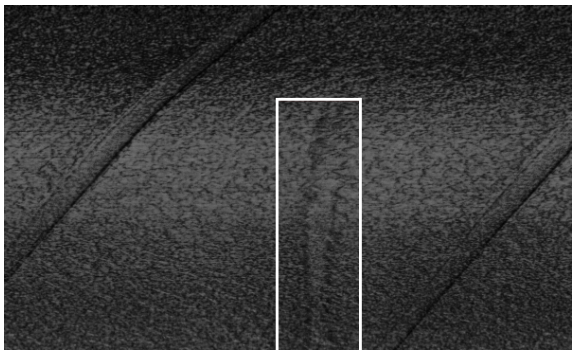
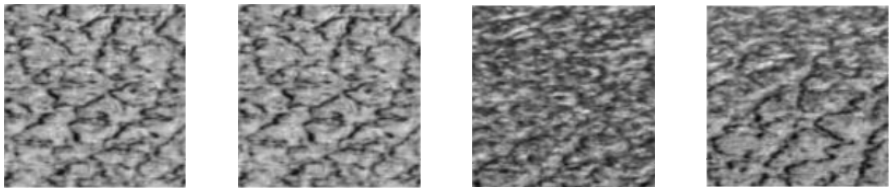
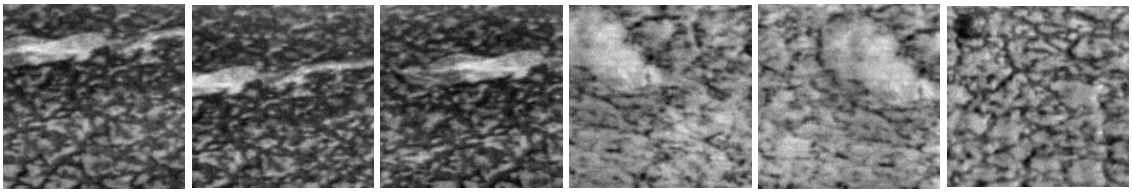


FIG. 2.4 – Image d'intensité lumineuse de laser d'une partie la surface interne d'un pneumatique. Dans la texture apparaît une anomalie encadrée en blanc. Source [225].



(a) Pas d'anomalie. (b) Pas d'anomalie. (c) Pas d'anomalie. (d) Pas d'anomalie.



(e) Anomalie. (f) Anomalie. (g) Anomalie. (h) Anomalie. (i) Anomalie. (j) Anomalie.

FIG. 2.5 – (a)-(d) Texture sans anomalie (e)-(j) Texture avec anomalies. Sources [225], [241].

En outre, la détection des anomalies devait être effectuée avec une contrainte de temps forte. Après analyse de la question pratique, il est apparu très rapidement que les signatures 3D devaient être détectées avec des outils de segmentation morphologique de type *top-hat* [277], [284] alors que les éléments texturés devaient être détectés par analyse de texture. Celle-ci devait s'appuyer sur de l'apprentissage statistique afin de s'adapter aux banques de données qui étaient en permanente évolution à cause de nouvelles acquisitions.

2.2 L'analyse d'image pour l'étude des performances des pneumatiques

Chez Michelin, de 2011 à 2014, j'ai travaillé sur divers projets de reconstruction tridimensionnelle (3D) de la surface et des composants des pneumatiques. L'objectif était de mieux comprendre leurs performances. Dans ce but, une image 3D de la surface et des composants permet de comprendre comment se déforme le pneu mis sous contrainte. Une comparaison avec la simulation numérique 3D par éléments finis devient alors possible.

La reconstruction 3D s'appuyait sur la stéréovision qui nécessite d'avoir au moins deux images du même objet numérisées sous deux poses différentes. En stéréovision, il est nécessaire de modéliser précisément le système de projection de la caméra et de connaître la matrice de transformation 3D entre les deux poses de la caméra [106]. Mes travaux et le stage de Thibaud Vilboux [312] ont permis de maîtriser la chaîne complète de reconstruction 3D qui comprend la prise d'image bidimensionnelle, le calibrage des caméras et l'algorithme de reconstruction tridimensionnel. La figure 2.6 illustre une mire de calibrage, dite « en damier ». La figure 2.7 montre un système de stéréovision et la reconstruction 3D d'un objet.

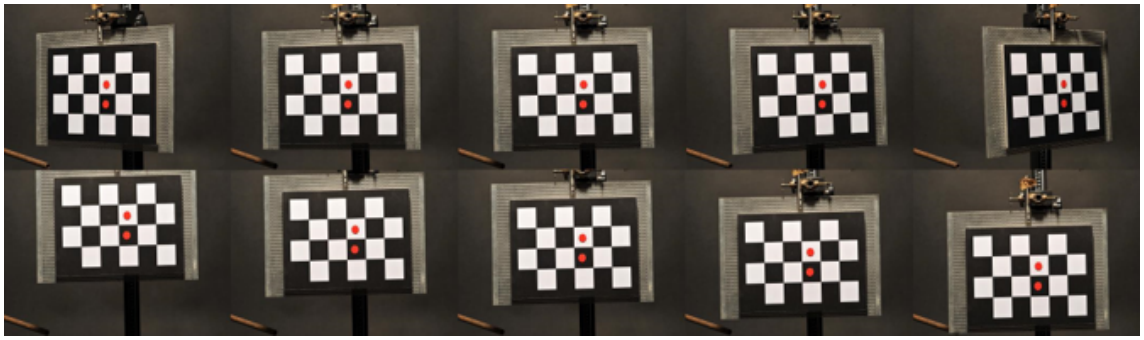
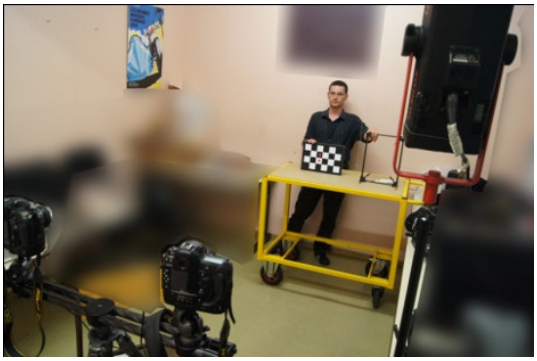


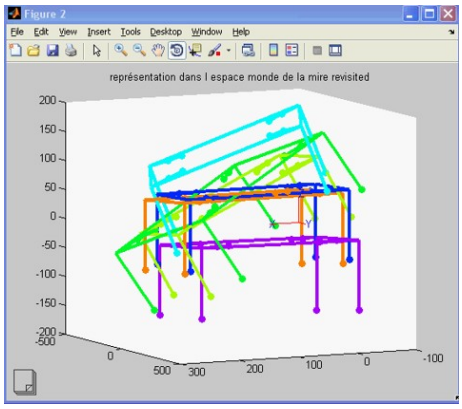
FIG. 2.6 – Acquisitions de différentes vues d'une mire damier. Source [312].

Afin de mieux comprendre les performances des pneumatiques, il était non seulement important de numériser la surface 3D mais aussi d'observer en profondeur les composants du pneumatique. Celui-ci pouvait être placé sous contrainte ou bien être au repos. La surface d'un pneumatique peut être reconstruite par des photographies dans le domaine du visible. En revanche, comme la partie interne est composée de différentes couches de gommes et de fils textiles ou métalliques, il est nécessaire d'utiliser des images acquises en rayons X afin d'observer les parties métalliques. Pour ce faire, j'ai développé un système complet de reconstruction 3D dans les domaines du visible et des rayons X. La figure 2.8 illustre la reconstruction 3D, dans ces deux domaines, de rayures et de petits disques métalliques collés à la surface d'un ballon gonflable. Dans cette expérience, seuls la ligne centrale des rayures et le centre des disques sont reconstruits. La figure 2.9 montre une nappe en acier d'un pneumatique extraite sur deux vues (ou clichés) acquises en rayons X et reconstruite en 3D.

L'extraction d'information sur les images acquises dans les domaines du visible et des rayons X nécessite d'avoir des outils d'analyse qui s'adaptent aux variations de contraste dans les images. La morphologie mathématique et un modèle qui tient compte de ces variations de contrastes tel



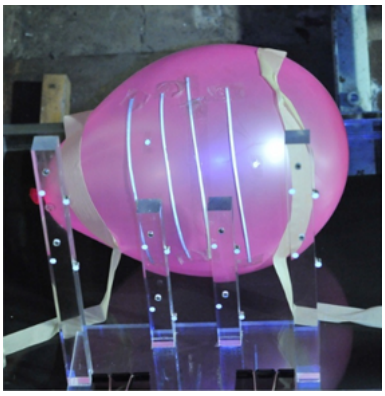
(a) Stéréovision.



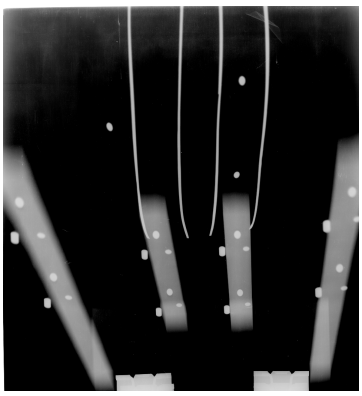
(b) Reconstruction 3D d'un objet.

FIG. 2.7 – (a) Système de stéréovision avec deux caméras. (b) On reconstruit en 3D un objet dans plusieurs positions. Source [312].

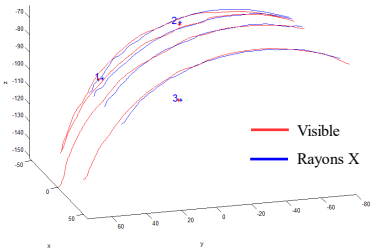
que le modèle *Logarithmic Image Processing* (LIP) [143], [144] sont bien adaptés. Mes travaux et le stage de Benoît Morel [198] sur le seuillage par contraste LIP et l'utilisation des métriques d'Asplund [142], [146] ont permis de mieux détecter les objets acquis en rayons X dans le pneumatique. La figure 2.10 montre deux cartes des distances d'Asplund entre une image acquise par rayons X et une sonde de forme cylindrique pour deux rayons différents. L'image représente une croix composée de deux bandes de plomb collées sur un pneumatique. La croix sert de repère pour faire de l'imagerie. Grâce aux rayons X, on voit les nappes métalliques des pneumatiques et la croix. La difficulté de détection de la croix est due au fait que les nappes métalliques apparaissent en superposition de la croix et agissent donc comme des objets confondants dans l'image. Les cartes des distances d'Asplund ont cependant permis de mettre en évidence les limites de la croix et de la détecter.



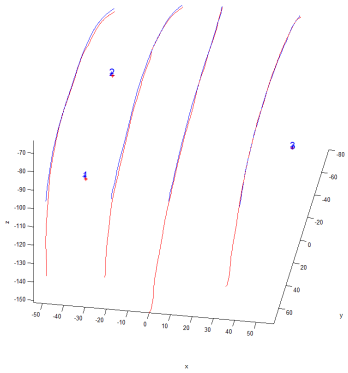
(a) Image acquise dans le visible.



(b) Image acquise en rayons X.

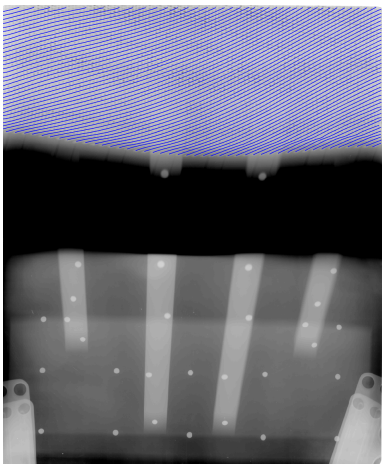


(c) Reconstruction 3D.

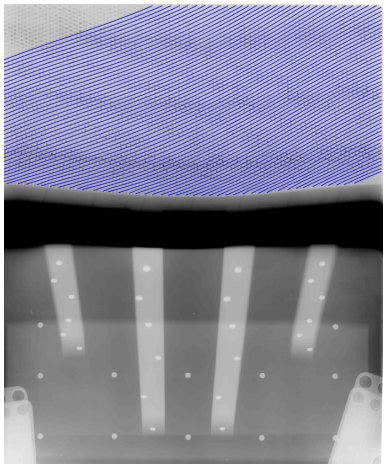


(d) Reconstruction 3D.

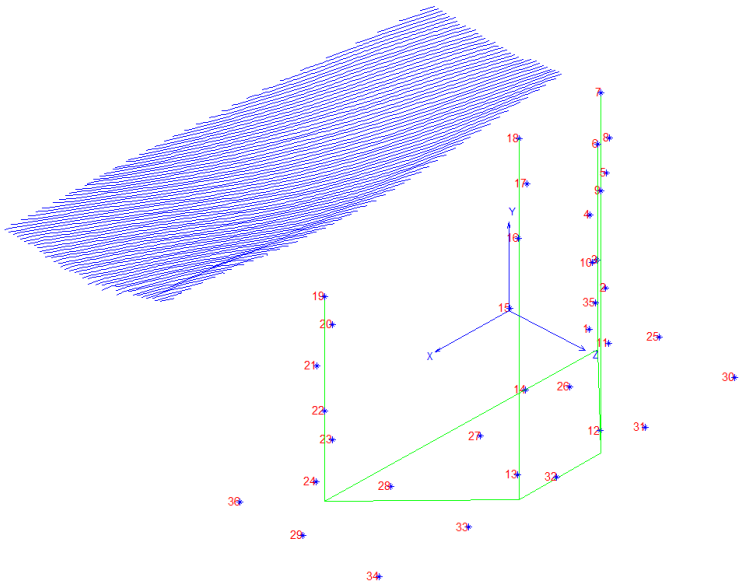
FIG. 2.8 – Images d'un ballon sur lequel sont collés des rayures et des petits disques (a) dans le domaine du visible et (b) dans celui des rayons X. (c) et (d) Deux points de vue différents, sur la reconstruction 3D des rayures et des disques dans le domaine du visible et dans celui des rayons X.



(a) Vue 1, rayons X.

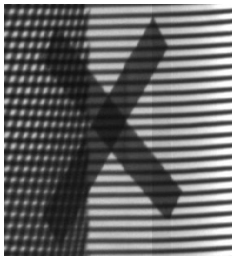


(b) Vue 2, rayons X.

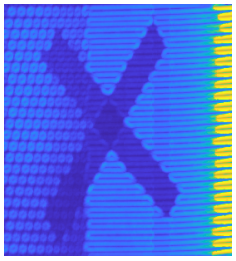


(c) Reconstruction 3D.

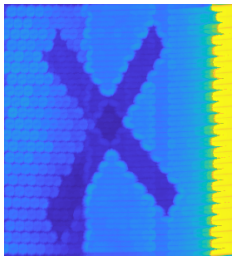
FIG. 2.9 – (a) et (b) Vues en rayons X des nappes d'un pneu et d'une mire de calibrage. Une des nappes a été extraite dans les deux images (courbes bleues). (c) Reconstruction 3D de la nappe et de la mire.



(a) Image f.



(b) Distance d'Asplund $r = 3$.



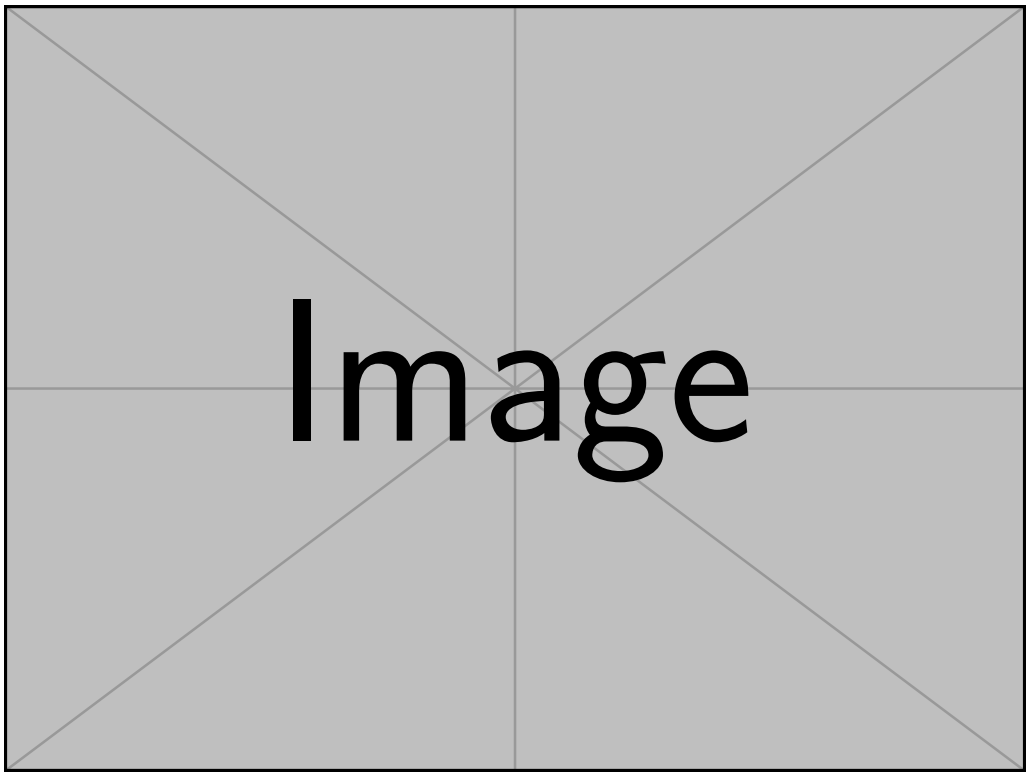
(c) Distance d'Asplund $r = 5$.

FIG. 2.10 – (a) Image f par rayons X d'un repère en forme de croix. Ce repère est composé de deux bandes de plomb collées sur un pneumatique. Les nappes métalliques du pneumatique apparaissent dans l'image et rendent difficile la détection du repère. (b) et (c) Carte des distances d'Asplund entre l'image f et un cylindre de rayon $r = 3$ et $r = 5$ pixels. Source [198].

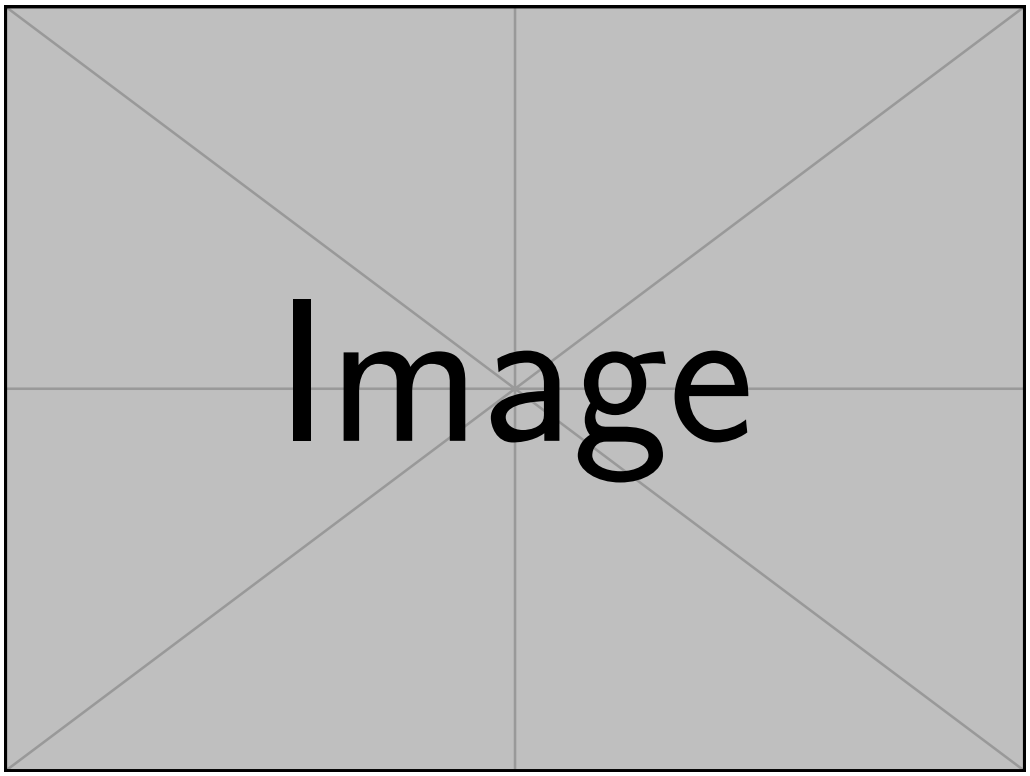
Une autre question pratique concernait la comparaison entre une épure théorique d'une coupe de pneumatique et une image haute définition de cette coupe réellement effectuée sur un pneumatique. Cette dernière est appelée coupe numérique et l'épure est un plan des frontières de certains composants du pneumatique. La coupe est faite selon un plan passant par l'axe de rotation. Comme le pneumatique est un objet souple, l'épure et l'image de la coupe ne peuvent être directement superposées. Mes travaux et le stage de Florian Rousset [265] ont montré qu'il est nécessaire de segmenter l'image puis de la recaler avec l'épure numérique de la coupe en utilisant notamment des déformations locales reposant sur des splines. La figure 2.11 montre la différence entre un recalage rigide et un recalage souple ou déformable de l'épure sur l'image de la coupe. Le pneumatique étant un objet déformable, un recalage déformable est nécessaire pour obtenir des résultats probants.

Lorsque l'épure théorique n'est pas disponible, il est possible d'extraire les frontières des composants des pneumatiques appelés semi-finis. La méthode disponible consistait en un opérateur qui sélectionnait plusieurs milliers de points reliés entre eux par une interpolation linéaire. Ce travail était donc très fastidieux et très long. Afin d'améliorer la situation, j'ai initié une collaboration avec Michel Jourlin et une PME innovante. L'objectif était de sélectionner un nombre bien plus réduit de points et d'utiliser des méthodes géodésiques [278], [279], [283] ou de percolation [123], [125], [180], [249], [250], [323], [324] pour suivre les frontières. La difficulté est que l'on travaille sur des surfaces faiblement contrastées. Pour résoudre ce problème, il a été proposé d'effectuer une détection de frontières à partir des métriques fonctionnelles d'Asplund qui sont robustes aux variations de contrastes et sensibles aux faibles contrastes. La figure 2.12 illustre la détection de frontières sur une coupe de pneu avec cette méthode.

Toutes ces méthodes de reconstruction et de segmentation avaient pour but de corroborer les résultats de déformations des pneumatiques obtenus par simulation numérique.



(a) Recalage rigide.



(b) Recalage souple.

FIG. 2.11 – (a) Recalage rigide de l'épure sur la coupe numérique. (b) Recalage souple (ou déformable) sur la coupe numérique [265].

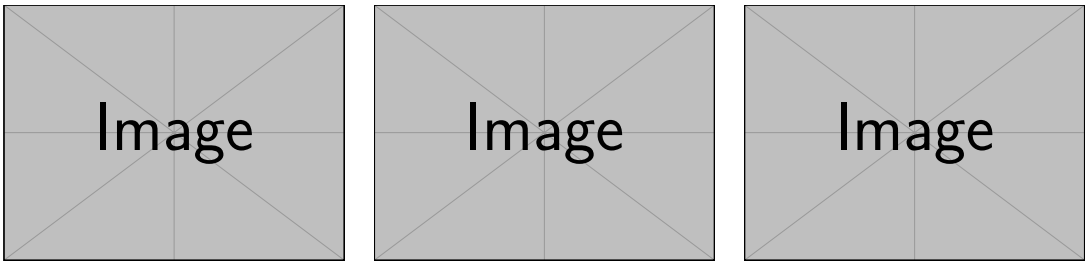


FIG. 2.12 – Trois exemples de détection de frontières sur une coupe de pneumatique.

2.3 L'analyse de banques d'images pour le diagnostic de la rétinopathie diabétique

Depuis 2014, à l'*International Prevention Research Institute* (IPRI), j'ai travaillé à l'amélioration du dépistage de la rétinopathie diabétique (RD). La RD est une complication microvasculaire des deux types 1 et 2 du diabète qui peut conduire à la cécité en l'absence de traitement précoce [175], [271]. C'est l'une des causes majeures de cécité dans le monde et un problème de santé publique de premier ordre. Le nombre de 463 millions de patients diabétiques aujourd'hui devrait augmenter de 51 % d'ici 2045, principalement dans les pays à faibles et moyens revenus [126]. On estime qu'environ un tiers des patients diabétiques développeront une forme de rétinopathie [127]. La carte de la figure 2.13 illustre la répartition et la progression du diabète dans le monde.

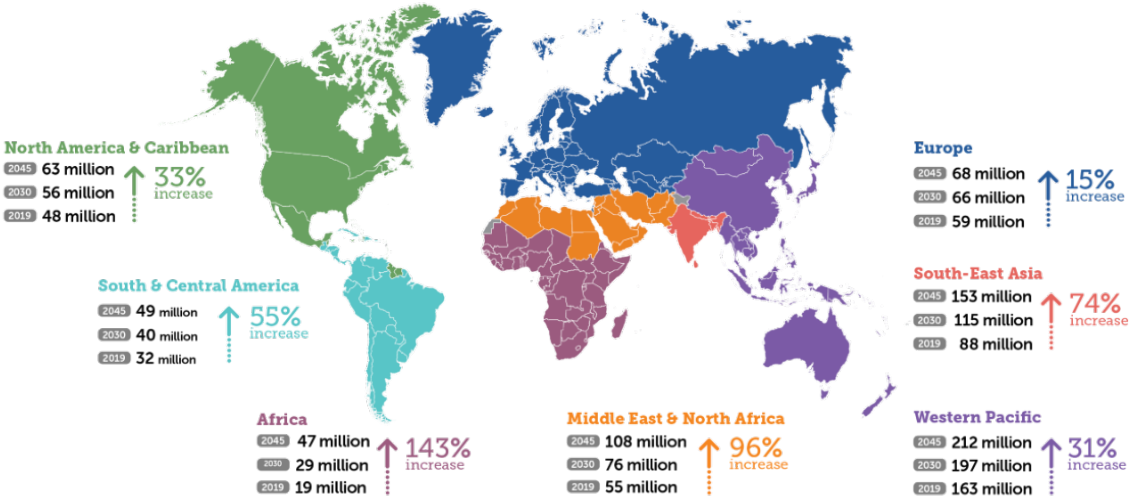


FIG. 2.13 – Nombre de personnes (entre 20 et 79 ans) atteintes du diabète au niveau mondial et régional. Pour chaque région la progression du nombre de cas est donnée entre 2019 et 2045. Source [126].

La rétinopathie cause des dommages aux vaisseaux sanguins qui approvisionnent la rétine [4], [175], i.e, la partie des yeux qui voit. Cela implique :

1. des occlusions de vaisseaux (i.e. des ischémies). De nouveaux vaisseaux vont apparaître et proliférer. Ils pourront saigner ou causer des détachements de la rétine. Cette étape correspond à la transition entre la rétinopathie non-proliférative et la rétinopathie proliférative ;
2. une rupture de la barrière hémato-rétinienne qui causera des pertes de fluides conduisant à l'œdème maculaire diabétique. Ces fuites vont endommager les photo-récepteurs.

Les dommages causés aux vaisseaux ne constituent pas la cause directe de la cécité mais sa conséquence. C'est la raison pour laquelle la rétinopathie diabétique n'est pas caractérisée par des symptômes évidents jusqu'à ce qu'elle soit à un stade avancé. Il est donc nécessaire qu'elle soit détectée le plus tôt possible. La seule manière de la détecter se fait par un examen régulier de la rétine, i.e. le fond d'œil, pour tout patient diabétique. Un moyen commun de diagnostic consiste à numériser une image couleur du fond de l'œil (Fig. 2.14). Cette image est ensuite analysée par un ophtalmologiste pour le diagnostic qui va rechercher des lésions telles que les micro-anévrysmes, les hémorragies, les anomalies microvasculaires intra-rétiniennes, etc. [175], [271]. Dans certains pays autres que la France, comme le Royaume-Uni, le dépistage est confié à des centres spécialisés où des techniciens spécialement formés font une première lecture des images pour un dépistage de la RD. Les cas les plus sévères sont ensuite transférés à un ophtalmologiste. La dilatation de la pupille peut améliorer la qualité de l'image et la sensibilité du dépistage. Cependant dans les pays où des personnels non-médecins peuvent réaliser des images du fond d'œil, l'administration par ces personnels d'un produit mydriatique (i.e. qui dilate la pupille) n'est pas toujours autorisée.

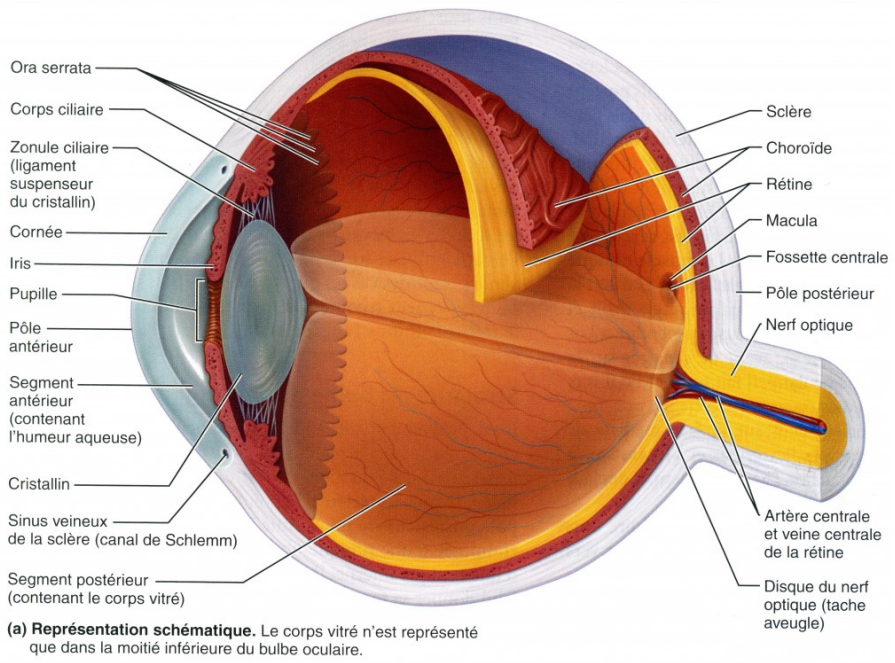


FIG. 2.14 – Représentation d'un œil en coupe.
Source : https://lewebpedagogique.com/brefjailuleblogduprofdesvt/files/2013/09/LOEIL1.pdf.

La figure 2.15 illustre les lésions caractérisant la rétinopathie diabétique. Cette maladie peut causer des changements dans les yeux tels que [127] :

- des micro-anévrismes - petits gonflements dans les vaisseaux sanguins de la rétine qui peuvent laisser fuir un fluide dans la rétine ;
- des hémorragies rétiniennes - toutes petites taches de sang qui fuient dans la rétine ;
- des exsudats secs - dépôts de lipides ;
- des nodules cotonneux - régions blanches cotonneuses et opaques liées à une ischémie nerveuse ;
- des dilatations veineuses et veines en « chapelet » ;
- des anomalies microvasculaires intra-rétiniennes (AMIR) - ramification anormale ou dilatation de vaisseaux sanguins existants ;
- de nouveaux vaisseaux anormaux (néovascularisation).

Au stade précoce, la RD est connue comme non proliférative avec des anomalies limitées à la rétine. Plus tard, elle devient proliférative à cause des anomalies microvasculaires qui restreignent le flux sanguin de la rétine et le prive d'oxygène. Dans une tentative d'apporter du sang dans les régions qui en sont privées, de nouveaux vaisseaux sanguins croissent depuis la rétine vers le vitré (une zone de l'œil). Cela peut causer une perte de vue sévère via les hémorragies vitreuses.

L'échelle internationale de sévérité de la maladie clinique [317], créée par un groupe d'experts, est fondée sur les découvertes de la *Wisconsin Epidemiologic Study of Diabetic Retinopathy* (WESDR) [154], [155], et l'*Early Treatment Diabetic Retinopathy Study Research group* (ETDRS) [72], [73]. Elle se compose de 5 stades de classification [127] :

1. « Pas de rétinopathie apparente » : pas d'anomalie.
2. « Rétinopathie Diabétique non proliférative légère » (RDNP) caractérisée par la présence de micro-anévrismes seuls.
3. La « RDNP modérée » est caractérisée par la présence d'autres lésions que des micro-anévrismes mais moins nombreuses que pour la RDNP sévère.

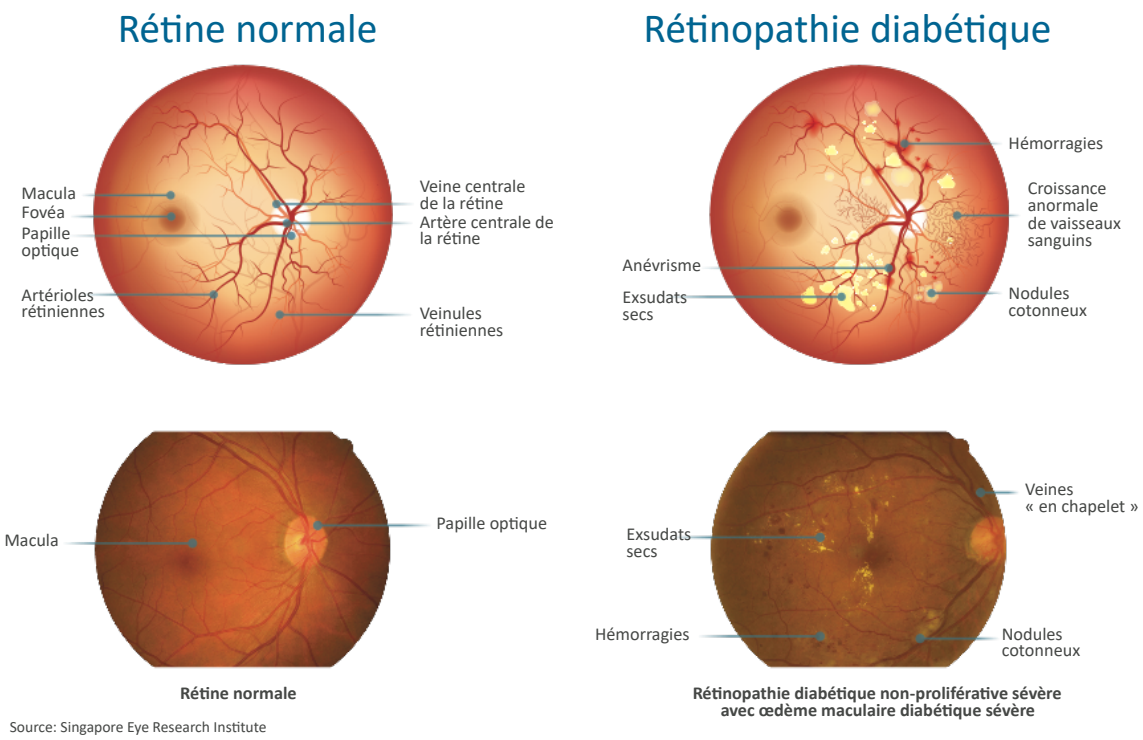


FIG. 2.15 – Comparaison entre une rétine normale et une rétine porteuse des symptômes de la rétinopathie diabétique. Source [127].

4. La « RDNP sévère » est le niveau clef défini par au moins l'une des caractéristiques suivantes : (i) hémorragies intra-rétiniennes ($\geq$ 20 dans chaque quadrant), (ii) veines en « chapelet » (dans deux quadrants), (iii) anomalies microvasculaires intra-rétiniennes (AMIR) (dans un quadrant), (iv) absence de signes de RD proliférative.
5. « La rétinopathie diabétique proliférative » (RDP) est caractérisée par l'un des points suivants : néovascularisation, hémorragies vitreuses/pré-rétiniennes.

La classification de la RD est faite conjointement avec celle de la maculopathie (œdème maculaire) sur une échelle de quatre niveaux. Il doit être noté (i) si l'œdème maculaire est présent ou absent. S'il est présent, il est alors classifié comme (ii) moyen (iii) modéré ou (iv) sévère selon la distance des lésions et de l'épaisseur de la rétine depuis le centre de la fovéa.

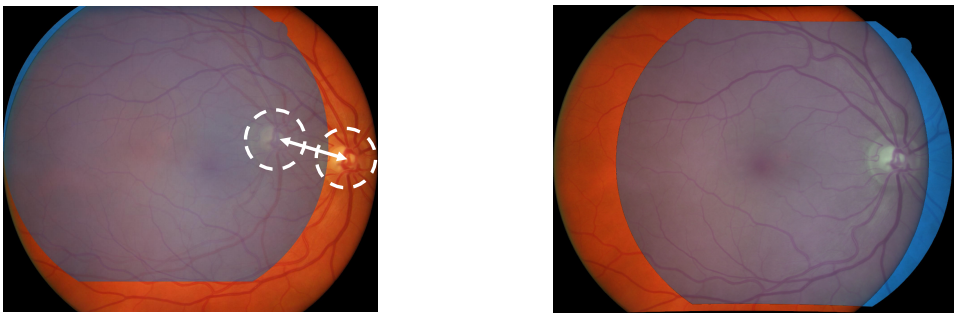
Les personnes diabétiques doivent être dépistées pour la RD de manière régulière. Il est recommandé un examen de l'œil annuel ou bisannuel selon les pays. Si cet examen révèle une rétinopathie diabétique et un œdème maculaire, le renvoi à un ophtalmologiste peut prévenir la perte de vision, ou stabiliser la vision. Le traitement est fait par photocoagulation au laser et/ou l'utilisation de traitements anti-VEGF (administration intraoculaire d'inhibiteurs du facteur de croissance endothélial vasculaire) [127].

L'existence de programmes de dépistage de la RD dans plusieurs pays comme le Royaume-Uni [102], les Etats-Unis, les Pays-Bas, la France, etc., a conduit à la création de grandes banques de données d'images de Santé Publique du fond d'œil qui peuvent être utilisées pour des études temporelles (i.e. longitudinales) de cohortes de patients. Une question commune qui apparaît dans l'analyse de ces banques de données de Santé Publique est comment rendre les images comparables ? Il existe notamment **trois problèmes**.

Premièrement, dans les grandes banques de données de Santé Publique, les images présentent des variations de contraste entre les différents examens. Elles peuvent être dues à une évolution de l'anatomie de l'œil (chirurgie de la cataracte, évolution de la rétinopathie à un stade plus sévère, etc.), au mouvement du patient, aux sources de lumière différentes lors de l'acquisition et aux différences d'absorption des tissus de l'œil [234]. Ces variations peuvent constituer une limite au

diagnostic de la rétinopathie car certaines lésions ne peuvent pas être prises en compte ou bien seule une partie de l'image peut être lue.

Deuxièmement, les images ne peuvent pas être directement superposées. Entre deux images du fond d'œil, les déformations sont causées par (i) la variation de la position de l'œil du patient modélisé par une rotation, une translation et une mise à l'échelle ; (ii) le changement d'appareil photographique caractérisé par une mise à l'échelle et une correction radiale ; (iii) la projection d'une scène 3D (la rétine de l'œil) qui est assimilée à une sphère sur le plan (2D) du capteur de la caméra et qui se modélise par une distorsion radiale ; (iv) la déformation radiale due à l'optique de la caméra et (v) la variabilité de couleur entre les images causées par l'intensité de la lumière et le capteur. La superposition ou recalage des images nécessite donc un modèle de déformation dont les paramètres doivent être estimés [234]. La figure 2.16 illustre la difficulté à recaler les images du fond d'œil numérisées avec un intervalle de 1,5 an et des appareils photographiques différents. Une superposition directe faite d'une rotation et d'une translation et sans modèle de déformation est insuffisante (Fig. 2.16a) alors qu'une superposition avec un modèle de déformation recale parfaitement les images (Fig. 2.16b).

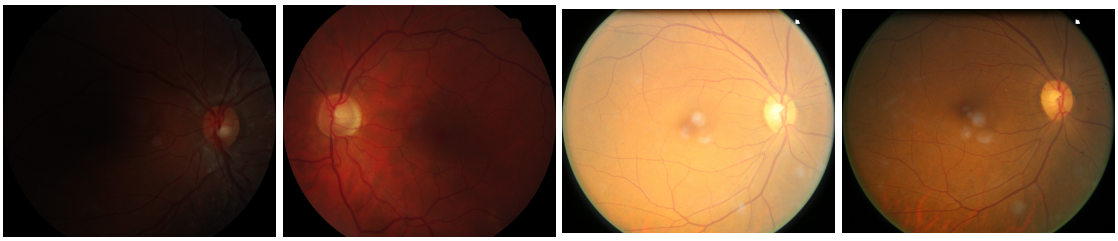


(a) Superposition directe. (b) Superposition avec modèle.

FIG. 2.16 – (a) Superposition directe et infructueuse due à l'absence de modèle adapté. (b) Superposition correcte grâce au modèle présenté plus loin (cf. section 6.2, p. 201). Afin de mieux visualiser les résultats, une des images du fond d'œil est représentée avec une fausse couleur bleutée.

Troisièmement, il est difficile de détecter des lésions ou d'autres parties anatomiques dans des images différentes qui présentent des contrastes différents. Pour la reconnaissance de formes, il est souvent intéressant d'extraire dans une image un gabarit de référence et d'en rechercher un autre qui lui est similaire dans une autre image. Cependant, dans les grandes banques de données, les conditions d'éclairement peuvent avoir changé entre les images et l'appariement du gabarit avec un motif similaire devient difficile. Pour surmonter cette difficulté des métriques robustes aux variations d'éclairement ont été étudiées. Elles sont appelées métriques d'Asplund fonctionnelles et ont été créées par Michel Jourlin [144]. Depuis ma thèse, j'ai utilisé ces métriques et étudié leurs propriétés, notamment leur lien avec la Morphologie Mathématique [231]. J'ai aussi développé un cadre de travail appelé *Morphologie Mathématique Logarithmique* [217] et qui utilise le *Logarithmic Image Processing*. Il s'est avéré très utile pour détecter les vaisseaux du fond d'œil dans des images faiblement contrastées [211].

La figure 2.17 montre que la couleur et le contraste changent dans chaque image (Fig. 2.17a et 2.17b). Le contraste diffère aussi entre deux examens du même patient prises à un intervalle de 1,5 an (Fig. 2.17c et 2.17d). La figure 2.18 illustre la détection des vaisseaux dans une image du d'œil faiblement contrastée par Morphologie Mathématique Logarithmique.



(a) Fond d'œil 1. (b) Fond d'œil 2. (c) Fond d'œil 3. (d) Fond d'œil 4.

FIG. 2.17 – (a) et (b) Images du fond d'œil acquises dans de bonnes conditions. (c) et (d) Images du fond d'œil du même patient acquises durant deux examens avec un intervalle de 1,5 an et dans des conditions plus difficiles.

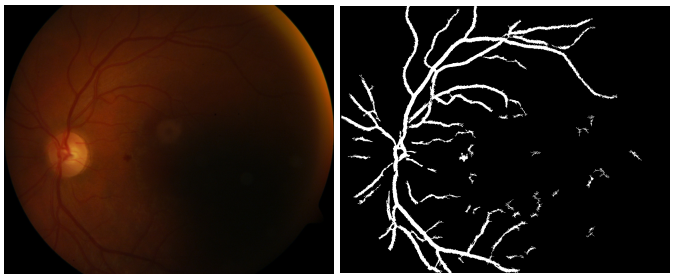


FIG. 2.18 – (a) Image du fond d'œil faiblement contrastée. (b) Détection des vaisseaux dans cette image par Morphologie Mathématique Logarithmique.

2.4 Conclusion du chapitre

Mon cadre de travail ayant porté mes recherches vers les applications industrielles et médicale, cela a influencé ma démarche de chercheur. Celle-ci a consisté à aller des applications vers la théorie. Dans ce chapitre, nous avons vu trois principales problématiques pratiques.

(i) La première porte sur la recherche d'anomalies dans les grandes banques d'images de pneumatiques pour faciliter le contrôle industriel. Cela nécessite de détecter des anomalies et re-

liefs et des anomalies texturées dans des images tridimensionnelles de surface et des images d'intensité d'un laser.

(ii) La deuxième est l'analyse des images de la surface et des composants internes des pneumatiques pour l'étude de leurs performances. Il est alors nécessaire de faire de la reconstruction multimodalités rayons X et visible et de comparer les coupes radiales de pneumatiques avec une épure numérique.

(iii) La troisième consiste à analyser les grandes banques d'images du fond d'œil de santé publique afin d'améliorer le diagnostic de la rétinopathie diabétique. Trois problèmes sont alors rencontrés :

a) les variations de contrastes entre les images de la rétine acquises avec une périodicité annuelle ;

b) la nécessité d'un modèle adapté pour superposer les images du fond d'œil ;

c) le besoin de méthode de reconnaissances de formes qui soient robustes aux variations de contrastes.

Dans les chapitres suivants, à partir de ces applications et problèmes pratiques, nous verrons les méthodes théoriques développées au fil du temps pour répondre à ces questions.

Chapitre 3

Segmentation morphologique et apprentissage statistique dans les grandes banques d'images

Au cours de mes expériences chez Michelin et à l'IPRI, j'ai travaillé sur l'analyse de grandes banques d'images. Au Centre de Recherche de Michelin, l'objectif était de détecter automatiquement des anomalies sur les pneumatiques fabriqués, à l'aide de la vision par ordinateur. *Le but était de développer un nouveau système de contrôle assisté par ordinateur des pneumatiques*. Comme tous les pneumatiques étaient contrôlés et comme les images avaient une très haute résolution, cela représentait beaucoup de données à traiter. J'ai donc conçu un premier système de détection des anomalies en relief à la surface des pneumatiques par Morphologie Mathématique ; puis un autre système détectant les anomalies texturées par apprentissage statistique. J'ai d'abord travaillé moi-même sur le sujet, puis j'ai créé une collaboration avec l'Ecole des Mines de Paris. Elle a permis de développer le système et d'accélérer le temps de traitement des algorithmes. La difficulté était de concevoir des algorithmes robustes et rapides sur ces très grandes banques d'images de la surface des pneumatiques.

A l'IPRI, j'ai analysé des banques de données de santé publique d'images de la rétine qui avaient été acquises *afin de suivre l'évolution de la rétinopathie diabétique* chez les patients porteurs de diabète. Durant une année, j'ai encadré un ingénieur-chercheur afin d'accélérer des algorithmes de Morphologie Mathématique par parallélisation des tâches sur plusieurs cœurs des processeurs et par vectorisation des données. Ce travail a notamment permis d'accélérer l'algorithme de seuillage de Köhler qui fonctionne par contraste sur les frontières des objets. En outre, mes travaux puis l'encadrement de deux stagiaires m'ont permis d'étudier :

i) un système de détection des images de qualité insuffisante par apprentissage statistique supervisé et

ii) un système de classification automatique des images du fond d'œil par apprentissage profond. Dans les images servant au dépistage de la rétinopathie, deux catégories ont été considérées : celles « à référer » à un ophtalmologiste et les autres, « à ne pas référer ». La conception de ces algorithmes pour l'analyse des images de la rétine s'inscrit un cadre de recherche. Leur application en clinique dépendra bien sûr des besoins spécifiques des ophtalmologistes et des habitudes propres à chaque pays.

3.1 Fondements de la Morphologie Mathématique

3.1.1 Morphologie Mathématique pour les ensembles

La Morphologie Mathématique (MM) a été initialement définie par Georges Matheron [176], [177] pour des ensembles, ou images binaires de l'espace euclidien $\mathcal{E}$. Dans cet espace, un ensemble B appelé élément structurant est utilisé pour analyser un autre ensemble X. Fixons l'origine o dans l'espace $\mathcal{E}$ et considérons une translation transformant o en un point p de $\mathcal{E}$. Cette translation transforme B en son translaté B_p qui interagit avec l'ensemble X selon deux opérations fondamentales [112].

- La première, la *dilatation*, aussi appelée *addition de Minkowski* [190] est définie par :

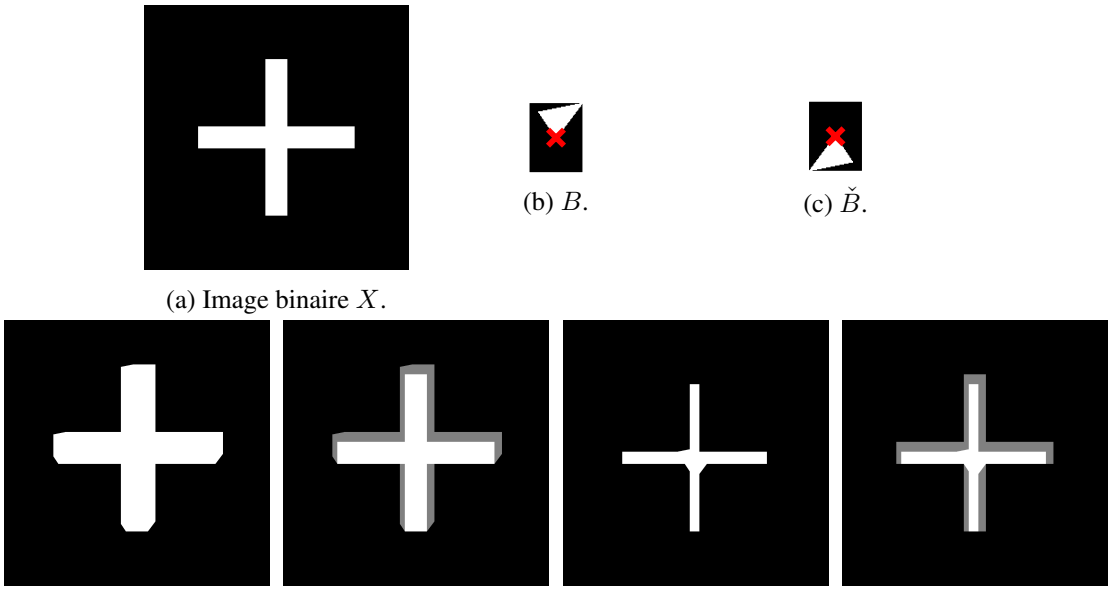
$$X \oplus B = \{x + b \mid x \in X, b \in B\} = \bigcup_{x \in X} B_x = \bigcup_{b \in B} X_b. \tag{3.1}$$

- La seconde, l'érosion, sa *duale*, introduite par Hadwiger [100], [101] sous le nom de *soustraction de Minkowski* est définie par :

$$X \ominus B = \{z \in \mathcal{E} \mid B_z \subseteq X\} = \bigcap_{b \in B} X_{-b}. \tag{3.2}$$

Remarque. La notation utilisée ici est celle de Heijmans et Ronse [111], de Angulo et Serra dans [309] ou de Najman et Talbot [201]. Elle diffère de la notation utilisée par Serra dans [275] et Matheron [177] qui a ensuite été reprise par Soille [284, p. 76]. Leur définition de l'addition de Minkowski $\oplus$ est la même qu'ici, mais ils appellent l'opérateur $X \to X \oplus \check{B}$ la dilatation par B, où $\check{B} = \{-b \mid b \in B\}$. Leur définition de la soustraction de Minkowski diffère de celle utilisée ici car ils écrivent $X \ominus B = \bigcap_{b \in B} X_b$, ce qui correspond à $X \ominus \check{B}$ dans notre notation. De plus, ils appellent $X \to X \ominus \check{B}$ l'érosion par B et cela coïncide avec notre nomenclature (bien que nous écrivions $X \ominus B$). L'intérêt de la notation employée ici est de définir des opérations duales par adjonction (voir ci-après) car cette dualité est générale alors que la dualité par complémentation n'est valable que dans le cas des ensembles booléens. Avec notre notation cette deuxième dualité donne : $X \ominus B = (X^c \oplus \check{B})^c$, où $X^c = \{x \in \mathcal{E} \mid x \notin X\}$.

La figure 3.1 illustre les opérations de MM sur les ensembles. En particulier, lorsque l'origine o de l'élément structurant B est incluse dans B, la dilatation de X par B revient à ajouter à X l'union des éléments structurants B dont l'origine o est déplacée le long de la frontière de X. L'érosion de X par B revient à enlever de X l'union des éléments structurants réfléchis $\check{B}$ dont l'origine est déplacée le long de la frontière de X.



(a) Image binaire X. (b) B. (c) $\check{B}$.

(d) Dilaté $X \oplus B$. (e) Parties ajoutées dans (d) en gris. (f) Erodé $X \ominus B$. (g) Parties enlevées dans (f) en gris.

FIG. 3.1 – Opérations morphologiques sur des ensembles. (a) Image binaire X. (b) Zoom sur l'élément structurant B et (c) l'élément structurant réfléchi $\check{B}$. L'origine des éléments structurants est indiquée par un point rouge. (d) Dilaté $X \oplus B$. (e) Parties en gris ajoutées lors de la dilatation. (f) Erodé $X \ominus B$. (e) Parties en gris enlevées lors de l'érosion.

3.1.2 Morphologie Mathématique pour les fonctions

La MM a ensuite été étendue aux fonctions à valeurs dans l'espace des réels $\mathbb{R}$ par Serra [275], Sternberg [288], [290] et Maragos [171]. Dans ce cas, une fonction $f : D \to \mathbb{R}$ définie sur le domaine $D \subset \mathbb{R}^n$ est analysée par une autre fonction $b : D \to \mathbb{R}$ appelée *fonction structurante*. L'extension de la MM aux fonctions réelles inclut les images à niveaux de gris dont les valeurs sont dans l'intervalle $[0; M[$ de l'espace des réels $\mathbb{R}$. Par exemple, pour les images numérisées sur 8 bits, M est égal à 256. D'un point de vue mathématique, la MM est définie sur des treillis complets [19], [275]. Présentons maintenant cette définition générale de la MM.

Définition 1 (Treillis complet)**.** *Etant donné un ensemble $\mathscr{L}$ et un ordre partiel $\leq$ sur $\mathscr{L}$, $\mathscr{L}$ est un treillis complet si tout sous-ensemble $\mathscr{X}$ de $\mathscr{L}$ possède un infimum et un supremum.*

L'infimum et le supremum de $\mathscr{X}$ seront respectivement notés $\bigwedge \mathscr{X}$ et $\bigvee \mathscr{X}$. L'ensemble des fonctions à valeur réelles $\overline{\mathbb{R}}^D$, avec $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, est un treillis complet selon l'ordre habituel $\leq$. Dans ce treillis, l'infimum et le supremum sont respectivement définis pour toute famille $\mathscr{X} \subset \overline{\mathbb{R}}^D$ par $(\wedge \mathscr{X})(x) = \wedge_{\overline{\mathbb{R}}} \{f(x) : f \in \mathscr{X},\ x \in D\}$ et $(\vee \mathscr{X})(x) = \vee_{\overline{\mathbb{R}}} \{f(x) : f \in \mathscr{X},\ x \in D\}$. Les deux opérations fondamentales de la MM, l'érosion et la dilatation, peuvent maintenant être définies dans le cas général des treillis.

Définition 2 (Erosion, dilatation [19], [277])**.** *Soient $\mathscr{L}_1$ et $\mathscr{L}_2$ deux treillis complets. Une application $\psi : \mathscr{L}_1 \to \mathscr{L}_2$ est*

- *une érosion ε : si et seulement si $\forall \mathscr{X} \subset \mathscr{L}_1$, $\psi(\bigwedge \mathscr{X}) = \bigwedge \psi(\mathscr{X})$;*
- *une dilatation δ : si et seulement si $\forall \mathscr{X} \subset \mathscr{L}_1$, $\psi(\bigvee \mathscr{X}) = \bigvee \psi(\mathscr{X})$.*

La paire (ε, δ) constitue une *adjonction* entre $\mathscr{L}_1$ et $\mathscr{L}_2$ si pour tout $X \in \mathscr{L}_1$, $Y \in \mathscr{L}_2$ il y a [112]

$$\delta(Y) \leq X \iff Y \leq \varepsilon(X). \tag{3.3}$$

Dans une *adjonction* (ε, δ), ε est une *érosion* et δ une *dilatation* [112]. Pour toute dilatation δ, il existe une unique érosion ε telle que (ε, δ) soit une adjonction et vice-versa. L'érosion ε et la dilatation δ sont duales par adjonction. De plus, si (ε, δ) est une *adjonction*, alors la composition :

$$\gamma = \delta\varepsilon \tag{3.4}$$

est une *ouverture* sur $\mathscr{L}_1$ et la composition :

$$\varphi = \varepsilon\delta \tag{3.5}$$

est une *fermeture* sur $\mathscr{L}_2$ [261]. *Ouverture* et *fermeture* sont des filtres morphologiques définis comme il s'ensuit [177], [261], [275].

Définition 3. *Un opérateur $\psi : \mathscr{L} \to \mathscr{L}$ sur le treillis complet $\mathscr{L}$ est appelé une* ouverture *si ψ est croissant ($\forall X, Y \in \mathscr{L}$, si $X \leq Y$ alors $\psi(X) \leq \psi(Y)$), anti-extensif ($\forall X \in \mathscr{L}$, $\psi(X) \leq X$) et idempotent ($\psi \circ \psi = \psi$). ψ est une* fermeture *s'il est croissant, extensif ($\forall X \in \mathscr{L}$, $X \leq \psi(X)$) et idempotent.*

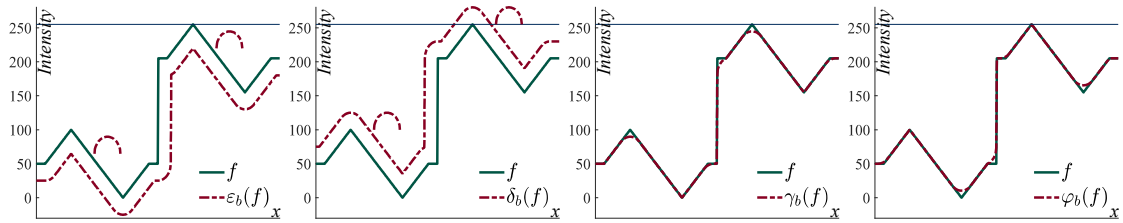
Dans le cas d'une fonction $f : D \to \overline{\mathbb{R}}$ et d'une fonction structurante $b : D \to \overline{\mathbb{R}}$, qui est invariante par translation dans le domaine D et l'espace $\overline{\mathbb{R}}$, la dilatation δ et l'érosion ε sont exprimées dans le treillis $(\overline{\mathbb{R}}^D, \leq)$ [112], [277] de la manière suivante :

$$(\delta_b(f))(x) = \bigvee_{h \in D} \{f(x-h) + b(h)\} = (f \oplus b)(x)\,; \tag{3.6}$$

$$(\varepsilon_b(f))(x) = \bigwedge_{h \in D} \{f(x+h) - b(h)\} = (f \ominus b)(x). \tag{3.7}$$

Les symboles $\oplus$ et $\ominus$ représentent l'extension aux fonctions des opérations de Minkowski entre ensembles [275]. Dans le cas d'expressions indéterminées, les conventions suivantes sont utilisées : $f(x-h)+b(h)=-\infty$ lorsque $f(x-h)=-\infty$ ou $b(h)=-\infty$, et $f(x+h)-b(h)=+\infty$ lorsque $f(x+h)=+\infty$ ou $b(h)=-\infty$ [112]. Cette définition de la MM pour les fonctions à valeurs à réelles inclut les images puisque leurs valeurs sont dans l'intervalle $[0;M[$.

La figure 3.2 illustre des opérations morphologiques sur des fonctions. L'image érodée $\varepsilon_b(f)$ (Fig. 3.2a) et l'image dilatée $\delta_b(f)$ (Fig. 3.2b) par une fonction structurante hémisphérique ont des valeurs qui excèdent l'intervalle $[0;M[$ d'une image. Pour les images, ceci est causé par la définition de la MM sur le treillis des fonctions à valeurs réelles $(\overline{\mathbb{R}}^D,\leq)$. L'érodé $\varepsilon_b(f)$ et le dilaté $\delta_b(f)$ d'une image peuvent donc avoir des valeurs dans tout l'espace des réels $\overline{\mathbb{R}}$.



(a) Erosion. (b) Dilatation. (c) Ouverture. (d) Fermeture.

FIG. 3.2 – Application d'opérateurs de MM fonctionnelle avec une fonction structurante b à une image à niveaux de gris f. (a) Erosion $\varepsilon_b(f)$, (b) dilatation $\delta_b(f)$, (c) ouverture $\gamma_b(f)$ et (d) fermeture $\varphi_b(f)$. (a) et (b) Pour les deux pics du signal, la fonction structurante hémisphérique b est représentée (après une translation horizontale).

L'ouverture et la fermeture sont des filtres morphologiques utilisés pour supprimer respectivement les pics clairs ou les creux sombres. Pour $h\in D$ et $v\in\mathbb{R}$, l'automorphisme $\tau_{h,v}$ des translations horizontales (i.e. selon x) et verticales (i.e. selon v) sur le treillis des fonctions $\overline{\mathbb{R}}^D$ est défini par $(\tau_{h,v}(b))(x)=b(x-h)+v$, pour la fonction $b\in\overline{\mathbb{R}}^D$. Soit $\mathbb{T}=\{\tau_{h,v}\mid h\in D,\ v\in\mathbb{R}\}$ l'ensemble des translations possibles. Pour une fonction $f:D\to\overline{\mathbb{R}}$ et une fonction structurante $b:D\to\overline{\mathbb{R}}$, l'ouverture $\gamma_b(f)$ et la fermeture $\varphi_b(f)$ (par adjonction) sont égales à [261] :

$$\gamma_b(f)=\sup\{\tau_{h,v}(b)\mid\tau_{h,v}\in\mathbb{T},\text{ et }\tau_{h,v}(b)\leq f\}\,; \tag{3.8}$$

$$\varphi_b(f)=\inf\{\tau_{h,v}(b)\mid\tau_{h,v}\in\mathbb{T},\text{ et }f\leq\tau_{h,v}(b)\}. \tag{3.9}$$

Ces équations montrent que l'ouverture correspond à un palpation par dessous, du graphe de la fonction f, par l'élément structurant b. La fermeture correspond à une palpation par-dessus, du graphe de la fonction f, par l'élément structurant b.

Preuve des équations (3.8) *et* (3.9). D'après l'équation (3.7), nous avons :

$$\begin{aligned}(\varepsilon_b(f))(x)&=\bigwedge_{h\in D}\{f(x+h)-b(h)\}\\&=\sup\Big\{v\in\overline{\mathbb{R}}\mid f(x+h)-b(h)\geq v,h\in D\Big\}\\&=\sup\Big\{v\in\overline{\mathbb{R}}\mid f(x+h)\geq b(h)+v,h\in D\Big\}\\&=\sup\Big\{v\in\overline{\mathbb{R}}\mid f(h)\geq b(h-x)+v,h\in D\Big\}.\end{aligned}$$

D'après les équations (3.4) et (3.6), nous avons pour tout $f \in \overline{\mathbb{R}}^D$ et $x \in D$:

$$\begin{aligned}
\gamma_b(f)(x) &= \delta_b(\varepsilon_b f)(x) \\
&= \bigvee_{h \in D} \{\varepsilon_b f(x-h) + b(h)\} \\
&= \bigvee_{h \in D} \{\varepsilon_b f(h) + b(x-h)\} \\
&= \bigvee_{h \in D} \left\{\sup\left\{v \in \overline{\mathbb{R}} \mid f(h') \geq b(h'-h) + v, h' \in D\right\} + b(x-h)\right\} \\
&= \bigvee_{h \in D} \left\{\sup\left\{v \in \overline{\mathbb{R}} \mid f(x) \geq b(x-h) + v\right\} + b(x-h)\right\} \quad \text{(on choisit } h' = x) \\
&= \bigvee_{h \in D} \{\sup\{v + b(x-h) \mid f(x) \geq b(x-h) + v\}\} \\
&= \sup\{\tau_{h,v}(b)(x) \mid \tau_{h,v} \in \mathbb{T} \text{ et } f(x) \geq \tau_{h,v}(b)(x)\}.
\end{aligned}$$

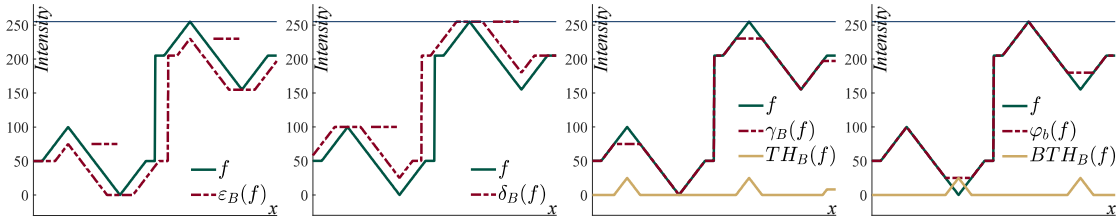
La démonstration est similaire pour la fermeture $\varphi_b(f)$ de l'équation (3.9). □

Pour les images à niveaux de gris, les *fonctions structurantes* les plus utilisées sont les *éléments structurants* plats notés B qui sont égaux à zéro sur un support compact D_b et prennent les valeurs $\wedge\overline{\mathbb{R}} = -\infty$ ailleurs. La différence entre une image f et son ouverture $\gamma_B(f)$ par un élément structurant plat B est appelée *top-hat* $TH_B(f)$ (ou *white-top-hat*) et permet de détecter les sommets. La différence entre la fermeture d'une image $\varphi_B(f)$ et l'image est appelée *black-top-hat* $BTH_B(f)$ et permet de détecter les vallées [277], [284]. Le *top-hat* $TH_B(f)$ et le *black-top-hat* $BTH_B(f)$ sont définis de la manière suivante :

$$TH_B(f) = f - \gamma_B(f), \tag{3.10}$$

$$BTH_B(f) = \varphi_B(f) - f. \tag{3.11}$$

La figure 3.3 montre des opérateurs morphologiques avec un élément structurant plat B. Le *top-hat* (Fig. 3.3c) et le *black-top-hat* (Fig. 3.3d) permettent d'extraire, respectivement, des sommets ou des vallées. On observe que les opérations morphologiques avec des éléments structurants plats appliquées à une image conservent une amplitude comprise dans l'intervalle $[0; M[$.



(a) Erosion. (b) Dilatation. (c) Ouverture. (d) Fermeture.

FIG. 3.3 – Application d'opérateurs de MM fonctionnelle avec un élément structurant plat B à une image à niveaux de gris f. (a) Erosion $\varepsilon_B(f)$, (b) dilatation $\delta_B(f)$, (c) ouverture $\gamma_B(f)$ et *top-hat* $TH_B(f)$, (d) fermeture $\varphi_B(f)$ et *black-top-hat* $BTH_B(f)$. (a) et (b) Pour les deux sommets du signal, l'élément structurant plat B est représenté (après une translation horizontale).

Propriété 1 (Propriété d'absorption pour les ouvertures et fermetures)**.** *Des familles d'ouvertures ou de fermetures peuvent être construites avec des éléments structurants de taille croissante. Ces familles sont indexées par un paramètre λ tel que :*

$$\lambda \geq \mu \implies \gamma_\lambda \leq \gamma_\mu \text{ et } \varphi_\lambda \geq \varphi_\mu. \tag{3.12}$$

Cette propriété constitue le fondement des familles granulométriques. L'idée est inspirée du tamisage du sable. Lorsque l'on tamise, les particules de plus faible taille passent à travers le tamis tandis que celles de plus grosse taille restent dessus. En utilisant des tamis de tailles de plus en plus petites, on peut trier le sable tamisé selon la taille de particules. En morphologie mathématique, on utilise des ouvertures et des fermetures de tailles croissantes pour construire des granulométries.

3.1.3 Filtrage morphologique

En MM, les filtres sont définis de la manière suivante.

Définition 4 (Filtre morphologique [201]). *Soient $\mathscr{L}$ un treillis complet et ψ un opérateur sur $\mathscr{L}$. Nous disons que ψ est un filtre morphologique s'il est croissant ($\forall f, g \in \mathscr{L}, f \leq g \implies \psi(f) \leq \psi(g)$) et idempotent ($\psi^2 = \psi$).*

Par conséquent, l'ouverture et la fermeture sont deux filtres morphologiques. La classe des filtres sur le treillis $\mathscr{L}$ forme un treillis complet. On peut obtenir des filtres par composition de deux ou trois opérateurs selon le théorème suivant.

Théorème 1 (Critère de Matheron [178]). *Soient ψ et ξ deux filtres tels que $\psi \leq \xi$. Alors :*

1) leurs produits de composition génèrent seulement les quatre filtres $\psi\xi$, $\psi\xi\psi$, $\xi\psi\xi$ et $\psi\xi$ qui sont ordonnés de la manière suivante :

$$\psi \leq \psi\xi\psi \leq \left\{ \begin{array}{c} \xi\psi \\ \psi\xi \end{array} \right\} \leq \xi\psi\xi \leq \xi.$$

2) $\xi\psi\xi$ est le plus petit filtre au-dessus de (i.e. plus grand que) $\xi\psi \vee \psi\xi$ et $\psi\xi\psi$ est le plus grand filtre sous $\xi\psi \wedge \psi\xi$.

3) Nous avons

$$\psi\xi\psi = \xi\psi \quad \iff \quad \xi\psi\xi = \psi\xi \quad \iff \quad \psi\xi \geq \xi\psi.$$

Ces produits de filtres sont particulièrement utilisés lorsque ψ est une ouverture et ξ une fermeture construites la plupart du temps avec des éléments structurants convexes. Chaque filtre $\xi\psi$ ou $\psi\xi$ élimine le bruit positif et les reliefs pointus ainsi que les bruits négatifs et les creux étroits. Il est important de noter que la composition de plus de trois opérateurs ne donne pas un nouveau filtre. Par contre, la composition de deux familles d'opérateurs permet de définir de nouveaux filtres appelés Filtres Alternés Séquentiels [277]. Ils servent à supprimer le bruit ou à faire une décomposition multi-échelles de l'image.

Définition 5 (Filtres Alternés Séquentiels - Serra [276]). *Soient γ_i une ouverture et φ_i une fermeture de taille i. Les quatre combinaisons suivantes sont des filtres morphologiques :*

$$\begin{aligned} m_i &= \gamma_i\varphi_i, \quad & r_i &= \varphi_i\gamma_i\varphi_i, \\ n_i &= \varphi_i\gamma_i, \quad & s_i &= \gamma_i\varphi_i\gamma_i. \end{aligned}$$

Les produits de compositions séquentiels suivants :

$$\begin{aligned} M_i &= m_i \cdots m_2 m_1, \quad & R_i &= r_i \cdots r_2 r_1, \\ N_i &= n_i \cdots n_2 n_1, \quad & S_i &= s_i \cdots s_2 s_1 \end{aligned}$$

sont des filtres morphologiques, appelés Filtres Alternés Séquentiels (FAS) de taille i.

3.1.4 Géodésie

La géodésie est très utilisée en MM, notamment pour les opérations de reconstruction géodésiques, l'algorithme de Ligne de Partage des Eaux, etc. Dans cette partie nous reprenons la présentation d'Angulo [10]. Définissons tout d'abord ce qu'est un chemin.

Définition 6 (Chemin)**.** *Un chemin C de longueur $l(C) = n+1$ ayant comme extrémités les points p et q (ou $C_{p,q}$) dans la trame V_G (de voisinage associé N_G) est un $(n+1)$-uplet $(p_0, p_1, \ldots, p_n)$ de pixels tels que :*

$$p_0 = p \text{ et } p_n = q$$
$$\forall i \in [1, n],\ p_i \in N_G(p_{i-1})\,.$$

Les chemins permettent de définir la géodésie sur un ensemble X.

Définition 7 (Distance géodésique binaire)**.** *Soit X un ensemble de $\mathbb{Z}^2$ (binaire). La distance géodésique entre deux points p et q à l'intérieur de l'ensemble X (p, $q \in X$) est la longueur du plus court chemin $C_{p,q}$ entre p et q restant à l'intérieur de X, i.e.*

$$d^g_X(p, q) = \inf\{l(C_{p,q})\}\,.$$

On considère par définition que cette distance vaut $+\infty$ s'il n'existe aucun chemin entre p et q à l'intérieur de X : i.e. si X n'est pas simplement connexe.

La boule géodésique de rayon $n \in \mathbb{N}$ et de centre $p \in X$ est l'ensemble $B_X(p, n)$ défini par : $B_X(p, n) = \{p' \in X, d_X(p', p) \leq n\}$. On peut maintenant définir la dilatation géodésique.

Définition 8 (Dilatation géodésique binaire)**.** *Soit $X \subset \mathbb{Z}^2$ un ensemble discret de $\mathbb{Z}^2$ et $Y \subset X$. La dilatation géodésique de taille $n \geq 0$ de Y dans X est l'ensemble des pixels de X dont la distance géodésique à Y est inférieure ou égale à n :*

$$d^n_X(Y) = \{p \in X \mid d_X(p, Y) \leq n\}\,.$$

La dilatation géodésique peut être exprimée grâce à des boules géodésiques :

$$\delta^n_X(Y) = \bigcup \{B_X(y, n), y \in Y\}\,.$$

Elle peut aussi être définie comme une combinaison de dilatations géodésiques unitaires :

$$\delta^n_X(Y) = \delta^1_X \circ \delta^1_X \circ \cdots \circ \delta^1_X(Y)_{(n fois)},$$

où la dilatation géodésique unitaire est définie par

$$\delta^1_X(Y) = (Y \oplus B) \cap X.$$

La figure 3.4 illustre une dilatation géodésique binaire.

L'opération de reconstruction qui permet d'extraire des objets marqués sans modifier leur forme peut maintenant être définie.

Définition 9 (Reconstruction binaire)**.** *La reconstruction de l'ensemble X à partir du marqueur $Y \subset X$ est obtenue par itération de la dilatation géodésique élémentaire de Y dans X jusqu'à l'idempotence :*

$$R^\delta_X(Y) = \delta^\infty_X(Y) = \delta^n_X(Y),\ \delta^n_X(Y) = \delta^{n+1}_X(Y).$$

La figure 3.5 illustre le filtrage par ouverture par reconstruction binaire. On érode un ensemble X puis on reconstruit X à partir de son érodé Y.

Pour des fonctions à valeurs réelles, on a des définitions similaires.

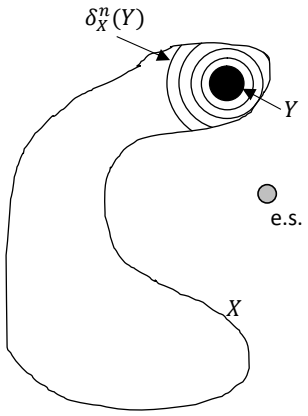


FIG. 3.4 – Dilatation géodésique binaire de Y dans X par un élément structurant (e.s.) B.

(a) Image binaire X. (b) Y : érodée de X. (c) Reconstruction $R_X^\delta(Y)$.

FIG. 3.5 – (c) Ouverture par reconstruction sur (a) une image binaire X. Suppression des plus petits ensembles par (b) une érosion par un disque suivie (c) d'une reconstruction binaire.

Définition 10 (Dilatation géodésique numérique)**.** *Soit $f : D \to \mathbb{R}$ une fonction (numérique) sous le masque géodésique $g : D \to \mathbb{R}$ (i.e. $g \geq f$), la dilatation géodésique de taille n est définie par :*

$$\delta_g^n(f) = \delta_g^1 \delta_g^1 \cdots \delta_g^1(f)_{(nfois)},$$

où la dilatation géodésique unitaire est donnée par :

$$\delta_g^1(f) = \delta_B^1(f) \wedge f.$$

La figure 3.6a illustre une dilatation géodésique numérique.

Définition 11 (Reconstruction numérique)**.** *La reconstruction de la fonction $g : D \to \mathbb{R}$ à partir de la fonction marqueur $f : D \to \mathbb{R}$, où $f \leq g$, est obtenue par itération de la dilatation géodésique élémentaire de f sous g jusqu'à l'idempotence :*

$$R_g^\delta(f) = \delta_g^\infty(f) = \delta_g^n(f),\ \delta_g^n(f) = \delta_g^{n+1}(f).$$

Comme illustré dans la figure 3.6b, la reconstruction numérique permet d'extraire une partie des structures d'une fonction g marquée par une autre fonction f. L'image f marque les structures ou régions de la fonction g à conserver (i.e. à préserver).

A partir de la reconstruction géodésique binaire, on peut définir le seuillage par hystérésis. Le problème du seuillage d'une fonction $f : D \to \mathbb{R}$ est de définir un seuil : trop haut et l'on risque de manquer des objets, trop bas et l'on risque d'ajouter du bruit de fond. Une solution consiste à choisir un ensemble binaire large X_{large} défini pour un seuil bas t_{low} par

(a) Dilatation géodésique numérique $\delta_g^1(f)$. (b) Reconstruction numérique $R_g^\delta(f)$.

FIG. 3.6 – (a) Dilatation géodésique numérique de f sous le masque g. (b) Reconstruction numérique $R_g^\delta(f)$ de g à partie de la fonction marqueur f.

$X_{large} = \{x \in D, f(x) \geq t_{low}\}$, et un ensemble étroit X_{narrow} défini pour un seuil haut t_{high} par $X_{narrow} = \{x \in D, f(x) \geq t_{high}\}$. On reconstruit ensuite l'ensemble X_{large} avec X_{narrow} comme marqueur pour obtenir l'image binaire X :

$$X = T_{\langle t_{low};t_{high}\rangle}(f) = R^\delta_{X_{large}}(X_{narrow}). \tag{3.13}$$

3.1.5 Ligne de Partage des Eaux

L'algorithme de segmentation le plus connu en MM est la Ligne de Partage des Eaux (LPE) [27], [32], [60], [187], [188]. Considérons l'image à niveaux de gris comme une surface topographique et supposons que des trous aient été percés en chaque minimum régional de la surface. La surface est alors lentement immergée dans l'eau. Partant des minima de plus faible altitude, l'eau va progressivement inonder les bassins versants de l'image. En outre, des barrages sont dressés aux lieux pour lesquels les eaux provenant de différents minima se rencontrent. A la fin de ce processus d'inondation, chaque minimum est entouré de barrages délimitant son bassin versant. L'ensemble des barrages, qui correspond à la ligne de partage des eaux, forme une partition de l'image d'entrée en ses différents bassins versants. La figure 3.7 illustre l'algorithme de Ligne de Partage des Eaux. Au lieu d'utiliser les minima régionaux de l'image pour simuler l'inondation on peut utiliser des marqueurs précédemment sélectionnés.

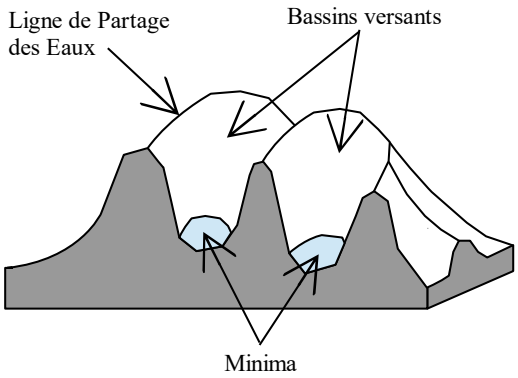


FIG. 3.7 – Relief à inonder par algorithme de Ligne de Partage des Eaux.

La LPE stochastique [8] consiste à calculer une LPE pour un jeu de marqueurs tirés aléatoirement sur l'image. On recommence le tirage des marqueurs un grand nombre de fois afin d'obtenir de multiples segmentations qui serviront à calculer une densité de probabilité de contours, grâce à une simulation par la méthode de Monte Carlo. Cette densité de contours est ensuite segmentée

par un seuillage ou une autre LPE. La LPE stochastique a été étendue au cas des images multispectrales ou multivariées dans [208], [212], [218], [219], [221]-[223] et dans [11]. Le calcul de la densité de probabilité de contours par simulation peut être remplacé par un calcul direct de la probabilité de contour entre des régions adjacentes [129], [189]. Cette méthode a été appliquée pour la segmentation 3D multi-échelles de milieux granulaires avec comme marqueurs des points [92] ou des lignes de Poisson orientées [91]. Des méthodes de calcul efficaces des cartes de probabilité de contours ont été proposées dans [170], [270]. Jeulin a aussi introduit des hiérarchies probabilistes morphologiques pour la segmentation de textures [131].

3.2 Segmentation morphologique d'objets en relief à la surface des pneumatiques

Le contrôle des pneumatiques nécessite de détecter les objets en relief à leur surface. Dans ce but, des acquisitions tridimensionnelles de surface sont effectuées par triangulation laser. L'acquisition d'image étant faite le long d'une ligne laser, on déplace cette ligne le long de la circonférence du pneumatique afin d'obtenir une image rectangulaire d'une partie de sa surface. Durant ce projet de recherche, j'ai été amené à travailler sur la détection des motifs en relief sur les flancs des pneumatiques ainsi que sur leur surface interne.

3.2.1 Détection des motifs sur la partie externe des pneumatiques

La partie externe et notamment les flancs présentent des motifs en relief. Ceux-ci ont été numérisés sur une image rectangulaire 3D $f : D \to [0; M[$ qui a été acquise le long de la circonférence des flancs. Le domaine $D \subset \mathbb{Z}$ est inclus dans l'ensemble des entiers relatifs et la valeur maximale M est égale à $2^8 = 256$ pour les images quantifiées sur 8 bits ou à 2^{16} pour les images quantifiées sur 16 bits. La taille du domaine de l'image pouvait être de l'ordre de $40\,000 \times 5\,000$ pixels pour un flanc [215]. La figure 3.8 montre un exemple d'image 3D d'un flanc. L'axe horizontal correspond à la direction circonférentielle du pneumatique et l'axe vertical à une direction radiale. La figure 3.9 présente une « forme bombée » sur les profils verticaux d'une image d'un flanc. Il sera nécessaire de corriger la courbure de cette forme afin d'aplanir la surface topographique de l'image 3D. Cette correction, appelée *mise à plat*, peut-être faite de manière simple par soustraction sur chaque ligne de sa moyenne comme on le fait pour l'affichage sur la figure 3.8. Cependant elle ne permet pas de corriger les variations circonférentielles de relief comme on en voit sur la figure 3.8a. Afin de de segmenter les objets en relief, j'ai préféré travailler à partir de techniques de type *top-hat* ou de différences d'ouvertures avec un élément structurant en forme de segment horizontal. Cette méthode avait l'avantage de corriger les variations de reliefs radiales et circonférentielles tout en étant simple et rapide pour ces images de très grandes dimensions. En effet, en 2008 sur ces images, la version de *MATLAB*® mettait plusieurs dizaines de minutes pour calculer une ouverture avec un segment de plusieurs centaines de pixels et au-delà de 1000 pixels, le calcul ne se terminait pas. Il a donc fallu utiliser une version accélérée des opérations morphologiques en utilisant le calcul vectoriel et les instructions SIMD (Single Instruction on Multiple Data) des processeurs [55]. Cette version était disponible dans le logiciel *Fast Morph-M* du Centre de Morphologie Mathématique de l'Ecole des Mines de Paris.

(a) Image 3D complète.

(b) Zoom.

FIG. 3.8 – (a) Image 3D de la surface complète d'un flanc. (b) Zoom sur une partie de l'image 3D. Afin de rendre visible les reliefs affichés, une correction du bombé radial a été faite sur chacune des lignes des images en lui retranchant sa moyenne.

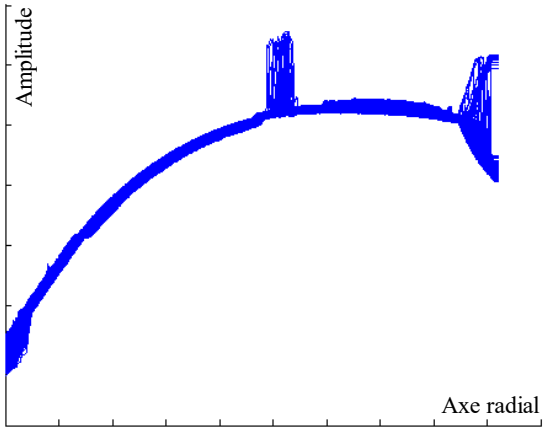


FIG. 3.9 – (a) Affichage de plusieurs profils verticaux de l'image 3.8a.

Les objets à détecter étant de taille variable et le pneu possédant des légères variations de relief selon la direction circonférentielle, il fallait choisir des éléments structurants suffisamment longs pour détecter tous les objets mais pas trop long afin d'éliminer les variations circonférentielles. Après étude des différents objets, il est apparu qu'une longueur maximale de 7 000 pixels permettait de répondre à ces deux conditions. Cependant, comme certains objets étaient beaucoup plus petits, j'ai choisi des ouvertures de plus petite taille afin de les détecter. Ce problème multi-échelles a nécessité trois longueurs d'éléments structurants en forme de segment : 150, 500 et 7 000 pixels [214]. A cause de la propriété d'absorption des ouvertures, une succession d'ouvertures de taille croissante peut être effectuée (Fig. 3.11a, 3.12a et 3.13a). Pour chacune de ces longueurs, une extraction des objets en relief par *résidus d'ouvertures* a été réalisée. Trois résidus d'ouvertures ont été calculés : $f - \gamma_{B_{150}}(f)$ (Fig. 3.11b), $\gamma_{B_{150}}(f) - \gamma_{B_{500}}(f)$ (Fig. 3.12b) et $\gamma_{B_{500}}(f) - \gamma_{B_{7000}}(f)$ (Fig. 3.13b). Les résidus d'ouvertures sont ensuite segmentés par seuillage afin d'extraire les objets. Un seuillage à 0,15 mm de hauteur a été appliqué aux deux premiers résidus d'ouvertures et un autre à 0,25 mm pour le dernier résidu. Les seuils ne devaient pas être trop bas pour éviter la segmentation de la moindre aspérité ni trop haut afin de ne pas manquer la détection de certains objets. Pour chacun des résidus, une segmentation est obtenue : seg_1 (Fig. 3.11c), seg_2 (Fig. 3.12c) et seg_3 (Fig. 3.13b). La segmentation du flanc seg est obtenue par l'union des trois segmentations $seg = seg_1 \bigcup seg_2 \bigcup seg_3$ (Fig. 3.13d). La méthode est illustrée sur la figure 3.10 et ses différentes

étapes sur les figures 3.11, 3.12 et 3.13.

Cette approche qui utilise les résidus d'ouvertures de tailles croissantes est proche de celle des *ouvertures ultimes* introduite par Serge Beucher [29], [30]. Cette dernière consiste à faire des ouvertures de taille croissante avec un pas unitaire afin de détecter toutes les tailles des objets. Appliquée à notre problème, cette méthode se serait avérée trop coûteuse en temps de calcul. C'est la raison pour laquelle, je me suis limité à trois tailles d'éléments structurants.

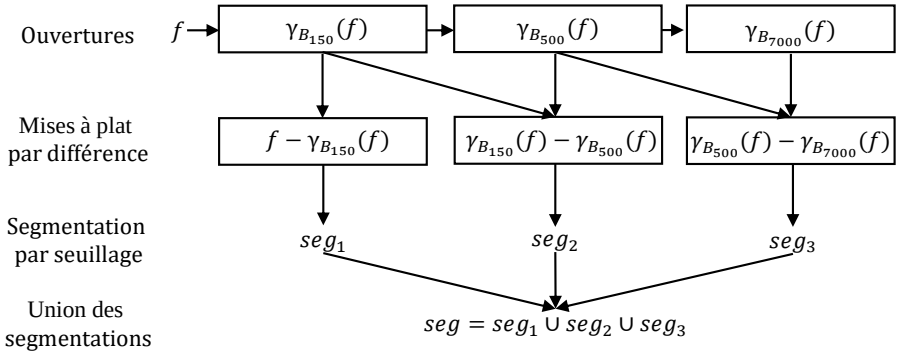
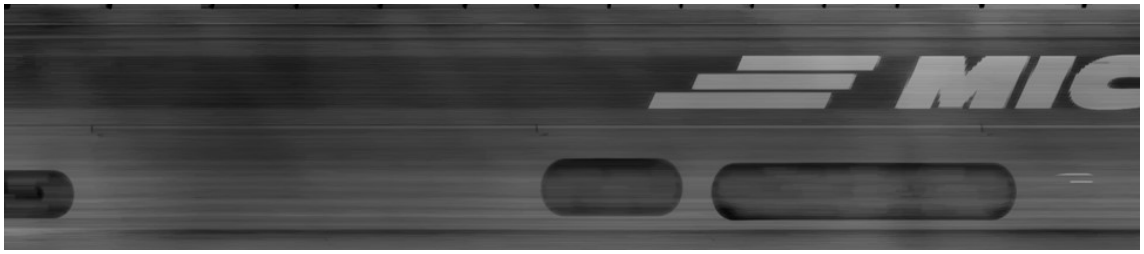
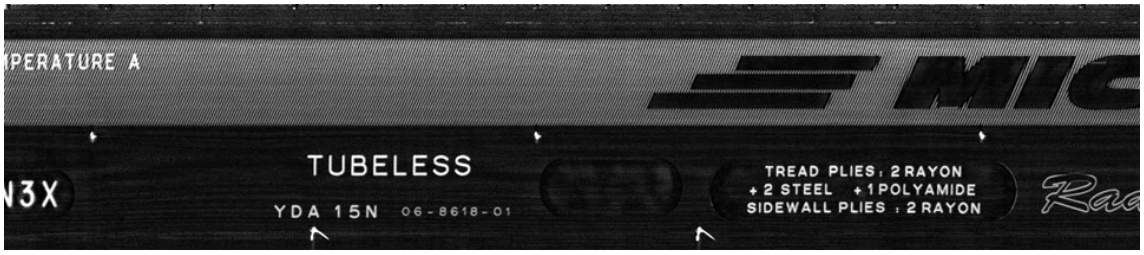


FIG. 3.10 – Méthode de segmentation des reliefs du flanc par résidus d'ouvertures.



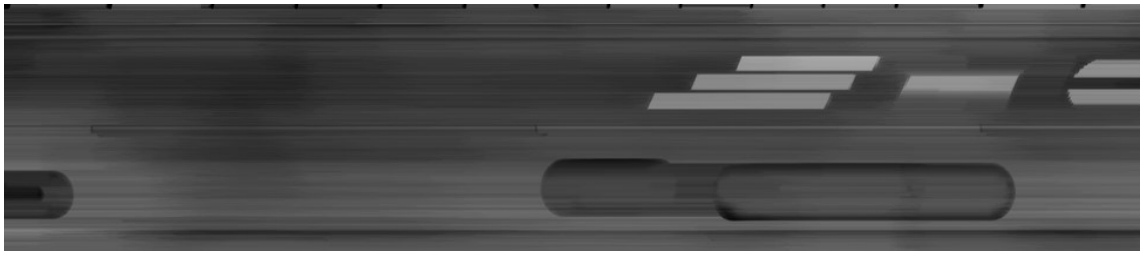
(a) Ouverture $\gamma_{B_{150}}(f)$.



(b) Résidu $f - \gamma_{B_{150}}(f)$.

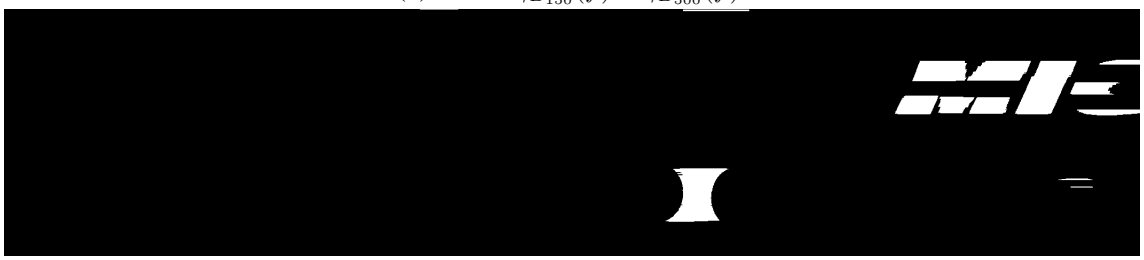
(c) Segmentation 1 seg_1.

FIG. 3.11 – Opérations pour segmenter l'image de la figure 3.8b. (a) Ouverture $\gamma_{B_{150}}(f)$. (b) Résidu d'ouvertures $f - \gamma_{B_{150}}(f) = \gamma_{B_0}(f) - \gamma_{B_{500}}(f)$, où $\gamma_{B_0}(f) = f$. (c) Segmentation 1 seg_1 du résidu (b).



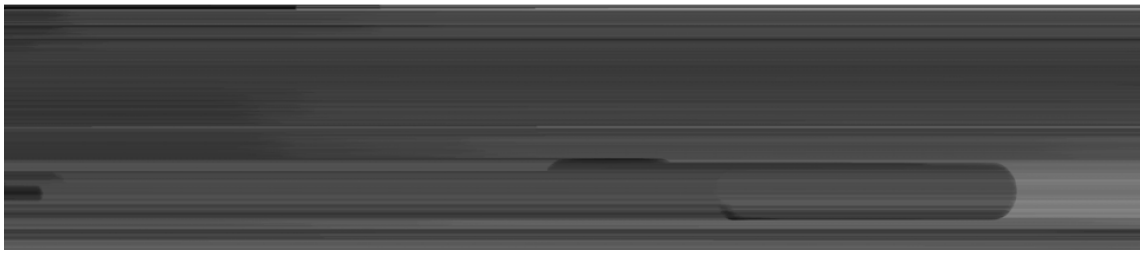
(a) Ouverture $\gamma_{B_{500}}(f)$.

(b) Résidu $\gamma_{B_{150}}(f) - \gamma_{B_{500}}(f)$.



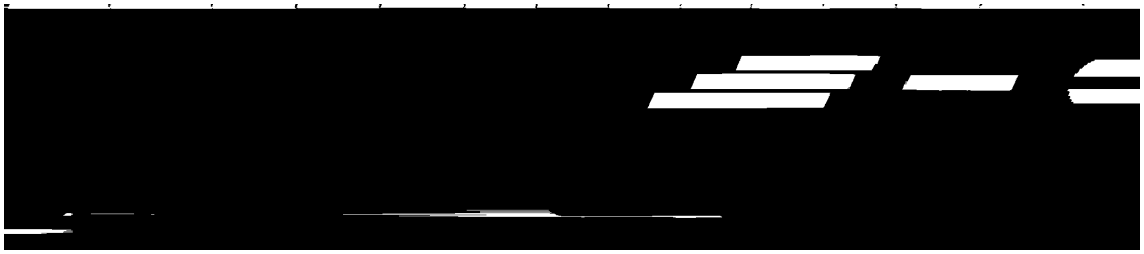
(c) Segmentation 2 seg_2.

FIG. 3.12 – Opérations pour segmenter l'image de la figure 3.8b. (a) Ouverture $\gamma_{B_{500}}(f)$. (b) Résidu d'ouvertures $\gamma_{B_{150}}(f) - \gamma_{B_{500}}(f)$. (c) Segmentation 2 seg_2 du résidu (e).



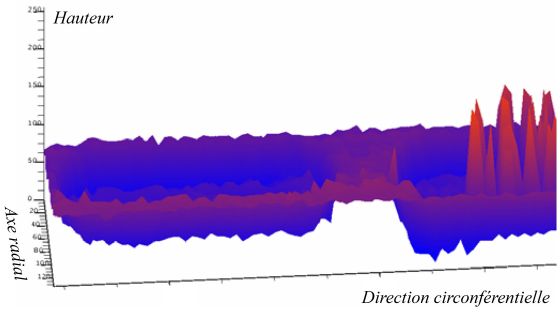
(a) Ouverture $\gamma_{B_{7000}}(f)$.

(b) Résidu $\gamma_{B_{500}}(f) - \gamma_{B_{7000}}(f)$.



(c) Segmentation 3 seg_3.

(d) Union des segmentations $seg = seg_1 \bigcup seg_2 \bigcup seg_3$.



(e) Zone inter-plaquettes en relief.

FIG. 3.13 – Suite des opérations pour segmenter l'image de la figure 3.8b. (a) Ouverture $\gamma_{B_{7000}}(f)$. (b) Résidu d'ouvertures $\gamma_{B_{500}}(f) - \gamma_{B_{7000}}(f)$. (c) Segmentation 3 seg_3 du résidu (b). (d) Union des segmentations $seg = seg_1 \bigcup seg_2 \bigcup seg_3$. Les formes détectées en trop sont entourées de rouge. Le cercle rouge correspond à une zone inter-plaquettes et l'ellipse rouge entoure une traînée. (e) Vue du relief de la zone entre deux plaquettes. Source [215].

L'union des segmentations $seg = seg_1 \bigcup seg_2 \bigcup seg_3$ (Fig. 3.13d) donne un bon résultat même si certains objets ont été détectés en trop comme : des traînées ou des zones inter-plaquettes ayant un relief positif entre deux zones en creux appelées « plaquettes » (Fig. 3.13e). Les traînées sont dues aux variations d'amplitude dans la direction circonférentielle. La difficulté est de séparer ces objets en trop du reste de la segmentation alors qu'ils apparaissent en relief.

Afin de supprimer ces objets détectés en trop, l'idée consiste à calculer pour chacun des objets de chaque segmentation des critères statistiques sur leur hauteur par rapport au fond. Seuls les objets vérifiant les critères statistiques seront conservés. Afin de calculer les hauteurs des objets, une image du fond est nécessaire. Elle ne doit contenir ni les objets en bosse (textes, symboles, etc.) ni ceux en creux (plaquettes). Les objets en bosses sont supprimés par une série d'ouvertures et ceux en creux sont supprimés par une série de fermetures qui sont combinées successivement comme des filtres alternés séquentiels. L'image filtrée du flanc de forme concave est appelée « forme bombée » ou « bombé » (Fig. 3.14a). Elle est définie par :

$$bombé(f) = \varphi_{B_{17000}} \gamma_{B_{7000}} \varphi_{B_{500}} \gamma_{B_{500}}(f). \tag{3.14}$$

Pour chacune des ouvertures $\gamma_{B_{150}}(f)$ et $\gamma_{B_{150}}(f)$, on calcule la différence $diff_B(f)$ entre l'ouverture et le bombé :

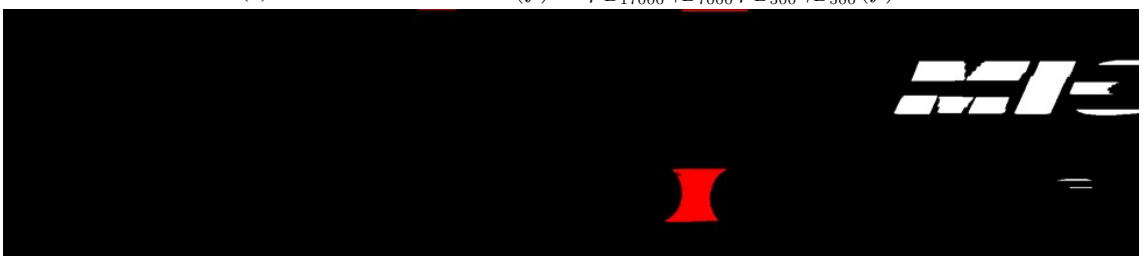
$$diff_B(f) = \gamma_B(f) - bombé(f). \tag{3.15}$$

Pour chacune des formes connexes (i.e. les objets) des segmentations seg_2 et seg_3, on calcule la moyenne et l'écart-type de leurs hauteurs sur l'image des différences $diff_B(f)$ correspondante. L'objet est conservé sur les segmentations seg_2 et seg_3 si : (i) sa moyenne est supérieure à 0,2 mm (ii) et son écart-type est inférieur à 0,2 mm. On obtient ainsi les segmentations corrigées notées seg_2^{cor} (Fig. 3.14b) et seg_3^{cor} (Fig. 3.14c). La correction n'est pas appliquée à la segmentation 1 seg_1 car l'élément structurant B_{150} étant de faible taille par rapport à l'image, il n'y avait pas de fausse détection observée à cette échelle. La segmentation finale seg^{cor} est obtenue comme l'union entre les segmentions : $seg^{cor} = seg_1 \bigcup seg_2^{cor} \bigcup seg_3^{cor}$ (Fig. 3.14d).

Cette méthode s'est avérée utile pour la segmentation des objets en relief sur différentes gammes de pneumatiques avec des marquages très différents. Elle a donné lieu à un brevet accepté dans plusieurs pays [215]. La segmentation du flanc peut ensuite être comparée à un plan théorique afin de vérifier la présence de tous les objets.

Remarque. Pour la correction de la forme bombée, une variante consiste à utiliser une décomposition par Analyse Factorielle des Correspondances [25] en considérant chaque colonne de l'image comme la réalisation d'une variable aléatoire. L'image est ensuite reconstruite à partir d'un nombre réduit d'axes factoriels, en général le premier.

(a) Forme bombée $bombé(f) = \varphi_{B_{17000}}\gamma_{B_{7000}}\varphi_{B_{500}}\gamma_{B_{500}}(f)$.



(b) Segmentation 2 corrigée seg_2^{cor}.

(c) Segmentation 3 corrigée seg_3^{cor}.

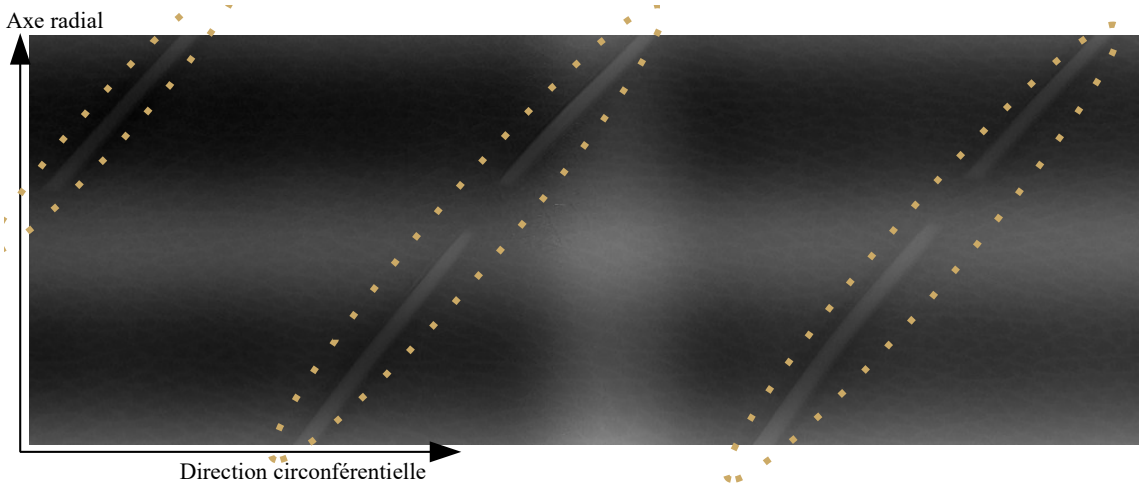
(d) Union des segmentations corrigées $seg^{cor} = seg_1 \bigcup seg_2^{cor} \bigcup seg_3^{cor}$.

FIG. 3.14 – Correction des segmentations des figures 3.12c et 3.13c. (a) Forme bombée $bombé(f)$. (b) Segmentation 2 corrigée seg_2^{cor}. (c) Segmentation 3 corrigée seg_3^{cor}. Les zones en rouge correspondent aux objets supprimés. (d) Union des segmentations corrigées $seg = seg_1 \bigcup seg_2^{cor} \bigcup seg_3^{cor}$. Source [215].

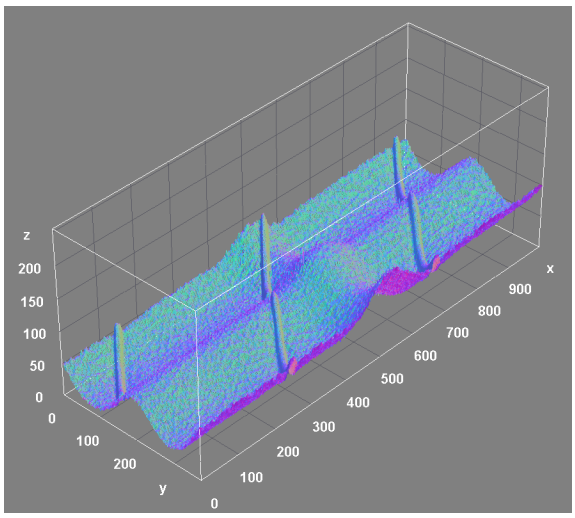
3.2.2 Détection des motifs sur la partie interne des pneumatiques

La partie interne des pneumatiques présente une texture 3D et des formes 3D allongées selon l'une des diagonales de l'image et que l'on appelle des « stries ». Ces stries normalement présentes à l'intérieur de certains pneumatiques doivent être détectées. La figure 3.15 montre une image 3D de la surface intérieure où apparaît une forme bombée (i.e. concave) selon l'axe radial. L'axe horizontal représente la direction circonférentielle et l'axe vertical représente l'axe radial. Une taille usuelle pour une image 3D de la surface intérieure pouvait aller jusqu'à $40\,000 \times 5\,000$ pixels [214].

Afin de supprimer la forme bombée tout en conservant les stries, on utilise des opérations de type *top-hat* avec des éléments structurants orientés dans la direction circonférentielle. Dans la pratique, ce seront des rectangles dont la longueur circonférentielle sera beaucoup plus grande que la longueur radiale. Afin de détecter les stries, un autre *top-hat* est utilisé avec un élément



(a) Image 3D de surface.



(b) Surface topographique.

FIG. 3.15 – (a) Image tridimensionnelle de la surface interne du pneumatique. Les niveaux de gris correspondent à une altitude. Les stries sont entourées en pointillés. (b) Surface topographique de (a). Source [214].

structurant rectangulaire selon l'axe radial.

Une des difficultés est de supprimer le relief dû à la texture que l'on peut assimiler à un bruit tout en conservant les stries. Le *top-hat* classique $TH_B(f) = f - \gamma_B(f)$ présente l'inconvénient de détecter le bruit et les stries. Afin de filtrer ce bruit, Salembier [268] a défini une variante du *top-hat* qui remplace l'ouverture γ_b, par l'opérateur $\gamma_{B_\gamma}\varphi_{B_\varphi} \wedge id$. Matheron et Serra ont démontré dans [277, p. 110, 120 et 192] que si γ est une ouverture et φ une fermeture alors $\gamma\phi \wedge id$ est une ouverture, où id représente l'identité. Nous avons appelé cette variante du *top-hat*, le *top-hat filtré* ou *filtered top-hat*. On la note FTH_{B_γ,B_φ} et elle est définie par :

$$FTH_{B_\gamma,B_\varphi}(f) = f - \left[\gamma_{B_\gamma}\varphi_{B_\varphi} \wedge id\right]. \tag{3.16}$$

La figure 3.16 illustre l'effet du *top-hat filtré*.

Sur l'image de la surface interne des pneumatiques, la longueur du bruit sur une ligne (Fig. 3.15a) a une taille bien plus petite que celle d'une strie. L'élément structurant B_φ de la fermeture aura donc une longueur inférieure à celle choisie pour l'élément structurant de l'ouverture B_γ. Dans cette application, comme les éléments structurants (e.s.) sont rectangulaires, on écrira B_{ij} pour un rectangle de i lignes selon l'axe radial et de j colonnes dans la direction circonférentielle.

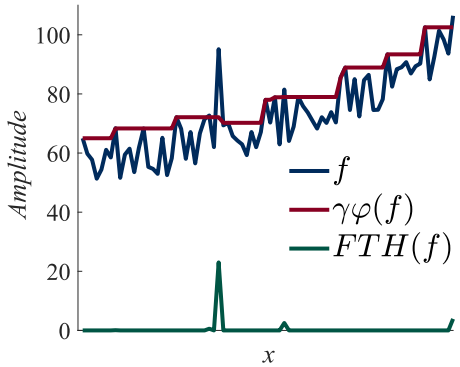


FIG. 3.16 – Illustration sur une fonction f d'une combinaison d'ouverture-fermeture $\gamma_B\varphi_B$ et du *top-hat filtré* $FTH(f) = f - [\gamma_B\varphi_B(f) \wedge id]$ avec le même élément structurant pour l'ouverture et la fermeture.

L'algorithme de traitement présenté sur la figure 3.17, est le suivant.

1) On effectue un premier *top-hat* avec un e.s. horizontal pour mettre à plat l'image, c'est-à-dire corriger la forme bombée radiale, $f_1 = TH_{3;151}(f)$.
2) Un deuxième *top-hat* avec filtrage permet d'éliminer le bruit, $f_2 = FTH_{(31;151),(31;51)}(f_1) = f_1 - [\gamma_{31;151}\varphi_{31;51} \wedge f_1]$.
3) On reconstruit le premier *top-hat* à partir du deuxième pour retrouver les structures partiellement éliminées par la suppression du bruit, $f_3 = R^{\delta}_{f_1}(f_2)$.
4) On effectue un troisième *top-hat filtré* identique au deuxième afin d'extraire les stries tout en éliminant la texture reconstruite précédemment, $f_4 = FTH_{(31;151),(31;51)}(f_3)$.
5) Un quatrième *top-hat filtré* avec un e.s. vertical permet de supprimer les bosses dans la direction circonférentielle, $f_5 = FTH_{(201;3),(11;3)}(f_4)$.
6) Un cinquième *top-hat filtré* avec un e.s. horizontal permet de supprimer le bombé radial restant, $f_6 = FTH_{(1;151),(1;51)}(f_5)$.
7) Un seuillage par hystérésis [284] permet d'extraire les stries et d'autres objets $seg_1 = T_{\langle 1;30\rangle}(f_6)$.
8) On supprime les objets ne correspondant pas aux stries sur des critères de forme.
9) On détermine l'angle des stries par moyenne des angles de chaque strie détectée.
10) Un algorithme similaire est appliqué sur les zones de recouvrements selon la direction circonférentielle afin d'obtenir la segmentation des stries seg_{stries}.

Les différentes étapes sont illustrées sur les figures 3.18 et 3.19. Cet algorithme a été testé avec succès non seulement sur l'image présentée mais aussi sur beaucoup d'autres.

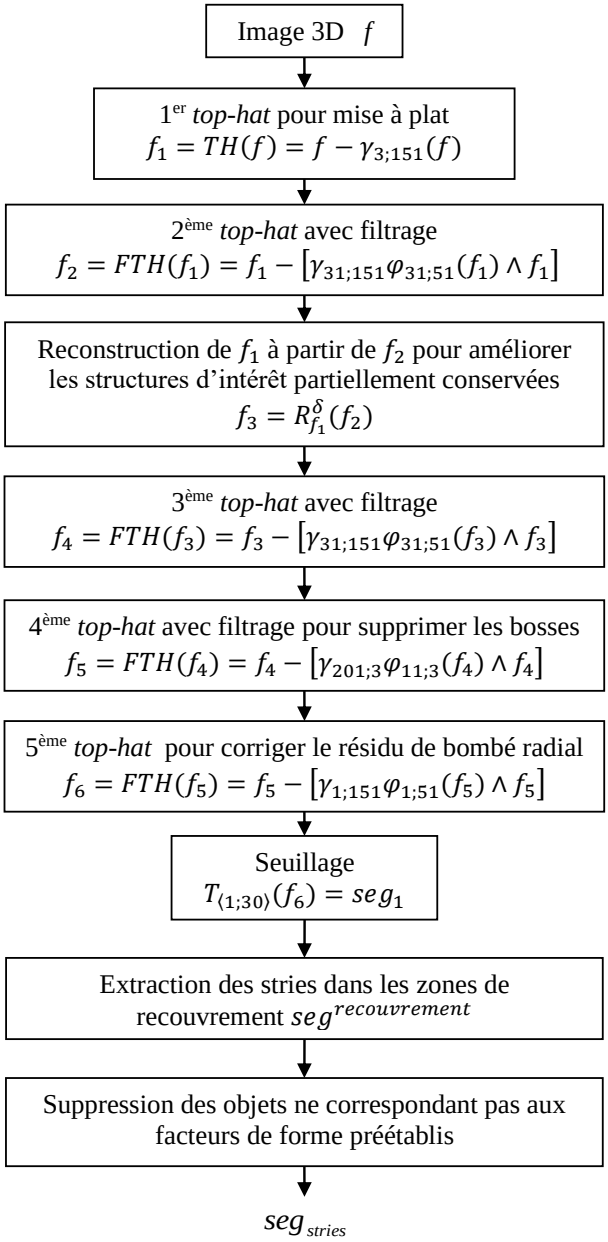
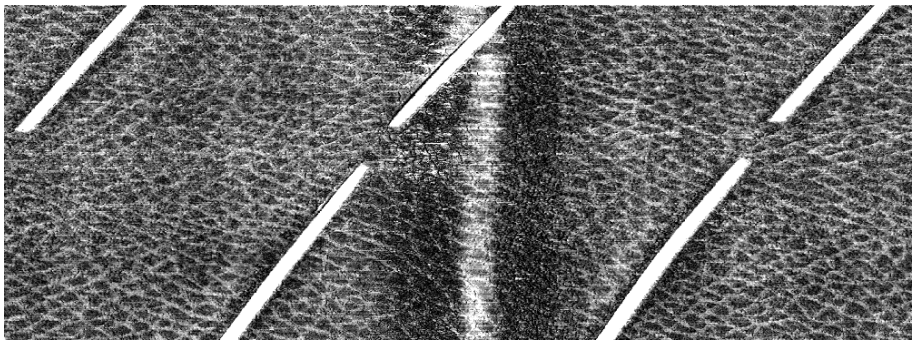
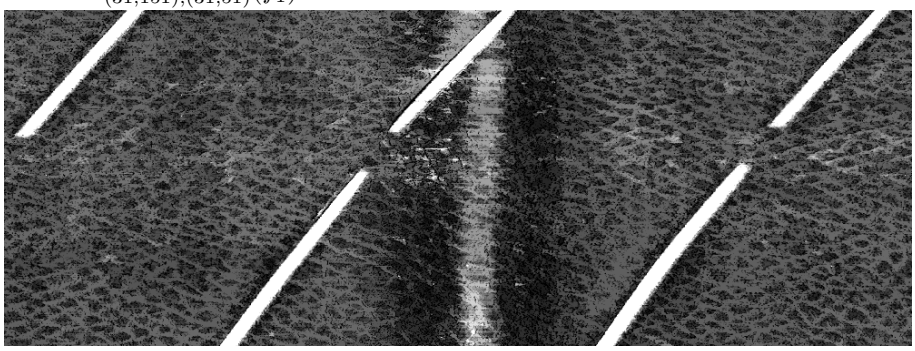


FIG. 3.17 – Méthode de segmentation des stries. Source [214].

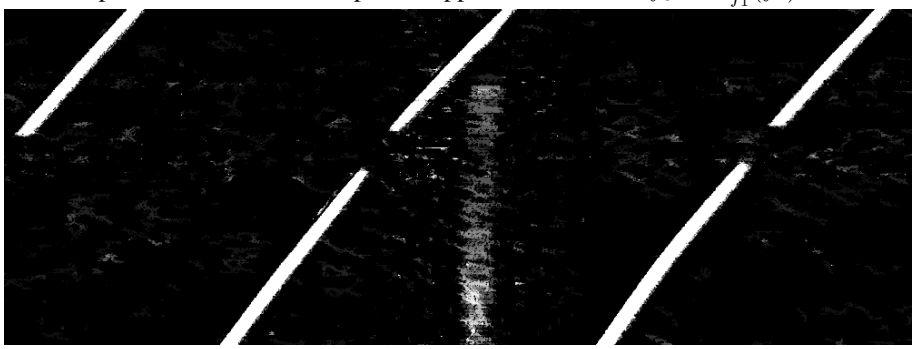


(a) Premier *top-hat* $f_1 = TH_{3;151}(f)$ pour mettre à plat l'image.

(b) Deuxième *top-hat* avec filtrage pour éliminer le bruit, $f_2 = FTH_{(31;151),(31;51)}(f_1)$.



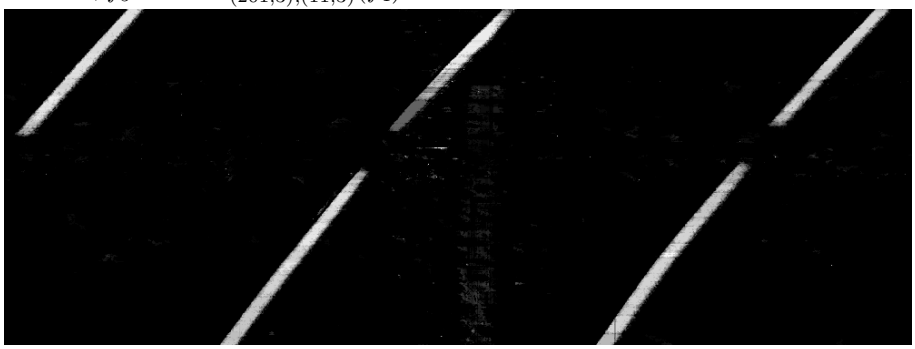
(c) Reconstruction du premier *top-hat* à partir du deuxième pour retrouver les structures partiellement éliminées par la suppression du bruit, $f_3 = R^{\delta}_{f_1}(f_2)$.



(d) Troisième *top-hat filtré* afin d'extraire les stries tout en éliminant la texture reconstruite précédemment, $f_4 = FTH_{(31;151),(31;51)}(f_3)$.

FIG. 3.18 – (a)-(d) Illustration des différentes étapes du traitement. Source [214].

(a) Quatrième *top-hat filtré* pour supprimer les bosses dans la direction circonférentielle, $f_5 = FTH_{(201;3),(11;3)}(f_4)$.



(b) Cinquième *top-hat filtré* avec un e.s. horizontal pour supprimer le bombé radial restant, $f_6 = FTH_{(1;151),(1;51)}(f_5)$.

(c) Seuillage par hystérésis $seg_1 = T_{\langle 1;30 \rangle}(f_6)$ et détermination de l'angle α des stries.

(d) Segmentation des stries après diverses corrections seg_{stries}.

FIG. 3.19 – (a)-(d) Suite des différentes étapes du traitement. Source [214].

J'ai aussi développé cinq autres algorithmes de segmentation d'anomalies visibles dans les images 3D ou d'intensité de laser. Le principe utilisé reposait sur des *top-hats filtrés* ou des filtres alternés séquentiels. Par choix de l'industriel, ils n'ont pas donné lieu à des brevets. Cependant, ces algorithmes sont pleinement fonctionnels. Ils ont été testés et validés sur de très nombreuses images et ont permis de lancer une collaboration de recherche avec l'Ecole des Mines de Paris dont certains résultats seront présentés dans la sous-section 3.3, p. 42.

J'ai aussi travaillé sur l'amélioration de la qualité des images 3D produites en supprimant notamment les zones de non-mesures et de fausses mesures du système d'acquisition par triangulation laser. Pour cette étude, j'ai notamment supervisé le travail du Dr. Yusi Shen (chercheur) afin de développer une méthode de corrections de ces zones par Morphologie Mathématique. Ces travaux ont donné lieu à trois brevets [195]-[197].

3.3 L'analyse de texture pour la recherche d'anomalies en contrôle industriel

Sur la partie interne des pneumatiques (i.e. la gomme intérieure), on recherchait des anomalies apparaissant dans la texture. Comme celles-ci ne présentaient que très peu de relief, elles étaient surtout visibles dans l'image d'intensité du laser. La figure 3.20 montre un exemple d'anomalie texturée.

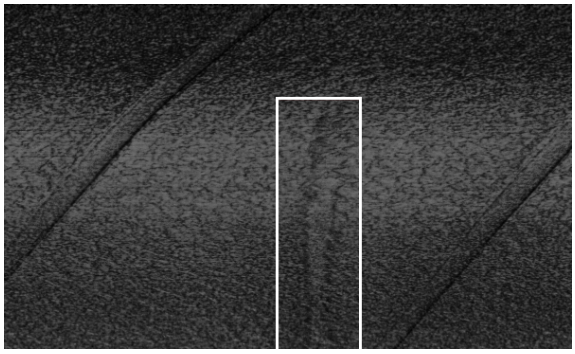


FIG. 3.20 – Image d'intensité lumineuse de laser d'une partie de la surface interne d'un pneumatique. Dans la texture apparaît une anomalie encadrée en blanc. Source [225].

Afin d'analyser les images texturées, il m'est apparu très rapidement qu'un système d'analyse de texture par apprentissage statistique devait être conçu. Comme les projets dans l'industrie sont balisés par des jalons stricts lors desquels un comité de pilotage décide la manière dont les chercheurs vont travailler durant une période de temps déterminée, l'étude s'est déroulée en deux étapes. Premièrement, à partir de travaux antérieurs [56], [57], [212], j'ai proposé une première méthode qui permette de reconnaître les anomalies dans des grandes banques d'images texturées de pneumatiques [213] acquises dans un environnement beaucoup plus contraint. Deuxièmement, une collaboration avec l'Ecole des Mines de Paris a permis d'améliorer la méthode et de la tester sur la banque de données.

3.3.1 Méthode initiale d'analyse de texture

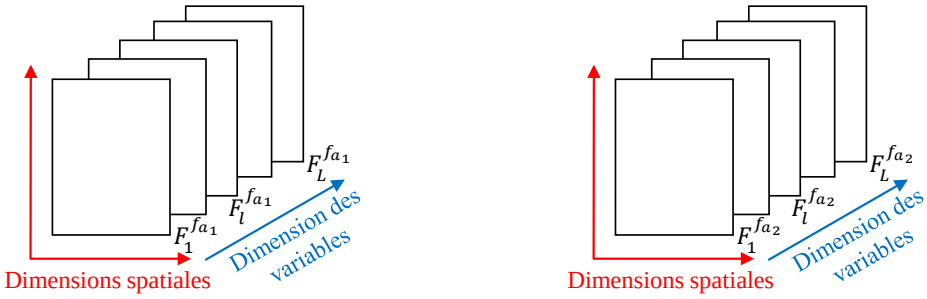
Présentons tout d'abord la première méthode qui a donné lieu à un brevet [225]. Comme toute méthode par apprentissage, l'approche proposée nécessitait un apprentissage statistique du classificateur sur une banque d'images d'entraînement et de validation puis son utilisation sur une banque de test. Pour ce faire, on disposait d'une banque d'images d'anomalies que l'on enrichissait au fur et à mesure des acquisitions et sur laquelle on allait pouvoir entraîner le classificateur.

Considérons que l'on a N images (ou réalisations) $f_{a_1} : D \to \mathbb{R}, \ldots, f_{a_n}, \ldots, f_{a_N}$ d'un même type d'anomalie a (Fig. 3.21). Ces images sont définies sur le domaine $D \subset \mathbb{Z}^2$ et ont leurs valeurs dans $\mathbb{R}$. Chacune contient P pixels, où P est égal au cardinal du domaine D, $P = card(D)$ (Fig. 3.21). Sur chaque image f_{a_n}, on applique une collection de filtres ou de transformées d'images $F_1 : \mathbb{R}^D \to \mathbb{R}^D$, $\ldots, F_l, \ldots, F_L$ pour obtenir une image multivariée $\mathbf{F}^{f_{a_n}} = (F_1^{f_{a_n}}, \ldots, F_l^{f_{a_n}}, \ldots, F_L^{f_{a_n}})$. La composante $F_l^{f_{a_n}}$ est égale à $F_l^{f_{a_n}} = F_l(f_{a_n})$ et l'image multivariée $\mathbf{F}^{f_{a_n}}$ est une fonction de D dans $\mathbb{R}^L$ (Fig. 3.22). Les filtres ou transformées peuvent être des érosions, dilatations, ouvertures, fermetures ou filtres en ondelettes, etc. Les tailles et orientations des éléments structurants ou des ondelettes peuvent bien sûr varier. Chaque composante (e.g. $F_l^{f_{a_1}}$) des images multivariées des anomalies $\mathbf{F}^{f_{a_1}} : D \to \mathbb{R}^L, \ldots, \mathbf{F}^{f_{a_n}}, \ldots, \mathbf{F}^{f_{a_N}}$ est mise en colonne puis concaténée avec les mêmes composantes des autres images (e.g. $F_l^{f_{a_2}}, \ldots, F_l^{f_{a_N}}$) pour former une colonne (e.g. $F_l^f = [F_l^{f_{a_1}}, \ldots, F_l^{f_{a_n}}, \ldots, F_l^{f_{a_N}}]^t$) d'une matrice ou tableau de contingence $\mathbf{F}^f = (F_1^f, \ldots, F_l^f, \ldots, F_L^f)$ de $N \cdot P$ lignes et L colonnes (Fig. 3.24). L'opérateur $[.]^t : \left(\mathbb{R}^D\right)^N \to \mathbb{R}^{N.P}$ transforme les images $F_l^{f_{a_n}} : D \to \mathbb{R}$ en vecteurs puis concatène les vecteurs en un seul en les mettant bout à bout.

(a) Anomalie 1, f_{a_1}.

(b) Anomalie n, f_{a_n}.

FIG. 3.21 – (a) et (b) Deux images d'un même type d'anomalie.



(a) Image multivariée de l'anomalie 1 : $\mathbf{F}^{f_{a_1}}$.

(b) Image multivariée de l'anomalie n : $\mathbf{F}^{f_{a_n}}$.

FIG. 3.22 – (a) et (b) Deux images multivariées $\mathbf{F}^{f_{a_1}}$ et $\mathbf{F}^{f_{a_n}}$ d'un même type d'anomalie obtenues comme la réponse à une collection de filtres $(F_1, \ldots, F_L)$ sur les images de départ f_{a_1} et f_{a_2}.

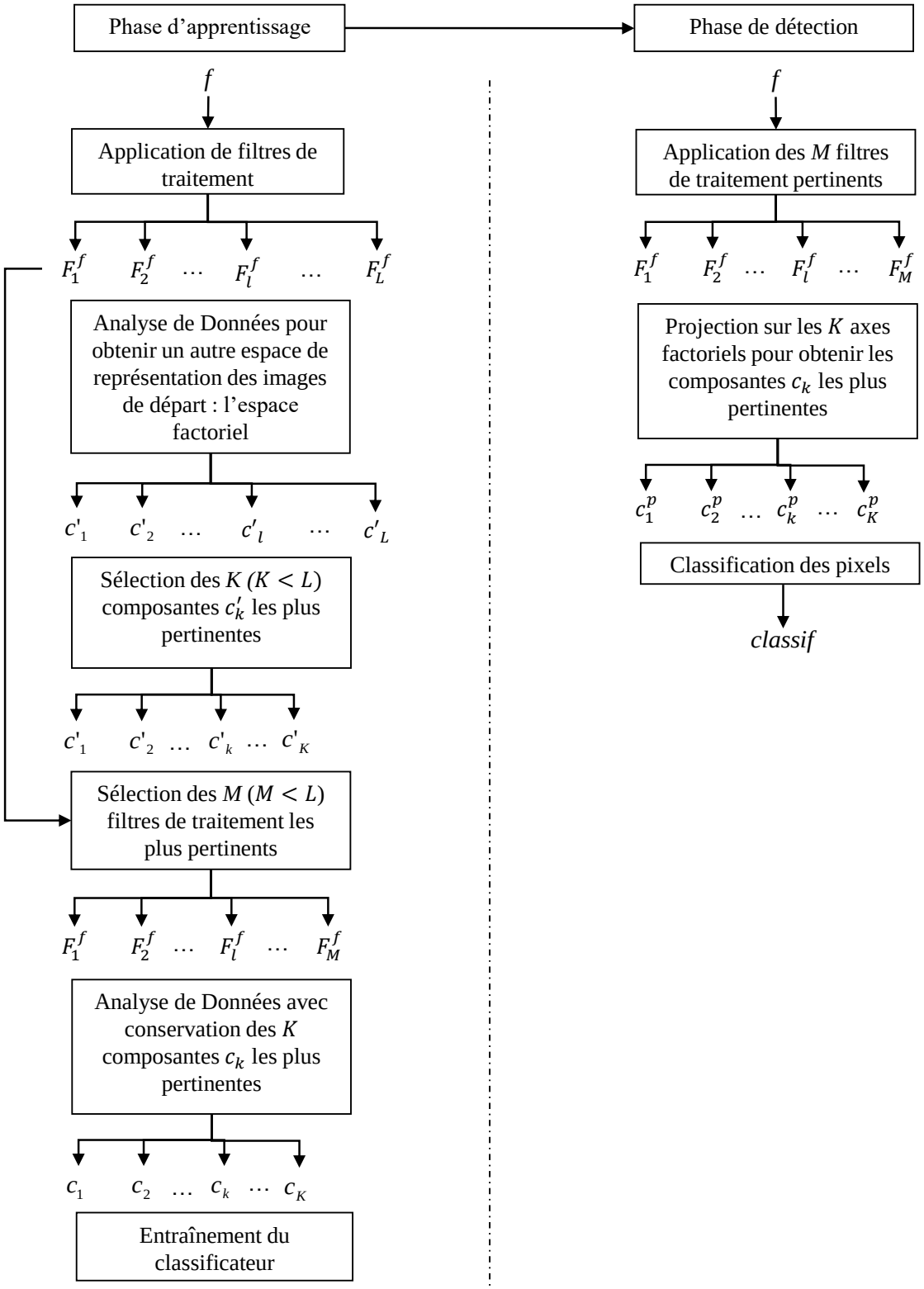


FIG. 3.23 – Méthode de détection des anomalies par analyse de texture.

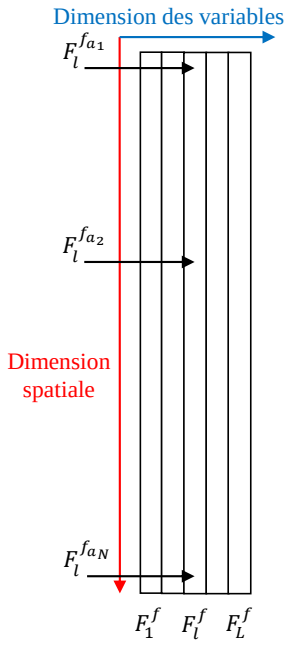


FIG. 3.24 – Regroupement des différentes images multivariées sous forme de tableau de contingence $\mathbf{F}^f$ de taille $N \cdot P \times L$. Le nombre d'images de l'anomalie est noté N, le nombre de pixels d'une image P et le nombre de filtres L.

L'ensemble de la méthode est présenté sur le schéma de la figure 3.23. A l'aide d'une transformée telle que l'Analyse en Composante Principale [108], [121], l'Analyse Factorielle des Correspondances [25] ou l'Analyse en Composantes Indépendantes [108], les composantes factorielles les plus représentatives de l'image multivariée $\mathbf{F}^f$ de la réponse des filtres sont extraites. Ces composantes factorielles sont obtenues par la minimisation d'une métrique caractérisant la dépendance entre les variables ou composantes de l'image $\mathbf{F}^f$. L'analyse de données revient à effectuer une transformation ζ d'un espace image à L dimensions dans un espace à K dimensions, avec K inférieur à L.

$$\zeta = \begin{cases} \mathbb{R}^L & \rightarrow \quad \mathbb{R}^K \mid K \leq L \\ \mathbf{F}^f(x) = (F_1^f(x), \ldots, F_L^f(x)) & \rightarrow \quad \mathbf{c}'(x) = (c'_1(x), \ldots, c'_K(x)) \end{cases} \tag{3.17}$$

On obtient une image multivariée des composantes factorielles $\mathbf{c}' = (c'_1, \ldots, c'_K)$ de plus petite dimension K que la dimension L de l'image multivariée $\mathbf{F}^f$. Chaque composante c'_k de l'image $\mathbf{c}'$ est une combinaison linéaire des composantes F_l^f de l'image des filtres $\mathbf{F}^f$. Le choix du nombre d'axes factoriels est fait selon leur inertie décroissante ou selon leur rapport signal à bruit décroissant comme expliqué dans [212], [219].

Une fois les composantes pertinentes c'_k déterminées, les filtres qui contribuent le plus à rendre les variables indépendantes sont sélectionnés. Seuls ces filtres seront ensuite utilisés. L'intérêt est de réduire le nombre de filtres à calculer tout en obtenant une bonne approximation de l'espace factoriel. Pour ce faire, on mesure la distance entre les axes factoriels Δ_k du nouvel espace et les variables (i.e. les réponses des filtres) représentées sous forme de points dans l'espace factoriel (Fig. 3.25). Les M filtres les plus proches des extrémités des axes factoriels sont alors retenus, où M est inférieur au nombre de filtres initiaux L.

Après avoir réduit le nombre de filtres, l'Analyse de Données est appliquée de nouveau sur l'image $\mathbf{F}^f$ avec M composantes afin d'obtenir l'espace factoriel de dimension réduite et l'image des composantes factorielles $\mathbf{c} = (c_1, \ldots, c_K)$. Sur cette image $\mathbf{c}$ de dimension réduite, on entraîne un classificateur de type *Analyse Discriminante Linéaire* ou *Support Vector Machine* [108] avec des images présentant des anomalies et d'autres ne présentant pas d'anomalies. Ce classificateur va

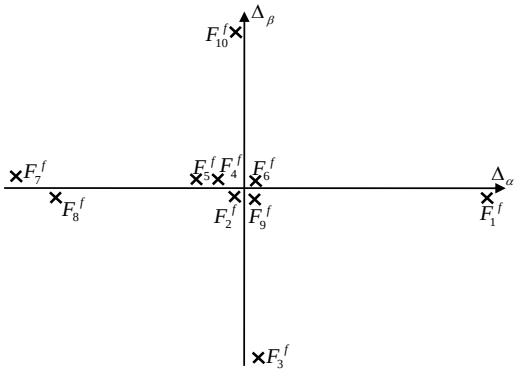


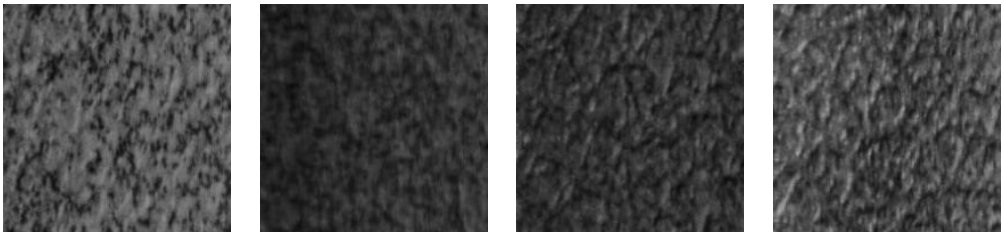
FIG. 3.25 – Espace factoriel à deux dimensions avec les axes factoriels Δ_k et les variables (i.e. les réponses des filtres) F_l^f représentées par des croix.

permettre de détecter la présence (ou l'absence) et la position de l'anomalie sur l'image $\mathbf{c}$. L'ensemble d'entraînement et de validation sont des sous-ensembles de l'ensemble d'apprentissage. Une fois que le classificateur statistique est entraîné, on peut classifier les images de l'ensemble de test en les projetant directement dans l'espace factoriel de l'image $\mathbf{c}$ pour obtenir une image $\mathbf{c}^p$ sur laquelle on applique le classificateur entraîné.

Cette méthode nécessite une annotation des anomalies sur les images au niveau des pixels. Comme cette annotation est difficile à obtenir, il a fallu se contenter d'une annotation des anomalies au niveau de sous-ensembles des images appelés « imagettes ». Il a donc été nécessaire de développer une méthode améliorée d'analyse de texture.

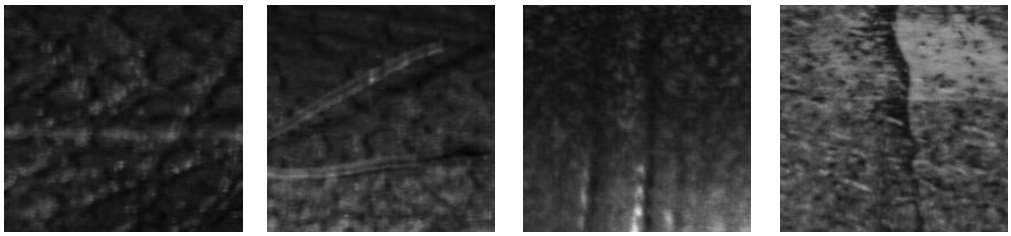
3.3.2 Méthode améliorée d'analyse de texture

La méthode améliorée classifie des imagettes de la surface d'un pneumatique. Ella été développée grâce à une collaboration entre Michelin et l'Ecole des Mines de Paris que j'ai initiée et organisée. Plusieurs chercheurs de l'Ecole des Mines ont participé à l'étude : le Professeur Dominique Jeulin (Directeur de Recherche), le Dr. Michel Bilodeau (Chercheur senior), le Dr. Estelle Parra-Denis (Post-Doc) puis un chercheur et deux autres Post-Docs. Cette collaboration a donné lieu à l'obtention de deux brevets : un sur la méthode améliorée d'Analyse de Texture [225] et un autre sur une méthode rapide de détection du relief [226] qui sera présentée dans la sous-section 3.4.1, page 52, ci-après. Une publication a aussi été faite dans une conférence par les partenaires académiques [241]. Les figures 3.26 et 3.27 illustrent des exemples de textures avec différentes rugosités et orientations. C'est dans ces textures que l'on a recherché des anomalies.



(a) (b) (c) (d)

FIG. 3.26 – Exemples de différentes textures de la moins rugueuse à la plus rugueuse. Source [241].



(a) (b) (c) (d)

FIG. 3.27 – Structures orientées avec différents arrière-plans texturés. Source [241].

Comme les anomalies n'ont pas toujours une frontière clairement définie, il est difficile de faire une classification au niveau des pixels. En revanche, définir si une imagette contient une anomalie ou pas est beaucoup plus aisé. La taille des imagettes est choisie par une étude préalable sur la Surface Elémentaire Représentative (SER) d'une anomalie. La présence d'un fond texturé permet l'utilisation d'une approche probabiliste afin de caractériser les structures aléatoires. La texture est alors vue comme la réalisation d'une fonction aléatoire [128], [177]. Cord et al. [56], [57] ont montré que les outils qui caractérisent les fonctions aléatoires peuvent directement servir pour l'analyse de texture. Sur une texture avec ou sans anomalie, l'imagette doit avoir une taille suffisante pour que la fonction de covariance atteigne la valeur nulle. Cette longueur pour laquelle la covariance devient nulle est appelée portée et correspond à la longueur d'échelle (i.e. la longueur de corrélation) du milieu. La fonction de covariance est définie comme l'espérance statistique du produit de l'image $f : D \to \mathbb{R}$ par son translaté par un vecteur $h \in D$. Le carré de la moyenne μ^2 de l'image est ensuite retranché du produit :

$$C(h) = E[f(x).f(x+h)] - \mu_f^2. \tag{3.18}$$

La covariance normée $Cr : \mathbb{R} \to [-1, +1]$ est définie comme le rapport entre la fonction de covariance et la variance σ_f^2 de l'image f. Les valeurs de la fonction de covariance normée sont comprises dans l'intervalle $[-1; +1]$:

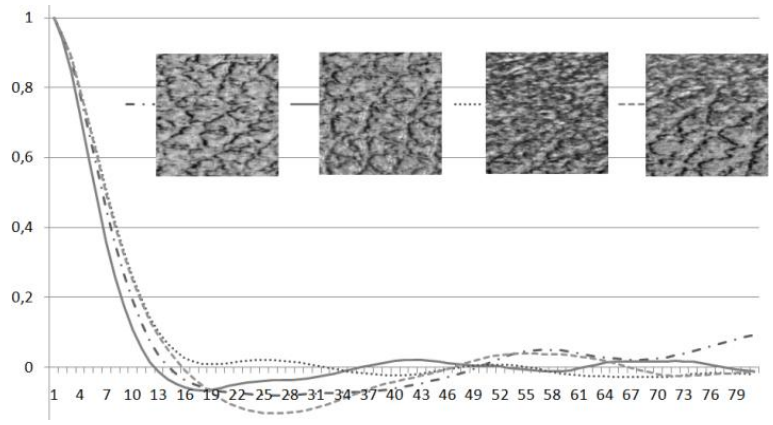
$$Cr(h) = \frac{C(h)}{\sigma_f^2}. \tag{3.19}$$

La figure 3.28 présente des fonctions de covariances normalisées pour des textures stationnaires sans anomalie (Fig. 3.28a) et d'autres avec anomalies (Fig. 3.28b). Pour les imagettes sans anomalies, la covariance décroît rapidement vers zéro, ce qui correspond à la longueur d'échelle de la texture. Par contre, pour les imagettes avec anomalies, la décroissance est beaucoup moins rapide et la longueur d'échelle peut aller jusqu'à 80 pixels. On choisit donc une taille pour les imagettes environ trois fois plus grande que la longueur maximale d'échelle et qui soit une puissance de deux afin d'accélérer les calculs effectués par les processeurs. Une taille de 256×256 pixels comme SER est retenue.

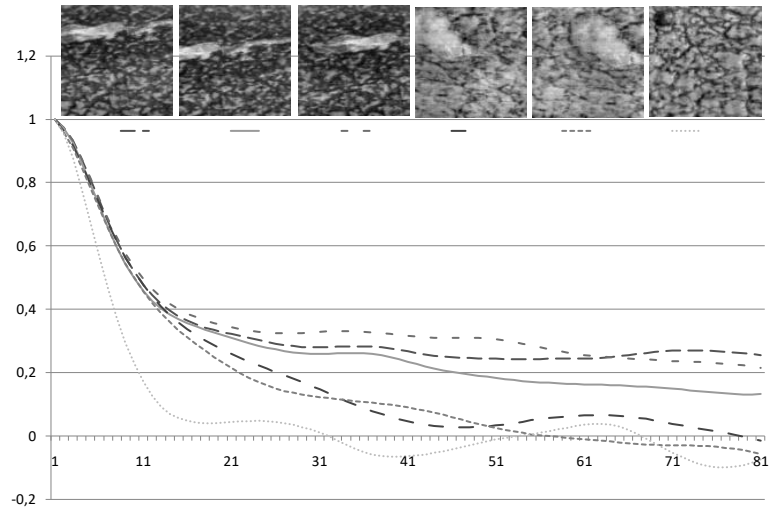
Puis l'image f est découpée en imagettes W_f^k, où $k \in \{1; \ldots; N_{im}\}$ est le numéro de l'imagette et N_{im} le nombre d'imagettes dans l'image f. Le découpage de l'image f en imagettes W_f^k peut être régulier en I_{im} par J_{im} imagettes (Fig. 3.29). Sur chaque imagette $\left\{W_f^k : D_{W^k} \to \mathbb{R}\right\}_{k=1}^{N_{im}}$, des descripteurs statistiques sont alors construits. Pour chaque imagette W_f^k, un descripteur multivarié est alors obtenu :

$$\mathbf{H}^{W_f} : \begin{cases} \left\{\mathbb{R}^{D_{W^k}}\right\}_{k=1}^{N_{im}} & \to & \mathbb{R}^{L_{\mathbf{H}}} \\ W_f^k & \to & \mathbf{H}^{W_f}(x_k), \end{cases} \tag{3.20}$$

où D_{W^k} est l'espace support d'une imagette W_f^k et $L_{\mathbf{H}}$ est le nombre de descripteurs pour la transformation $\mathbf{H}^{W_f}$. A chaque imagette $W_f^k : D_{W^k} \to \mathbb{R}$, le descripteur associé est un vecteur



(a) Covariances pour des imagettes avec une texture stationnaire (sans anomalie).



(b) Covariances pour des imagettes avec anomalies.

FIG. 3.28 – Covariances normalisées selon la direction h horizontale. Source [241].

$\mathbf{H}^{W_f}(x_k)$ à valeurs réelles de longueur $L_{\mathbf{H}}$ à l'emplacement $x_k \subset \mathbb{R}^2$ de l'imagette. Afin d'alléger les notations, on note également $\mathbf{H}^{W_f} : D_{N_{im}} \to \mathbb{R}^{L_{\mathbf{H}}}$ l'image multivariée obtenue en calculant les descripteurs pour chaque imagette (Fig. 3.29). $D_{N_{im}} \subset \mathbb{Z}$ est l'espace support discret avec un cardinal de N_{im}.

Pour cette étude, une grande banque d'environ dix mille imagettes était disponible. Environ dix pourcents des imagettes contenaient des structures orientées de type anormale sur un fond texturé [241]. Chaque imagette était annotée comme normale ou anormale. Nous disposions donc d'une vérité terrain. Cette banque d'imagettes a été découpée en un ensemble d'entraînement et un autre de test.

Sur chaque imagette W_f^k de taille $I \times J$ pixels, plusieurs descripteurs $\mathbf{H}^{W_f}$, $\mathbf{G}^{W_f}$, etc. ont été calculés [225]. En voici leur liste :

- l'histogramme des niveaux de gris d'une imagette. Si l'imagette est à valeurs dans $[\![0; 255]\!] \subset \mathbb{N}$, on obtient un vecteur composé des 256 fréquences de chaque niveau de gris. Pour l'ensemble des imagettes, l'image de la pile de descripteurs est notée $\mathbf{H}^{W_f}$;
- la covariance normalisée horizontale et verticale. Pour l'ensemble des imagettes, les images des piles de descripteurs sont notées $\mathbf{G}^{W_f}$ et $\mathbf{K}^{W_f}$.

Puis on calcule :

- l'écart à la moyenne pour chaque ligne i_n, $\Delta_{i_n}(W_f^k) = E\left[W_f^k(i_n, j) - E[W_f^k], j = [\![1; J]\!]\right]$, et chaque colonne j_n, $\Delta_{j_n}(W_f^k)$;

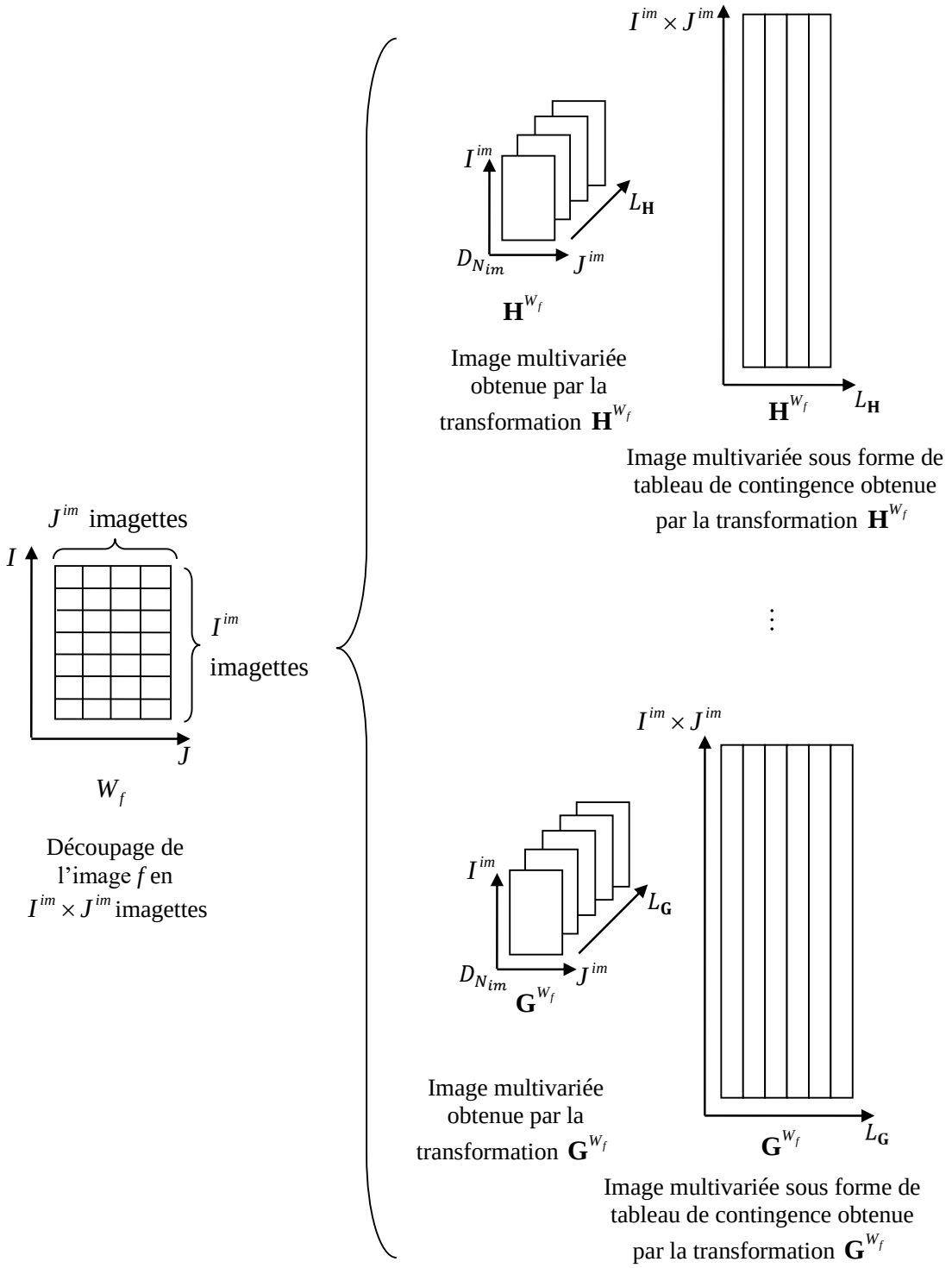


FIG. 3.29 – Création d'images multivariées de différentes dimensions L par application de transformation sur un pavage de l'image de départ en imagettes.

- la variance pour chaque ligne et chaque colonne : $V_{i_n}(W_f^k)$ et $V_{j_n}(W_f^k)$.

Un profil (i.e. un vecteur) est obtenu pour les écarts à la moyenne pour toutes les lignes $\Delta_i(W_f^k)$: $[\![1; J]\!] \to \mathbb{R}$, et pour toutes les colonnes $\Delta_j W(_f^k)$: $[\![1; I]\!] \to \mathbb{R}$. Des profils de variances $V_i(W_f^k)$: $[\![1; J]\!] \to \mathbb{R}$ et $V_j(W_f^k)$: $[\![1; I]\!] \to \mathbb{R}$ sont aussi obtenus pour les lignes et les colonnes. Pour les signaux (i.e. les vecteurs) correspondant à l'écart à la moyenne pour les lignes $\Delta_i(W_f^k)$ et pour les colonnes $\Delta_j W(_f^k)$, on calcule :

- les écart-types $\sigma[\Delta_i(W_f^k)]$ et $\sigma[\Delta_j W(_f^k)]$;
- les extrema : minima et maxima : $\min[\Delta_i(W_f^k)]$, $\max[\Delta_i(W_f^k)]$, $\min[\Delta_j W(_f^k)]$ et $\max[\Delta_j W(_f^k)]$;
- les écarts à l'amplitude maximale :

$$\zeta[\Delta_i(W_f^k)] = 2\sigma[\Delta_i(W_f^k)] - \Big|\max[\Delta_i(W_f^k)] - \min[\Delta_i(W_f^k)]\Big|, \qquad (3.21)$$
$$\zeta[\Delta_j W(_f^k)];$$

- les traversées moyennes positives pour un seuil fixé à $s = 2\sigma[\Delta_i(W_f^k)]$:

$$tr_{s=2\sigma[\Delta_i(W_f^k)]}[\Delta_i(W_f^k)] = \frac{card(T_s[\Delta_i(W_f^k)])}{card\Big(connex(T_s[\Delta_i(W_f^k)])\Big)}, \qquad (3.22)$$
$$tr_{s=2\sigma[\Delta_j W(_f^k)]}[\Delta_j W(_f^k)],$$

  où $T_s(f)$ est le signal binaire obtenu par seuillage du signal f à la valeur s, $card$ le cardinal d'un ensemble discret et $connex$ les composantes connexes non nulles du signal binaire obtenu par seuillage ;
- les traversées moyennes négatives pour un seuil fixé à $s = -2\sigma[\Delta_i(W_f^k)]$: $tr_{s=-2\sigma[\Delta_i(W_f^k)]}[\Delta_i(W_f^k)]$ et $tr_{s=-2\sigma[\Delta_j W(_f^k)]}[\Delta_j W(_f^k)]$.

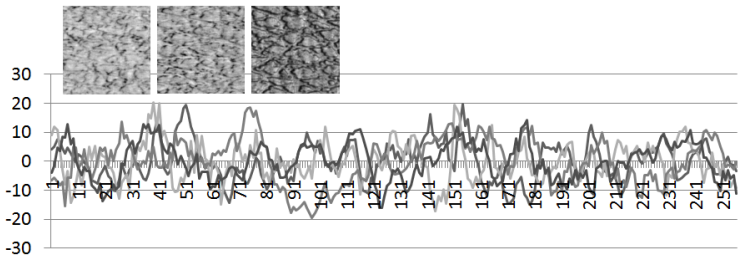
Pour les profils (i.e. les vecteurs) correspondant à la variance à la moyenne pour chaque ligne $V_i(W_f^k) : [\![1; J]\!] \to \mathbb{R}$, et pour chaque colonne $V_j(W_f^k) : [\![1; I]\!] \to \mathbb{R}$, on calcule :

- les écart-types $\sigma[V_i(W_f^k)]$ et $\sigma[V_j(W_f^k)]$;
- les extrema : minima et maxima : $\min[V_i(W_f^k)]$, $\max[V_i(W_f^k)]$, $\min[V_j(W_f^k)]$ et $\max[V_j(W_f^k)]$;
- les écarts à l'amplitude maximale : $\zeta[V_i(W_f^k)]$ et $\zeta[V_j(W_f^k)]$;
- les traversées moyennes positives $tr_{s=2\sigma[V_i(W_f^k)]}[V_i(W_f^k)]$ et $tr_{s=2\sigma[V_j(W_f^k)]}[V_j(W_f^k)]$;
- les traversées moyennes négatives $tr_{s=-2\sigma[V_i(W_f^k)]}[V_i(W_f^k)]$ et $tr_{s=-2\sigma[V_j(W_f^k)]}[V_j(W_f^k)]$.

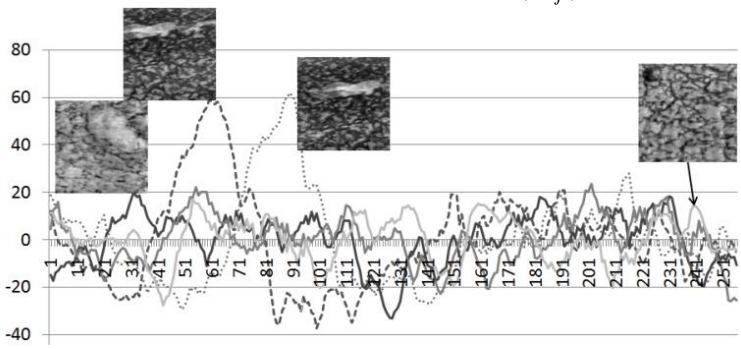
L'ensemble des descripteurs sur les lignes et sur les colonnes sont ensuite regroupés. Pour chaque imagettes, les vecteurs de descripteurs sont mis bout à bout afin de former un seul vecteur de descripteurs. L'image multivariée résultante est notée $\mathbf{L}^{W_f}$. La figure 3.30 compare les profils des écarts à la moyenne sur les lignes avec ou sans anomalie. On constate que les profils des imagettes avec anomalies sont bien différents de ceux sans anomalie

Afin de créer de nouveaux descripteurs, sur chaque imagette des « pyramides » d'érosions ε_B, dilatation δ_B, ouverture γ_B et fermeture φ_B sont calculées avec des éléments structurants en forme de segments orientés selon l'axe horizontal puis vertical et avec des tailles croissantes [241]. Sur ces transformées d'imagettes, les descripteurs précédemment cités sont ensuite appliqués. Les figures 3.31 et 3.32 illustrent les piles de dilatations et de fermetures. L'image multivariée des descripteurs est notée $\mathbf{M}^{W_f}$ pour les érosions et dilatations et celles pour les ouvertures fermetures $\mathbf{N}^{W_f}$.

Chaque image multivariée est ensuite transformée en un tableau de contingence en rangeant les pixels en colonnes et leurs descripteurs en lignes comme illustré sur la figure 3.29. Les tableaux de contingences vont servir à l'Analyse de Données. Sur chaque tableau, on applique une Analyse en Composante Principale (ACP) afin de déterminer les facteurs principaux qui maximisent l'énergie du nuage [108], [121], [212]. Pour les tableaux de la covariance $\mathbf{G}^{W_f}$ et $\mathbf{K}^{W_f}$, on passe de 80 descripteurs à 3 descripteurs. Pour le tableau des profils sur la pyramide des érosions et dilatations $\mathbf{M}^{W_f}$ on passe de 1760 à 10 descripteurs. On obtient la même réduction pour le tableau des profils

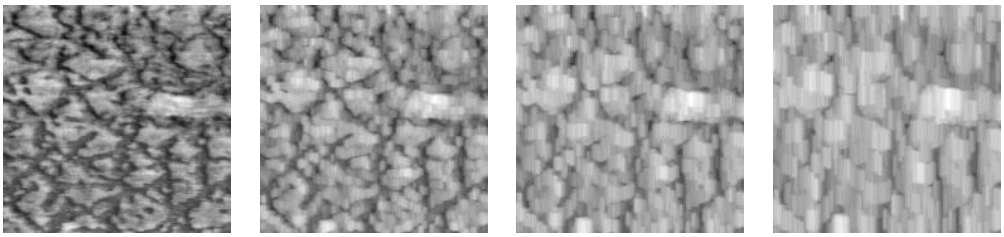


(a) Profils des écarts à la moyenne sur les lignes $\Delta_i(W_f^k)$ (sans anomalies).



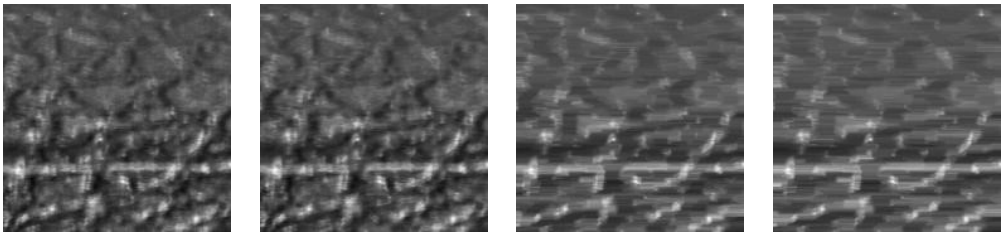
(b) Profils des écarts à la moyenne sur les lignes $\Delta_i(W_f^k)$ (avec anomalies).

FIG. 3.30 – Comparaison des profils des écarts à la moyenne sur les lignes $\Delta_i(W_f^k)$: (a) avec ou (b) sans anomalie. Source [241].



(a) (b) (c) (d)

FIG. 3.31 – Dilatations par un élément structurant vertical de taille croissante. Source [241].



(a) (b) (c) (d)

FIG. 3.32 – Fermetures par un élément structurant horizontal de taille croissante. Source [241].

de la pyramide d'ouvertures et de fermetures $\mathbf{N}^{W_f}$. Chaque tableau de contingence $\mathbf{H}^{W_f}$, $\mathbf{G}^{W_f}$, ..., $\mathbf{N}^{W_f}$ donne donc une image multivariée dans l'espace factoriel qui est notée respectivement $\mathbf{c}^{\mathbf{H}^{W_f}} : D_{N_{im}} \rightarrow \mathbb{R}^{L_{\mathbf{H}}}$, $\mathbf{c}^{\mathbf{G}^{W_f}}$, ... $\mathbf{c}^{\mathbf{N}^{W_f}}$ [241]. Comme les espaces factoriels sont des espaces normés dont la métrique est euclidienne, les tableaux de contingence sont regroupés en un seul, noté $\mathbf{c}^{W_f}$, par concaténation selon l'axe des variables. Chaque imagette est ainsi décrite par un ensemble de 26 descripteurs (obtenus par combinaison linéaire des descripteurs initiaux). La tableau $\mathbf{c}^{W_f}$ a donc une taille de $N_{im} \times 26$. Sur ce tableau de contingence $\mathbf{c}^{W_f}$, un classificateur supervisé de type Analyse Discriminante Linéaire [108] est alors entraîné.

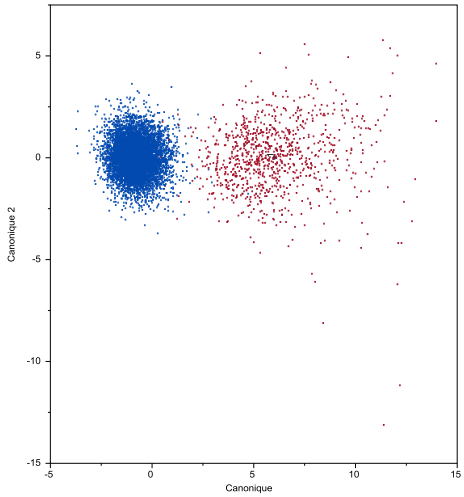


FIG. 3.33 – Classification par Analyse Discriminante Linéaire sur les deux axes principaux de l'espace factoriel des descripteurs. Chaque point représente une imagette. En rouge les imagettes avec anomalies et en bleu celles sans anomalie. Source [241].

Sur les imagettes de test, les mêmes descripteurs sont calculés afin d'obtenir un tableau de contingence similaire qui est ensuite projeté dans l'espace du classificateur. La classification obtenue a une sensibilité Se de $99,45\%$ et une spécificité Sp de $99,61\%$. La sensibilité ou taux de vrais positifs est définie comme le rapport des vrais positifs sur l'ensemble des échantillons positifs. La spécificité ou taux de vrais négatifs est définie comme le rapport des vrais négatifs sur l'ensemble des échantillons négatifs. La figure 3.33 illustre la classification par Analyse Discriminante Linéaire en imagettes avec et sans anomalie dans l'espace factoriel des descripteurs. Une séparation nette des imagettes avec et sans anomalies en deux nuages est observée. La méthode développée a ensuite été testée dans un environnement industriel.

3.4 L'accélération d'algorithmes

L'accélération d'algorithmes est nécessaire pour analyser les grandes banques d'images. Du fait de leur taille, l'utilisation d'algorithmes de traitement d'images sur ces banques est souvent limitée par les temps de calcul. Par exemple, pour le contrôle des pneumatiques, des anomalies sont recherchées dans de très grandes images qui sont aussi très nombreuses. Sans accélération, l'exécution d'un programme même assez simple devient vite très longue. Il existe plusieurs axes pour réduire le temps de calcul dans un programme. Un premier axe consiste à utiliser une autre méthode moins gourmande en quantité de données à traiter. C'est l'axe indispensable dans l'accélération d'un programme et qui permet de faire les gains les plus importants. Un deuxième axe consiste à accélérer un algorithme existant en le programmant d'une manière différente qui utilise mieux les capacités des processeurs. Cet axe permet lui aussi des gains significatifs, mais dans un deuxième temps. J'ai eu l'occasion de travailler à deux reprises à l'accélération d'algorithmes : chez Michelin pour le contrôle industriel des pneumatiques et à l'IPRI sur des algorithmes morphologiques et de segmentation.

3.4.1 Segmentation rapide d'anomalies pour le contrôle industriel

Comme la surface interne des pneumatiques présente des reliefs 3D, il est nécessaire de les segmenter afin de pouvoir les contrôler. J'ai créé un algorithme d'extraction des stries présenté dans la sous-section 3.2.2, page 36, ainsi qu'un autre qui segmente des anomalies en relief. Leur principe repose sur des *top-hats filtrés* ou des filtres alternés séquentiels. Par choix de l'industriel, ils n'ont

pas donné lieu à des brevets. Comme expliqué précédemment, ils ne seront pas présentés dans ce mémoire, même s'ils sont pleinement fonctionnels et largement validés. Ces algorithmes développés sur une année et demie m'ont permis d'initier et d'organiser une collaboration avec l'Ecole des Mines de Paris afin d'accélérer les traitements. Cette accélération était essentielle à cause de la taille des images qui pouvait aller jusqu'à $40\,000 \times 5\,000$ pixels [214]. Cette collaboration a impliqué le Professeur Dominique Jeulin (Directeur de Recherche), le Dr. Michel Bilodeau (Chercheur senior), le Dr. Estelle Parra-Denis (Post-Doctorante) ainsi qu'un autre chercheur le Dr. Petr Dokladal et deux autres Post-Doctorants. Un brevet a été obtenu [226] sur la méthode. Les partenaires académiques ont aussi fait une publication dans un congrès [33]. Les résultats de ce papier seront reproduits ci-après.

Les détecteurs de type *top-hat* présentés précédemment ont l'inconvénient d'être coûteux en temps de calcul car ils nécessitent pour chaque pixel de l'image d'explorer tous les pixels de son voisinage. Avec des éléments structurants en forme de segments horizontaux de plusieurs centaines ou plusieurs milliers de pixels, le nombre de pixels à explorer représente donc une multiplication du nombre de pixels de l'image par plusieurs centaines ou plusieurs milliers. Le temps de calcul augmente donc significativement. Une des premières solutions que j'ai testée a été la bibliothèque *Fast Morph-M* du Centre de Morphologie Mathématique de l'Ecole des Mines de Paris qui utilisait le calcul vectoriel et les instructions SIMD (Single Instruction on Multiple Data) des processeurs [55]. Cette bibliothèque m'a permis de travailler pendant un an et demi. La collaboration avec l'Ecole des Mines a ensuite permis de proposer une solution encore plus rapide grâce à un nouvel algorithme qui explore beaucoup moins de points dans l'image. Cet algorithme est fondé sur les gradients multipoints qui ne nécessitent qu'un calcul sur quelques voisins plus ou moins éloignés. Le choix de la distance entre ces voisins est en lien avec les objets que l'on veut détecter et ceux que l'on veut supprimer.

Soit $g : D \to \mathbb{R}$ une fonction (ou profil) à valeurs réelles. Le domaine D est un sous-ensemble de l'axe des réels $\mathbb{R}$. Trois types de gradients multipoints ont été proposés par Jeulin [130] avec $a > b > 0$:

— le 2-points qui correspond à un gradient calculé entre deux points de distance a :

$$f(x, a) = |g(x) - g(x + a)| \; ; \tag{3.23}$$

— le 4-points :

$$f(x, a, b) = |g(x) - g(x + b)| - |g(x + a) - g(x + a + b)| \tag{3.24}$$

qui correspond à la comparaison de deux gradients entre deux points x et $x + a$ distants de a et entre leurs voisins $x + b$ et $x + a + b$ distants de b. Plus la différence des points avec leurs voisins est grande, plus la réponse de l'opérateur 4-points est marquée ;

— le 5-points où $p \in [0; 1]$:

$$\begin{aligned} f(x, a, b, p) = p\,(|g(x - a) - g(x - b)| + |g(x + a) - g(x + b)|) \\ - (1 - p)\,(|g(x) - g(x - b)| + |g(x) - g(x + b)|) \; ; \end{aligned} \tag{3.25}$$

qui consiste à comparer le gradient entre les voisins a (respectivement $-a$) et b (resp. $-b$) du point x avec le gradient entre le point x et son voisin b (resp. $-b$).

L'opérateur 5-points peut être interprété de la manière suivante en monodimensionnel (1D). L'explication peut être aussi étendue au cas bidimensionnel. A chaque point x de l'image, on considère les 4 points translatés $\{x + a\}$, $\{x - a\}$, $\{x + b\}$ et $\{x - b\}$. Supposons que x soit l'abscisse du sommet d'une perturbation que l'on cherche à extraire comme dans la figure 3.34. Dans ce but, l'entrée et la sortie de la perturbation vont être prises en compte ainsi que son sommet. L'entrée et la sortie de la perturbation doivent correspondre à un gradient élevé du profil g. Cette information est accessible à partir des écarts $|g(x - a) - g(x - b)|$ et $|g(x + a) - g(x + b)|$. A contrario, le

voisinage du sommet de la perturbation doit correspondre à un gradient plus faible de g qui est accessible à partir des écarts $|g(x) - g(x-b)|$ et $|g(x) - g(x+b)|$. La position x du sommet de la perturbation peut être détectée en maximisant l'expression :

$$f_5(x,a,b) = (|g(x-a) - g(x-b)| + |g(x+a) - g(x+b)|) \\ - (|g(x) - g(x-b)| + |g(x) - g(x+b)|)\,. \quad (3.26)$$

Afin de modifier le contraste entre le gradient d'entrée/sortie et le gradient du sommet, on pondère par les poids p et $1-p$, où $p \in [0;1]$, pour obtenir l'équation (3.25).

Comme pour le *top-hat* calculé ligne à ligne, ces opérateurs ne sont pas sensibles à une forme bombée (ou courbure) selon la direction radiale (i.e. la direction verticale sur l'image 3D de la surface). A la différence du *top-hat*, l'opérateur à 5 points extrait simultanément les perturbations en creux ou en relief que l'on peut séparer a posteriori à partir du signe de $g(x) - g(x-b)$ et de $g(x) - g(x+b)$. Pour réaliser cette opération avec le *top-hat*, il faudrait calculer le *white-top-hat* et le *black-top-hat* ; ce qui serait encore plus coûteux en temps de calcul.

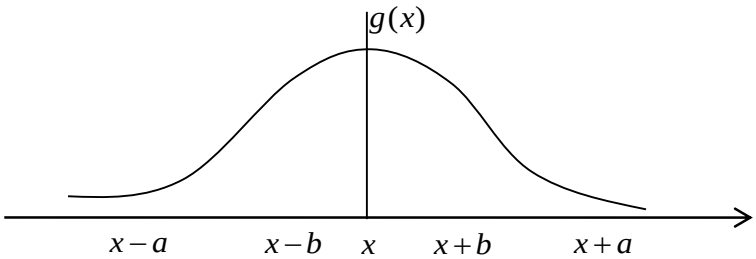


FIG. 3.34 – Perturbation sur un profil unidimensionnel.

L'équation (3.26) peut aussi s'exprimer à l'aide de quatre convolutions avec une matrice éparse :

$$f_5(x,a,b) = |g * K_{a,b}| + |g * K_{-a,-b}| - |g * K_b| - |g * K_{-b}|\,, \quad (3.27)$$

où K_{xx} sont les noyaux éparses composés des poids suivants :

$$K_{-a,-b} = (1 \ldots\ldots -1 \ldots\ldots 0 \ldots\ldots\ldots\ldots),$$
$$K_{-b} = (0 \ldots\ldots -1 \ldots\ldots 1 \ldots\ldots\ldots\ldots).$$

Les autres noyaux sont obtenus par transposition. Les poids peuvent être remplacées par p ou $1-p$ afin d'obtenir l'équation (3.25).

La figure 3.35 illustre le détecteur 5-points de l'équation (3.26) avec $a = 40$, $b = 8$. Le signal original en bleu est composé de trois gaussiennes avec des largeurs σ de 21 ; 50 et 100. La réponse du filtre apparaît en noir. Le paramètre $a = 40$ est choisi pour être un peu plus large que la largeur du pic de gauche. On observe que le détecteur réagit bien pour ce pic. Afin d'améliorer la détection du pic de droite, on augmente le paramètre a à la valeur $a = 100$ sur la figure 3.36d. Par contre, la réponse sur le pic de gauche s'élargit à cause du dôme central. Les figures 3.36a, 3.36b et 3.36c, montrent la réponse du filtre pour chacune des gaussiennes constituant le signal. Afin de réduire l'influence du dôme central, on va augmenter le paramètre b à la valeur $b = 50$ comme l'illustre la figure 3.37.

Le détecteur peut être appliqué ligne par ligne sur une image 3D de surface comme illustré sur la figure 3.38a où la bosse à détecter est entourée de rouge. La figure 3.38b montre la réponse du *Laplacian of Gaussian* (*LoG*) [174] définie pour une image $f : D \to \mathbb{R}$, où $D \subset \mathbb{R}^2$, comme :

$$LoG = \sigma^2 \nabla^2 G_\sigma * f, \quad (3.28)$$

où ∇^2 est le Laplacien. G_σ est une gaussienne de largeur σ définie comme :

$$G_\sigma = \frac{1}{2\pi\sigma^2} \exp^{\frac{x^2+y^2}{2\sigma^2}}\,. \quad (3.29)$$

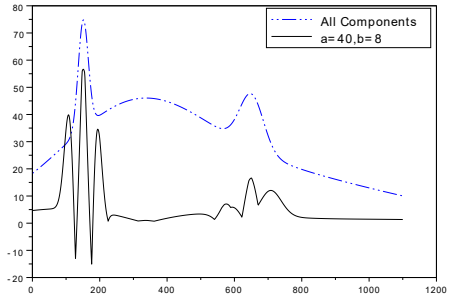
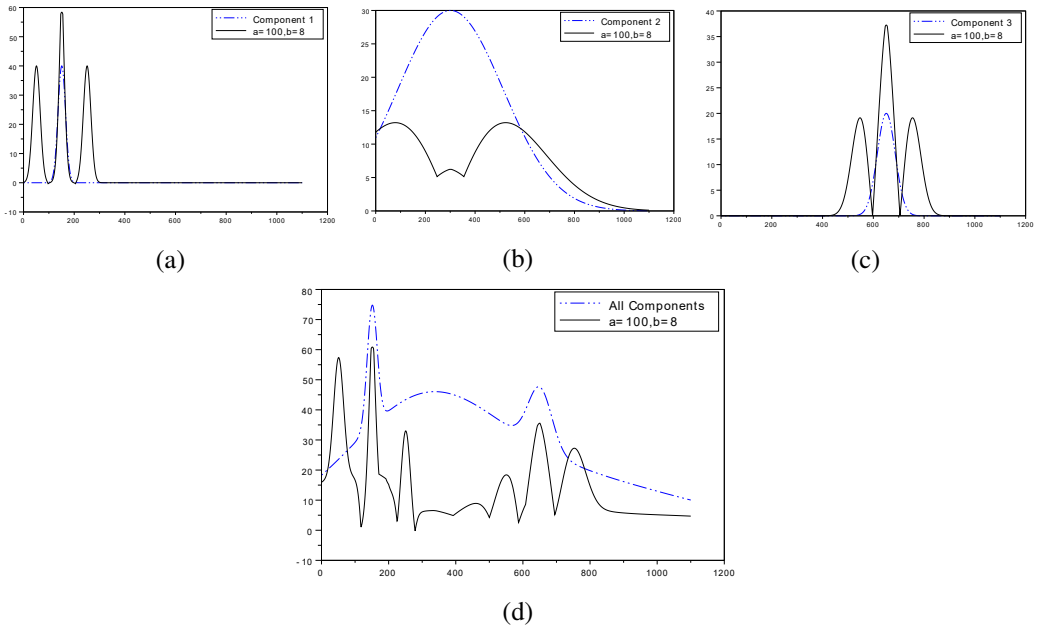


FIG. 3.35 – Détecteur 5-points $f_5(x, a, b)$ avec $a = 40$ et $b = 8$. Le signal est en bleu et la réponse du filtre en noir. Source [33].

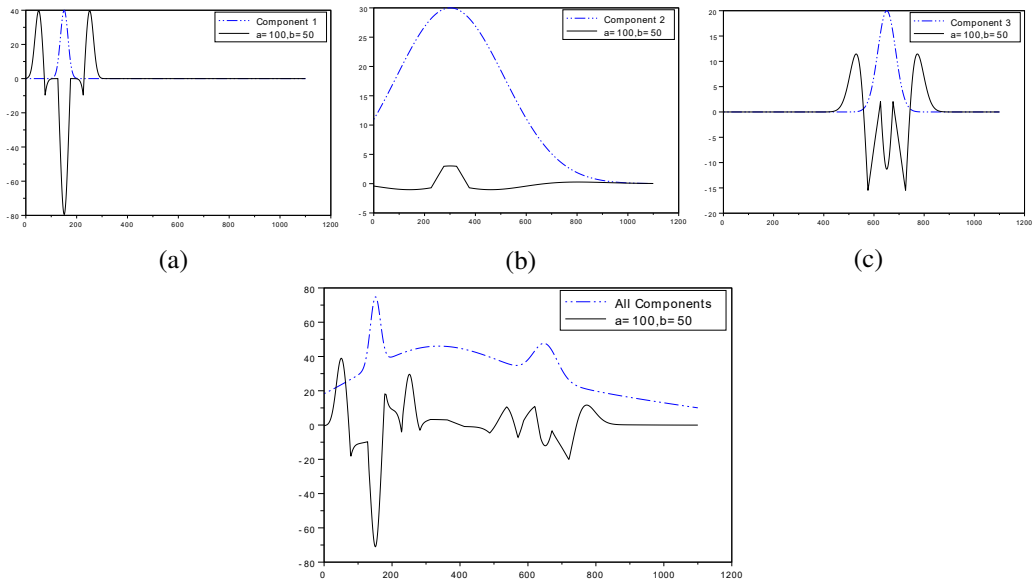


(a) (b) (c)

(d)

FIG. 3.36 – Détecteur 5-points $f_5(x, a, b)$ avec $a = 100$ et $b = 8$. Les signaux sont en bleu et les réponses du filtre en noir. En haut : détecteur appliqué sur chacune des gaussiennes (avec des largeur σ de 21 ; 50 et 100) composant le signal du bas. Source [33].

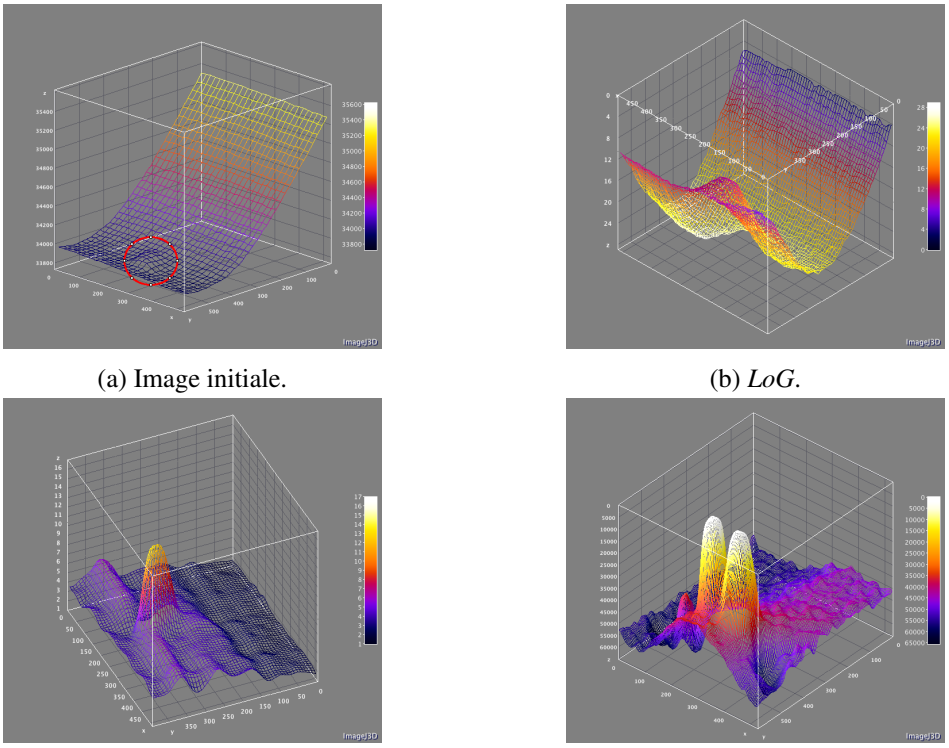
Avec le *LoG* la bosse est un peu plus visible mais toujours difficile à détecter. La figure 3.38c montre que le *top-hat* extrait bien la bosse mais avec un temps de calcul beaucoup plus long qu'un détecteur 5-points. Ce dernier détecte très bien la bosse (Fig. 3.38d).



(a) (b) (c)

(d)

FIG. 3.37 – Détecteur 5-points $f_5(x, a, b)$ avec $a = 100$ et $b = 50$. Les signaux sont en bleu et les réponses du filtre en noir. En haut : détecteur appliqué sur chacune des gaussiennes composant le signal du bas. Source [33].

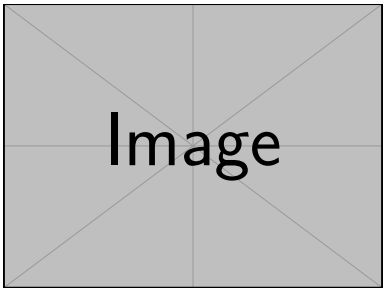


(a) Image initiale. (b) *LoG*.

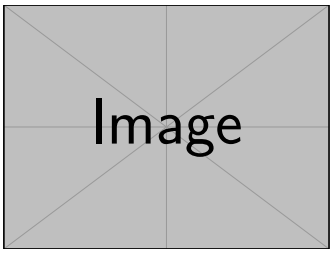
(c) *Top-hat*. (d) Détecteur 5-points.

FIG. 3.38 – Surface 3D des résultats. (a) Image initiale où la bosse à détecter est entourée en rouge. (b) *LoG* avec $\sigma = 100$. (c) Transformée par *top-hat* avec un segment de longueur 100 comme élément structurant. (d) Détecteur 5-points $f_5(x, a, b)$ avec $a = 200$ et $b = 100$. Source [33].

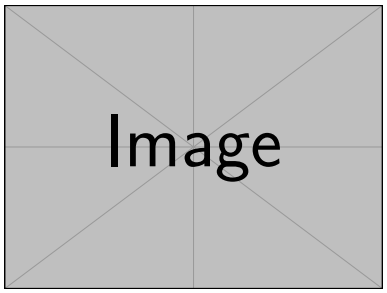
Les détecteurs multipoints ne sont efficaces qu'avec un choix judicieux des paramètres a et b. Avec le détecteur bipoints (éq. (3.23)), le paramètre a sera égal à la distance inter-stries. Pour le détecteur quadri-points (éq. (3.24)), les paramètres a et b devront être plus grands que la largeur d'une strie. Pour le paramètre a, on choisira par exemple quatre fois la largeur d'une strie et pour le paramètre b deux fois la largeur d'une strie. Pour le détecteur 5-points (éq. (3.25)), les paramètres a et b auront des valeurs proches de la largeur des stries afin de les faire ressortir. Des valeurs plus grandes que la largeur d'une strie pourront aussi être choisies afin de faire ressortir les anomalies de taille plus grande et de pente moins forte. La figure 3.39 illustre le choix des paramètres des détecteurs 4-points et 5-points pour une anomalie et une strie sur une image 3D de la surface intérieure d'un pneumatique.



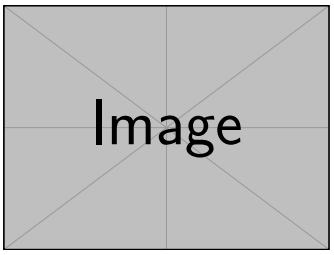
(a) Image initiale.



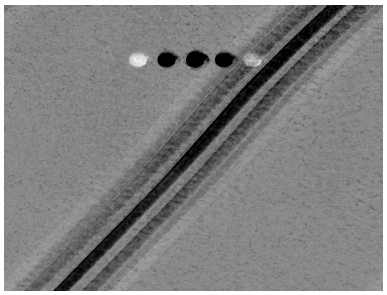
(b) 4-points : $a = 50$, $b = 25$.



(c) 4-points : $a = 200$, $b = 100$.



(d) 5-points : $a = 50$, $b = 25$, $p = 1/5$.



(e) 5-points : $a = 200$, $b = 100$, $p = 1/5$.

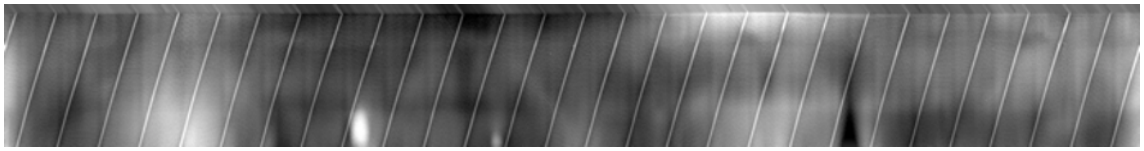
FIG. 3.39 – (a) Image 3D de la surface initiale avec une strie et une anomalie circulaire. (b) et (c) Détecteur 4-points avec différents paramètres. (d) et (e) Détecteur 5-points avec différents paramètres. Source [226].

La figure 3.40, montre l'application du détecteur 5-points à l'ensemble de l'image. Lorsque le profil par moyenne est calculé sur les colonnes du détecteur, seules certaines anomalies apparaissent (Fig. 3.40d). D'autres anomalies sont masquées par les creux du signal causés par les stries. Lorsque le profil est calculé sur l'image seuillée, les anomalies deviennent plus visibles (Fig. 3.40e). Cependant des reliefs normaux peuvent aussi être détectés. Afin d'améliorer la détection sur ce profil, il est nécessaire de démoduler les variations qui sont causées par les stries. L'information de périodicité due aux stries apparaît comme une fréquence porteuse qu'il convient de démoduler pour récupérer le reste du signal qui contient les anomalies. Deux techniques simples

de démodulation ont été testées.

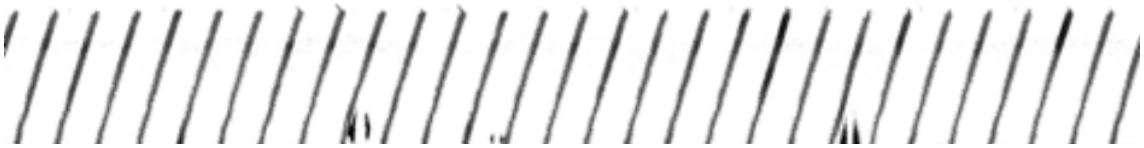
- La première consiste à soustraire au profil, un filtrage passe-bas de ce profil qui ne laisse passer que le signal de modulation (i.e. le signal correspondant aux stries).
- Lorsque la périodicité des stries est quasiment parfaite, la deuxième méthode consiste à soustraire au profil l'infimum du profil avec son translaté. La démodulation g est obtenue pour tout $x \in D$ par l'opérateur $g(x) = f(x) - \wedge(f(x), f(x-a))$. Le vecteur $a \in D$ représente la translation correspondant à la période des stries.

Sur le profil de la figure 3.41a, l'anomalie est entourée en rouge et les reliefs normaux en vert. Ce signal est ensuite démodulé par décalage (Fig. 3.41b) ou par filtrage passe-bas (Fig. 3.41c). Dans les deux cas, l'anomalie (en rouge) ressort clairement. Elle sera ensuite séparée des autres reliefs normaux (en vert), grâce à une analyse des pics par d'autres filtres.

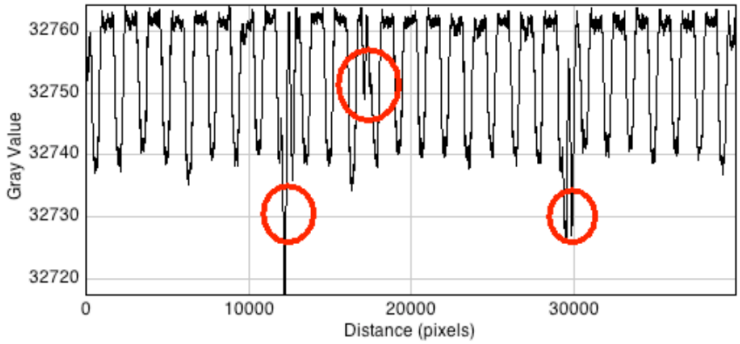


(a) Image initiale en 3D de la surface.

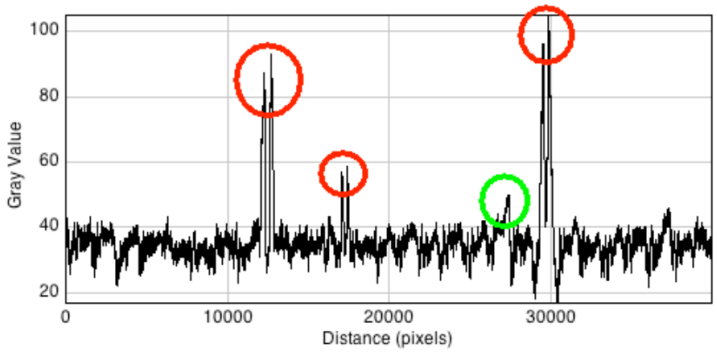
(b) Détecteur 5-points : $a = 50$, $b = 25$, $p = 1/2$.



(c) Seuillage de l'image du détecteur 5-points.

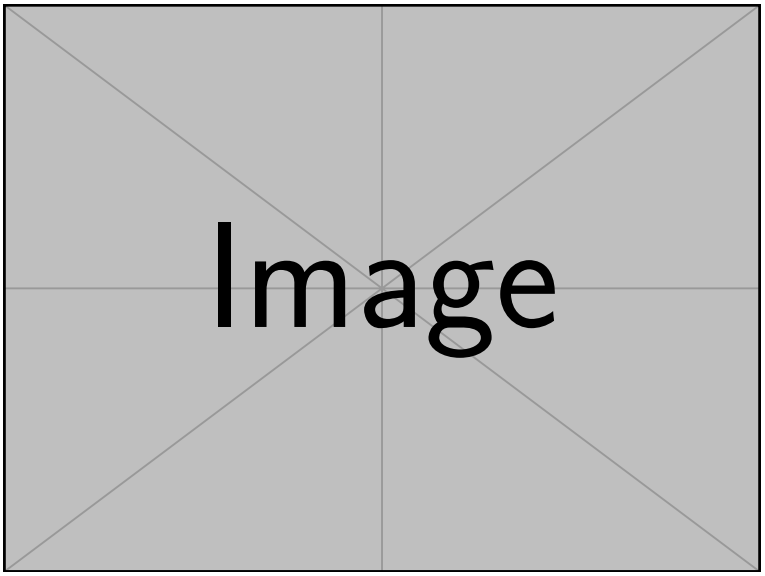


(d) Profil (sur les colonnes) du détecteur 5-points (b). Les anomalies sont entourées en rouge.

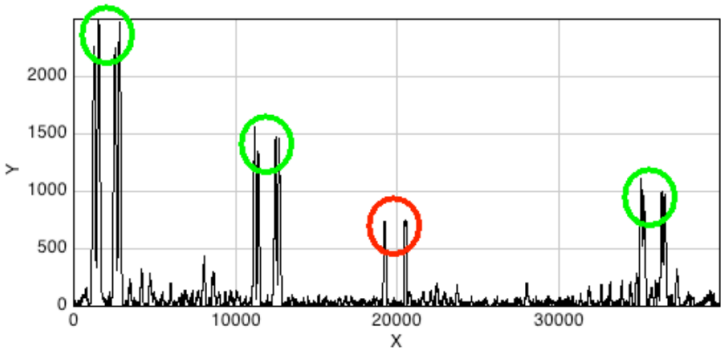


(e) Profil (sur les colonnes) du détecteur 5-points seuillé (c). Les anomalies sont entourées en rouge et les reliefs normaux en vert.

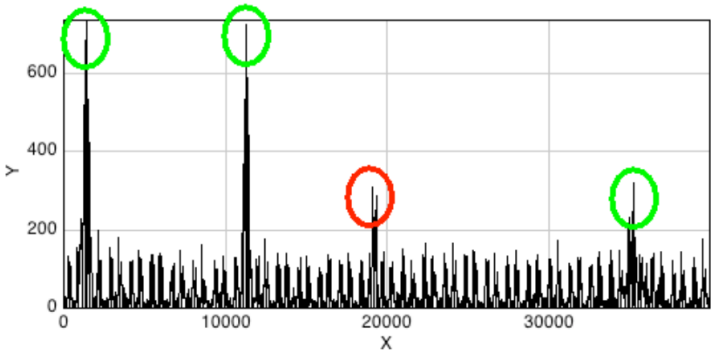
FIG. 3.40 – (a) Image originale. (b) Sortie d'un 5-points. (c) Seuillage du 5-points. (d) Profil des moyennes calculé sur le tiers bas de (b). (e) Profil du nombre de pixels par colonne calculé sur l'image seuillée (c). Source [226].



(a) Profil (sur les colonnes) du détecteur 5-points (b).



(b) Démodulation par décalage $g(x) = f(x) - \wedge(f(x), f(x-a))$.



(c) Démodulation par filtrage passe-bas.

FIG. 3.41 – (a) Profil colonnes du détecteur 5-points d'une image 3D de surface. (b) Démodulation par décalage. (c) Démodulation par filtrage passe-bas. Les reliefs normaux sont entourés en vert et l'anomalie en rouge. Source [226].

3.4.2 Algorithmes rapides en Morphologie Mathématique

Dès le début de mon travail de recherche à l'IPRI sur l'analyse des images du fond d'œil pour le suivi de la rétinopathie diabétique, il m'est apparu nécessaire d'avoir des algorithmes rapides pour analyser ces images. En particulier, les métriques d'Asplund couleurs et multivariées [209] (cf. section 4.3, p. 102) sont vites apparues très longues à calculer lorsque qu'elles ont été programmées par approche directe. Comme ces métriques ont un lien étroit avec la morphologie mathématique, je me suis intéressé à l'accélération des algorithmes morphologiques. J'ai commencé seul puis j'ai encadré durant une année un jeune ingénieur-chercheur, Pierre Guilbert, de mi-2015 à mi-2016.

Calcul parallèle et vectoriel pour la morphologie

Etant donné un ensemble compact et convexe B d'un espace euclidien $\mathcal{E}$, Matheron a montré que B est infiniment divisible par rapport à la dilatation [177, chap 1, p. 22] :

$$B = \left(\frac{1}{k}B\right) \oplus \left(\frac{1}{k}B\right) \oplus \cdots \oplus \left(\frac{1}{k}B\right) (k \text{ facteurs}), \tag{3.30}$$

où aB est l'homothétique de B par un facteur $a \in \mathbb{R}$. L'accélération d'opérations morphologique élémentaires est donc essentielle. De nombreux algorithmes rapides existaient mais en 2014-15 peu de bibliothèques de Morphologie Mathématique utilisaient la méthode de programmation multicœurs parallèle *OpenMP* (Open Multi-processing) [49] et le calcul vectoriel des processeurs grâce aux instructions SIMD (Single Instruction on Multiple Data) [55]. Nous allons regarder comment utiliser ces méthodes de programmation sur les opérations morphologiques.

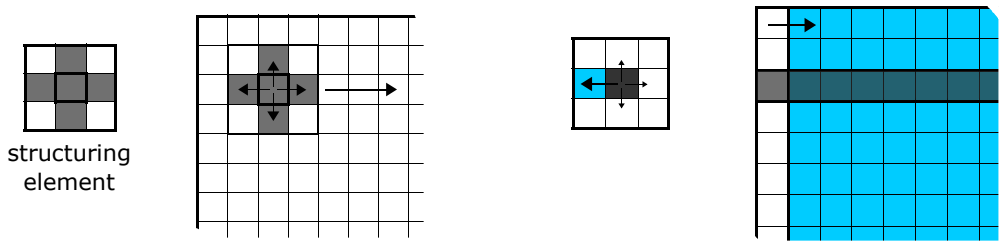
La dilatation morphologique correspond à l'addition de Minkowski

$$X \oplus B = \bigcup_{b \in B} X_b = \{x + b : x \in X, b \in B\}. \tag{3.31}$$

$X_b = \{x + b : x \in X\}$ est l'ensemble $X \subset \mathcal{E}$ translaté par le vecteur b. Comme illustré sur la figure 3.42a et dans l'algorithme 1, une mise en œuvre (i.e. une *implementation*) directe par calcul de l'union sur le voisinage de chaque pixel aura une complexité en $\mathcal{O}(N \times |B|)$ [86]. N est le nombre de pixels de l'image et $|B|$ le cardinal de l'ensemble B. Cependant, l'addition de Minkowski (i.e. la dilatation) a la propriété d'être distributive sur l'union [100], [275] : $X \oplus (B \cup B') = (X \oplus B) \cup (X \oplus B')$. Cette propriété a des conséquences technologiques importantes puisque la dilatation peut être calculée éléments par éléments de l'élément structurant B avant de combiner les résultats intermédiaires par l'union. Nous avons donc

$$\begin{aligned} X \oplus B &= \{x \in D \mid \exists b \in B, x - b \in X\} \\ &= \bigcup_{b \in B} \{x \in D \mid x - b \in X\}. \end{aligned} \tag{3.32}$$

Une mise en œuvre par translation de tous les pixels $x \in X$ par les vecteurs $b \in B$ et en vérifiant s'ils appartiennent à l'ensemble X aura une complexité en $\mathcal{O}(|X \oplus B|)$ [86]. Comme les images sont stockées dans la mémoire de ordinateurs comme des tableaux unidimensionnels, une mise en œuvre efficace consiste à translater les lignes au lieu des pixels comme illustré dans la figure 3.42b et dans l'algorithme 2 [76].



(a) Calcul sur le voisinage.

(b) Translation par lignes.

FIG. 3.42 – Calcul d'une dilatation (a) en utilisant un itérateur de voisinage ou (b) par translation par lignes. Source [207].

Algorithme 1 Dilatation par un itérateur sur le voisinage

1: $NbLines$, $NbCol$ ▷ nombre de lignes et colonnes
2: b ▷ $b = (b.x, b.y)$ itérateur (ensemble des coordonnées centrées) du voisinage B
3: $inPix$, $outPix$ ▷ images d'entrée et de sortie
4: **for all** $i \in [1; NbLines]$ **do**
5: **for all** $j \in [1; NbCol]$ **do**
6: $maxN = 0$
7: **for all** $b \in B$ **do**
8: $outPix(i, j) = \max(maxN, inPix(\text{i+b.y,j+b.x}))$
9: **end for**
10: $outPix(i, j) = maxN$
11: **end for**
12: **end for**

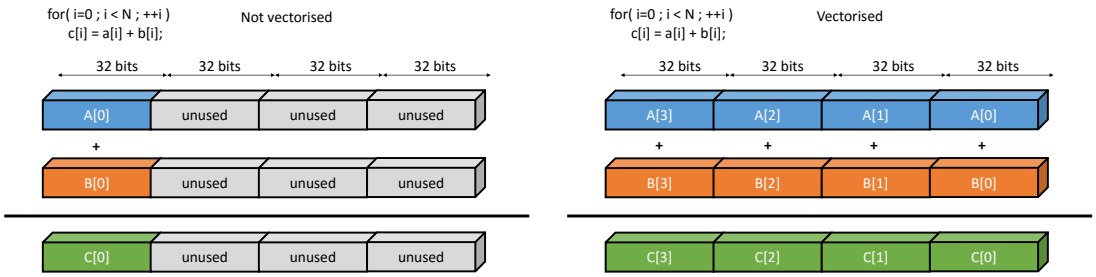
Algorithme 2 Dilatation par translation par lignes

1: $NbLines$, $NbCol$ ▷ nombre de lignes et colonnes
2: $inLine$, $outLine$ ▷ lignes des images d'entrée et de sortie
3: $tmpLine$ ▷ tableau de longueur $NbCol$
4: b ▷ $b = (b.x, b.y)$ coordonnées centrées du voisinage B
5: **for all** $i \in [1; NbLines]$ **do**
6: fill($outLine(i, :)$, 0)
7: **for all** $b \in B$ **do**
8: $tmpLine = translate(inLine(i + b.y, :), b.x)$ ▷ voir alg. 3
9: **for all** $j \in [1; NbCol]$ **do**
10: $outLine(i, j) = \max(outLine(i, j), tmpLine(j))$
11: **end for**
12: **end for**
13: **end for**

Algorithme 3 Translation d'une ligne

1: **function** TRANSLATE($inLine$,dx)
2: fill($outLine$,0) ▷ $outLine$ tableau de même longueur L que $inLine$
3: **for** $n \in [max(0; -dx); min(L - dx; L)[$ **do**
4: $outLine(n + dx) = inLine(n)$
5: **end for**
6: **end function**

Comme les opérations sur chaque ligne sont indépendantes de celles sur les autres lignes (éq. (3.32)), une méthode de programmation parallèle est utilisée, à savoir *OpenMP*, qui est conçue pour les ordinateurs multicœurs à mémoire partagée. Chaque *thread* (fil) est calculé par un seul cœur et tous les cœurs partagent une mémoire commune [49]. De plus, les données vectorisées sont traitées avec les instructions SIMD dans chaque thread [55]. Les opérations vectorielles avec les instructions SIMD permettent à des données multiples d'être traitées par une seule instruction du processeur alors que les opérations scalaires utilisent une instruction pour traiter chaque donnée individuelle. La taille des registres SIMD étant de 128 bits, 16 opérations sur des entiers encodés sur 8 bits sont effectuées en même temps au lieu d'une seule. De même, 4 opérations sur des entiers encodés sur 32 bits sont calculées en une fois comme illustré sur la figure 3.43. Dans les compilateurs récents, la vectorisation s'effectue automatiquement après activation de l'option adéquate. Faessel a montré dans [76] que dans une image numérisée sur 8 bits et de taille $1\,024 \times 1\,024$ pixels, une dilatation par un élément structurant carré de taille 3 pixels est accélérée par un facteur 136 avec une mise en œuvre par ligne avec parallélisation et vectorisation par comparaison avec une mise en œuvre par voisinage (0,33 ms versus 45 ms, Intel®Core™ i3 CPU M330 ; 2,13 GHz ; 2 cœres ; 4 threads). La figure 3.43 montre une comparaison entre les différentes méthodes de dilatation mises en œuvre par Pierre Guilbert pour un élément structurant carré de taille 100 pixels. A l'époque la meilleure mise en œuvre avec décomposition de l'élément structurant, parallélisation et vectorisation était bien plus rapide que celle de *MATLAB*®.



(a) Calcul sans vectorisation.

(b) Calcul avec vectorisation.

FIG. 3.43 – Calcul de la somme c de deux vecteurs a et b (a) sans et (b) avec vectorisation automatique pour des entiers 32 bits sur des registres SIMD de 128 bits.

Lorsque les opérations morphologiques élémentaires peuvent être calculées de manière rapide, il est aisé d'obtenir des opérations rapides pour des éléments structurants plus complexes. Ceci a constitué le cœur du travail de Pierre Guilbert pendant une année. Il a programmé ces opérations pour des lignes orientées, des rectangles et des disques en utilisant les accélérations au niveau des processeurs mais aussi des algorithmes astucieux de décomposition des éléments structurants. Cette bibliothèque informatique a été programmée en langage C++ et interfacée avec *MATLAB*®. Une programmation générique des fonctions a été faite grâce à des *Templates* (i.e. des modèles) réutilisables pour tous les types d'entiers de 8 à 64 bits mais aussi pour des nombres à virgules flottantes en simple précision (encodés sur 32 bits) et en double précision (encodés sur 64 bits). Pierre Guilbert a aussi réalisé des tests unitaires et des tests d'intégration de cette bibliothèque ainsi qu'une documentation automatique avec le générateur de documentation *Doxygen* [1].

Une décomposition des érosions et dilatations avec un élément structurant ligne par la méthode des doublets de points a d'abord été mise en œuvre. Cette méthode est présentée par Beucher dans [31] et sa première mise en œuvre serait due à Sternberg dans [289]. Puis une autre mise en œuvre a été réalisée par la méthode de Van Herk [304] et Gil and Werman [89] qui est notamment présentée dans [90]. La méthode de Van-Herk-Gil-Werman permet un calcul des érosion/dilatations avec un élément structurant linéaire en seulement 3 opérations de minimum/maximum sur un pixel. La

1. `https://www.doxygen.nl`

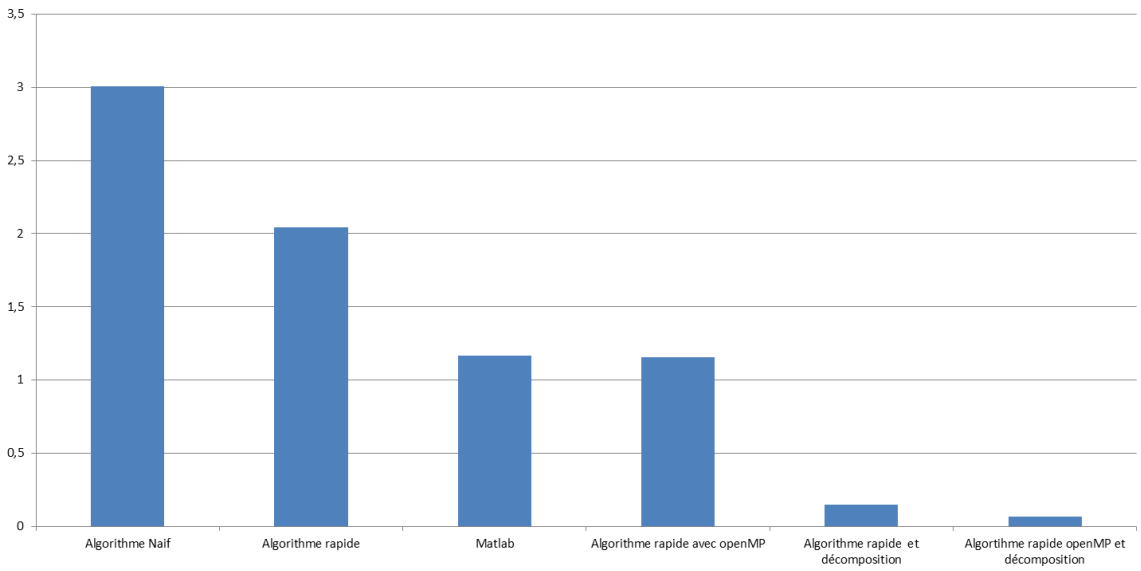


FIG. 3.44 – Comparaison des temps de calcul (en s) d'une dilatation par un élément structurant de taille 100×100 sur une image de taille $768 \times 1\ 024$ pixels par (a) la méthode naïve par voisinage, (b) l'algorithme rapide par translation, (b) *MATLAB*® 2015, (b) l'algorithme rapide avec parallélisation par OpenMP et vectorisation, (e) l'algorithme rapide avec décomposition et (f) l'algorithme rapide avec décomposition, parallélisation et vectorisation.

programmation des érosions et dilatations sur des segments orientés a ensuite été réalisée par la méthode récursive de Soille [282] qui généralise l'algorithme de Van Herk grâce à un *padding* astucieux sur les bords des images. Les segments discrets ont été obtenus soit par l'algorithme de Bresenham [38], soit par les *Lignes Périodiques* de Jones et Soille [134]. Les érosions/dilatations avec des segments orientés sont ensuite utilisées pour faire des érosions/dilatations avec des disques grâce à une décomposition radiale approchée. Deux méthodes de décomposition ont été mises en œuvre : celle d'Adams avec des *lignes de Bresenham* [7] et celles de Soille [134] avec les *Lignes Périodiques*. L'avantage de la méthode de décomposition par *Lignes Périodiques* est de pouvoir être utilisée dans des ouvertures granulométriques ; i.e. pour deux disques de rayons λ_i, λ_j la propriété suivante est vérifiée : $\gamma_{\lambda_i}\gamma_{\lambda_j} = \gamma_{\lambda_j}\gamma_{\lambda_i} = \gamma_{\max(\lambda_i,\lambda_j)}$ [284].

Accélération de la méthode de Köhler sur le seuillage par contraste

Ces méthodes de calcul rapides par décomposition d'éléments structurants et par parallélisation et vectorisation m'ont inspiré un nouvel algorithme pour le seuillage de Köhler par contraste. Celui-ci a donné lieu à une publication au congrès ICIP 2017 (*IEEE International Conference on Image Processing*) [207]. Le seuillage adaptatif est l'une des techniques de segmentation les plus utilisées dans de nombreuses applications car elle est rapide à calculer et quand elle combinée avec un filtrage précédent, elle donne des règles de décision robustes pour la reconnaissance de forme.

Il existe de nombreuses techniques de seuillage [41]. La méthode de Köhler a la particularité d'utiliser la courbe de contraste sur les frontières des régions de l'image [156]. Les sauts de contraste correspondent aux maxima locaux de la courbe de contraste et peuvent être extraits par un seuillage multiple de l'image. Cette méthode a été utilisée dans de nombreuses applications : industrielles, biomédicales, vidéo, etc. [109], [140], [146]. Cependant, son calcul est coûteux en temps de traitement ; presque une minute avec une mise en œuvre en C++ sur un ordinateur actuel pour une image acquise par un appareil numérique récent sur 18 millions de pixels, comme illustrée sur la figure 3.45a. J'ai donc proposé une mise en œuvre plus rapide que les algorithmes existants afin de la rendre utilisable pour des applications nécessitant du calcul rapide ou du temps

réel (e.g. seuillage vidéo, analyse de grandes banques d'images). L'ingénieur-chercheur que je supervisais, Pierre Guilbert, a réalisé la mise en œuvre rapide en C++.

Précédemment, deux tentatives pour réduire les temps de calcul de la méthode de Köhler avaient été faites. Zeboudj et Jourlin [59], [333] utilisaient des opérations morphologiques pour arriver à une version similaire de la méthode. Cependant leur approche était efficace sur des systèmes électroniques qui ne sont plus disponibles aujourd'hui. Dans l'autre, Hautière [109], [110] fait le calcul sur une partie réduite du voisinage et précalcule des images de minimum et maximum entre l'image translatée horizontalement et verticalement afin d'obtenir la courbe de contraste. Cependant, cet algorithme ne permet pas la parallélisation. Ma méthode introduit un algorithme parallèle du seuillage de Köhler grâce à des translations par lignes sur l'image. Je calcule aussi le contraste sur un voisinage réduit comme dans la méthode d'Hautière [109]. Ce nouvel algorithme permet de tirer parti de la vectorisation grâce aux instructions SIMD des processeurs et du calcul multicœur (i.e. parallèle) grâce à OpenMP.

Présentons tout d'abord la méthode de Köhler [146], [156]. Soient une image à niveaux de gris $f : D \to [0; M[$, où $D \subset \mathbb{Z}^2$, $[0; M[\subset \mathbb{R}$ et $M = 256$ pour numérisation sur 8 bits, ainsi qu'un seuil $t \in [0; M[$. Afin de réduire la longueur des notations, on note pour tout $x \in D$, $f_x = f(x)$. Deux classes sont générées par le seuil t :

$$C_0^t(f) = \{x \in D, f_x \leq t\} \text{ et} \tag{3.33}$$

$$C_1^t(f) = \{x \in D, f_x > t\} \tag{3.34}$$

ainsi qu'une frontière :

$$B(t) = \left\{(x_0, x_1) \in D^2, x_0 \in C_0^t(f), x_1 \in C_1^t(f) \text{ et } x_1 \in N_4(x_0)\right\}. \tag{3.35}$$

Le voisinage à 4 pixels du point x est noté $N_4(x)$. Pour les images bidimensionnelles, des voisinages à 6 ou 8 pixels pourraient être utilisés avec des différences insignifiantes [109]. A chaque couple de pixels (x_0, x_1) de la frontière $B(t)$, Köhler associe un contraste $C_K^t(x_0, x_1)$ défini comme :

$$C_K^t(x_0, x_1) = \min\left(f_{x_1} - t, t - f_{x_0}\right). \tag{3.36}$$

Il est le minimum des deux sauts (en matière de contraste) générés par le seuil t entre les valeurs $f_{x_0} = f(x_0)$ et $f_{x_1} = f(x_1)$. Le contraste moyen de la frontière $B(t)$ est défini comme :

$$C_K(B(t)) = \frac{1}{card\,\{B(t)\}} \cdot \sum_{(x_0,x_1)\in B(t)} C_K^t(x_0, x_1), \tag{3.37}$$

où $card\,\{B(t)\}$ est le cardinal (i.e. le nombre d'éléments) de l'ensemble $B(t)$ constituant la frontière. La sommation est faite pour les couples de pixels (x_0, x_1) appartenant à la frontière $B(t)$. Ceci crée une courbe de contrastes $C_K(B(t))$ pour tous les seuls possibles $t \in [0; M[$. Le seuil optimal t_0 est choisi comme :

$$C_K(B(t_0)) = \max_{t\in[0;M[} \left(C_K(B(t)\right). \tag{3.38}$$

Dans la figure 3.45c, six seuils significatifs, correspondant aux minima locaux, ont été extraits de la courbe de contraste. Ces seuils multiples permettent une simplification (i.e. une compression) des niveaux de gris de l'image : en passant de 256 à 7 niveaux de gris (Fig. 3.45a et 3.45b). Les sept niveaux de gris correspondent à la valeur moyenne des pixels de chacune des classes de la segmentation.

Une mise en œuvre directe consiste à calculer pour chaque seuil $t \in [0; M[$, le contraste $C_K^t(B(t))$ de la frontière $B(t)$. En C++, celle-là a une durée de 53 s sur une image de taille 3672×4096 pixels en utilisant un processeur Intel®Core™ i7 CPU 4702HQ ; 2,20 GHz ; 4 cœurs ; 8 threads. Pour des applications temps-réel, ou pour des grandes banques de données, un algorithme plus rapide est nécessaire.

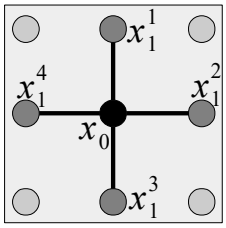
(a) Image initiale (256 classes). (b) Image initiale (7 classes). (c) Courbe de contraste.

FIG. 3.45 – Seuillage multiple par la méthode de Köhler (a) d'une image en une (b) image segmentée (la valeur de chaque classe est la moyenne des valeurs de ses pixels) par (c) les sept seuils sélectionnés sur la courbe de contraste.

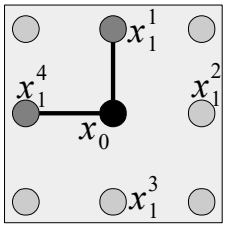
Cependant l'approche directe n'est pas conçue pour le traitement parallèle et vectoriel. Le contraste de Köhler (éq. (3.36)) est sommé sur les frontières $B(t)$ qui sont obtenues par la différence ensembliste entre une dilatation morphologique de l'ensemble $C_1^t(f)$ et ce même ensemble : $\{x_0 \in D, (x_0, x_1) \in B(t)\} = (C_1^t(f) \oplus A) \setminus C_1^t(f)$. L'élément structurant A correspond au voisinage à 4 pixels. La complexité de la mise en œuvre directe est en $\mathcal{O}(N^2)$ (i.e. la complexité d'une dilatation, $\mathcal{O}(N)$, multipliée par la complexité du parcours des couples de pixels de la frontière $B(t)$, $\mathcal{O}(M \cdot card\{B(t)\}) = \mathcal{O}(N)$). Une accélération directe consisterait à calculer la frontière avec une dilatation morphologique accélérée comme présenté précédemment. Cependant, une telle approche ne réduirait pas la complexité de l'algorithme. Dans ce but, nous proposons d'effectuer d'abord la translation des lignes de l'image, qui est adaptée au calcul parallèle, puis de calculer la contribution au contraste de chaque couple de pixels, pour chaque seuil. La complexité décroît alors à $\mathcal{O}(NM)$, i.e. le produit du nombre de pixels par le nombre de niveaux de gris. L'algorithme 4 présente notre approche.

Algorithme 4 Algorithme rapide pour la méthode de Köhler

1: I, J ▷ nombre de lignes et colonnes
2: $inLine(i, :)$ ▷ ligne i de l'image d'entrée
3: a ▷ $a = (a.x, a.y)$ translation du voisinage $A = N_4^*$
4: $curLine, nLine_1 \ldots nLine_a$ ▷ tableaux de longueur J
5: $C, card$ ▷ tableaux de longueur M initialement mis à zéro
6: **for all** $i \in [1, I]$ **do**
7: $curLine \leftarrow inLine(i, :)$
8: **for all** $a \in A$ **do**
9: $nLine_a \leftarrow translate(inLine(i + a.y, :), a.x)$
10: **for all** $j \in [1, J]$ **do**
11: $mini \leftarrow \min(curLine(j), nLine_a(j))$
12: $maxi \leftarrow \max(curLine(j), nLine_a(j))$
13: **for all** $t \in [mini, maxi - 1]$ **do**
14: $C(t) \leftarrow C(t) + \min(maxi - t, t - mini)$
15: $card(t) \leftarrow card(t) + 1$
16: **end for**
17: **end for**
18: **end for**
19: **end for**
20: **for all** $t \in [0, M[$ **do**
21: $C(t) \leftarrow C(t)/card(t)$
22: **end for**



(a) Voisinage N_4.



(b) « Demi »-voisinage N_4*.

FIG. 3.46 – (a) Le voisinage à 4 pixels N_4 sa (b) « moitié » N_4^*.

Afin de gagner un facteur 2 supplémentaire, montrons qu'il est équivalent de faire le calcul sur un « demi »-voisinage N_4* ou sur le voisinage à 4 pixels N_4 (fig. 3.46). Cette idée a été introduite par Hautière dans [109], [110]. Ici, l'égalité des deux approches est démontrée. Lorsque l'on scanne la frontière $B(t)$ avec le 4-voisinage N_4, seuls les pixels x_1, tels que $f_{x_0} < f_{x_1}$, contribuent au contraste de la frontière. Cependant, les deux pixels x_0 et x_1 sont voisins : $x_0 \in N_4(x_1) \iff x_1 \in N_4(x_0)$. Donc, les deux couples (ordonnés) (x_0, x_1) et (x_1, x_0) sont parcourus mais un seul couple contribue au contraste de la paire (désordonnée) de points $\{x_0, x_1\}$,

$$\begin{aligned}\Gamma_K^t(\{x_0,x_1\}_{N_4}) &= \begin{cases}\min\left(f_{x_0}-t, t-f_{x_1}\right), & \text{si} \quad f_{x_1} \le t < f_{x_0}\\ \min\left(f_{x_1}-t, t-f_{x_0}\right), & \text{sinon} \quad f_{x_0} \le t < f_{x_1}\end{cases}\\ &= \min\left(|f_{x_0}-t|, |t-f_{x_1}|\right). \end{aligned} \tag{3.39}$$

Supprimons la condition d'ordre entre les niveaux de gris, $f_{x_0} \le t < f_{x_1}$ et définissons le contraste absolu du couple (x_0, x_1) comme

$$\overline{C}_K^t(x_0,x_1) = \min\left(|f_{x_0}-t|, |t-f_{x_1}|\right).$$

Remarquons que $\overline{C}_K^t(x_0,x_1) = \overline{C}_K^t(x_1,x_0)$. Lorsque les deux couples sont parcourus sans l'ordre des niveaux de gris, le contraste de la paire de points $\{x_0, x_1\}$ est égal à :

$$\begin{aligned}\overline{\Gamma}_K^t(\{x_0,x_1\}_{N_4}) &= \overline{C}_K^t(x_0,x_1) + \overline{C}_K^t(x_1,x_0)\\ &= 2\min\left(|f_{x_0}-t|, |t-f_{x_1}|\right)\\ &= 2\Gamma_K^t(\{x_0,x_1\}_{N_4}) \qquad \text{(éq. (3.39))}. \end{aligned} \tag{3.40}$$

L'utilisation du « demi »-voisinage N_4^* permet de parcourir un seul couple de pixels. Le résultat suivant est donc obtenu :

$$\overline{\Gamma}_K^t(\{x_0,x_1\}_{N_4^*}) = \Gamma_K^t(\{x_0,x_1\}_{N_4}). \tag{3.41}$$

Afin de rendre parallèle l'algorithme, le calcul du contraste peut être effectué indépendamment ligne par ligne. Pour chaque thread parallèle k, deux tableaux de longueur M, notés C_k pour le contraste et $Card_k$ pour le comptage des pixels, ont besoin d'être créés au début du processus parallèle (ligne 6, alg. 4). A la fin du processus parallèle (ligne 19), ils sont groupés par sommation dans les tableaux de contraste C et de comptage $Card$. Au lieu d'être effectuées entre des nombres seuls, plusieurs opérations peuvent être faites sur des tableaux ce qui permet la vectorisation des données. Les opérations suivantes sont vectorisées et traitées grâce aux instructions SIMD :

i) la translation par ligne (ligne 9),

ii) le calcul du minimum (ligne 11) et du maximum (ligne 12) entre les tableaux $curLine$ et $nLine_a$,

iii) le calcul du contraste C_k (ligne 14) et du comptage $card_k$ (ligne 15) et

iv) la normalisation du contraste (ligne 20).

Comparons maintenant la durée de la mise en œuvre directe à celle de l'algorithme rapide. Un processeur Intel®Core™ i7 CPU 4702HQ ; 2,20 GHz ; 4 cœurs ; 8 threads avec 16 Go de RAM a été utilisé. Sur l'image « Tulipe » (Fig. 3.45a) avec un appareil photographique actuel ayant une résolution de $3\,672 \times 4\,896$ pixels en utilisant l'algorithme rapide avec parallélisation, le calcul de la méthode de Köhler prend 0,13 s (Table 3.1) au lieu de 53 s, soit un facteur de gain de 405. Avec des images d'ancienne résolution (512×512 pixels), tels que l'image « Lenna » ; la méthode directe prend 0,69 s et la méthode rapide 0,005 s, i.e. un facteur de gain de 126. Pour traiter des images avec une résolution actuelle, il est donc essentiel d'utiliser un algorithme plus rapide au lieu de la mise en œuvre directe. D'autres expériences ont confirmé ce résultat.

Name	Lenna	Tulipe
Taille (pixels)	512×512	$3\,672 \times 4\,896$
Directe (D)	$6{,}90 \times 10^{-1}$ s	$5{,}29 \times 10^{1}$ s
Rapide (B)	$1{,}62 \times 10^{-2}$ s	$4{,}86 \times 10^{-1}$ s
Rapide (A)	$5{,}46 \times 10^{-3}$ s	$1{,}30 \times 10^{-1}$ s
Gain (B vs. D)	42	109
Gain (A vs. D)	126	405

TABLE 3.1 – Comparaison de la durée des mises en œuvre pour des images de différentes tailles : mise en œuvre directe (D), rapide sans parallélisation (B) et rapide avec parallélisation (A). Les facteurs de gain sont calculés entre les mises en œuvre (B) et (D) et entre les mises en œuvre (A) et (D).

La méthode de Köhler a ensuite été testée sur une vidéo de voiture [2] issue de la banque de données YFCC100M (Yahoo Flickr Creative Commons 100M) [299]. Dans la figure 3.47, deux frames (i.e. plans) et leurs segmentations sont montrées. La mise en œuvre directe segmente la vidéo à un taux de 1 frame par seconde, ce qui est plus rapide que les 25 frames/s de la vidéo (Tab. 3.2). L'algorithme rapide de la méthode de Köhler est donc adapté au traitement des vidéos en temps-réel.

Nom	YFCC100M (car)
taille (pixels)	502×480
Nombre de frames	640
Directe (D)	0,99 frames/s
Rapide (A)	96,96 frames/s
Gain (A vs. D)	98

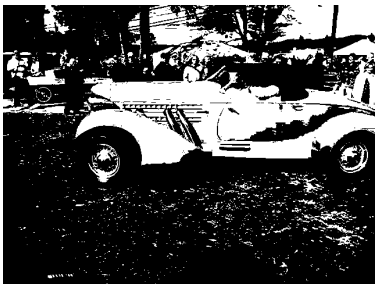
TABLE 3.2 – Nombre de frames par secondes segmentées par la méthode de Köhler avec différentes mises en œuvre appliquées sur une vidéo : directe (D), et rapide avec parallélisation (A). Les facteurs de gain entre les mises en œuvre (A) et (D) sont donnés.

L'algorithme rapide pour le seuillage de Köhler a une complexité en $\mathcal{O}(NM)$ qui est plus faible que celle de l'approche directe en $\mathcal{O}(N^2)$. Sa conception tire profit des capacités des processeurs actuels : le traitement multicœur avec OpenMP et le calcul vectoriel avec les instructions SIMD. Cela le rend utilisable pour les applications temps-réel ou le traitement des grandes banques d'images. Son intérêt est de le combiner avec d'autres transformées précédentes sur les

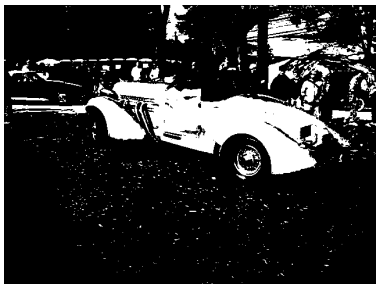
2. Vidéo de voiture `https://www.flickr.com/photos/32109282@N00/4078577031`

(a) Frame 101.

(b) Frame 251.



(c) Frame segmentée 101.



(d) Frame segmentée 251.

FIG. 3.47 – (a), (b) Deux frames d'une vidéo de voiture de la banque de donnée YFCC100M et (c), (d) leurs segmentations en deux classes par la méthode de Köhler.

images telles qu'un filtre passe-bas, un *top-hat* ou une carte des distances d'Asplund couleurs [209], [229]. Dans le futur, d'autres contrastes pourraient être testés tels que ceux définis avec le modèle *Logarithmic Image Processing* [140], [146]. De plus, les méthodes de calcul multicœurs parallèles et vectorielles pourront servir à accélérer d'autres algorithmes sur des voisinages tels que les cartes des distances d'Asplund (cf. section 4.3, p. 102).

3.5 Classification d'images du fond d'œil

A l'IPRI, j'avais accès à de grandes banques de données d'images du fond d'œil qui devaient être traitées automatiquement afin de suivre l'évolution de la rétinopathie chez les patients. Comme les images étaient de qualité variable, il convenait de séparer celles de qualité insuffisante des autres. Pour ce faire, j'ai encadré le stage de Sabina Stefan au cours du premier semestre 2016. Puis afin de classifier la sévérité de la rétinopathie diabétique chez les patients en deux groupes, ceux qui devaient être adressés à un ophtalmologiste des autres, j'ai encadré le stage de Lucas Arnaldo-Alves durant un semestre en 2019.

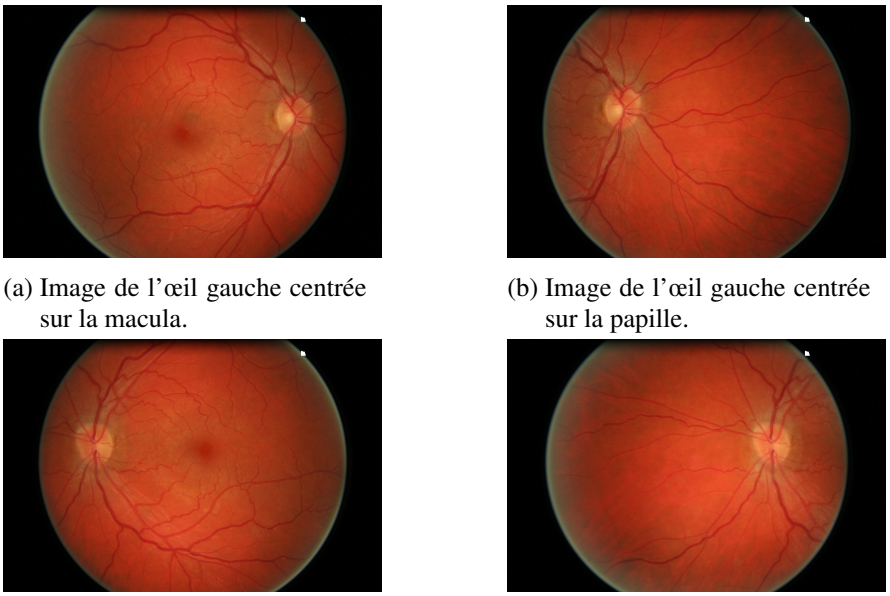
3.5.1 Evaluation de la qualité des images

Entre janvier et juillet et 2016, Sabina Stefan diplômée d'un *Bachelor* (Bac +4) de *University of Cape Town* en Afrique du Sud a effectué son stage à l'IPRI sous ma supervision. Son travail a consisté à analyser la qualité des images du fond d'œil. Nous avions accès à une banque d'images de santé publique provenant de *Diabetic Eye Screening Wales* (DESW) et fournie par le Professeur David Owens diabétologue à *Swansea University* au Pays de Galles, Royaume-Uni. Dans cette banque, les images étaient très majoritairement de haute qualité. Dans d'autres banques que nous avons utilisées, les images avaient une qualité moindre. Dans toutes les banques, les images de trop faible qualité devaient être détectées. Ceci a constitué la première partie du stage de Sabina Stefan.

En outre, j'avais développé un algorithme qui permettait d'améliorer le contraste des images du fond d'œil [234]. Cependant certaines images ressortaient de l'algorithme avec des couleurs beau-

coup plus vives que les autres car les patients venaient d'être opérés de la cataracte. L'algorithme d'amélioration de l'époque était donc différent pour ces images. Par conséquent, il fallait les détecter automatiquement afin de savoir quel algorithme appliquer. Ceci a constitué la deuxième partie du stage. Depuis, l'algorithme d'amélioration a été revu et s'adapte automatiquement à tous les types d'images.

Le dépistage annuel des patients souffrant du diabète par des non médecins consiste à prendre une image de chaque œil dans deux positions : une centrée sur la papille optique et l'autre sur la macula (Fig. 3.48). La faible qualité de certaines images peut être causée par [206] :



(a) Image de l'œil gauche centrée sur la macula.

(b) Image de l'œil gauche centrée sur la papille.

(c) Image de l'œil droit centrée sur la macula.

(d) Image de l'œil droit centrée sur la papille.

FIG. 3.48 – Images du fond d'œil utilisées pour le dépistage de la rétinopathie diabétique. Pour chaque œil deux images sont acquises : l'une centrée sur la macula et l'autre sur la papille. (a) et (b) œil gauche. (c) et (d) œil droit.

- une perte de contraste due au mouvement du patient ou au mouvement de l'œil ;
- une illumination non-uniforme de la rétine due à i) une taille insuffisante de la pupille ou ii) l'acquisition d'une partie de la paupière due au clignement de l'œil ;
- des médias opaques tels que la cataracte ou les hémorragies vitreuses qui empêchent une acquisition de qualité normale de la rétine.

La figure 3.49 montre des images de faible qualité et des images de qualité normale.

Afin de trouver les images de trop faible qualité, Sabina Stefan s'est appuyée sur la méthode développée par Niemeijer et al. [205], [206]. Durant son stage, elle a programmé et testé cette méthode dont les principales étapes sont illustrées par le schéma de la figure 3.50.

Le principe en est le suivant. La structure de chaque image est classifiée par une technique appelée *Image Structure Clustering (ISC)* qui est proche de la description de texture par *textons* de Malik et al. [169]. Celle-là détermine le plus important ensemble de structures présentes dans des images de qualité normale grâce à une classification des variables générées par une banque de filtres. La présence de cet ensemble de structures indique la qualité de l'image. Pour ce faire, une série de filtres gaussiens est utilisée avec une largeur σ de plus en plus grande (1, 2, 4, 8, 16). Cependant ces filtres doivent être invariants aux rotations et translations afin de générer des réponses similaires aux structures spécifiques des images quelles que soient leurs orientations et positions dans l'image. Pour cette raison, les coordonnées de jauge (*Gauge coordinates*) [99] sont employées. Soit $L : D \to \mathbb{R}$, où $D \subset \mathbb{R}^2$, une image. Celle-ci peut être vue comme une

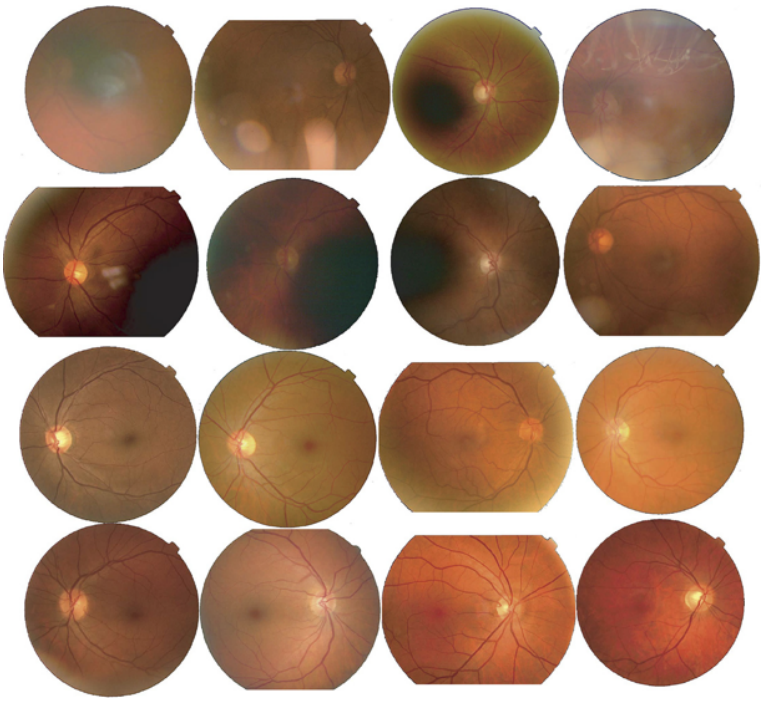


FIG. 3.49 – Comparaison des images de faible qualité sur les deux premières lignes avec des images de qualité normale sur les deux dernières lignes. Source [206].

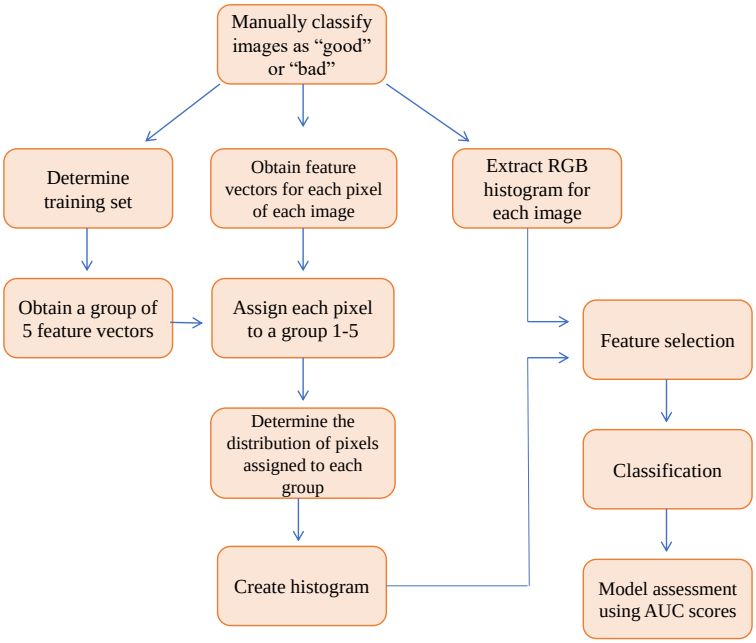


FIG. 3.50 – Schéma de la méthode de détection de la qualité d'image programmée par Sabina Stefan. AUC : *Area Under the Curve* (aire sous la courbe). RGB : *Red, Green, Blue*.

collection de courbes de mêmes valeurs, les isophotes. Les coordonnées de jauge sont telles que la direction v est tangente à l'isophote et la direction w pointe dans la direction du gradient $\left(\frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}\right)$ (Fig. 3.51). Le gradient étant invariant aux rotations et translations, les dérivées exprimées dans les coordonnées de jauge le sont également.

Les dérivées jusqu'au second ordre sont utilisées. Cinq filtres sont retenus L, L_w, L_{vv}, L_{vw} et L_{ww} dont les indices indiquent une dérivée selon une certaine direction, e.g. L_{ww} est la dérivée

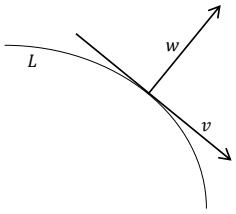


FIG. 3.51 – Coordonnées de jauge. Le vecteur v est tangent à l'image L et le vecteur w pointe dans la direction du gradient $\left(\frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}\right)$.

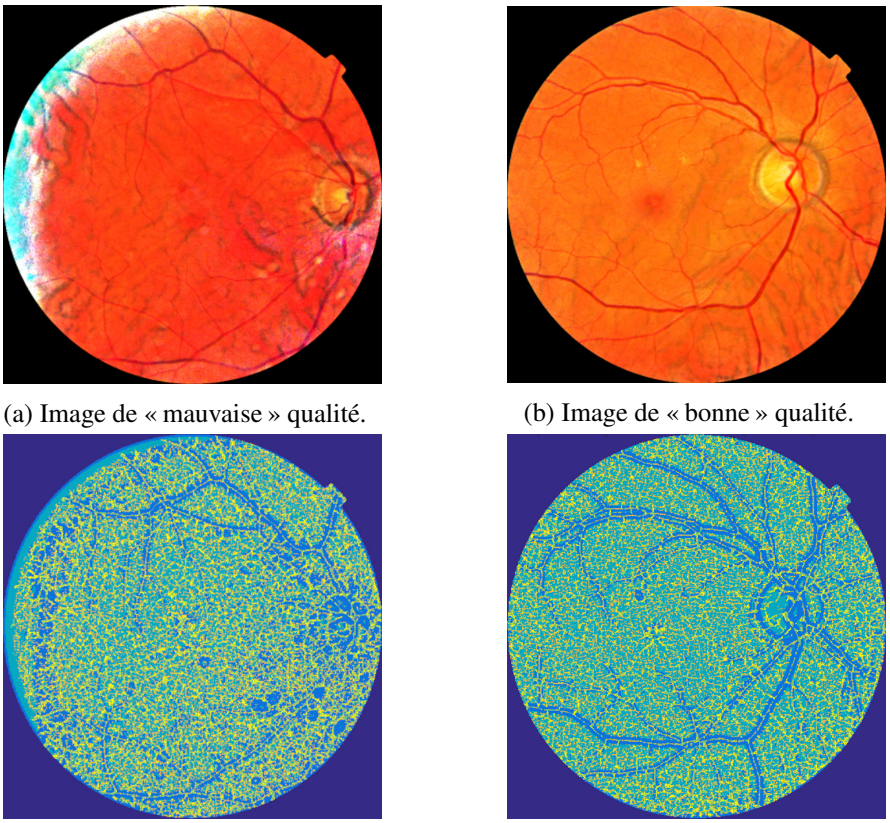
seconde selon le vecteur w. Ces cinq dérivées constituent un ensemble irréductible d'invariants de second ordre. Tous les autres invariants jusqu'au second ordre peuvent être exprimés comme des fonctions polynomiales de celles-ci [99], [119]. Les cinq dérivées peuvent être écrites à l'aide des dérivées sur l'image L, L_x, L_y, L_{xx} et L_{xy}, comme montré dans le tableau 3.3. En pratique, des

Caractéristique	Expression
L	L
L_w	$\sqrt{L_x^2 + L_y^2}$
L_{vv}	$\frac{-2L_xL_{xy}L_y+L_{xx}L_y^2+L_x^2L_{yy}}{L_x^2+L_y^2}$
L_{vw}	$\frac{-L_x^2L_{xy}+L_y^2L_{xy}+L_xL_y(L_{xx}-L_{yy})}{L_x^2+L_y^2}$
L_{ww}	$\frac{L_x^2L_{xx}+2L_xL_{xy}L_y+L_y^2L_{yy}}{L_x^2+L_y^2}$

TABLE 3.3 – Ensemble irréductible des invariants de second ordre de la structure de l'image. Dans la colonne de gauche, les dérivées en coordonnées de jauge (v, w) sont données et dans celle de droite leur expression dans les coordonnées (x, y) de l'image.

filtres dérivateurs gaussiens servent à calculer les dérivées selon x et y [82]. Les structures de différentes échelles spatiales sont analysées en faisant varier la largeur σ du filtre gaussien (1, 2, 4, 8, 16). Pour chaque pixel on obtient un vecteur de 25 caractéristiques. Afin de constituer un ensemble d'entraînement, plusieurs images sont choisies par un observateur et classifiées comme étant de « bonne » ou de « mauvaise » qualité. Puis, dans chaque image de « bonne qualité », 150 vecteurs de caractéristiques sont extraits aléatoirement et classifiés en 5 classes en utilisant la méthode *k-means* [69], [103]. Les autres pixels sont ensuite agrégés aux cinq classes ISC comme illustré sur la figure 3.52. Les variables avaient été centrées et réduites afin de faciliter leur comparaison lors de la classification.

Pour chaque image, un histogramme des cinq classes ISC est alors construit. De même, pour chacune des composantes Rouge, Vert et Bleu (RVB) des images, un histogramme à cinq classes est obtenu. Chaque image est donc décrite par un total de 20 paramètres. Pour chaque méthode de classification, un sous-ensemble optimal de paramètres est sélectionné en utilisant : l'algorithme *Sequential Forward Floating Selection* (SFFS) [248] ou l'Analyse en Composantes Principales. Dans l'algorithme SFFS, les paramètres sont ajoutés séquentiellement, un par un, à un ensemble de paramètres initialement vide mais à chaque ajout plusieurs paramètres peuvent être enlevés si cela améliore la performance du classificateur. La performance du classificateur est évaluée en utilisant l'aire mesurée sous les courbes ROC (*Receiver-Operating Characteristic*). Ces courbes représentent la sensibilité (i.e. le taux de vrais positifs) d'un classificateur en fonction de un moins sa spécificité (i.e. le taux de faux positifs) [185]. Plusieurs classificateurs ont été testés : i) les *Support Vector Machine* (SVM) avec un noyau à base radiale [108], ii) l'Analyse Discriminante Quadratique [71], iii) l'Analyse Discriminante Linéaire [71] et iv) les k-plus proches voisins [108]. Pour les SVM avec une combinaison des paramètres ISC et RVB, le papier de Niemeijer et al.



(a) Image de « mauvaise » qualité.

(b) Image de « bonne » qualité.

(c) Classes ISC de l'image de « mauvaise » qualité.

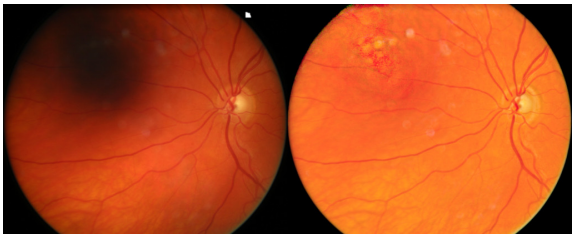
(d) Classes ISC de l'image de « bonne » qualité.

FIG. 3.52 – Images de « mauvaise » et « bonne » qualité classifiées par l'algorithme *k-means* sur les descripteurs ISC.

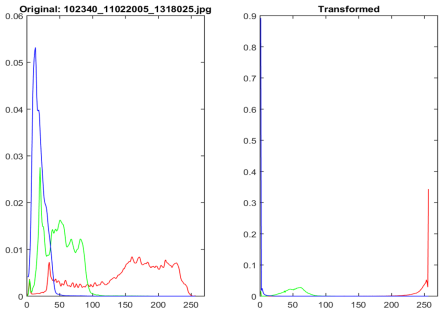
[206] obtenait une aire sous la courbe de 99,7 %. Avec notre banque de données qui était différente, Sabina Stefan a obtenu 95,17 %. En utilisant un algorithme d'amélioration de contraste sur les images [234] avant la détection de qualité, l'aire sous la courbe est de 90 %. La méthode de détection de qualité est donc sensible aux variations de contraste et d'intensité lumineuse dans les images.

En 2016, pour les images des patients ayant subi une opération de la cataracte, l'algorithme d'amélioration du contraste donnait des couleurs trop saturées (i.e. trop vives) pour être facilement analysées (Fig. 3.53c). Il fallait donc détecter ces images pour réduire la saturation des couleurs. La figure 3.53 compare l'amélioration des images de la rétine entre un patient n'ayant pas subi d'opération de la cataracte (Fig. 3.53a et 3.53b) et un patient l'ayant subie (Fig. 3.53c et 3.53d).

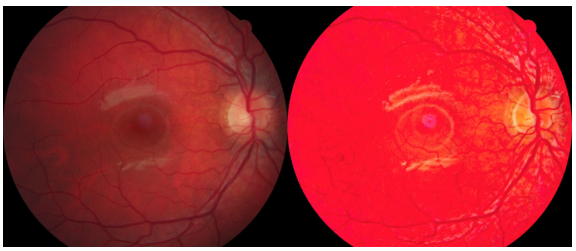
Afin de séparer les deux types d'images, Sabina Stefan a commencé par examiner plusieurs variables sur les histogrammes : le nombre de pics, leur hauteur, la différence de position entre deux pics, etc. Finalement l'algorithme qui fonctionnait le mieux était celui qui déterminait la position du pic de l'histogramme de la composante verte. La figure 3.54 compare l'histogramme d'une image améliorée d'un patient n'ayant pas subi l'opération de la cataracte avec celui de l'image améliorée d'un patient ayant eu cette opération. Pour ce dernier, le pic sur l'histogramme de la composante verte est situé proche de zéro alors que ce n'est pas le cas pour le patient n'ayant pas été opéré. Sur un ensemble d'entraînement constitué manuellement, un seuil sur la position du pic a été déterminé par un algorithme non supervisé (*k-means*). Comme illustré sur la figure 3.55, les pics dont la position est inférieure au seuil (en rouge) correspondent aux images des patients ayant eu une opération de la cataracte. Avec ce seuil une précision (i.e. *accuracy*) de 98 % a été obtenue. La précision correspond au rapport de la somme des vrais positifs et vrais négatifs par la



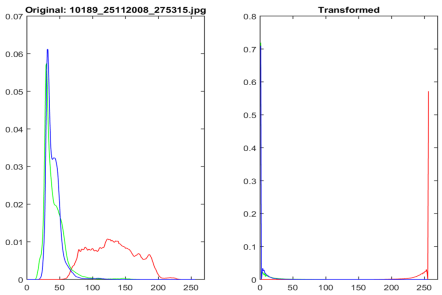
(a) Image initiale et améliorée d'un patient n'ayant pas été opéré de la cataracte.



(b) Histogrammes de l'image initiale et améliorée.



(c) Image initiale et améliorée d'un patient ayant subi une opération de la cataracte.

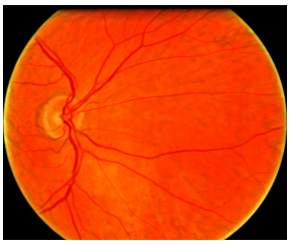


(d) Histogrammes de l'image initiale et améliorée.

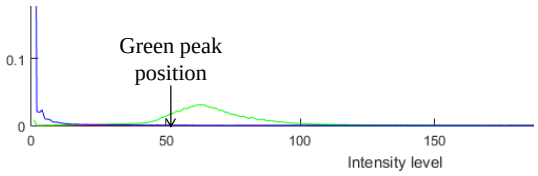
FIG. 3.53 – (a) Image initiale et améliorée de la rétine d'un patient n'ayant pas été opéré de la cataracte. (b) Histogrammes correspondants. (c) Image initiale et améliorée de la rétine d'un patient ayant été opéré de la cataracte. (d) Histogrammes correspondants.

somme de tous les éléments. La banque de données comportait 69 patients choisis aléatoirement dans une banque de donnée de santé publique. Chaque patient avait eu sept examens de la rétine avec environ un an d'intervalle entre les examens [234].

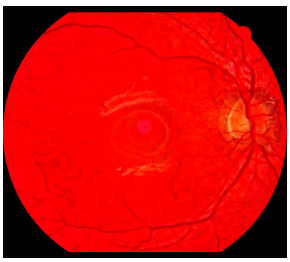
Les résultats obtenus pour la détection des images de qualité insuffisante ainsi que pour la détection des images des patients ayant été opérés de la cataracte ont donné des résultats très satisfaisants sur des banques de d'images dont nous disposions. Je me suis ensuite intéressé à la classification automatique de la rétinopathie dans les images du fond d'œil.



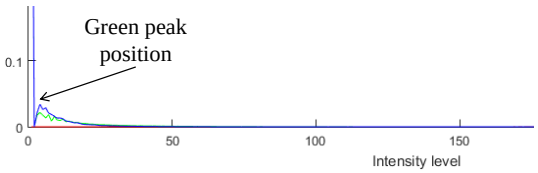
(a) Image améliorée d'un patient n'ayant pas été opéré de la cataracte.



(b) Histogramme de l'image améliorée.



(c) Image améliorée d'un patient ayant été opéré de la cataracte.



(d) Histogrammes de l'image améliorée.

FIG. 3.54 – (a) Image améliorée de la rétine d'un patient n'ayant pas été opéré de la cataracte. (b) Histogramme correspondant et pic de la composante verte. (c) Image améliorée de la rétine d'un patient ayant été opéré de la cataracte. (d) Histogramme correspondant et pic de la composante verte.

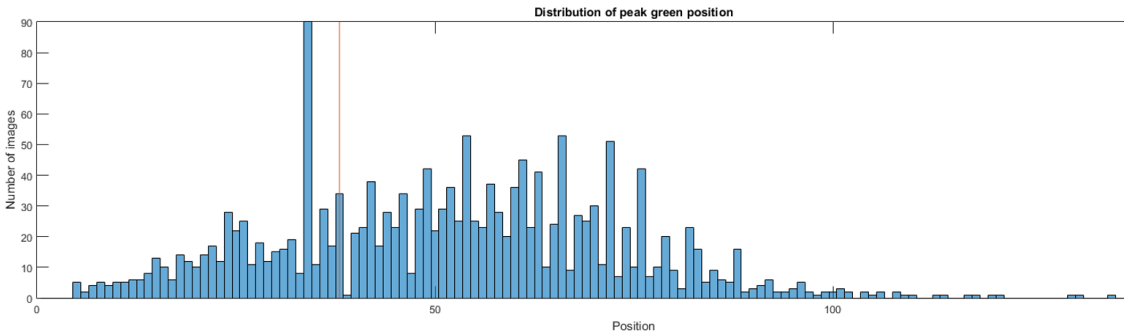


FIG. 3.55 – Distribution des positions des pics des composantes vertes des histogrammes. Les pics dont la position est inférieure au seuil (en orange) correspondent aux images des patients ayant eu une opération de la cataracte.

3.5.2 Classification automatique des images du fond d'œil

De février à août 2019, Lucas Arnaldo Alves a effectué son stage d'élève-ingénieur et de Master 2 Recherche à l'IPRI sous ma supervision [14]. L'objectif du stage était d'étudier la classification automatique par *Machine Learning* des images du fond d'œil pour le dépistage de la rétinopathie diabétique. *Le dépistage consiste en un examen visuel de la rétine (i.e. le fond de l'œil) afin de déterminer quels patients souffrant du diabète doivent être adressés à un ophtalmologiste.* En Grande Bretagne [102], le dépistage est fait de manière visuelle par des non-médecins. En France, il est fait par des ophtalmologistes. Comme tous les patients diabétiques devraient être dépistés cela constitue un travail très important. C'est pourquoi les algorithmes automatiques de classifications automatique ont un intérêt grandissant. Plusieurs algorithmes ont été développés afin de détecter les lésions ou la sévérité de la rétinopathie diabétique [5], [23], [80], [83], [97], [98], [157], [252], [258], [300], [303], [320]. Actuellement, les systèmes fondés sur les réseaux de neurones convolutifs (*ConvNet*) [158] donnent des performances de détection automatique élevées.

Afin de classifier automatiquement les images fond d'œil, Lucas Arnaldo s'est intéressé à la méthode par apprentissage profond (i.e. *deep learning*) proposée par Quellec et al. [252]. Il a mis en œuvre une partie de l'approche et l'a testée sur une banque d'images. Grâce à la bibliothèque d'apprentissage *TensorFlow*, Lucas a programmé en langage Python l'architecture présentée sur la figure 3.56. Celle-ci compte plus de $4{,}6 \times 10^6$ paramètres.

L'apprentissage supervisé consiste à entraîner un système de prédiction sur un *ensemble d'entraînement*. Cet ensemble d'entraînement est constitué d'observations (e.g. des images de patients) pour lesquelles une mesure du résultat est connue (e.g. « à référer », « à ne pas référer » à un ophtalmologiste). Grâce à ces données, on construit un modèle de prédiction qui sera capable de prédire le résultat pour un ensemble de données non encore vues par le système qui est appelé *ensemble de test* [108]. Dans notre cas, le système est constitué : i) d'une entrée, l'image de couches dites *cachées* des réseaux de neurones et ii) d'une sortie, la probabilité d'appartenance à chacune classes « à référer à un ophtalmologiste » (*REF*) et « à ne pas référer à un ophtalmologiste » (*NonREF*). Sur l'échelle de sévérité internationale de la rétinopathie diabétique (RD), la classe « à ne pas référer » (*NonREF*) correspond aux catégories « Pas de RD apparente » et « RDNP légère ». La classe « à référer » (*REF*) correspond aux catégories « RDNP modérée » ou plus (cf. section 2.3, p. 15).

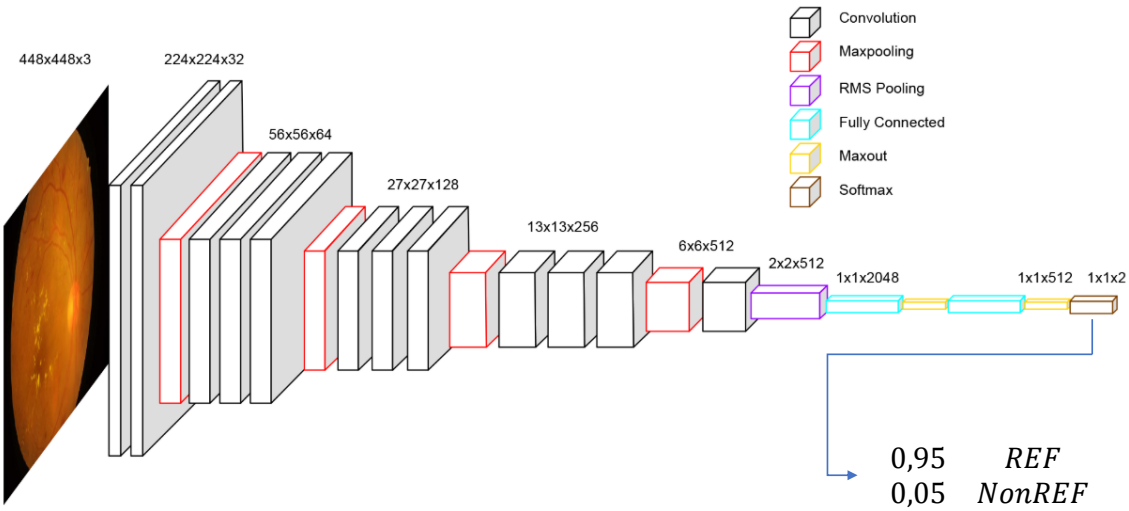


FIG. 3.56 – Architecture du système d'apprentissage profond par réseau de neurones convolutifs. L'entrée du système est l'image du fond d'œil. La sortie est la probabilité que l'image soit classée comme devant être référée à un ophtalmologiste, *REF*, ou comme ne devant pas lui être référée, *NonREF*.

Rappels sur l'apprentissage en profondeur supervisé

Une présentation détaillée de l'apprentissage profond est faite dans l'ouvrage de Goodfellow et al. [94] ou dans celui de Nielsen [204]. Nous nous limiterons ici à quelques rappels succincts. Dans un réseau de neurones artificiels, l'unité fondamentale de traitement est le *perceptron* également appelé neurone (Fig. 3.57). Celui-ci a été introduit par Rosenblatt en 1958 [262]. Dans un réseau, chaque neurone est connecté à ceux de la couche antérieure par ses entrées $x_i \in \mathbb{R}$ pondérées par les poids ω_i. Un biais b est ajouté à la somme pondérée des entrées, $S = \sum_{i=1}^{n} \omega_i x_i + b$, afin de déterminer si le neurone est activé, $f(x) = x$ si $S \geq 0$, ou pas, $f(x) = 0$ si $S < 0$. L'étape d'activation permet d'ajouter la non-linéarité au modèle, ce qui augmente fortement la capacité du système à se rapprocher d'une fonction quelconque [61]. La sortie de la fonction d'activation est ensuite reliée à la couche suivante afin de contribuer au réseau. Le perceptron est répété autant de fois qu'il y a de couches. Le perceptron multicouches ou MLP (*Multilayer Perceptron*) est principalement utilisé pour des tâches de classification. En effet, comme toutes les sorties de la

couche précédente sont connectées à son entrée, il possède une connexion globale (i.e. complète) à cette couche.

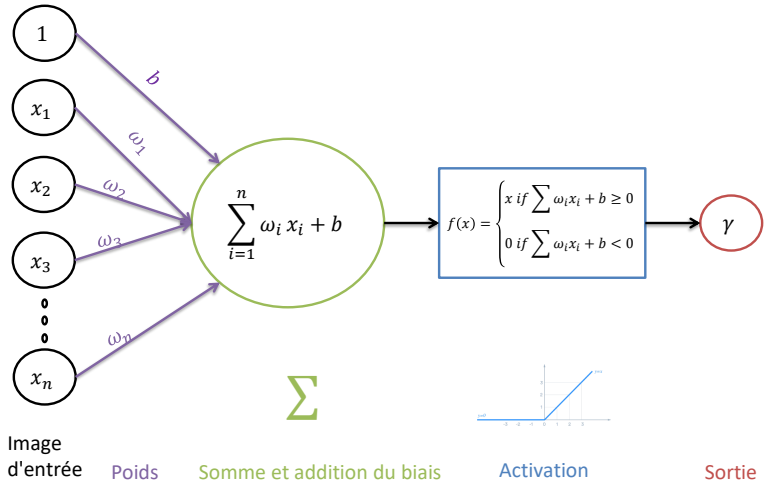


FIG. 3.57 – Schéma d'un perceptron.

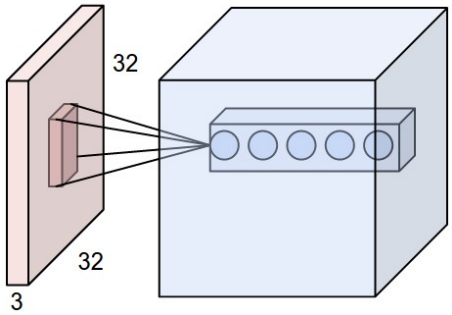
A la différence du perceptron, les couches de réseaux de neurones convolutifs permettent surtout d'extraire l'information sur l'image d'entrée. En apprentissage profond, le terme convolution dénote en fait la corrélation croisée. Soit une image $I : D \to \mathbb{R}$, où $D \subset \mathbb{N}^2$, et un noyau $\Omega : D_\Omega \to \mathbb{R}$, dont le domaine $D_\Omega \subset D$ est de taille $M \times N$. La corrélation croisée entre l'image I et le noyau Ω est définie de la manière suivante :

$$S(x, y) = (I * \Omega)(x, y) = \sum_{m=1}^{M} \sum_{n=1}^{N} I(x + m, y + n)\Omega(m, n). \tag{3.42}$$

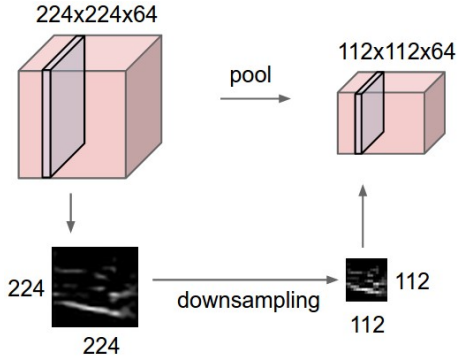
Le neurone convolutif a une connexion locale de taille $M \times N$ qui est plus petite que la connexion globale du neurone MLP. Dans une couche de convolution, les paramètres à apprendre sont les poids du noyau de convolution. Le noyau de convolution reste cependant le même dans une même couche afin d'assurer la propriété d'invariance par translation. La sortie d'une couche de convolution est appelée « carte d'activation ». La carte d'activation est donc la convolution de l'entrée de la couche par son noyau. Comme chaque carte possède un noyau qui lui est propre, le nombre de cartes d'activations détermine la profondeur d'une couche de convolution ; c'est à dire le nombre de noyaux à apprendre lors de l'entraînement. Pour une image couleur $I : D \to \mathbb{R}^3$, chaque couche de convolution transforme un volume d'entrée 3D en un volume de sortie 3D en utilisant les cartes d'activations. Dans l'exemple de la figure 3.58a, la couche convolutive se compose de 5 cartes d'activations et sa profondeur d'entrée est 3 à cause des composantes rouge, vert et bleu de l'image. L'opération dite de convolution sur l'image couleur I à $L = 3$ composantes et un tenseur Ω composé de C_Ω filtres ayant chacun une dimension de $w \times h \times L$ est décrite de manière complète par l'équation suivante :

$$S_{x,y,c} = a\left(\sum_{u=-\frac{w}{2}}^{w-\frac{w}{2}} \sum_{v=-\frac{h}{2}}^{h-\frac{h}{2}} \sum_{d=1}^{L} I_{s_l x+u, s_l y+v, d}\Omega_{u,v,d,c} + b_{x,y,c}\right). \tag{3.43}$$

Les variables sont notées en indices. La valeur de la sortie $S_{x,y,c}$ à la position $(x, y) \in D$ et pour la composante $c \in [\![1, C_\Omega]\!]$ est le produit de convolution entre l'image I, dans une fenêtre glissante centrée en $(s_l x, s_l y)$, et le tenseur Ω. s_l est le pas entre chaque glissement de la fenêtre, i.e. le décalage entre deux positions de la fenêtre glissante. $S_w = \frac{I_w - w + 2P}{s_l} + 1$ et $S_h = \frac{I_h - h + 2P}{s_l} + 1$ représentent la largeur et la hauteur de la sortie de l'opération de convolution. P est le paramètre



(a) Couche de convolution.



(b) Couche de sous-échantillonnage.

FIG. 3.58 – (a) Illustration d'un volume d'entrée en rouge (e.g. une image de taille $32 \times 32 \times 3$) et d'un volume de neurones représentées par les boules bleues dans la première couche de convolution. Dans la couche de convolution, chaque neurone (représenté par une boule bleue) est seulement connecté à une région locale dans le domaine spatial, mais à la profondeur complète (i.e. toutes les composantes couleur). Il y a cinq neurones selon la profondeur et tous regardent la même région de l'entrée. (b) Couche de sous-échantillonnage (i.e. *pooling layer*). Le volume d'entrée de taille $224 \times 224 \times 64$ est sous-échantillonné par un filtre de taille 2, avec un pas de 2, en un volume de sortie de taille $112 \times 112 \times 64$. La profondeur du volume (64) est préservée. Habituellement on utilise comme filtre le maximum (i.e. *max pooling*).

Source : https://cs231n.github.io/convolutional-networks/

de remplissage (i.e. *padding*). I_w, I_h sont la largeur et la hauteur de l'image d'entrée. $a : D \to \mathbb{R}$ est la fonction d'activation non linéaire. $b_{x,y,c}$ est le biais qui est généralement attaché $b_{x,y,c} = b_c$.

La couche de convolution est généralement suivie d'une couche de sous-échantillonnage (i.e. *pooling layer*) (Fig. 3.58b). Cette couche permet de réduire la dimension spatiale de l'image ainsi que le nombre de paramètres et le temps de traitement du réseau de neurones, tout en conservant l'information utile grâce à une opération statistique comme par exemple le maximum (*Max pooling*) ou la moyenne (*RMS pooling*).

La couche *Maxout* [95] est introduite en complément de la couche *Dropout* [286] afin de faciliter l'optimisation et d'éviter le phénomène de sur-apprentissage. A chaque itération, des filtres de la couche de *Dropout* sont sélectionnées aléatoirement et désactivés (un filtre parmi k dans I sous-ensembles de filtres). L'intérêt est de tester si un modèle avec des filtres en moins donne une erreur plus faible. La couche *Maxout* suit la couche *Dropout*. Elle consiste à prendre le maximum dans chacun des I sous-ensembles de k filtres z_j de la couche précédente :

$$g(z)_i = \max_{j \in \mathbb{G}^{(i)}} z_j. \tag{3.44}$$

où $\mathbb{G}^{(i)}$ est l'ensemble des indices dans le groupe de filtres i, $\{(i-1)k+1, \ldots, ik\}$. Chaque unité *Maxout* $g(z)_i$ peut être interprétée comme faisant l'approximation linéaire par morceaux d'une fonction convexe arbitraire.

La fonction coût représente l'erreur de prédiction de notre modèle $\hat{y}_i$ par rapport à la vérité y_i. L'erreur quadratique moyenne a été choisie comme fonction coût :

$$L = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2. \tag{3.45}$$

Comme nous faisons de la classification, l'entropie croisée, $-\sum_i y_i \ln(\hat{y}_i)$, aurait aussi pu être utilisée [108]. Cette fonction coût doit être optimisée. C'est la méthode de *descente de gradient*

stochastique qui est souvent employée [108]. Il s'agit d'une méthode itérative qui permet de minimiser une fonction objectif grâce au gradient de la fonction coût par rapport à ses paramètres. Comme illustré dans la figure 3.59, à chaque itération, un échantillon est tiré aléatoirement dans l'ensemble des valeurs du gradient. La mise à jour des paramètres (ou poids) ρ est faite à partir de l'équation suivante :

$$\rho_{t+1} = \rho_t - \eta \frac{\partial L_t}{\partial \rho}. \tag{3.46}$$

η représente le taux d'apprentissage qui peut varier afin de choisir une vitesse d'apprentissage et ainsi influencer la convergence de la fonction. Une pratique courante en apprentissage des réseaux de neurones consiste à réduire sa valeur de façon dynamique lorsque le modèle stagne afin d'avoir un ajustement plus fin. Le calcul de la dérivée du coût par rapport à chaque poids de l'architecture ($\partial L_t / \partial \rho$) sera calculé par rétropropagation [108].

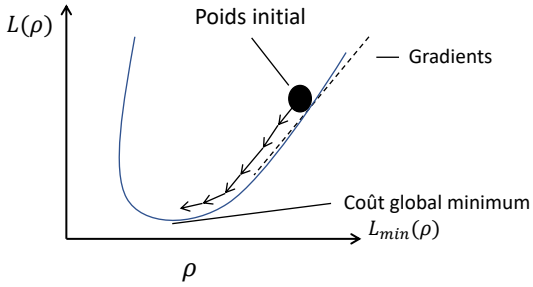


FIG. 3.59 – Méthode de descente de gradient stochastique.

La rétropropagation s'appuie sur la règle de dérivation en chaîne. Soit deux fonctions f, g : $\mathbb{R} \to \mathbb{R}$. Supposons que $y = g(x)$ et $z = f(g(x)) = f(y)$, alors la règle de dérivation en chaîne permet d'écrire :

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}. \tag{3.47}$$

L'algorithme commence par calculer une propagation directe (*forward propagation*) en alimentant le réseau avec un lot d'images (*batch*) et en calculant ses prédictions $\hat{y}_i$. Le coût total est calculé à partir de l'équation (3.45). Cette fonction coût est ensuite dérivée par rapport à la variable de sortie du réseau afin d'effectuer la rétropropagation. La figure 3.60 illustre application de l'algorithme de rétropropagation au calcul de la dérivée du coût par rapport au poids ω_1, $dL_{Total}/d\omega_1$. L_{o1} et L_{o2} représentent les coûts (i.e. les erreurs) associés aux neurones o_1 et o_2 de la couche qui suit le neurone analysé h_1. $\Sigma = \sum_i \omega_i x_i + b$ représente la somme des poids et biais associés à un neurone (Fig. 3.57). γ est la sortie du neurone h_1 après l'application de la fonction non-linéaire d'activation.

La dérivée de l'erreur par rapport à une couche $l-1$ peut être calculée à partir de celle de la couche suivante l en utilisant l'équation (3.48). La dérivée de l'erreur par rapport à un paramètre ρ est liée à la dérivée de la sortie de la couche par rapport au paramètre grâce à l'équation (3.49).

$$\frac{\partial L_L}{\partial D^{(l-1)}} = \frac{\partial L_L}{\partial D^{(l)}} \frac{\partial D^{(l)}}{\partial D^{(l-1)}} \tag{3.48}$$

$$\frac{\partial L_L}{\partial \rho} = \frac{\partial L_L}{\partial D^{(l)}} \frac{\partial D^{(l)}}{\partial \rho} \tag{3.49}$$

$l \in [\![0, L]\!]$ est le numéro de la couche. $D^{(l)}$ est le vecteur de sortie de la couche l du réseau. $D^{(0)}$ représente l'entrée du réseau et $D^{(L)}$ le vecteur de sortie du réseau avec la prédiction associée à chaque classe. ρ est un des paramètres de la couche et peut correspondre au poids ω, au biais b ou à l'un des noyaux de convolution Ω.

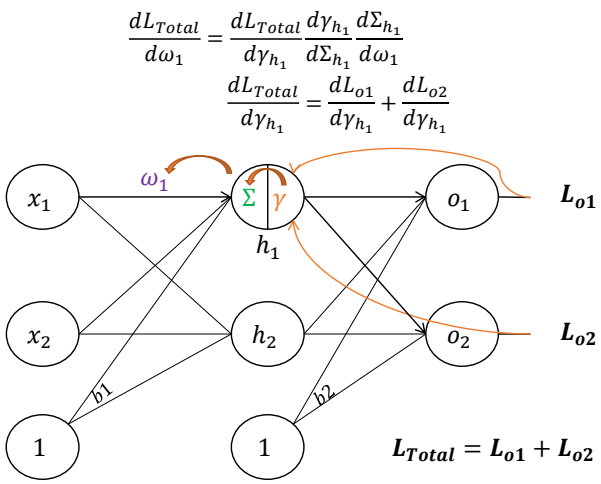
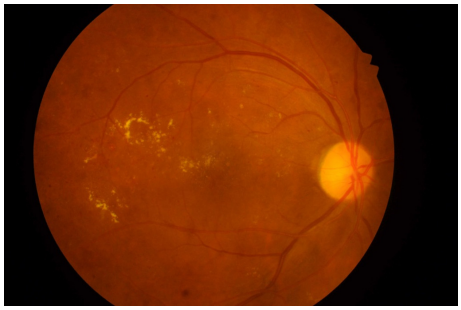


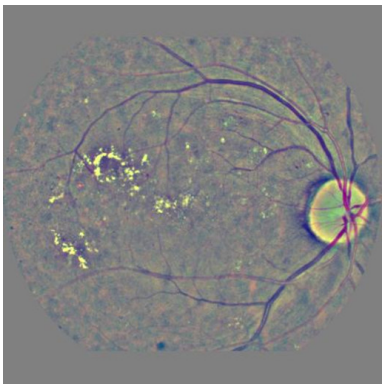
FIG. 3.60 – Méthode de rétropropagation de la dérivée de l'erreur L_{Total} par rapport au poids ω_1.

Spécificités de la méthode

Afin de résoudre les problèmes de bord et de taille des images, un pré-traitement a été appliqué comme illustré sur la figure 3.61. Il consiste à estimer le champ de vision de la caméra sur l'image, puis à faire un sous-échantillonnage de cette zone afin que sa largeur soit égale à 512 pixels. Un *padding* avec la valeur moyenne de chaque composante couleur est ensuite effectué afin que la hauteur de l'image soit aussi de 512 pixels. Les valeurs en dehors du masque du champ de vision sont aussi mises à la valeur moyenne. Afin de réduire les variations du fond, un filtrage gaussien d'une taille de 8,5 pixels est appliqué sur chaque composante de l'image. Puis, pour enlever les artefacts du bord, une érosion morphologique est appliquée avec un disque de rayon égal à 5 % de celui du champ de vision. Cette opération est suivie d'un deuxième sous-échantillonnage et du *padding* afin que l'image ait une taille de $448 \times 448 \times 3$ pixels.



(a) Image initiale.



(b) Image prétraitée.

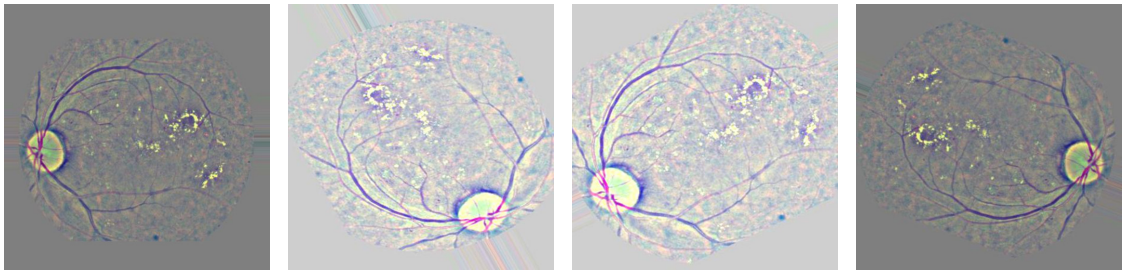
FIG. 3.61 – (a) Image de la banque IDRiD (*Indian Diabetic Retinopathy Image Dataset*) [245], [246]. (b) Pré-traitement sur cette image.

Afin de réduire le problème de sur-apprentissage, l'augmentation de données par création artificielle de nouveaux échantillons est employée. A chaque itération d'entraînement, l'ensemble d'entraînement est agrandi en réalisant des transformations de paramètres aléatoires sur les images. Les opérations réalisées sont les suivantes :

- translation dans l'intervalle $[-10 \text{ pixels}; 10 \text{ pixels}]$;
- rotation dans l'intervalle $[0°; 360°]$;
- changement de contraste par multiplication par un scalaire dans l'intervalle $[0{,}6; 1{,}67]$;

- changement d'échelle dans l'intervalle $[85\ \%; 115\ \%]$ de la taille initiale.

La figure 3.62 illustre quatre réalisations de l'augmentation de données sur une image prétraitée.



(a) (b) (c) (d)

FIG. 3.62 – Quatre images générées par augmentation de données à partir de l'image prétraitée de la figure 3.61b.

Afin de visualiser l'impact de chaque pixel d'entrée sur la décision finale, un critère de sensibilité a été introduit par Simonyan et al. [281]. Ce critère utilise les gradients de la sortie associée à une classe par rapport à l'entrée :

$$\bar{\omega}_{x,y} = \left\| \frac{\partial sortie_+}{\partial entree} \right\|_\infty = \left\| \frac{\partial D^{(L)}_{x,y,c}}{\partial D^{(0)}_{x,y,c}} \right\|_\infty . \tag{3.50}$$

Le symbole $+$ représente la classe *à référer* REF. (x, y) sont les coordonnées d'un pixel de l'image et c l'indice de la composante couleur de l'image. Le critère de sensibilité permet de calculer une *carte de chaleur* de même domaine que l'image d'entrée. Ses pixels représentent l'amplitude du changement de la probabilité de la classe *à référer* par rapport au pixel analysé, c'est-à-dire l'importance de ce pixel dans le choix de la classe. Le critère de sensibilité apporte une information très importante qui servira pour l'apprentissage. La contrainte suivante a été ajoutée à ce critère. Une somme est réalisée sur la dérivée des composantes couleur de l'image afin de réduire les artefacts liés à la teinte :

$$\pi_{x,y} = \left| \sum_c \frac{\partial D^{(L)}_{x,y,c_+}}{\partial D^{(0)}_{x,y,c}} \right| . \tag{3.51}$$

Dans l'ensemble d'entraînement, les classes *à référer* REF et *à ne pas référer* $NonREF$ ont des fréquences très différentes. Ceci peut biaiser l'apprentissage. Il est donc nécessaire de repondérer les classes grâce à un facteur de correction α qui permet de surpondérer la classe la moins représentée. Une « époque » t correspond au parcours de l'ensemble des images lors de l'entraînement du réseau de neurones. Plusieurs « époques » sont nécessaires pour réaliser l'entraînement complet. Au début de l'entraînement, le facteur α_0 donne un poids équivalent aux deux classes grâce à l'équation suivante :

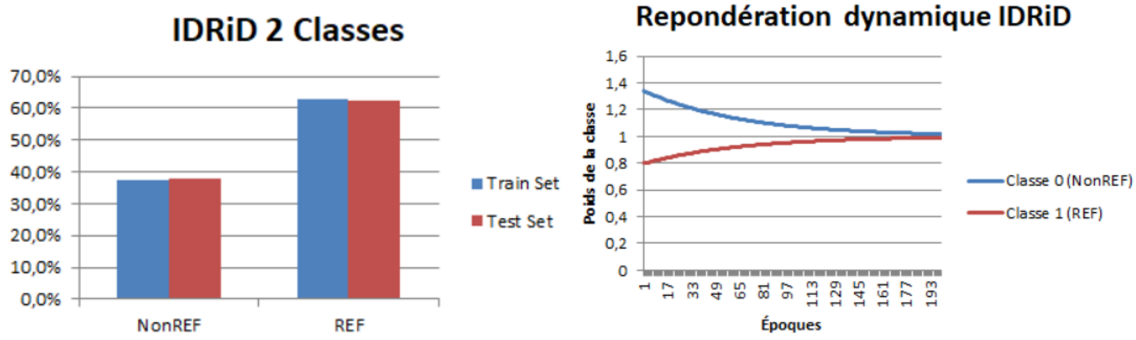
$$\alpha_{0_i} = \frac{Card\{images\}}{Nb_{classes} \times Card\{classe_i\}} . \tag{3.52}$$

$Card\{images\}$ correspond au nombre total d'images et $Nb_{classes}$ au nombre de classes. $Card\{classe_i\}$ est le cardinal de la classe i. Puis, au fur et à mesure des époques t, le paramètre α est progressivement réduit jusqu'à atteindre la valeur $\alpha_f = 1$ lors de la dernière époque ; i.e. l'entraînement est fait sur la distribution réelle des données. La repondération dynamique de la classe i à l'époque t est décrite par l'équation suivante :

$$\alpha_{t_i} = r^{t-1}\alpha_{0_i} + (1 - r^{t-1})\alpha_f , \tag{3.53}$$

où $r = 0{,}975$ est une constante fixée empiriquement.

La repondération dynamique a permis d'améliorer la convergence de l'apprentissage des paramètres lors de la phase d'entraînement. La distribution de la banque d'images analysée et sa repondération dynamique sont illustrées sur la figure 3.63.



(a) Fréquence de chaque classe.

(b) Repondération dynamique.

FIG. 3.63 – (a) Distribution des fréquences des classes de la banques de données IDRiD (*Indian Diabetic Retinopathy Image Dataset*) [245], [246]. (b) Repondération dynamique des classes.

Afin de réduire le problème du sur-apprentissage [108], la régularisation des paramètres est utilisée. Deux méthodes sont les plus employées, la régularisation L_1 en norme 1 et la régularisation L_2 en norme 2 qui sont définies de la manière suivante :

$$E_1 = \nu \sum_{\rho} |\rho| \,, \tag{3.54}$$

$$E_2 = \nu \sum_{\rho} \rho^2. \tag{3.55}$$

ν est une constante qui représente l'intensité de régularisation. Pour la régularisation, L_1 est plus discriminant que L_2. Elle donne une solution qui est plus éparse que L_2, i.e. plus de poids seront proches de zéro. Cela crée une plus grande discrimination sur les paramètres en aidant le réseau à « oublier » des caractéristiques moins significatives. L_2 est moins discriminant et conserve plus de paramètres.

Le critère de sensibilité présenté dans l'équation (3.51), nécessite des informations sur la dérivée de la sortie de réseau. Comme la bibliothèque *TensorFlow* ne possède pas de routine qui permette d'inclure des dérivées dans la rétropropagation, il a fallu en créer une. La fonction coût de l'ensemble du système est définie comme il s'ensuit :

$$\mathcal{L} = L_L + L_0. \tag{3.56}$$

L_L représente la fonction de coût associée à la sortie du réseau :

$$L_L = \frac{1}{n} \sum_{i=1}^{n} \alpha_{t_i} (y_i - \hat{y}_i)^2 + \beta \sum_{\rho} \rho^2. \tag{3.57}$$

Elle inclut la repondération dynamique α_{t_i} selon les époques (éq. (3.53)) et la régularisation L_2 pour chaque paramètre ρ. L_0 est le coût lié au critère de sensibilité par la dérivée $\frac{\partial L_L}{\partial D^{(0)}}$, dans la sens rétrograde. Il s'agit d'une régularisation L_1 du gradient L_L par rapport à la couche d'entrée $D^{(0)}$:

$$L_0 = \nu \sum_{(x,y)} \left| \sum_{c} \frac{\partial L_L}{\partial D^{(0)}_{x,y,c}} \right| (x, y). \tag{3.58}$$

Afin de mettre en œuvre la rétropropagation du coût L_L par rapport à la couche d'entrée $D^{(0)}$, la chaîne de dérivation de l'équation (3.48) a été utilisée. La chaîne complète s'écrit comme il s'ensuit :

$$\frac{\partial L_L}{\partial D^{(0)}} = \frac{\partial L_L}{\partial D^{(L)}} \prod_{i=1}^{L} \frac{\partial D^{(l)}}{\partial D^{(l-1)}}. \tag{3.59}$$

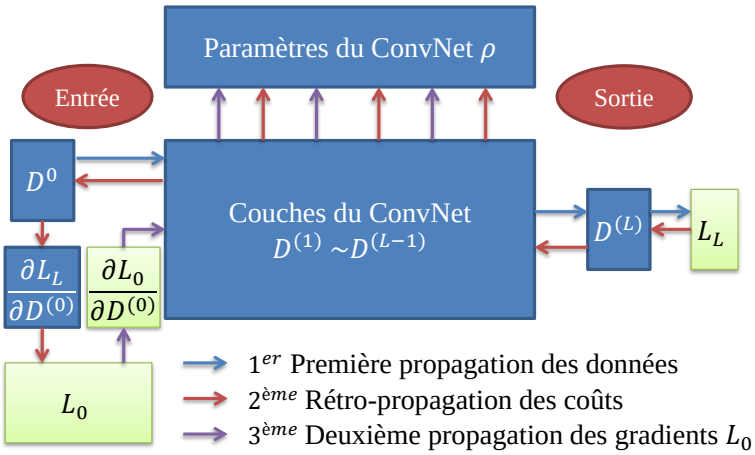


FIG. 3.64 – Représentation des itérations de l'algorithme durant la phase d'apprentissage.

Comme illustré dans la figure 3.64, La routine complète d'entraînement se compose de quatre étapes :

1. le paquet d'images est propagé sur le réseau de $D^{(0)}$ à $D^{(L)}$ (dans le sens direct) afin de calculer le coût L_L ;
2. le gradient $\frac{\partial L_L}{\partial D^{(L-1)}}$ est rétro-propagé grâce aux équations (3.48) et (3.49) afin de calculer $\frac{\partial L_L}{\partial \rho}$. Le gradient $\frac{\partial L_L}{\partial D^{(0)}}$ est déterminé par l'équation (3.59) , ce qui permet d'évaluer le coût L_0 (éq. (3.58)) ;
3. le gradient $\frac{\partial L_0}{\partial D^{(0)}}$ est propagé dans le sens direct sur le réseau de $D^{(0)}$ à $D^{(L)}$ de façon à calculer $\frac{\partial L_0}{\partial \rho}$;
4. finalement, chaque paramètre du ConvNet est mis à jour dans l'optimisation proportionnellement à $\frac{\partial L_L}{\partial \rho} + \frac{\partial L_0}{\partial \rho}$.

$\frac{\partial L_0}{\partial D^{(0)}}$ représente le coût associé au critère de sensibilité (éq. (3.51)) dans le sens direct. Il dépend de l'information acquise dans la rétropropagation. L'image du gradient $\frac{\partial L_0}{\partial D^{(0)}}$ peut être facilement propagée dans le ConvNet puisqu'elle a la même dimension que l'image entrée $D^{(0)}$. Nous pouvons remarquer que si le paramètre ν est mis à la valeur nulle dans l'équation (3.58), la régularisation est supprimée et nous retrouvons la rétropropagation classique programmée par défaut dans la bibliothèque *TensorFlow*.

Résultats

L'architecture du réseau de neurones convolutifs a été entraînée sur la banque de données publique Indian Diabetic Retinopathy Image Dataset (IDRiD) [245], [246]. Elle se compose de 516 images classées selon diverses sévérités de la RD. Les données ont été divisées en un ensemble de 413 images d'entraînement et un ensemble de 103 images de test. Comme illustré sur la figure 3.63a, les distributions statistiques des classes « à référer », *REF* (62 %), et « à ne pas référer », *NonREF* (38 %), ne respectent pas la distribution de la RD dans la population générale des diabétiques, qui est d'environ 30 %. La banque d'images est donc statistiquement biaisée par une sur-représentation de la classe la plus sévère, « à référer » *REF*.

L'apprentissage a été réalisé en 3 étapes au cours desquelles le nombre de couches de l'architecture a progressivement augmenté selon les itérations. A chaque itération, les paramètres ont été initialisés avec ceux du modèle précédent. L'architecture finale se compose d'un total de 14,5 millions de paramètres répartis sur 25 couches. à la fin de chaque étape d'entraînement, Lucas a sélectionné l'époque qui a donné la meilleure valeur de précision sur l'ensemble de test. En toute rigueur statistique, l'époque avec la meilleure valeur de précision aurait dû être sélectionnée sur un ensemble de validation extrait de l'ensemble d'entraînement.

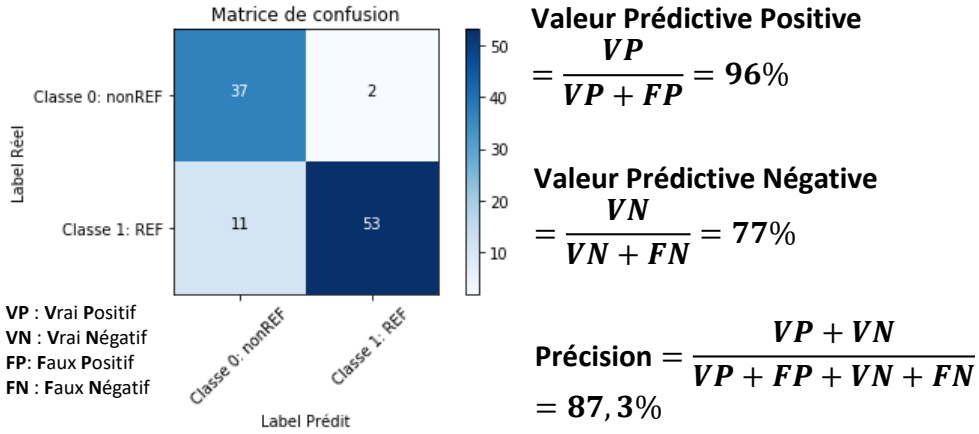


FIG. 3.65 – Matrice de confusion et résultats du classificateur pour la banque d'images IDRiD.

Comme illustré sur la figure 3.65, l'architecture entraînée atteint une précision de 87,3 %. Sur la classe référable, 96 % des images ayant un stade « à référer » à un ophtalmologiste ont été détectées correctement ; c'est la valeur prédictive positive. Dans le cas « à ne pas référer » à un ophtalmologiste, 77 % des images ont été bien classifiées ; c'est la valeur prédictive négative. Ces résultats indiquent que le système fournit une bonne classification. Cependant, il existe un certain nombre de faux négatifs (11) qui ne sont pas négligeables. Ces cas correspondent à des patients atteint d'une sévérité « à référer » et que le système classifie comme « à ne pas référer ». D'un point de vue médical c'est très problématique. En outre, dans de nombreux dispositifs, on considère que la valeur prédictive négative acceptable doit être supérieure à 95 %, alors qu'elle n'est ici que de 77 %. La performance du système est ici limitée par le faible volume de données. C'est une contrainte critique dans un problème d'apprentissage. Donc, étant donné la taille de la banque de données qui n'est que de 413 images pour l'apprentissage, ce résultat est encourageant. Il indique une bonne direction en ce qui concerne l'architecture retenue et la classification de la rétinopathie diabétique en deux stades.

Lucas a aussi tracé la courbe ROC (*Receiver Operating Characteristic*) du système avec ou sans régularisation (Fig. 3.66). L'aire sous la courbe avec la régularisation est de 0,89 contre 0,73 sans régularisation ; ce qui démontre l'intérêt de régulariser.

Afin de vérifier de manière visuelle l'aptitude du modèle à prendre sa décision sur des critères anatomiques, les cartes de chaleur (éq. (3.51)) sont représentées dans la figure 3.67. Les pixels activés (en blanc) correspondent aux lésions de la RD ; ce qui signifie que le système s'appuie plutôt sur des lésions pour prendre sa décision.

Quellec et al. [252] ont testé leur méthode sur la banque de données *Kaggle* pour la détection de la RD [68]. Celle-ci contient un total 88 702 images du fond d'œil provenant de 44 351 patients. A la différence de la banque d'images IDRiD, les images ont été acquises avec plusieurs résolutions et des types de caméras différentes. La distribution des classes « à ne pas référer » (*NonREF*) et « à référer » (*REF*) est d'environ 80 % et 20 % ce qui correspond aux proportions dans la population générale des patients souffrants de diabètes. Sur cette banque, Quellec et al. ont utilisé la méthode décrite ici ainsi que d'autres classificateurs qui recherchent les lésions de la rétinopathie. Ils ont obtenu des résultats très encourageants avec une aire sous la courbe ROC de 0,954. Par manque

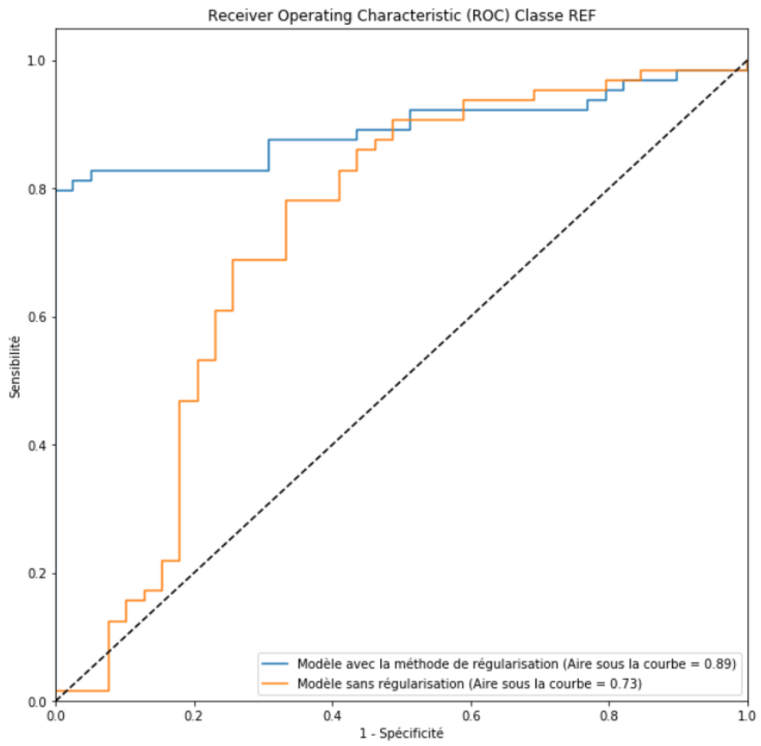


FIG. 3.66 – Courbe ROC (sensibilité en fonction de 1-spécificité) avec ($\nu = 1 \times 10^{-3}$) ou sans régularisation ($\nu = 0$) à l'entraînement pour la banque de données IDRiD.

de temps, Lucas a effectué des tests sur un sous-ensemble de 1 000 images de la banque d'images Kaggle. Cependant ceux-ci n'ont pas été concluants puisque du fait du faible nombre d'images et du faible nombre de lésions, un classificateur presque aléatoire a été obtenu.

Comme l'architecture de la figure 3.56, p. 76, compte plus de $4{,}6 \times 10^6$ paramètres et que les banques d'entraînement se composent au plus de quelques dizaines de milliers d'images [252], il y a donc une redondance des paramètres. Cette redondance fait l'objet d'études dans la communauté d'apprentissage profond, e.g. [251], [294].

La classification automatique de la RD reste donc un champ de recherche encore important. Des résultats très encourageants ont été obtenus par différentes équipes [5], [23], [80], [83], [97], [98], [157], [252], [258], [300], [303], [320]. Cependant les risques de ne pas détecter certains patients avec des formes sévères ne sont pas toujours nuls. En outre, il est souhaitable que les décisions prises par le classificateur soient explicables d'un point de vue anatomique ; ce qui n'est malheureusement pas toujours le cas. A ce jour, plusieurs systèmes de classification automatique de la RD ont reçu une certification comme des dispositifs médicaux européens (*Idx-DR* [3], [5], *RetmarkerDR* [258], *EyeArt* [303], *Retinalyze* [157]) et pour l'instant un seul a reçu un *approval* de la *Food and Drug Administration* (FDA) américaine (*Idx-DR*) [97].

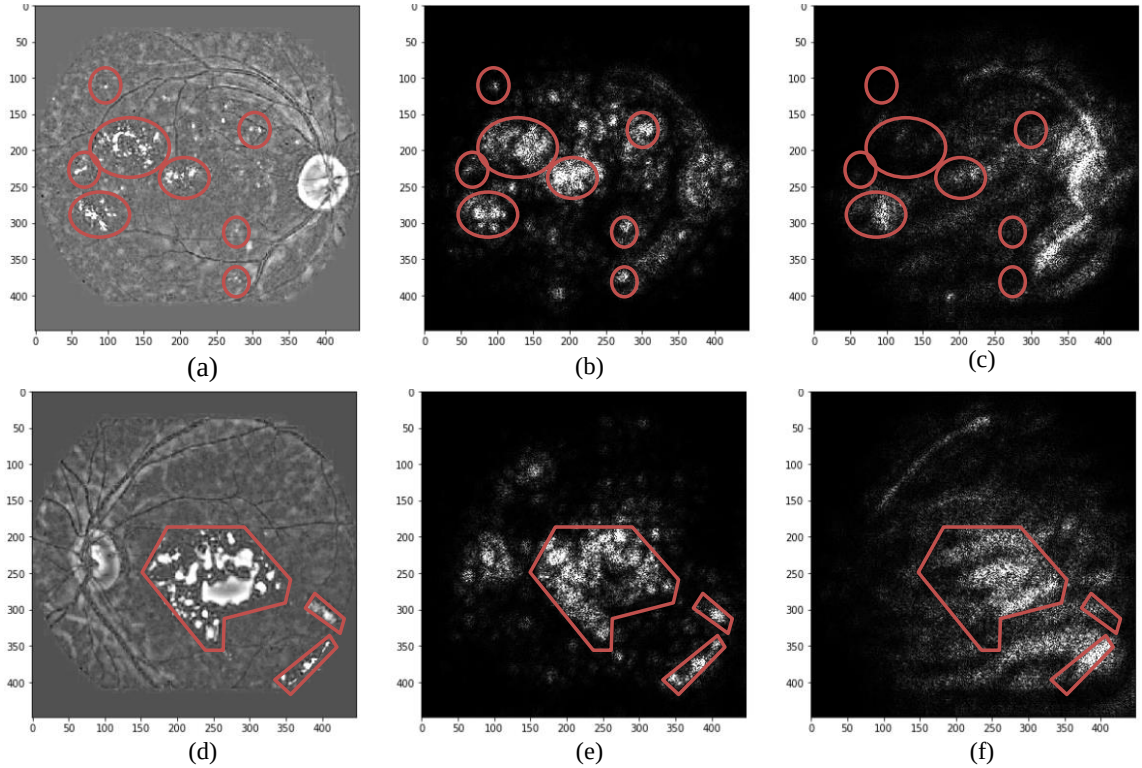


FIG. 3.67 – Analyse visuelle du critère de sensibilité par rapport aux lésions. (a) et (d) Images prétraitées. (b) et (e) Cartes de chaleur de l'architecture entraînée avec le critère de sensibilité avec la contrainte de teinte. Les exsudats secs sont encadrés en rouge. (c) et (f) Cartes de chaleur de l'architecture entraînées sans prendre en compte le critère de sensibilité dans l'apprentissage.

3.6 Conclusion du chapitre

Ce chapitre concernait l'analyse des grandes banques d'images. Un rappel a été fait sur les fondements de la Morphologie Mathématique. C'est un corpus mathématique solidement établi qui fournit des outils pratiques et robustes pour de nombreuses applications. Au Centre de Recherche de Michelin, la première application portait sur l'analyse des grandes banques d'images de la surface des pneumatiques dans le but de contrôler par vision automatique tous les pneumatiques fabriqués. Plusieurs méthodes de détection morphologique des anomalies en relief à la surface des pneumatiques ont été présentées ainsi qu'une méthode d'analyse des anomalies texturées par apprentissage statistique supervisé. Ces travaux ont notamment été réalisés grâce à une collaboration que j'ai initiée avec l'Ecole des Mines de Paris. Du côté académique, six chercheurs y ont participé : le Professeur Dominique Jeulin, le Dr. Michel Bilodeau, le Dr. Petr Dokladal, le Dr. Estelle Parra-Denis (à l'époque Post-Doctorante) et deux autres Post-Doctorants. Comme, la quantité de données à traiter était très élevée, la réduction des temps de traitement constituait un point critique. Sur ce point, plusieurs axes ont été étudiés. i) Les algorithmes de détection des anomalies en relief ont été accélérés grâce à une approche qui utilisait beaucoup moins de points sur les voisinages des pixels. Cette méthode a été développée avec l'Ecole des Mines de Paris. ii) A l'IPRI, j'ai travaillé sur la réduction de la complexité algorithmique et l'accélération d'algorithmes morphologiques par parallélisation des tâches sur plusieurs cœurs des processeurs et la vectorisation des données. Ces études ont notamment été possibles grâce à un ingénieur-chercheur, Pierre Guilbert, que j'ai encadré pendant une année. Elles ont permis d'accélérer la méthode de seuillage de Köhler qui utilise le contraste des frontières. Puis, une stagiaire Sabina Stephan, a étudié durant six mois une méthode de détection des images du fond d'œil de qualité insuffisante par apprentissage supervisé.

Le stage de Lucas Arnaldo-Alves a aussi permis le développement d'une approche de classification automatique par apprentissage profond des images de la rétine acquises lors du dépistage de la rétinopathie diabétique.

Dans ce chapitre l'analyse des grandes banques d'images a été faite avec des outils morphologiques mais aussi des méthodes d'apprentissage statistiques. La rapidité des algorithmes a aussi été étudiée. Dans le futur, il sera nécessaire de développer des algorithmes plus robustes pour l'analyse des images médicales notamment pour le suivi et la prédiction de la rétinopathie diabétique chez les patients. Ces algorithmes devront aussi être plus explicables d'un point de vue anatomique. Les techniques d'apprentissage profond, même si elles sont très performantes, ne doivent pas faire oublier l'importance de travailler à l'amélioration et la standardisation des images. C'est ce qui sera étudié dans le prochain chapitre.

Chapitre 4

Traitement d'Image Logarithmique et métriques d'Asplund pour l'analyse des images à faible contraste

4.1 Traitement d'Image Logarithmique

4.1.1 Traitement d'Image Logarithmique pour les images à niveaux de gris

Le Traitement d'Image Logarithmique, plus communément appelé *Logarithmic Image Processing* (LIP), a été introduit par Michel Jourlin [143], [150] à partir du raisonnement suivant [144]. Il est bien connu que la *Morphologie Mathématique à niveaux de gris* [111], [112], [201], [261], [275], [277], [284] consiste en une extension aux images à niveaux de gris de la *Morphologie Mathématique binaire* [176], [177], [275]. Cette dernière propose en particulier la dilatation $A \oplus B$ de la forme binaire analysée A par une forme particulière B, appelée élément structurant. Habituellement, la forme B utilisée est un disque afin de préserver l'isotropie de la dilatation. Remarquons qu'un opérateur comme celui-là nécessite deux objets de même nature, i.e. deux formes binaires. Or, ce n'est plus le cas lorsque l'on considère la dilatation à niveaux de gris $f \oplus B$ par un élément structurant binaire, i.e. plat. Au tout début, l'objectif du contexte LIP était de proposer de nouveaux opérateurs « $f \triangle g$ » d'une image à niveaux de gris f, par une autre fonction structurante à niveaux de gris g. La première étape a été de choisir un cas pratique ou l'« addition » de deux images à niveaux de gris possède une signification physique. Pour cette raison, Jourlin et al. [150] ont d'abord considéré les images acquises en transmission, lorsque l'objet est situé entre la source et le capteur. Dans cette situation, il est évidemment possible d'empiler des objets semi-transparents qui sont considérés comme une « addition » d'obstacles dont l'effet est de faire décroître l'intensité de la source de lumière qui les traverse.

Une image à niveaux de gris f est définie sur un domaine $D \subset \mathbb{R}^n$ et prend ses valeurs dans l'échelle de gris $[0; M[\subset \mathbb{R}$. L'image f est un élément de l'ensemble des images $\mathcal{I} = [0; M[^D$. Interprétons maintenant la signification des extrémités 0 et M de l'échelle de gris LIP qui est inversée par rapport à l'échelle habituelle. En effet, comme le modèle LIP a été initialement défini pour les images acquises en transmission, l'extrémité 0 correspond à l'intensité maximale acquise par le capteur (le blanc), i.e. lorsqu'aucun obstacle n'est situé entre la source et le capteur. Dans ce cas, l'objet observé peut être considéré comme entièrement transparent et la valeur 0 apparaît comme l'élément neutre de l'addition logarithmique $\triangle$: $f \triangle 0 = f$. L'autre extrémité de l'échelle, M, correspond au cas limite où aucun élément de la source n'est transmis (le noir). Cette valeur limite est exclue de l'échelle des niveaux de gris. Pour des images numérisées sur 8-bits, M est égal à 256. Après ces considérations sur les échelles de gris, les deux opérations fondamentales du modèle LIP peuvent être définies : l'addition de deux images $f \triangle g$ et la multiplication scalaire $\lambda \triangle f$ d'une image par un nombre réel λ. Ces deux opérations sont déduites de la fameuse loi optique des transmittances [138] :

$$T_{f \triangle g} = T_f \cdot T_g \,. \tag{4.1}$$

La transmittance $T_f(x)$ d'une image à niveaux de gris f en un point x de D est le rapport du flux sortant en x par le flux entrant. Le flux entant correspond à l'intensité de la source. D'un

point de vue mathématique, la transmittance $T_f(x)$ représente la probabilité pour une particule de la source arrivant en x d'être acquise par le capteur. L'équation (4.1) signifie que la probabilité pour une particule émise par la source de passer à travers la « somme » des obstacles générant l'image $f \triangle g$ est égale au produit des probabilités de passer à travers l'obstacle générant f et de celui générant g. La transmittance $T_f(x)$ est liée au niveau de gris $f(x)$, pour tout $x \in D$, par l'équation suivante :

$$T_f(x) = 1 - \frac{f(x)}{M} \,. \tag{4.2}$$

En remplaçant dans l'équation (4.1) les transmittances T_f, T_g et $T_{f \triangle g}$ par leurs expressions de l'équation (4.2), la loi d'addition-LIP de deux images $f \triangle g$ est déduite et s'exprime de la manière suivante :

$$f \triangle g = f + g - \frac{fg}{M} \,. \tag{4.3}$$

En considérant que l'addition $f \triangle f$ s'écrit comme $2 \triangle f$ selon l'équation 4.3, la multiplication d'une image f par un nombre réel λ peut alors être exprimée comme :

$$\lambda \triangle f = M - M\,(1 - f/M)^{\lambda} \,. \tag{4.4}$$

La fonction opposée $\triangle f$ de f se déduit aisément de l'égalité $f \triangle (\triangle f) = 0$, ainsi que la différence $f \triangle g$ entre deux fonctions à niveaux de gris f et g :

$$\triangle f = (-f)/(1 - f/M), \tag{4.5}$$

$$f \triangle g = (f - g)/(1 - g/M). \tag{4.6}$$

Remarquons que $\triangle f$ n'est pas une image, car elle prend des valeurs négatives et $f \triangle g$ est une image si et seulement si $f \geq g$. La fonction opposée $\triangle f$ de l'image $f \in \mathcal{I}$ appartient à l'ensemble $\mathcal{F}_M = \,]-\infty; M[^D$ des fonctions réelles h dont les valeurs sont majorées par M, $h : D \to \,]-\infty; M[$.

Le cadre de travail LIP possède les propriétés fondamentales qui sont listées ci-après.

Propriété 2 (Le cadre de travail LIP est cohérent avec la vision humaine et adapté aux images acquises en lumière réfléchie)**.** *La réponse de l'œil humain aux variations lumineuses δI est connue pour être logarithmique. Elle suit la loi de* Weber-Fechner *définie par*

$$W = \frac{\delta I}{I} \,. \tag{4.7}$$

W est le rapport de Weber qui reste constant pour un grand intervalle d'intensités I. Lorsque l'on observe un objet homogène sur un fond homogène, il signifie que la réponse du système visuel humain apparaît comme proportionnelle aux variations d'intensité δI divisées par l'intensité lumineuse I. Ce ratio est exprimé comme la dérivée d'une fonction logarithmique ce qui indique que la vision humaine est de nature logarithmique. Brailean et al. [37] ont montré expérimentalement que le modèle LIP est cohérent avec la vision humaine. Ce résultat a pour conséquence d'élargir le champ d'application du cadre de travail LIP aux images acquises en lumière réfléchie, sur lesquelles on cherche généralement à appliquer des traitements qui reproduisent ce qu'un œil humain ferait. Les lois du modèle LIP sont donc adaptées au traitement de ces images et simulent l'interprétation d'une image par le système visuel humain [138], [150].

Propriété 3 (De fortes propriétés physiques)**.** *Pour les images acquises en transmission, l'addition-LIP (ou la soustraction-LIP) d'une constante c à une (ou d'une) image f consiste en l'ajout (ou en la soustraction) d'un objet uniforme semi-transparent de niveaux de gris c, qui résulte en un assombrissement ou (un éclaircissement) de f. Ces opérations sont utiles pour corriger les variations d'illumination. Les images acquises en lumière transmise ou réfléchie ont toutes deux les propriétés suivantes.*

• *L'addition (ou la soustraction) d'une constante c à (ou de) f simule la décroissance (ou la croissance) du temps d'exposition de la caméra lors de l'acquisition [45], [65]. Si les valeurs de $f \triangle c$ deviennent strictement négatives, elles se comportent comme des intensificateurs de lumière [136], [137] (Fig. 4.1).*

• *La multiplication scalaire $\lambda \triangle f$ de f par un nombre positif réel λ signifie que l'épaisseur (ou l'absorption) de l'objet semi-transparent qui génère f est LIP-multipliée par λ. L'image est assombrie si $\lambda \geq 1$ ou éclaircie si $\lambda \in [0; 1[$ et reste inchangée lorsque $\lambda = 1$ (Fig. 4.2).*

(a) Image initiale f. (b) Soustraction d'une constante $c = 100$ à f : $f \triangle c$. (c) Ajout d'une constante $c = 100$ à f : $f \triangle c$.

FIG. 4.1 – (a) Image initiale. (b) Eclaircissement de l'image f par soustraction-LIP d'une constante positive, $c = 100$: $f \triangle c$. (c) Assombrissement de l'image f par addition-LIP d'une constante positive, $c = 100$: $f \triangle c$.

(a) Image initiale f. (b) Multiplication de f par un scalaire $\lambda = 0{,}1$: $\lambda \triangle f$. (c) Multiplication de f par un scalaire $\lambda = 3$: $\lambda \triangle f$.

FIG. 4.2 – (a) Image initiale. (b) Eclaircissement de l'image f par multiplication-LIP de f par un scalaire $\lambda = 0{,}1$ compris dans l'intervalle $[0; 1[$: $\lambda \triangle f$. (c) Assombrissement de l'image f par multiplication-LIP de f par un scalaire $\lambda = 3$ supérieur à 1 : $\lambda \triangle f$.

Propriété 4 (De fortes propriétés mathématiques)**.** *Soit $\mathcal{F}_M =]-\infty; M[^D$, l'espace des fonctions définies sur D dont les valeurs sont dans l'intervalle $]-\infty; M[$. $(\mathcal{F}_M, \triangle, \triangle)$ est un espace vectoriel réel et l'espace des images $(\mathcal{I}, \triangle, \triangle)$ représente son cône positif [138], [143].*

La figure 4.3 illustre plusieurs feuilles semi-transparentes empilées les unes sur les autres entre une source de lumière et une caméra [141], [181]. La caméra acquiert une image de la source de lumière à travers les feuilles. L'intensité perçue $f(x)$ par la caméra, en un point x, est tracée comme une fonction du nombre de feuilles. L'échelle grise inversée est utilisée sur l'axe d'intensité. Lorsqu'il n'y a pas d'obstacle, l'intensité est égale à zéro (i.e. le blanc). Elle croît d'une

manière logarithmique avec le nombre de feuilles et atteint un maximum M (i.e. noir), lorsqu'aucune lumière n'est perçue à travers les feuilles. La non-linéarité est prise en compte dans le modèle LIP. Le modèle LIP a été utilisé dans de nombreuses applications médicales et industrielles [137].

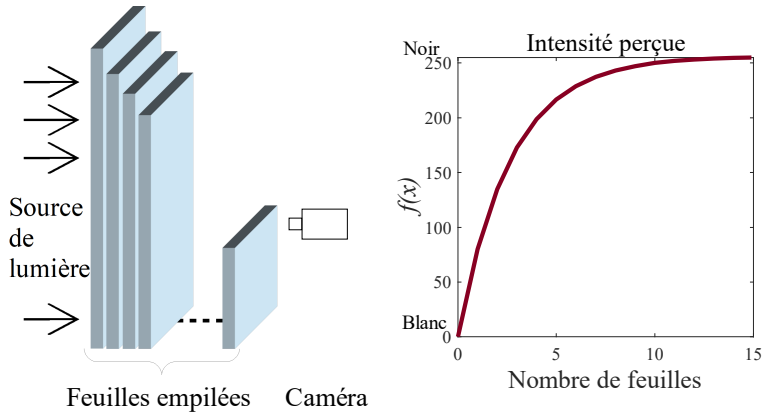


FIG. 4.3 – Les rayons d'une source de lumière passent à travers plusieurs feuilles semi-transparentes empilées. L'intensité $f(x)$ acquise par la caméra est tracée comme une fonction du nombre de feuilles en utilisant l'échelle de gris LIP (i.e. l'échelle inversée).

Remarque. Il existe une version symétrique du LIP, à savoir le *LIP symétrique* [203]. Cependant, ce modèle n'est pas physiquement justifié bien qu'il soit intéressant d'un point de vue mathématique pour sa symétrie. Il permet e.g. de proposer une version LIP de l'opérateur laplacien et de créer des ondelettes logarithmiques [202].

4.1.2 Traitement d'Image Logarithmique pour les images couleur

Jourlin et al. [145] ont défini un modèle de Traitement d'Image Logarithmique couleur, noté LIPC pour *LIP couleur*. Dans cette sous-section, seule une présentation succincte en sera donnée.

Une image couleur $\mathbf{f}$ est définie sur un domaine $D \subset \mathbb{R}^2$, avec des valeurs dans l'intervalle $[0; M[^3$, où l'intervalle $[0; M[$ est un sous-ensemble de l'axe des réels $\mathbb{R}$. Pour des images couleur numérisées sur 24 bits (i.e. 3×8 bits), M est égal à 256. L'image couleur $\mathbf{f}$ s'écrit comme :

$$\mathbf{f} : \begin{cases} D & \rightarrow & [0; M[^3 \\ x & \rightarrow & \mathbf{f}(x) = (f_R(x), f_G(x), f_B(x)) \,. \end{cases} \tag{4.8}$$

f_R , f_G et f_B sont les composantes Rouge, Vert et Bleu de l'image $\mathbf{f}$. $\mathbf{f}(x)$ est un pixel vecteur. $x = x_j$, $j \in \{1; 2...; P\}$, est la coordonnée dans le domaine D du vecteur pixel. P est le cardinal de D, i.e. le nombre de pixels. Une image couleur est un cas particulier d'une image multivariée qui est définie comme $\mathbf{f}_\lambda : D \rightarrow [0; M[^L$, où $L \in \mathbb{N}$ désigne le nombre de composantes [220]. La matrice $\mathbf{F} \in \mathcal{M}_{3,P}([0; M[)$, de taille de $3 \times P$ et à valeurs dans l'intervalle $[0; M[$, est associée à l'image couleur $\mathbf{f}$ et s'écrit :

$$\mathbf{F} = \begin{bmatrix} f_R(x_1) & f_R(x_2) & ... & f_R(x_P) \\ f_G(x_1) & f_G(x_2) & ... & f_G(x_P) \\ f_B(x_1) & f_B(x_2) & ... & f_B(x_P) \end{bmatrix} . \tag{4.9}$$

Chaque ligne F_i, où $i \in [\![1; 3]\!]$, correspond aux pixels de la composante couleur $f_i \in \{f_R, f_G, f_B\}$ de l'image $\mathbf{f}$. Les éléments de la matrice $\mathbf{F}$ sont notés $F_{ij} = f_i(x_j)$, où i est l'indice de la composante et $j \in [\![1; P]\!]$ est l'indice du pixel vecteur. Le mot « image » désigne aussi bien la fonction $\mathbf{f}$ que sa matrice $\mathbf{F}$.

L'espace des images couleur pour les matrices **F** à 24-bits est noté $\mathcal{I}^3(D, [0; M[^3)$ ou plus simplement $\mathcal{I}^3$. D'un point de vue mathématique, le surespace $\mathcal{F}^3(D, [0; +\infty[^3)$, également noté $\mathcal{F}^3$, devient un espace vectoriel pour les lois d'addition $\triangle_c$ et de multiplication $\triangle_c$ du modèle couleur LIPC. L'espace des images $\mathcal{I}^3$ représente son cône positif et les valeurs dans $[M, +\infty[^3$ sont l'opposé des valeurs des couleurs de l'intervalle $[0; M[^3$ [145].

Comme pour le modèle LIP à niveaux de gris, le modèle couleur LIPC repose sur la même approche du signal transmis. Il est donc cohérent avec la vision humaine et adapté aux images acquises par réflexion [145]. Dans ce modèle, une scène est typiquement acquise sur un capteur électronique d'une caméra à travers un filtre de Bayer et enregistrée sur trois composantes, Rouge, Vert et Bleu (R, G et B). L'image est ensuite affichée sur un écran d'ordinateur 8-bits avec les trois mêmes canaux. Dans ce contexte-ci, la transmittance couleur est liée à deux modèles :

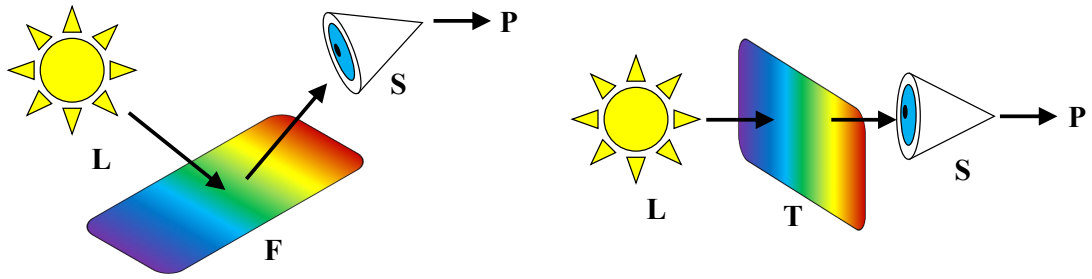
i) d'un côté la perception oculaire d'une image sur un écran d'ordinateur et

ii) de l'autre la perception oculaire de la transmittance grâce à une source de lumière placée derrière la couche semi-transparente.

En outre, afin de conserver l'approche du système visuel humain, le cadre de travail LIPC prend en compte :

- la sensibilité de l'œil humain dans le domaine du visible en utilisant des tables de fonctions de correspondances couleur établies par Stiles et Burch (1959) [292], [293], [310] ;
- la distribution spectrale de la lumière caractérisée par l'illuminant D65 [53], [54], [272].

La figure 4.4, illustre la perception par l'œil humain d'une scène par réflexion ou par transmission.



(a) Lumière réfléchie. (b) Lumière transmise.

FIG. 4.4 – (a) Perception **P** d'une image par l'œil humain : l'image **F** est illuminée par la lumière **L** et vue par le capteur **S**. (b) Système équivalent d'observation par transmission où la perception **P** dépend de la lumière **L**, la transmittance **T** et le capteur **S**.

Contrairement au modèle LIP à niveaux de gris dans lequel l'élément neutre de l'addition est 0 [143], l'élément neutre de l'addition-LIPC est le vecteur constant $[255; 255; 255]^t$ par construction du modèle [145]. La figure 4.5 compare l'échelle classique des gris avec celles des modèles LIP et LIPC.

En lumière réfléchie, l'œil perçoit une image affichée par un dispositif électronique tel qu'un écran d'ordinateur (Fig. 4.4a). L'œil est assimilé à un capteur noté **S** ayant trois canaux S_R, S_G et S_B. La *perception* **P** prend en compte l'image observée $\mathbf{F} \in \mathcal{I}^3$ et la source de lumière L. Le modèle LIPC est défini de manière équivalente en lumière transmise (Fig. 4.4b). Il repose sur la *perception* **P** par un œil humain de la *transmittance* **T** associée à un médium semi-transparent. Il est donc nécessaire de définir la matrice de *perception* **P** et la matrice *transmittance* **T**. Seul le principe est ici rappelé. Le lecteur intéressé par l'explication complète pourra se référer à [145].

Premièrement, la matrice de *perception* **P** est définie pour une image acquise par réflexion (Fig. 4.4a). Le capteur **S** (i.e. l'œil) est divisé en plusieurs cellules $(x, y) \in D$ dont le nombre est égal à celui des pixels dans l'image. Chacune est divisée en trois canaux (R, G et B) et reçoit une part de la lumière L en provenance de l'image **F**. La perception P_i sur le canal i, avec $i \in \{R, G, B\}$, peut être définie comme l'intégration le long du spectre visible de ces différents

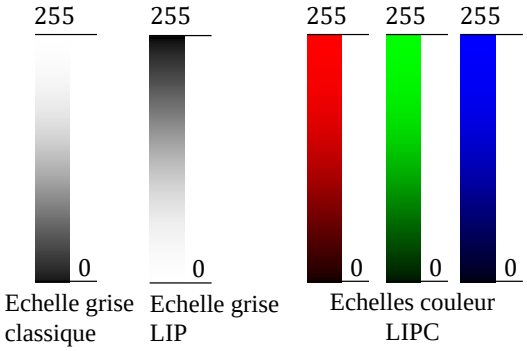


FIG. 4.5 – Différences entre les échelles à niveaux de gris des modèles classique, LIP et LIPC.

éléments :

$$P_i(x,y) = \int_{\lambda=390}^{780} L(\lambda)F_i(x,y)S_i(\lambda)d\lambda, \qquad i \in \{R,G,B\}\,. \tag{4.10}$$

Le comportement de l'œil S dépend des tables de fonctions de correspondances couleur établies par Stiles et Burch (1959). Après intégration, la matrice de *perception* $\mathbf{P}$ est liée à l'image $\mathbf{F}$ par la matrice $\acute{\mathbf{K}}$ de taille 3×3 avec des valeurs réelles, $\acute{\mathbf{K}} \in \mathcal{M}_{3,3}(\mathbb{R})$, grâce à l'équation suivante :

$$\mathbf{P} = \acute{\mathbf{K}}\mathbf{F}. \tag{4.11}$$

Deuxièmement, la matrice de *transmittance* $\mathbf{T}$ est définie en exprimant la matrice de *perception* $\mathbf{P}$ pour une image acquise par transmission. La *perception du capteur* P_i sur le canal i, avec $i \in \{R,G,B\}$, peut être définie comme l'intégration le long du spectre visible du produit : du spectre de la source L (i.e. l'illuminant D65), de la sensibilité du capteur S_i et de la transmittance T :

$$P_i = (M-1)\int_{\lambda=390}^{780} L(\lambda)T(\lambda)S_i(\lambda)d\lambda, \qquad i \in \{R,G,B\}\,. \tag{4.12}$$

Le facteur $(M-1)$ permet d'ajuster l'amplitude de la transmittance comprise dans l'intervalle $[0;1]$ avec la dynamique du capteur comprise dans l'intervalle $[0;M-1]$. Après intégration, la *transmittance* $\mathbf{T} \in \mathcal{F}^3$ d'une image $\mathbf{F}$ est liée à la *perception* $\mathbf{P}$ par la matrice $\acute{\mathbf{U}} \in \mathcal{M}_{3,3}(\mathbb{R})$ grâce à l'équation suivante :

$$\mathbf{T} = \acute{\mathbf{U}}^{-1}\mathbf{P}. \tag{4.13}$$

étudions maintenant quelques propriétés du modèle couleur LIPC. En utilisant les équations (4.11) et (4.13), l'expression de la matrice de *transmittance* $\mathbf{T}$ dépend exclusivement de l'image $\mathbf{F}$ et des matrices $\acute{\mathbf{U}}$ et $\acute{\mathbf{K}}$:

$$\mathbf{T} = \acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{F}. \tag{4.14}$$

Le modèle LIPC dépend donc de deux matrices $\acute{\mathbf{K}}$ et $\acute{\mathbf{U}}$ qui ont été calculées une fois pour toutes dans [145] avec l'illuminant D65 et les tables de fonctions de correspondances couleur établies par Stiles et Burch (1959) :

$$\acute{\mathbf{U}} = \begin{bmatrix} 25{,}044\,0 & 53{,}141\,6 & 176{,}814\,4 \\ 21{,}300\,2 & 185{,}974\,4 & 47{,}725\,4 \\ 229{,}247\,4 & 19{,}994\,4 & 5{,}758\,3 \end{bmatrix}, \tag{4.15}$$

$$\acute{\mathbf{K}} = \begin{bmatrix} 0{,}699\,1 & 0{,}210\,9 & 0{,}089\,9 \\ 0{,}194\,7 & 0{,}800\,2 & 0{,}004\,9 \\ 0{,}068\,1 & 0{,}000\,2 & 0{,}931\,5 \end{bmatrix}. \tag{4.16}$$

Dans le cadre de travail LIPC, et d'une manière générale en signal transmis, la transmittance de la somme de deux images $\mathbf{F}$ et $\mathbf{G} \in \mathcal{I}^3$ est égale au produit de leur transmittance $\mathbf{T_F}$ et $\mathbf{T_G}$:

$$\mathbf{T}_{\mathbf{F}\mathbb{A}_c G} = \mathbf{T_F} \cdot \mathbf{T_G}. \tag{4.17}$$

L'addition-LIPC est notée $\triangle\!\!\!\!+_c$. Le symbole $\cdot$ désigne le produit matriciel de Hadamard également appelé produit « point à point ». De l'équation (4.13) est déduite la matrice de *perception* pour la somme-LIPC de deux images $\mathbf{F}$ et $\mathbf{G} \in \mathcal{I}^3$:

$$\mathbf{P}_{\mathbf{F} \triangle\!\!\!\!+_c G} = \acute{\mathbf{U}}(\mathbf{T}_{\mathbf{F}} \cdot \mathbf{T}_{\mathbf{G}}).$$

Grâce aux équations (4.11) et (4.14), l'expression de l'addition-LIPC, $\mathbf{F} \triangle\!\!\!\!+_c \mathbf{G}$, entre deux images $\mathbf{F}$ et $\mathbf{G}$ est obtenue :

$$\mathbf{F} \triangle\!\!\!\!+_c \mathbf{G} = \acute{\mathbf{K}}^{-1}\acute{\mathbf{U}}(\acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{F} \cdot \acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{G}). \tag{4.18}$$

De cette loi d'addition, une loi de multiplication par un scalaire $\triangle\!\!\!\!+_c$ est déduite pour chaque nombre réel α :

$$\alpha \triangle\!\!\!\!+_c \mathbf{F} = \acute{\mathbf{K}}^{-1}\acute{\mathbf{U}}(\acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{F})^{\alpha}. \tag{4.19}$$

La soustraction $\mathbf{F} \triangle_c \mathbf{G}$ entre deux images $\mathbf{F}$ et $\mathbf{G}$ est définie par :

$$\mathbf{F} \triangle_c \mathbf{G} = \mathbf{F} \triangle\!\!\!\!+_c (-1 \triangle\!\!\!\!+_c \mathbf{G}) = \acute{\mathbf{K}}^{-1}\acute{\mathbf{U}} \left(\frac{\acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{F}}{\acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{G}} \right). \tag{4.20}$$

La fraction entre les deux matrices correspond à la division entre les éléments des deux matrices (i.e. la division de Hadamard). Une condition pour que la différence-LIPC, $\mathbf{F} \triangle_c \mathbf{G}$, soit une image couleur est que les valeurs de $\mathbf{G}$ soient plus grandes que celles de $\mathbf{F}$, sinon le résultat n'est pas dans l'intervalle des valeurs des images $[0; M[^3$ mais dans celui des fonctions tri-variées à valeurs positives $[0; +\infty[^3$. Dans ce dernier cas, le résultat peut être tronqué à la valeur $M - 1$, bien que de l'information soit alors perdue. Présentons maintenant quelques propriétés du modèle LIPC.

- L'élément neutre de l'addition-LIPC est défini par $\mathbf{F} \triangle\!\!\!\!+_c \mathbf{G}_0 = \mathbf{F}$ où $\mathbf{G}_0 = \mathbf{255}_{(3,P)}$ est une image constante qui est égale 255 (i.e. $M - 1$) pour les images couleur numérisées sur 24 bits (i.e. 3×8 bits).
- L'élément opposé de l'addition-LIPC est défini par $\forall \mathbf{F} \in \mathcal{I}^3$, $\mathbf{F} \triangle\!\!\!\!+_c (\triangle_c \mathbf{F}) = \mathbf{G}_0$ où $\triangle_c \mathbf{F} = \acute{\mathbf{K}}^{-1}\acute{\mathbf{U}} * \left(\frac{1}{\acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{F}}\right)$. Les valeurs de $\triangle_c \mathbf{F}$ sont dans l'intervalle $[M; +\infty[^3$.

Comme pour le modèle LIP à niveaux de gris, la soustraction-LIPC d'une constante de l'intervalle $[0; M[$ à une image permet d'éclaircir l'image en simulant une augmentation du temps d'exposition de la caméra. De même, l'addition-LIPC de cette constante à l'image l'assombrie en simulant une réduction du temps d'exposition (Fig. 4.6). La multiplication-LIPC d'une image par un scalaire $\lambda \in [0; 1[$ permet d'éclaircir cette image en simulant un amincissement (ou une diminution de l'opacité) de la lame semi-transparente associée à la scène. De même, la multiplication-LIPC par un scalaire $\lambda > 1$ l'assombrit en simulant un épaississement de la lame semi-transparente (Fig. 4.7).

(a) Image initiale f. (b) Soustraction d'une constante $c = 200$ à f : $f \triangle_c c$. (c) Ajout d'une constante $c = 200$ à f : $f \triangle_c c$.

FIG. 4.6 – (a) Image initiale. (b) Eclaircissement de l'image f par soustraction-LIPC d'une constante positive, $c = 200$: $f \triangle_c c$. (c) Assombrissement de l'image f par addition-LIPC d'une constante positive, $c = 200$: $f \triangle_c c$.

(a) Image initiale f. (b) Multiplication de f par un scalaire $\lambda = 0{,}5$: $\lambda \triangle_c f$. (c) Multiplication de f par un scalaire $\lambda = 3$: $\lambda \triangle_c f$.

FIG. 4.7 – (a) Image initiale. (b) Eclaircissement de l'image f par multiplication-LIPC de f par un scalaire $\lambda = 0{,}5$ compris dans l'intervalle $[0; 1[$: $\lambda \triangle_c f$. (c) Assombrissement de l'image f par multiplication-LIPC de f par un scalaire $\lambda = 3$ supérieur à 1 : $\lambda \triangle_c f$.

4.2 Amélioration des contrastes

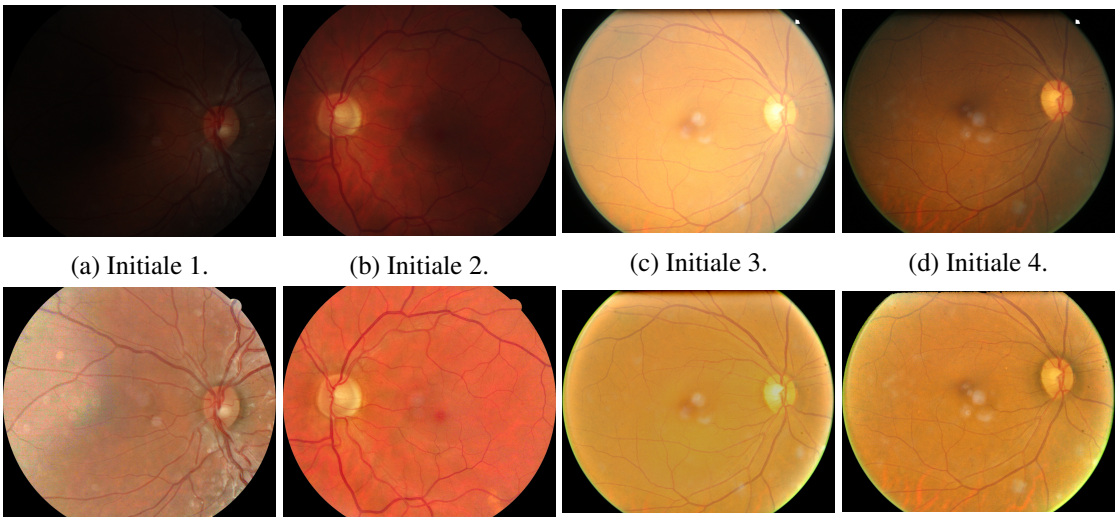
Afin d'analyser les grandes banques de données d'images du fond d'œil, il est rapidement apparu que nos outils devaient être robustes aux variations d'éclairement. En effet, ces conditions varient et dépendent de l'illuminant lors de la prise de vue, de la dynamique du capteur utilisée, de la lumière absorbée par l'objet, etc.

4.2.1 Application aux images du fond d'œil

Dans les banques de données de Santé Publique, les photographies de la rétine sont prises à une fréquence annuelle chez les patients diabétiques. Dans de nombreux pays, l'existence de programme de dépistages depuis plus de dix ans a conduit à la création de grandes banques d'images [102]. Cependant la luminosité des images du fond d'œil n'est pas uniforme pour différentes raisons : des maladies telles que la cataracte, les mouvements du patient, les conditions d'acquisitions différentes (e.g. caméras différentes, éclairement différents, dilatation ou non de la pupille) et des différences dans l'absorption de la lumière dans l'œil [313]. *La normalisation des contrastes dans les images est donc cruciale pour pouvoir leur appliquer des traitements informatiques comparables.* Dans les figures 4.8a, 4.8b, 4.8c et 4.8d on peut remarquer que certaines parties des images

apparaissent claires tandis que d'autres sont sombres. En outre, le changement éventuel de la caméra du fond d'œil entre deux examens séparés peut conduire à un changement de la couleur entre deux images du même œil. E.g. les examens des figures 4.8c et 4.8d illustrent ce changement de couleur car ils ont été effectués sur le même œil avec un intervalle de 1,5 an.

J'ai développé une méthode pour normaliser ces changements de contraste dans les images du fond d'œil [234]. Cette méthode utilise le modèle LIP qui est cohérent avec les principes optiques de la formation de l'image. Les contrastes des parties sombres et claires des images sont ajustés d'une façon qui donne un aspect couleur similaire aux lésions telles que les micro-anévrismes ou aux structures anatomiques tels que les vaisseaux ou les veines. Après normalisation du contraste (Fig. 4.8e, 4.8f, 4.8g et 4.8h), les images améliorées deviennent plus faciles à lire par un spécialiste ou plus faciles à traiter par un algorithme. Notre méthode couleur va plus loin que les méthodes précédentes à niveaux de gris qui utilisaient un ajustement polynomial [314], la morphologie mathématique [336] ou l'égalisation d'histogramme [340].



(a) Initiale 1. (b) Initiale 2. (c) Initiale 3. (d) Initiale 4.

(e) Améliorée 1. (f) Améliorée 2. (g) Améliorée 3. (h) Améliorée 4.

FIG. 4.8 – (a) et (b) Images du fond d'œil acquises dans de bonnes conditions. (c) et (d) Images du fond d'œil du même patient acquises durant deux examens avec un intervalle de 1,5 an et dans des conditions plus difficiles. (e), (f), (g) et (h) Images améliorées après normalisation de contraste des images (a), (b), (c) et (d).

Les résultats de la méthode ont été présentés dans deux conférences médicales internationales : *World Diabetes Congress – IDF 2015* de l'*International Diabetes Federation* [233] et l'*American Diabetes Association 77th Scientific Sessions – ADA2017* [232]. Un article a aussi été publié dans un journal international [234]. De plus, grâce à cette méthode, j'ai obtenu trois financements (cf. sous-section 2.2, p. 239) :

- deux de la Banque Publique d'Investissement France (BPI France) de 5 k€ et 18 k€ ;
- un projet européen, *Eye-Light* du programme Horizon2020 de l'Union Européenne (financement de 50 k€) [1]. L'objectif de ce projet était de réaliser une étude de faisabilité pour un service d'amélioration d'image pour le diagnostic de la rétinopathie diabétique. Une attention particulière était portée aux pays en voie de développement.

Afin de transférer cette méthode en clinique et d'obtenir une certification européenne comme *dispositif médical*, une étude clinique a été planifiée. Le but était de montrer que ce système rend le diagnostic de la rétinopathie meilleur (i.e. plus évident) et plus rapide. Pour ce projet, j'ai d'abord collaboré avec l'équipe du Professeur David Owens, diabétologue à *Swansea University* et à *Diabetic Eye Screening Wales* au Pays de Galles, Royaume-Uni. Puis, j'ai établi une collaboration

1. `https://cordis.europa.eu/project/id/717108/fr`

avec l'équipe du Professeur Laurent Kodjikian, ophtalmologiste aux Hospices Civils de Lyon et à l'Université Claude Bernard Lyon 1.

4.2.2 Amélioration du contraste global des images à niveaux de gris

L'approche présentée par Jourlin et al. dans [137], [145] consiste à étendre la dynamique globale d'une image en retranchant au sens LIP une constante à cette image. Ceci a pour effet d'améliorer son contraste. La soustraction-LIP d'une constante $c \in [0; M[$ agit donc comme un intensificateur de lumière.

Soient $f : D \to [0; M[$ une image, $a = f(x_a) = \sup\{f(x), x \in D\}$ son supremum et $b = f(x_b) = \inf\{f(x), x \in D\}$ son infimum. x_a et $x_b \in D$ sont les coordonnées respectives de ces extrema. La dynamique de l'image f est définie par [137] :

$$DR(f) = a - b. \tag{4.21}$$

Comme $a > b$ cela implique que la dynamique de l'image à améliorer est strictement positive. Lorsqu'une constante $c \in [0; M[$ est soustraite au sens LIP, la dynamique est égale à :

$$DR(f \triangle c) = (a \triangle c) - (b \triangle c) = \frac{a-b}{1 - \frac{c}{M}}\,. \tag{4.22}$$

Comme la constante c est dans l'intervalle $[0; M[$, la dynamique de l'image translatée $f \triangle c$ est supérieure ou égale à la dynamique de l'image initiale f : $DR(f \triangle c) \geq DR(f)$. Lorsque c tend vers M, la dynamique de l'image translatée $DR(f \triangle c)$ tend vers l'infini. Il existe donc une constante c telle que la dynamique soit égale à une valeur Γ donnée, avec e.g. $\Gamma = M - 1$:

$$DR(f \triangle c) = M - 1 = \Gamma. \tag{4.23}$$

D'après l'équation (4.22), cela implique que :

$$c = M\left(1 - \frac{a-b}{\Gamma}\right). \tag{4.24}$$

Le minimum $b \triangle c$ de l'image $f \triangle c$ étant négatif et afin d'avoir des valeurs dans l'intervalle $[0; M[$, l'image améliorée f_E est définie comme :

$$f_E = f \triangle c - b \triangle c. \tag{4.25}$$

La figure 4.9 montre un exemple d'amélioration de la dynamique par soustraction de constante. L'histogramme est plus diffus pour l'image améliorée (Fig. 4.9c et 4.9d) que pour l'image initiale (Fig. 4.9a et 4.9b). Les distances entre les barres de l'histogramme ont augmenté parce que les niveaux de gris sont distribués différemment sur l'échelle de gris.

4.2.3 Amélioration du contraste global des images couleur

L'approche d'amélioration du contraste par amélioration de la dynamique a été étendue aux images couleur par Jourlin et al. dans [137], [145]. Tout d'abord les extrema de l'image sont calculés :

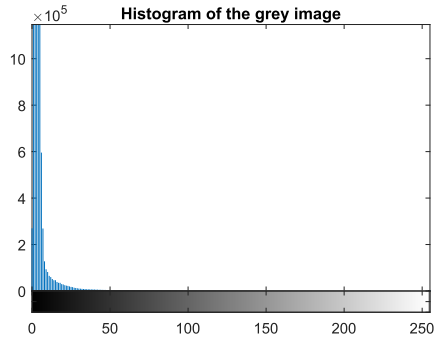
$$\sup(\mathbf{f}) = A = \sup_{x \in D} \{F_R(x), F_G(x), F_B(x)\}\,, \tag{4.26}$$

$$\inf(\mathbf{f}) = B = \inf_{x \in D} \{F_R(x), F_G(x), F_B(x)\}\,. \tag{4.27}$$

La dynamique d'une image couleur $\mathbf{f} \in \mathcal{I}^3$ est définie par :

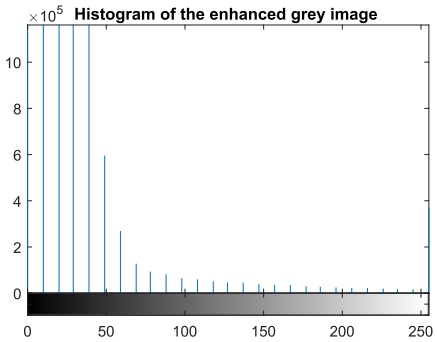
$$DR(\mathbf{f}) = \sup(\mathbf{f}) - \inf(\mathbf{f}). \tag{4.28}$$

(a) Image initiale.



(b) Histogramme de l'image initiale.

(c) Image améliorée.



(d) Histogramme de l'image améliorée.

FIG. 4.9 – (a) Image sombre initiale et (b) son histogramme. (c) Image améliorée par soustraction de constante et (c) son histogramme.

Elle est égale à :

$$DR(\mathbf{f}) = \sup\{DR(f_R), DR(f_G), DR(f_B)\}. \tag{4.29}$$

Elle fait partie de l'intervalle $[0; M[$.

Définissons maintenant les opérateurs de « duplication » suivants. L'opérateur $[.]_N :\rightarrow \mathbb{R}^N$ transforme un scalaire réel x en un vecteur de valeur constante égale à x et de taille $N \in \mathbb{N}$:

$$[.]_N : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R}^N \\ x & \rightarrow & [x]_N = [x, x...x]^t_{(N \text{ fois})}. \end{cases} \tag{4.30}$$

L'opérateur ^{t} note la transposition d'un vecteur ou d'une matrice. L'opérateur $[.]_{N,M} : \mathbb{R} \rightarrow \mathcal{M}_{N,M}(\mathbb{R})$ transforme un scalaire réel x en une matrice de valeur constante égale à x et de taille $N \times M$:

$$[.]_{N,M} : \begin{cases} \mathbb{R} & \rightarrow & \mathcal{M}_{N,M}(\mathbb{R}) \\ x & \rightarrow & [x]_{N,M} = \begin{bmatrix} x^{1,1} & x^{1,2} & \dots & x^{1,M} \\ \vdots & & \ddots & \vdots \\ x^{N,1} & x^{N,2} & \dots & x^{N,M} \end{bmatrix}. \end{cases} \tag{4.31}$$

Soient une matrice constante $\mathbf{C}_i = [C]_{3,P} \in \mathcal{M}_{3,P}([0; M[)$, et un vecteur constant $\mathbf{C}_v = [C]_3 \in [0; M[^3$. Lorsqu'une matrice constante $\mathbf{C}_i$ est soustraite au sens LIPC d'une image couleur $\mathbf{F} \in \mathcal{I}^3$, la dynamique de l'image devient :

$$DR(\mathbf{F} \triangle_c \mathbf{C}_i) = (\mathbf{A} \triangle_c \mathbf{C}_v)(1) - (\mathbf{B} \triangle_c \mathbf{C}_v)(1), \tag{4.32}$$

où $\mathbf{A} = [A]_3$, $\mathbf{B} = [B]_3$ et $\mathbf{C}_v(1) = C$. A des fins de simplification, nous utiliserons $\mathbf{C}$ au lieu de la matrice constante $\mathbf{C}_i$ ou du vecteur constant $\mathbf{C}_v$

$$DR(\mathbf{F} \triangle_c \mathbf{C}) = (\mathbf{A} \triangle_c \mathbf{C})(1) - (\mathbf{B} \triangle_c \mathbf{C})(1). \tag{4.33}$$

Comme les vecteurs $\mathbf{A}$, $\mathbf{B}$ et $\mathbf{C}$ ont des valeurs constantes, leur première composante sera utilisée (e.g. $(\mathbf{A} \triangle_c \mathbf{C})(1)$), sans réduire le caractère général de l'expression de $DR(\mathbf{F} \triangle_c \mathbf{C})$. De plus, nous avons $(\mathbf{A} \triangle_c C)(1) = \frac{A}{C}(M-1)$ (cf. démonstration ci-dessous). L'expression de la dynamique peut donc être simplifiée :

$$DR(\mathbf{F} \triangle_c \mathbf{C}) = \frac{A-B}{C}(M-1). \tag{4.34}$$

Comme $C \in [0; M-1]$, la dynamique de l'image translatée $f \triangle_c C$ est supérieure ou égale à la dynamique de l'image initiale f : $DR(f \triangle C) \geq DR(f) = A - C$.

Preuve de l'expression de la différence LIPC entre deux vecteurs constants. Soient $\mathbf{A} = [A]_3$ et $\mathbf{C} = [C]_3$ deux vecteurs constants de l'intervalle $[0; +\infty[^3$ et la matrice $\mathbf{\Omega} = \acute{\mathbf{K}}^{-1}\acute{\mathbf{U}} = [\omega_{ij}]_{i,j \in [\![1;3]\!]}$. D'après l'équation (4.20), p. 95, nous avons :

$$\begin{aligned}
\mathbf{A} \triangle_c \mathbf{C} &= \acute{\mathbf{K}}^{-1}\acute{\mathbf{U}} \left(\frac{\acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{A}}{\acute{\mathbf{U}}^{-1}\acute{\mathbf{K}}\mathbf{C}} \right) \\
&= \mathbf{\Omega} \left(\frac{\mathbf{\Omega}^{-1}\mathbf{A}}{\mathbf{\Omega}^{-1}\mathbf{C}} \right) \\
&= \frac{A}{C}\mathbf{\Omega} \left(\frac{\mathbf{\Omega}^{-1}[1]_3}{\mathbf{\Omega}^{-1}[1]_3} \right) \\
&= \frac{A}{C}\mathbf{\Omega}[1]_3 \\
&= \left[\frac{A}{C} \sum_{j=1}^{3} \omega_{1,j} \right]_3 \\
&= \left[\frac{A}{C}(M-1) \right]_3 .
\end{aligned} \tag{4.35}$$

Sans réduire le caractère général de l'équation $\mathbf{\Omega}[1]_3 = \left[\sum_{j=1}^{3} \omega_{1,j}\right]_3$, nous choisissons la somme de la première ligne de la matrice $\mathbf{\Omega}$ car $\forall i \in [\![1;3]\!]$, $\sum_{j=1}^{3} \omega_{ij} = M-1$. Ce résultat est dû aux propriétés de normalisation des matrices $\acute{\mathbf{K}}$ et $\acute{\mathbf{U}}$ [145]. □

La dynamique $DR(\mathbf{F} \triangle_c \mathbf{C})$ de l'image translatée (éq. (4.34)) est étendue à l'amplitude maximale, $\Gamma = M-1$, en soustrayant une matrice constante $\mathbf{C}_0 = [C_0]_{3,P}$ telle que :

$$DR(\mathbf{F} \triangle_c \mathbf{C}_0) = \Gamma. \tag{4.36}$$

D'après les équations (4.36) et (4.34), l'expression de la constante C_0 est égale à :

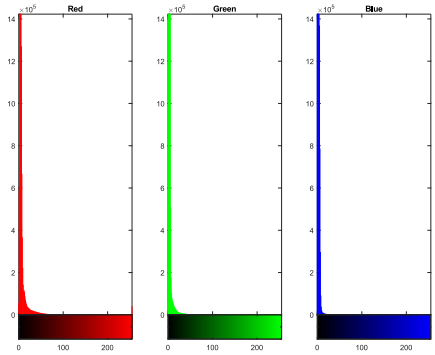
$$\begin{aligned}
C_0 &= \frac{A-B}{\Gamma}(M-1) \\
&= A - B.
\end{aligned} \tag{4.37}$$

Afin d'être affiché dans l'intervalle $[0; M[^3$, le résultat $\mathbf{F} \triangle_c \mathbf{C}_0$, est translaté par son minimum $-\inf(\mathbf{F} \triangle_c \mathbf{C}_0)$. Ce minimum est aussi égal à $-(\mathbf{B} \triangle_c \mathbf{C}_0)(1)$, car $\mathbf{B}$ et $\mathbf{C}_0$ sont des vecteurs constants. Ceci conduit à la définition de l'image couleur améliorée :

$$\begin{aligned}
\mathbf{F}_E &= \mathbf{F} \triangle_c \mathbf{C}_0 - (\mathbf{B} \triangle_c \mathbf{C}_0) \\
&= \mathbf{F} \triangle_c \mathbf{C}_0 - \left[\frac{B}{C_0}(M-1) \right]_{3,P} \\
&= \mathbf{F} \triangle_c \mathbf{C}_0 - \left[\frac{B}{A-B}(M-1) \right]_{3,P} .
\end{aligned} \tag{4.38}$$

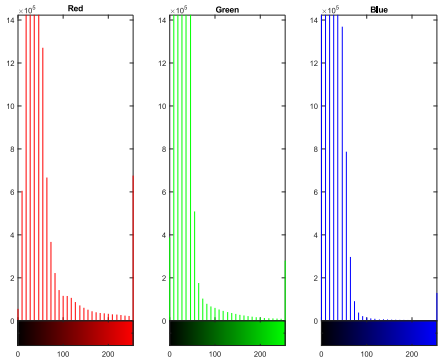
La figure 4.10 montre un exemple d'amélioration de la dynamique de l'image par soustraction-LIPC d'une constante. Les histogrammes des composantes couleur de l'image améliorée sont plus diffus (Fig. 4.10c et 4.10d) que ceux de l'image originale (Fig. 4.10a et 4.10b). Comme pour la méthode d'amélioration à niveaux de gris, les distances entre les barres de l'histogramme ont augmenté car les couleurs sont distribuées différemment le long des échelles de gris de chaque composante.

(a) Image initiale.



(b) Histogrammes de l'image initiale.

(c) Image améliorée.



(d) Histogrammes de l'image améliorée.

FIG. 4.10 – (a) Image couleur sombre initiale et (b) ses trois histogrammes de chacune des composantes couleur. (c) Image couleur améliorée par soustraction de constante et (c) ses trois histogrammes.

Afin d'améliorer le contraste des images du fond d'œil présentées sur la figure 4.8, p. 97, j'ai appliqué la méthode par amélioration de la dynamique de manière locale sur les images. La mise au point de cette méthode a néanmoins nécessité des développements mathématiques supplémentaires ainsi que de nombreux tests avant d'obtenir une amélioration efficace des images du fond d'œil.

4.2.4 Amélioration du contraste local des images à niveaux de gris

Cette sous-section n'est pas présentée dans le mémoire en ligne.

4.2.5 Amélioration du contraste local des images couleur

Cette sous-section n'est pas présentée dans le mémoire en ligne.

4.3 Les métriques d'Asplund

Les métriques d'Asplund ont été initialement définies sur les ensembles par Edgar Asplund [15], [96]. Elles ont ensuite été étendues aux images à niveaux de gris par Jourlin et al. [140], [142], [146] dans le cadre du Traitement d'Image Logarithmique. Avec Michel Jourlin, j'ai proposé leur extension aux images couleur et multivariées [209], [229] et nous avons démontré le lien existant avec la Morphologie Mathématique [216], [227], [228], [231].

Les métriques ou leurs valeurs, à savoir les distances, jouent un rôle central en analyse d'image comme objets de comparaison. Elles possèdent de fortes propriétés mathématiques (symétrie, positivité, séparation et inégalité triangulaire), mais elles ne sont généralement pas fondées sur des propriétés optiques. Elles ne sont donc pas adaptées aux images capturées dans des conditions variables d'éclairement. Ce problème affecte la plupart des métriques classiques telles que : les distances de type euclidien [162], [316], la métrique intégrale (correspondant à la norme un), la métrique uniforme (correspondant à la norme infinie), la distance de Stepanov (ou distance de Minkowski) [66] (correspondant à la norme p), la métrique d'Hausdorff [70], [124] ainsi que beaucoup d'autres [66]. Le but de cette section est de proposer des « outils métriques » qui sont robustes aux variations d'éclairement.

Nous partons d'une métrique peu connue définie pour une paire (A, B) de formes binaires, à savoir la métrique d'Asplund [15], [96]. Elle consiste en une double palpation d'une des formes par l'autre ; e.g. B est choisie pour effectuer la double palpation de A par le moyen de deux formes *homothétiques* de B. Pour ce faire, deux nombres réels sont calculés : le plus petit nombre λ_0 tel que $\lambda_0 B$ contienne A et le plus grand nombre μ_0 telle que A contienne $\mu_0 B$. La distance d'Asplund $d_{asp}(A, B)$ entre A et B est alors définie comme :

$$d_{asp}(A, B) = \ln\left(\frac{\lambda_0}{\mu_0}\right). \tag{4.39}$$

Remarque. Cette définition implique que d_{asp} reste inchangée lorsqu'une forme est agrandie ou réduite par tout ratio $k \in \mathbb{R}^+$.

La figure 4.11 illustre la métrique d'Asplund binaire où une forme A est palpée sur ces deux faces par une forme de référence B.

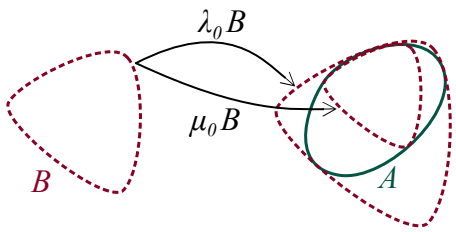


FIG. 4.11 – Métrique d'Asplund pour des formes binaires. La forme A est palpée sur ces deux faces par la forme de référence B grâce à ses deux homothétiques $\lambda_0 B$ et $\mu_0 B$.

4.3.1 Travaux connexes

Les méthodes de mise en correspondance de formes par double palpation ne sont pas toujours efficaces sur des images acquises sous éclairement variable. Parmi la littérature existante, nous pouvons notamment citer les deux papiers de Barat et al. [20], [21]. Le premier propose le concept de palpation afin d'unifier les méthodes suivantes :

(i) l'approche morphologique proposée par Banon et Faria [18] ;

(ii) l'extension de la transformée « hit-or-miss » [275] aux images à niveaux de gris introduite par Khosravi et Schafer [152] ;

(iii) l'approche inspirée du calcul de la distance d'Hausdorff de Odone et al. [236].

Dans le deuxième papier, Barat et al. [21] introduisent la palpation double-face d'une fonction à niveaux de gris f, où une paire de sondes entoure localement la surface représentative de f. Cette technique détecte les lieux possibles de la forme recherchée. Cependant, les sondes sont choisies arbitrairement et simplement translatées le long de l'axe des niveaux de gris pour rester en contact avec la fonction f. Ces translations, qui assombrissent (ou éclaircissent) les sondes, semblent tenir compte des variations d'éclairement de l'image. Cependant, comme elles ne sont pas fondées au sens physique, elles ne modélisent pas correctement ces variations.

4.3.2 Les métriques d'Asplund à niveaux de gris

La métrique d'Asplund binaire a la propriété remarquable d'être *insensible au grossissement de l'objet* [142]. La motivation de Jourlin et al. a été d'étendre cette propriété à des métriques de type Asplund dédiées aux images à niveaux de gris. Ces extensions de la métrique d'Asplund binaire au cas fonctionnel nécessitent l'usage d'*images homothétiques* ou de fonctions qui sont *insensibles aux variations d'éclairement*. Ce concept est mathématiquement bien défini et physiquement justifié dans le cadre du Traitement d'Image Logarithmique (LIP). Deux extensions seront étudiées :

i) la métrique d'Asplund LIP-multiplicative définie grâce à l'opération de LIP-multiplication $\triangle$ d'une image par un nombre réel ;

ii) la métrique d'Asplund LIP-additive définie grâce à l'opération de LIP-addition $\triangle$ d'une constante à une image [138].

Le modèle LIP donne aux métriques fonctionnelles d'Asplund une forte propriété optique, i.e. une très faible sensibilité aux variations d'éclairement, particulièrement pour les images sous-éclairées. La métrique d'Asplund LIP-multiplicative est ainsi théoriquement insensible aux variations d'absorption (ou d'opacité) de l'objet acquis en lumière transmise. La métrique LIP-additive est quant à elle théoriquement insensible aux variations d'intensité de la source (ou du temps d'exposition de la caméra). Afin de reconnaître des formes, une carte des distances d'Asplund est calculée entre un gabarit et une image. Les formes de l'image les plus proches du gabarit correspondent aux valeurs minimales de la carte. Les formes sont alors détectées en trouvant les minima de la carte.

Métrique d'Asplund LIP-multiplicative

Soient $]0; M[$ l'axe des niveaux de gris sans la valeur nulle et $\mathcal{I}^* =]0; M[^D$ l'ensemble des images avec des valeurs strictement positives.

Définition 12 (Métrique d'Asplund LIP-multiplicative [142], [146])**.** *Soient f et $g \in \mathcal{I}^*$ deux images à niveaux de gris. Comme pour les formes binaires, nous choisissons une fonction de palpation ou sonde, e.g. g, et nous définissons les deux nombres : $\lambda = \inf\{\alpha, f \leq \alpha \triangle g\}$ et $\mu = \sup\{\alpha, \alpha \triangle g \leq f\}$. La métrique d'Asplund LIP-multiplicative $d_{asp}^{\triangle}$ est définie par :*

$$d_{asp}^{\triangle}(f, g) = \ln\left(\frac{\lambda}{\mu}\right). \tag{4.40}$$

La figure 4.12 illustre la double palpation d'une image f par une sonde g. La sonde inférieure $\mu \triangle g$ est en contact avec la face inférieure de la fonction f tandis que la sonde supérieure $\lambda \triangle g$ est en contact avec la face supérieure de f. Par comparaison avec la forme de la sonde, les formes des sondes inférieures et supérieures sont déformées par la multiplication-LIP $\triangle$. Cette déformation, qui dépend de la valeur de gris de g, est caractéristique de la propriété suivante de la métrique.

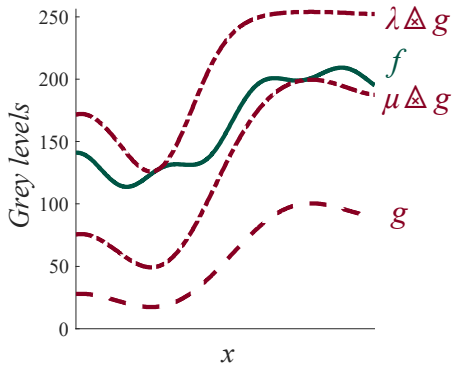


FIG. 4.12 – Palpation double-face d'une fonction f par une sonde g effectuée par la métrique d'Asplund LIP-multiplicative. $\mu \triangle g$ est la sonde inférieure et $\lambda \triangle g$ est la sonde supérieure.

Propriété 5 (Invariance par multiplication-LIP d'une constante)**.** *L'avantage le plus important de la métrique LIP-multiplicative est de rester inchangée lorsqu'une image (e.g. f) est remplacée par tout homothétique $\alpha \triangle f$, $\forall \alpha \in \mathbb{R}^{*+}$:*

$$d_{asp}^{\triangle}(f, g) = d_{asp}^{\triangle}(\alpha \triangle f, g). \tag{4.41}$$

Cette dernière équation démontre l'insensibilité de la métrique $d_{asp}^{\triangle}$ aux variations d'illuminations modélisées par la loi LIP-multiplicative $\triangle$, i.e. celles qui correspondent à un changement d'absorption ou d'opacité de l'objet.

Remarque (*Annexe mathématique [140]*). Etant donnée la propriété précédente, il serait plus rigoureux d'expliquer la métrique d'Asplund LIP-multiplicative comme il s'ensuit.

a) Une relation d'équivalence $\mathscr{R}$ est définie sur l'espace des images à niveaux de gris. Deux images f et $g \in \mathcal{I}^*$ sont dites en relation si elles satisfont :

$$(f \mathscr{R} g) \iff \exists \alpha > 0,\ f = \alpha \triangle g. \tag{4.42}$$

La relation précédente $\mathscr{R}$ satisfait manifestement les propriétés d'une relation d'équivalence :

i) réflexivité : $\forall f \in \mathcal{I}, f \mathscr{R} f$,

ii) symétrie : $\forall (f, g) \in \mathcal{I}^2$, $f \mathscr{R} g \iff g \mathscr{R} f$ et

iii) transitivité : $\forall (f, g, h) \in \mathcal{I}^3$, $(f \mathscr{R} g$ et $g \mathscr{R} h) \implies f \mathscr{R} h$.

b) A chaque image $f \in \mathcal{I}^*$, une classe d'équivalence $f^{\triangle}$ est associée : $f^{\triangle} = \{g,\ g \mathscr{R} f\}$.

c) Une définition rigoureuse de la métrique d'Asplund LIP-multiplicative est alors donnée dans l'espace des classes d'équivalence $\mathcal{I}^{\triangle}$ par :

$$\forall \left(f^{\triangle}, g^{\triangle}\right) \in \left(\mathcal{I}^{\triangle}\right)^2, \quad d_{asp}^{\triangle}(f^{\triangle}, g^{\triangle}) = d_{asp}^{\triangle}(f_1, g_1), \tag{4.43}$$

où $d_{asp}^{\triangle}(f_1, g_1)$ est la distance d'Asplund quels que soient les éléments f_1 et g_1 des classes d'équivalences $f^{\triangle}$ et $g^{\triangle}$.

Dans [228], [231], les propriétés de la métrique d'Asplund (i.e. symétrie, positivité, séparation et inégalité triangulaire) sont vérifiées dans l'espace des classes d'équivalences $\mathcal{I}^{\triangle}$.

Métrique d'Asplund LIP-additive

Soit $\overline{\mathcal{F}}_M = [-\infty; M]^D$ l'ensemble des fonctions dont les valeurs sont dans l'intervalle $[-\infty; M]$.

Définition 13 (Métrique d'Asplund LIP-additive [140])**.** *Soient f et $g \in \mathcal{F}_M$ deux fonctions. Nous choisissons une fonction de palpation, e.g. g, et nous définissons les deux nombres : $c_1 = \inf\{c, f \leq c \triangle g\}$ et $c_2 = \sup\{c, c \triangle g \leq f\}$, où c est dans l'intervalle $]-\infty; M[$. La métrique d'Asplund LIP-additive $d^{\triangle}_{asp}$ est définie par :*

$$d^{\triangle}_{asp}(f, g) = c_1 \triangle c_2. \tag{4.44}$$

Remarque. Par définition, les valeurs de c_1 et c_2 sont dans l'intervalle $]-\infty; M[$, ce qui implique que les fonctions de palpation $g \triangle c_1$ et $g \triangle c_2$ ne sont pas toujours des images, i.e. leurs valeurs peuvent être dans l'intervalle $]-\infty; M[$. Cependant, c_1 est toujours plus grande que c_2. Toutefois, la métrique d'Asplund a la propriété suivante.

Proposition 1. *La distance $d^{\triangle}_{asp}(f, g)$ est dans l'intervalle $[0; M[$ [231, p. 16]). La distance d'Asplund LIP-additive est donc une image (réduite à un seul pixel).*

Propriété 6 (Invariance par addition-LIP d'une constante)**.** *La métrique d'Asplund LIP-additive reste inchangée lorsque la fonction $f \in \mathcal{F}_M$ (ou g) est remplacée par toute fonction translatée $f \triangle k$, où $k \in]-\infty; M[$. En effet, les constantes c_1 et c_2 deviennent respectivement $c_1 \triangle k$ et $c_2 \triangle k$. Ceci implique que $d^{\triangle}_{asp}(f \triangle k, g) = d^{\triangle}_{asp}(f, g)$ et que $\forall f, \forall k$, $d^{\triangle}_{asp}(f, f \triangle k) = 0$. Sachant que l'addition d'une constante à une fonction est équivalente à la variation du temps d'exposition de la caméra [45], [65], [136], nous avons le résultat fondamental suivant : la métrique d'Asplund LIP-additive est insensible au changement de temps d'exposition.*

Remarque (*Annexe mathématique*). Comme pour la métrique d'Asplund LIP-multiplicative, une définition mathématique rigoureuse est obtenue en remplaçant chaque image f par sa classe d'équivalence $f^{\triangle}$, qui représente l'ensemble des fonctions h telles que $h = f \triangle k$, pour une constante k dans l'intervalle $]-\infty; M[$. Toutefois, si nous conservons les notations f et g, il n'y a pas d'ambiguïté. En effet, pour tout couple $(f^{\triangle}, g^{\triangle})$ de classes d'équivalence, nous avons :

$$d^{\triangle}_{asp}(f^{\triangle}, g^{\triangle}) = d^{\triangle}_{asp}(f_1, g_1), \tag{4.45}$$

où f_1 et g_1 sont respectivement des éléments des classes $f^{\triangle}$ et $g^{\triangle}$. Dans [140], [231], les propriétés de la métrique d'Asplund (i.e. symétrie, positivité, séparation, inégalité triangulaire) sont vérifiées dans l'espace des classes d'équivalences $\mathcal{I}^{\triangle}$.

4.3.3 Analyse de forme avec la métrique d'Asplund LIP-multiplicative

Jourlin et al. [146] ont introduit une carte des distances d'Asplund LIP-multiplicatives. Avec Jourlin, je l'ai ensuite étudiée et j'ai établi son lien avec la Morphologie Mathématique (MM) dans différents articles [209], [227], [228], [231].

Cartes des distances d'Asplund LIP-multiplicatives

Soient $]0; M[$ l'axe des niveaux de gris sans la valeur nulle et $\mathcal{I}^* =]0; M[^D$ l'ensemble des images avec des valeurs strictement positives.

Définition 14 (Carte des distances d'Asplund LIP-multiplicatives [146])**.** *Soit $f \in \mathcal{I}^*$ une image à niveaux de gris et $b \in]0; M[^{D_b}$ une sonde. La carte des distances d'Asplund $Asp^{\triangle}_b : \mathcal{I}^* \to (\mathbb{R}^+)^D$ est définie par :*

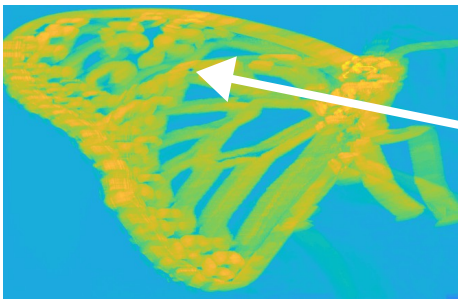
$$Asp^{\triangle}_b f(x) = d^{\triangle}_{asp}(f_{|D_b(x)}, b). \tag{4.46}$$

Pour chaque point $x \in D$, la distance $d^{\triangle}_{asp}(f_{|D_b(x)}, b)$ est calculée sur le voisinage $D_b(x)$ centré en x et le gabarit b agit tel un élément structurant. $f_{|D_b(x)}$ est la restriction de l'image f au domaine $D_b(x)$. $Asp^{\triangle}_b f : D \to \mathbb{R}^+$ est la carte des distances d'Asplund entre l'image f et la sonde b.

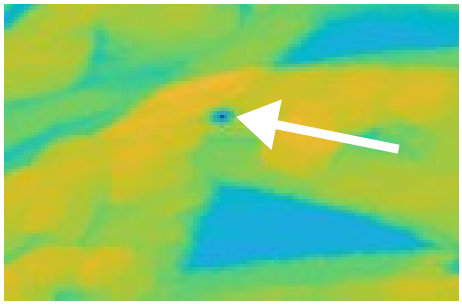
La figure 4.13 illustre la carte des distances d'Asplund avec une image d'un papillon extraite du jeu de données *Yahoo Flickr Creative Commons 100 Million* [299]. Le traitement est effectué sur son image de luminance f qui est en niveaux de gris même si l'image est affichée en couleur. Sur l'image de luminance f du papillon (Fig. 4.13a), une tache blanche est sélectionnée pour servir de sonde b (Fig. 4.13b). La carte des distances d'Asplund $Asp^{\triangle}_b f$ entre l'image f et la sonde b (Fig. 4.13c) présente un minimum qui correspond à la sonde que nous recherchons (Fig. 4.13d). Une carte des distances d'Asplund nous permet donc de chercher une forme de référence ou une sonde dans une image. Cependant, comme les images peuvent présenter du bruit d'acquisition, une version robuste au bruit de la métrique est utile pour la reconnaissance de forme.

(a) Image couleur et la sonde.

(b) Zoom sur la sonde.



(c) Carte des distances $Asp^{\triangle}_b f$ sur la luminance de (a).



(d) Zoom sur la carte des distances.

FIG. 4.13 – (a) Version couleur de l'image de papillon. Dans son image de luminance f, une tache blanche – entourée de vert – est sélectionnée comme fonction sonde b. Source [327]. (b) Zoom sur la version couleur de la sonde de luminance b. (c) Carte des distances d'Asplund LIP-multiplicatives $Asp^{\triangle}_b f$ entre l'image f et la sonde b. Son minimum est indiqué par la flèche blanche. (d) Zoom sur la carte des distances d'Asplund de l'image. La flèche blanche pointe sur son minimum.

Version robuste au bruit

La métrique d'Asplund est définie grâce à des extrema, ce qui la rend sensible au bruit. Afin de repousser cette limite, Jourlin [142] a proposé une extension qui enlève du domaine D les points les plus pénalisants. Cette idée est liée à la *topologie de la convergence en mesure* [36, chap. 4] dans le contexte des images à niveaux de gris.

Rappelons brièvement la théorie de la mesure dans le contexte des images à niveaux de gris. Etant données une mesure μ sur $\mathbb{R}^n$ et une métrique d entre des images à niveaux de gris, un voisinage $N_{\mu,d,\varepsilon,\varepsilon'}$ de l'image $f \in \mathcal{I}$ peut être défini grâce à μ et deux petits nombres réels positifs

et arbitraires ε et ε' selon :

$$N_{\mu,d,\varepsilon,\varepsilon'} = \{g, \mu(x \in D, d(f(x), g(x)) > \varepsilon) < \varepsilon'\}. \tag{4.47}$$

Ceci signifie que la mesure de l'ensemble des points x, où $d(f(x), g(x))$ excède la tolérance ε, satisfait une autre tolérance ε'. Cela peut être simplifié dans le contexte de la métrique d'Asplund. Comme l'image est numérisée, le nombre de pixels du domaine D est fini et la « mesure » d'un sous-ensemble de D est liée au cardinal de ce sous-ensemble, e.g. le pourcentage P de ses éléments par rapport à D (ou à une région d'intérêt $R \subset D$).

Nous recherchons donc un sous-ensemble D' de D, tel que :

i) $f_{|D'}$ et $g_{|D'}$ soient voisins (pour la métrique d'Asplund) ;

ii) l'ensemble complémentaire $D \setminus D'$ de D' par rapport à D soit de petite taille lorsqu'il est comparé à D.

Cette dernière condition s'écrit : $P(D \setminus D') = \frac{\#(D \setminus D')}{\#D} \leq p$, où p représente un pourcentage acceptable et $\#D$ est le nombre d'éléments de D. Un voisinage de l'image $f \in \mathcal{I}$ peut être défini grâce à un petit nombre réel ε comme :

$$N_{P,d^{\triangle}_{asp},\varepsilon,p}(f) = \left\{g \mid \exists D' \subset D,\ d^{\triangle}_{asp}(f_{|D'}, g_{|D'}) < \varepsilon \text{ et } P(D \setminus D') \leq p\right\}. \tag{4.48}$$

L'ensemble D' correspond aux pixels de bruit qu'il convient d'écarter. Comme ces pixels-ci sont les plus proches de la sonde, ils sont sélectionnés grâce à une fonction $\gamma^{\triangle}_{(f,g)} : D \to \mathbb{R}$ caractérisant le contraste LIP-multiplicatif entre les fonctions f et g qui est définie par :

$$\forall x \in D,\ \gamma^{\triangle}_{(f,g)}(x) \triangle g(x) = f(x) \iff \gamma^{\triangle}_{(f,g)} = \frac{\ln(1 - f/M)}{\ln(1 - g/M)}; \tag{4.49}$$

où $\gamma^{\triangle}_{(f,g)}(x)$ est le réel par lequel chaque valeur de l'image $f(x)$ est LIP-multipliée pour être égale à la valeur de la sonde $g(x)$. Les expressions des sondes $\lambda \triangle g$ et $\mu \triangle g$ (Déf. 12) peuvent être écrites avec la fonction de contraste $\gamma^{\triangle}_{(f,g)}$ comme il s'ensuit : $\lambda = \inf\{\alpha, \forall x, \gamma^{\triangle}_{(f,g)}(x) \leq \alpha\}$ et $\mu = \sup\{\alpha, \forall x, \alpha \leq \gamma^{\triangle}_{(f,g)}(x)\}$. Ceci explique que les valeurs des images les plus proches de la sonde supérieure $\lambda \triangle g$, ou de la sonde inférieure $\mu \triangle g$, correspondent aux plus grandes, ou plus petites valeurs de $\gamma^{\triangle}_{(f,g)}$, respectivement. En utilisant cette propriété, de nouvelles sondes $\lambda' \triangle g$, ou $\mu' \triangle g$, peuvent être définies sur le domaine $D \setminus D'$ obtenus en écartant un pourcentage $(1-p)/2$ des pixels avec les plus grandes, ou les plus petites, valeurs de contraste $\gamma^{\triangle}_{(f,g)}(x)$, respectivement. Le domaine restreint $D \setminus D'$ a ainsi un cardinal égal à un pourcentage p du cardinal de D.

Définition 15 (Métrique d'Asplund LIP-multiplicative avec tolérance)**.** *Soient $(1-p)$ un pourcentage de points de D à écarter et D' l'ensemble de ces points écartés. La métrique d'Asplund (LIP-multiplicative) avec tolérance entre deux images à niveaux de gris f et $g \in \mathcal{I}$ est définie par :*

$$d^{\triangle}_{asp,p}(f, g) = \ln\left(\frac{\lambda'}{\mu'}\right). \tag{4.50}$$

Les facteurs λ' et μ' sont égaux à $\lambda' = \inf\{\alpha, \forall x \in D,\ \gamma^{\triangle}_{(f_{|D \setminus D'}, g_{|D \setminus D'})}(x) \leq \alpha\}$ et $\mu' = \sup\{\alpha, \forall x \in D, \alpha \leq \gamma^{\triangle}_{(f_{|D \setminus D'}, g_{|D \setminus D'})}(x)\}$. Un pourcentage $(1-p)/2$ des points $x \in D$ avec les plus grandes, respectivement les plus petites, valeurs de contraste $\gamma^{\triangle}_{(f,g)}(x) = \frac{\ln(1-f(x)/M)}{\ln(1-g(x)/M)}$ est écarté.

En utilisant les équations (4.49) et (4.4), la fonction de contraste $\gamma^{\triangle}_{(f,g)}$ est proportionnelle à la fonction de contraste $\gamma^{\triangle}_{(f,\alpha \triangle g)}$ entre l'image f et la sonde LIP-multipliée $\alpha \triangle g$: $\gamma^{\triangle}_{(f,\alpha \triangle g)}(x) = (1/\alpha)\gamma^{\triangle}_{(f,g)}(x)$, avec $\alpha > 0$. Il s'ensuit la propriété suivante qui est démontrée dans [231, supp. mat.].

Propriété 7 (Invariance par multiplication-LIP)**.** *La métrique* $d^{\triangle}_{asp,p}$ *est invariante par multiplication-LIP par un scalaire :* $\forall \alpha \in \mathbb{R}$, $\forall (f,g) \in \mathcal{I}^2$, $d^{\triangle}_{asp,p}(f, \alpha \triangle g) = d^{\triangle}_{asp,p}(f,g)$.

Remarque. Une carte des distances d'Asplund avec tolérance peut être définie comme dans la définition 14 : $Asp^{\triangle}_{b,p} f(x) = d^{\triangle}_{asp,p}(f_{|D_b(x)}, b)$.

La figure 4.14a illustre la capacité de la métrique d'Asplund avec une tolérance à écarter les valeurs extrémales $f(x_\lambda)$ et $f(x_\mu)$ associée au bruit. Sur la figure 4.14b, un plan est choisi comme la sonde g et une image f est obtenue en additionnant un bruit blanc à la sonde avec une densité spatiale de $\rho = 8\%$. Le plan de la sonde g est déformé par la multiplication-LIP utilisée pour le calcul de la sonde supérieure $\lambda \triangle g$ et pour celui de la sonde inférieure $\mu \triangle g$. La distance d'Asplund qui est initialement égale à $d^{\triangle}_{asp}(f,g) = 3{,}75$ décroît à $d^{\triangle}_{asp,p}(f,g) = 0{,}79$, avec une tolérance de $p = 97$ %. Les expériences ont montré qu'avec $p = 90$ % ($\leq (1-\rho)$), la distance d'Asplund est égale à zéro.

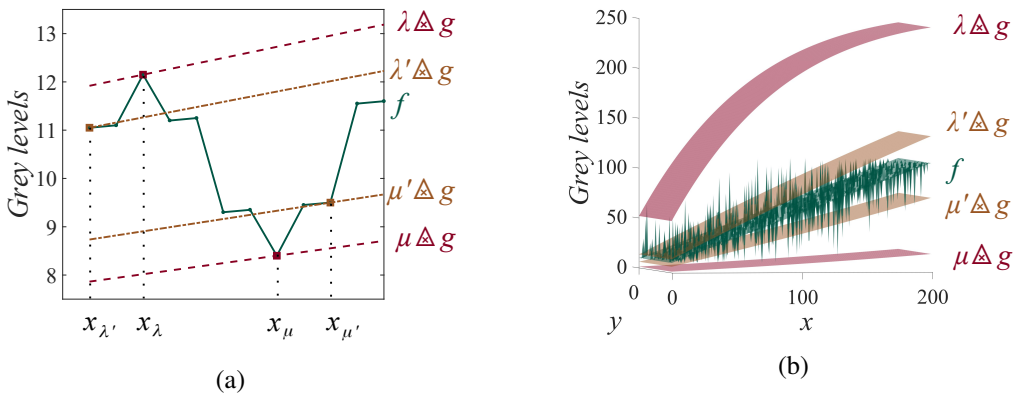


FIG. 4.14 – Illustration de la métrique d'Asplund LIP-multiplicative où l'image g est utilisée pour palper l'image f. $\lambda \triangle g$ (respectivement $\lambda' \triangle g$) est la sonde supérieure et $\mu \triangle g$ (resp. $\mu' \triangle g$) est la sonde inférieure de la métrique d'Asplund $d^{\triangle}_{asp}(f,g)$ (resp. de la métrique d'Asplund avec une tolérance p, $d^{\triangle}_{asp,p}(f,g)$). (a) Sans paramètre de tolérance, la métrique d'Asplund est égale à $d^{\triangle}_{asp}(f,g) = 0{,}42$ tandis qu'avec une tolérance de $p = 80\%$ elle décroît à la valeur $d^{\triangle}_{asp,p}(f,g) = 0{,}24$. (b) L'image f est obtenue en ajoutant à la sonde g un bruit blanc gaussien de moyenne 0, de variance 5 et de densité spatiale $\rho = 8$ %.

Lien avec la Morphologie Mathématique

Les cartes des distances seront d'abord présentées avec des opérations de voisinage. Le lien avec la MM sera ensuite établi dans le cas spécifique des éléments structurants plats et dans le cas général d'une fonction structurante. J'ai initialement présenté ces résultats dans [231].

Expression générale des cartes des distances d'Asplund avec des opérations de voisinage A partir de l'équation (4.40), pour chaque $x \in D$, nous avons $d^{\triangle}_{asp}(f_{|D_b(x)}, b) = \ln(\lambda_b f(x)/\mu_b f(x))$, où $\lambda_b f(x) = \inf\left\{\alpha, f_{|D_b(x)} \leq \alpha \triangle b\right\}$ et $\mu_b f(x) = \sup\left\{\alpha, \alpha \triangle b \leq f_{|D_b(x)}\right\}$. Ces expressions montrent que la carte des distances d'Asplund consiste en une palpation double-face, à chaque point x, par le plus petit majorant $\lambda_b f(x) \triangle b$ et par le plus grand minorant $\mu_b f(x) \triangle b$. Comme $\lambda_b f(x)$ et $\mu_b f(x)$ existent pour tout $x \in D$, les cartes suivantes peuvent alors être définies [228].

Définition 16 (Les cartes LIP-multiplicatives des plus petits majorants et des plus grands minorants [228])**.** *Etant donné l'espace fermé des réels positifs* $\overline{\mathbb{R}}^+ = [0, +\infty]$, *soient* $f \in \overline{\mathcal{I}}$ *une image*

et $b \in (]0; M[)^{D_b}$ *une sonde. Leur carte des plus petits majorants (cmaj)* $\lambda_b : \overline{\mathcal{I}} \to (\overline{\mathbb{R}}^+)^D$ *et leur carte des plus grands minorants (cmin)* $\mu_b : \overline{\mathcal{I}} \to (\overline{\mathbb{R}}^+)^D$ *sont définies par :*

$$\lambda_b f(x) = \inf_{h \in D_b} \{\alpha, f(x+h) \leq \alpha \triangle b(h)\}, \tag{4.51}$$

$$\mu_b f(x) = \sup_{h \in D_b} \{\alpha, \alpha \triangle b(h) \leq f(x+h)\}. \tag{4.52}$$

Définissons $\tilde{f} = \ln(1 - f/M)$, avec $f \in \overline{\mathcal{I}}$, et introduisons les expressions générales des cartes des majorants et des minorants. Les propositions 2, 4 et 5 ont été démontrées dans [228], [231].

Proposition 2. *Les cartes des plus petits majorants* λ_b *et des plus grands minorants* μ_b *sont égales à :*

$$\lambda_b f(x) = \bigvee \{\tilde{f}(x+h)/\tilde{b}(h), h \in D_b\}, \tag{4.53}$$

$$\mu_b f(x) = \bigwedge \{\tilde{f}(x+h)/\tilde{b}(h), h \in D_b\}. \tag{4.54}$$

Corollaire 1. *Etant donnée* $f \in \overline{\mathcal{I}}$ *une image à valeurs strictement positives* $f > 0$*, l'expression de la carte des distances d'Asplund LIP-multiplicatives devient :*

$$Asp_b^{\triangle} f = \ln\left(\frac{\lambda_b f}{\mu_b f}\right). \tag{4.55}$$

Cas particulier d'un élément structurant plat Dans le cas d'un élément structurant plat, les expressions des différentes cartes peuvent être simplifiées comme il s'ensuit.

Proposition 3. *Soit* $b = b_0 \in]0; M[^{D_b}$ *un élément structurant plat (*$\forall x \in D_b$*,* $b(x) = b_0$*). Les cartes des plus petits majorants* λ_{b_0}*, des plus grands minorants* μ_{b_0} *et des distances d'Asplund* $Asp_{b_0}^{\triangle}$ *sont égales à :*

$$\lambda_{b_0} f = (1/\tilde{b}_0) \ln[1 - (\delta_{\overline{D}_b} f)/M], \tag{4.56}$$

$$\mu_{b_0} f = (1/\tilde{b}_0) \ln[1 - (\varepsilon_{D_b} f)/M], \tag{4.57}$$

$$Asp_{b_0}^{\triangle} f = \ln\left[\frac{\ln[1 - (\delta_{\overline{D}_b} f)/M]}{\ln[1 - (\varepsilon_{D_b} f)/M]}\right], \text{ où } f > 0 \tag{4.58}$$

et $\overline{D}_b = \{-h, h \in D_b\}$ *est le domaine réfléchi (ou transposé) de l'élément structurant.*

Ce résultat montre que la carte des distances d'Asplund est une combinaison de logarithmes, d'une érosion ε_b et d'une dilatation δ_b de l'image f par l'élément structurant plat b. Comme de nombreux logiciels de traitement d'image incluent des opérations de dilatation et d'érosion, la programmation de cette carte de distances s'en trouve facilitée.

Cas général : une fonction structurante

Proposition 4. *Les cartes des plus petits majorants* λ_b *et des plus grands minorants* μ_b *sont une dilatation et une érosion, respectivement, entre les deux treillis complets* $\mathscr{L}_1 = (\overline{\mathcal{I}}, \leq)$ *et* $\mathscr{L}_2 = ((\overline{\mathbb{R}}^+)^D, \leq)$.

Déterminons les expressions de cette dilatation et de cette érosion en utilisant la MM avec des fonctions structurantes dans le cadre de la morphologie mathématique multiplicative qui a été introduite par Heijmans [112], [113].

Définition 17 (Erosion et dilatation avec une fonction structurante multiplicative [112], [113])**.** *Soient une fonction $f \in (\overline{\mathbb{R}}^+)^D$ et $b \in (\overline{\mathbb{R}}^+)^{D_b}$ une fonction structurante multiplicative. L'opérateur $f \dot{\oplus} b$ défini par*

$$\bigvee_{h \in D_b} \{f(x-h).b(h)\} = (f \dot{\oplus} b)(x) \tag{4.59}$$

est une dilatation et l'opérateur $f \dot{\ominus} b$ défini par

$$\bigwedge_{h \in D_b} \{f(x+h)/b(h)\} = (f \dot{\ominus} b)(x) \tag{4.60}$$

est une érosion, avec la convention que $f(x-h).b(h) = 0$ lorsque $f(x-h) = 0$ ou $b(h) = 0$ et que $f(x+h)/b(h) = +\infty$ lorsque $f(x+h) = +\infty$ ou $b(h) = 0$. Les symboles $\dot{\oplus}$ et $\dot{\ominus}$ représentent l'extension aux fonctions structurantes multiplicatives des opérations de Minkowski entre ensembles [275].

Il existe une relation entre l'érosion ou la dilatation multiplicative et leurs pendants additifs de la section 3.1.2, p. 23 [112], [113] :

$$f \dot{\oplus} b = \exp(\ln f \oplus \ln b), \tag{4.61}$$

$$f \dot{\ominus} b = \exp(\ln f \ominus \ln b). \tag{4.62}$$

La proposition suivante donne les expressions morphologiques des différentes cartes.

Proposition 5. *Soient une fonction structurante $b \in [0; M[^{D_b}$ et une image $f \in \overline{\mathcal{I}}$. Les expressions de la carte des plus petits majorants λ_b, qui est une dilatation, et de la carte des plus grands minorants μ_b, qui est une érosion, sont égales à :*

$$\lambda_b f = (-\tilde{f}) \dot{\oplus} (-1/\tilde{\overline{b}}) = \exp(\hat{f} \oplus (-\hat{\overline{b}})), \tag{4.63}$$

$$\mu_b f = (-\tilde{f}) \dot{\ominus} (-\tilde{b}) \quad = \exp(\hat{f} \ominus \hat{b}). \tag{4.64}$$

La fonction structurante réfléchie $\overline{b}$ est définie par $\forall x \in \overline{D}_b$, $\overline{b}(x) = b(-x)$ [284] et $\tilde{f}$ est la fonction $\tilde{f} = \ln(1 - f/M)$, avec $\tilde{f} \in [-\infty, 0]$. $\hat{f} = \ln(-\tilde{f}) = \ln(-\ln(1 - f/M))$ est une transformée de f.

La carte des distances d'Asplund est la différence entre une dilatation et une érosion dont l'expression est :

$$Asp_b^{\triangle} f = \left[\hat{f} \oplus (-\hat{\overline{b}})\right] - \left[\hat{f} \ominus \hat{b}\right] = \delta_{-\hat{\overline{b}}} \hat{f} - \varepsilon_{\hat{b}} \hat{f}. \tag{4.65}$$

Nous remarquons que la carte des distances d'Asplund LIP-multiplicatives est similaire à la norme du gradient morphologique ϱ_B appelé aussi gradient de Beucher [277], [284], mais qui est quant à lui défini pour un élément structurant plat et symétrique B par l'expression : $\varrho_B \hat{f} = \delta_B \hat{f} - \varepsilon_B \hat{f}$. La carte des distances d'Asplund LIP-multiplicatives est en fait l'extension du gradient morphologique aux fonctions dans le cas d'opérations LIP-multiplicatives. Cette similarité montre que la carte des distances d'Asplund agit comme un opérateur de dérivation.

Remarque. La carte des distances d'Asplund avec tolérance $Asp_{b,p}^{\triangle}$ peut être calculée en remplaçant la dilatation et l'érosion par des filtres de rang [277].

La figure 4.15a illustre la palpation sur les deux faces d'une image f par une sonde b en utilisant la loi LIP-multiplicative $\triangle$ qui modifie l'amplitude des sondes supérieures $\lambda_b f(x) \triangle b$ et inférieures $\mu_b f(x) \triangle b$. Les deux pics ont des amplitudes différentes causées par une dérive d'éclairement créée par la loi LIP-multiplicative. Dans la figure 4.15b, lorsque la sonde b est similaire à une forme dans f (selon la distance d'Asplund), la carte des distances d'Asplund de f, $Asp_b^{\triangle} f$, présente un minimum local. Ici, les lieux des deux pics sont situés aux minima les plus profonds de la carte de f par la sonde b, $Asp_b^{\triangle} f$. Ce résultat montre que la carte des distances d'Asplund LIP-multiplicatives est insensible à une variation d'éclairement correspondant à une variation de l'absorption (ou de l'opacité) de l'objet.

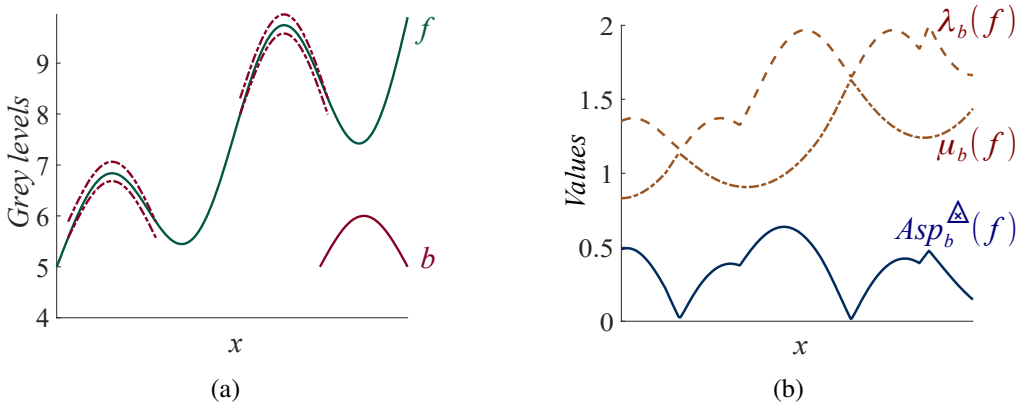


FIG. 4.15 – (a) L'image f est palpée sur ses deux faces par la sonde b (de couleur rouge) grâce à la loi LIP-multiplicative $\triangle$. La double palpation est illustrée à deux endroits par les courbes en tirets rouges. (b) $\lambda_b f$ et $\mu_b f$ sont respectivement les cartes des plus petits majorants et plus grands minorants de f. $Asp_b^{\triangle} f$ est la carte des distances d'Asplund entre f et b.

4.3.4 Analyse de forme avec la métrique LIP-additive de type Asplund

La métrique LIP-additive d'Asplund est utile pour les images acquises avec une faible intensité de la source ou un faible temps d'exposition de la caméra [140]. Dans cette sous-section, à cause des similarités importantes avec le cas LIP-multiplicatif (cf. sous-section 4.3.3), nous montrerons uniquement les nouveaux résultats et équations. Nous considérerons l'ensemble des fonctions $\mathcal{F}_M =]-\infty; M[^D$ (ou $\overline{\mathcal{F}}_M = [-\infty; M]^D$) avec des valeurs dans $]-\infty; M[$ (ou dans $[-\infty; M]$).

Cartes des distances d'Asplund LIP-additives

Définition 18 (Carte des distances d'Asplund LIP-additives)**.** *Soient $f \in \mathcal{F}_M$ une fonction et $b \in]-\infty; M[^{D_b}$ une sonde. La carte des distances d'Asplund est l'application $Asp_b^{\triangle} : \mathcal{F}_M \to \mathcal{I}$ définie par :*

$$Asp_b^{\triangle} f(x) = d_{asp}^{\triangle}(f_{|D_b(x)}, b). \tag{4.66}$$

L'addition LIP $\triangle$ rend la carte des distances robuste aux variations de contraste dues aux changements de temps d'exposition.

Version robuste au bruit

Afin de surmonter la sensibilité au bruit de la métrique d'Asplund, un voisinage $N_{P,d_{asp}^{\triangle},\varepsilon,p}(f)$ de la fonction f est défini en remplaçant dans l'équation (4.48) la métrique d'Asplund LIP-multiplicative $d_{asp}^{\triangle}$ par la LIP-additive $d_{asp}^{\triangle}$. Les pixels de bruit qui sont les plus proches de la sonde sont sélectionnés en utilisant une fonction $\gamma_{(f,g)}^{\triangle} : D \to]-\infty; M[$ qui caractérise le contraste (LIP-additif) entre les fonctions f et g :

$$\forall x \in D,\ \gamma_{(f,g)}^{\triangle}(x) \triangle g(x) = f(x) \iff \gamma_{(f,g)}^{\triangle} = f \triangle g, \tag{4.67}$$

où $\gamma_{(f,g)}^{\triangle}(x) \in]-\infty; M[$ est la valeur LIP-ajoutée à chaque valeur de la fonction $f(x)$ afin qu'elle devienne égale à la valeur de la sonde $g(x)$. Les expressions des sondes $c_1 \triangle g$ et $c_2 \triangle g$ (Déf. 13) sont égales à : $c_1 = \inf\{c, \forall x, \gamma_{(f,g)}^{\triangle}(x) \leq c\}$ et $c_2 = \sup\{c, \forall x, c \leq \gamma_{(f,g)}^{\triangle}(x)\}$. Ces sondes sont utilisées pour introduire les définitions suivantes.

Définition 19 (Métrique d'Asplund LIP-additive avec tolérance)**.** *Soit* $(1-p)$ *un pourcentage de points de* D *à écarter et* D' *l'ensemble de ces points écartés. La métrique d'Asplund LIP-additive avec tolérance entre deux fonctions* f *et* $g \in \mathcal{F}_M$ *est définie par :*

$$d^{\triangle}_{asp,p}(f,g) = c'_1 \triangle c'_2. \tag{4.68}$$

Les constantes c'_1 *et* c'_2 *sont égales à* $c'_1 = \inf\{c, \forall x \in D, \gamma^{\triangle}_{(f_{|D\setminus D'}, g_{|D\setminus D'})}(x) \leq c\}$ *et* $c'_2 = \sup\{c, \forall x \in D, c \leq \gamma^{\triangle}_{(f_{|D\setminus D'}, g_{|D\setminus D'})}(x)\}$*. Un pourcentage* $(1-p)/2$ *des points* $x \in D$ *avec les plus grandes, respectivement les plus petites, valeurs* $\gamma^{\triangle}_{(f,g)}(x) = f(x) \triangle g(x)$ *est écarté.*

A partir des équations (4.67) et (4.6), la fonction de contraste $\gamma^{\triangle}_{(f,g)}$ est reliée à la fonction de contraste $\gamma^{\triangle}_{(f,c\triangle g)}$ entre f et la sonde $c \triangle g$, où $\forall x \in D, c(x) = c, c \in]-\infty; M[$, par l'équation suivante : $\gamma^{\triangle}_{(f,c\triangle g)}(x) = \gamma^{\triangle}_{(f,g)}(x) \triangle c$. Ceci conduit à la propriété suivante démontrée dans [231, supp. mat.].

Propriété 8 (Invariance par addition-LIP d'une constante)**.** *La métrique* $d^{\triangle}_{asp,p}$ *est invariante par addition-LIP (ou soustraction-LIP) d'une constante :* $\forall k \in]-\infty; M[$, $\forall (f,g) \in (\mathcal{F}_M)^2$, $d^{\triangle}_{asp,p}(f, g \triangle k) = d^{\triangle}_{asp,p}(f,g)$ *et* $d^{\triangle}_{asp,p}(f, f \triangle k) = 0$.

Remarque. Une carte des distances d'Asplund avec tolérance peut être introduite comme dans la définition 18 : $Asp^{\triangle}_{b,p} f(x) = d^{\triangle}_{asp,p}(f_{|D_b(x)}, b)$.

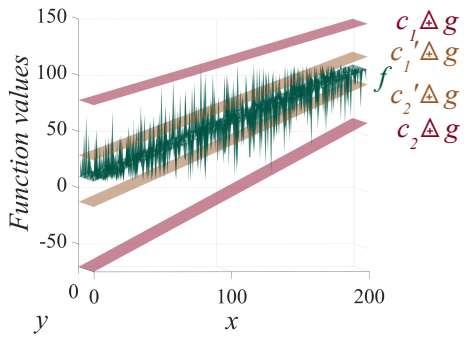


FIG. 4.16 – Illustration de la métrique LIP-additive pour les fonctions où g est utilisée pour palper f. $c_1 \triangle g$ (respectivement $c'_1 \triangle g$) est la sonde supérieure et $c_2 \triangle g$ (resp. $c'_2 \triangle g$) est la sonde inférieure de la métrique d'Asplund $d^{\triangle}_{asp}(f,g)$ (resp. de la métrique d'Asplund avec une tolérance p, $d^{\triangle}_{asp,p}(f,g)$). La fonction f est obtenue en ajoutant un bruit blanc gaussien à la sonde plane g (moyenne 0, variance 5 et densité spatiale $\rho = 8\,\%$).

La figure. 4.16 illustre que la métrique d'Asplund LIP-additive $d^{\triangle}_{asp,p}(f,g)$ est robuste au bruit. Un plan est choisi comme sonde g et la fonction f est obtenue en ajoutant un bruit à la sonde g avec une densité spatiale de $\rho = 8\%$. Nous remarquons que la sonde supérieure $c_1 \triangle g$ et la sonde inférieure $c_2 \triangle g$ peuvent prendre des valeurs négatives. Après l'addition-LIP, la surface plane de la sonde g est toujours un plan mais avec une pente différente. La distance d'Asplund qui est initialement égale à $d^{\triangle}_{asp}(f,g) = 116$ décroît à $d^{\triangle}_{asp,p}(f,g) = 40$, avec une tolérance de $p = 97\,\%$. Les expériences ont montré qu'avec $p = 90\,\%$ ($\leq (1-\rho)$), la distance d'Asplund est égale à zéro.

Lien avec la Morphologie Mathématique

Expression générale des cartes des distances d'Asplund avec des opérations de voisinage Afin d'établir le lien avec la MM, nous exprimerons les cartes des distances d'Asplund avec des opérations sur les voisinages. A partir de l'équation (4.44), pour chaque

$x \in D$, l'expression de la carte devient $d^{\triangle}_{asp}(f_{|D_b(x)}, b) = c_{1_b} f(x) \triangle c_{2_b} f(x)$, où $c_{1_b} f(x) = \inf \left\{ c, f_{|D_b(x)} \leq c \triangle b \right\}$ et $c_{2_b} f(x) = \sup \left\{ c, c \triangle b \leq f_{|D_b(x)} \right\}$. Cela conduit à la définition suivante.

Définition 20 (Cartes LIP-additives des plus petits majorants et des plus grands minorants)**.** *Soient $f \in \overline{\mathcal{F}}_M$ une fonction et $b \in]-\infty; M[^{D_b}$ une sonde. Leur carte des plus petits majorants (cmaj) $c_{1_b} : \overline{\mathcal{F}}_M \to \overline{\mathcal{F}}_M$ et leur carte des plus petits minorants (cmin) $c_{2_b} : \overline{\mathcal{F}}_M \to \overline{\mathcal{F}}_M$ sont définies par :*

$$c_{1_b} f(x) = \inf_{h \in D_b} \{c, f(x+h) \leq c \triangle b(h)\}, \tag{4.69}$$

$$c_{2_b} f(x) = \sup_{h \in D_b} \{c, c \triangle b(h) \leq f(x+h)\}. \tag{4.70}$$

Les propositions 6 à 9 ont été démontrées dans [231]. Les expressions générales des cartes des majorants, des minorants et des distances d'Asplund seront données ci-après.

Proposition 6. *Les cartes des plus petits majorants c_{1_b} et des plus grands minorants c_{2_b} sont égales à :*

$$c_{1_b} f(x) = \bigvee \{f(x+h) \triangle b(h), h \in D_b\}, \tag{4.71}$$

$$c_{2_b} f(x) = \bigwedge \{f(x+h) \triangle b(h), h \in D_b\}. \tag{4.72}$$

Corollaire 2. *La carte des distances d'Asplund entre la fonction f et la sonde b est égale à :*

$$Asp^{\triangle}_b f = c_{1_b} f \triangle c_{2_b} f. \tag{4.73}$$

Cas particulier d'un élément structurant plat

Proposition 7. *Soit $b = b_0 \in]-\infty; M[^{D_b}$ un élément structurant plat ($\forall x \in D_b$, $b(x) = b_0$). Les cartes des plus petits majorants $c_{1_{b_0}}$, des plus petits minorants $c_{2_{b_0}}$ et la carte des distances d'Asplund $Asp^{\triangle}_{b_0}$ sont égales à :*

$$c_{1_{b_0}} f = \delta_{\overline{D}_b} f \triangle b_0, \tag{4.74}$$

$$c_{2_{b_0}} f = \varepsilon_{D_b} f \triangle b_0, \tag{4.75}$$

$$Asp^{\triangle}_{b_0} f = \delta_{\overline{D}_b} f \triangle \varepsilon_{D_b} f. \tag{4.76}$$

Avec une sonde plate b, la carte des distances d'Asplund est une différence-LIP entre une dilatation $\delta_{\overline{D}_b}$ et une érosion ε_{D_b} de la fonction f par le domaine D_b de l'élément structurant b. Elle est alors similaire à un gradient morphologique avec une différence LIP. Cette similarité montre que la carte des distances d'Asplund agit comme un opérateur de dérivation.

Cas général : une fonction structurante

Proposition 8. *La carte des plus petits majorants c_{1_b} (resp. des plus grands minorants c_{2_b}) est une dilatation (resp. une érosion) dans le même treillis $\mathscr{L}_1 = \mathscr{L}_2 = (\overline{\mathcal{F}}_M, \leq)$.*

Remarque. Dans l'équation (4.73), on remarque que la carte des distances d'Asplund LIP-additives est une différence entre une dilatation et une érosion, qui correspond à une extension de la version LIP du gradient morphologique aux fonctions structurantes.

Etablissons la relation de la dilatation c_{1_b} ou de l'érosion c_{2_b} avec la dilatation usuelle δ_b (éq. (3.6)) ou l'érosion usuelle ε_b (éq. (3.7)), respectivement. Pour ce faire, un isomorphisme est nécessaire entre le treillis $\overline{\mathbb{R}}^D$ de $\delta_b f$, ou $\varepsilon_b f$, et le treillis $\overline{\mathcal{F}}_M$ de $c_{1_b} f$, ou $c_{2_b} f$. Jourlin et al. [148], [203] ont défini cet isomorphisme $\xi : \overline{\mathcal{F}}_M \to \overline{\mathbb{R}}^D$ et son inverse $\xi^{-1} : \overline{\mathbb{R}}^D \to \overline{\mathcal{F}}_M$ par :

$$\xi(f) = -M \ln(1 - f/M), \tag{4.77}$$

$$\xi^{-1}(f) = M(1 - \exp(-f/M)). \tag{4.78}$$

Remarque. Comme ξ^{-1} et ξ sont des bijections croissantes, elles distribuent sur les infima, ainsi que sur les suprema.

Proposition 9. *Etant données deux fonctions $f, g \in \overline{\mathcal{F}}_M$, l'isomorphisme ξ a la propriété de transformer la différence-LIP, $\triangle$, en la différence canonique, $-$: $\xi(f \triangle g) = \xi(f) - \xi(g)$.*

La dilatation c_{1_b} (éq. (4.71)) et l'érosion c_{2_b} (éq. (4.72)) peuvent donc être exprimées comme :

$$\begin{aligned} c_{1_b} f(x) &= \xi^{-1} \circ \xi\Big(\bigvee_{h \in \overline{D}_b} \Big\{ f(x-h) \triangle \overline{b}(h) \Big\}\Big) \\ &= \xi^{-1}\Big[\bigvee_{h \in \overline{D}_b} \{\xi(f)(x-h) - \xi(\overline{b})(h)\}\Big] \\ &= M(1 - e^{-\bigvee_{h \in \overline{D}_b} \{-\tilde{f}(x-h) + \tilde{\overline{b}}(h)\}}), \end{aligned} \tag{4.79}$$

$$\begin{aligned} c_{2_b} f(x) &= \xi^{-1}\Big[\bigwedge_{h \in D_b} \{\xi(f)(x+h) - \xi(b)(h)\}\Big] \\ &= M(1 - e^{-\bigwedge_{h \in D_b} \{-\tilde{f}(x+h)) - [-\tilde{b}(h)]\}}), \end{aligned} \tag{4.80}$$

où $\xi(f) = -M\tilde{f}$.

Proposition 10. *Soient $b \in]-\infty; M[^{D_b}$ une fonction structurante et $f \in \overline{\mathcal{F}}_M$ une fonction. Les expressions de leur carte des plus petits majorants c_{1_b}, qui est une dilatation, et de leur carte des plus grands minorants c_{2_b}, qui est une érosion, sont égales à :*

$$\begin{aligned} c_{1_b} f &= \xi^{-1}[\xi(f) \oplus (-\xi(\overline{b}))] \\ &= M(1 - e^{-[\acute{f} \oplus (-\acute{\overline{b}})]}), \end{aligned} \tag{4.81}$$

$$\begin{aligned} c_{2_b} f &= \xi^{-1}[\xi(f) \ominus \xi(b)] \\ &= M(1 - e^{-[\acute{f} \ominus \acute{b}]}). \end{aligned} \tag{4.82}$$

La fonction $\acute{f} \in \overline{\mathbb{R}}^D$ est définie par $\acute{f} = -\tilde{f} = -\ln(1 - f/M) = [\xi(f)]/M$.

Leur carte des distances d'Asplund est liée à la différence entre une dilatation et une érosion (avec un fonction structurante additive) par :

$$Asp_b^{\triangle} f = \xi^{-1}\Big[[\xi(f) \oplus (-\xi(\overline{b}))] - [\xi(f) \ominus \xi(b)]\Big] \tag{4.83}$$

$$\begin{aligned} &= M(1 - \exp(-[(\acute{f} \oplus (-\acute{\overline{b}})) - (\acute{f} \ominus \acute{b})])) \\ &= M(1 - \exp(-[\delta_{-\acute{\overline{b}}} \acute{f} - \varepsilon_{\acute{b}} \acute{f}])). \end{aligned} \tag{4.84}$$

Remarque. En remplaçant la dilatation et l'érosion par des filtres de rang [277], on peut calculer la carte des distances d'Asplund avec une tolérance $Asp_{b,p}^{\triangle}$.

La figure 4.17a illustre la palpation double face d'une image f par une sonde (ou une fonction structurante) b. Les amplitudes des deux pics sont liées par l'addition-LIP d'une constante. Dans la figure 4.17b, les deux minima les plus profonds de la carte des distances de f, $As_b^{\triangle} f$, correspondent aux lieux des deux pics similaires à la sonde b. Ceci illustre l'insensibilité de la carte des distances d'Asplund LIP-additives à une variation de l'intensité lumineuse ou du temps d'exposition de la caméra.

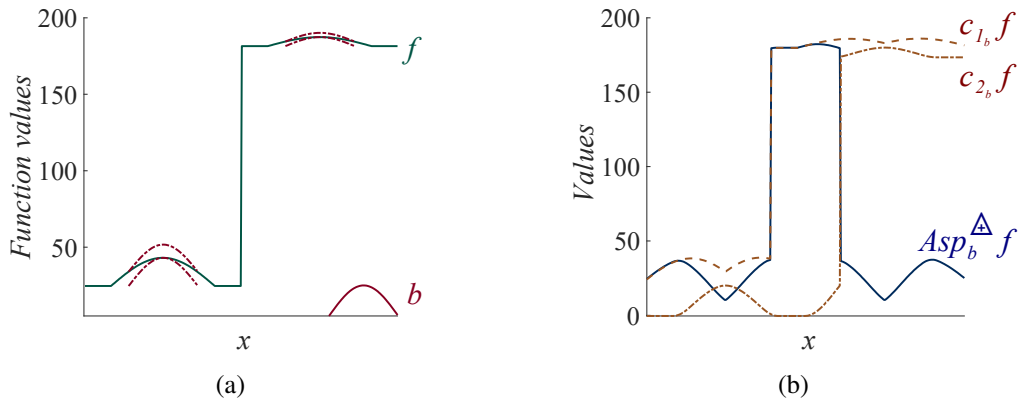


FIG. 4.17 – (a) La fonction f est palpée sur ses deux faces (haute et basse) par la sonde b (en rouge) grâce à la loi LIP-additive $\triangle$. La double palpation est illustrée à deux endroits par les courbes en tirets rouges. (b) $c_{1_b} f$ et $c_{2_b} f$ sont les cartes des plus petits majorants et des plus grands minorants de f, respectivement. $Asp_b^{\triangle} f$ est la carte des distances d'Asplund entre f et b.

4.3.5 Expériences et résultats

Mise en œuvre informatique

Les cartes des distances d'Asplund LIP-multiplicatives et LIP-additives de l'image f, $Asp_b^{\triangle} f$ et $Asp_b^{\triangle} f$, respectivement, ont été toutes deux programmées en langage *MATLAB*® en utilisant leurs expressions directes (éq. (4.46), (4.66)) et morphologiques (éq. (4.65), (4.84)). La même programmation a été faite pour les cartes des distances d'Asplund avec tolérance $Asp_{b,p}^{\triangle} f$ et $Asp_{b,p}^{\triangle} f$ de f. Du fait de l'existence dans de nombreux logiciels d'analyse d'image (e.g. *MATLAB*®, *Python scikit-images*) de versions rapides des opérations morphologiques de dilatation $\oplus$, d'érosion $\ominus$ et de filtres de rangs, la mise en œuvre morphologique est plus facile et plus rapide que celle qui est directe. Dans le tableau 4.1, les durées des mises en œuvre morphologiques et directes (avec parallélisation multicœurs) des cartes des distances d'Asplund sans et avec tolérance sont comparées en utilisant l'exemple de la figure 4.18. La taille de l'image est de 1 224 $\times$ 918 pixels et la sonde contient 285 pixels (Fig. 4.18d). *La mise en œuvre morphologique est toujours plus rapide que l'approche directe*, avec un facteur de gain compris entre 5,9 et 11 (processeur Intel® Core™ i7 CPU 4702HQ ; 2,20 GHz ; 4 cœurs ; 8 threads avec 16 Go de RAM).

Cas simulés

Nous évaluons les cartes des distances LIP-multiplicatives et LIP-additives pour détecter des boules dans une image claire et dans deux versions assombries obtenues par simulation (Fig. 4.18). L'image claire est acquise en couleur avec des contrastes « normaux » – choisis automatiquement par la caméra (Fig. 4.18a) – et convertie en une image de luminance à niveaux de gris, notée f. La première version assombrie $f_{dk,\triangle}$ est obtenue par la multiplication-LIP de f par un scalaire λ (Fig. 4.18b). La deuxième version assombrie $f_{dk,\triangle}$ est obtenue par addition-LIP d'une constante k (Fig. 4.18e). A cause de la réflexion de la lumière à la surface des boules et d'autres objets

TABLE 4.1 – Comparaisons entres les durées des mises en œuvre directes et morphologiques pour les cartes d'Asplund de distances LIP-multiplicatives, $Asp_b^{\triangle}$ et $Asp_{b,95\%}^{\triangle}$, et LIP-additives, $Asp_b^{\triangle}$ et $Asp_{b,94\%}^{\triangle}$. Les comparaisons sont réalisées avec l'exemple de la figure 4.18. Dernière ligne : facteur de gain entre les mises en œuvre directes et morphologiques.

	$Asp_b^{\triangle} f$	$Asp_{b,95\%}^{\triangle} f$	$Asp_b^{\triangle} f$	$Asp_{b,94\%}^{\triangle} f$
Morpho.	1,4 s	4,2 s	1,4 s	4,3 s
Directe	16,1 s	24,6 s	15,4 s	25,5 s
Facteur de gain	10,9	5,9	11,0	5,9

confondants, la détection des boules représente une tâche difficile à accomplir dans ces images. Dans ce but, nous utilisons une sonde b composée d'un anneau entourant un cylindre (Fig. 4.18d). L'anneau a un rayon extérieur de rayon de 15 pixels, une épaisseur de 3 pixels, une valeur de gris de 18. Le disque a un rayon de 2 pixels et un niveau de gris de 190.

Remarque. Comme l'échelle de gris est complémentée dans le modèle LIP, la carte des distances LIP-multiplicatives $As_{b^c,p}^{\triangle} f^c$ est calculée en utilisant le complément de l'image $f^c = M - f \in]0; M[^D$, et le complément de la sonde b^c. Cependant, l'image $f_{dk,\triangle} \in]-\infty; M[^D$, assombrie par l'addition-LIP d'une constante, présente des valeurs négatives et son complément $f_{dk,\triangle}^c \in]0, +\infty[^D$ est en dehors de la dynamique permise pour la carte $[-\infty; M]^D$. Pour cette raison, l'image assombrie $f_{dk,\triangle}$ n'est pas complémentée.

Dans l'image f (Fig. 4.18a) et dans sa version assombrie $f_{dk,\triangle}$ (Fig. 4.18b) obtenue par multiplication-LIP, toutes les boules sont détectées par seuillage de leur carte des distances d'Asplund $Asp_{b^c,p}^{\triangle} f^c$ (Fig. 4.18c). La carte est la même pour les deux images f et $f_{dk,\triangle}$ à cause de son invariance par multiplication-LIP par un scalaire. Des résultats similaires sont obtenus pour l'image f (Fig. 4.18a), sa version assombrie $f_{dk,\triangle}$ par addition-LIP (Fig. 4.18e) et leur carte des distances $Asp_{b,p}^{\triangle} f$ (Fig. . 4.18f) qui est invariante par addition-LIP d'une constante. Les mêmes paramètres de la sonde sont employés pour toutes les images f, $f_{dk,\triangle}$ et $f_{dk,\triangle}$. Ces résultats montrent que les cartes des distances d'Asplund sont capables de détecter des cibles sous différentes conditions d'illumination modélisées par la multiplication-LIP $\triangle$ ou par l'addition-LIP $\triangle$.

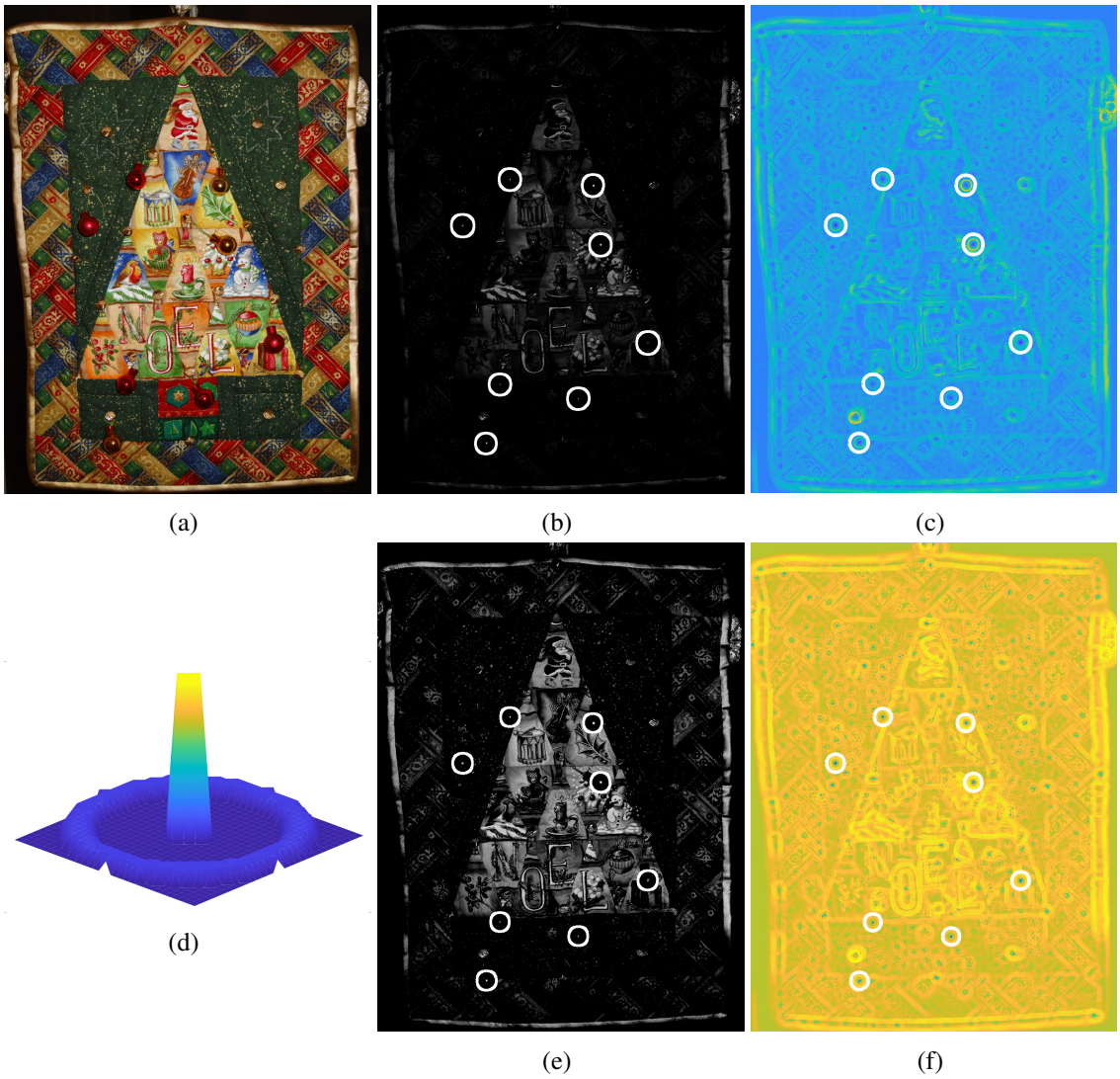


FIG. 4.18 – (a) Version couleur de l'image de luminance f et (b) sa version assombrie $f_{dk,\triangle} = (\lambda \triangle f^c)^c$ obtenue par multiplication-LIP ($\lambda = 5$) de l'image complémentée de la luminance f^c. Les boules sont détectées par seuillage de (c) la carte des distances d'Asplund LIP-multiplicatives de f^c avec une tolérance de $p = 95$ %, $Asp^{\triangle}_{b^c,p} f^c$. (d) Sonde b. (e) Version assombrie $f_{dk,\triangle} = k \triangle f$ de l'image de luminance f obtenue par addition-LIP de la constante $k = \triangle 100 = -164{,}10$. Les boules sont segmentées par seuillage de (f) la carte des distances LIP-additives d'Asplund de f avec une tolérance $p = 94$ %, $Asp^{\triangle}_{b,p} f$.

Cas réels

Les cartes des distances d'Asplund LIP-additives et LIP-multiplicatives sont ensuite illustrées sur des images réelles. Pour une meilleure interprétation visuelle, les résultats sont présentés avec des images couleur même si le traitement est effectué sur leur luminance.

Métrique LIP-additive : images acquises avec un temps d'exposition variable Sur la figure 4.19, la même scène est capturée avec trois différents temps d'exposition de la caméra (ou vitesse d'obturation). Cela donne trois images, chacune avec une clarté différente : i) une claire f (Fig. 4.19a), ii) une mi-sombre f_{dk_1} (Fig. 4.19b) and iii) une sombre f_{dk_2} (Fig. 4.19c). La scène est composée de boules brillantes sur fond multicolore avec petites boules apparaissant comme

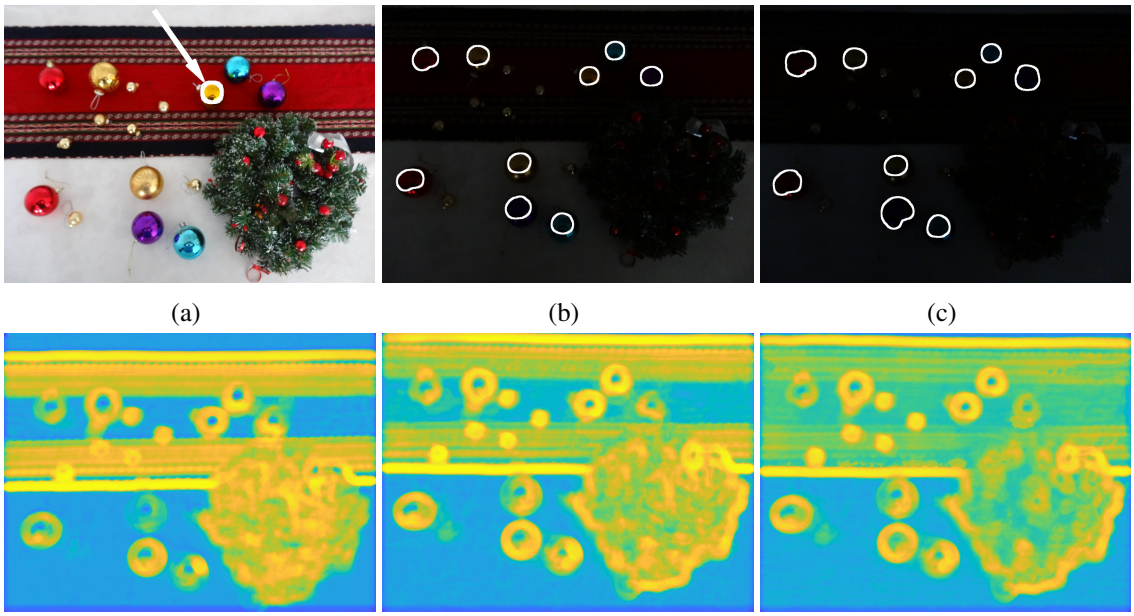
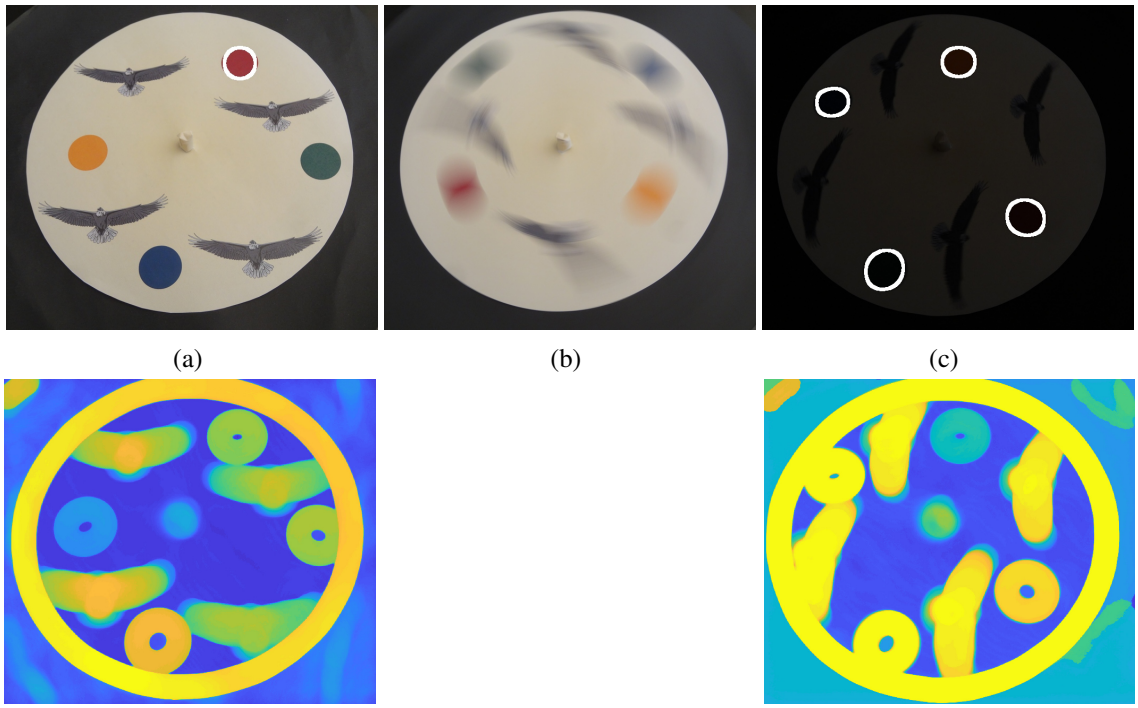


FIG. 4.19 – (a) Image claire f acquise avec un temps d'exposition de 1/5 s. La fonction sonde b est pointée par la flèche banche. (b) Détection des boules dans une image mi-sombre f_{dk_1} (temps d'exposition de 1/80 s). (c) Détection des boules dans une image sombre f_{dk_2} (temps d'exposition de 1/160 s). (d) Cartes des distances d'Asplund LIP-additives $Asp^{\triangle}_{b^c,p} f^c$ de l'image f, (e) carte $Asp^{\triangle}_{b^c,p} f^c_{dk_1}$ de l'image f_{dk_1} et (f) carte $Asp^{\triangle}_{b^c,p} f^c_{dk_2}$ de l'image f_{dk_2}. Le paramètre de tolérance p est fixé à 70 %.

des objets confondants. Afin de rendre la détection des boules plus ardue, la caméra n'est pas exactement dans la même position pour capturer les images. De plus, les boules sont de différentes couleurs, sur différents fonds et présentent plusieurs réflexions. Un disque de diamètre 55 pixels est choisi manuellement à l'intérieur d'une boule de l'image claire f (Fig. 4.19a) afin de servir de fonction sonde b. Une carte des distances d'Asplund est calculée sur le complément de chacune des trois images avec la même sonde b et en utilisant le même paramètre de tolérance $p = 70$ %. Dans les trois cartes des distances des images $Asp^{\triangle}_{b^c,p} f^c$ (Fig. 4.19d), $Asp^{\triangle}_{b^c,p} f^c_{dk_1}$ (Fig. 4.19e) et $Asp^{\triangle}_{b^c,p} f^c_{dk_2}$ (Fig. 4.19f), on peut remarquer que les amplitudes sont similaires. Ceci est causé par la faible sensibilité de la métrique d'Asplund LIP-additive aux variations d'éclairement dues aux différents temps d'exposition. Afin d'extraire la position des grosses boules, les cartes de l'image mi-sombre $Asp^{\triangle}_{b^c,p} f^c_{dk_1}$ et de l'image sombre $Asp^{\triangle}_{b^c,p} f^c_{dk_2}$ sont segmentées en utilisant la même technique – un seuil (au 33ème centile) et une reconstruction des h-minima régionaux [284] – suivi par un post-traitement morphologique (cf. remarque 4.19 ci-après). Une telle technique permet de détecter toutes les grosses boules dans les deux images sombres f_{dk_1} (Fig. 4.19b) et f_{dk_2} (Fig. 4.19c), en utilisant la même sonde b extraite sur l'image claire f (Fig. 4.19a). Ceci illustre la robustesse de la carte des distances d'Asplund LIP-additives aux différents temps d'exposition de la caméra.

Remarque (Détails sur la méthode de segmentation de la figure 4.19). Les h-minima ont une hauteur plus grande que les 1,6 premiers centiles de la carte. Seuls sont conservés les minima avec une aire plus petite que celle de la sonde et ayant une forme circulaire.

La carte des distances d'Asplund LIP-additive est aussi utile pour détecter les objets en mouvement. Sur la figure 4.20, un disque blanc avec des formes est monté sur la platine d'un tourne-disque. Les formes incluent quatre petits disques colorés et des formes confondantes (i.e. des aigles). Tout d'abord, une image f avec de bons contrastes est capturée avec un temps d'expo-



(a) (b) (c)

(d) (e)

FIG. 4.20 – (a) Image f de l'objet fixe acquise avec un temps d'exposition de 1/13 s. La fonction sonde b est entourée en blanc. (d) Carte des distances d'Asplund $Asp^{\triangle}_{b^c,p} f^c$ du complément de f. (b) Image floue $f_{mov,bl}$ de l'objet en mouvement acquise avec le même temps d'exposition de 1/13 s. (c) Détection de disque dans une image sombre $f_{mov,dk}$ de l'objet en mouvement acquis avec un temps d'exposition plus court de 1/160 s. Les boules (entourées en blanc) sont détectées en trouvant les minima régionaux de (e) sa carte des distances d'Asplund LIP-additives $Asp^{\triangle}_{b^c,p} f^c_{mov,dk}$. Pour chaque carte, le paramètre de tolérance p est fixé à 95 %.

sition approprié, 1/13 s (Fig. 4.20a). Dans cette image, une sonde circulaire b est sélectionnée à l'intérieur d'un disque coloré. Le tourne-disque est alors démarré à une vitesse de 45 tours/min et deux images sont alors capturées. La première image $f_{mov,bl}$, qui est acquise avec le même temps d'exposition que celui de f, est correctement exposée mais floue (Fig. 4.20b). Comme elle est floue, il est inutile de détecter les disques colorés dessus. Une deuxième image $f_{mov,dk}$ est acquise avec un temps d'exposition plus court de 1/160 s. Cette deuxième image n'est pas floue mais apparaît cependant plus sombre que f (Fig. 4.20c). La carte des distances d'Asplund LIP-additives de son complément, $Asp^{\triangle}_{b^c,p} f^c_{mov,dk}$ (Fig. 4.20e), est utile pour la détection des disques colorés. Ceux-ci sont détectés en trouvant les minima régionaux de la carte avec une hauteur suffisante grâce à la transformée par h-minima [284]. Les minima régionaux avec des aires trop larges ou trop petites sont aussi enlevés. En comparant la carte de l'image du disque en mouvement $Asp^{\triangle}_{b^c,p} f^c_{mov,dk}$ (Fig. 4.20e) à la carte de l'image du disque fixe $Asp^{\triangle}_{b^c,p} f^c$ (Fig. 4.20d), on peut remarquer qu'elles présentent toutes deux des formes et des amplitudes similaires bien que les images f et $f_{mov,dk}$ aient été capturées dans des conditions différentes. En particulier, les cinétiques de la scène sont différentes car le disque est fixe ou tournant et les conditions d'éclairement diffèrent significativement car les temps d'exposition varient d'un facteur 12.

La figure 4.19 montre la robustesse des métriques d'Asplund LIP-additives aux variations d'intensité causées par des temps d'expositions variables. La figure 4.20 montre son efficacité à détecter des objets en mouvement dans des images sombres acquises avec un faible temps d'exposition qui est nécessaire pour capturer une vue non floue de l'objet. De tels cas apparaissent dans de nombreuses applications comme des images médicales [234] ou industrielles [213], [225]. E.g.

en contrôle industriel les objets sont souvent présentés à la caméra sur un convoyeur (linéaire, ou circulaire, etc.) dont la vitesse varie avec le taux de production.

Métrique LIP-multiplicative : images acquises avec une absorption variable du milieu L'indépendance des cartes des distances d'Asplund LIP-multiplicatives aux variations d'éclairements dues à une absorption (ou opacité) différente de l'objet semi-transparent est vérifiée grâce à un montage que nous avons réalisé (Fig. 4.21). Il se compose d'un réservoir transparent. Un papier avec des motifs est collé sur une de ces faces. Sur la face opposée, une caméra est disposée pour capturer une image du papier à travers un milieu composé du réservoir et de son contenu : un colorant vert dilué dans de l'eau. Trois images f_Γ (Fig. 4.22a), $f_{3\Gamma}$ (Fig. 4.22b) et $f_{12\Gamma}$ (Fig. 4.22c) sont acquises pour des concentrations croissantes du colorant : Γ, 3Γ et 12Γ, où Γ est la concentration initiale du colorant dans le réservoir. On peut remarquer que la clarté des images décroît avec la croissance de la concentration du colorant. Une sonde circulaire b est manuellement sélectionnée dans l'image la plus claire f_Γ afin de détecter des formes similaires dans les deux autres images $f_{3\Gamma}$ et $f_{12\Gamma}$. Une carte des distances d'Asplund est calculée sur le complément de chaque image avec la même sonde b et le même paramètre de tolérance $p = 90\ \%$. On peut remarquer que les trois cartes des images $Asp^{\triangle}_{b^c,p} f^c_\Gamma$ (Fig. 4.22d), $Asp^{\triangle}_{b^c,p} f^c_{3\Gamma}$ (Fig. 4.22e) et $Asp^{\triangle}_{b^c,p} f^c_{12\Gamma}$ (Fig. 4.22f) présentent des amplitudes similaires. Elles sont segmentées par la même technique de seuillage – au 33ème centile – suivie de deux ouvertures aréolaires afin d'enlever les régions trop petites ou trop larges. Les régions sélectionnées sont ensuite dilatées à des fins d'affichage. On peut remarquer que tous les disques sont détectés dans les deux images les plus sombres $f_{3\Gamma}$ (Fig. 4.22b) et $f_{12\Gamma}$ (Fig. 4.22c) en utilisant la sonde b extraite dans l'image claire f_Γ. Ces résultats montrent la faible sensibilité des cartes des distances d'Asplund LIP-multiplicatives à des variations d'éclairement causées par des absorptions différentes. Cette situation se produit dans les images acquises par transmission (e.g. les radiographies rayons X, la tomographie, la spectrophotométrie, etc.) [144].

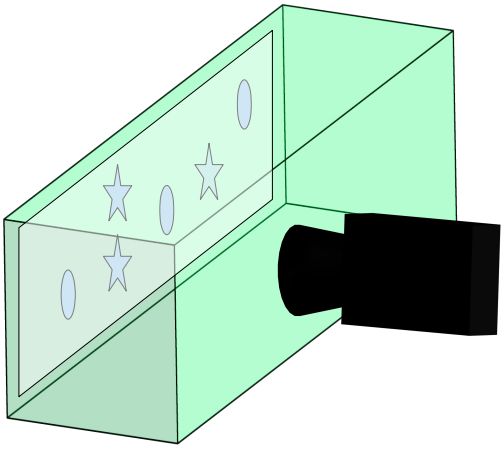


FIG. 4.21 – Montage comportant d'une caméra qui acquiert une image à travers un milieu d'absorption variable composé d'un réservoir transparent rempli d'eau avec un colorant vert. La concentration du colorant est augmentée par ajout dans la solution avant l'acquisition de chaque image.

Remarque. Dans cette sous-section, des sondes avec invariance circulaire ont été utilisées pour faciliter la présentation. Cependant, les métriques d'Asplund sont aussi efficace pour détecter les objets non-circulaires avec des sondes adaptées. De plus, ces métriques peuvent être comparées au détecteur SIFT (*Scale-Invariant Feature Transform*) qui est robuste aux variations d'éclairement [166]. La comparaison n'a pas été effectuée dans le présent mémoire, car la robustesse aux va-

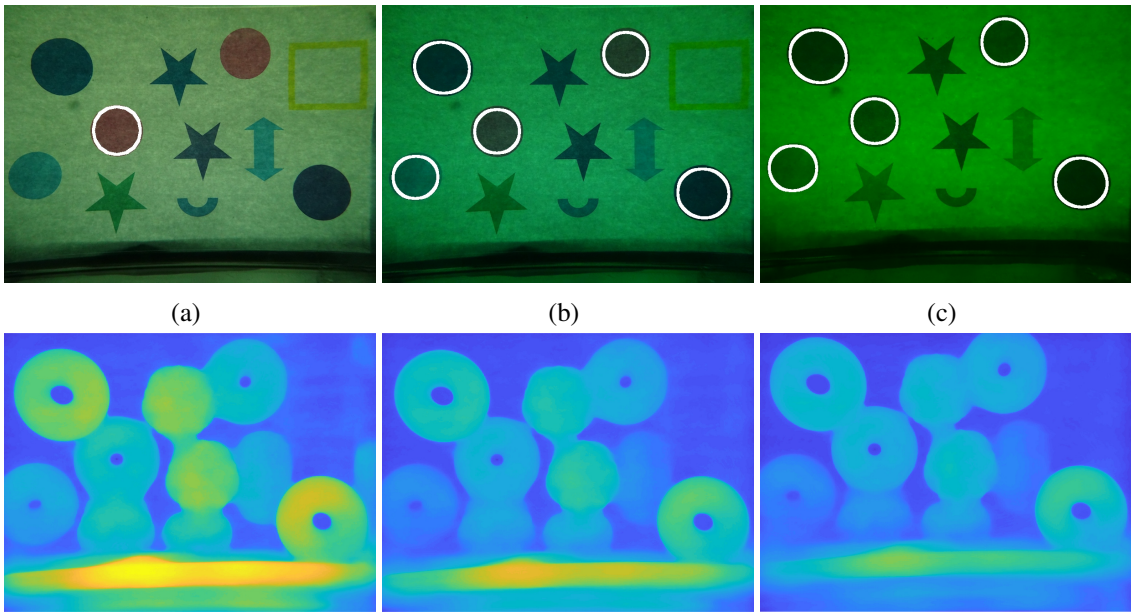


FIG. 4.22 – (a) Image f_Γ acquise avec une faible concentration, Γ, de colorant. La fonction sonde b est entourée en blanc. (b) Détection des disques dans une image $f_{3\Gamma}$ acquise avec une concentration intermédiaire, 3Γ, de colorant. (c) Détection des disques dans une image $f_{12\Gamma}$ acquise avec une concentration élevée, 12Γ, de colorant. (d) Cartes des distances d'Asplund LIP-multiplicatives $Asp^{\triangle}_{b^c,p} f^c_\Gamma$ de l'image f_Γ. (e) Carte $Asp^{\triangle}_{b^c,p} f^c_{3\Gamma}$ de l'image $f_{3\Gamma}$. (f) Carte $Asp^{\triangle}_{b^c,p} f^c_{12\Gamma}$ de l'image $f_{12\Gamma}$. Le paramètre de tolérance p est fixé à 90 %.

riations d'éclairement du détecteur SIFT n'est pas fondée sur une loi optique contrairement aux métriques d'Asplund. Cependant, nous avons montré dans [234] qu'améliorer les contrastes des images donne un meilleur recalage pour une méthode utilisant les points SIFT [166] (cf. section 6.2, p. 201). Nous pourrions développer un SIFT dans le cadre de travail LIP. Ces découvertes seront étudiées dans des travaux futurs.

4.3.6 Lien entre les deux métriques d'Asplund fonctionnelles LIP-multiplicative et LIP-additive

Dans [216], j'ai démontré le lien entre l'opérateur des cartes des distances d'Asplund LIP-multiplicatives $Asp^{\triangle}_b :]0; M[^D \to (\mathbb{R}^+)^D$ et l'opérateur des cartes des distances d'Asplund LIP-additives $Asp^{\triangle}_{b_1} :]-\infty; M[^D \to]-\infty; M[^D$. b et b_1 sont les fonctions structurantes avec des valeurs appartenant respectivement à des intervalles différents : $b : D_b \to]0; M[$ et $b_1 : D_b \to]-\infty; M[$, avec $D_b \subset D$.

Pour relier ces deux opérateurs, un isomorphisme est nécessaire entre le treillis $(\overline{\mathbb{R}}^+)^D$ des cartes des majorants et minorants LIP-multiplicatives, $\lambda_b f$ et $\mu_b f$, et le treillis $[-\infty; M]^D$ des cartes des majorants et minorants LIP-additives $c_{1_b} f$ et $c_{2_b} f$. Il s'agit de l'isomorphisme $\xi : [-\infty; M]^D \to \overline{\mathbb{R}}^D$ et de son inverse $\xi^{-1} : \overline{\mathbb{R}}^D \to [-\infty; M]^D$ définis dans les équations (4.77) et (4.78), p. 114, par $\xi(f) = -M \ln(1 - f/M)$ et $\xi^{-1}(f) = M(1 - \exp(-f/M))$.

Dans [216], j'ai démontré la relation suivante entre les deux opérateurs.

Proposition 11. *Soient $f \in]0; M[^D$ une image, $f^c = M - f$ son complément et $b \in]0; M[^{D_b}$ une fonction structurante. L'opérateur de* carte des distances d'Asplund LIP-additives *est lié à celui*

de carte des distances LIP-multiplicatives *par la relation :*

$$Asp_b^{\triangle} f = \frac{1}{M}\xi\left(Asp_{(\xi(b))^c}^{\triangle}(\xi(f))^c\right), \tag{4.85}$$

où $(\xi(f))^c = M - \xi(f) \in]-\infty; M[^D$.

Corollaire 3. *En utilisant le changement de variable suivant* $f_1 = (\xi(f))^c$ *et* $b_1 = (\xi(b))^c$*, avec* $f_1 \in \mathcal{F}_M$ *et* $b_1 \in]-\infty; M[^{D_b}$*, une relation équivalente est obtenue :*

$$Asp_{b_1}^{\triangle} f_1 = \xi^{-1}\left(M.Asp_{\xi^{-1}(b_1^c)}^{\triangle}\xi^{-1}(f_1^c)\right). \tag{4.86}$$

Remarque. L'équation (4.85) peut être aussi écrite comme :

$$\begin{aligned} Asp_{(\xi(b))^c}^{\triangle}(\xi(f))^c &= \xi^{-1}(M.Asp_b^{\triangle} f) \\ &= M(1 - \exp(-Asp_b^{\triangle} f)). \end{aligned} \tag{4.87}$$

Comme la carte des distances d'Asplund $Asp_b^{\triangle} f$ de f est un élément de $[0; +\infty[^D$, la carte des distances d'Asplund LIP-additives $Asp_{(\xi(b))^c}^{\triangle}(\xi(f))^c$ de $(\xi(f))^c$ est un élément de l'ensemble $[0; M[^D = \mathcal{I}$ et est donc une image.

Remarque. La même relation existe entre les deux métriques d'Asplund fonctionnelles LIP-multiplicative et LIP-additive :

$$d_{asp}^{\triangle}(f, g) = \frac{1}{M}\xi\Big[d_{asp}^{\triangle}([\xi(f)]^c, [\xi(g)]^c)\Big]. \tag{4.88}$$

Les équations (4.85) et (4.86) montrent que les opérateurs des cartes des distances d'Asplund LIP-multiplicatives et LIP-additives ainsi que leurs métriques associées sont liés par l'isomorphisme ξ. Ces relations permettent de calculer une carte des distances d'une image par le moyen de l'autre carte des distances d'une transformée de cette image. E.g. la carte des distances LIP-additives $Asp_{b_1}^{\triangle} f_1$ de la fonction f_1 peut être calculée en utilisant l'opérateur (et le programme informatique) de la carte des distances LIP-multiplicatives $Asp_{\xi^{-1}(b_1^c)}^{\triangle}\xi^{-1}(f_1^c)$ de la fonction transformée $\xi^{-1}(f_1^c)$. Cependant, les deux équations ne lient pas directement les deux cartes des distances de l'image f. E.g. la carte des distances LIP-multiplicatives de l'image f, $Asp_b^{\triangle} f$, n'est pas directement liée à la carte des distances LIP-additives de l'image f, $Asp_b^{\triangle} f$.

La relation donnée dans les équations (4.85) et (4.86) n'est pas surprenante. En effet, la carte des distances d'Asplund LIP-additives $Asp_{(\xi(b))^c}^{\triangle}(\xi(f))^c$ est une image qui appartient à l'ensemble $[0; M[^D$, tandis que la carte des distances d'Asplund LIP-multiplicatives $Asp_b^{\triangle} f$ appartient à l'ensemble $(\mathbb{R}^+)^D$. L'isomorphisme ξ permet de passer de l'espace image, de la carte des distances LIP-additives $Asp_{(\xi(b))^c}^{\triangle}(\xi(f))^c$, à l'espace des fonctions réelles positives, de la carte des distances LIP-multiplicatives $Asp_b^{\triangle} f$.

4.3.7 Conclusion sur les métriques d'Asplund fonctionnelles

Avec Michel Jourlin, nous avons présenté dans [231] un nouveau cadre de mise en correspondance de formes robustes aux variations d'éclairement entre des images faiblement et fortement contrastées. Celui-là est composé de deux métriques. Premièrement, la métrique d'Asplund LIP-multiplicative est robuste aux changements d'illumination dus aux variations d'absorption ou d'opacité de l'objet. Deuxièmement, la métrique d'Asplund LIP-additive est robuste aux changements d'illuminations causés par les variations de l'intensité de la source ou du temps d'exposition de la caméra. Les deux métriques sont fondées respectivement sur les lois multiplicative et additive

du modèle LIP qui leur donnent de fortes propriétés optiques. Les deux métriques fonctionnelles sont ainsi théoriquement *insensibles à des variations d'éclairement spécifiques*. Elles étendent aux images la propriété d'*insensibilité au grossissement de l'objet* de la métrique d'Asplund entre des formes binaires [15], [96]. Après une présentation des métriques fonctionnelles et de leurs propriétés, nous avons introduit des versions robustes au bruit. Nous avons démontré que les équations des cartes des distances d'Asplund entre une image et une fonction sonde sont composées d'opérations morphologiques. Les deux cartes des distances sont en particulier liées aux opérations de dilatation et d'érosion pour des fonctions. Cette relation facilite la programmation des cartes des distances car ces opérations existent dans de nombreuses bibliothèques de traitement d'images. Les propriétés des deux métriques ont alors été vérifiées sur des cas réels et simulés. Les résultats montrent que les cartes des distances d'Asplund LIP-multiplicatives et LIP-additives sont respectivement capables de détecter des formes dans des images acquises avec des illuminations différentes, causées par une forte absorption de l'objet ou par un temps d'exposition de la caméra plus court. Surtout, la sonde peut être extraite dans une image fortement contrastée et la détection effectuée dans une image faiblement contrastée. *Ces propriétés ouvrent la voie à de nombreuses applications où les variations d'éclairement ne sont pas contrôlées* : e.g. dans l'industrie [225], en médecine [219], [234], en gestion du trafic routier [85], en sécurité et surveillance [81], en imagerie des objets en mouvement [213], etc. Ces métriques peuvent aussi être étendues à l'analyse des images couleur et multivariées [209], [229] et sont liées à la *Morphologie Mathématique Logarithmique* [217] comme nous le verrons ci-après. Il serait aussi intéressant de faire une étude des métriques fonctionnelles d'Asplund pour différents types de bruit.

4.3.8 Les métriques d'Asplund pour les images couleur et multivariées

Avec Michel Jourlin, j'ai étendu les métriques d'Asplund fonctionnelles aux cas des images couleur par deux approches vectorielles (i.e. combinant l'ensemble des composantes) [212], [219] qui consistent en une double palpation sur l'ensemble des composantes couleur.

i) La première utilise le modèle LIP [209].

ii) La seconde s'appuie sur le modèle LIPC [229].

Approche avec le modèle LIP

Avant de définir une métrique entre des images couleur, il est tout d'abord nécessaire de définir une métrique entre deux vecteurs de couleur.

Définition 21. *Soient deux couleurs* $C_1 = (R_1, G_1, B_1)$ *et* $C_2 = (R_2, G_2, B_2) \in]0; M[^3$. *Leur distance d'Asplund est égale à :*

$$d^{\triangle}_{asp,C}(C_1, C_2) = \ln\left(\frac{\lambda}{\mu}\right), \tag{4.89}$$

où $\lambda = \inf\{\alpha \in \mathbb{R} \mid \alpha \triangle R_1 \geq R_2, \alpha \triangle G_1 \geq G_2, \alpha \triangle B_1 \geq B_2\}$
et $\mu = \sup\{\alpha \in \mathbb{R} \mid \alpha \triangle R_1 \leq R_2, \alpha \triangle G_1 \leq G_2, \alpha \triangle B_1 \leq B_2\}$.

Remarque. Cette approche est un bon compromis entre l'utilisation du modèle LIPC (que nous verrons ci-après) et une solution triviale consistant à calculer une distance sur chaque composante, comme nous l'avons fait à niveaux de gris, et à les combiner de différentes manières [140].

Remarque. La métrique d'Asplund entre deux couleurs C_1 et C_2, $d^{\triangle}_{asp,C}$, reste inchangée lorsqu'une des couleurs (e.g. C_1) est remplacée par tout homothétique $k \triangle C_1$, pour $k > 0$. Pour rappel, la couleur $k \triangle C_1$ est plus claire que C_1 lorsque $k \in [0, 1[$ et plus sombre lorsque $k \in [1, +\infty[$. Cette propriété montre l'insensibilité de la métrique d'Asplund LIP-multiplicative entre deux couleurs à une variation importante d'illumination modélisée par la loi LIP-multiplicative $\triangle$.

D'un point de vue mathématique, $d^{\triangle}_{asp,C}$ est une métrique si les couleurs $C_n = (R_n, G_n, B_n)$, avec $n \in \{1, 2\}$, sont remplacées par leurs classes d'équivalence $\tilde{C}_n = \left\{C \in]0; M[^3 \mid C\mathscr{R}C_n\right\}$. Deux couleurs $C_1 = (R_1, G_1, B_1)$ et $C_2 = (R_2, G_2, B_2)$ sont dites « en relation » $\mathscr{R}$ si elles satisfont :

$$C_1\mathscr{R}C_2 \iff \exists\alpha \in \mathbb{R}^+, \alpha \triangle C = C_n,$$

i.e. $(\alpha \triangle R = R_n, \alpha \triangle G = G_n, \alpha \triangle B = B_n)$.

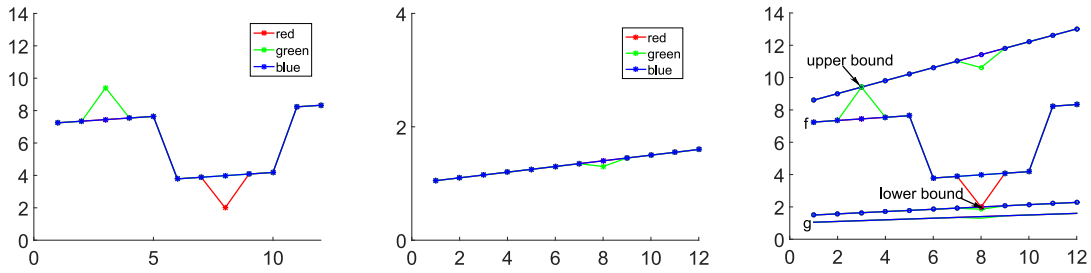
Nous pouvons maintenant introduire la définition suivante.

Définition 22. *La métrique d'Asplund LIP-multiplicative entre deux images couleur* $\mathbf{f}$ *et* $\mathbf{g} : D \to]0; M[^3$ *est égale à :*

$$d^{\triangle}_{asp}(\mathbf{f}, \mathbf{g}) = \ln\left(\frac{\lambda}{\mu}\right), \tag{4.90}$$

où $\lambda = \inf\{\alpha \in \mathbb{R} \mid \alpha \triangle f_R \geq g_R, \alpha \triangle f_G \geq g_G, \alpha \triangle f_B \geq g_B\}$
et $\mu = \sup\{\alpha \in \mathbb{R} \mid \alpha \triangle f_R \leq g_R, \alpha \triangle f_G \leq g_G, \alpha \triangle f_B \leq g_B\}$.

Sur la figure 4.23, la métrique d'Asplund a été testée entre la sonde couleur $\mathbf{g}$ et l'image couleur $\mathbf{f}$ sur leur domaine de définition D. Dans ce cas, la distance d'Asplund est égale à $d^{\triangle}_{asp}(\mathbf{f}, \mathbf{g}) = 1{,}76$. La sonde inférieure $\mu \triangle \mathbf{g} = (\mu \triangle g_R, \mu \triangle g_G, \mu \triangle g_B)$ et la sonde supérieure $\lambda \triangle \mathbf{g}$ sont respectivement en contact par-dessous et par-dessus avec la courbe représentative de l'image f comme illustré dans la figure 4.23c.



(a) Image couleur $\mathbf{f}$. (b) Sonde couleur $\mathbf{g}$. (c) Sonde supérieure et inférieure.

FIG. 4.23 – Illustration de la distance d'Asplund entre deux images couleur $d^{\triangle}_{asp}(\mathbf{f}, \mathbf{g}) = 1{,}76$. (a) Image $\mathbf{f}$ et (b) sonde $\mathbf{g}$ en couleur. Chaque composante couleur est représentée par une courbe de la couleur correspondante. (c) Double palpation par la sonde inférieure $\mu \triangle \mathbf{g}$ et la sonde supérieure $\lambda \triangle \mathbf{g}$.

Comme pour les images à niveaux de gris, une carte des distances d'Asplund peut être définie pour les images couleur. Soit $(\mathcal{I}^*)^3 = (]0; M[^3)^D$ l'ensemble des images couleur avec des valeurs strictement positives.

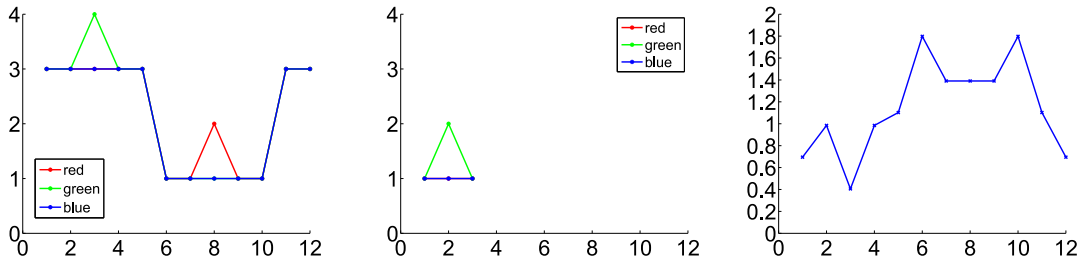
Définition 23. *Soient* $\mathbf{f} : D \to]0; M[^3$ *une image couleur et* $\mathbf{b} : D_b \to]0; M[^3$ *une sonde couleur, avec* $D_b \subset D$. *La carte des distances d'Asplund* $Asp^{\triangle}_b : (\mathcal{I}^*)^3 \to (\mathbb{R}^+)^D$ *est définie pour tout* $x \in D$ *par :*

$$Asp^{\triangle}_{\mathbf{b}}\mathbf{f}(x) = d^{\triangle}_{asp}(\mathbf{f}_{|D_b(x)}, \mathbf{b}), \tag{4.91}$$

où $D_b(x)$ *est le voisinage associé à* D_b *et centré au point* $x \in D$.

Sur la figure 4.24, la carte des distances d'Asplund est calculée entre une fonction couleur et une sonde. Le minimum de la carte correspond au lieu où une forme couleur est similaire à la sonde.

Comme expliqué précédemment (cf. sous-sous-section « Version robuste au bruit », p. 106), la métrique d'Asplund est très sensible au bruit car elle utilise des calculs d'extrema. Il est donc



(a) Image couleur $\mathbf{f}$. (b) Sonde couleur $\mathbf{b}$. (c) Carte des distances d'Asplund $Asp_{\mathbf{b}}^{\triangle}\mathbf{f}$.

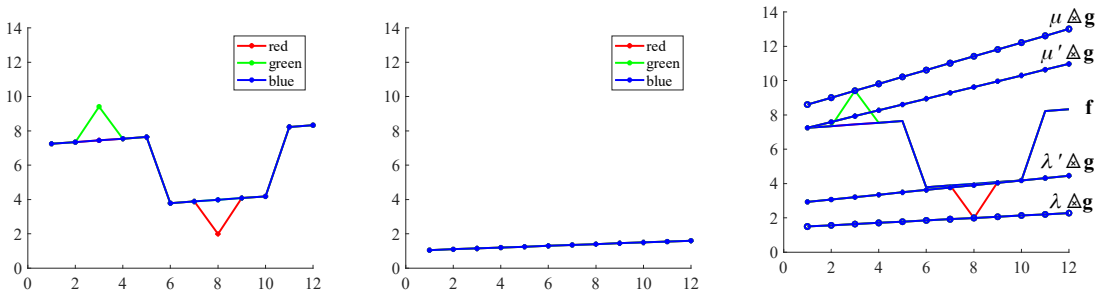
FIG. 4.24 – (c) Carte des distances d'Asplund $As_{\mathbf{b}}^{\triangle}\mathbf{f}$ entre (a) une image couleur $\mathbf{f}$ et (b) une sonde couleur $\mathbf{b}$. Chaque composante couleur est représentée par une courbe de la même couleur.

intéressant d'introduire une version robuste au bruit dans le contexte de la « théorie de la mesure ». Sur le même principe que celui utilisé pour les images à niveaux de gris, une métrique d'Asplund $d_{asp,p}^{\triangle}(\mathbf{f},\mathbf{g})$ avec un paramètre de tolérance p peut être définie entre deux images $\mathbf{f}$ et $\mathbf{g}$. Le but est d'écarter les points extrémaux correspondant au bruit. Pour ce faire, un voisinage de la fonction $\mathbf{f}$ s'écrit comme :

$$N_{P,d_{asp}^{\triangle},\varepsilon,p}(\mathbf{f}) = \left\{\mathbf{g} \mid \exists D' \subset D, d_{asp}^{\triangle}(\mathbf{f}_{|D'},\mathbf{g}_{|D'}) < \varepsilon \text{ et } P(D \setminus D') \leq p\right\}, \tag{4.92}$$

où ε représente une tolérance sur la distance d'Asplund et $P(D\setminus D') = \frac{\#(D\setminus D')}{\#D}$ est le pourcentage de points conservés.

Sur la figure 4.25, une tolérance de $p = 80\ \%$ est utilisée. Elle revient à écarter deux points extrémaux représentant le bruit. Ainsi la distance d'Asplund décroît de $d_{asp}^{\triangle}(\mathbf{f},\mathbf{g}) = 1{,}76$ à $d_{asp,p=80\%}^{\triangle}(\mathbf{f},\mathbf{g}) = 0{,}91$.

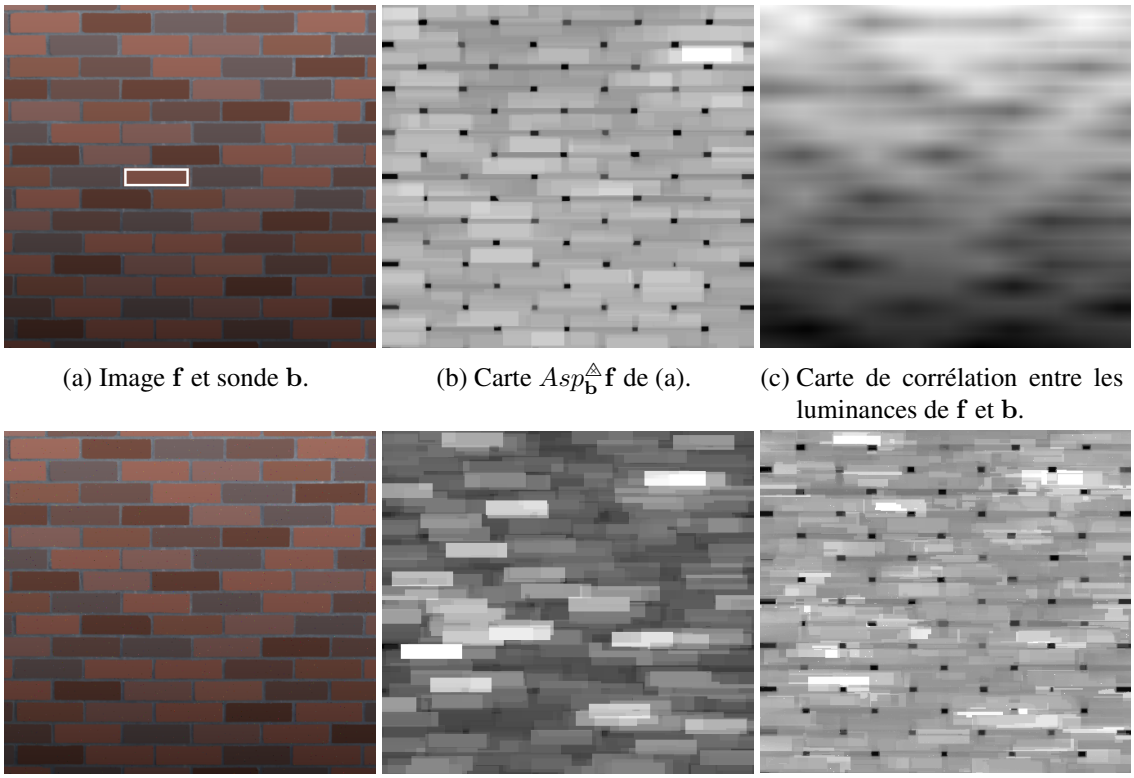


(a) Image couleur $\mathbf{f}$. (b) Sonde couleur $\mathbf{g}$. (c) Sonde supérieure et inférieure, $p = 80\ \%$.

FIG. 4.25 – Distance d'Asplund avec une tolérance de $p = 80\ \%$ entre deux images couleur, $d_{asp,p}^{\triangle}(\mathbf{f},\mathbf{g}) = 0{,}91$, et sans tolérance, $d_{asp}^{\triangle}(\mathbf{f},\mathbf{g}) = 1{,}76$. (a) Image $\mathbf{f}$ et (b) sonde $\mathbf{g}$ en couleur. Chaque composante couleur est représentée par un courbe de la couleur correspondante. (c) Double palpation par la sonde inférieure $\mu \triangle \mathbf{g}$ et la sonde supérieure $\lambda \triangle \mathbf{g}$. μ et λ sont les scalaires qui multiplient la sonde sans la tolérance sur la distance. μ' et λ' sont les scalaires qui multiplient la sonde avec la tolérance sur la distance.

Une carte des distances d'Asplund avec tolérance $Asp^{\triangle}_{\mathbf{b},p} : (\mathcal{I}^*)^3 \to (\mathbb{R}^+)^D$ est alors définie pour tout $x \in D$ comme :

$$Asp^{\triangle}_{\mathbf{b},p}\mathbf{f}(x) = d^{\triangle}_{asp,p}(\mathbf{f}_{|D_b(x)}, \mathbf{b}). \tag{4.93}$$



(a) Image $\mathbf{f}$ et sonde $\mathbf{b}$. (b) Carte $Asp^{\triangle}_{\mathbf{b}}\mathbf{f}$ de (a). (c) Carte de corrélation entre les luminances de $\mathbf{f}$ et $\mathbf{b}$.

(d) Image $\tilde{\mathbf{f}}$ avec du bruit. (e) Carte $Asp^{\triangle}_{\mathbf{b}}\tilde{\mathbf{f}}$ de (d). (f) Carte $Asp^{\triangle}_{\mathbf{b},p=98\%}\tilde{\mathbf{f}}$ de (d).

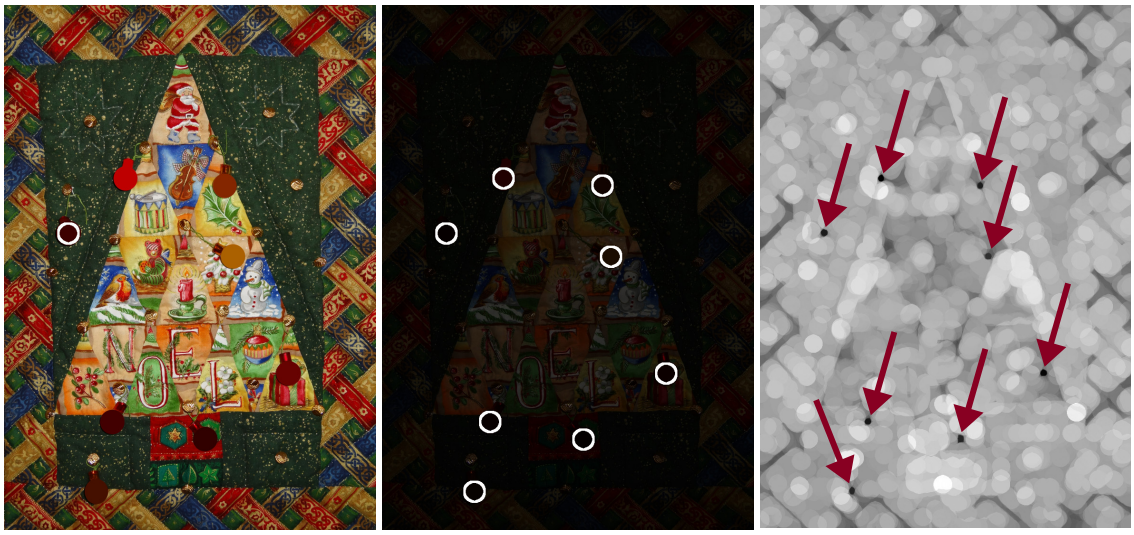
FIG. 4.26 – (a) Image couleur $\mathbf{f}$ acquise sans bruit et sonde couleur $\mathbf{b}$ (entourée de blanc). (b) Carte des distances d'Asplund $Asp^{\triangle}_{\mathbf{b}}\tilde{\mathbf{f}}$ entre l'image $\mathbf{f}$ et la sonde $\mathbf{b}$. (c) Carte de corrélation entre les images (à niveaux de gris) de luminance de l'image couleur $\mathbf{f}$ et de la sonde couleur $\mathbf{b}$. (d) Image $\tilde{\mathbf{f}}$ bruitée par ajout sur l'image $\mathbf{f}$ d'un bruit blanc gaussien avec une moyenne de 0, une variance de 2,6 et une densité spatiale de $1\ \%$. (e) Carte des distances d'Asplund sans tolérance $Asp^{\triangle}_{\mathbf{b}}\tilde{\mathbf{f}}$ et (f) avec une tolérance $Asp^{\triangle}_{\mathbf{b},p}\tilde{\mathbf{f}}$ de $p = 98\ \%$ entre l'image bruitée $\tilde{\mathbf{f}}$ et la sonde $\mathbf{b}$.

Sur la figure 4.26, on cherche à détecter des briques de couleur homogène à l'intérieur d'une image couleur d'un mur de briques. Il est à noter que l'image $\mathbf{f}$ présente une dérive d'éclairement selon l'axe vertical (Fig. 4.26a). Sur la carte des distances d'Asplund $Asp^{\triangle}_{\mathbf{b}}\mathbf{f}$ de l'image $\mathbf{f}$ acquise (quasiment) sans bruit (Fig. 4.26b), les minima régionaux correspondent aux lieux où la sonde est similaire à l'image (selon la distance d'Asplund). A titre de comparaison, la carte de corrélation, entre les images (à niveaux de gris) des luminances de l'image $\mathbf{f}$ et de la sonde $\mathbf{b}$, ne permet pas de détecter précisément la position des briques (Fig. 4.26c). Sur l'image avec du bruit $\tilde{\mathbf{f}}$ (Fig. 4.26d), la carte des distances d'Asplund est très sensible au bruit (Fig. 4.26e). Il est donc nécessaire d'utiliser une tolérance dans la carte $Asp^{\triangle}_{\mathbf{b},p}\tilde{\mathbf{f}}$ (Fig. 4.26f), afin de trouver les minima correspondant à la position des briques. On peut remarquer que :

— pour l'image bruitée $\tilde{\mathbf{f}}$, les minima sont préservés dans la carte des distances d'Asplund avec une tolérance $Asp^{\triangle}_{\mathbf{b},p}\tilde{\mathbf{f}}$ (Fig. 4.26f) comparée à la carte sans tolérance $Asp^{\triangle}_{\mathbf{b}}\tilde{\mathbf{f}}$ (Fig. 4.26e) ;

— pour les deux images sans bruit $\mathbf{f}$ et avec du bruit $\tilde{\mathbf{f}}$, les cartes des distances d'Asplund sont insensibles à une dérive d'éclairement : $Asp^{\triangle}_{\mathbf{b}}\mathbf{f}$ (Fig. 4.26b), $Asp^{\triangle}_{\mathbf{b}}\tilde{\mathbf{f}}$ (Fig. 4.26e) et $Asp^{\triangle}_{\mathbf{b},p}\tilde{\mathbf{f}}$ (Fig. 4.26f).

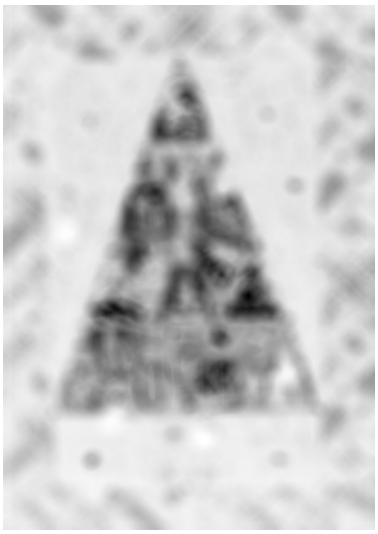
Les minima régionaux des cartes peuvent être extraits en utilisant des opérations usuelles de Morphologie Mathématique comme la transformée h-minima [284].



(a) Image initiale claire $\mathbf{f}$ et sonde $\mathbf{b}$.

(b) Boules détectées dans l'image sombre $\tilde{\mathbf{f}}$.

(c) Carte $Asp_{\mathbf{b}}^{\triangle}\tilde{\mathbf{f}}$.



(d) Carte des suprema des distances $L^*a^*b^*$.

(e) Carte des moyennes des distances $L^*a^*b^*$.

FIG. 4.27 – Détection de boules de Noël colorées dans (b) une image sombre $\tilde{\mathbf{f}}$ à partir d'une sonde $\mathbf{b}$ extraite dans (a) une image claire $\mathbf{f}$. La délinéation de la sonde $\mathbf{b}$ apparaît en blanc dans (a) de même que les boules détectées dans (b). (c) Carte des distances d'Asplund $Asp_{\mathbf{b}}^{\triangle}\tilde{\mathbf{f}}$ entre l'image sombre $\tilde{\mathbf{f}}$ et la sonde $\mathbf{b}$ extraite dans l'image claire. Les minima sont indiqués par des flèches. Comparaison avec les cartes (d) des suprema et (e) des moyennes des distances entre l'image et la sonde dans le contexte $L^*a^*b^*$ CIE 1976.

La figure 4.27 montre deux images couleur de la même scène, une image claire $\mathbf{f}$ (Fig. 4.27a) et une image sombre $\tilde{\mathbf{f}}$ (Fig. 4.27b). Cette dernière est obtenue par multiplication LIP de chacune des composantes par un scalaire $\lambda > 1$ qui simule sur l'intensité de l'image l'effet d'une augmentation d'opacité du milieu entre l'objet et la caméra. La sonde $\mathbf{b}$ représentant une boule de Noël est extraite dans l'image claire et utilisée pour calculer la carte des distances d'Asplund $Asp_{\mathbf{b}}^{\triangle}\tilde{\mathbf{f}}$ dans l'image sombre (Fig. 4.27c). Grâce à l'extraction des minima régionaux de la carte, les boules sont détectées et leurs contours sont ajoutés à l'image sombre de la figure 4.27b. On constate que

la carte des distances d'Asplund est robuste aux variations d'éclairement correspondant à une variation d'opacité du milieu semi-transparent entre la scène et la caméra. A titre de comparaison, les cartes des distances entre l'image $\mathbf{f}$ et la sonde ont été calculées dans le contexte CIE 1976 $L^*a^*b^*$ [272]. La distance $L^*a^*b^*$au point $x \in D$ entre un pixel-vecteur de la sonde $\mathbf{b} = (b_L, b_a, b_b)$ et de l'image $\mathbf{f} = (f_{L)}, f_a, f_b)$ est égale à :

$$d_{Lab}(\mathbf{f}(x), \mathbf{b}(x)) = \sqrt{\frac{1}{3}\left[(f_L(x) - b_L(x))^2 + (f_a(x) - b_a(x))^2 + (f_b(x) - b_b(x))^2\right]}. \quad (4.94)$$

Pour tous les points x du domaine D de l'image $\mathbf{f}$, on calcule le supremum ou la moyenne des distances $L^*a^*b^*$, d_{Lab}, entre la sonde $\mathbf{b} : D_b \to [0; M[^3$ et la restriction de l'image $\mathbf{f}_{|D_b(x)} : D_b(x) \to [0; M[^3$ à la fenêtre glissante $D_b(x)$ du domaine de la sonde centrée en $x \in D$:

$$\sup_{y \in D_b(x)} \left\{ d_{Lab}\left(\mathbf{f}_{|D_b(x)}(y), \mathbf{b}(y)\right)\right\}, \quad (4.95)$$

$$\sqrt{\frac{1}{\#D_b} \sum_{y \in D_b(x)} d_{Lab}\left(\mathbf{f}_{|D_b(x)}(y), \mathbf{b}(y)\right)^2}. \quad (4.96)$$

On constate que ni la carte des suprema des distances $L^*a^*b^*$ (Fig. 4.27d), ni la carte de leur moyenne (Fig. 4.27e) ne permettent de détecter efficacement les boules colorées. La métrique d'Asplund LIP-multiplicative couleur est donc efficace pour de la mise en correspondance de formes couleur sous éclairement variable. Elle est aussi valide pour les images hyperspectrales en appliquant la même définition sur plus de trois composantes.

Remarque. De la même manière que nous venons de procéder pour la métrique d'Asplund LIP-multiplicative dans la présente sous-sous-section et dans [209], Jourlin a introduit la métrique d'Asplund LIP-additive pour les images couleur dans [140]. Celle-ci présente la propriété d'être robuste aux variations d'éclairement modélisées par la soustraction LIP d'une constante et correspondant à une variation de l'intensité de la source ou du temps d'exposition de la caméra.

Approche avec le modèle LIPC

Avec Michel Jourlin, j'ai aussi défini une métrique d'Asplund LIPC-multiplicative avec le modèle couleur LIPC [229].

Etant donnés deux vecteurs de couleurs $C_1 = (R_1, G_1, B_1)$ et $C_2 = (R_2, G_2, B_2) \in [0; M[^3$ et comme nous cherchons des minorants et des majorants, il est nécessaire d'ordonner les vecteurs de couleur. Pour ce faire, un ordre marginal [22] est utilisé :

$$C_1 \geq C_2 \iff \{R_1 \geq R_2 \text{ et } G_1 \geq G_2 \text{ et } B_1 \geq B_2\}. \quad (4.97)$$

Définissons ensuite la distance entre deux couleurs de le cadre de travail couleur LIPC.

Définition 24. *Soient deux couleurs C_1 et $C_2 \in]0; M[^3$. Leur distance d'Asplund (LIPC-multiplicative) est égale à :*

$$d_{asp}^{\triangle_c}(C_1, C_2) = \ln\left(\frac{\mu}{\lambda}\right), \quad (4.98)$$

où $\lambda = \inf\{\alpha \in \mathbb{R}, \alpha \triangle_c C_2 \geq C_1\}$ et $\mu = \sup\{\alpha \in \mathbb{R}, \alpha \triangle_c C_2 \leq C_1\}$.

De manière rigoureuse, $d_{asp}^{\triangle_c}$ est une métrique si les couleurs C_n sont remplacées par leurs classes d'équivalence $\tilde{C}_n = \{C \in]0; M[^3 \mid \exists \alpha \in \mathbb{R}^+, \alpha \triangle_c C = C_n\}$.

Remarque. Dans l'équation (4.98), contrairement à la distance d'Asplund (LIP-multiplicative) définie dans [209] (cf. éq. 4.90), nous avons $\lambda \leq \mu$. En effet, par définition, les échelles de gris du modèle couleur LIPC sont inversées par rapport au modèle LIP à niveaux de gris [145] (cf. Fig. 4.5, p. 94).

La métrique d'Asplund entre deux images couleur peut alors être définie comme il s'ensuit.

Définition 25. *La métrique d'Asplund (LIPC-multiplicative) entre deux images couleur* $\mathbf{f}$ *et* $\mathbf{g}$: $D \to]0; M[^3$ *est égale à :*

$$d_{asp}^{\triangle_c}(\mathbf{f}, \mathbf{g}) = \ln\left(\frac{\mu}{\lambda}\right), \tag{4.99}$$

où $\lambda = \inf\{\alpha \in \mathbb{R} \mid \forall x \in D, \alpha \triangle_c \mathbf{g}(x) \geq \mathbf{f}(x)\}$
et $\mu = \sup\{\alpha \in \mathbb{R} \mid \forall x \in D, \alpha \triangle_c \mathbf{g}(x) \leq \mathbf{f}(x)\}$.

La métrique fonctionnelle d'Asplund $d_{asp}^{\triangle_c}$ peut être adaptée au traitement local pour calculer une carte des distances entre une image et une sonde. Soit $(\mathcal{I}^*)^3 = (]0; M[^3)^D$ l'ensemble des images couleur avec des valeurs strictement positives.

Définition 26. *Soient une image couleur* $\mathbf{f} : D \to]0; M[^3$ *et une sonde couleur* $\mathbf{b} : D_b \to]0; M[^3$, *avec* $D_b \subset D$. *La carte des distances d'Asplund LIPC-multiplicative* $Asp_b^{\triangle} : (\mathcal{I}^*)^3 \to (\mathbb{R}^+)^D$ *est définie pour tout* $x \in D$ *par :*

$$Asp_{\mathbf{b}}^{\triangle_c}\mathbf{f}(x) = d_{asp}^{\triangle_c}(\mathbf{f}_{|D_t(x)}, \mathbf{b}), \tag{4.100}$$

où $D_b(x)$ *est le voisinage associé à* D_b *et centré au point* $x \in D$.

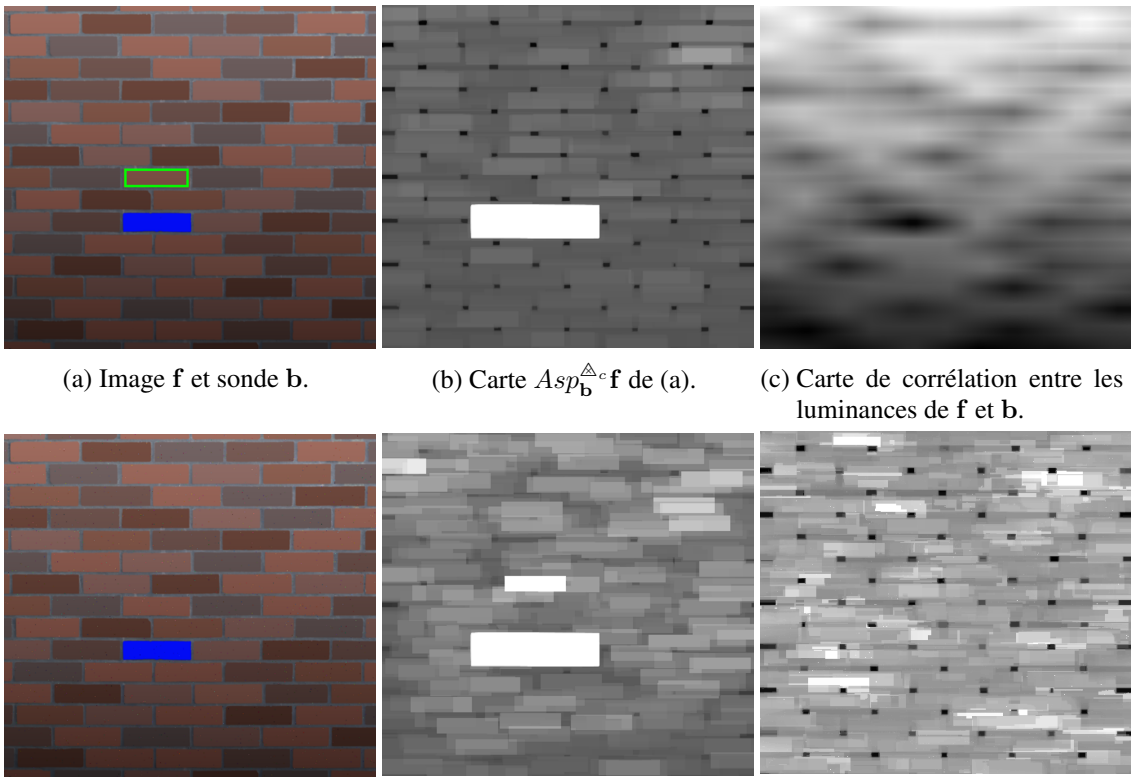
Comme expliqué précédemment (cf. sous-sous-section « Version robuste au bruit », p. 106), la métrique d'Asplund est très sensible au bruit car elle utilise des calculs d'extrema. Il est donc intéressant d'introduire une version robuste au bruit dans le contexte de la « théorie de la mesure ». Sur le même principe que celui utilisé pour les images à niveaux de gris, une métrique d'Asplund $d_{asp,p}^{\triangle_c}(\mathbf{f}, \mathbf{g})$ avec un paramètre de tolérance p peut être définie entre deux images $\mathbf{f}$ et $\mathbf{g}$. Le but est d'écarter les points extrémaux correspondant au bruit. Pour ce faire, un voisinage de la fonction $\mathbf{f}$ s'écrit comme :

$$N_{P,d_{asp}^{\triangle_c},\varepsilon,p}(\mathbf{f}) = \left\{\mathbf{g} \mid \exists D' \subset D, d_{asp}^{\triangle_c}(\mathbf{f}_{|D'}, \mathbf{g}_{|D'}) < \varepsilon \text{ et } P(D \setminus D') \leq p\right\}, \tag{4.101}$$

où ε représente une tolérance sur la distance d'Asplund et $P(D \setminus D') = \frac{\#(D \setminus D')}{\#D}$ est le pourcentage de points conservés. Une carte des distances d'Asplund avec tolérance $Asp_{\mathbf{b},p}^{\triangle_c} : (\mathcal{I}^*)^3 \to (\mathbb{R}^+)^D$ est alors définie pour tout $x \in D$ comme :

$$Asp_{\mathbf{b},p}^{\triangle_c}\mathbf{f}(x) = d_{asp,p}^{\triangle_c}(\mathbf{f}_{|D_b(x)}, \mathbf{b}). \tag{4.102}$$

Sur la figure 4.28, nous recherchons des briques d'un mur similaire à une sonde couleur. Une brique de couleur bleue a été ajoutée dans l'image du mur de brique. Cette image acquise (quasiment) sans bruit présente aussi une dérive d'éclairement selon l'axe vertical (Fig. 4.28a). Sur cette image $\mathbf{f}$, une carte des distances d'Asplund LIPC-multiplicatives $Asp_{\mathbf{b}}^{\triangle_c}\mathbf{f}$ est calculée. Ses minima régionaux (i.e. les points sombres dans la figure 4.28b) correspondent aux centres des briques similaires à la sonde (selon la distance d'Asplund). Le rectangle blanc correspond aux maxima des distances entre la brique bleue et la sonde. La distance est donc sensible à la couleur (i.e. la teinte). Dans l'image avec du bruit $\tilde{\mathbf{f}}$ (Fig. 4.28d), la carte sans tolérance $Asp_{\mathbf{b}}^{\triangle_c}\tilde{\mathbf{f}}$ est plus sensible au bruit (Fig. 4.28e) que la carte avec tolérance $Asp_{\mathbf{b},p}^{\triangle_c}\tilde{\mathbf{f}}$ (Fig. 4.28f). En effet, les minima sont préservés dans la carte avec tolérance (Fig. 4.28f) par comparaison avec la carte sans tolérance (Fig. 4.28e). Les minima peuvent être extraits en utilisant des opérations morphologiques. Surtout, les cartes des distances d'Asplund sont insensibles à la dérive d'éclairement. En outre, une carte de corrélation entre les images de luminance de l'image couleur $\mathbf{f}$ et de la sonde $\mathbf{b}$ ne permet de trouver les positions des briques (Fig. 4.28c).



(a) Image $\mathbf{f}$ et sonde $\mathbf{b}$.

(b) Carte $Asp_{\mathbf{b}}^{\triangle_c}\mathbf{f}$ de (a).

(c) Carte de corrélation entre les luminances de $\mathbf{f}$ et $\mathbf{b}$.

(d) Image $\tilde{\mathbf{f}}$ avec du bruit.

(e) Carte $Asp_{\mathbf{b}}^{\triangle_c}\tilde{\mathbf{f}}$ de (d).

(f) Carte $Asp_{\mathbf{b},p=98\%}^{\triangle_c}\tilde{\mathbf{f}}$ de (d).

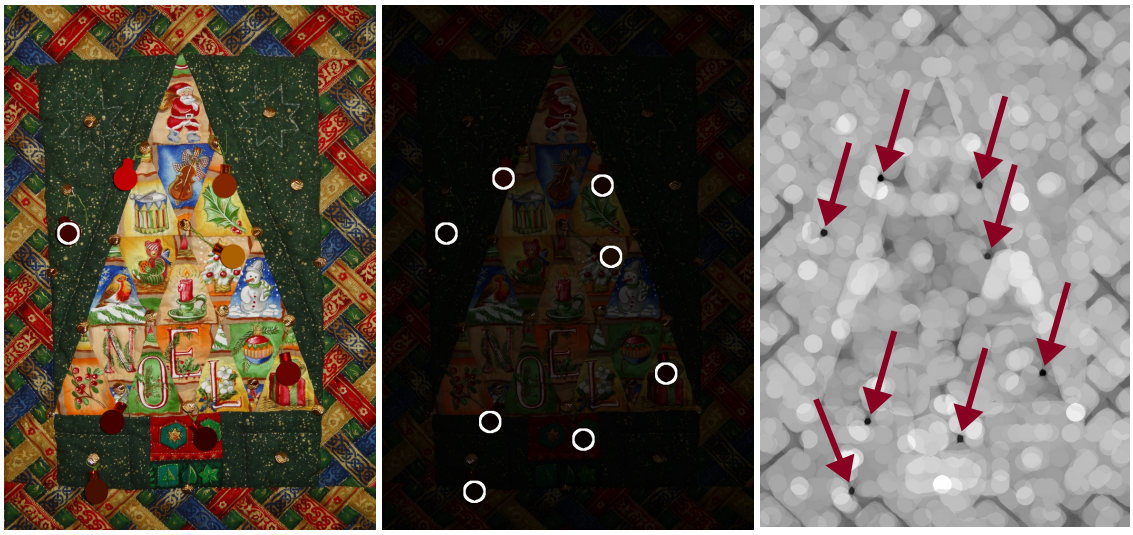
FIG. 4.28 – (a) Image couleur $\mathbf{f}$ acquise sans bruit et sonde couleur $\mathbf{b}$ (entourée de vert). Une brique bleue a été ajoutée à l'image. (b) Carte des distances d'Asplund $Asp_{\mathbf{b}}^{\triangle_c}\tilde{\mathbf{f}}$ entre l'image $\mathbf{f}$ et la sonde $\mathbf{b}$. (c) Carte de corrélation entre les images (à niveaux de gris) de luminance de l'image couleur $\mathbf{f}$ et de la sonde couleur $\mathbf{b}$. (d) Image $\tilde{\mathbf{f}}$ bruitée par ajout sur l'image $\mathbf{f}$ d'un bruit blanc gaussien avec une moyenne de 0, une variance de 2,6 et une densité spatiale de $1\ \%$. (e) Carte des distances d'Asplund sans tolérance $Asp_{\mathbf{b}}^{\triangle_c}\tilde{\mathbf{f}}$ et (f) avec une tolérance $Asp_{\mathbf{b},p}^{\triangle_c}\tilde{\mathbf{f}}$ de $p = 98\ \%$ entre l'image bruitée $\tilde{\mathbf{f}}$ et la sonde $\mathbf{b}$.

La figure 4.29 montre deux images couleur de la même scène, une image claire $\mathbf{f}$ (Fig. 4.29a) et une image sombre $\tilde{\mathbf{f}}$ (Fig. 4.29b). Cette dernière est obtenue par multiplication LIPC par un scalaire $\lambda > 1$ qui simule l'effet sur l'intensité de l'image d'une augmentation d'opacité du milieu entre l'objet et la caméra. La sonde $\mathbf{b}$ représentant une boule de Noël est extraite dans l'image claire et utilisée pour calculer la carte des distances d'Asplund $Asp_{\mathbf{b}}^{\triangle_c}\tilde{\mathbf{f}}$ dans l'image sombre (Fig. 4.29c). Grâce à l'extraction des minima régionaux de la carte, les boules sont détectées et leurs contours sont ajoutés à l'image sombre de la figure 4.29b. On constate que la carte des distances d'Asplund est robuste aux variations d'éclairement correspondant à une variation d'opacité du milieu semi-transparent entre la scène et la caméra.

La figure 4.30 illustre la détection de grains de maïs colorés grâce à la métrique d'Asplund LIPC-multiplicative à partir d'une sonde extraite sur un grain.

La métrique d'Asplund LIPC-multiplicative entre les images couleur est donc efficace pour détecter les formes en couleur acquises sous éclairement variable.

Remarque. De la même manière que pour la métrique d'Asplund LIP-additive pour les images couleur, une métrique d'Asplund LIPC-additive peut être définie. Celle-ci présente la propriété d'être robuste aux variations d'éclairement modélisées par la soustraction LIPC d'une constante et correspondant à une variation de l'intensité de la source ou du temps d'exposition de la caméra. Des expériences complémentaires devront être menées pour vérifier cette propriété.



(a) Image initiale claire $\mathbf{f}$ et sonde $\mathbf{b}$.

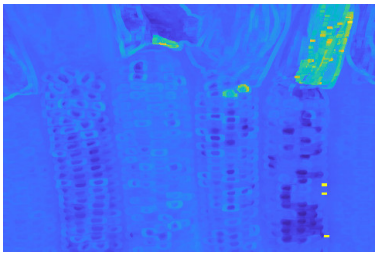
(b) Boules détectées dans l'image sombre $\tilde{\mathbf{f}}$.

(c) Carte $Asp_{\mathbf{b}}^{\triangle_c}\tilde{\mathbf{f}}$.

FIG. 4.29 – Détection de boules de Noël colorées dans (b) une image sombre $\tilde{\mathbf{f}}$ à partir d'une sonde $\mathbf{b}$ extraite dans (a) une image claire $\mathbf{f}$. La délinéation de la sonde $\mathbf{b}$ apparaît en blanc dans (a) de même que les boules détectées dans (b). (c) Carte des distances d'Asplund $Asp_{\mathbf{b}}^{\triangle_c}\tilde{\mathbf{f}}$ entre l'image sombre $\tilde{\mathbf{f}}$ et la sonde $\mathbf{b}$ extraite dans l'image claire. Les minima sont indiqués par des flèches.

(a) Image initiale $\mathbf{f}$ et sonde $\mathbf{b}$.

(b) Sonde couleur $\mathbf{b}$ agrandie.



(c) Carte $Asp_{\mathbf{b},p=95\%}^{\triangle_c}\mathbf{f}$.

(d) Grains détectés dans l'image $\mathbf{f}$.

FIG. 4.30 – Détection de grains de maïs colorés. (a) Dans l'image couleur $\mathbf{f}$ une sonde $\mathbf{b}$ est sélectionnée et entourée en cyan. (b) Sonde couleur $\mathbf{b}$ agrandie. (c) Carte des distances d'Asplund avec une tolérance de 95 % entre l'image $\mathbf{f}$ et la sonde $\mathbf{b}$, $Asp_{\mathbf{b},p=95\%}^{\triangle_c}\mathbf{f}$. (d) Grains de maïs détectés et entourés en bleu.

Les métriques d'Asplund LIP et LIPC couleur ont donc été définies. Dans le futur, il sera intéressant de vérifier leur lien avec la Morphologie Mathématique et leur propriétés optiques. De la même manière que pour les images couleur, il est possible d'envisager de définir un modèle LIP pour les images hyperspectrales ainsi que des métriques hyperspectrales d'Asplund. Celles-ci

seront très importantes en imagerie satellitaire ou microscopique où les conditions d'éclairement peuvent varier.

4.3.9 Application aux images rayons X des pneumatiques

Avant les travaux précédemment présentés sur les cartes des distances d'Asplund, j'ai commencé à m'intéresser à ces métriques en encadrant le stage de Benoît Morel [198], au Centre de Recherche de Michelin, d'avril à septembre 2013. Benoît effectuait son stage de fin d'études pour l'école d'ingénieur CPE-Lyon et pour un Master recherche de l'INSA de Lyon. Son mémoire s'intitulait « Rehaussement de contraste en imagerie Rayons-X des pneumatiques ». Pour son travail, il s'est inspiré de l'article de Jourlin et al. [146].

Contexte du projet

Sur des images du pneumatique acquises en rayons X, Benoît Morel cherchait à extraire des repères collés à sa surface. Ceux-ci étaient constitués de bandes de plomb atténuant fortement les rayons X, ce qui les rendait visibles parmi les câbles métalliques composant les pneumatiques. La figure 4.31 montre un exemple d'image acquise en rayons X sur lesquelles sont visibles des câbles et des repères. Comme ces derniers atténuent l'intensité de la source et gênent la détection des câbles, il est intéressant de les supprimer numériquement sur l'image.

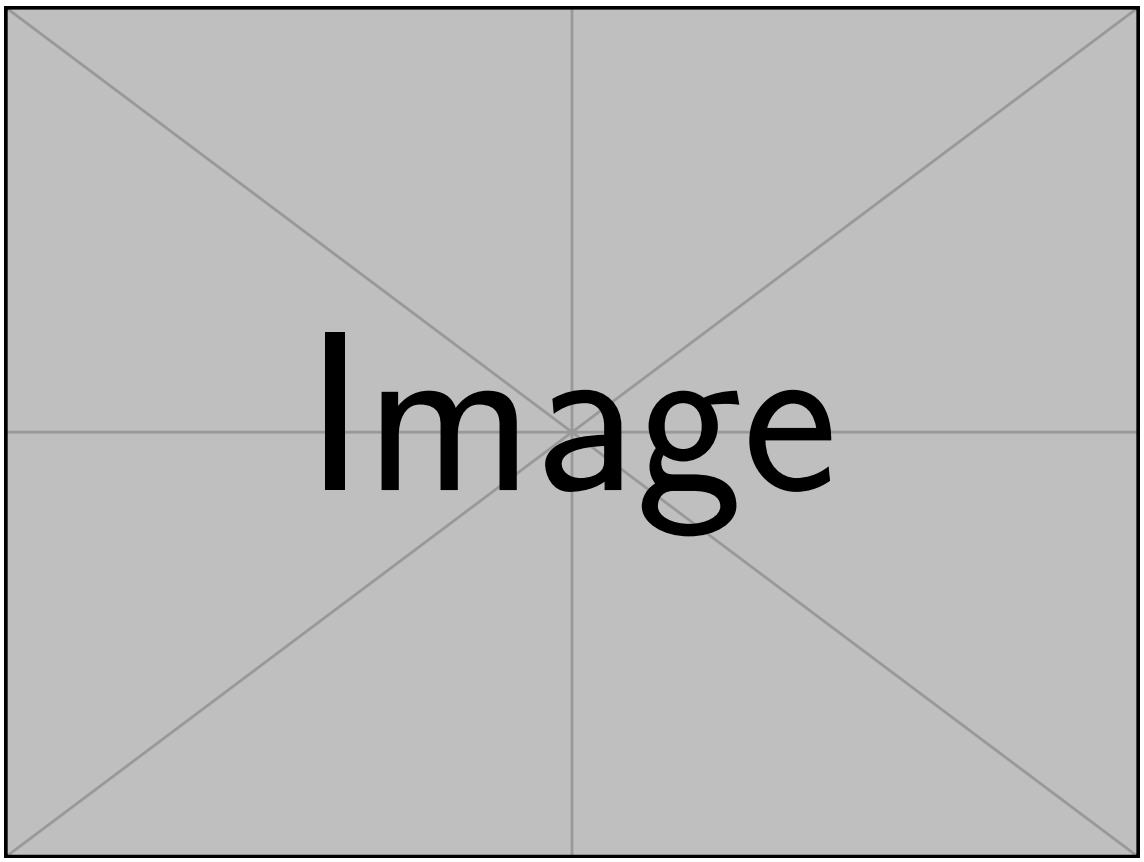
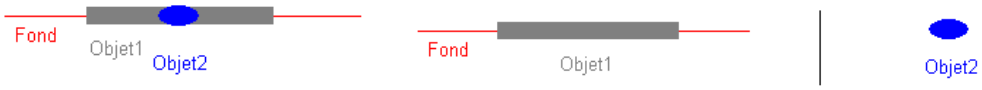


FIG. 4.31 – A gauche : image acquise en rayons X d'un pneumatique avec des repères collés à sa surface. Ceux-ci apparaissent comme des bandes plus sombres que leur voisinage. A droite : grossissement sur une partie de l'image acquise en rayons X.

Dans ce but, on va venir extraire leur forme puis rehausser le contraste entre l'objet et son voisinage par soustraction LIP d'une constante. Ceci est rendu possible parce que le principe de superposition d'objets semi-transparents acquis par transmission est modélisé par le modèle LIP

(Fig. 4.32). Soient c_1 et f_2 les images de deux objets semi-transparents acquis en transmission (e.g. par rayons X). L'image de leur superposition f_s est égale à la somme LIP des images c_1 et f_2 : $f_s = c_1 \triangleup f_2$. L'image f_2 est ainsi égale à la différence LIP entre f_s et c_1 : $f_2 = f_s \triangle c_1$.

La figure 4.33a montre une image d'un repère en forme de croix superposé avec les câbles métalliques de la nappe carcasse d'un pneumatique. L'image du repère est filtrée par une fermeture morphologique et segmentée par seuillage (Fig. 4.33b). La différence LIP (i.e. le contraste) entre un point de la croix et un point du fond blanc est calculée pour obtenir la constante LIP, c_1. Cette constante c_1 est soustraite au sens LIP de l'image initiale f_s sur la zone segmentée, $f_2 = f_s \triangle c_1$ (Fig. 4.33c). Au croisement des deux bras de la croix, la constante c_1 est deux fois LIP-soustraite de l'image f_s : $f_2 = f_s \triangle (c_1 \triangleup c_1)$ (Fig. 4.33d).



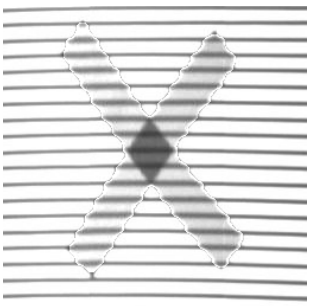
(a) Superposition des objets 1 et 2.

(b) Objets 1 et 2 générant les images c_1 et f_2.

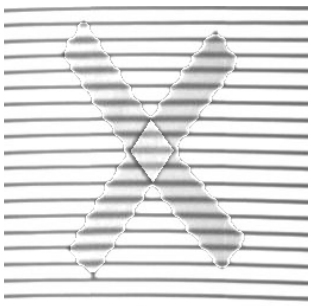
FIG. 4.32 – (a) Superposition des objets 1 et 2 pour obtenir l'image f_s acquise en transmission. (b) Objets 1 et 2 générant les images f_1 et f_2 acquises séparément. Grâce au modèle LIP, l'image finale f_s correspond à la somme LIP des images c_1 et f_2 : $f_s = c_1 \triangleup f_2$.

(a) Image du repère.

(b) Masque du repère.



(c) Rehaussement de contraste sur les bras de la croix.



(d) Rehaussement de contraste sur les bras et au centre de la croix.

FIG. 4.33 – (a) Repère en forme de croix et les deux points sélectionnés (en rouge) pour déterminer la constante c_1. (b) Segmentation de la croix. (c) Soustraction LIP d'une constante c_1 sur le masque. (d) Soustraction au sens LIP d'une constante c_1 sur le masque. Soustraction au sens LIP de deux fois la constante (i.e. $c_1 \triangleup c_1$) à l'intersection des bras de la croix.

Seuillage de Köhler avec un contraste LIP

Le modèle LIP permet de calculer un contraste additif entre deux niveaux de gris $f(x)$ et $f(y)$ d'une image $f : D \to [0, M[$. Ce contraste, noté $C^{\triangleup}_{(x,y)(f)}$, est appelée LAC (*Logarithmic Additive*

Contrast) et a été introduit par Jourlin et al. [149]. Il est défini par :

$$\min(f(x), f(y)) \triangle C^{\triangle}_{(x,y)(f)} = \max(f(x), f(y)). \tag{4.103}$$

Remarque. Interprétation optique. Ce contraste représente le niveau de gris qui doit être ajouté (i.e. superposé) au point le plus clair (i.e. de plus faible niveau de gris dans l'échelle LIP) pour obtenir le plus sombre (i.e. de plus haut niveau de gris). Ce contraste correspond alors à un niveau de gris de l'intervalle $[0, M[$.

Le contraste LAC peut alors s'exprimer comme :

$$C^{\triangle}_{(x,y)(f)} = \max(f(x), f(y)) \triangle \min(f(x), f(y)) \tag{4.104}$$

$$= \frac{|f(x) - f(y)|}{1 - \frac{\min(f(x), f(y))}{M}}. \tag{4.105}$$

Le lecteur intéressé par les propriétés du LAC et son lien avec le contraste de Michelson pourra se référer à [141], [146].

La méthode de seuillage de Köhler, qui a été présentée dans la sous-sous-section « Accélération de la méthode de Köhler sur le seuillage par contraste » (p. 64), nécessite de calculer un contraste entre un couple de pixels (x_0, x_1) de la frontière $B(t)$ qui a été obtenue par seuil à la valeur t de l'image f. Le contraste LIP de Köhler entre les points de ce couple $C^{\triangle,t}_K(x_0, x_1)$ est défini comme [146] :

$$C^{\triangle,t}_K(x_0, x_1) = \min\left(f(x_1) \triangle t, t \triangle f(x_0)\right). \tag{4.106}$$

Le contraste moyen LIP-additif de la frontière $B(t)$ est alors égal à :

$$C^{\triangle}_K(B(t)) = \frac{1}{card\{B(t)\}} \triangle \underset{(x_0,x_1)\in B(t)}{\triangle} C^{\triangle,t}_K(x_0, x_1). \tag{4.107}$$

La figure 4.34 représente la courbe des contrastes de Köhler pour un contraste linéaire et un contraste LIP. Afin d'obtenir une image multi-seuillée, 8 seuils sont sélectionnés sur la courbe de contrastes LIP, à l'emplacement des maxima régionaux. Une image seuillée en 9 classes est alors obtenue sur laquelle on distingue bien le repère (Fig. 4.35).

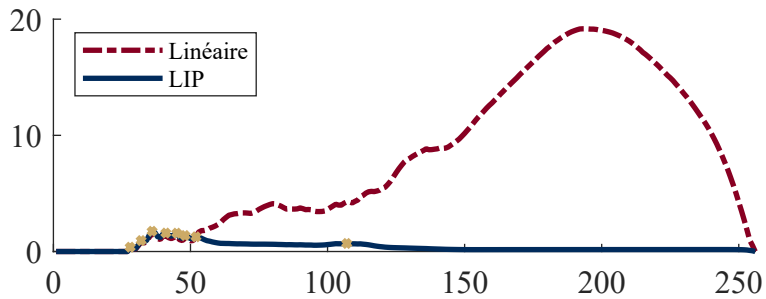


FIG. 4.34 – Courbes des contrastes de Köhler pour le contraste linéaire (en rouge) et pour le contraste LIP (en bleu) calculées sur l'image du repère (Fig. 4.33a). Les 8 maxima régionaux sont indiqués par des points de couleur beige.

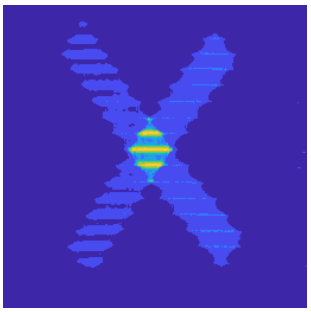


FIG. 4.35 – Résultat du multi-seuillage en 8 seuils correspondant aux 8 pics les plus significatifs sur la courbe de contraste LIP de Köhler LIP de la figure 4.34. L'image compte donc 9 classes.

Détection par carte des distances d'Asplund LIP-multiplicatives

En imagerie rayons X des pneumatiques, plusieurs nappes métalliques peuvent être superposées (Fig. 4.36a). La détection d'un repère s'avère alors plus difficile. L'image f d'un repère peut être extraite grâce à une carte des distances d'Asplund LIP-multiplicatives qui est insensible aux variations de contrastes causées par une variation d'opacité du milieu semi-transparent. Pour ce faire on utilise une sonde cylindrique b de hauteur égale à 1 niveau de gris. Quatre essais ont été réalisés :

i) pour un rayon du cylindre de $r = 3$ pixels, $Asp^{\triangle}_{b_{r=3}} f$ (Fig. 4.36b),

ii) pour un rayon de $r = 5$ pixels, $Asp^{\triangle}_{b_{r=5}} f$ (Fig. 4.36c),

iii) pour un rayon de $r = 15$ pixels, $Asp^{\triangle}_{b_{r=15}} f$ (Fig. 4.36d) et

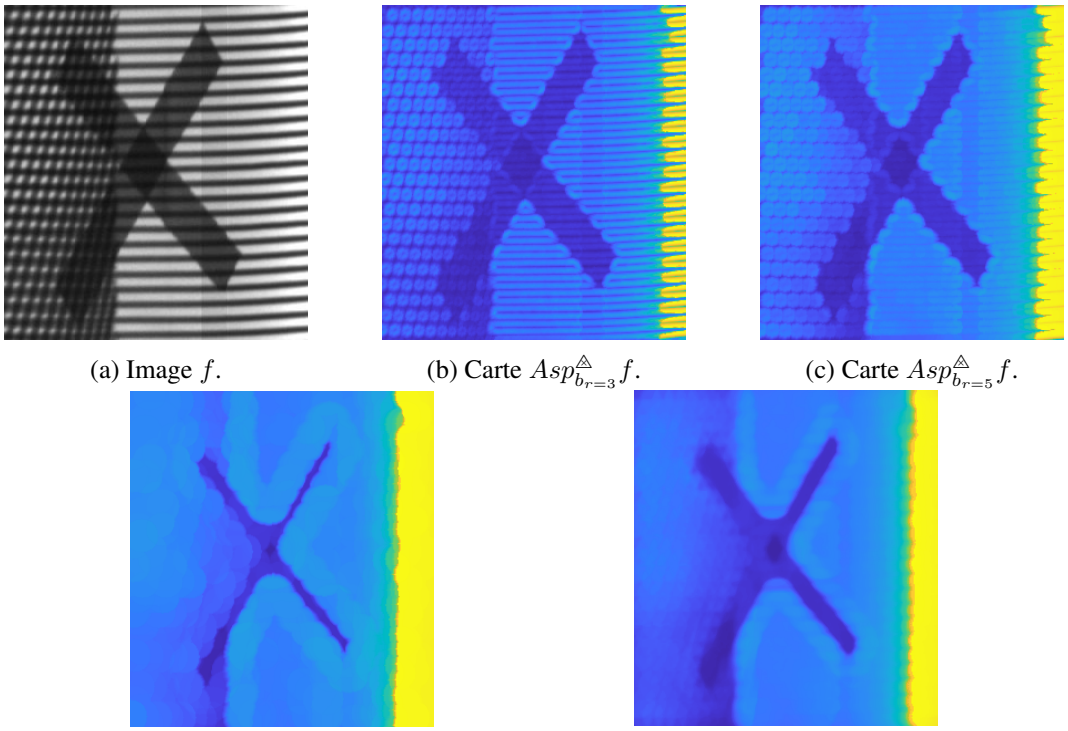
iv) pour un rayon de $r = 15$ pixels avec une tolérance de $p = 90$ % sur les points de contacts conservés, $Asp^{\triangle}_{b_{r=15},p=90\%} f$ (Fig. 4.36e).

On constate que l'utilisation de cartes des distances d'Asplund LIP-multiplicatives permet de détecter la croix malgré la superposition des nappes métalliques. Cette méthode s'avère donc efficace en imagerie rayons X des pneumatiques.

Benoît Morel a aussi réalisé une mise en œuvre informatique :

- du seuillage de Köhler en langage C++ et
- des cartes des distances d'Asplund LIP-multiplicatives avec la librairie NT2 [77]. Celle-ci permet un prototypage rapide depuis *MATLAB*® avec des temps de calcul plus courts que ceux de *MATLAB*®.

Les résultats obtenus étaient plus rapide que l'implémentation en *MATLAB*®. Cependant, la méthode rapide de calcul du seuillage de Köhler que j'ai introduite dans [207] (cf. sous-sous-section « Accélération de la méthode de Köhler sur le seuillage par contraste », p. 64), s'avère être d'une complexité plus faible. Il en est de même pour la mise en œuvre des carte des distances d'Asplund à partir d'opérations de Morphologie Mathématique efficacement programmées (cf. sous-section 4.3.3, p. 105).



(a) Image f. (b) Carte $Asp^{\triangle}_{b_{r=3}} f$. (c) Carte $Asp^{\triangle}_{b_{r=5}} f$.

(d) Carte $Asp^{\triangle}_{b_{r=15}} f$. (e) Carte $Asp^{\triangle}_{b_{r=15},p=90\%} f$.

FIG. 4.36 – (a) Image f d'un repère en forme de croix situé au niveau de la superposition de deux nappes métalliques d'un pneumatique. Cartes des distances d'Asplund LIP-multiplicatives pour une sonde cylindrique b (b) de rayon $r = 3$ pixels, $Asp^{\triangle}_{b_{r=3}} f$, (c) de rayon $r = 5$ pixels, $Asp^{\triangle}_{b_{r=5}} f$, (d) de rayon $r = 15$ pixels, $Asp^{\triangle}_{b_{r=15}} f$, et (e) de rayon $r = 15$ pixels avec une tolérance de $p = 90$ %, $Asp^{\triangle}_{b_{r=15},p=90\%} f$.

4.4 Morphologie Mathématique Logarithmique

Dans [217], j'ai introduit un nouveau cadre de travail appelé *Logarithmic Mathematical Morphology* (LMM) dont les opérateurs sont adaptatifs aux variations d'intensité et de temps d'exposition de la caméra.

4.4.1 Motivation

L'extension de la Morphologie Mathématique (MM) aux fonctions inclut les images à niveaux de gris dont les valeurs sont dans l'intervalle borné $[0; M[$ de l'ensemble des réels $\mathbb{R}$. Généralement, les fonctions structurantes et les opérateurs morphologiques sont invariants aux translations horizontales (i.e. dans l'espace du domaine de définition) et aux translations verticales (i.e. en intensité) [111]. Cependant, l'application de la MM aux images à niveaux de gris présente deux limitations.

1) Premièrement, bien que les images à niveaux de gris aient des valeurs majorées par M, la MM fonctionnelle a été définie pour des fonctions à valeurs réelles [112], [113]. En effet, étant données une image f et une fonction structurante b, toutes deux définies sur le même domaine $D \subset \mathbb{R}^n$ et dont les valeurs sont dans la gamme $[0; M[$, leur somme $f + b$ n'appartient pas à l'intervalle $[0; M[$. Les solutions pratiques à ce problème consistent à utiliser soit : (i) une fonction structurante dont le supremum est égal à 0, soit (ii) un élément structurant plat dont les valeurs sont égales à 0, soit (iii) de tronquer les valeurs de l'image résultante [112].

2) Deuxièmement, l'ajout d'une fonction structurante à une image sans tenir compte de son intensité dans l'amplitude de la fonction structurante n'est pas optiquement justifié. Dans la vision

humaine, la réponse de l'œil aux variations d'intensité est connue pour être de type logarithmique [37], [138], [150], [296], [305]. Il s'ensuit donc que les variations de contraste dans les images sont aussi logarithmiques et que les variations les plus sombres sont plus atténuées que les plus claires. L'amplitude de la fonction structurante doit donc dépendre de l'intensité de l'image, i.e. du niveau de gris. Une telle fonction structurante sera seulement invariante sous translation horizontale mais pas sous translation verticale dans le domaine de l'intensité $[0; M[$.

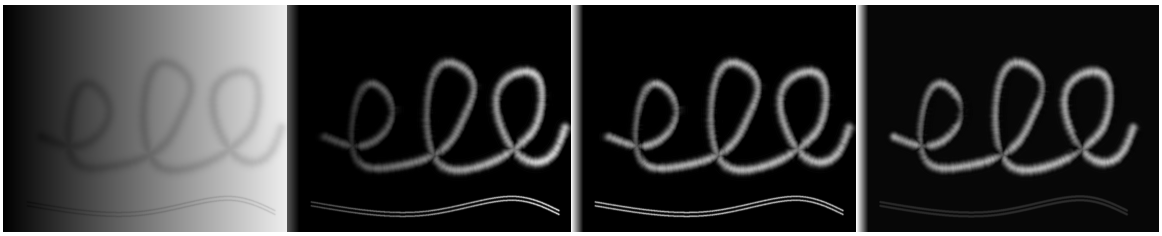
En parallèle de la genèse de la MM à niveaux de gris, Jourlin a développé le modèle de Traitement d'Image Logarithmique (LIP) qui est adapté à la vision humaine [138], [143], [150] et qui permet de traiter les images comme un œil humain le ferait. Mathématiquement bien défini, il possède de fortes propriétés optiques. En particulier, l'addition-LIP $\triangle\!\!\!+$ de deux images donne une image. Il permet aussi, dans les images, de simuler les variations d'intensité lumineuse ou du temps d'exposition de la caméra.

Le but de cette section est de traiter les deux limitations précédemment listées de l'application de la MM aux images à niveaux de gris en présentant le cadre de travail appelé *Morphologie Mathématique Logarithmique* (ou LMM pour *Logarithmic Mathematical Morphology*) [217]. La LMM est adaptée au traitement des images acquises en lumière réfléchie. Elle permet d'ajuster l'amplitude de la fonction structurante selon l'intensité de l'image grâce à l'addition-LIP $\triangle\!\!\!+$ entre l'image et la fonction structurante.

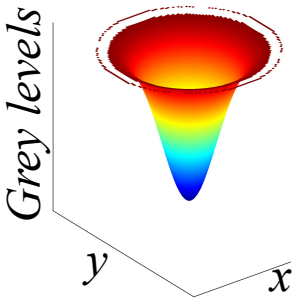
4.4.2 Travaux connexes : opérateurs de détection robustes aux variations d'éclairement

Précédemment, des tentatives ont eu lieu dans la littérature pour créer des opérateurs qui sont théoriquement robustes aux changements d'éclairement. Cependant, ces opérateurs ne tiennent généralement pas compte des causes optiques de ces changements d'éclairement telles que des variations d'intensité lumineuses ou de manière équivalente des variations du temps d'exposition de la caméra. En LMM, ces causes sont modélisées par la loi LIP-additive $\triangle\!\!\!+$ [136], [137]. La figure 4.37 montre une image composée d'une spirale, d'une dérive d'éclairement et de deux courbes confondantes (Fig. 4.37a). Sur cette image, les opérations de LMM détectent mieux la spirale (Fig. 4.37d) que les méthodes usuelles avec un *top-hat* (Fig. 4.37b) ou un *top-hat-logarithmique* (Fig. 4.37c).

Soit $f : D \to \overline{\mathbb{R}}$ une fonction définie sur un domaine $D \subset \mathbb{R}^n$ dont les valeurs sont dans l'ensemble $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Soit $b : D \to \overline{\mathbb{R}}$ une fonction structurante. En dehors du domaine $D_b \subset D$, toutes ses valeurs sont égales à $-\infty$, $\forall x \in D \setminus D_b$, $b(x) = -\infty$. De plus, si ses valeurs dans D_b sont égales à 0, la fonction structurante est appelée un élément structurant (plat) et est notée par la lettre majuscule B.



(a) Image d'entrée. (b) *Top-hat*. (c) *LIP-top-hat*. (d) LMM.



(e) Sonde.

FIG. 4.37 – Comparaison entre des méthodes usuelles et la LMM pour détecter une spirale dans (a) une image simulée contenant aussi deux courbes rapprochées et une dérive d'éclairement. (b) Le *top-hat* avec un disque plat extrait la spirale et les deux courbes mais est sensible à la dérive d'éclairement. (c) Le *LIP-top-hat* avec un disque plat extrait la spirale et les deux courbes sans la dérive d'éclairement. (d) Les opérations de LMM consistent en une différence-LIP entre une ouverture-LIP par (e) une sonde composée d'un anneau contenant une gaussienne et une ouverture-LIP par le même anneau. Les opérations de LMM extraient avec succès la spirale et atténuent fortement les deux courbes rapprochées.

Opérateurs de top-hat et gradient morphologique

Meyer [186], [275] créa l'opérateur de *top-hat* $TH_B(f)$ pour détecter les pics dans une fonction $f : D \to \mathbb{R}$. Il est égal à la différence entre la fonction et son ouvert morphologique $\gamma_B(f)$ par un élément structurant plat B, $TH_B(f) = f - \gamma_B(f)$. L'opération complémentaire pour la détection des vallées est le *top-hat-noir* $BTH_B(f)$ défini comme la différence entre la fermeture morphologique $\phi_B(f)$ de la fonction et la fonction elle-même, $BTH_B(f) = \phi_B(f) - f$ (cf. sous-section 3.1.2, p. 23). Beucher [32] a défini le gradient morphologique $\varrho_B(f)$ comme la différence entre la dilatation $\delta_B(f)$ et l'érosion $\varepsilon_B(f)$ d'une fonction f par un élément structurant plat B, $\varrho_B(f) = \delta_B(f) - \varepsilon_B(f)$. Le gradient morphologique correspond en fait à la norme du gradient usuel d'une fonction [28]. Afin de respecter cette propriété, il doit être défini avec un élément structurant plat.

Les deux *top-hats* et le gradient morphologique sont invariants aux variations artificielles d'intensité causées par l'addition ou la soustraction de toute constante réelle c à une fonction f : $TH_B(f + c) = TH_B(f)$, $BTH_B(f + c) = BTH_B(f)$ et $\varrho_B(f + c) = \varrho_B(f)$. Cependant, ils ne sont pas invariants à toute variation d'éclairement avec une cause optique modélisée par une addition-LIP $\triangleup$ ou une soustraction-LIP $\triangle$ d'une constante à une image $f : D \to [0; M[$. Afin de résoudre ce problème, Jourlin et al. [135], [143] ont introduit des *LIP-top-hats* et un gradient morphologique-LIP où la différence-LIP $\triangle$ remplace la différence habituelle « $-$ ». Zaharescu [330] a proposé des variantes des *LIP-top-hats*. Cependant, ces *LIP-top-hats* et gradient morphologique-LIP sont toujours définis avec un élément structurant plat.

Scale Invariant Feature Transform (SIFT)

Lowe a introduit le *Scale Invariant Feature Transform* pour détecter des caractéristiques des images qui sont « partiellement invariantes à un changement d'illumination » [166]. Dans le SIFT, les points saillants sont tout d'abord détectés comme des extrema d'un espace d'échelle de fonctions obtenu grâce à des différences entre des filtrages gaussiens de l'image. Un descripteur local de l'image est associé à chaque point saillant. Ce descripteur d'image est fondé sur l'histogramme d'orientation du gradient de l'image filtrée par une gaussienne. Comme le gradient est calculé par des différences entre les valeurs de l'image, cela le rend insensible aux changements d'illumination causés par l'addition d'une constante. L'orientation est pondérée par la norme du gradient qui est égale au gradient morphologique et a donc la même invariance. En outre, une normalisation entre 0 et 1 des histogrammes d'orientation rend le descripteur SIFT invariant à la multiplication de l'image par une constante. Cependant, ces invariances n'ont pas de cause optique.

Arbres des composantes connexes

Les arbres de composantes connexes sont fondés sur les *ensembles de niveau*. Un *ensemble de niveau* d'une image est l'ensemble de ses pixels dont les valeurs sont supérieures ou égales à une valeur donnée de seuil. En augmentant la valeur de seuil, les composantes connexes d'un plus haut niveau sont incluses dans la composante connexe d'un plus bas niveau. Ces relations d'inclusion peuvent être représentées par des arbres des composantes connexes. Plusieurs types d'arbres peuvent être construits selon les relations d'inclusion entre les ensembles de niveaux. On peut citer e.g. l'*arbre des composantes*, appelé aussi *max-tree* [267] ou l'*arbre des formes*, appelé aussi *arbre d'inclusion* [191]. Des méthodes récentes de segmentation par des arbres de composantes connexes ont été présentées dans [242], [322]. Les arbres des composantes d'une fonction réelle sont théoriquement invariants aux changements d'intensité causés par l'application à la fonction d'une transformée continue et croissante. Cependant, comme l'intensité d'une image est quantifiée en valeurs discrètes de l'ensemble $[\![0, 1, \ldots, M-1]\!]$ avec un pas constant, les intensités faibles sont mal représentées par ces valeurs discrètes. Pour cette raison, dans les images faiblement éclairées, les arbres des composantes connexes peuvent présenter une robustesse limitée aux variations d'éclairement [230].

Morphologie Mathématique dépendant de l'intensité

Heijmans [111] a nommé les opérateurs morphologiques usuels *t-opérateurs* à cause de leur invariance par translation horizontale (i.e. dans l'espace de leur domaine de définition) et par translation verticale (i.e. en intensité). Il a aussi défini une classe d'opérateurs morphologiques, les *h-opérateurs* qui sont invariants par translation *horizontale* mais adaptatif dans le domaine vertical (i.e. en intensité). Bien que les exemples de *h-opérateurs* donnés dans [111] soient mathématiquement bien définis pour des fonctions réelles, ils ne sont pas justifiés d'un point de vue optique. Les opérateurs de LMM constituent un cas particulier de *h-opérateurs* mais qui sont définis pour des fonctions avec des valeurs majorées par M et qui sont justifiés d'un point de vue optique.

4.4.3 Les opérateurs de Morphologie Mathématique Logarithmique

Le nouveau cadre de travail

La LMM est définie sur le treillis $\overline{\mathcal{F}}_M = [-\infty; M]^D$ des fonctions à valeurs dans $[-\infty; M]$ avec l'ordre usuel $\leq$. L'infimum $\bigwedge$ et le supremum $\bigvee$ sont respectivement définis pour toute famille

$\mathscr{X} \subset \overline{\mathcal{F}}_M$ par :

$$\left(\bigwedge \mathscr{X}\right)(x) = \wedge \{f(x) \in [-\infty; M] \mid f \in \mathscr{X},\ x \in D\}, \tag{4.108}$$

$$\left(\bigvee \mathscr{X}\right)(x) = \vee \{f(x) \in [-\infty; M] \mid f \in \mathscr{X},\ x \in D\}. \tag{4.109}$$

Le plus petit et le plus grand élément, O et I, sont les fonctions constantes égales pour tout $x \in D$ à $O(x) = -\infty$ et $I(x) = M$, respectivement. La loi LIP-additive $\triangle\!\!\!\!+$ et la loi LIP-négative $\triangle$ permettront d'effectuer des transformations morphologiques qui sont compatibles avec le système visuel humain. La LMM repose sur les opérateurs adjoints d'érosion et de dilatation qui sont introduits comme il s'ensuit.

Soient $f : D \to [-\infty; M]$ une fonction et $b : D \to [-\infty; M]$ une fonction structurante dont les valeurs en dehors du domaine $D_b \subset D$ sont égales à $-\infty$. Soient les applications $\delta_b^{\triangle\!\!\!\!+} : \overline{\mathcal{F}}_M \to \overline{\mathcal{F}}_M$ et $\varepsilon_b^{\triangle\!\!\!\!+} : \overline{\mathcal{F}}_M \to \overline{\mathcal{F}}_M$ définies par :

$$\delta_b^{\triangle\!\!\!\!+}(f)(x) = \bigvee \{f(x-h) \triangle\!\!\!\!+\, b(h), h \in D\}, \tag{4.110}$$

$$\varepsilon_b^{\triangle\!\!\!\!+}(f)(x) = \bigwedge \{f(x+h) \triangle\, b(h), h \in D\}. \tag{4.111}$$

Dans le cas d'expressions indéterminées, les conventions suivantes sont utilisées : $f(x-h) \triangle\!\!\!\!+\, b(h) = -\infty$ lorsque $f(x-h) = -\infty$ ou $b(h) = -\infty$, et $f(x+h) \triangle\, b(h) = M$ lorsque $f(x+h) = M$ ou $b(h) = -\infty$.

Proposition 12. *Le couple d'applications* $(\varepsilon_b^{\triangle\!\!\!\!+}, \delta_b^{\triangle\!\!\!\!+})$ *constitue une* adjonction*, où* $\varepsilon_b^{\triangle\!\!\!\!+}$ *est une* érosion *et* $\delta_b^{\triangle\!\!\!\!+}$ *une* dilatation*.*

Définition 27. $\varepsilon_b^{\triangle\!\!\!\!+}$ *est appelée* érosion-logarithmique *et* $\delta_b^{\triangle\!\!\!\!+}$ dilatation-logarithmique

Démonstration de la proposition 12. • Démontrons que $\delta_b^{\triangle\!\!\!\!+}$ est une dilatation (cf. def. 2, p. 23).

Comme la loi $\triangle\!\!\!\!+$ préserve l'ordre $\leq$ [143], nous avons $\forall f, g \in \overline{\mathcal{F}}_M, \forall x \in D$,

$$\begin{aligned}
\delta_b^{\triangle\!\!\!\!+}(f \vee g)(x) &= \vee_{h\in D} \{((f \vee g)(x-h)) \triangle\!\!\!\!+\, b(h)\} \\
&= \vee_{h\in D} \{(f(x-h) \triangle\!\!\!\!+\, b(h)) \vee (g(x-h) \triangle\!\!\!\!+\, b(h))\} \\
&= [\vee_{h\in D} \{f(x-h) \triangle\!\!\!\!+\, b(h)\}] \vee [\vee_{h\in D} \{g(x-h) \triangle\!\!\!\!+\, b(h)\}] \\
&= \delta_b^{\triangle\!\!\!\!+}(f)(x) \vee \delta_b^{\triangle\!\!\!\!+}(g)(x).
\end{aligned}$$

Pour le plus petit élément O de $\overline{\mathcal{F}}_M$, nous avons $\forall x \in D$:

$$\begin{aligned}
\delta_b^{\triangle\!\!\!\!+}(O)(x) = \delta_b^{\triangle\!\!\!\!+}(f_{-\infty})(x) &= \vee_{h\in D}\{(-\infty(x-h) \triangle\!\!\!\!+\, b(h))\} \\
&= \vee_{h\in D}\{-\infty(1 - b(h)/M) + b(h)\} = -\infty \\
&= O(x).
\end{aligned}$$

• De façon similaire, nous avons $\forall f, g \in \overline{\mathcal{F}}_M$,

$$\begin{aligned}
\varepsilon_b^{\triangle\!\!\!\!+}(f \wedge g) = \varepsilon_b^{\triangle\!\!\!\!+}(f) \wedge \varepsilon_b^{\triangle\!\!\!\!+}(g) \ \text{ et} \\
\varepsilon_b^{\triangle\!\!\!\!+}(I) = \varepsilon_b^{\triangle\!\!\!\!+}(f_M) = M = I.
\end{aligned}$$

Par conséquent, $\varepsilon_b^{\triangle\!\!\!\!+}$ est une érosion.

• Le couple $(\varepsilon_b^{\triangle\!\!\!\!+}, \delta_b^{\triangle\!\!\!\!+})$ forme une adjonction car, $\forall f, g \in \overline{\mathcal{F}}_M$, nous avons :

$$\begin{aligned}
\delta_b^{\triangle\!\!\!\!+}(f) \leq g &\iff \forall x \in D,\ \vee_{h\in D}\{f(x-h) \triangle\!\!\!\!+\, b(h)\} \leq g(x) \\
&\iff \forall x \in D, \forall h \in D,\ f(x-h) \triangle\!\!\!\!+\, b(h) \leq g(x) \\
&\iff \forall y \in D, \forall h \in D,\ f(y) \leq g(y+h) \triangle\, b(h) \\
&\iff \forall y \in D,\ f(y) \leq \wedge_{h\in D}\{g(y+h) \triangle\, b(h)\} \\
&\iff f \leq \varepsilon_b^{\triangle\!\!\!\!+}(g)
\end{aligned}$$

□

Comme le couple $(\varepsilon_b^{\triangle}, \delta_b^{\triangle})$ est une adjonction, une *ouverture* et une *fermeture* peuvent être définies par combinaison des deux opérateurs d'*érosion-logarithmique* $\varepsilon_b^{\triangle}$ et de *dilatation-logarithmique* $\delta_b^{\triangle}$ (cf. sous-section 3.1.2, p. 23). Les opérateurs $\gamma_b^{\triangle}$ et $\varphi_b^{\triangle}$ définis par

$$\gamma_b^{\triangle} = \delta_b^{\triangle}\varepsilon_b^{\triangle}, \tag{4.112}$$

$$\varphi_b^{\triangle} = \varepsilon_b^{\triangle}\delta_b^{\triangle} \tag{4.113}$$

sont respectivement une ouverture et une fermeture (par adjonction).

Définition 28. $\gamma_b^{\triangle}$ *est appelée* ouverture-logarithmique *et* $\varphi_b^{\triangle}$ fermeture-logarithmique.

Une autre propriété intéressante en MM est que l'érosion d'une fonction est égale à la dilatation de sa *fonction négative*, et vice versa. Cette propriété constitue la dualité par *fonction négative* des deux opérateurs et, pour la LMM, elle est établie dans la proposition 13. La *fonction négative* f^* de f est égale à $f^* = \triangle f$ parce que nous avons $(f^*)^* = f$ [113].

Proposition 13. *Soit* $\overline{b} \in \overline{\mathcal{F}}_M$ *la fonction structurante réfléchie de* b*, où* $\forall x \in D$, $\overline{b}(x) = b(-x)$. *L'érosion-logarithmique* $\varepsilon_b^{\triangle}$ *et la dilatation-logarithmique* $\delta_b^{\triangle}$ *sont duales par leur* fonction négative *:*

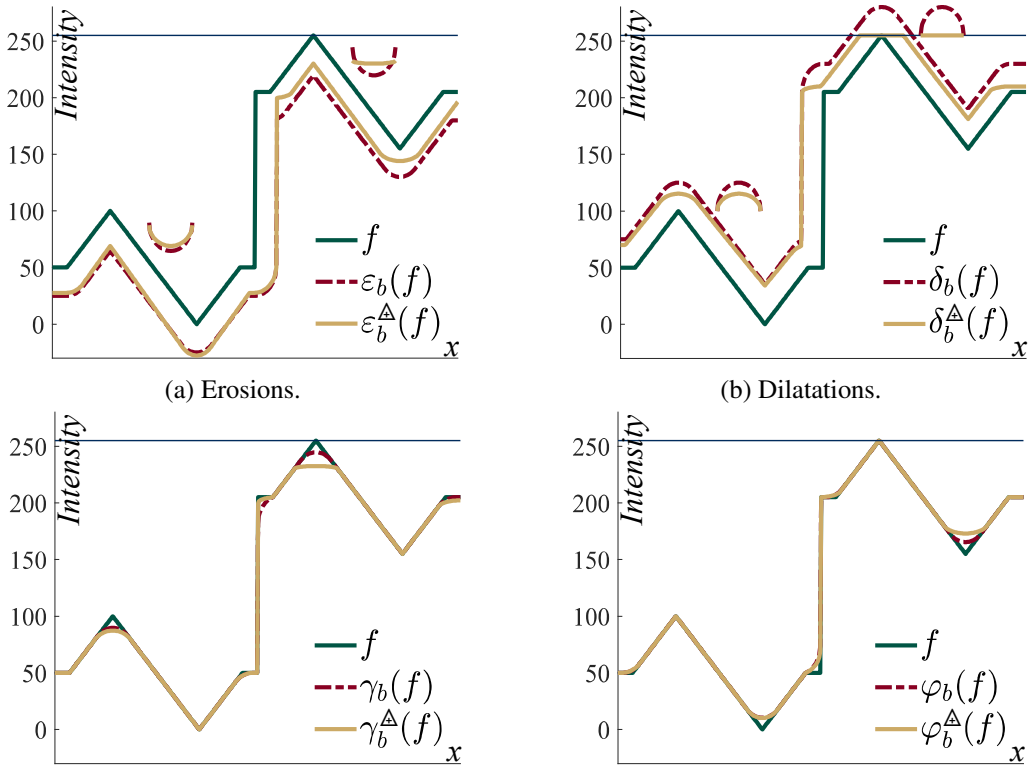
$$(\delta_b^{\triangle}(f^*))^* = \varepsilon_{\overline{b}}^{\triangle}(f) \quad \textit{et} \quad (\varepsilon_b^{\triangle}(f^*))^* = \delta_{\overline{b}}^{\triangle}(f). \tag{4.114}$$

Démonstration. L'érosion $\varepsilon_b^{\triangle}$ et la dilatation logarithmiques $\delta_b^{\triangle}$ sont duales par leur *fonction négative* parce que pour $b \in \overline{\mathcal{F}}_M$ nous avons, $\forall f, g \in \overline{\mathcal{F}}_M$, $\forall x \in D$,

$$\begin{aligned}
(\delta_b^{\triangle}(f^*))^*(x) &= \triangle(\vee_{h\in D}\{\triangle f(x-h) \triangle b(h)\}) \\
&= \wedge_{h\in D}\{f(x-h) \triangle b(h)\} \\
&= \wedge_{h\in D}\{f(x+h) \triangle \overline{b}(h)\} \\
&= \varepsilon_{\overline{b}}^{\triangle}(f)(x).
\end{aligned}$$

De façon similaire, nous avons $(\varepsilon_b^{\triangle}(f^*))^* = \delta_{\overline{b}}^{\triangle}(f)$. □

Sur le signal de la figure 4.38, les opérateurs de MM fonctionnelle (i.e. usuelle) sont comparés à ceux de LMM. Un demi-disque sert de fonction structurante (fs) b. En LMM, l'amplitude de la fs change selon les valeurs de l'image (Fig. 4.38a et 4.38b). Ceci est causé par les lois LIP, $\triangle$ ou $\triangle$, utilisées dans les équations (4.110) et (4.111). Les opérateurs de LMM sont donc des *h-opérateurs* qui sont uniquement invariants par translation horizontale (cf. sous-sous-section « Morphologie Mathématique dépendant de l'intensité », p. 139). Cependant, en MM fonctionnelle, l'amplitude de la fs reste la même. Cela génère des *t-opérateurs* qui sont invariants par translations horizontales et verticales. De plus, sur la figure 4.38b, la dilatation-logarithmique $\delta_b^{\triangle}(f)$ de f est toujours en dessous de la borne supérieure $M = 256$, tandis que la dilatation fonctionnelle $\delta_b(f)$ de f peut dépasser cette borne. Cette propriété est due à l'addition LIP $\triangle$. Sur la figure 4.38a, les valeurs négatives de l'érosion fonctionnelle $\varepsilon_b(f)$ n'ont pas de justification physique, tandis que celles de l'érosion $\varepsilon_b^{\triangle}(f)$ correspondent à des intensificateurs de lumière. Sur la figure 4.38c, la différence entre l'ouverture fonctionnelle $\gamma_b(f)$ et l'ouverture-logarithmique $\gamma_b^{\triangle}(f)$ est plus grande pour les intensités proches de M que pour celles proches de 0. En effet, en LMM, l'amplitude de la fs est plus élevée pour les plus hautes intensités de l'image que pour les plus faibles intensités à cause de la non-linéarité des lois LIP $\triangle$ et $\triangle$. La même observation peut être faite entre la fermeture fonctionnelle $\varphi_b(f)$ et la fermeture logarithmique $\varphi_b^{\triangle}(f)$ (Fig. 4.38d).



(a) Erosions. (b) Dilatations.

(c) Ouvertures. (d) Fermetures.

FIG. 4.38 – Sur une image f, comparaison entre la MM fonctionnelle et la LMM pour les (a) érosions $\varepsilon_b(f)$, $\varepsilon_b^{\triangle}(f)$, (b) dilatations $\delta_b(f)$, $\delta_b^{\triangle}(f)$, (c) ouvertures $\gamma_b(f)$, $\gamma_b^{\triangle}(f)$ et (d) fermetures $\varphi_b(f)$, $\varphi_b^{\triangle}(f)$. (a) et (b) Pour les deux pics du signal, la fonction structurante b est représentée (après une translation horizontale) pour les érosions $\varepsilon_b(f)$, $\varepsilon_b^{\triangle}(f)$ et les dilatations $\delta_b(f)$, $\delta_b^{\triangle}(f)$.

Lien avec la Morphologie Mathématique fonctionnelle

La LMM est définie dans le treillis $\overline{\mathcal{F}}_M$, tandis que la MM fonctionnelle est définie dans le treillis $\overline{\mathbb{R}}^D$. Afin de relier la LMM à la MM fonctionnelle, un isomorphisme entre les deux treillis est nécessaire. Cet isomorphisme $\xi : \overline{\mathcal{F}}_M \to \overline{\mathbb{R}}^D$ et son inverse $\xi^{-1} : \overline{\mathbb{R}}^D \to \overline{\mathcal{F}}_M$ ont été tous deux définis dans [148] par $\xi(f) = -M \ln(1 - f/M)$ et $\xi^{-1}(f) = M(1 - \exp(-f/M))$. Avec cet isomorphisme ξ, la proposition suivante peut être établie.

Proposition 14. *Soit $f \in \overline{\mathcal{F}}_M$ une fonction et $b \in \overline{\mathcal{F}}_M$ une fonction structurante toutes deux à valeurs majorées par M. La dilatation logarithmique $\delta_b^{\triangle}$ et l'érosion logarithmique $\varepsilon_b^{\triangle}$ sont respectivement liées à la dilatation fonctionnelle δ_b, ou $\oplus b$, et à l'érosion fonctionnelle ε_b, ou $\ominus b$, par les équations :*

$$\begin{aligned}\delta_b^{\triangle}(f) &= \xi^{-1}\left(\delta_{\xi(b)}[\xi(f)]\right) = \xi^{-1}\left[\xi(f) \oplus \xi(b)\right] \\ &= M[1 - \exp(-\delta_{\acute{b}}(\acute{f}))], \end{aligned} \tag{4.115}$$

$$\begin{aligned}\varepsilon_b^{\triangle}(f) &= \xi^{-1}\left(\varepsilon_{\xi(b)}[\xi(f)]\right) = \xi^{-1}\left[\xi(f) \ominus \xi(b)\right] \\ &= M[1 - \exp(-\varepsilon_{\acute{b}}(\acute{f}))], \end{aligned} \tag{4.116}$$

où $\acute{f} : D \to \overline{\mathbb{R}}$ est égale à $\acute{f} = -\ln(1 - f/M)$.

Ces relation facilitent la mise en œuvre informatique des opérateurs de LMM car ceux de MM fonctionnelle existent déjà dans de nombreux logiciels d'analyse d'image.

Pour prouver la proposition 14, il est nécessaire de démontrer le lemme suivant.

Lemme 1. *Soient $f, g \in \overline{\mathcal{F}}_M$ deux fonctions à valeurs majorées par M. L'isomorphisme ξ transforme respectivement la somme-LIP $\triangleup$ et la différence-LIP $\triangle$ entre des fonctions en la somme usuelle $+$ et la différence usuelle $-$. Nous avons*

$$\xi(f \triangleup g) = \xi(f) + \xi(g), \tag{4.117}$$
$$\xi(f \triangle g) = \xi(f) - \xi(g). \tag{4.118}$$

Démonstration. Soient $f, g \in \overline{\mathcal{F}}_M$ deux fonctions. Nous avons :

$$\begin{aligned}
\xi(f \triangleup g) &= -M \ln(1 - (f \triangleup g)/M) \\
&= -M \ln((1 - f/M)(1 - g/M)) \\
&= -M \ln(1 - f/M) - M \ln(1 - g/M) \\
&= \xi(f) + \xi(g), \\
\xi(\triangle g) &= -M \ln(1 + g/(M - g)) \\
&= -M \ln(M/(M - g)) \\
&= M \ln(1 - g/M) \\
&= -\xi(g), \\
\xi(f \triangle g) &= \xi(f \triangleup (\triangle g)) = \xi(f) + \xi(\triangle g) \\
&= \xi(f) - \xi(g).
\end{aligned}$$

□

Nous pouvons maintenant établir la preuve de la proposition 14.

Démonstration de la proposition 14. Soient f une fonction et b une fonction structurante de $\overline{\mathcal{F}}_M$. Soit $\acute{f} : D \to \overline{\mathbb{R}}$ une fonction égale à $\acute{f} = -\ln(1 - f/M)$. En tant que bijections croissantes, les isomorphismes ξ et ξ^{-1} distribuent sur les infima et suprema. La dilatation $\delta_b^{\triangleup}$ peut donc s'exprimer pour tout $x \in D$ par :

$$\begin{aligned}
\delta_b^{\triangleup} f(x) &= \xi^{-1} \circ \xi[\vee_{h \in D} \{f(x - h) \triangleup b(h)\}] \\
&= \xi^{-1} [\vee_{h \in D} \{\xi[f(x - h) \triangleup b(h)]\}] \\
&= \xi^{-1} [\vee_{h \in D} \{\xi[f(x - h)] + \xi[b(h)]\}] \\
&= \xi^{-1} \left[\delta_{\xi(b)}[\xi(f)]\right]
\end{aligned} \tag{4.119}$$

A partir de l'équation (4.119), nous avons :

$$\begin{aligned}
\delta_b^{\triangleup} f(x) &= M \left[1 - e^{\frac{-1}{M} \vee_{h \in D} \left\{-M \ln\left[1 - \frac{f(x-h)}{M}\right] - M \ln\left[1 - \frac{b(h)}{M}\right]\right\}}\right] \\
&= M \left[1 - e^{\frac{-1}{M} \vee_{h \in D} \left\{M\left(-\ln\left[1 - \frac{f(x-h)}{M}\right] - \ln\left[1 - \frac{b(h)}{M}\right]\right)\right\}}\right] \\
&= M \left[1 - e^{\frac{-M}{M} \vee_{h \in D} \left\{-\ln\left[1 - \frac{f(x-h)}{M}\right] - \ln\left[1 - \frac{b(h)}{M}\right]\right\}}\right] \\
&= M \left[1 - e^{-\vee_{h \in D} \left\{\acute{f}(x-h) + \acute{b}(h)\right\}}\right] \\
&= M \left[1 - e^{-\delta_{\acute{b}}(\acute{f})(x)}\right].
\end{aligned}$$

De manière similaire, nous avons :

$$\varepsilon_b^{\triangleup}(f) = \xi^{-1} \left[\varepsilon_{\xi(b)}[\xi(f)]\right] = M[1 - e^{-\varepsilon_{\acute{b}}(\acute{f})}].$$

□

Filtres de rang

Les dilatations fonctionnelles δ_b et les érosions ε_b sont fondées sur des opérations de supremum et d'infimum. Comme le supremum et l'infimum sont très sensibles au bruit de *speckle* [261], ils peuvent être remplacés par des *filtres de rang* [284] appelés aussi *filtres d'ordre statistique* [172], *filtres de centiles* ou *filtres de rang d'ordre*. Le filtre de rang k sélectionne les k plus petits éléments d'un ensemble. Il correspond au k^{e} *minimum* représenté par $\wedge^k$. Son dual, le k^{e} *maximum* $\vee^k$, sélectionne le k^{e} plus grand élément d'un ensemble. Dans l'équation (3.7), p. 23, un k^{e} *filtre de minimum* $\zeta_{b,k} : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ est défini avec le k^{e} minimum $\wedge^k$ au lieu de l'infimum $\wedge$. De façon similaire, dans l'équation (3.6), p. 23, un k^{e} *filtre de maximum* $\vartheta_{b,k}$ est défini avec le k^{e} maximum $\vee^k$. Si $k = 0$, le k^{e} *filtre de minimum* $\zeta_{b,k}$ est égal à l'*érosion* fonctionnelle ε_b. De même, le k^{e} *filtre de maximum* $\vartheta_{b,k}$ est alors égal à la *dilatation* fonctionnelle δ_b. En LMM, le k^{e} *filtre logarithmique de minimum* $\zeta^{\triangle}_{b,k} : \overline{\mathcal{F}}_M \to \overline{\mathcal{F}}_M$ et le k^{e} *filtre logarithmique de maximum* $\vartheta^{\triangle}_{b,k}$ peuvent être aussi définis en utilisant le k^{e} minimum $\wedge^k$ et le k^{e} maximum $\vee^k$ dans les équations (4.111) et (4.110), p. 140, respectivement. Les k^{e} *filtres logarithmiques de minimum et de maximum* $\zeta^{\triangle}_{b,k}$ et $\vartheta^{\triangle}_{b,k}$ sont aussi liés aux k^{e} *filtres de minimum et de maximum* $\zeta_{b,k}$ et $\vartheta_{b,k}$ en remplaçant les érosions ε_b, $\varepsilon^{\triangle}_b$ et les dilations δ_b, $\delta^{\triangle}_b$ par leurs filtres de rang correspondant dans les équations (4.116) et (4.115), respectivement.

4.4.4 Opérateurs robustes aux variations d'éclairement

Dans cette sous-section, je donnerai des exemples d'opérateurs robustes aux changements d'éclairement causés par des variations du temps d'exposition de la caméra ou de l'intensité lumineuse de la source. Ces changements d'éclairement sont modélisés par l'addition-LIP d'une constante.

Cartes des distances d'Asplund LIP-additives

Lien avec la LMM Il existe un lien fort entre les cartes des distances d'Asplund LIP-additives et la LMM. Ce lien sera détaillé dans les propositions suivantes.

Proposition 15. *Soit $f \in \overline{\mathcal{F}}_M$ une fonction et $b \in \overline{\mathcal{F}}_M$ une fonction structurante, où pour tout $x \in D_b$, $D_b \subset D$, $b(x) > -\infty$. La carte des distances d'Asplund entre la fonction f et la fonction structurante b est égale à :*

$$Asp^{\triangle}_b f = \delta^{\triangle}_{\triangle\bar{b}}(f) \triangle \varepsilon^{\triangle}_b(f). \tag{4.120}$$

Pour la carte des plus petits majorants de f, $c_{1_b} f$, et pour la carte des plus grands minorants de f, $c_{2_b} f$, nous avons :

$$c_{1_b} f = \delta^{\triangle}_{\triangle\bar{b}}(f), \tag{4.121}$$

$$c_{2_b} f = \varepsilon^{\triangle}_b(f). \tag{4.122}$$

Dans le cas d'expressions ambiguës, les conventions suivantes sont utilisées : $Asp^{\triangle}_b f(x) = M$ lorsque $\delta^{\triangle}_{\triangle\bar{b}}(f)(x) = M$ ou $\varepsilon^{\triangle}_b(f)(x) = -\infty$, et $Asp^{\triangle}_b f(x) = 0$ lorsque $\delta^{\triangle}_{\triangle\bar{b}}(f)(x) = \varepsilon^{\triangle}_b(f)(x)$.

Démonstration. A partir de l'équation (4.71), p. 113, nous avons $\forall x \in D$:

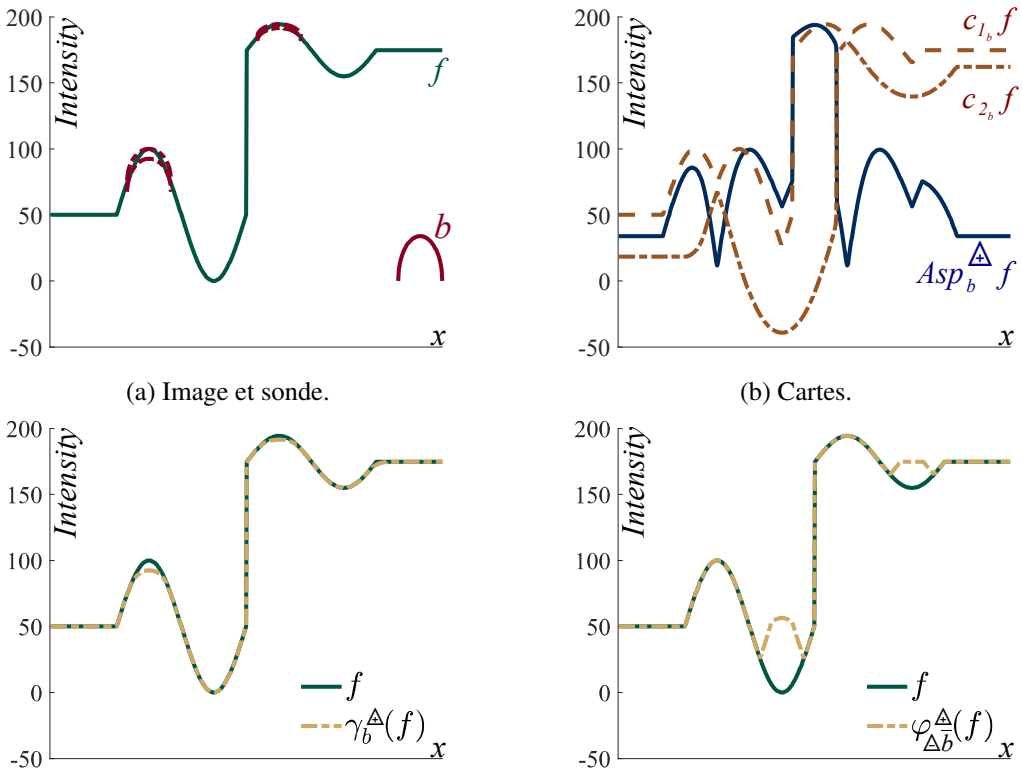
$$\begin{aligned} c_{1_b} f(x) &= \vee \{f(x+h) \triangle b(h), h \in D_b\} \\ &= \vee \{f(x-h) \triangle b(-h), -h \in D_b\} \\ &= \vee \left\{f(x-h) \triangle (\triangle\bar{b}(h)), h \in D_{\bar{b}}\right\} \\ &= \delta^{\triangle}_{\triangle\bar{b}}(f)(x). \end{aligned}$$

En comparant l'équation (4.111), p. 140, avec l'équation (4.72), p. 113, nous avons pour tout $x \in D$:

$$c_{2_b} f(x) = \varepsilon_b^{\triangle}(f)(x).$$

A partir de l'équation (4.73), p. 113, nous déduisons que $Asp_b^{\triangle} f = \delta_{\triangle\bar{b}}^{\triangle}(f) \triangle \varepsilon_b^{\triangle}(f)$. □

La carte des distances d'Asplund consiste en une double palpation d'une fonction f par la même sonde b par-dessus et par-dessous comme illustré dans la figure 4.39. La *carte des plus petits majorants* $c_{1_b} f$ et la *carte des plus grands minorants* $c_{2_b} f$ correspondent respectivement à une dilatation $\delta_{\triangle\bar{b}}^{\triangle}(f)$ et à une érosion $\varepsilon_b^{\triangle}(f)$ de l'image f par leurs fonctions structurantes respectives $\triangle\bar{b}$ ou b (Fig. 4.39b). Lorsque la sonde b est similaire à l'image f (selon la métrique d'Asplund), la carte des distances d'Asplund $Asp_b^{\triangle} f$ de l'image présente un minimum. Dans la figure 4.39b, les deux minima de la carte des distances correspondent aux deux bosses de l'image f (Fig. 4.39a).



(a) Image et sonde. (b) Cartes.

(c) Ouverture-logarithmique. (d) Fermeture-logarithmique.

FIG. 4.39 – (a) Une image f est analysée par une sonde b par-dessus et par-dessous. (b) La *carte des plus petits majorants* $c_{1_b} f$, la *carte des plus grands minorants* $c_{2_b} f$ et la carte des distances d'Asplund $Asp_b^{\triangle} f$ entre l'image et la sonde. (c) Le supremum de la sonde b tel que la sonde est en contact avec la fonction f par-dessus est l'ouverture-logarithmique $\gamma_b^{\triangle}$ de la fonction f. (d) L'infimum de la sonde b tel que la sonde est en contact avec la fonction f par-dessous est la fermeture-logarithmique $\varphi_{\triangle\bar{b}}^{\triangle}(f)$ de la fonction f par la fonction structurante $\triangle\bar{b}$.

Lien entre double-palpation et ouverture et fermeture logarithmiques Etudions le lieu du supremum ou de l'infimum des sondes translatées $b \in \overline{\mathcal{F}}_M$ telles qu'elles soient en contact avec la fonction f par-dessus ou par-dessous, respectivement. De façon similaire à celle d'Heijmans et al. dans [112], nous définissons l'automorphisme $\tau_{h,v}$ des translations horizontales et

verticales sur $\overline{\mathcal{F}}_M$ par

$$(\tau_{h,v}(f))(x) = f(x-h) \triangle v, \tag{4.123}$$

pour $h \in D$ et $v \in]-\infty; M[$. Pour les sondes inférieures, le supremum des sondes translatées b qui sont en contact avec la fonction f par-dessous est le supremum des sondes translatées $\tau_{h,v}(b)$ qui sont inférieures ou égales à la fonction f. Il est égal à :

$$\Gamma_b^{\triangle}(f) = \sup\{\tau_{h,v}(b) \mid h \in D, v \in]-\infty; M[\,, \tau_{h,v}(b) \leq f\}. \tag{4.124}$$

Comme démontré dans [261], $\Gamma_b^{\triangle}(f)$ est une ouverture. De façon similaire pour les sondes supérieures, l'infimum des sondes translatées b qui sont en contact avec la fonction f par-dessus est égal à

$$\Phi_b^{\triangle}(f) = \inf\{\tau_{h,v}(b) \mid h \in D, v \in]-\infty; M[\,, \tau_{h,v}(b) \geq f\}. \tag{4.125}$$

Comme démontré dans [261], $\Phi_b^{\triangle}(f)$ est une fermeture.

Proposition 16. *Le supremum des sondes translatées b qui sont en contact avec la fonction f par-dessous, $\Gamma_b^{\triangle}(f)$, est égal à l'ouverture-logarithmique de f, $\gamma_b^{\triangle}(f)$:*

$$\Gamma_b^{\triangle} f(x) = \gamma_b^{\triangle}(f)(x). \tag{4.126}$$

L'infimum des sondes translatées b qui sont en contact avec la fonction f par-dessus, $\Phi_b^{\triangle} f$, est égal à la fermeture-logarithmique de f, $\varphi_{\triangle\bar{b}}^{\triangle}(f)$:

$$\Phi_b^{\triangle} f(x) = \varphi_{\triangle\bar{b}}^{\triangle}(f)(x). \tag{4.127}$$

Démonstration. Soient f une fonction et b une sonde de $\overline{\mathcal{F}}_M$. Pour $h \in D$ et $v \in]-\infty; M[$, l'automorphisme $\tau_{h,v}$ des translations horizontales et verticales sur $\overline{\mathcal{F}}_M$ est défini par $(\tau_{h,v}(f))(x) = f(x-h) \triangle v$. Le supremum des sondes translatées b qui sont en contact avec la fonction f par-dessous est égal à :

$$\begin{aligned}
\Gamma_b^{\triangle} f(x) &= \sup\{\tau_{h,v}(b)(x) \mid h \in D, v \in]-\infty; M[\,, \tau_{h,v}(b) \leq f\} \\
&= \sup\{b(x-h) \triangle v \mid h \in D, v \in]-\infty; M[\,, x' \in D,\ b(x'-h) \triangle v \leq f(x')\} \\
&= \vee_{h\in D}\left\{b(x-h) \triangle v(h) \mid v(h) = \sup_{x'\in D}\{v, b(x'-h) \triangle v \leq f(x')\}\right\} \\
&= \vee_{h\in D}\left\{b(x-h) \triangle v(h) \mid v(h) = \sup_{x'\in D}\{v, b(x') \triangle v \leq f(h+x')\}\right\} \\
&= \vee_{h\in D}\{b(x-h) \triangle c_{2_b} f(h)\} \quad \text{(éq. (4.70), p. 113)} \\
&= \vee_{h\in D}\left\{b(x-h) \triangle \varepsilon_b^{\triangle}(f)(h)\right\} \quad \text{(Prop. 15, p. 144)} \\
&= \vee_{h\in D}\left\{\varepsilon_b^{\triangle}(f)(h) \triangle b(x-h)\right\} \\
&= \vee_{h\in D}\left\{\varepsilon_b^{\triangle}(f)(x-h) \triangle b(h)\right\} \\
&= \delta_b^{\triangle}\varepsilon_b^{\triangle}(f)(x) \quad \text{(éq. (4.110), p. 140)} \\
&= \gamma_b^{\triangle}(f)(x). \quad \text{(éq. (4.112), p. 141)}
\end{aligned}$$

Cette relation est celle d'une ouverture.

L'infimum des sondes translatées b qui sont en contact avec la fonction f par-dessus est égal à :

$$\begin{aligned}
\Phi_b^{\triangle\!\!\!\!+} f(x) &= \inf\{\tau_{h,v}(b)(x) \mid h \in D, v \in]-\infty; M[\,, \tau_{h,v}(b) \geq f\} \\
&= \inf\{b(x-h) \triangle\!\!\!\!+ v \mid h \in D, v \in]-\infty; M[\,, x' \in D,\ b(x'-h) \triangle\!\!\!\!+ v \geq f(x')\} \\
&= \wedge_{h \in D_b}\Big\{b(x-h) \triangle\!\!\!\!+ v(h) \mid v(h) = \inf_{x' \in D_b}\{v, b(x'-h) \triangle\!\!\!\!+ v \geq f(x')\}\Big\} \\
&= \wedge_{h \in D_b}\Big\{b(x-h) \triangle\!\!\!\!+ v(h) \mid v(h) = \inf_{x' \in D_b}\{v, b(h) \triangle\!\!\!\!+ v \geq f(x'+h)\}\Big\} \\
&= \wedge_{h \in D_b}\{b(x-h) \triangle\!\!\!\!+ c_{1_b} f(h)\} \quad \text{(éq. (4.69), p. 113)} \\
&= \wedge_{h \in D_b}\Big\{b(x-h) \triangle\!\!\!\!+ \delta_{\triangle \bar{b}}^{\triangle\!\!\!\!+}(f)(h)\Big\} \quad \text{(Prop. 15, p. 144)} \\
&= \wedge_{h \in D_b}\Big\{\delta_{\triangle \bar{b}}^{\triangle\!\!\!\!+}(f)(h) \triangle\!\!\!\!+ b(x-h)\Big\} \\
&= \wedge_{h \in D_b}\Big\{\delta_{\triangle \bar{b}}^{\triangle\!\!\!\!+}(f)(x-h) \triangle\!\!\!\!+ b(h)\Big\} \\
&= \wedge_{-h \in D_b}\Big\{\delta_{\triangle \bar{b}}^{\triangle\!\!\!\!+}(f)(x+h) \triangle\!\!\!\!+ b(-h)\Big\} \\
&= \wedge_{h \in D_{\bar{b}}}\Big\{\delta_{\triangle \bar{b}}^{\triangle\!\!\!\!+}(f)(x+h) \triangle (\triangle \bar{b}(h))\Big\} \\
&= \varepsilon_{\triangle \bar{b}}^{\triangle\!\!\!\!+} \delta_{\triangle \bar{b}}^{\triangle\!\!\!\!+}(f)(x) \quad \text{(éq. (4.111), p. 140)} \\
&= \varphi_{\triangle \bar{b}}^{\triangle\!\!\!\!+}(f)(x) \quad \text{(éq. (4.113), p. 141).}
\end{aligned}$$

Cette relation est celle d'une fermeture. □

Les figures 4.39c et 4.39d illustrent le supremum ou l'infimum des sondes b translatées et qui sont respectivement en contact avec la fonction f par-dessous ou par-dessus.

Lien avec le gradient morphologique-LIP Soit $b_0 : D \to \overline{\mathbb{R}}$ un élément structurant symétrique et constant qui est défini pour tout $x \in D_{b_0}$, où $D_{b_0} \subset D$, par $b_0(x) = b_0$ et $b_0(-x) = b_0(x)$. Dans le cas d'un élément structurant symétrique et constant b_0, la carte des distances d'Asplund LIP-additives $Asp_{b_0}^{\triangle\!\!\!\!+}$ est égale au gradient morphologique-LIP $\varrho_{B_0}^{LIP}$. Pour tout $x \in D$, nous avons :

$$Asp_{b_0}^{\triangle\!\!\!\!+} f(x) = \varrho_{B_0}^{LIP}(f)(x). \tag{4.128}$$

B_0 est un élément structurant plat avec le même support D_{b_0} que b_0.

Démonstration. Soit B_0 un élément structurant plat avec le même support D_{b_0} que l'élément structurant symétrique et constant b_0. $b_0 : D \to \overline{\mathbb{R}}$ est défini pour tout $x \in D_{b_0}$, où $D_{b_0} \subset D$, par $b_0(x) = b_0$ et $b_0(-x) = b_0(x)$. Pour tout $x \in D$, nous avons :

$$\begin{aligned}
Asp_{b_0}^{\triangle\!\!\!\!+} f(x) &= \delta_{\triangle \bar{b}_0}^{\triangle\!\!\!\!+}(f)(x) \triangle \varepsilon_{b_0}^{\triangle\!\!\!\!+}(f)(x) \\
&= \vee_{h \in D_{\bar{b}_0}}\{f(x-h) \triangle\!\!\!\!+ (\triangle \bar{b}_0)\} \triangle \wedge_{h \in D_{b_0}}\{f(x+h) \triangle b_0\} \\
&= \Big(\vee_{h \in D_{b_0}}\{f(x-h)\} \triangle b_0\Big) \triangle \Big(\wedge_{h \in D_{b_0}}\{f(x+h)\} \triangle b_0\Big) \\
&= \vee_{h \in D_{b_0}}\{f(x-h)\} \triangle \wedge_{h \in D_{b_0}}\{f(x+h)\} \\
&= \delta_{B_0}(f)(x) \triangle \varepsilon_{B_0}(f)(x) \\
&= \varrho_{B_0}^{LIP}(f)(x).
\end{aligned}$$

□

Lorsque la fonction structurante b est non plate, la carte des distances d'Asplund LIP-additives $Asp_b^{\triangle\!\!\!\!+}$ est une extension du *gradient morphologique-LIP* ϱ_B^{LIP}. Cependant, contrairement au gradient morphologique, la carte des distances d'Asplund n'est plus la norme d'un gradient. E.g. soit

$Y \subset D$ une zone constante (i.e. plate) d'une fonction f et soit $X = Y \ominus D_b$ la zone plate érodée par le domaine D_b de la fonction structurante b. $\ominus$ représente l'érosion binaire (éq. (3.2), p. 22). Dans la zone plate érodée X, la carte des distances d'Asplund est égale à une constante, tandis qu'un gradient et sa norme sont égaux à 0. Nous avons, pour tout $x \in X$:

$$Asp_b^{\triangle} f(x) = b_{max} \triangle b_{min}. \tag{4.129}$$

b_{max} et b_{min} sont respectivement le supremum et l'infimum, de la fonction structurante b.

Démonstration. Soit $Y \subset D$ une zone constante (i.e. plate) d'une fonction $f : D \to [-\infty; M]$ et soit $X = Y \ominus D_b$ la zone plate érodée par le domaine D_b de la fonction structurante b. Nous avons pour tout $x \in X$:

$$\begin{aligned}
Asp_b^{\triangle} f(x) &= c_{1_b} f(x) \triangle c_{2_b} f(x) \\
&= \vee_{h \in D_b} \{f(x+h) \triangle b(h)\} \triangle \wedge_{h \in D_b} \{f(x+h) \triangle b(h)\} \\
&= (f(x) \triangle \vee_{h \in D_b} \{b(h)\}) \triangle (f(x) \triangle \wedge_{h \in D_b} \{b(h)\}) \\
&= \triangle \vee_{h \in D_b} \{b(h)\} \triangle \wedge_{h \in D_b} \{b(h)\} \\
&= \wedge_{h \in D_b} \{b(h)\} \triangle \vee_{h \in D_b} \{b(h)\} \\
&= b_{max} \triangle b_{min}.
\end{aligned}$$

b_{max} et b_{min} sont respectivement le supremum et l'infimum, de la fonction structurante b. Dans la zone plate érodée X, la carte des distances d'Asplund de la fonction f est donc égale à la constante $b_{max} \triangle b_{min}$. □

Dans la figure 4.40, le gradient morphologique $\varrho_B(f)$ de l'image f et son gradient morphologique-LIP $\varrho_B^{LIP}(f)$, tous deux avec un élément structurant plat B, sont comparés à la carte des distances d'Asplund $Asp_b^{\triangle} f$ avec une fonction structurante b. Cette dernière a une forme de bosse qui a été choisie pour détecter les bosses de l'image f. Pour les deux bosses de l'image, la carte des distances d'Asplund $Asp_b^{\triangle} f$ présente deux minima profonds et de même dynamique, tandis que le gradient morphologique-LIP $\varrho_B^{LIP}(f)$ et le gradient morphologique $\varrho_B(f)$ ont deux minima régionaux avec une dynamique plus faible. En outre, les minima régionaux du gradient morphologique $\varrho_B(f)$ n'ont pas la même profondeur. Pour les zones plates de l'image f, les deux gradients $\varrho_B(f)$ et $\varrho_B^{LIP}(f)$ sont égaux à 0, tandis que la carte des distances d'Asplund distances $Asp_b^{\triangle} f$ est égale à une constante positive définie par l'équation (4.129).

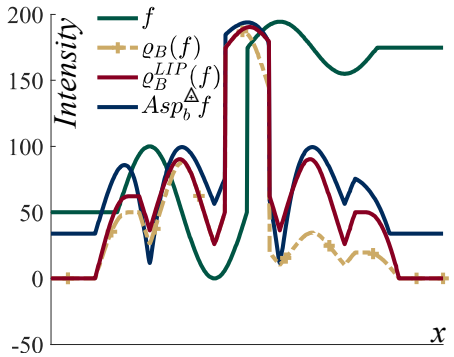


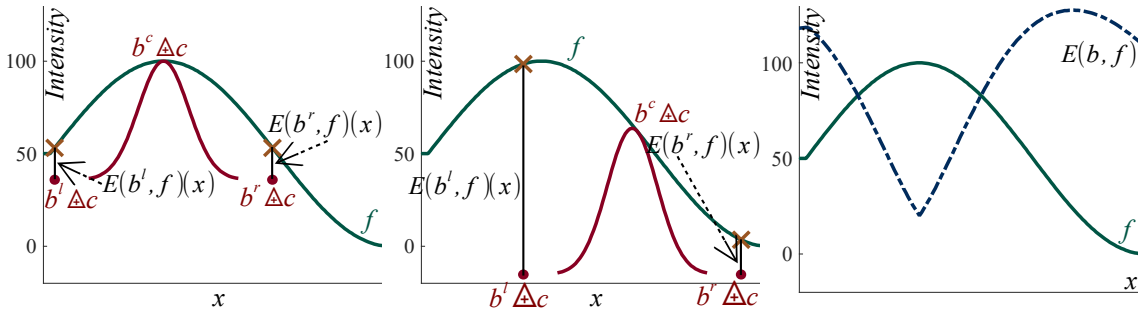
FIG. 4.40 – Comparaison entre le gradient morphologique $\varrho_B(f)$, le gradient morphologique-LIP $\varrho_B^{LIP}(f)$ et la carte des distances d'Asplund $Asp_b^{\triangle} f$ de l'image f. L'élément structurant plat B a le même support D_B que la fonction structurante b de la figure 4.39a.

La carte des distances d'Asplund LIP-additive est donc l'extension du gradient morphologique-LIP aux fonctions structurantes non plates. Elle donne à ce gradient des propriétés de métrique robuste aux variations d'éclairement. Le gradient morphologique-LIP est ainsi une double palpation d'une image par un élément structurant plat.

Différence-LIP entre des érosions-LIP

Sur la figure 4.41, une sonde b a été conçue pour détecter une bosse (Fig. 4.41a) mais pas une transition (Fig. 4.41b) sur une image monodimensionnelle f. La sonde b se compose de trois éléments : le point gauche b^l, le point droit b^r et une bosse centrale b^c avec approximativement la même amplitude que la bosse à détecter mais avec une largeur plus petite. Pour chaque point x du domaine D, la sonde est mise en contact avec l'image f par-dessous en LIP-ajoutant une constante c. Celle-ci est égale à la valeur de la *carte des plus grands minorants* $c_{2_b} f(x) = \bigwedge \{f(x+h) \triangle b(h), h \in D_b\}$ qui a été définie dans l'équation (4.72), p. 113. Nous calculerons la différence-LIP (ou contraste) entre l'image f et les points gauches et droits, b^l et b^r, de la sonde $b \triangle c_{2_b} f(x)$ qui est contact avec l'image f. Les détecteurs gauche et droit, $E(b^l, f)$ et $E(b^r, f) : D \to]-\infty; M[$, sont définis comme il s'ensuit :

$$\begin{aligned} E(b^l, f)(x) &= \wedge_{h \in D_b} \{f(x+h) \triangle [b^l(h) \triangle c_{2_b} f(x)]\} \\ &= \wedge_{h \in D_b} \{f(x+h) \triangle b^l(h)\} \triangle c_{2_b} f(x), \qquad (4.130) \\ E(b^r, f)(x) &= \wedge_{h \in D_b} \{f(x+h) \triangle [b^r(h) \triangle c_{2_b} f(x)]\} \\ &= \wedge_{h \in D_b} \{f(x+h) \triangle b^r(h)\} \triangle c_{2_b} f(x). \qquad (4.131) \end{aligned}$$



(a) Bosse et sonde. (b) Transition et sonde. (c) Détecteur.

FIG. 4.41 – Une image monodimensionnelle f est analysée par une sonde b par-dessous. Nous voulons détecter (a) une bosse mais pas (b) une transition. Les détecteurs gauche et droit $E(b^l, f)(x)$ et $E(b^r, f)(x)$ sont les différences-LIP entre f et l'élément gauche b^l de la sonde b et son élément droit b^r, respectivement ; après l'addition-LIP d'une constante c. (c) Détecteur $E(b, f)$ sur une bosse et une transition.

En cas de bosse similaire à la sonde, les détecteurs gauche $E(b^l, f)$ et droit $E(b^r, f)$ auront des valeurs proches (Fig. 4.41a), tandis qu'en cas de transition, un des détecteurs aura une valeur beaucoup plus élevée que celle de l'autre (Fig. 4.41b). Cette propriété permet de séparer les bosses (similaires à la sonde) des transitions. Le détecteur de bosse $E(b, f) : D \to]-\infty; M[$ est donc défini comme le supremum $\vee$ point à point entre les détecteurs gauche et droit :

$$E(b, f) = \vee\{E(b^l, f), E(b^r, f)\}. \qquad (4.132)$$

Comme illustré dans la figure 4.41c, en cas de bosse, la détecteur présente un minimum profond, tandis que dans le cas d'une transition, ce minimum disparaît.

Propriété 9. *Soient $f \in \overline{\mathcal{F}}_M$ une fonction et b^l, $b^r \in \overline{\mathcal{F}}_M$ des fonction structurantes. Les détecteurs gauche $E(b^l, f)$ et droit $E(b^r, f)$ sont égaux à des différences-LIP entre des érosions-logarithmiques :*

$$E(b^l, f) = \varepsilon_{b^l}^{\triangle}(f) \triangle \varepsilon_{b}^{\triangle}(f), \qquad (4.133)$$

$$E(b^r, f) = \varepsilon_{b^r}^{\triangle}(f) \triangle \varepsilon_{b}^{\triangle}(f). \qquad (4.134)$$

Démonstration. D'après l'équation (4.111), p. 140, nous avons $\forall x \in D$, $\wedge_{h\in D_b}\{f(x+h) \triangle b^l(h)\} = \varepsilon_{b^l}^{\triangle}(f)(x)$ et d'après l'équation (4.122), p. 144, $c_{2_b} f = \varepsilon_b^{\triangle}(f)$. De l'équation (4.130), on déduit que :

$$\begin{aligned} E(b^l, f)(x) &= \wedge_{h\in D_b}\{f(x+h) \triangle b^l(h)\} \triangle c_{2_b} f(x) \\ &= \varepsilon_{b^l}^{\triangle}(f)(x) \triangle \varepsilon_b^{\triangle}(f)(x). \end{aligned}$$

De façon similaire, nous avons $E(b^r, f)(x) = \varepsilon_{b^r}^{\triangle}(f)(x) \triangle \varepsilon_b^{\triangle}(f)(x)$. □

Propriété 10. *Les détecteurs gauche $E(b^l, f)$ et droit $E(b^r, f)$ et le détecteur de bosse $E(b, f)$ sont insensibles à l'addition-LIP (ou à la soustraction-LIP) de tout constante $c \in]-\infty; M[$: $E(b^l, f \triangleq c) = E(b^l, f)$, $E(b^r, f \triangleq c) = E(b^r, f)$ et $E(b, f \triangleq c) = E(b, f)$.*

Démonstration. D'après l'équation (4.111), p. 140, nous avons $\forall x \in D$:

$$\begin{aligned} \varepsilon_b^{\triangle}(f \triangleq c)(x) &= \wedge_{h\in D_b}\{(f \triangleq c)(x+h) \triangle b(h)\} \\ &= \wedge_{h\in D_b}\{f(x+h) \triangleq c \triangle b(h)\} \\ &= \wedge_{h\in D_b}\{f(x+h) \triangle b(h)\} \triangleq c \\ &= \varepsilon_b^{\triangle}(f) \triangleq c. \end{aligned}$$

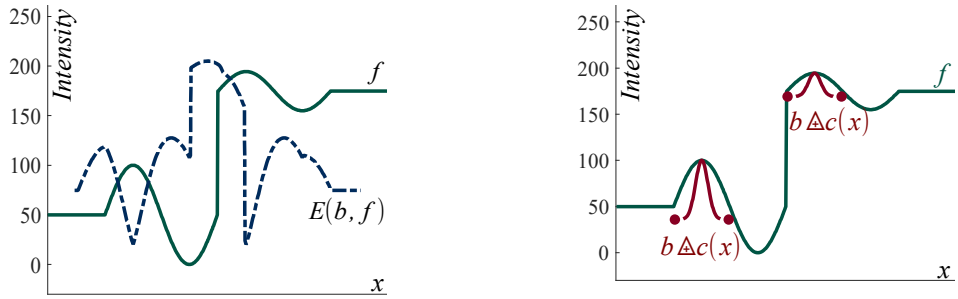
De l'équation (4.133), on déduit que :

$$\begin{aligned} E(b^l, f \triangleq c) &= \varepsilon_{b^l}^{\triangle}(f \triangleq c) \triangle \varepsilon_b^{\triangle}(f \triangleq c) \\ &= (\varepsilon_{b^l}^{\triangle}(f) \triangleq c) \triangle (\varepsilon_b^{\triangle}(f) \triangleq c) \\ &= \varepsilon_{b^l}^{\triangle}(f) \triangle \varepsilon_b^{\triangle}(f) \triangleq c \triangle c \\ &= \varepsilon_{b^l}^{\triangle}(f) \triangle \varepsilon_b^{\triangle}(f). \end{aligned}$$

De façon similaire, nous obtenons $E(b^r, f \triangleq c) = E(b^r, f)$. Par conséquent, d'après l'équation (4.132), nous avons $E(b, f \triangleq c) = E(b, f)$. □

La figure 4.42a illustre l'application du détecteur de bosse sur une image monodimensionnelle f. Le détecteur présente des minima de même profondeur pour les deux bosses de l'image. Celles-ci sont reliées par l'addition-LIP d'une constante c qui modélise un changement d'intensité de l'image causé par une variation d'intensité lumineuse ou du temps d'exposition de la caméra. Le détecteur par différence-LIP entre des érosions-logarithmiques est donc robuste à des variations d'éclairement modélisées par l'addition-LIP d'une constante. A cause de l'addition-LIP de la carte des plus grands minorants $c_{2_b} f(x)$ de l'image, à la sonde b, celle-ci à une amplitude qui s'adapte à l'intensité de l'image (Fig. 4.42b). La méthode présentée sera notamment utilisée dans la sous-section 4.5.1, p. 153, afin de détecter les vaisseaux du fond d'œil [211].

La différence-LIP $\triangle$ entre deux opérations morphologiques logarithmiques serait aussi robuste à des variations d'éclairement. E.g. la figure 4.37d, p. 138 illustre la robustesse de la différence-LIP entre une ouverture-logarithmique $\gamma_b^{\triangle} f$ de l'image f par une sonde gaussienne b entourée d'un anneau (Fig. 4.37e) et une ouverture-logarithmique $\gamma_{b_a}^{\triangle} f$ par ce même anneau b_a. Le résultat, $\gamma_b^{\triangle} f \triangle \gamma_{b_a}^{\triangle} f$, est insensible à une dérive d'éclairement causée par l'addition LIP d'une fonction linéaire (i.e. un plan). En outre, la forme gaussienne de la sonde permet de détecter la spirale sans détecter les deux courbes rapprochées. Cette détection n'aurait pas pu être réalisée en MM fonctionnelle usuelle.

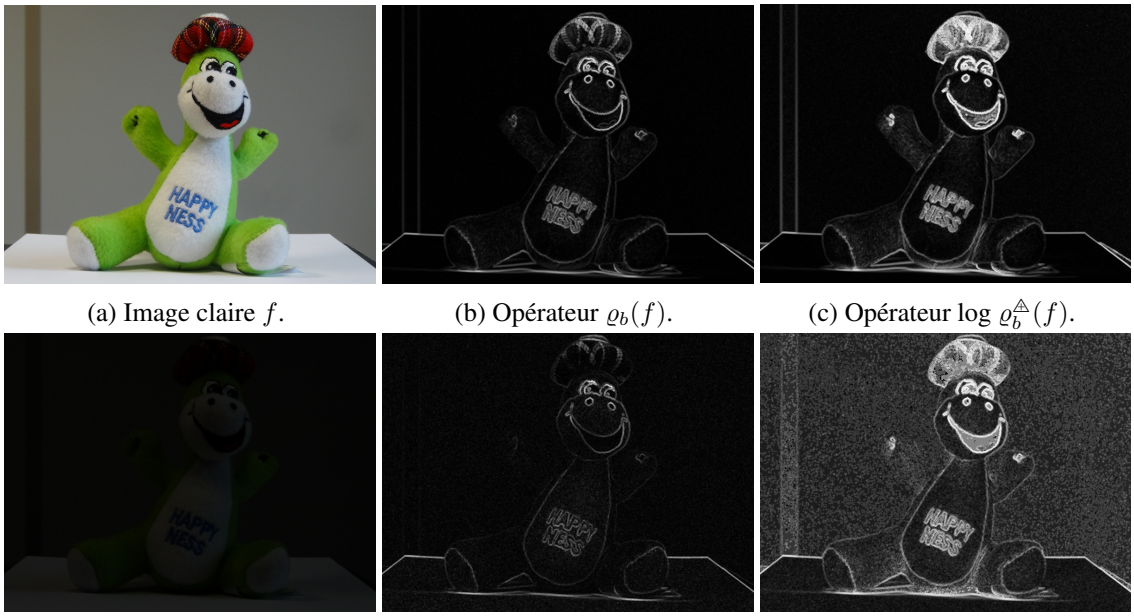


(a) Image et détecteur. (b) Image et sonde à deux endroits.

FIG. 4.42 – (a) Image et détecteur de bosse $E(b, f)$. (b) Image monodimensionnelle f et sonde b pour deux points $x \in D$.

4.4.5 Illustrations

Illustrons maintenant la différence entre MM fonctionnelle et LMM. Sur la figure 4.43, deux images de la même scène sont acquises avec des temps d'exposition de la caméra différents. Une image claire f est capturée avec un temps d'exposition de 1/40 s (Fig. 4.43a) et une image sombre f^d est acquise avec un temps d'exposition plus court, 1/800 s (Fig. 4.43d). Considérons une fonction structurante b hémisphérique. La différence entre une érosion et une dilatation $\varrho_b(f) = \delta_b f - \varepsilon_b f$ n'est pas un gradient morphologique puisque que la fonction structurante n'est pas plate. Le calcul de cet opérateur ϱ_b pour les deux images sombre f^d et claire f donne des résultats différents qui dépendent du temps d'exposition utilisé (Fig. 4.43b et 4.43e). Cependant, l'opérateur logarithmique $\varrho_b^{\triangle} = \delta_b^{\triangle} f \triangle \varepsilon_b^{\triangle} f$ conduit à des résultats similaires quel que soit le temps d'exposition (Fig. 4.43c et 4.43f) même s'il y a un peu de bruit lors de l'acquisition avec un faible temps d'exposition. L'opérateur logarithmique $\varrho_b^{\triangle}$ est ainsi robuste aux variations du temps d'exposition de la caméra contrairement à l'opérateur morphologique usuel ϱ_b.



(a) Image claire f. (b) Opérateur $\varrho_b(f)$. (c) Opérateur log $\varrho_b^{\triangle}(f)$.

(d) Image sombre f^d. (e) Opérateur $\varrho_b(f^d)$. (f) Opérateur log $\varrho_b^{\triangle}(f^d)$.

FIG. 4.43 – Comparaison entre (b, e) l'opérateur morphologique ϱ_b et (c, f) l'opérateur logarithmique $\varrho_b^{\triangle}$ sur une image claire f (a) (acquise avec un temps d'exposition de 1/40 s) et sur une image sombre de la même scène (d) (temps d'exposition de 1/800 s). La fonction structurante b est un hémisphère de rayon 2 pixels.

Sur la figure 4.44, une ouverture $\gamma_b(f^d)$ et une fermeture $\varphi_b(f^d)$ sont comparées à l'ouverture-logarithmique $\gamma_b^{\triangle}(f^d)$ et à la fermeture-logarithmique $\varphi_b^{\triangle}(f^d)$ sur l'image sombre f^d capturée avec un temps d'exposition de 1/800 s (Fig. 4.43d). La fonction structurante est un hémisphère de rayon 15 pixels. L'ouverture $\gamma_b(f^d)$ et la fermeture $\varphi_b(f^d)$ usuelles (Fig. 4.44a et 4.44c) ont un effet limité en ce qui concerne la modification de l'image tandis que l'ouverture et la fermeture logarithmiques ont un effet plus important (Fig. 4.44b et 4.44d). E.g. sur le corps du monstre en peluche, les mots sont enlevés avec l'ouverture et la fermeture logarithmiques tandis qu'ils ne le sont pas avec l'ouverture et la fermeture usuelles.

Ces exemples illustrent la propriété de la Morphologie Mathématique Logarithmique (LMM) de s'adapter à des changements importants de contraste causés par une variation de l'intensité lumineuse de la source ou de temps d'exposition de la caméra.

(a) Ouverture $\gamma_b(f^d)$.

(b) Ouverture-logarithmique $\gamma_b^{\triangle}(f^d)$.

(c) Fermeture $\varphi_b(f^d)$.

(d) Fermeture-logarithmique $\varphi_b^{\triangle}(f^d)$.

FIG. 4.44 – (a) Ouverture $\gamma_b(f^d)$ et (b) ouverture-logarithmique $\gamma_b^{\triangle}(f^d)$ sur l'image sombre f^d de la figure 4.43d. (c) Fermeture $\varphi_b(f^d)$ et (d) fermeture-logarithmique $\varphi_b^{\triangle}(f^d)$. La fonction structurante b est un hémisphère de rayon 15 pixels. A des fins d'affichage, les niveaux de gris ont été mis à l'échelle entre 0 et 255.

4.4.6 Perspectives

La LMM constitue un cadre de travail bien défini d'un point de vue mathématique et avec des propriétés optiques très intéressantes. Elle permet d'étendre le gradient morphologique logarithmique aux fonctions structurantes par la carte des distances d'Asplund LIP-additives. Elle fournit aussi de nouveaux outils de détection robustes aux variations d'éclairement et qui dépendent de la forme des fonctions structurantes choisies. Il reste maintenant à étudier plus en détail les opérations de LMM notamment celles avec deux sondes différentes. De même que j'ai défini une LMM pour la loi LIP-additive $\triangle$, des opérations morphologiques peuvent être définies avec la loi LIP-multiplicative $\triangle$ et reliées à la carte des distances d'Asplund LIP-multiplicatives. Celle-ci aura un intérêt pour les images d'objets acquis en transmission et dont l'opacité peut varier. *Toutefois, la LMM s'avère être de première importance pour de nombreuses applications et notamment en analyse d'images médicales où les images présentent souvent des variations de contraste dues aux variations d'éclairement lors de l'acquisition [211].*

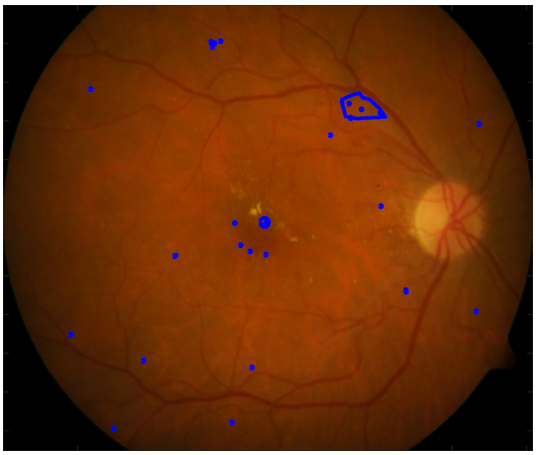
4.5 Segmentation robuste aux variations d'éclairement

Grâce au Traitement d'Image Logarithmique, aux métriques d'Asplund fonctionnelles et à la Morphologie Mathématique Logarithmique, j'ai pu introduire des méthode de segmentations robustes au variations d'éclairement.

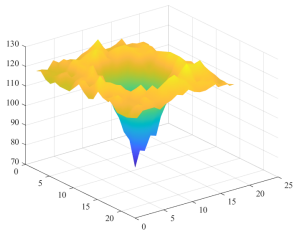
4.5.1 Segmentation des micro-anévrismes et des vaisseaux du fond d'œil

Segmentation des micro-anévrismes

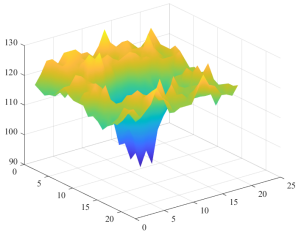
Les images rétiniennes peuvent présenter des micro-anévrismes qui apparaissent comme des petits points de couleur rouge (cf. section 2.3, p. 15). La figure 4.45 illustre des exemples de la surface topographique de plusieurs micro-anévrismes sur l'image de luminance (à niveaux de gris) d'une image couleur du fond d'œil. Cette image est extraite de la banque de données publiques DIARETDB1 [151].



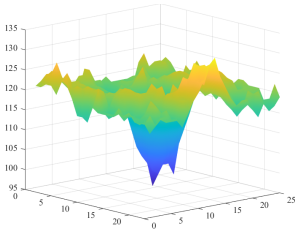
(a) Image avec micro-anévrismes.



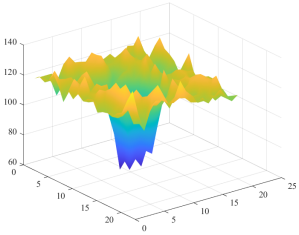
(b) Micro-anévrisme 1.



(c) Micro-anévrisme 2.



(d) Micro-anévrisme 3.

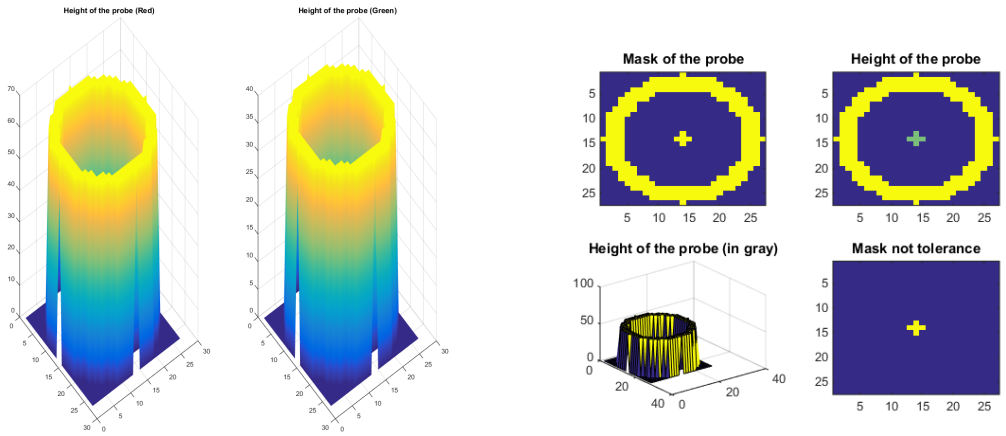


(e) Micro-anévrisme 4.

FIG. 4.45 – (a) Image du fond d'œil avec des micro-anévrismes entourés en bleu par un spécialiste. Source : banque de données DIARETDB1 [151]. (b), (c), (d) et (e) Surfaces topographiques de la luminance de plusieurs micro-anévrismes extraits dans l'image (a).

Afin de détecter les micro-anévrismes, une sonde $\mathbf{b}$ est conçue comme un pic entouré par un anneau. La figure 4.46 illustre cette sonde qui ne contient que deux des composantes couleurs : rouge et verte, $\mathbf{b} = (b_R, b_G)$. En effet, dans les images du fond d'œil, la composante bleue ne comporte que très peu d'information et peut être négligée. En outre, le pic central de la sonde n'est pas soumis à la tolérance pour supprimer les points extrémaux dans le calcul de la métrique d'Asplund. En effet, ce point central est nécessaire pour bien différencier les micro-anévrismes des autres formes présentes dans l'image.

La sonde est ensuite utilisée pour une double palpation d'un micro-anévrisme par métrique d'Asplund uniquement sur la composante verte (Fig. 4.47a) ou simultanément sur les deux composantes rouge et verte (Fig. 4.47b). Dans ce deuxième cas, c'est la métrique multivariée présentée



(a) Amplitude de la sonde multivariée pour les composantes rouge et verte.

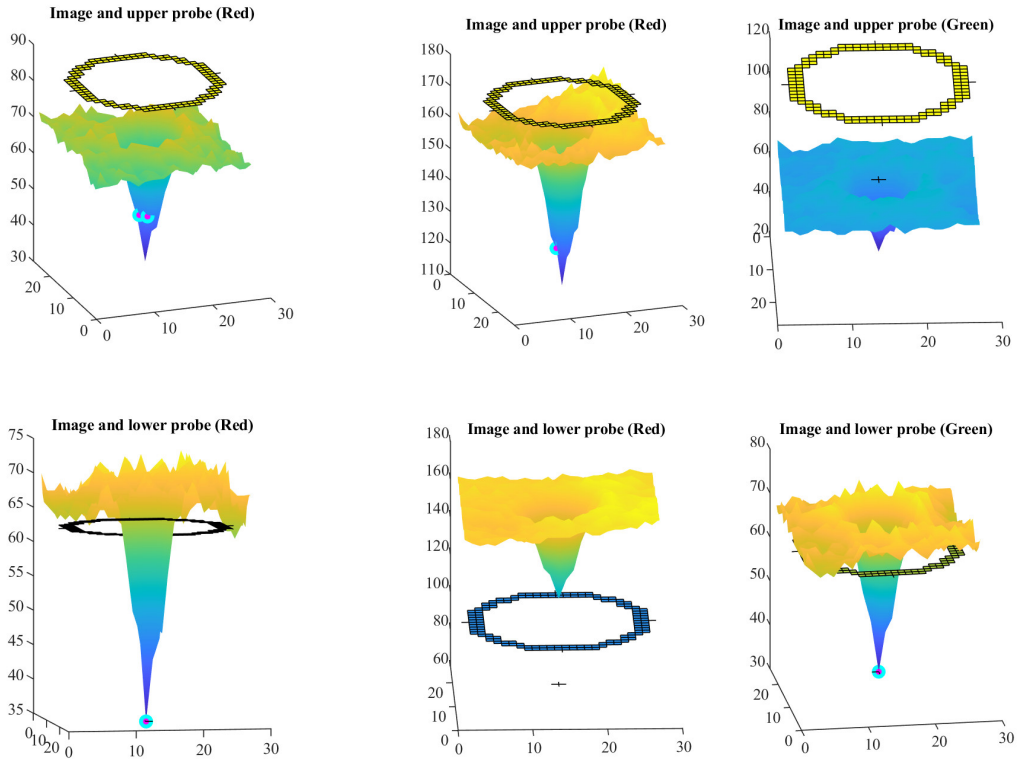
(b) Masque de la sonde.

FIG. 4.46 – (a) Amplitude de la sonde multivariée $\mathbf{b}$ pour les composantes rouge (à gauche) et verte (à droite). (b) Masque et amplitude à niveaux de gris de la sonde. Le « masque de non tolérance » (*mask not tolerance*) correspond au pic central de la sonde pour lequel la tolérance sur la métrique d'Asplund n'est pas appliquée. L'amplitude de la sonde est donnée en niveaux de gris comme la moyenne des composantes rouge et verte.

dans la sous-sous-section « Approche avec le modèle LIP », p. 123, qui est utilisée. Comme la méthode multivariée est définie grâce à une palpation conjointe sur les deux composantes, seules la sonde supérieure rouge et la sonde inférieure verte sont, dans cet exemple, en contact avec l'image du micro-anévrisme (Fig. 4.47b).

Une carte des distances d'Asplund LIP-multiplicatives avec tolérance $Asp^{\triangle}_{b_G,p=97\%} f_G$ est ensuite calculée sur les composantes vertes de la sonde b_G et de l'image f_G (Fig. 4.48a). Le paramètre de tolérance est fixé à $p = 97\ \%$. Une comparaison est effectuée avec la carte des distances d'Asplund $Asp^{\triangle}_{\mathbf{b},p=97\%}\mathbf{f}$ sur les composantes rouges et vertes de la sonde $\mathbf{b} = (b_R, b_G)$ et de l'image $\mathbf{f} = (f_R, f_G)$ (Fig. 4.48b).

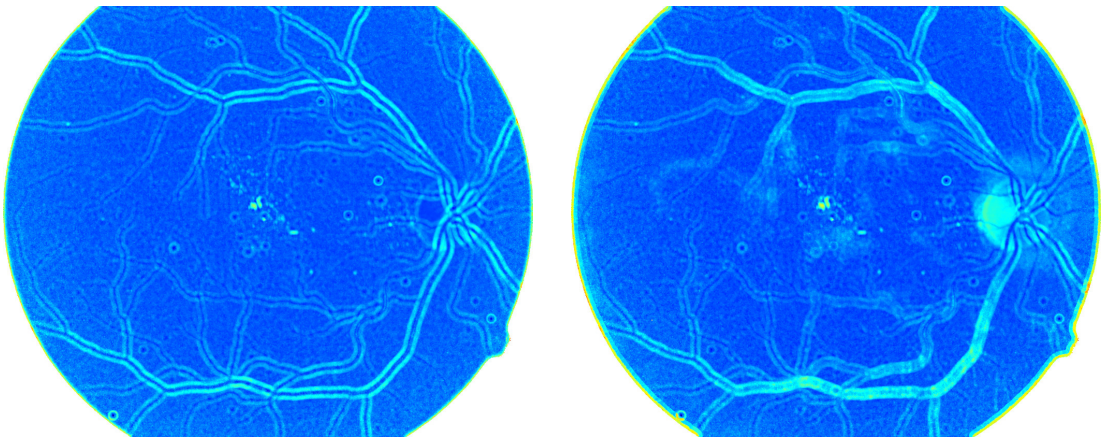
Les cartes des distances d'Asplund sont ensuite segmentées par le moyen d'un seuil sur un *black-top-hat*. Celui-ci est calculé grâce à un infimum sur des fermetures directionnelles comme expliqué dans [331]. Afin de rendre la fermeture φ_{b_φ} plus robuste, elle est précédée d'une ouverture γ_{b_γ} avec un élément structurant b_γ de plus petite longueur que celui de la fermeture b_φ, $long(b_\gamma) \leq long(b_\varphi)$. Comme présenté dans la sous-section 3.2.2, p. 36, la fermeture s'écrit donc $\wedge(g, \varphi_{b_\varphi}\gamma_{b_\gamma} g)$, pour la carte des distances $g = Asp^{\triangle}_{b_G} f_G$ ou $g = Asp^{\triangle}_{\mathbf{b}}\mathbf{f}$. Les résultats de segmentation des deux cartes des distances d'Asplund sont présentées sur la figure 4.49. Les deux méthodes extraient des micro-anévrismes mais aussi d'autres zones à l'intersection des vaisseaux. Il est à noter que la vérité terrain est ici approximative et que tous les micro-anévrismes n'ont pas été entourés. Ces segmentations constituent des résultats préliminaires encourageants qu'il conviendra d'affiner en tenant compte notamment d'une segmentation des vaisseaux et de la papille optique. En outre, il sera intéressant de tester la carte des distances d'Asplund LIP-additive qui est bien adaptée pour les images acquises par réflexion de la lumière sur l'objet d'intérêt.



(a) Double palpation sur la composante verte.

(b) Double palpation sur les composantes verte et rouge.

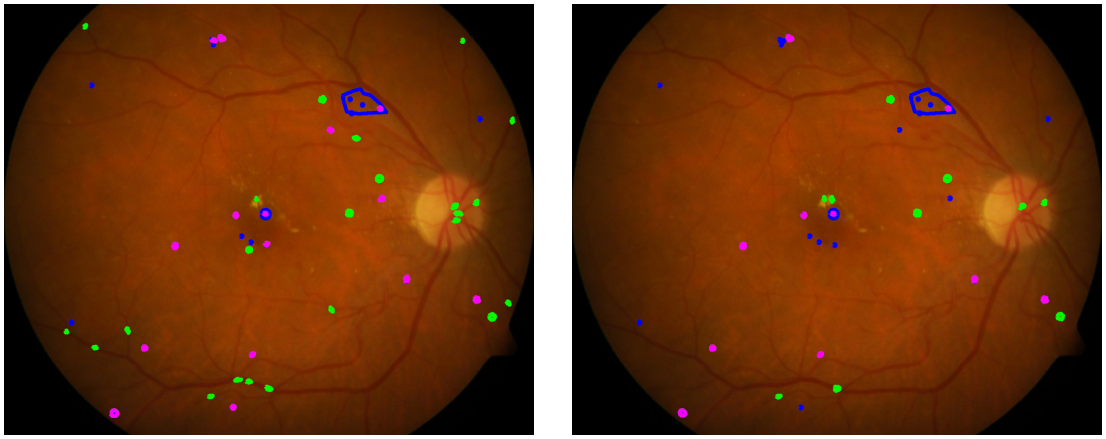
FIG. 4.47 – (a) Double palpation (supérieure et inférieure) d'un micro-anévrisme sur la composante verte par une sonde à niveaux de gris b. (b) Double palpation (supérieure et inférieure) d'un micro-anévrisme sur les composantes verte et rouge par une sonde multivariée à deux composantes $\mathbf{b} = (b_R, b_G)$. La sonde supérieure rouge et la sonde inférieure verte sont en contact avec l'image.



(a) Carte des distances d'Asplund $Asp^{\triangle}_{b_G,p} f_G$ sur la composante verte.

(b) Carte des distances d'Asplund $Asp^{\triangle}_{\mathbf{b},p} \mathbf{f}$ sur les composantes rouge et verte.

FIG. 4.48 – (a) Carte des distances d'Asplund $Asp^{\triangle}_{b_G,p=97\%} f_G$ sur les composantes vertes de la sonde b_G et de l'image f_G. (b) Carte des distances d'Asplund $Asp^{\triangle}_{\mathbf{b},p=97\%} \mathbf{f}$ sur les composantes rouges et vertes de la sonde $\mathbf{b} = (b_R, b_G)$ et de l'image $\mathbf{f} = (f_R, f_G)$.



(a) Segmentation sur la composante verte de l'image.

(b) Segmentation sur les composantes rouge et verte de l'image.

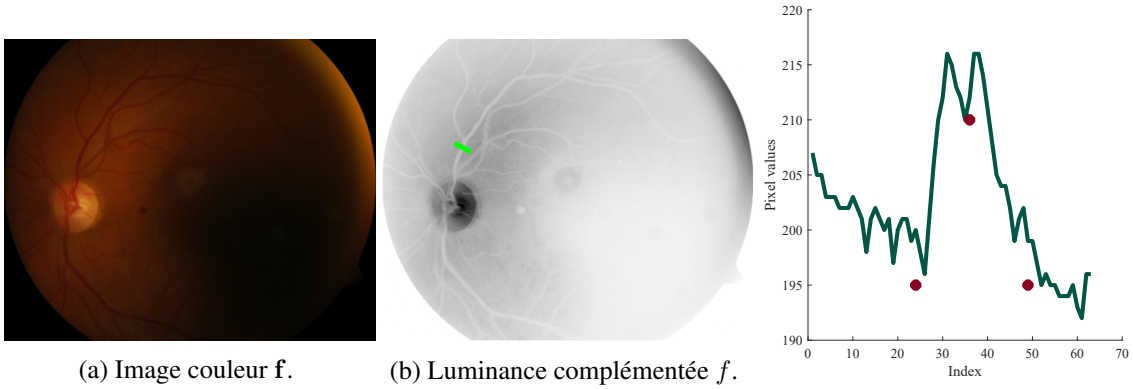
FIG. 4.49 – (a) Seuillage d'un *black-top-hat* multidirectionnel sur la carte des distances d'Asplund $Asp^{\triangle}_{b_G} f_G$ de la composante verte de l'image. (b) Seuillage d'un *black-top-hat* multi-directionnel sur la carte des distances d'Asplund $Asp^{\triangle}_{\mathbf{b}} \mathbf{f}$ des composantes rouge et verte de l'image. Les zones segmentées apparaissent en vert (faux positifs) ou en magenta (vrais positifs) et la vérité terrain (approximative) en bleu.

Segmentation des vaisseaux

Motivation Afin de détecter les vaisseaux des images du fond d'œil, j'ai proposé une nouvelle méthode reposant sur un détecteur robuste aux variations d'éclairement [211]. Ce travail a été réalisé en collaboration avec l'équipe du Pr. Laurent Kodjikian, ophtalmologiste aux *Hospices Civils de Lyon* et à l'Université Claude Bernard, Lyon 1.

L'extraction de vaisseaux (veines ou artères) dans les images du fond d'œil a été explorée dans de nombreux papiers, e.g. [16], [122], [182], [287], [338], [339]. Cependant, ces méthodes peuvent présenter des limites lorsqu'il existe de fortes variations d'éclairement dans les images. L'existence de programmes de dépistage pour la rétinopathie diabétique a conduit à la création de grandes banques d'images du fond d'œil qui contiennent des variations de contraste. Elles peuvent être dues : à l'absorption inhomogène des tissus de l'œil ou à différentes conditions d'éclairement [234]. Notre but est d'introduire une méthode des segmentation qui soit adaptative aux variations d'éclairement dans les images couleur du fond d'œil. Après avoir complémenté la luminance de ces images, les vaisseaux apparaissent comme un relief positif (i.e. a « une chaîne de montagne ») sur la surface topographique des images. Leur détection peut être réalisée par une sonde composée de trois segments parallèles, où le segment central a une plus forte intensité que les deux autres. Lorsque la sonde est à l'intérieur d'un vaisseau (i.e. une « montagne »), la différence d'intensité entre ses segments extérieurs et le pied de la montagne est minimal, tandis que lorsque la sonde est en dehors d'un vaisseau, la différence d'intensité devient plus grande. Ce principe sera utilisé pour détecter les vaisseaux. L'adaptabilité aux variations d'éclairement provient quant à elle du modèle LIP.

Méthode Dans cette partie, j'utilise la méthode introduite dans la sous-section 4.4.4, p. 149. Considérons une image couleur $\mathbf{f} = (f_R, f_G, f_B)$ et sa luminance à niveaux de gris $\bar{f}$ qui est égale à $0{,}299 \times f_R + 0{,}587 \times f_G + 0{,}114 \times f_B$ (Fig. 4.50a). Pour être dans l'échelle LIP, la luminance $\bar{f}$ est complémentée, $f = (M-1) - \bar{f}$ (Fig. 4.50b), et un profil f_p est extrait sur un vaisseau (Fig. 4.50c). Comme ce vaisseau apparaît comme une « bosse » ou une « montagne », une sonde b composée de 3 points est conçue pour palper son profil par-dessous. Le point central de la sonde b^c présente une plus forte intensité que les points de gauche b^l et de droite b^r, situés en bas de la sonde. La distance entre les points du bas (i.e. la largeur de la sonde) est plus grande que la largeur du vaisseau (Fig. 4.50c).

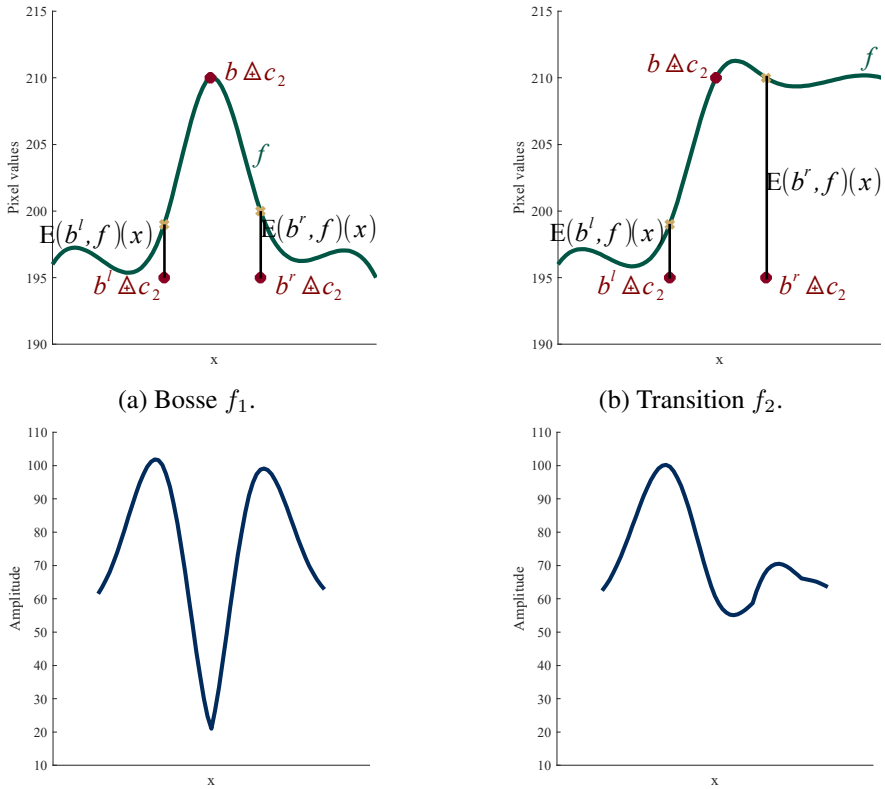


(a) Image couleur $\mathbf{f}$. (b) Luminance complémentée f. (c) Profil f_p et sonde b.

FIG. 4.50 – (a) Image faiblement contrastée de la banque DIARETDB1 [151]. (b) Sur son image de luminance après complémentation f, un profil de vaisseau est extrait (segment vert). (c) Profil f_p et sonde b.

Considérons les profils de la figure 4.51, où nous voulons détecter une bosse f_1 (Fig. 4.51a) mais pas une transition f_2 (Fig. 4.51b). Afin de mettre en contact la sonde 3-points b avec le profil f_p, une constante c_2 est LIP-ajoutée à la sonde (éq. (4.130) et (4.131), p. 149) et un détecteur

$E(b, f)$ de bosse est créé (éq. (4.132), p. 149). Dans le cas d'une bosse, le détecteur présente un minimum profond (Fig. 4.51c), tandis que dans le cas d'une transition, ce minimum disparaît (Fig. 4.51d).



(a) Bosse f_1.

(b) Transition f_2.

(c) Détecteur $E(b, f_1)$ de la transition f_1.

(d) Détecteur $E(b, f_2)$ de la transition f_2.

FIG. 4.51 – (a) Palpation d'une bosse f_1. (b) Palpation d'une transition f_2. Les détecteurs gauche et droit, $E(b^l, f_p)(x)$ et $E(b^r, f_p)(x)$ sont montrés par des flèches noires verticales pour des points $x \in D$. Cartes des détecteurs de bosse pour (c) la bosse, $E(b, f_1)$, et pour (d) la transition, $E(b, f_2)$.

Comme les images rétiniennes sont des fonctions 2D, la sonde b_θ, définie sur le domaine $D_{b_\theta} \subset D$, se compose de 3 segments parallèles de même longueur l et de même orientation θ (Fig. 4.52a). L'origine de la sonde correspond à une des extrémités du segment central b^c_θ. Son intensité est plus grande que l'intensité des segments gauche et droit b^l_θ et b^r_θ (Fig. 4.52b). Ces deux segments sont équidistants de celui du centre et leur distance, i.e. la largeur de la sonde, est égale à w. Afin de définir des opérateurs robustes au bruit, nous utiliserons le k^{e} minimum, $\wedge^k$, défini comme le k^{e} élément de l'ensemble $X = \{x_1, \ldots x_k, \ldots x_n\}$ trié par ordre décroissant, $\wedge^k(X) = x_k$.

Soit f l'image complémentée de luminance d'une image rétinienne. Les cartes des détecteurs gauche et droit, $E_{b^l_\theta} f$ et $E_{b^r_\theta} f$, sont pour tout $x \in D$ égales à :

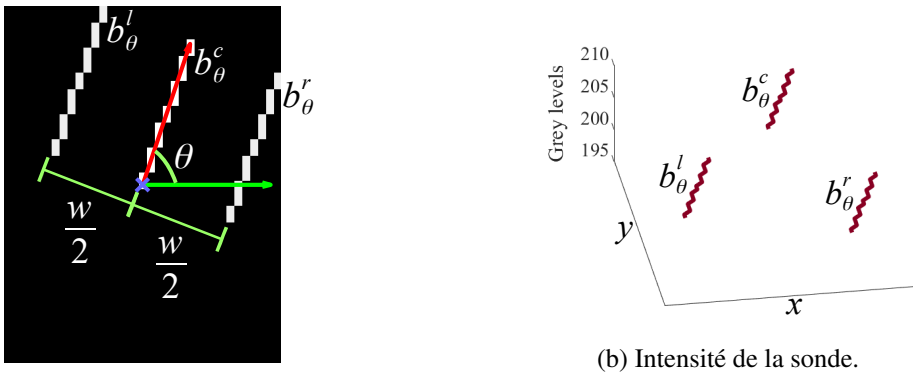
$$E_{b^l_\theta,k} f(x) = \wedge^k_{h \in D_{b^l_\theta}} \{f(x+h) \triangle b^l_\theta(h)\} \triangle \grave{c}_{b_\theta,k} f(x) \tag{4.135}$$

$$= \zeta^{\triangle}_{b^l_\theta,k} f(x) \triangle \grave{c}_{b_\theta,k} f(x) \tag{4.136}$$

$$E_{b^r_\theta,k} f(x) = \wedge^k_{h \in D_{b^r_\theta}} \{f(x+h) \triangle b^r_\theta(h)\} \triangle \grave{c}_{b_\theta,k} f(x) \tag{4.137}$$

$$= \zeta^{\triangle}_{b^r_\theta,k} f(x) \triangle \grave{c}_{b_\theta,k} f(x) \tag{4.138}$$

La carte de constante $\grave{c}_{b_\theta,k} f$ est l'infimum point à point $\wedge$ des cartes de constantes $c_{b^c_\theta} f$, $c_{b^l_\theta,k} f$ et



(a) Domaine de la sonde.

(b) Intensité de la sonde.

FIG. 4.52 – (a) Sonde b 2D d'orientation θ et de largeur w. (b) Le segment central b^c_θ a un intensité plus élevée que celle des deux autres b^l_θ et b^r_θ.

$c_{b^r_\theta,k}f$ pour chaque segment b^c_θ, b^l_θ et b^r_θ de la sonde :

$$\grave{c}_{b_\theta,k}(f) = \wedge\{\, c_{b^c_\theta}f,\ \wedge[c_{b^l_\theta,k}f, c_{b^r_\theta,k}f]\,\}. \tag{4.139}$$

$$= \wedge\{\, \varepsilon^{\triangle}_{b^c_\theta}f,\ \wedge[\zeta^{\triangle}_{b^l_\theta,k}f, \zeta^{\triangle}_{b^r_\theta,k}f]\,\}. \tag{4.140}$$

Comme le segment central b^c_θ doit entrer entièrement dans le relief du vaisseau, l'infimum doit être exact et la carte $c_{b^c_\theta}f = \wedge_{h\in b^c_\theta}\{f(x+h)\triangle b^c_\theta(h)\} = \varepsilon^{\triangle}_{b^c_\theta}f$ est utilisée. Par contre, pour les segments gauche et droit, b^l_θ et b^r_θ, les cartes robustes au bruit $c_{b^l_\theta,k}f = \wedge^k_{h\in b^l_\theta}\{f(x+h)\triangle b^l_\theta(h)\} = \zeta^{\triangle}_{b^l_\theta,k}f$ et $c_{b^r_\theta,k}f = \zeta^{\triangle}_{b^r_\theta,k}f$ sont utilisés. $\zeta^{\triangle}_{b^l_\theta,k}$ est le k^{e} *filtre logarithmiques de minimum* pour la sonde b^l_θ (cf. sous-sous-section « Filtres de rang », p. 144).

La carte du détecteur de bosse selon l'orientation θ, $E_{b_\theta,k}f$, est définie comme le supremum point à point $\vee$ de :

$$E_{b_\theta,k}f = \vee\{E_{b^l_\theta,k}f, E_{b^r_\theta,k}f\}. \tag{4.141}$$

La carte du détecteur de bosse s'exprime comme l'infimum point à point des cartes $E_{b_\theta,k}f$ selon toutes les orientations $\theta \in \Theta$:

$$E_{b,k}f = \wedge\{E_{b_\theta,k}f \mid \theta \in \Theta\}. \tag{4.142}$$

Comme la détection de vaisseaux est un problème multi-échelles, I sondes différentes $\{b_i\}_{i\in[\![1\ldots I]\!]}$, de largeurs $\{w_i\}_i$ et longueur $\{l_i\}_i$ sont utilisées. Les cartes de détecteur de bosse $E_{b_i,k}f$ pour les sondes b_i sont alors combinées par infimum point à point :

$$e^I_{b,k}f = \wedge\{E_{b_i,k}f \mid i \in [\![1\ldots I]\!]\} \tag{4.143}$$

Dans la *carte d'appartenance aux vaisseaux* (i.e. *map of vesselness*) $e^I_{b,k}f$ (Fig. 4.53a), les vaisseaux apparaissent comme des vallées et peuvent être segmentés par un seuil. Pour une meilleure représentation, la *carte d'appartenance aux vaisseaux* peut être normalisée comme il s'ensuit. Comme les valeurs des vaisseaux sont inférieures à la médiane μ de la carte $e^I_{b,k}f$ (Fig. 4.53b), nous définissons une nouvelle carte par :

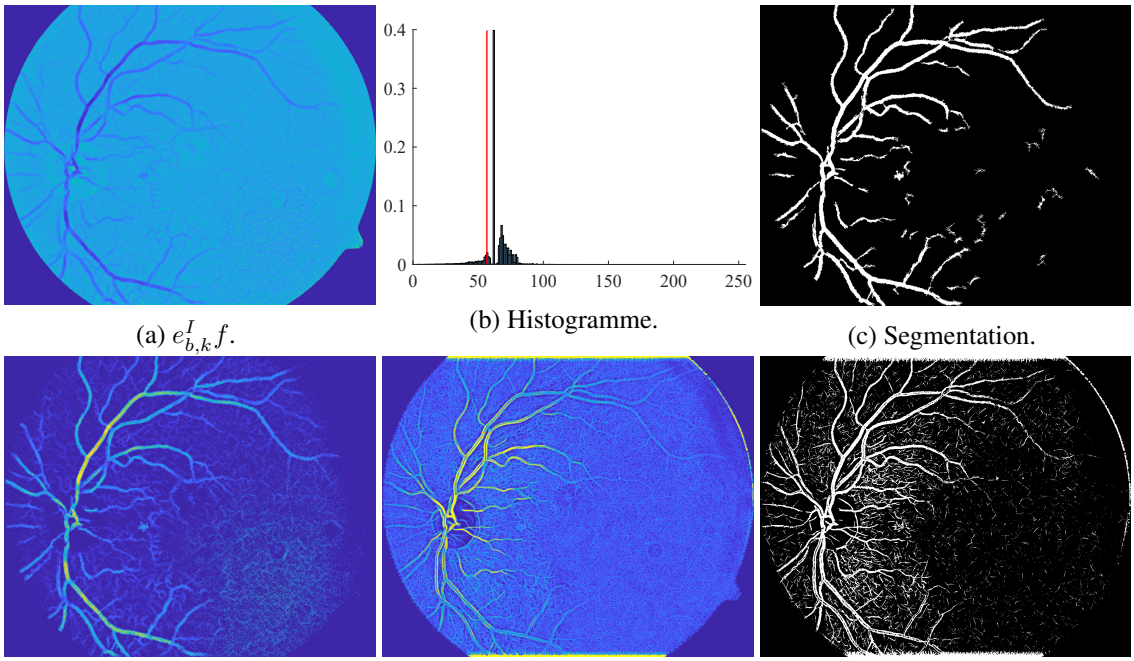
$$\bar{e}_{b,k}f(x) = \begin{cases} e^I_{b,k}f(x) & \text{si } e^I_{b,k}f(x) \le \mu, \\ \mu & \text{sinon.} \end{cases} \tag{4.144}$$

Les valeurs de la carte $\bar{e}_{b,k}$ sont mis à l'échelle dans l'intervalle $[0;1]$ afin de définir la carte normalisée $\Phi_{b,k}f$ (Fig. 4.53d), pour tout $x \in D$, par :

$$\Phi_{b,k}f(x) = 1 - \frac{\bar{e}_{b,k}(x) - \min\{\bar{e}_{b,k}\}}{\max\{\bar{e}_{b,k}\} - \min\{\bar{e}_{b,k}\}}. \tag{4.145}$$

Expériences pour l'estimation des paramètres Afin d'estimer les valeurs des paramètres, des expériences ont été menées sur des images faiblement contrastées de la banque de données DIARETDB1 [151] (Fig. 4.50a) et dans des images fortement contrastées provenant de la banque de données DRIVE [287]. Les images de la banque DIARETDB1 ont été capturées avec un Champ De Vision (CDV) de $50°$ [151], tandis que celle de la banque DRIVE avec un angle du CDV de $45°$. Les paramètres sont normalisés pour être les mêmes pour toutes les images. Chaque paramètre est choisi avec soin afin d'obtenir les meilleurs résultats de segmentation. Une image de DIARETDB1 est utilisée pour une évaluation qualitative, tandis que les images de DRIVE servent à une évaluation quantitative. En effet, cette banque contient 20 images avec une référence donnée par un expert. Les paramètres sont les suivants. Le k^{e} minimum, $\wedge^k$, est choisi tel que 20% des points minimaux d'un ensemble soient écartés. 18 orientations θ entre $0°$ et $360°$ ont été retenues comme suffisantes. Un nombre maximum de 3 sondes b_1, b_2 et b_3 est utilisé. Leurs largeurs sont reliées au diamètre du CDV, D_{CDV}, de l'image et du ratio entre l'angle du CDV de la caméra de référence, $\alpha_r = 45°$, et à l'angle du CDV de la caméra de l'image, α. Comme la largeur w_1 doit être plus grande que le diamètre des vaisseaux les plus larges, la valeur $w_1 = (D_{CDV}/50)(\alpha_r/\alpha)$ est un choix approprié. Les largeurs w_2 et w_3 sont égales à $w_2 = 0{,}75w_1$ et $w_3 = 0{,}5w_1$. Comme les plus petits vaisseaux peuvent être plus tortueux que les plus larges, la longueur l_i d'une sonde doit être plus petite que sa largeur w_i. Nous utiliserons $l_i = 0{,}75w_i$. L'intensité des sondes dépend de la valeur moyenne de l'image m_f. Initialement, l'intensité du segment central est mise à la valeur $h^c_{ref} = 215$ et les intensités des segments gauche et droit à $h^r_{ref} = h^l_{ref} = 225$. Pour chaque image f, l'intensité du segment central est alors égale à $h^c = m_f$ et les intensités des segments gauche et droit à $h^r = h^l = h^r_{ref} \triangleq (h^c \triangleq h^c_{ref})$. La *carte d'appartenance aux vaisseaux* $e^I_{b,k}f$ est segmentée avec un seuil tel que $12\ \%$ de l'aire du CDV soit considérée comme des vaisseaux (Fig. 4.53a). Afin d'éviter la segmentation des zones de bruit, moins de 3 sondes peuvent être utilisées. Le nombre I de sondes est choisi en vérifiant que le nombre de pixels dont la classe change entre les segmentations de la *carte d'appartenance aux vaisseaux* avec I sondes $seg(e^I_{b,k}f)$ et avec la première sonde $seg(e^1_{b,k}f)$ n'excède pas $40\ \%$ de l'aire des vaisseaux de cette dernière segmentation $seg(e^1_{b,k}f)$. La segmentation retenue $seg(e^I_{b,k}f)$ est alors filtrée : les régions dont l'aire est inférieure à $(w_1/2)^2$ sont enlevées et les petits trous de la segmentation des vaisseaux sont remplis (i.e. le complément de l'image segmentée est érodé par un carré unitaire et reconstruit par dilatation sous le complément de l'image segmentée). De plus, pour la *carte d'appartenance aux vaisseaux* $e^I_{b,k}f$ et la carte normalisée $\Phi_{b,k}f$, seules les valeurs qui sont à l'intérieur du masque de CDV sont considérées. Dans la banque DIARETDB1, les masques du CDV sont segmentés par un seuil, tandis qu'ils sont disponibles dans la banque de données DRIVE.

Résultats qualitatifs sur une image faiblement contrastée La *carte d'appartenance aux vaisseaux* $e^I_{b,k}f$ (Fig. 4.53a) est calculée pour l'image de la figure 4.50a. Deux sondes sont automatiquement sélectionnées. On peut noter que le seuil ($12\ \%$ de l'aire du CDV) est en dessous de la médiane (Fig. 4.53b). La segmentation (Fig. 4.53c) est visuellement bonne et permet de détecter des vaisseaux qui sont à peine visibles dans l'image originale (Fig. 4.50a). La carte normalisée $\Phi_{b,k}f$ (Fig. 4.53d) est comparée au détecteur de vaisseaux B-COSFIRE [16] (Fig. 4.53e) dont le code est disponible publiquement. Dans les parties les plus claires de l'image, le filtre B-COSFIRE est très efficace pour trouver les vaisseaux et donne plus de détails que notre méthode. Cependant, comparé à notre méthode et dans les parties sombres, le filtre B-COSFIRE est plus sensible au bruit et améliore moins les vaisseaux. En utilisant la même seuil sur l'aire, sa segmentation détecte beaucoup de bruit en plus des vaisseaux (Fig. 4.53f).



(a) $e^I_{b,k}f$. (b) Histogramme. (c) Segmentation.

(d) $\Phi_{b,k}f$. (e) B-COSFIRE. (f) Segmentation de (e).

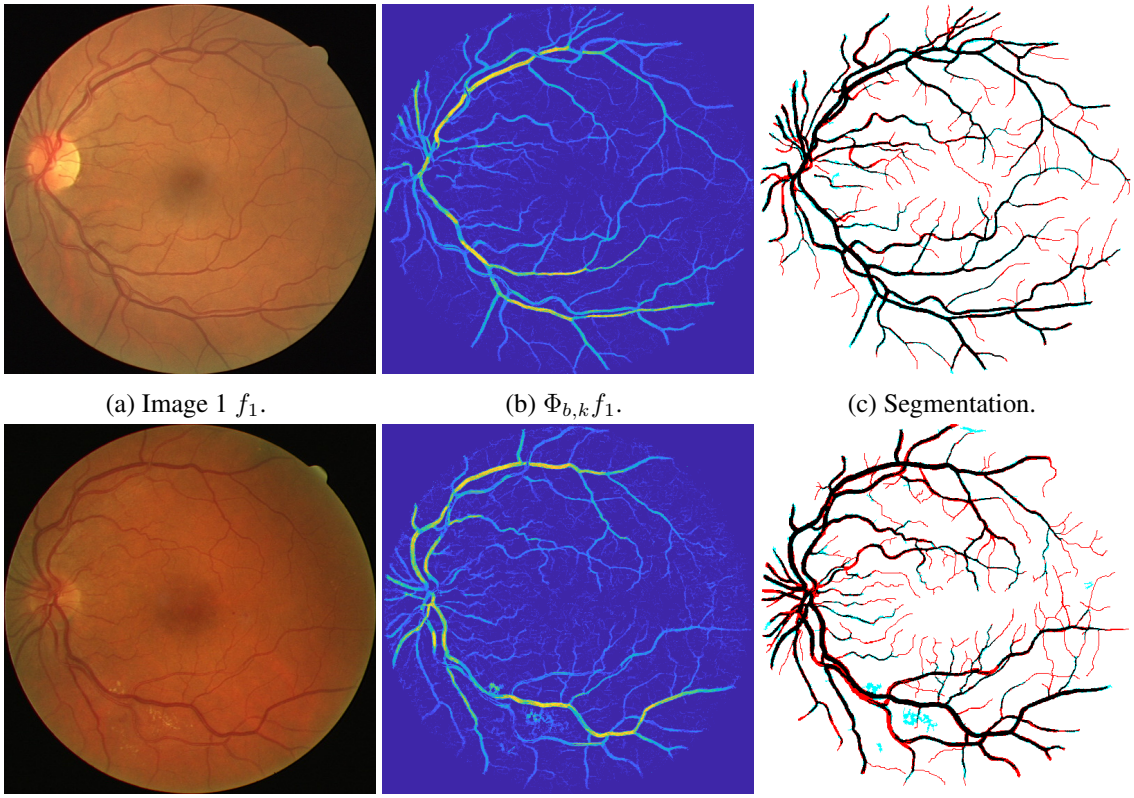
FIG. 4.53 – (a) *Carte d'appartenance aux vaisseaux* $e^I_{b,k}f$. (b) Histogramme de $e^I_{b,k}f$ valeur de seuil en rouge. (c) Segmentation des vaisseaux. (d) *Carte d'appartenance aux vaisseaux* normalisée $\Phi_{b,k}f$. (e) Image filtrée par la méthode B-COSFIRE. (f) Segmentation de (e).

Résultats quantitatifs sur une banque d'images fortement contrastées Dans la banque de données DRIVE, comme une segmentation de référence est disponible, nous comparons les résultats de notre méthode à ceux de la segmentation du 2e expert (fournie avec la banque de données) et à ceux de six autres méthodes de l'état de l'art [16], [122], [182], [287], [338], [339] (Tab. 4.2). Nous utilisons les critères suivants moyennés sur la banque de données : la sensibilité (Se), la spécificité (Sp) et la précision (Acc pour *Accuracy*) [287]. En utilisant le critère de précision, notre méthode est cinquième sur sept méthodes automatiques. Cependant, lorsque l'on prend en compte l'écart-type ; la nôtre, les segmentations du 2e expert et les méthodes [16], [182], [287] sont dans le même intervalle de confiance. Trois méthodes [122], [338], [339] sont au-dessus des autres et du 2e expert. Sur deux images, la figure 4.54 montre que notre méthode est bonne pour trouver les vaisseaux principaux (Fig. 4.54c et 4.54f). Cependant, elle est toujours limitée pour segmenter les plus petits. Dans la partie inférieure de la figure 4.54f, on peut remarquer que des lésions de la rétinopathie comme les exsudats créent des faux positifs (en cyan). En effet, une zone fine entre deux exsudats peut être confondue, par l'algorithme, avec un vaisseau. Ceci sera amélioré dans des travaux futurs. Cependant, ces résultats préliminaires sont encourageants parce que notre méthode est autonome et sans pré-traitement tel que de l'amélioration de contraste [234].

Perspectives Nous avons introduit avec succès une méthode entièrement automatique pour extraire les vaisseaux dans les images couleur de la rétine et qui est adaptative aux variations d'éclairement. Elle est fondée sur une palpation, par-dessous, de la surface topographique d'une image par une sonde à 3 segments. Une différence LIP est alors mesurée localement entre l'image et la sonde. Cette différence correspond à une érosion-logarithmique. La constante nécessaire à la mise en contact de la sonde avec l'image par-dessous est elle aussi obtenue par une érosion-logarithmique ou un filtre de rang sur les minima. Une *carte d'appartenance aux vaisseaux* est alors obtenue sur laquelle les vaisseaux peuvent être extraits par un seuil. Sur une image faiblement contrastée, les résultats ont montré que notre méthode détecte mieux les vaisseaux qu'une

Méthode	Se	Sp	Acc	std
Zhu [339]	0,714 0	0,986 8	0,960 7	0,004 0
Zhao [338]	0,742	0,982	0,954	
Hu [122]	0,777 2	0,979 3	0,953 3	
2e expert	0,776 0	0,972 5	0,947 3	0,004 8
Mendonça [182]	0,734 4	0,976 4	0,946 3	0,006 5
La nôtre	0,735 8	0,976 5	0,945 4	0,006 0
Azzopardi [16]	0,765 5	0,970 4	0,944 2	
Staal [287]			0,944 1	0,006 5

TABLE 4.2 – Moyennes de la sensibilité (Se), spécificité (Sp), précision (Acc) et son écart-type (std) pour différentes méthodes dans la banque de données DRIVE.



(a) Image 1 f_1. (b) $\Phi_{b,k} f_1$. (c) Segmentation.

(d) Image 3 f_3. (e) $\Phi_{b,k} f_3$. (f) Segmentation.

FIG. 4.54 – (a,d) Images rétiniennes. (b,e) Cartes normalisées. (c,f) Comparaison des segmentations avec la référence. Les pixels noirs sont les vrais positifs, les pixels blancs sont les vrais négatifs, les pixels cyans sont les faux positifs et les pixels rouges sont les faux négatifs.

approche de l'état de l'art [16]. Dans une banque d'images fortement contrastée (DRIVE), notre méthode donne des résultats similaires ou meilleurs que trois autres de l'état de l'art [16], [182], [287] et la segmentation manuelle d'un second expert. Trois méthodes [122], [338], [339] sont au-dessus des autres et du second expert. Néanmoins notre méthode est très simple et robuste. Elle est liée à la Morphologie Mathématique Logarithmique et montre son intérêt dans l'analyse des banques d'images de santé publique. Dans le futur, nous rendrons la méthode plus robuste aux lésions. Il serait aussi intéressant de la comparer au détecteur 5-points défini à l'équation (3.25), p. 53.

4.5.2 Critères d'homogénéité de région

Avec Michel Jourlin, j'ai proposé des critères d'homogénéité de régions qui soient robustes aux variations d'éclairement [147], [230]. Ceux-ci sont appliqués à la segmentation par croissance de région.

Motivation

La segmentation des images acquises avec différentes conditions d'éclairement est une tâche difficile en analyse d'image qui peut se produire dans différents cadres tels que l'inspection visuelle pour l'industrie [56], [213], [225], [241], les images médicales [219], [234], la sûreté [81], l'assistance à la conduite [109], etc. Les variations d'éclairement représentent de manière évidente un obstacle conséquent pour le développement d'algorithmes de traitement d'image robustes. Afin de dépasser cette difficulté, nous pouvons retenir les deux approches suivantes.

(i) Premièrement, une approche *pratique* consiste à contourner le problème, e.g. par un processus d'apprentissage statistique. C'est souvent le cas pour la reconnaissance faciale, où une banque de données est créée en faisant varier conjointement l'angle d'acquisition et l'intensité de la source d'éclairement [254]. Une alternative aux processus par apprentissage statistique consiste à extraire des images, des caractéristiques indépendantes des conditions d'éclairement [280], [315]. Une alternative a pour but d'atténuer les effets des variations d'éclairement par le moyen de différentes méthodes comme la suppression d'ombre [335].

(ii) Deuxièmement, une approche *fondamentale* consiste à créer des modèles et des algorithmes presque insensibles à ces variations. A cet égard, Chen et al. [51] ont défini le modèle de *Logarithmic Total Variation* (LTV) afin d'enlever l'illumination variable dans les images de visages. Yu et al. [329] regardent une image dans l'espace à trois dimensions comme une onde grise composée de pics et de creux qui divisent l'image en plusieurs sous-régions locales. Le modèle *Retinex* prend en compte la perception humaine des couleurs et contrastes. Il a été passé en revue de manière extensive par Elad et al. [74] à côté de méthodes apparentées sur la compensation d'illumination. Nous pouvons aussi mentionner les méthodes de représentation des ensembles de niveaux des images par les arbres de leurs composantes connexes, tels que : l'*arbre des composantes* (en anglais *component-tree*), aussi appelé *max-tree* [267], ou l'*arbre des formes*, aussi appelé *arbre d'inclusion* [191]. Ils sont invariants aux changements de contraste causés par une fonction continue et croissante appliquée à l'image [191]. Plus récemment, Passat et al. [242] ont introduit une méthode de segmentation par *arbre des composantes* présentant une certaine robustesse à l'inhomogénéité de l'intensité pourvu que le contraste maximal de l'image soit dans une dynamique intermédiaire : pas trop faible ou pas trop élevée. Le filtrage connexe sur les espaces de formes fondés sur des arbres de Xu et al. [322] est aussi théoriquement invariant à ces changements de contraste.

Néanmoins, seulement quelques papiers se focalisent sur les causes des variations d'éclairement. A cet effet, nous avons introduit dans [147], [230] deux nouveaux critères d'homogénéité dans le cadre de travail LIP, à savoir les critères homogénéités LIP-additif et LIP-multiplicatif. Chaque critère est défini : soit avec la loi d'addition-LIP $\triangle$ de deux images ; soit avec la loi de multiplication-LIP $\triangle$ d'une image par un scalaire. Ces lois sont connues pour modéliser les variations d'éclairement qui affectent l'étape d'acquisition de l'image [136]. En particulier, la loi LIP-additive $\triangle$ modélise les changements d'intensité de la source ou du temps d'exposition de la caméra tandis que la loi LIP-multiplicative $\triangle$ modélise les changements d'épaisseur ou d'opacité de l'objet observé.

Le but de cette partie est d'étudier ces critères d'homogénéité en détail. Au niveau théorique, nous démontrons leur insensibilité aux variations d'éclairement qui sont modélisées par les lois LIP correspondantes. Au niveau applicatif, nous prouvons leur robustesse aux changements d'éclairement à la fois sur des images réelles et simulées. A des fins de segmentation, nous les appliquons dans l'algorithme de Revol et Jourlin [257] qui appartient à la classe des méthodes

par *croissance de régions*. Cet algorithme a la spécificité d'évaluer l'homogénéité d'une région d'une image à chaque étape. Ses résultats de segmentation seront comparés à ceux obtenus par la méthode des *arbres de composantes* de Passat et al. [242].

Critères d'homogénéité LIP-additif et LIP-multiplicatif

Soit $f : D \to [0; M[$ une image à niveaux de gris définie sur le support spatial $D \subset \mathbb{R}^n$, et dont les valeurs sont dans l'échelle de gris $[0; M[$. Considérons une région R du domaine D. Il existe des paramètres classiques évaluant l'homogénéité de l'image f sur la région R, tels que l'écart-type $\sigma_f(R)$ ou le « diamètre » de R défini au sens mathématique par $\sup_{x\in R} f(x) - \inf_{x\in R} f(x)$ et qui n'est rien d'autre que la dynamique de l'image sur la région R. Ces paramètres d'homogénéité sont à l'évidence sensibles aux variations d'éclairement, ce qui justifie l'introduction de nouveaux critères : l'homogénéité LIP-additive et l'homogénéité LIP-multiplicative.

Remarque. Lorsqu'il n'y a pas d'ambiguïté, nous considérerons que l'homogénéité d'une image f sur une région R est équivalente à l'homogénéité de la région R. En effet, la région R est supposée être associée à l'image f.

Homogénéité LIP-additive L'homogénéité $H_f^{\triangle}(R)$ de la région R est définie comme sa dynamique au sens LIP :

$$H_f^{\triangle}(R) = \sup_{x\in R} f(x) \triangle \inf_{x\in R} f(x). \tag{4.146}$$

Remarquons que $H_f^{\triangle}(R)$ appartient à l'intervalle $[0; M[$ et représente donc un niveau de gris qui peut être interprété comme le *Contraste Logarithmique Additif* maximal observé sur la région [143].

La plus importante propriété de ce critère d'homogénéité est son insensibilité aux variations du temps d'exposition de la caméra qui sont modélisées par l'addition/soustraction-LIP d'une constante $C \in]-\infty; M[$:

$$H_{f\triangle C}^{\triangle}(R) = H_f^{\triangle}(R). \tag{4.147}$$

Démonstration. Remarquons tout d'abord que

$$\sup_{x\in R} \{(f \triangle C)(x)\} = \sup_{x\in R} \{f(x)\} \triangle C.$$

En fait, l'équation (4.3) donne :

$$\begin{aligned}
&\sup\{f\} \triangle C = \sup\{f\} + C - (C \cdot \sup\{f\})/M \\
&= (1 - C/M)\sup\{f\} + C = \sup\{f\,(1 - C/M)\} + C \\
&= \sup\{f\,(1 - C/M) + C\} = \sup\{f + C - f \cdot C/M\} \\
&= \sup\{f \triangle C\}.
\end{aligned}$$

De la même manière, nous avons : $\inf_{x\in R}\{(f \triangle C)(x)\} = \inf_{x\in R}\{f(x)\} \triangle C$. Dans ces conditions, l'équation (4.146) peut être écrite :

$$H_{f\triangle C}^{\triangle}(R) = (\sup_{x\in R}\{f(x)\} \triangle C) \triangle (\inf_{x\in R}\{f(x)\} \triangle C).$$

Le calcul suivant montre que $(A \triangle C) \triangle (B \triangle C)$ est égale à $A \triangle B$:

$$\begin{aligned}
(A \triangle C) \triangle (B \triangle C) &= \frac{A + C - \frac{A.C}{M} - \left(B + C - \frac{B.C}{M}\right)}{1 - \frac{B+C-\frac{B.C}{M}}{M}} \\
&= \frac{(A - B)\left(1 - \frac{C}{M}\right)}{1 - \frac{B}{M} - \frac{C}{M} - \frac{B.C}{M^2}} = \frac{(A - B)\left(1 - \frac{C}{M}\right)}{\left(1 - \frac{B}{M}\right)\left(1 - \frac{C}{M}\right)} = \frac{(A - B)}{\left(1 - \frac{B}{M}\right)} \\
&= A \triangle B.
\end{aligned}$$

Les deux résultats précédents démontrent l'équation (4.147). □

Homogénéité LIP-multiplicative Maintenant, nous proposons un critère d'homogénéité fondé sur la multiplication-LIP par un scalaire et plus précisément sur la notion de *Contraste Logarithmique Multiplicatif*, noté par son acronyme anglais LMC pour *Logarithmic Multiplicative Contrast* [146]. Rappelons que le LMC, associé à une paire de niveaux de gris g_1 et g_2, où g_1 et g_2 sont strictement positifs, est le nombre réel défini par la formule :

$$LMC(g_1, g_2) \triangle \min(g_1, g_2) = \max(g_1, g_2). \tag{4.148}$$

Dans le contexte d'images acquises en transmission, cela signifie que le LMC de deux niveaux de gris représente le nombre de fois que le plus clair d'entre eux doit être empilé sur lui-même afin d'obtenir le plus sombre. Dans ces conditions, l'homogénéité LIP-multiplicative d'une région R est calculée selon :

$$H_f^{\triangle}(R) = LMC(\sup_{x\in R}\{f(x)\}, \inf_{x\in R}\{f(x)\}) \tag{4.149}$$

$$= \frac{\ln(1 - \sup_{x\in R}\{f(x)\}/M)}{\ln(1 - \inf_{x\in R}\{f(x)\}/M)}. \tag{4.150}$$

Remarque. Dans le cas où $\inf_{x\in R} f(x) = 0$, on peut remplacer cette valeur par 1 afin d'éviter une valeur infinie pour $H_f^{\triangle}(R)$.

Ce critère d'homogénéité possède la propriété fondamentale d'être insensible aux variations d'éclairement modélisées par la loi LIP-multiplicative $\triangle$, telles que l'opacité ou les variations d'épaisseur :

$$H_{\lambda\triangle f}^{\triangle}(R) = H_f^{\triangle}(R). \tag{4.151}$$

Démonstration. Supposons que $H_f^{\triangle}(R) = \mu$, ce qui signifie que $\sup_{x\in R} f(x) = \mu \triangle \inf_{x\in R} f(x)$. Nous devons démontrer l'égalité :

$$\sup_{x\in R} \lambda \triangle f(x) = \mu \triangle \inf_{x\in R} \lambda \triangle f(x). \tag{4.152}$$

Etant donné $\lambda \geq 0$, il est facile d'établir que :

$$\begin{aligned}
\sup_{x\in R}\{\lambda \triangle f(x)\} &= \sup_{x\in R}\{M(1 - (1 - f(x)/M)^{\lambda})\} \\
&= M(1 - \inf_{x\in R}\{(1 - f(x)/M)^{\lambda}\}) \\
&= M(1 - (\inf_{x\in R}\{1 - f(x)/M\})^{\lambda}) \\
&= M(1 - (1 - \sup_{x\in R}\{f(x)\}/M)^{\lambda}) \\
&= \lambda \triangle \sup_{x\in R} f(x).
\end{aligned}$$

De la même manière, $\inf_{x\in R} \lambda \triangle f(x) = \lambda \triangle \inf_{x\in R} f(x)$. Par conséquent, l'équation (4.152) est établie ; ce qui termine la démonstration. □

Un algorithme par croissance de régions dépendant de leur homogénéité

Revol et Jourlin [257] ont présenté une méthode par croissance de régions qui minime la variance des régions. Dans [230], nous avons proposé de remplacer la variance par un des critères d'homogénéité $H^{\triangle}$ ou $H^{\triangle}$. Présentons les principes de cet algorithme. Les détails sont donnés dans [257].

Soit $dyn_f(R)$ la dynamique de l'image f sur la région R. Elle est définie comme l'intervalle

$$dyn_f(R) = \left[\min_{x\in R}\{f(x)\}, \max_{x\in R}\{f(x)\}\right]. \tag{4.153}$$

A chaque étape n, une région R_n est dilatée par un élément structurant unitaire N, e.g. un carré de côté 3 pixels. Cela donne une région $D_{n+1} = R_n \oplus N$ dont l'homogénéité $H_f(D_{n+1})$ est estimée. Selon une valeur de seuil t, la région est considérée comme homogène ou pas.

1. Si $H_f(D_{n+1}) \leq t$, la région D_{n+1} est alors considérée comme homogène. La nouvelle région R_{n+1} devient D_{n+1}.
2. Si $H_f(D_{n+1}) > t$, la région D_{n+1} est alors inhomogène. Un processus de *réduction* $Rdct_f(D_{n+1})$ tronque les classes extrémales de l'histogramme de D_{n+1} en gardant uniquement les pixels appartenant à la dynamique de R_n, $dyn_f(Rdct_f(D_{n+1})) = dyn_f(R_n)$. Cela donne une région D'_{n+1}.
   - 2.a Si $H_f(D'_{n+1}) \leq t$, la région D'_{n+1} est alors modifiée par un processus d'*extension* qui ajoute les pixels voisins à la région selon les deux conditions suivantes. i) La nouvelle valeur de pixels a une différence-LIP inférieure ou égale à 1 avec le maximum ou le minimum de f sur D'_{n+1}. ii) Les nouveaux pixels donnent une région homogène. La nouvelle région R_{n+1} devient l'extension de D'_{n+1} : $E(D'_{n+1})$.
   - 2.b Si $H_f(D'_{n+1}) > t$, un processus de *contraction* réduit la dynamique de la région D'_{n+1} en supprimant les classes extrémales de son histogramme jusqu'à ce que la région devienne homogène. L'addition-LIP et la soustraction-LIP remplacent l'addition et la soustraction usuelles utilisées dans [257]. La nouvelle région R_{n+1} est la contraction de D'_{n+1} : $Ctrct(D'_{n+1})$.

L'algorithme s'arrête lorsque la région R_{n+1} ne croît plus.

Afin d'obtenir une méthode de segmentation entièrement automatique, les points germes (i.e. de départ) peuvent être choisis comme les minima d'une fonction comme le gradient morphologique d'une image – comme pour la transformée par Ligne de Partage des Eaux [32] – ou une carte des distances d'Asplund de l'image. Le point germe peut être automatiquement déterminé par une classification précédente supervisée [219], [222] ou non supervisée [220]. Comme le critère d'homogénéité est invariant aux changements de contraste avec une cause optique, le seuil peut être appris à partir d'un ensemble d'entraînement [167]. Il peut aussi être automatiquement choisi comme le contraste le plus fréquent sur tous les voisinages de pixels [139].

Une méthode de l'état de l'art qui utilise l'arbre des composantes de l'image

Notre méthode est comparée à celle de Passat et al. [242] qui repose sur les ensembles de niveaux et présente une certaine robustesse aux variations d'éclairement. Cette dernière consiste en une segmentation interactive de l'*arbre des composantes*. L'*arbre des composantes* est l'arbre formé par la relation d'inclusion entre les composantes binaires des ensembles de niveau successifs de l'image. Un nœud N de l'*arbre des composantes* est une composante connexe de l'ensemble de niveau $X_v(f) = \{x \in D \mid f(x) \geq v\}$ de l'image. Soit $\mathcal{K}$ l'ensemble de toutes les composantes connexes de tous les ensembles de niveaux de l'image. Le sous-ensemble $\mathcal{K}' \subset \mathcal{K}$ des composantes connexes qui ajuste au mieux la région G, sélectionnée par l'utilisateur dans le domaine D, est automatiquement déterminé par le problème de minimisation suivant :

$$\widehat{\mathcal{K}} = \arg\min_{\mathcal{K}'\in\mathcal{P}(\mathcal{K})}\{d^\alpha(\cup_{N\in\mathcal{K}'}N, G)\},$$

où $\mathcal{P}(\mathcal{K})$ est l'ensemble des parties de $\mathcal{K}$. La pseudodistance d^α est définie par

$$d^\alpha(X,Y) = \alpha|X \setminus Y| + (1-\alpha)|Y \setminus X|,$$

où $\alpha \in [0;1]$. $d^0(X,Y) = |Y \setminus X|$ et $d^1(X,Y) = |X \setminus Y|$ sont respectivement la quantité de faux négatifs et de faux positifs dans X par rapport à Y. Une solution à ce problème de minimisation peut être trouvée par programmation dynamique. Les détails sont donnés dans [242]. Dans [230], nous avons utilisé le code disponible publiquement de Naegel et Passat [199].

Validation expérimentale

Nous venons d'établir que le critère d'homogénéité LIP-additif, $H^{\triangle\!\!\!+}$, est théoriquement insensible à une addition-LIP $\triangle\!\!\!+$ ou à une soustraction-LIP $\triangle\!\!\!-$ d'une constante d'une image, respectivement. Ces opérations modélisent les changements d'éclairement causés par une variation de l'intensité de la source ou du temps d'exposition de la caméra. Comme pour le critère d'homogénéité LIP-multiplicatif, $H^{\triangle\!\!\!\times}$, il est insensible à une multiplication-LIP d'une image f par un scalaire qui modélise les variations de l'épaisseur de l'objet ou de son opacité. Effectuons une validation expérimentale de ces insensibilités sur des images simulées et réelles. Comparons aussi l'approche présentée à celle de l'état de l'art de Passat et al. [242].

Images simulées Pour cette expérience, nous avons sélectionnés une image couleur d'un papillon (Fig. 4.55a) dans l'ensemble de données YFCC100M (*Yahoo Flickr Creative Commons 100M*) [299], [327]. Nous avons extrait sa luminance comme une image f à niveaux de gris (Fig. 4.55b). Afin de vérifier les insensibilités respectives des critères $H^{\triangle\!\!\!+}$ et $H^{\triangle\!\!\!\times}$ aux variations d'éclairement simulées par une addition-LIP $\triangle\!\!\!+$ ou une multiplication-LIP $\triangle\!\!\!\times$, nous avons assombri et éclairci cette image de luminance f en utilisant chacune des opérations LIP.

Pour le critère d'homogénéité LIP-additif, $H^{\triangle\!\!\!+}$, une image sombre $f_{dk,a}$ est obtenue par addition-LIP d'une constante $k = 120$ à l'image f, $f_{dk,a} = f \triangle\!\!\!+ k$ (Fig. 4.55c). Une image claire $f_{br,a}$ est aussi simulée par la soustraction-LIP de la même constante : $f_{br,a} = f \triangle\!\!\!- k$ (Fig. 4.55d).

Pour le critère d'homogénéité LIP-multiplicatif, $H^{\triangle\!\!\!\times}$, une image sombre $f_{dk,m}$ est obtenue par une multiplication-LIP par un scalaire $\lambda_{dr} = 4$ de l'image f, $f_{dk,m} = \lambda_{dr} \triangle\!\!\!\times f$ (Fig. 4.55e). Une image claire $f_{br,m}$ est aussi simulée par une multiplication-LIP par un scalaire $\lambda_{br} = 0{,}1$, $f_{br,m} = \lambda_{br} \triangle\!\!\!\times f$ (Fig. 4.55f).

Dans ces cinq images f, $f_{dk,a}$ $f_{br,a}$, $f_{dk,m}$ et $f_{br,m}$ la même région R est sélectionnée. Les critères d'homogénéité $H^{\triangle\!\!\!+}$ et $H^{\triangle\!\!\!\times}$ sont alors calculés sur la région R de ces images f, $f_{dk,a}$, $f_{br,a}$ et des images f, $f_{dk,m}$, $f_{br,m}$, respectivement. Leurs valeurs sont données dans la sous-sous-section « Résultats » (p. 168).

Images réelles La robustesse aux variations d'éclairement des critères d'homogénéité LIP-additif $H^{\triangle\!\!\!+}$ et LIP-multiplicatif $H^{\triangle\!\!\!\times}$ a besoin d'être vérifiée sur des images réelles en utilisant du matériel expérimental spécifique.

Le critère d'homogénéité LIP-additif $H^{\triangle\!\!\!+}$ est présumé avoir une faible sensibilité aux variations de l'intensité de la source d'éclairement ou au temps d'exposition de la caméra. Afin de vérifier cette hypothèse, nous avons conduit une expérience. Une image de la même scène est acquise par une caméra avec des temps d'exposition significativement différents et des poses légèrement différentes. La scène se compose d'un monstre en peluche appelé « Nessie » (Fig. 4.56). Trois images couleur sont capturées avec des temps d'exposition différents 1/40 s (Fig. 4.56a et 4.56b), 1/400 s (Fig. 4.56c et 4.56d) et 1/800 s (Fig. 4.56c et 4.56e). Elles sont converties en image de luminance à niveaux de gris : f_0, f_1 et f_2, respectivement. Plus le temps d'exposition est court, plus l'image devient sombre. Une segmentation est effectuée par l'algorithme de Revol et al. [257] en utilisant le critère d'homogénéité LIP-additif $H^{\triangle\!\!\!+}$ et un seuil de 200 pour décider si une région R est homogène (i.e. $H_f^{\triangle\!\!\!+}(R) \leq 200$) ou pas. L'initialisation de l'algorithme se fait par un point germe, sélectionné manuellement à l'intérieur des lettres situées sur le corps de « Nessie ». La segmentation est effectuée sur les trois images de luminance f_0, f_1 et f_2. Trois régions segmentées R_0 (Fig. 4.56a), R_1 (Fig. 4.56c) et R_2 (Fig. 4.56e) sont ainsi obtenues dans les images f_0, f_1

et f_2, respectivement. L'homogénéité LIP-additive de ces régions est alors calculée. A des fins de comparaison, la méthode de segmentation de Passat et al. [242] est appliquée aux images f_0 (Fig. 4.56b), f_1 (Fig. 4.56d) et f_2 (Fig. 4.56f). Les régions sélectionnées par l'utilisateur G sont constituées des mêmes points germes qui sont utilisés pour notre approche après une dilatation par un carré de 3 pixels de côté. Comme les régions sélectionnées par l'utilisateur G sont plus petites que les lettres que nous voulons segmenter, nous choisissons un paramètre $\alpha = 0$ afin de minimiser le nombre de faux-négatifs $d^0(\cup_{N\in\mathcal{K}'} N, G)$ dans ces régions G. Après minimisation, ce nombre est égal à 0.

Afin de vérifier la faible sensibilité à l'opacité de l'objet du critère d'homogénéité LIP-multiplicatif $H^{\triangle}$, nous avons sélectionné trois images d'un bagage acquises par rayons X, sous différents angles d'exposition, à des fins de sûreté (Fig. 4.57). Les images proviennent d'une banque de données disponible publiquement et appelée GDXray [84], [184]. L'objet à détecter est une lame de rasoir. Selon son orientation, les rayons X passent à travers différentes épaisseurs de matière et sont donc plus ou moins absorbés. Dans la première image f_0 (Fig. 4.57a), la lame de rasoir est perpendiculaire aux rayons X provenant de la source. L'image de la lame de rasoir est évidemment plus claire dans cette image f_0 que dans les deux autres f_1 (Fig. 4.57c) et f_2 (Fig. 4.57e) où elle est oblique par rapport aux rayons X. Pour détecter cet objet, l'approche de segmentation de Revol et al. [257] est utilisée avec un seuil de 2,7 sur le critère d'homogénéité LIP-multiplicatif $H^{\triangle}$ d'une région. L'algorithme est initialisé par des points germes manuellement sélectionnés à l'intérieur de la lame de rasoir (Fig. 4.57a, 4.57c et 4.57e). Trois régions R_0 (Fig. 4.57b), R_1 (Fig. 4.57c) et R_2 (Fig. 4.57e) correspondant à l'objet sont alors obtenues dans les images f_0, f_1 et f_2, respectivement. L'homogénéité LIP-multiplicative de ces régions est alors estimée. Ses valeurs sont présentées dans la sous-sous-section suivante.

Résultats

Les résultats expérimentaux sont donnés pour les images simulées et réelles.

Images simulées Le critère d'homogénéité LIP-additif $H^{\triangle}$ de la région R est calculé pour l'image f (Fig. 4.55b), sa version assombrie $f_{dk,a}$ par addition-LIP (Fig. 4.55c) et sa version éclaircie $f_{br,a}$ par soustraction-LIP (Fig. 4.55d). Pour chaque image, le critère $H^{\triangle}$ de la région R est égal à :

$$H_f^{\triangle}(R) = H_{f_{dk,a}}^{\triangle}(R) = H_{f_{br,a}}^{\triangle}(R) = 150{,}7.$$

Ce résultat vérifie ainsi l'insensibilité du critère d'homogénéité LIP-additif $H^{\triangle}$ à l'addition-LIP $\triangle$ ou à la soustraction-LIP $\triangle$ d'une constante d'une image, respectivement.

Le critère d'homogénéité LIP-multiplicatif $H^{\triangle}$ est aussi estimé dans l'image f (Fig. 4.55b) et dans ses versions assombrie et éclaircie par multiplication-LIP $f_{dk,m}$ (Fig. 4.55e) et $f_{br,m}$ (Fig. 4.55f), respectivement. Pour chaque image, le critère $H^{\triangle}$ de la région R est égal à :

$$H_f^{\triangle}(R) = H_{f_{dk,m}}^{\triangle}(R) = H_{f_{br,m}}^{\triangle}(R) = 3{,}33.$$

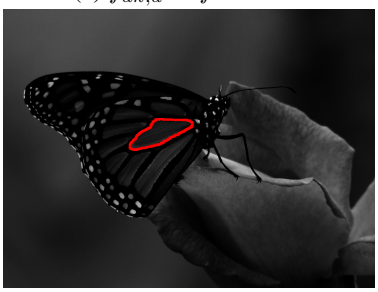
Ce résultat montre l'insensibilité du critère LIP-multiplicatif $H^{\triangle}$ à la multiplication-LIP $\triangle$ d'une image par un scalaire.

(a) Image couleur.

(b) Luminance f.

(c) $f_{dk,a} = f \triangle 120$.

(d) $f_{br,a} = f \triangle 120$.



(e) $f_{dk,m} = 4 \triangle f$.

(f) $f_{br,m} = 0{,}1 \triangle f$.

FIG. 4.55 – (a) Image couleur d'un papillon et (b) son image de luminance f avec une région R en rouge. (c) Image assombrie $f_{dk,a}$ obtenue par addition-LIP d'une constante $k = 120$. (d) Image éclaircie $f_{br,a}$ obtenue par soustraction-LIP de k. (e) Image assombrie $f_{dk,m}$ obtenue par multiplication-LIP de f par $\lambda_{dr} = 4$. (f) Image éclaircie $f_{br,m}$ obtenue par une multiplication-LIP par $\lambda_{br} = 0{,}1$.

Images réelles Donnons maintenant les résultats obtenus sur les images réelles des figures 4.56 et 4.57.

Le critère d'homogénéité LIP-additif $H^{\triangle}$ est illustré sur la figure 4.56. On peut remarquer que les segmentations des lettres sur le corps de « Nessie » sont très similaires dans les trois images f_0 (Fig. 4.56a), f_1 (Fig. 4.56c) et f_2 (Fig. 4.56e) en utilisant ce critère $H^{\triangle}$ dans la méthode de Revol et al. [257] (i.e. notre approche). En outre, les valeurs du critère d'homogénéité sont les suivantes :

$$H^{\triangle}_{f_0}(R_0) = 198{,}5,$$
$$H^{\triangle}_{f_1}(R_1) = 197{,}5 \text{ et}$$
$$H^{\triangle}_{f_2}(R_2) = 197{,}0.$$

Ces valeurs sont très proches pour les trois régions R_0, R_1 et R_2 de ces images qui ont été acquises avec des temps d'exposition significativement différents. Ces résultats démontrent donc la robustesse du critère d'homogénéité LIP-additif $H^{\triangle}$ dans la méthode de segmentation de Revol et al. [257] aux variations du temps d'exposition de la caméra. Cette approche est aussi comparée à celle de Passat et al. [242]. On peut remarquer que les segmentations avec cette dernière approche sont

différentes entre les trois images f_0 (Fig. 4.56b), f_1 (Fig. 4.56d) et f_2 (Fig. 4.56f). Ces résultats montrent que la méthode de Passat et al. [242] manque de robustesse à des variations importantes du temps d'exposition de la caméra, tandis que la nôtre est robuste à ces variations. En outre, les lettres ne sont pas entièrement extraites avec la méthode de Passat et al. [242] tandis qu'avec la nôtre, leurs segmentations sont meilleures (Fig. 4.56a, 4.56c et 4.56e).

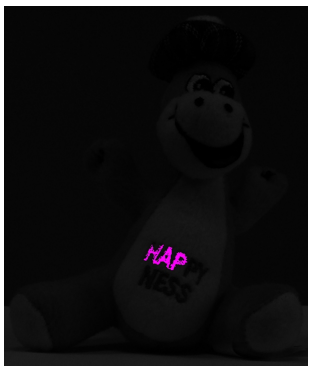
(a) f_0, $H^{\triangle}_{f_0}(R_0) = 198{,}5$.

(b) f_0, Passat et al. [242].

(c) f_1, $H^{\triangle}_{f_1}(R_1) = 197{,}5$.

(d) f_1, Passat et al. [242].

(e) f_2, $H^{\triangle}_{f_2}(R_2) = 197{,}0$.



(f) f_2, Passat et al. [242].

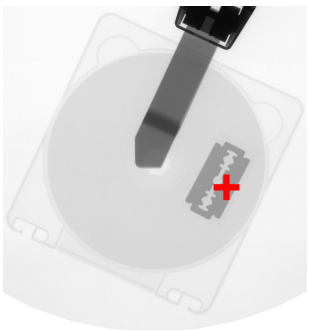
FIG. 4.56 – Comparaison de la robustesse à trois variations du temps d'exposition (a, c, e) de la méthode de Revol et al. [257] avec un critère LIP-additif d'homogénéité $H^{\triangle}$ et (b, d, f) de la méthode de Passat et al. [242]. Versions couleur des images de luminance (a, b) f_0, (c, d) f_1 et (e, f) f_2 acquises à 1/40 s, 1/400 s et 1/800 s, respectivement. Les régions segmentées sont peintes en couleur magenta. Les valeurs du critère d'homogénéité $H^{\triangle}_{f_0}(R_0)$, $H^{\triangle}_{f_1}(R_1)$, $H^{\triangle}_{f_2}(R_2)$ des régions R_0, R_1 et R_2 sont données sous les images (a), (c) et (e), respectivement.

La figure 4.57 présente les résultats obtenus pour le critère LIP-multiplicatif $H^{\triangle}$. Les segmentations de la lame de rasoir sont très similaires dans ces trois images (Fig. 4.57b, 4.57d et 4.57f) en utilisant ce critère $H^{\triangle}$ dans la méthode de Revol et al. [257]. Comme les trois images ont été

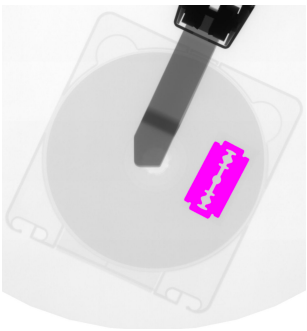
capturées avec différents angles d'exposition, les rayons X sont absorbés différemment par l'objet. En outre, les valeurs du critère d'homogénéité sont les suivantes pour chaque région segmentée :

$$H_{f_0}^{\triangle}(R_0) = 2{,}11,$$
$$H_{f_1}^{\triangle}(R_1) = 2{,}50 \text{ et}$$
$$H_{f_2}^{\triangle}(R_2) = 1{,}94.$$

On peut remarquer que les valeurs du critère sont d'amplitudes similaires pour les trois régions d'un objet présentant différentes épaisseurs qui absorbent les rayons X. Cette expérience illustre la robustesse du critère LIP-multiplicatif $H^{\triangle}$, dans la méthode de segmentation de Revol et al. [257], à différentes opacités (ou absorption) de l'objet.

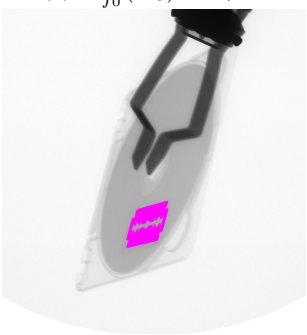


(a) f_0.

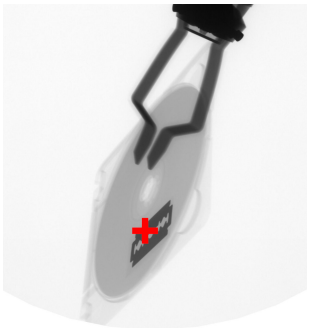


(b) $H_{f_0}^{\triangle}(R_0) = 2{,}11$.

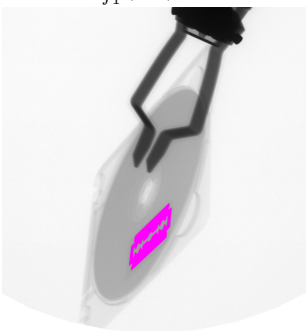
(c) f_1.



(d) $H_{f_1}^{\triangle}(R_1) = 2{,}50$.



(e) f_2



(f) $H_{f_2}^{\triangle}(R_2) = 1{,}94$.

FIG. 4.57 – Robustesse aux variations d'opacité de l'objet, de la segmentation par critère d'homogénéité LIP-multiplicatif $H^{\triangle}$. (a) (b) et (c) Trois images f_0, f_1 et f_2 acquises pour différents temps d'exposition d'un bagage contenant une lame de rasoir. Les points germes sont indiqués par une croix rouge. (d) (e) et (f) Les régions segmentées R_0, R_1 et R_2 des images f_0, f_1 et f_2, respectivement, sont peintes en couleur magenta. Les valeurs du critère d'homogénéité $H_{f_0}^{\triangle}(R_0)$, $H_{f_1}^{\triangle}(R_1)$, $H_{f_2}^{\triangle}(R_2)$ sont données sous les images (b), (d) et (f).

Discussion

Les résultats obtenus avec les images simulées montrent que le critère d'homogénéité LIP-additif $H^{\triangle}$ est insensible à l'addition-LIP $\triangle$ ou à la soustraction-LIP $\triangle$ d'une constante de l'image f. Ces opérations $\triangle$ et $\triangle$ modélisent un assombrissement ou un éclaircissement de l'image par une variation de l'intensité de la source d'éclairement ou du temps d'exposition de la caméra [136]. Ce résultat démontre l'insensibilité théorique du critère LIP-additif $H^{\triangle}$ à ces variations. De façon similaire, le critère d'homogénéité LIP-multiplicatif $H^{\triangle}$ est théoriquement insensible aux variations d'opacité ou d'épaisseur de l'objet qui sont modélisées par une multiplication-LIP $\triangle$ de l'image par un scalaire [136].

Ces simulations sont confirmées par les résultats expérimentaux. La figure 4.56 démontre la robustesse du critère d'homogénéité LIP-additif $H_f^{\triangle}(R)$ aux variations d'éclairement causées par une variation du temps d'exposition de la caméra. Cette variation d'éclairement est équivalente à une variation de l'intensité de la source. Bien que les méthodes fondées sur les ensembles de niveaux soient théoriquement robustes à un changement de contraste causé par une fonction continue et croissante appliquée à l'image, les résultats obtenus sur la figure 4.56 prouvent que l'approche par ensembles de niveaux de Passat et al. [242] manque de robustesse pour des variations importantes du temps d'exposition de la caméra. Les résultats montrés dans la figure 4.57 prouvent la robustesse du critère d'homogénéité LIP-multiplicatif $H_f^{\triangle}(R)$ aux variations d'éclairement causées par des changements d'opacité de l'objet.

Les résultats précédents montrent que les méthodes par ensembles de niveaux utilisant les *arbres de composantes* [242], [267] ou les *arbres d'inclusion* [191], [322] renforceront leur robustesse aux forts changements de contraste – dus à une cause optique – en utilisant un critère d'homogénéité LIP. La cause optique peut être soit une variation de l'intensité de la source, qui sera corrigée par le critère LIP-additif $H^{\triangle}$, soit une variation de l'opacité de l'objet, qui sera corrigée par le critère LIP-multiplicatif $H^{\triangle}$. Ces découvertes seront étudiées dans de prochains travaux.

Perspectives

Avec Michel Jourlin, nous avons donc introduit dans [147], [230] deux nouveaux critères d'homogénéité de régions qui sont robustes aux changements d'éclairement, à savoir le critère d'homogénéité LIP-additif $H^{\triangle}$ et le critère d'homogénéité LIP-multiplicatif $H^{\triangle}$. Grâce à des expériences sur des images simulées et réelles, nous avons montré que le critère d'homogénéité LIP-additif $H^{\triangle}$ est robuste aux changements causés par des variations de l'intensité de la source (ou du temps d'exposition de la caméra), tandis que le critère LIP-multiplicatif $H^{\triangle}$ est robuste aux changements dus aux variations d'opacité (ou d'épaisseur) de l'objet. L'introduction de ces critères dans la méthode de segmentation de Revol et al. [257] lui donne la même robustesse. Comparé à la méthode de Passat et al. [242] fondée sur l'*arbre des composantes* des ensembles de niveaux de l'image, la nôtre est plus robuste à de forts changements d'intensité dus à l'une des causes optiques précédentes. Pour autant que nous le sachions, c'est la première fois qu'une méthode de segmentation utilisant un paramètre d'homogénéité de région, robuste à ces types de changements d'éclairement, a été présentée. Cette nouvelle méthode de segmentation ouvre la voie à de nombreuses applications présentant de fortes variations d'éclairement.

4.6 Percolation

Dans cette section, la méthode de *percolation* sera d'abord présentée avant d'être appliquée aux images de pneumatiques

4.6.1 Percolation robuste aux variations d'intensité lumineuse

Ce travail a été fait en lien avec Michel Jourlin.

Motivations

La *percolation* nécessite une direction favorite et une *fonction de coût* [139] qui guide la trajectoire des pixels dans l'image comme il s'ensuit : un pixel bouge vers la position suivante qui minimise la fonction de coût. La direction favorisée est utilisée au moins pour initialiser le processus mais peut aussi servir à restreindre le nombre de positions possibles pour le déplacement du pixel.

Au niveau mathématique, la théorie de la percolation a été introduite par Broadbent et Hammersley [39] et étudiée en profondeur par Babalievski [17]. Cependant, en traitement d'image, la percolation n'a été utilisée que dans quelques travaux tels que : la segmentation [123], la caractérisation de texture [125], [180] ou la détection de fissures. Dans ce dernier cas, Yamaguchi et al. [323], [324] effectuent une percolation par une méthode de croissance de région tandis que Puntous et al. [249], [250] suivent les courbes des fissures par le moyen d'un contraste perceptuel logarithmique. Dans [180], la percolation a été comparée au temps géodésique défini par Soille dans [283]. Cependant, pour les temps géodésiques, les plus courts chemins sont calculés d'un point à tous les autres ; tandis que pour la percolation, seuls un nombre très réduit de chemins est exploré. Une autre méthode importante de propagation géodésique est le *Fast Marching* de Sethian [278]. Dans [224], j'ai proposé un calcul efficace de toutes les paires de distances géodésiques d'un ensemble de points d'une image hyperspectrale. Dans [47], [243], la tortuosité des chemins géodésiques est calculée. Cependant, ces travaux ont été dédiés à l'analyse des images à niveaux de gris. En outre, les variations d'éclairement constituent un obstacle notable pour le développement d'algorithmes de percolation fiables.

Le but de cette partie est de résoudre ce problème en introduisant une méthode de percolation pour les images couleur qui soit robuste aux variations d'intensité lumineuse. Dans un algorithme de percolation, nous ajouterons : une « tête de guidage » avec des fonctions coût en couleur qui possèdent cette propriété de robustesse. Celle-ci résultera de la définition des fonctions de coût dans le cadre du Traitement d'Image Logarithmique (LIP). Deux cas seront considérés.

i) Premièrement, une approche *marginale* (i.e. composante couleur par composante) combinera le Contraste Logarithmique Additif (en anglais *Logarithmic Additive Contrast* - LAC) de chaque composante couleur.

ii) Deuxièmement, une approche *vectorielle* (toutes les composantes ensemble) utilisera une version couleur du LAC, le CLAC. Elle sera définie dans le cadre de travail LIPC pour les images couleur.

Percolation couleur

Le contraste entre deux pixels $f(x)$ et $f(y)$ d'une image à niveaux de gris $f : D \to [0; M[$ peut être estimé par le Contraste Logarithmique Additif (en anglais *Logarithmic Additive Contrast* - LAC), $C^{LP}_{(x,y)}(f)$. Il est défini comme le niveau de gris à ajouter au sens LIP à la plus faible valeur de pixel (i.e. la plus claire dans l'échelle LIP) afin d'obtenir la plus élevée (i.e. la plus sombre) [144] :

$$C^{\triangle}_{(x,y)}(f) = \max((f(x), f(y)) \triangle \min((f(x), f(y)). \tag{4.154}$$

Le LAC possède la propriété de robustesse aux variations d'intensité lumineuse modélisée par l'addition LIP d'une constante $c \in]-\infty; M[: C^{\triangle}_{(x,y)}(f \triangle c) = C^{\triangle}_{(x,y)}(f)$.

Fonctions couleur de coût Soient $\mathbf{f}(x) = (f_R(x), f_G(x), f_B(x))$ et $\mathbf{f}(y) = (f_R(y), f_G(y), f_B(y)) \in [0; M[^3$ deux couleurs d'une image couleur $\mathbf{f} : D \to [0; M[^3$. Leur contraste peut être défini comme il s'ensuit :

— par une approche marginale, avec une métrique $d_{(x,y)}$ définie comme une moyenne quadratique du LAC, $C^{\triangle}_{(x,y)}(f_i)$, de chaque composante f_i

$$d_{(x,y)}(\mathbf{f}) = \frac{1}{\sqrt{3}}\sqrt{\sum_{i\in\{R,G,B\}} [C^{\triangle}_{(x,y)}(f_i)]^2}\,; \tag{4.155}$$

— par une approche marginale, avec une métrique $d^{\triangle}_{(x,y)}$ définie comme la moyenne LIP du LAC de chaque composante

$$d^{\triangle}_{(x,y)}(\mathbf{f}) = \frac{1}{3} \triangle \left[\triangle_{i\in\{R,G,B\}} C^{\triangle}_{(x,y)}(f_i) \right]\,; \tag{4.156}$$

— par la métrique euclidienne dans l'espace couleur CIE 1976 L*a*b*

$$d^{lab}_{(x,y)}(\mathbf{f}) = \frac{1}{\sqrt{2}}\sqrt{\sum_{i\in\{L,a,b\}} [f_i(x) - f_i(y)]^2}\,; \tag{4.157}$$

— par une approche vectorielle, avec le minimum, $C^{\triangle_c min}_{(x,y)}(\mathbf{f})$, du Contraste Logarithmique Additif Couleur (en anglais *Colour Logarithmic Additive Contrast* - CLAC) [145]

$$\begin{aligned} C^{\triangle_c min}_{(x,y)}(\mathbf{f}) &= \min\,[\mathbf{C}^{\triangle_c}_{(x,y)}(\mathbf{f})] \\ &= \min\,[\wedge(\mathbf{f}(x),\mathbf{f}(y)) \triangle_c \vee(\mathbf{f}(x),\mathbf{f}(y))]. \end{aligned} \tag{4.158}$$

Le CLAC, $\mathbf{C}^{\triangle_c}_{(x,y)}(\mathbf{f})$, est un vecteur couleur de $[0; M[^3$ à ajouter au sens LIPC au maximum couleur point par point $\vee(\mathbf{f}(x),\mathbf{f}(y))$ (i.e. la couleur la plus claire) afin d'obtenir la couleur minimale point par point $\wedge(\mathbf{f}(x),\mathbf{f}(y))$ (i.e. la couleur la plus sombre). Le minimum CLAC $C^{\triangle_c min}_{(x,y)}(\mathbf{f})$ est un niveau de gris dans l'intervalle $[0; M[$. A cause de l'échelle couleur LIPC, le maximum $\vee$ est soustrait du minimum $\wedge$. C'est le contraire du LAC à niveaux de gris (éq. (4.154)) qui est défini dans l'échelle LIP opposée (Fig. 4.5, p. 94). Comme les métriques $d_{(x,y)}$ et $d^{\triangle}_{(x,y)}$ sont définies avec le LAC, elles sont dans l'échelle LIP. Cependant, le contraste minimum CLAC, $C^{\triangle_c min}_{(x,y)}$, est dans l'échelle usuelle. Afin d'utiliser la même échelle pour l'algorithme de percolation, nous remplacerons CLAC par la fonction coût $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f})$ définie par la complémentation suivante : $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f}) = (M-1) - C^{\triangle_c min}_{(x,y)}((M-1)-\mathbf{f})$.

La figure 4.58 illustre les courbes contrastes (ou de distances) entre deux vecteurs couleurs constants $\mathbf{f}(x) = [c,c,c]^t$ et $\mathbf{f}(y) = \mathbf{f}(x)+1$, où la valeur c varie dans l'intervalle $[0; M-1[$. On constate que : le minimum complémenté du CLAC, $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f})$, le LAC, $C^{\triangle}_{(x,y)}(f)$, la moyenne quadratique du LAC, $d_{(x,y)}(\mathbf{f})$ et la moyenne LIP du LAC, $d^{\triangle}_{(x,y)}(\mathbf{f})$ ont des courbes qui se superposent ; tandis que la courbe des distances Lab, $d^{lab}_{(x,y)}(\mathbf{f})$, a une forme différente.

Remarque. On ne peut cependant pas conclure que tous les contrastes logarithmiques sont égaux, mais que leurs comportements sont similaires sur une image couleur $\mathbf{f}$ ayant les trois mêmes composantes $\mathbf{f} = (f,f,f)$.

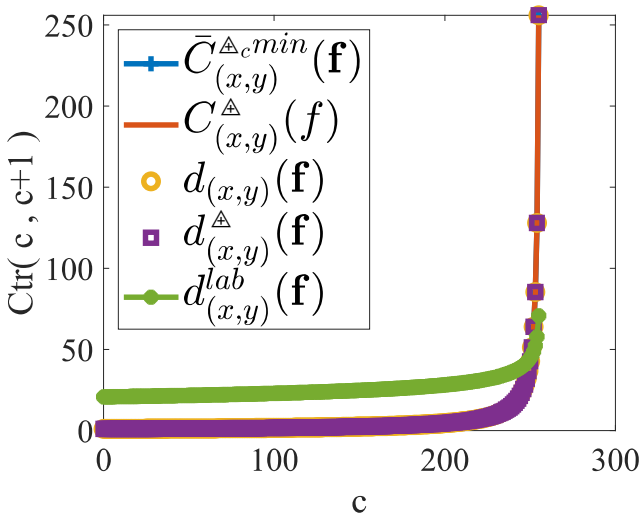
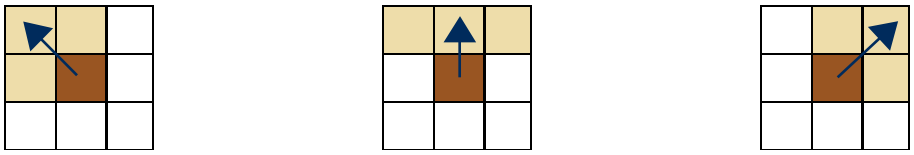


FIG. 4.58 – Comparaison des courbes de contraste avec $\mathbf{f}(x) = [c, c, c]^t$ un vecteur constant, où la valeur c varie dans l'intervalle $[0; M-1[$, et $\mathbf{f}(y) = \mathbf{f}(x) + 1$. On compare le minimum complémenté du CLAC, $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f})$, avec le LAC, $C^{\triangle}_{(x,y)}(f)$, la moyenne quadratique du LAC, $d_{(x,y)}(\mathbf{f})$, la moyenne LIP du LAC, $d^{\triangle}_{(x,y)}(\mathbf{f})$, et la distance Lab, $d^{lab}_{(x,y)}(\mathbf{f})$.

Percolation Sans perte de généralité, notre algorithme de percolation suivra les vallées sombres. Ces vallées ont des valeurs couleur plus proches de 0 que les valeurs des pixels environnants (dans l'échelle de gris usuelle). C'est le cas lorsque l'on considère e.g. l'ordre partiel $\leq$ entre les deux couleurs $\mathbf{c}_1 = (r_1, g_1, b_1)$ et $\mathbf{c}_2 = (r_2, g_2, b_2) \in [0; M[^3$: $\mathbf{c}_1 \leq \mathbf{c}_2 \iff \min\{r_1, g_1, b_1\} \leq \min\{r_2, g_2, b_2\}$. Notre algorithme est inspiré par la percolation à niveaux de gris de Puntous et al. [249], [250] qui utilise le LAC, $C^{\triangle}_{(x,y)}$, comme fonction de coût. Au-delà de ce travail, nous créons un algorithme couleur en utilisant une *fonction couleur de coût* et nous ajoutons une « tête de guidage » pour conduire la percolation. L'algorithme est défini comme il s'ensuit.

Soient S_p le point choisi à l'intérieur d'une vallée, m la taille d'un tableau de mémoire, th un seuil, $\{\vec{v_i}\}_{i=[1;N]}$ un ensemble de N directions possibles contenant la direction de départ $\vec{v_s}$ et $\mathcal{N}_{\vec{v_i}}(p)$ un voisinage d'un point p dépendant de la direction $\vec{v_i}$. L'ensemble de ces voisinages est appelé « tête de guidage ». La figure 4.59 montre un exemple de « tête de guidage ».

1. Le pixel de départ S_p est ajouté à la courbe de percolation $\mathcal{C}$ et devient le pixel courant C_p. Dans le voisinage de la « tête de guidage » $\mathcal{N}_{\vec{v_s}}(C_p)$ du pixel C_p, le pixel de plus faible intensité L_p est sélectionné. Ceci définit la nouvelle direction $\vec{v_i} = L_p - C_p$.
2. Le voisinage de la « tête de guidage » de C_p est mis dans la nouvelle direction $\vec{v_i}$.
3. Le pixel de plus faible intensité du voisinage $\mathcal{N}_{\vec{v_i}}(C_p)$ est sélectionné et devient le nouveau pixel N_p. Une nouvelle direction $\vec{v_i} = N_p - C_p$ est définie.
4. Une *fonction couleur de coût* est estimée entre l'intensité moyenne des m derniers pixels de la courbe $\mathcal{C}$ et N_p.
5. Si le coût est inférieur ou égal au seuil th, le pixel N_p est ajouté à la courbe $\mathcal{C}$, C_p est remplacé par N_p et le processus va à l'étape 2. Sinon, il s'arrête.



(a) $\mathcal{N}_{\vec{v}_1}(p)$. (b) $\mathcal{N}_{\vec{v}_2}(p)$. (c) $\mathcal{N}_{\vec{v}_3}(p)$.

FIG. 4.59 – Voisinages $\mathcal{N}_{\vec{v}_i}(p)$ de la « tête de guidage » dans 3 directions $\vec{v}_1$, $\vec{v}_2$ et $\vec{v}_3$ indiquée par les flèches. Les pixels en marron clair forment le voisinage $\mathcal{N}_{\vec{v}_i}(p)$ autour du pixel central p.

Illustration sur une image simulée

Dans cette partie la percolation couleur est illustrée sur une image simulée. L'objectif est de suivre de fines bandes blanches sur un fond rouge qui présente des reflets. Une image d'un panneau publicitaire couvert d'une vitre (Fig. 4.60a, 4.60b, 4.60c et 4.60d) [326] a été choisie dans la banque de données *Yahoo Flickr Creative Commons 100 Million Dataset* (YFCC100M) [299]. La vitre crée des reflets qui modifient localement le contraste. Afin de tester la robustesse de notre méthode de percolation aux variations d'éclairement, nous avons créé une image sombre $\mathbf{f}^{dk}$ en simulant une variation de l'intensité lumineuse (ou du temps d'exposition de la caméra) [144] (Fig. 4.60e, 4.60f, 4.60g et 4.60h). Dans ce but, l'image sombre $\mathbf{f}^{dk} = (f_r \triangle k, f_g \triangle k, f_b \triangle k)$ a été obtenue en ajoutant au sens LIP une constante $k = 160$ à chaque composante de l'image originale $\mathbf{f} = (f_r, f_g, f_b)$ représentée dans l'échelle LIP. Les points de départ ont été sélectionnés comme des maxima d'intensité de la dernière ligne de l'image et la « tête de guidage » a été celle représentée sur la figure 4.59. L'algorithme de percolation a été appliqué aux deux images $\mathbf{f}$ et $\mathbf{f}^{dk}$ représentées dans l'échelle LIP. Quatre fonctions couleur de coût ont été testées :

(i) la métrique de la moyenne quadratique du LAC $d_{(x,y)}$ (Fig. 4.60a et 4.60e) ;

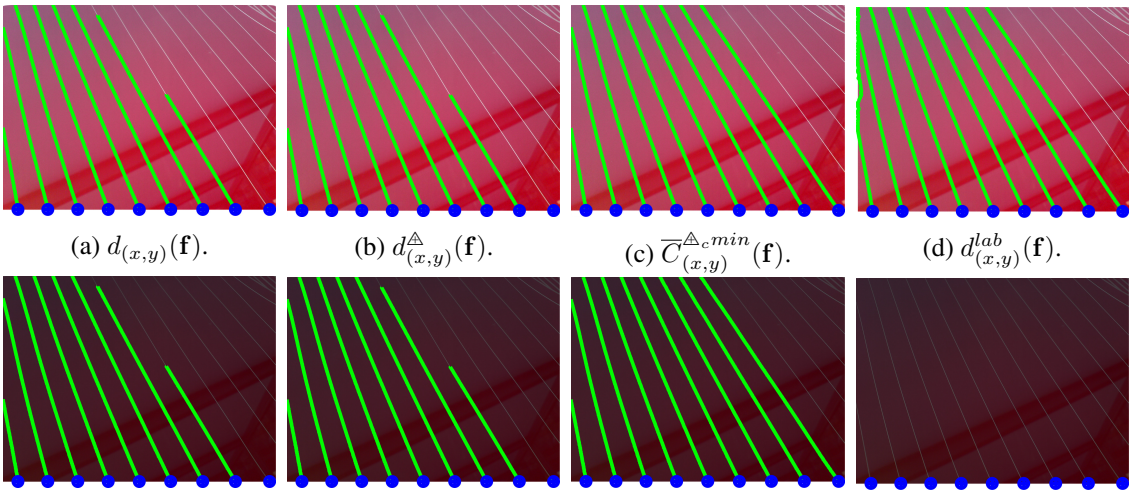
(ii) la métrique de la moyenne-LIP du LAC $d^{\triangle}_{(x,y)}$ (Fig. 4.60b et 4.60f) ;

(iii) le minimum du CLAC $\overline{C}^{\triangle_c min}_{(x,y)}$ (Fig. 4.60c et 4.60g) ;

(iv) la métrique euclidienne dans l'espace couleur CIE 1976 L*a*b* $d^{lab}_{(x,y)}(\mathbf{f}) = 1/\sqrt{2}\sqrt{\sum_{j\in\{l,a,b\}}[f_j(x) - f_j(y)]^2}$ (Fig. 4.60d et 4.60h).

Pour chaque fonction de coût couleur, le même seuil de percolation th a été utilisé sur l'image originale $\mathbf{f}$ et sur l'image assombrie $\mathbf{f}^{dk}$. On peut remarquer que pour chacune des trois fonctions logarithmiques couleur de coût $d_{(x,y)}$ (Fig. 4.60a et 4.60e), $d^{\triangle}_{(x,y)}$ (Fig. 4.60b et 4.60f) et $\overline{C}^{\triangle_c min}_{(x,y)}$ (Fig. 4.60c et 4.60g), les résultats sont les mêmes dans les deux images $\mathbf{f}$ et $\mathbf{f}^{dk}$. Cela montre la robustesse de ces trois fonctions de coût à une variation de l'intensité lumineuse ou du temps d'exposition de la caméra. Cependant, pour la métrique euclidienne Lab $d^{lab}_{(x,y)}$, la percolation diffère entre l'image originale $\mathbf{f}$ et l'image assombrie $\mathbf{f}^{dk}$(Fig. 4.60d, 4.60h). La métrique euclidienne Lab $d^{lab}_{(x,y)}$ n'est donc pas robuste aux changements d'intensité lumineuse.

De plus, le seuil de percolation $th = 65$ était le même pour les trois fonctions logarithmiques de coût $d_{(x,y)}$ (Fig. 4.60a et 4.60e), $d^{\triangle}_{(x,y)}$ (Fig. 4.60b et 4.60f) et $\overline{C}^{\triangle_c min}_{(x,y)}$ (Fig. 4.60c et 4.60g). Dans cet exemple, on peut remarquer que la métrique de la moyenne quadratique du LAC $d_{(x,y)}$ et la métrique de la moyenne-LIP du LAC $d^{\triangle}_{(x,y)}$ donnent des résultats similaires.



(a) $d_{(x,y)}(\mathbf{f})$. (b) $d^{\triangle}_{(x,y)}(\mathbf{f})$. (c) $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f})$. (d) $d^{lab}_{(x,y)}(\mathbf{f})$.

(e) $d_{(x,y)}(\mathbf{f}^{dk})$. (f) $d^{\triangle}_{(x,y)}(\mathbf{f}^{dk})$. (g) $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f}^{dk})$. (h) $d^{lab}_{(x,y)}(\mathbf{f}^{dk})$.

FIG. 4.60 – Courbes de percolation (en vert) dans une image $\mathbf{f}$ (a, b, c, d) et dans sa version assombrie $\mathbf{f}^{dk}$ (e, f, g, h). Quatre fonctions couleur de coût ont été utilisées : (a, e) $d_{(x,y)}$, (b, f) $d^{\triangle}_{(x,y)}$, (c, g) $\overline{C}^{\triangle_c min}_{(x,y)}$ et (d, h) $d^{lab}_{(x,y)}$. Le seuil est de $th = 65$ dans (a, e), (b, f) et (c, g), et de $th = 30$ dans (d, h). Les points de départ sont représentés en bleu et la direction initiale est du bas vers le haut. La taille du tableau de mémoire est de $m = 30$ pixels pour une image de taille $2\,048 \times 1\,536$ pixels.

Un algorithme de percolation pour une image du fond d'œil

Sur une image couleur du fond d'œil $\mathbf{f} : D \to [0; M[^3$ (Fig. 4.61a), j'ai proposé une méthode inspirée de la croissance de région et de la percolation afin d'extraire les vaisseaux. Le contraste minimum du CLAC $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f})$ a été utilisé.

L'algorithme utilise la méthode de la *carte d'appartenance aux vaisseaux* de la sous-sous-section « Segmentation des vaisseaux », p. 157, afin de déterminer les directions de percolation. Les sondes $b_{\theta,i}$, où $\theta \in \Theta$ et $i \in [\![1 \ldots I]\!]$ sont employées. Elles se composent de trois segments parallèles d'orientations θ : celui de gauche $b^l_{\theta,i}$, le central $b^c_{\theta,i}$ et le droit $b^r_{\theta,i}$. Ces segments ont des intensités et des longueurs l_i différentes. La largeur entre les segments gauche et droit est égale à w_i. Dans l'algorithme de percolation (alg. 5), le choix des directions se fait grâce à un seuil th fixé empiriquement. Celui-ci est égal à la multiplication-LIP par le facteur 0,85, de la différence-LIP entre le maximum et le minimum de la sonde, i.e. son amplitude, $th = 0{,}85 \triangle (b^c_{\theta,i} \triangle b^l_{\theta,i})$. Lorsqu'une direction θ est choisie pour un pixel P, on regarde ses voisins dans les deux sens de la direction. Pour les opérations de voisinage, une 8-connexité est utilisée. La percolation est arrêtée grâce à un seuil th_{ctr} de 220 (dans l'échelle inversée des gris) sur le minimum du CLAC, $\overline{C}^{\triangle_c min}_{(x,y)}(\mathbf{f})$.

La figure 4.61b montre les résultats préliminaires de la méthode sur une image présentant de fortes variations de contraste. Les résultats sont encourageants et restent à valider sur plus de données.

Algorithme 5 Algorithme de détection des vaisseaux par percolations successives.

1: Soit seg le masque de segmentation des vaisseaux, de la taille du domaine D de l'image $\mathbf{f}$ qui est initialisé à la valeur $False$. ▷ seg est la sortie de l'algorithme
2: Soit $imWork$ un tableau de taille égale à celle de D et prenant 3 valeurs $LabelCandidate$, $LabelQueued$, $LabelDone$. $imWork$ est initialisé à la valeur $LabelCandidate$.
3: Soit $th_{ctr} = 220$ le seuil sur le contraste (dans l'échelle LIP).
4: Dans l'image $\mathbf{f}$, on sélectionne un point de départ S_p.
5: Soit $\left\{E_{b_{\theta,i},k}\right\}_{\theta\in\Theta;i\in[\![1\ldots I]\!]}$ l'ensemble des cartes détecteurs de bosses de l'équation (4.141), p. 159 pour la sonde $b_{\theta,i}$ composée de trois segments.
6: Soit PQ une file d'attente hiérarchique [201].
7: $I_{max} \leftarrow S_p$ ▷ Couleur maximale appartenant à $[0; M[^3$
8: $I_{min} \leftarrow S_p$ ▷ Couleur minimale appartenant à $[0; M[^3$
9: $seg(S_p) \leftarrow True$
10: $imWork(S_p) \leftarrow LabelDone$
11: $LNeigh \leftarrow S_p + \{v_\alpha\}_{\alpha\in A}$, où $A = \left\{\{\theta; \theta+\pi\} \mid \{E_{b_{\theta,i},k}(S_p), i \in [\![1\ldots I]\!]\} < th\}\right\}$ ▷ $LNeigh$ liste des pixels voisins, v_α vecteur d'orientation α
12: $[PQ, imWork] \leftarrow$ AddNeighbours1($LNeigh$, A, C_p, PQ, $imWork$, I_{max}, I_{min}, th, th_{ctr}, $\mathbf{f}$, $\{E_{b_{\theta,i},k}\}_{\theta\in\Theta;i\in[\![1\ldots I]\!]}$) ▷ alg. 7
13: **while** PQ n'est pas vide **do**
14: $[seg, imWork] \leftarrow$ PercolationProcess(PQ, I_{max}, I_{min}, seg, $\mathbf{f}$, $imWork$, th_{ctr}, th $\{E_{b_{\theta,i},k}\}_{\theta\in\Theta;i\in[\![1\ldots I]\!]}$) ▷ alg. 6
15: $[I_{max}, I_{min}, seg] \leftarrow$ ContractionProcess(PQ , seg, $\mathbf{f}$, th_{ctr}) ▷ alg. 9
16: $LNeigh$ liste des voisins du masque seg obtenue par dilatation
17: $[PQ, imWork] \leftarrow$ AddNeighbours2($LNeigh$, PQ, $imWork$, I_{max}, I_{min}, th, th_{ctr}, $\mathbf{f}$, $\{E_{b_{\theta,i},k}\}_{\theta\in\Theta;i\in[\![1\ldots I]\!]}$) ▷ alg. 8
18: **end while**

Algorithme 6 $[seg, imWork] \leftarrow$ PercolationProcess(PQ, I_{max}, I_{min}, seg, $\mathbf{f}$, $imWork$, th_{ctr}, th, $\{E_{b_{\theta,i},k}\}_{\theta\in\Theta;i\in[\![1\ldots I]\!]}$)

1: **while** PQ n'est pas vide **do**
2: $C_p \leftarrow PQ.pop()$ ▷ C_p point courant
3: $imWork(C_p) \leftarrow LabelDone$
4: **if** $\overline{C}^{\triangle_c min}(\vee\{\mathbf{f}(C_p), I_{max}\}, \wedge\{\mathbf{f}(C_p), I_{min}\}) < th_{ctr}$ **then**
5: $I_{max} \leftarrow \vee\{\mathbf{f}(C_p), I_{max}\}$ ▷ Supremum point à point
6: $I_{min} \leftarrow \wedge\{\mathbf{f}(C_p), I_{min}\}$ ▷ Infimum point à point
7: $seg(C_p) \leftarrow True$
8: $imWork(C_p) \leftarrow LabelDone$
9: $LNeigh \leftarrow C_p + \{v_\alpha\}_{\alpha\in A}$, où $A = \left\{\{\theta; \theta+\pi\} \mid \{E_{b_{\theta,i},k}(C_p), i \in [\![1\ldots I]\!]\} < th\}\right\}$ ▷ $LNeigh$ liste des pixels voisins, v_α vecteur d'orientation α
10: $[PQ, imWork] \leftarrow$ AddNeighbours1($LNeigh$, A, C_p, PQ, $imWork$, I_{max}, I_{min}, th, th_{ctr}, $\mathbf{f}$, $\{E_{b_{\theta,i},k}\}_{\theta\in\Theta;i\in[\![1\ldots I]\!]}$) ▷ alg. 7
11: **end if**
12: **end while**

Algorithme 7 $[PQ, imWork] \leftarrow$ AddNeighbours1($LNeigh$, A, C_p, PQ, $imWork$, I_{max}, I_{min}, th, th_{ctr}, $\mathbf{f}$, $\{E_{b_{\theta,i},k}\}_{i\in[\![1\ldots I]\!]}$)

1: $P \leftarrow length(LNeigh)$
2: **for all** $p \in [\![1 \ldots P]\!]$ **do**
3: $N_p \leftarrow LNeigh(p)$
4: $\alpha_p \leftarrow A(p)$
5: $prio \leftarrow \min\{E_{b_{\alpha_p,i},k}(C_p) \mid i \in [\![1 \ldots I]\!]\}$ ▷ Priorité du pixel
6: **if** $(imWork(N_p) = LabelCandidate)$ et $(prio < th)$ et $(\overline{C}^{\triangle_c min}(\vee\{\mathbf{f}(N_p), I_{max}\}, \wedge\{\mathbf{f}(N_p), I_{min}\}) < th_{ctr})$ **then**
7: $PQ.insert \leftarrow (prio, N_P)$
8: $imWork(N_p) \leftarrow LabelQueued$
9: **end if**
10: **end for**

Algorithme 8 $[PQ, imWork] \leftarrow$ AddNeighbours2($LNeigh$, PQ, $imWork$, I_{max}, I_{min}, th, th_{ctr}, $\mathbf{f}$, $\{E_{b_{\theta,i},k}\}_{\theta\in\Theta; i\in[\![1\ldots I]\!]}$)

1: **for all** $N_p \in LNeigh$ **do**
2: $prio \leftarrow \min\{E_{b_{\theta,i},k}(N_p) \mid \theta \in \Theta, i \in [\![1 \ldots I]\!]\}$ ▷ Priorité du pixel
3: **if** $(imWork(N_p) = LabelCandidate)$ et $(prio < th)$ et $(\overline{C}^{\triangle_c min}(\vee\{\mathbf{f}(N_p), I_{max}\}, \wedge\{\mathbf{f}(N_p), I_{min}\}) < th_{ctr})$ **then**
4: $PQ.insert \leftarrow (prio, N_P)$
5: $imWork(N_p) \leftarrow LabelQueued$
6: **end if**
7: **end for**

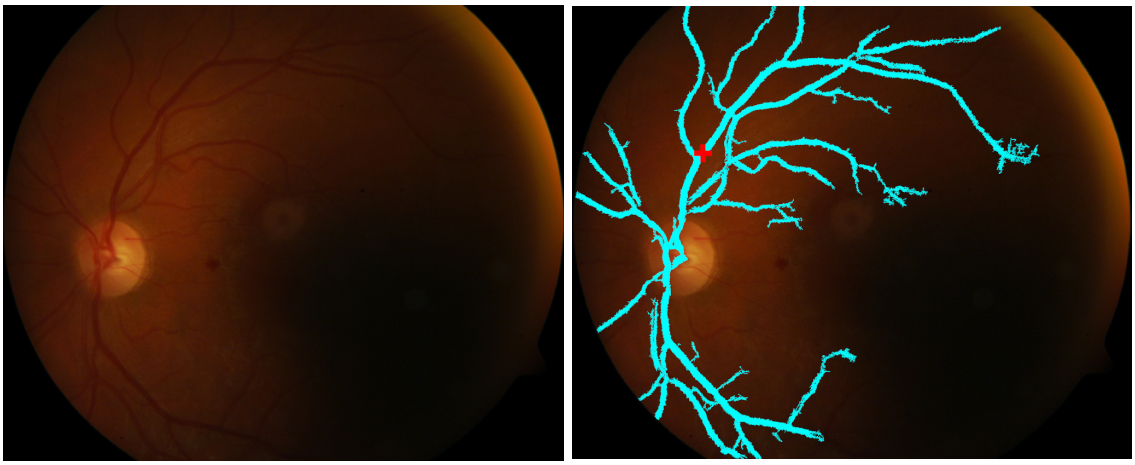
Algorithme 9 $[I_{max}, I_{min}, seg] \leftarrow$ ContractionProcess(PQ , seg, $\mathbf{f}$, th_{ctr})

1: $k \leftarrow 1$
2: $Maxi \leftarrow 0$
3: **while** $Maxi = 0$ **do**
4: **for all** $i \in [\![0; k]\!]$ **do**
5: $[mskCtrct] \leftarrow$ ithcontraction(i, k , $\mathbf{f}$, seg) ▷ alg. 10
6: **if** $mskCtrct$ n'est pas vide **then**
7: $I_{max} \leftarrow \vee\{\mathbf{f}(x \in D \mid mskCtrct(x)$ est égal à $True)\}$ ▷ Supremum point à point
8: $I_{min} \leftarrow \wedge\{\mathbf{f}(x \in D \mid mskCtrct(x)$ est égal à $True)\}$ ▷ Infimum point à point
9: $Card_C \leftarrow Card\{x \mid mskCtrct(x)\}$
10: **if** $(\overline{C}^{\triangle_c min}(I_{max}, I_{min}) < th_{ctr})$ et $(Card_C > Maxi)$ **then**
11: $Maxi \leftarrow Card_C$
12: $seg \leftarrow mskCtrct$
13: **end if**
14: **else**
15: $Maxi \leftarrow 1$
16: $I_{max} \leftarrow \vee\{\mathbf{f}(x \in D \mid seg(x)$ est égal à $True)\}$ ▷ Supremum point à point
17: $I_{min} \leftarrow \wedge\{\mathbf{f}(x \in D \mid seg(x)$ est égal à $True)\}$ ▷ Infimum point à point
18: **end if**
19: **end for**
20: **if** $Maxi$ est égal à 0 **then**
21: $k \leftarrow k + 1$
22: **end if**
23: **end while**

Algorithme 10 $[mskCtrct] \leftarrow$ ithcontraction(i, k , $\mathbf{f}$, seg)

Permet de supprimer les extrémités de l'histogramme des pixels de $\mathbf{f}$ appartenant au masque seg.

1: $E \leftarrow \{x \in D \mid seg(x)$ est égal à $True\}$ ▷ Ensemble des positions des pixels égaux à $True$ dans le masque seg
2: $\forall x \in E$, $f_{max}(x) \leftarrow \sup\left[\{f_R(x), f_G(x), f_B(x)\}\right]$ ▷ Image des maxima des 3 composantes pixels à pixels
3: $\forall x \in E$, $f_{min}(x) \leftarrow \inf\left[\{f_R(x), f_G(x), f_B(x)\}\right]$ ▷ Image des minima des 3 composantes pixels à pixels
4: $maxi \leftarrow \sup\{f_{max}(x), x \in E\}$ ▷ Maximum des valeurs des pixels
5: $mini \leftarrow \inf\{f_{min}(x), x \in E\}$ ▷ Minimum des valeurs des pixels
6: $i' \leftarrow k \triangle i$
7: $mini_{shift} \leftarrow mini \triangle i$
8: $maxi_{shift} \leftarrow maxi \triangle i$
9: Initialisons les masques $mask_{min_{kept}}$ et $mask_{max_{kept}} : D \rightarrow \{True, False\}$ à la valeur $False$
10: $\forall x \in E$, $mask_{min_{kept}}(x) \leftarrow \{True;$ si $f_{min}(x) \geq mini_{shift};$ sinon $False\}$ ▷ Masque des minima conservés
11: $\forall x \in E$, $mask_{max_{kept}}(x) \leftarrow \{True;$ si $f_{max}(x) \leq maxi_{shift};$ sinon $False\}$ ▷ Masque des maxima conservés
12: $mskCtrct \leftarrow \vee\{mask_{max_{kept}}, mask_{min_{kept}}\}$ ▷ Masque de la contraction



(a) Image couleur $\mathbf{f}$.

(b) Percolation dans les vaisseaux.

FIG. 4.61 – (a) Image faiblement contrastée de la banque DIARETDB1 [151]. (b) Extraction des vaisseaux par percolation. Le point de départ de l'algorithme est indiqué par une croix rouge.

Perspectives

Nous avons introduit une nouvelle méthode de percolation couleur qui est robuste aux changements de contraste causés par une variation de l'intensité lumineuse ou du temps d'exposition de la caméra. Cette propriété est due aux *fonctions couleur de coût* qui ont été définies d'une manière marginale dans le cadre de travail du Traitement d'Image Logarithmique (LIP) ou d'une manière vectorielle dans le cadre LIPC pour les images couleur.

- Deux fonctions de coût sont marginales, à savoir la métrique de la moyenne quadratique du *Logarithmic Additive Contrast* (LAC) et la métrique de la moyenne-LIP du LAC.
- Une fonction de coût est vectorielle, à savoir le minimum du *Colour* LAC (CLAC).

Ces trois *fonctions couleur de coût* sont robustes aux changements de contraste qui sont modélisés par une addition-LIP d'une constante. Les résultats obtenus sur une image simulée sont meilleurs que ceux obtenus avec la distances euclidienne dans l'espace couleur $L^*a^*b^*$ qui ne présente pas cette robustesse. Les résultats obtenus sur une image couleur du fond d'œil montrent qu'il est possible de détecter des structures fines telles que des vaisseaux avec la *fonction couleur de coût* correspondant au minimum du *Colour* LAC. La percolation couleur ouvre la voie à la détection de structures fines présentant de fortes variations de contrastes causées par des changements lumineux et qui existent dans de nombreux cadres comme la télédétection [222], les images médicales [234], etc. La percolation couleur peut aussi être combinée avec des filtres de structures curvilignes [183].

4.6.2 Percolation pour l'analyse des images de pneumatique

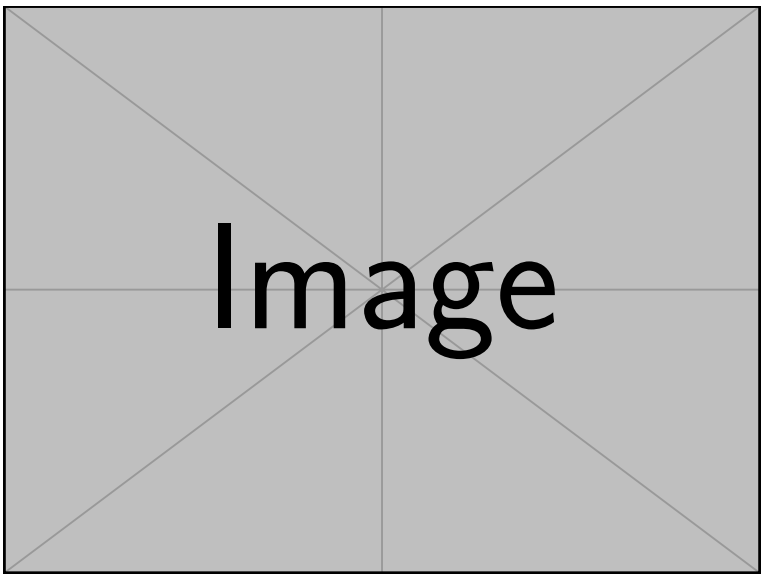
Au Centre de Recherche de Michelin, j'ai utilisé la percolation à niveaux de gris avec le contraste LAC (*Logarithmic Additive Contrast*) sur des images du pneumatique acquises par rayons X. L'objectif était de suivre les fils métalliques de la nappe carcasse sur un pneumatique. La méthode de percolation a permis de résoudre plusieurs cas d'extraction où les fils de deux nappes métalliques sont superposés, avec des angles différents comme illustré sur la figure 4.62.

En outre, des sauts de contrastes pouvaient exister sur des images car l'acquisition était faite avec plusieurs clichés comme on peut le voir sur la figure 4.63a. La soustraction-LIP d'une constante égale au contraste LAC au niveau du saut a permis de corriger ce dernier (Fig. 4.63b).

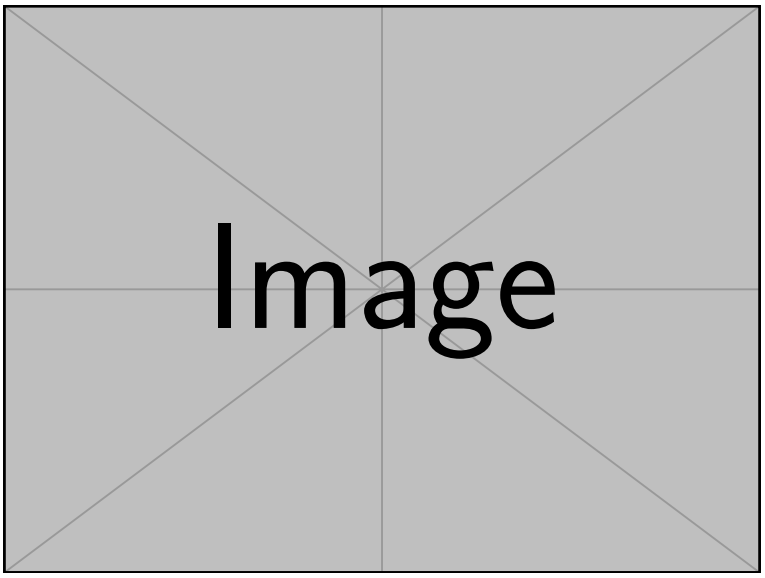
Avec un autre système d'acquisition par rayons X, j'ai extrait les fils métalliques par filtrage puis percolation sur les deux Nappes Sommet de Travail d'un pneumatique (Fig. 4.64). Sur l'image de départ (Fig. 4.64a), les fils sont difficiles à détecter. Cependant, grâce à du filtrage directionnel adaptatif et à de la percolation, j'ai pu détecter les deux nappes acquises selon des poses différentes (Fig. 4.64a et 4.64b). Ce suivi de fils sur plusieurs poses du pneumatique, a ensuite servi à faire de la reconstruction tri-dimensionnelle (cf. section 5.2, p. 200).

La percolation a aussi été utilisée pour suivre les limites des composants des pneumatiques sur une coupe selon un plan radial (Fig. 4.65). J'ai ainsi supervisé une collaboration avec Michel Jourlin et une PME innovante qui a abouti à un outil de détection de la limite des produits. En pratique, l'opérateur choisit un nombre réduit de points sur l'image et l'algorithme est capable de tracer la frontière entre les points. Une carte des distances d'Asplund entre une sonde et l'image permet de réduire les itinéraires possibles pour la percolation (Fig. 4.66a). La difficulté est que l'on recherche des frontières sur des images faiblement contrastées. Grâce à la carte des distances d'Asplund et à la percolation, l'algorithme réussit à trouver un chemin entre des points distants (Fig. 4.66b). La méthode introduite a permis de réaliser un gain de temps appréciable du fait de la réduction drastique de la quantité de points sélectionnés.

L'utilisation du Traitement d'Image Logarithmique (LIP) s'avère donc très pertinent dans de nombreuses applications où le contraste de l'acquisition n'est pas maîtrisé.

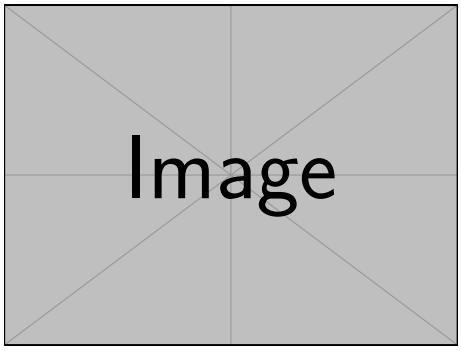


(a) Extraction des fils par percolation.

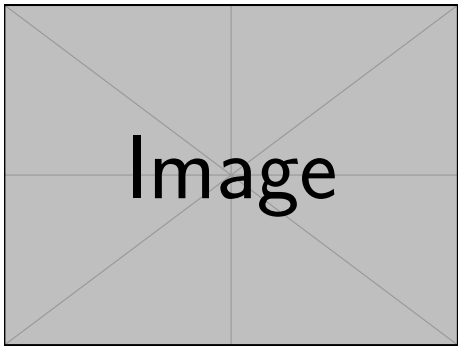


(b) Zoom sur une partie de la nappe carcasse.

FIG. 4.62 – (a) Extraction des fils de la nappe carcasse par percolation logarithmique. (b) Zoom sur une partie de la nappe carcasse. Les fils sont aussi extraits par percolation.



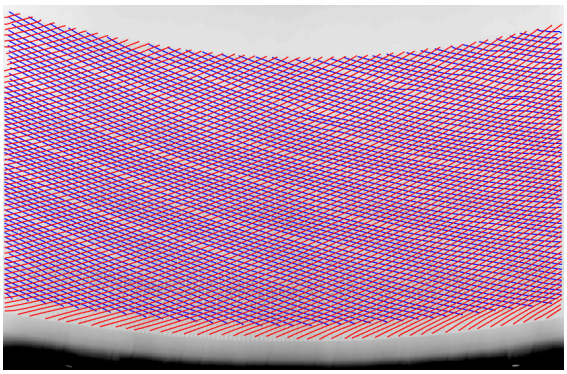
(a) Saut de contraste.



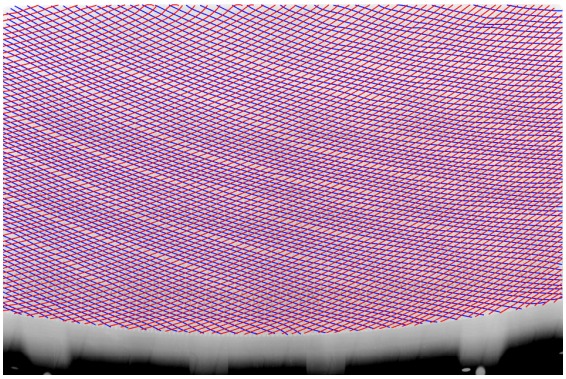
(b) Suppression du saut de contraste.

FIG. 4.63 – (a) Image initiale en rayons X avec un saut de contraste. (b) Correction du saut de contraste par soustraction d'une constante LIP égale au LAC du saut de contraste.

(a) Image par rayons X.

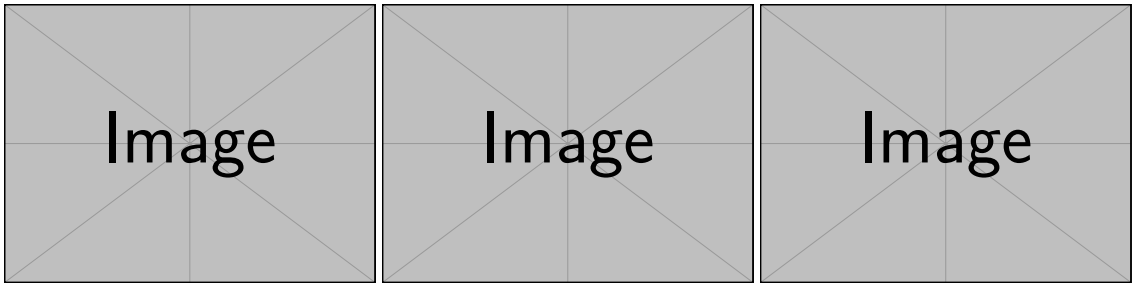


(b) Extraction des fils sur la vue 1.



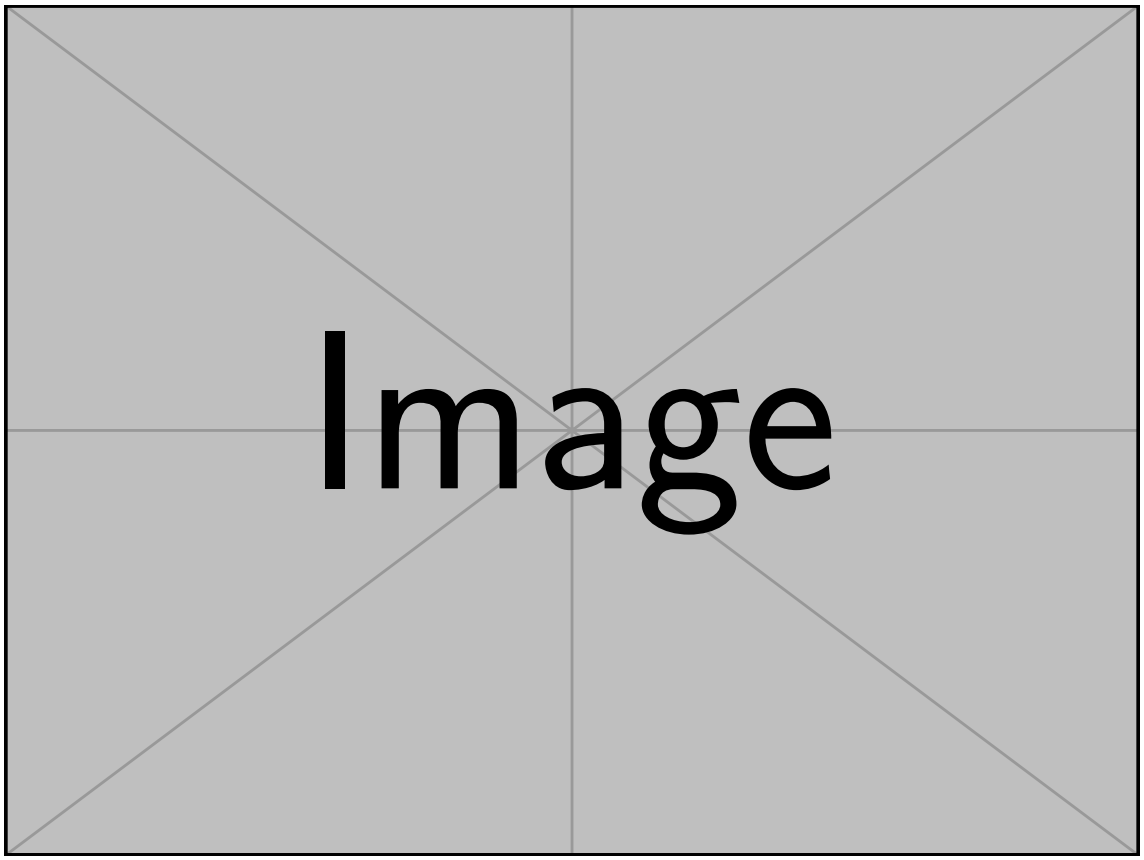
(c) Extraction des fils sur la vue 2.

FIG. 4.64 – (a) Images des nappes sommets de travail d'un pneumatique acquises en imagerie par rayons X. (b) Extraction des fils métalliques des deux nappes sommets de travail par filtrage et percolation sur la vue dans la pose n° 1. (b) Idem sur la vue dans la pose n° 2.

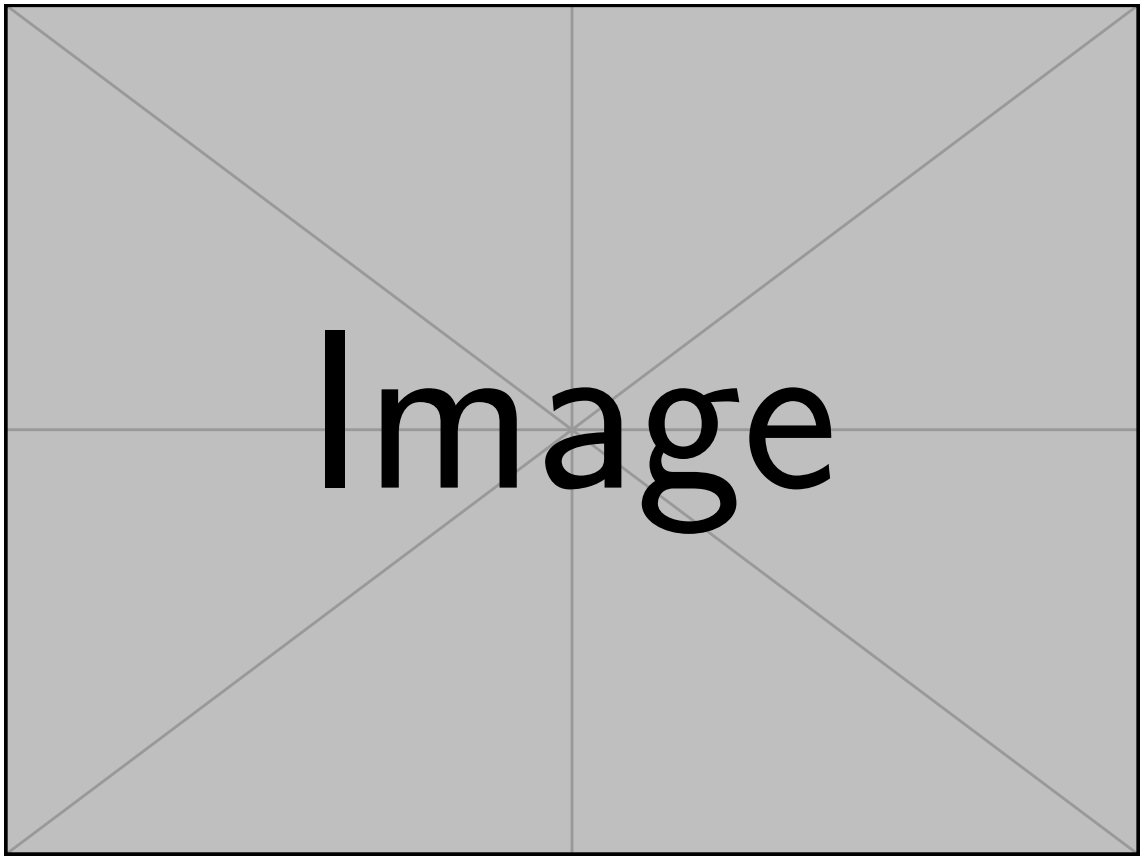


(a) (b) (c)

FIG. 4.65 – (a), (b) et (c) Trois exemples de détection de frontières sur une coupe de pneumatique grâce à la percolation sur une carte des distances d'Asplund.



(a) Carte des distances d'Asplund.



(b) Sélection de deux points et percolation.

FIG. 4.66 – (a) Carte des distances d'Asplund sur une image de coupe de pneumatique selon un plan radial. (b) Percolation entre deux points sélectionnés sur la coupe.

4.7 Conclusion du chapitre

Dans ce chapitre, j'ai présenté mes travaux sur le Traitement d'Image Logarithmique (LIP). Ceux-ci sont de première importance pour les images présentant de forts changements de contraste dus à des causes physiques. Les effets de la variation de l'intensité de la source d'éclairement ou du temps d'exposition de la caméra sont corrigés par la soustraction-LIP d'une constante d'une image. Les effets de la variation de l'opacité ou de l'absorption de l'objet acquis par la transmission d'une source d'éclairement à travers cet objet sont corrigés par la multiplication-LIP par un scalaire d'une image. Ces opérations fonctionnent aussi bien pour les images en niveaux de gris que pour les images couleur.

- J'ai proposé une nouvelle méthode d'amélioration des contrastes couleur des images du fond d'œil. Celle-ci améliore les variations locales de contraste afin de rendre plus visibles les lésions ou les structures anatomiques tels que les vaisseaux. La classification de la sévérité de la rétinopathie diabétique devient ainsi plus évidente et plus rapide. Ces travaux ont été initialement développés avec l'équipe du Professeur David Owens, diabétologue à *Swansea University*, et en lien avec le *Diabetic Eye Screening Wales*, au Royaume-Uni, dirigé à l'époque par Andrew Crowder. Des tests ont été réalisés sur plusieurs milliers d'images du fond d'œil de qualités variées. Le travail s'est poursuivi avec l'équipe du Professeur Laurent Kodjikian, ophtalmologiste aux Hospices Civils de Lyon. Nous avons ainsi pu voir que les lésions et notamment les plus petites (telles que les anomalies microvasculaires intra-rétiniennes dont la détection est fondamentale) deviennent plus visibles. Une étude clinique a été planifiée afin de valider le système d'amélioration d'un point de vue médical. En outre, cet algorithme de normalisation de contraste facilite le traitement informatique des images comme nous le verrons dans la section 6.2, p. 201.
- Avec Michel Jourlin, j'ai démontré le lien entre les cartes des distances fonctionnelles d'Asplund et la Morphologie Mathématique. Ce lien est intéressant d'un point de vue théorique car il permet l'introduction de nouveaux opérateurs morphologiques (i.e. les cartes des distances d'Asplund) avec de fortes propriétés optiques. Il est aussi intéressant d'un point de vue pratique, puisqu'il permet de plus facilement programmer ces cartes parce que les opérateurs morphologiques sont présents dans de nombreux logiciels de traitement d'image. En ce qui concerne les propriétés optiques, les cartes des distances d'Asplund LIP-multiplicatives sont robustes aux changements de contraste causés par une variation d'opacité ou d'absorption de l'objet et modélisés par la loi LIP-multiplicative $\triangle$. Quant aux cartes des distances d'Asplund LIP-additives, elles sont robustes aux changements de contraste causés par une variation de l'intensité de la source d'éclairement ou du temps d'exposition de la caméra et modélisés par la loi LIP-additive $\triangle$. J'ai aussi défini les cartes des distances d'Asplund LIP-multiplicatives pour les images couleur et multivariées.
- J'ai défini le cadre théorique de la Morphologie Mathématique Logarithmique (LMM) entre des fonctions. Grâce à la loi LIP-additive $\triangle$, elle permet l'introduction d'opérateurs adaptatifs ou robustes aux variations de l'intensité de la source d'éclairement ou du temps d'exposition de la caméra. La LMM constitue un cadre de travail mathématiquement bien établi et justifié d'un point de vue optique. Elle permet d'interpréter l'opérateur de la *carte des distances d'Asplund LIP-additives* comme une extension du gradient morphologique logarithmique aux fonctions structurantes (i.e. les éléments structurants non plats). En outre, des opérateurs robustes aux variations de contrastes modélisés par la loi LIP-additive $\triangle$ peuvent être définis par différence logarithmique entre des opérateurs de LMM. Celle-ci ouvre la voie à la création de nouveaux opérateurs fonctionnels robustes aux variations de contrastes qui seront très utiles dans de nombreuses applications.
- Grâce aux cartes des distances d'Asplund et à la LMM, j'ai pu proposer en lien avec l'équipe du Professeur Laurent Kodjikian, des méthodes de segmentation des micro-anévrismes et des vaisseaux dans les images du fond d'œil. Pour la détection des vaisseaux, les premiers résultats

sont comparables aux meilleurs de l'état de l'art. Ceci est rendu possible grâce à la LMM qui permet de définir des opérateurs robustes aux variations de contraste. En outre, avec Michel, Jourlin, j'ai proposé deux critères d'homogénéité de région à partir du modèle LIP. Le premier, défini grâce au Contraste Logarithmique Additif, est robuste aux variations d'intensité de la source d'éclairement ou du temps d'exposition de la caméra. Le deuxième, défini grâce au Contraste Logarithmique Multiplicatif, est robuste aux variations d'opacité ou d'absorption de l'objet acquis par une source d'éclairement transmise.

- J'ai utilisé la percolation logarithmique pour le suivi de fils métalliques dans des images de pneumatiques acquises par rayons X. Avec Michel Jourlin, j'ai étudié la percolation logarithmique avec des *fonctions couleur de coût*. Les résultats obtenus sont très prometteurs et permettent e.g. d'extraire l'arbre vasculaire dans les images du fond d'œil présentant de fortes variations de contraste. Cette approche sera très utile pour l'analyse d'images rétinienne provenant de grandes banques de données de santé publique.

Les méthodes et outils développés autour du Traitement d'Image Logarithmique s'avèrent être cruciaux pour l'analyse des grandes banques de données d'images industrielles ou médicales. Ces nouveaux outils permettent de rendre les images comparables entre elles et de faire de la segmentation reproductible entre des images acquises dans des conditions différentes et présentant des variations de contrastes globaux et locaux. Les approches proposées, combinées avec d'autres méthodes d'analyse d'image, telles que l'apprentissage statistique, permettront une analyse robuste au changement de contraste et reproductible de ces grandes banques d'images.

Chapitre 5

Stéréovision pour l'analyse des performances des pneumatiques

Afin d'analyser les performances des pneumatiques, la connaissance tridimensionnelle de leur surface topographique ou des positions des fils métalliques les composant s'avère intéressante. En effet, cette mesure peut permettre une comparaison avec les simulations numériques. Cependant, la reconstruction d'un volume 3D à partir de photographies bidimensionnelles, nécessite de calibrer le système d'acquisition, i.e. de connaître les matrices de projection permettant le passage du 3D au 2D. Grâce à au moins deux projections selon des poses différentes de la caméra et grâce à la matrice de passage entre ces deux poses, on pourra reconstruire une information 3D. Les acquisitions d'images ont été faites dans deux domaines de longueurs d'ondes :

i) le visible, dans l'intervalle 380 nm à 780 nm, grâce à la photographie qui capture la surface des pneumatiques et

ii) les rayons X, dans l'intervalle 10×10^{-3} nm à 10 nm, grâce à la radiographie qui capture les fils métalliques composant les pneumatiques.

Une fois le calibrage effectué, i.e. lorsque l'on connaît les paramètres du système pour remonter à l'information 3D, on peut effectuer la reconstruction 3D par triangulation sur des paires d'images.

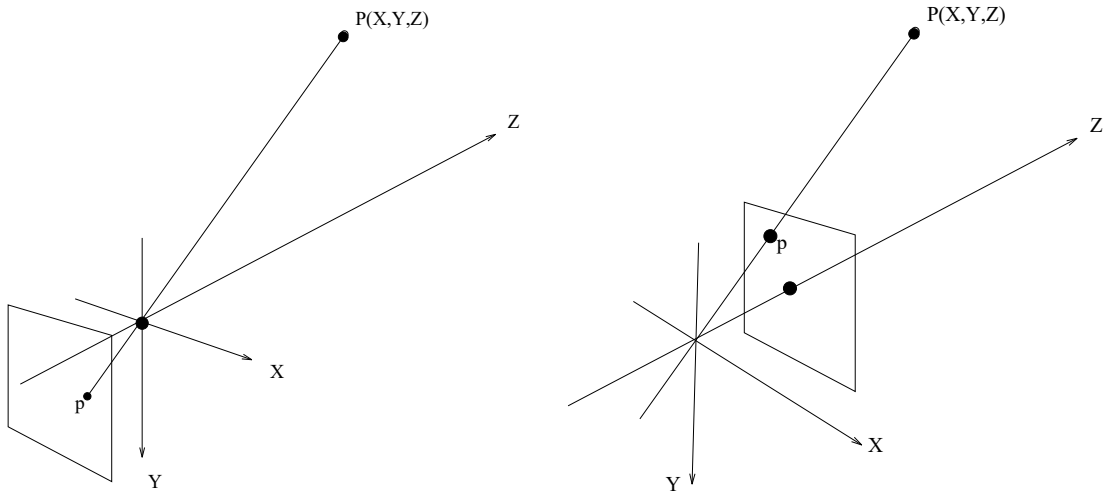
Le stage de Thibaud Vilboux [312], que j'ai co-encadré, a permis d'étudier les algorithmes de calibrage et de reconstruction 3D dans le visible. Mes travaux ont ensuite permis de concevoir la chaîne complète de reconstruction 3D dans le visible et en imagerie rayons X. Celui-ci comprend la prise d'image bidimensionnelle, le calibrage des caméras et l'algorithme de reconstruction 3D.

5.1 Système de stéréovision par photographie

Le calibrage d'un système de stéréovision consiste à déterminer une équation entre les coordonnées tridimensionnelles de l'objet et les coordonnées bidimensionnelles de son projeté dans une image. Sur cette partie, j'ai co-encadré le stage de Thibaud Vilboux de mars à septembre 2011 pour son stage d'élève-ingénieur qui s'est déroulé au centre de Recherche de Michelin [312]. Durant le stage, nous avons eu l'occasion d'échanger avec Jean-Marc Lavest, Professeur à l'Université Blaise-Pascal de Clermont-Ferrand et spécialiste des problèmes de stéréovision. Le travail du stagiaire a été inspiré par le cours de Dhome et Lavest [67]. Une partie de ce cours sera reprise ci-après. Les ouvrages de Hartley et Zisserman [106] et de Trucco et Verri [302] constituent aussi des références dans le domaine, sur lesquelles je me suis appuyé.

5.1.1 Modèle de transformation

L'appareil photographique et son objectif sont modélisés par le modèle sténopé (en anglais *pinhole*). Dans ce modèle, tous les rayons passent par un seul et même point, le centre optique (i.e. un trou de faible diamètre percé dans une chambre noire et qui permet de capturer une image). Le capteur de l'appareil se situe à une distance f de ce centre optique et représente la distance focale (Fig. 5.1a). Afin d'éviter le problème d'inversion de l'image, on modélise le capteur comme étant devant le centre optique, ce qui est équivalent à la situation précédente (Fig. 5.1b).



(a) Modèle de sténopé.

(b) Plan image défini devant le centre optique.

FIG. 5.1 – (a) Géométrie de la formation d'une image avec le modèle sténopé. (b) Modèle équivalent où le plan image est défini devant le centre optique. Source [67].

Une projection perspective de l'objet sur le plan image est utilisée, i.e. les rayons se projettent dans l'image via un point unique appelé le centre optique [67], [106]. Les notations suivantes sont employées (Fig. 5.2) :

- R_c est le *repère caméra* dans lequel l'axe z est confondu avec l'axe optique ;
- R_w est le *repère monde* lié à la modélisation de l'objet ;
- $(0, u, v)$ est le *repère image* qui prend en compte l'effet de discrétisation.

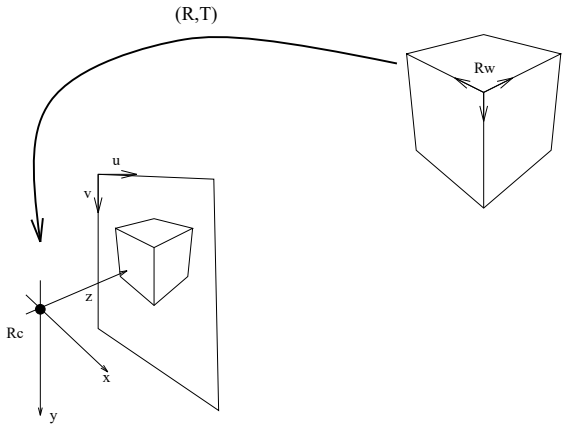


FIG. 5.2 – Représentation du repère monde R_w, du repère caméra R_c et du repère image $(0, u, v)$. Source [67].

Afin de passer les coordonnées 3D des points, $\mathbf{P}_w = [X_w, Y_w, Z_w]^t \in \mathbb{R}^3$, du repère objet au repère caméra, $\mathbf{P}_c = [X_c, Y_c, Z_c]^t \in \mathbb{R}^3$, une transformation est nécessaire. Elle se compose d'une matrice de rotation sur $\mathbb{R}^3$, $\mathbf{R} \in \mathcal{M}_{(3,3)(\mathbb{R})}$, et d'un vecteur de translation, $\mathbf{T} \in \mathbb{R}^3$ (Fig. 5.2). C'est donc une isométrie qui s'écrit avec les points en coordonnées homogènes,

$\mathbf{P}_w = [X_w, Y_w, Z_w, 1]^t \in \mathbb{R}^4$ et $\mathbf{P}_c = [X_c, Y_c, Z_c, 1]^t \in \mathbb{R}^4$, comme il s'ensuit :

$$\mathbf{P}_c = \mathbf{M}_e \cdot \mathbf{P}_w, \tag{5.1}$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{pmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{pmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{5.2}$$

Les notations des points en coordonnées 3D (i.e. dans $\mathbb{R}^3$) ou en coordonnées homogènes (i.e. dans $\mathbb{R}^4$) sont volontairement confondues afin de ne pas alourdir inutilement les notations. La matrice $\mathbf{M}_e$ est appelée matrice des paramètres extrinsèques (à la caméra). La matrice de rotation $\mathbf{R}$ peut s'exprimer avec les 3 angles d'Euler α (rotation autour de l'axe x), β (rotation autour de l'axe y) et γ (rotation autour de l'axe z) :

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{5.3}$$

$$= \begin{bmatrix} \cos(\gamma)\cos(\beta) & \cos(\gamma)\sin(\beta)\sin(\alpha) - \sin(\gamma)\cos(\alpha) & \cos(\gamma)\sin(\beta)\cos(\alpha) + \sin(\gamma)\sin(\alpha) \\ \sin(\gamma)\cos(\beta) & \cos(\gamma)\cos(\alpha) + \sin(\gamma)\sin(\beta)\sin(\alpha) & \sin(\gamma)\sin(\beta)\cos(\alpha) - \cos(\gamma)\sin(\alpha) \\ -\sin(\beta) & \cos(\beta)\sin(\alpha) & \cos(\beta)\cos(\alpha) \end{bmatrix}. \tag{5.4}$$

Le vecteur de translation est égal à $\mathbf{T} = [T_x, T_y, T_z]^t$. Six paramètres extrinsèques seront à estimer : α, β, γ, T_x, T_y et T_z.

Le changement du repère caméra au repère image s'écrit :

$$\begin{cases} x &= X_c f / Z_c \\ y &= Y_c f / Z_c \\ z &= f \end{cases}, \tag{5.5}$$

ou de manière équivalente en coordonnées homogènes :

$$\begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{P}_c. \tag{5.6}$$

Le facteur multiplicatif s est introduit par les notations homogènes.

Dans l'équation (5.6), les coordonnées (x, y) dans le plan image ont comme origine l'axe optique. Un changement de coordonnées s'avère donc nécessaire afin de fixer l'origine dans le coin supérieur gauche de l'image discrétisée par le capteur (Fig. 5.3). Ces coordonnées en pixels dans l'image sont notées (u, v).

La transformation s'écrit :

$$\begin{cases} u &= x/dx + u_0 \\ v &= y/dy + v_0 \end{cases}, \tag{5.7}$$

où (u_0, v_0) représentent les coordonnées (en pixels) dans l'image, de l'intersection de l'axe optique avec le plan image. (dx, dy) sont les dimensions d'un pixel selon les axes x et y de l'image numérisée (Fig. 5.3). En coordonnées homogènes, l'équation (5.7) s'écrit :

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} sx \\ sy \\ s \end{bmatrix}. \tag{5.8}$$

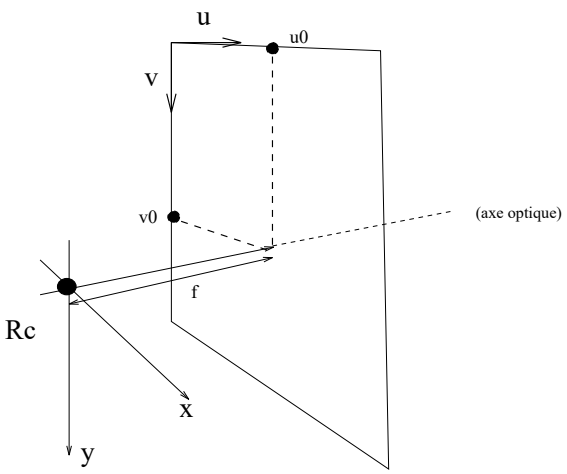


FIG. 5.3 – Changement de repère dans le plan image entre son intersection avec l'axe optique et l'origine de l'image discrétisée par le capteur. Source [67].

En combinant les équations (5.6) et (5.8), on obtient la matrice $\mathbf{M}_i$ de passage entre les coordonnées dans le repère caméra et celles dans le repère image (en pixels) :

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}, \tag{5.9}$$

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \mathbf{M}_i \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}. \tag{5.10}$$

La matrice $\mathbf{M}_i$ est appelée matrice intrinsèque de la caméra car ses valeurs sont intrinsèques à la caméra. En posant $f_x = f/dx$ et $f_y = f/dy$, on a $f_x/f_y = dx/dy$ où le rapport dx/dy représente le rapport pixel. Comme dx est connu (car il est généralement donné par le fabriquant de la caméra), il ne reste que quatre paramètres intrinsèques à estimer : $\{fx, fy, u_0, v_0\}$.

En combinant les équations (5.2), (5.3) et (5.9), le modèle de formation d'image s'écrit donc :

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, \tag{5.11}$$

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \mathbf{M}_i \mathbf{M}_e \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{5.12}$$

5.1.2 Estimation des paramètres du modèle

L'équation (5.11) peut être résolue de manière linéaire. Cependant, une approche non-linéaire issue de la photogrammétrie est préférée, comme présentée dans [67]. Celle-là va permettre d'introduire une correction de la distorsion radiale due à l'objectif de la caméra.

Soit une projection perspective entre un objet 3D observé et son image 2D (avec une hypothèse de modèle sténopé) (Fig. 5.4). L'équation entre un point de l'objet, $\mathbf{P}_{w,i} = [X_i, Y_i, Z_i]^t$ et sa

projection dans l'image $[x_i, y_i, z_i]^t$, où $z_i = f$, s'écrit de la manière suivante :

$$\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \lambda_i \left[\mathbf{R} \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} + \mathbf{T} \right], \tag{5.13}$$

où λ_i est un facteur d'échelle introduit lors du passage du 3D au 2D.

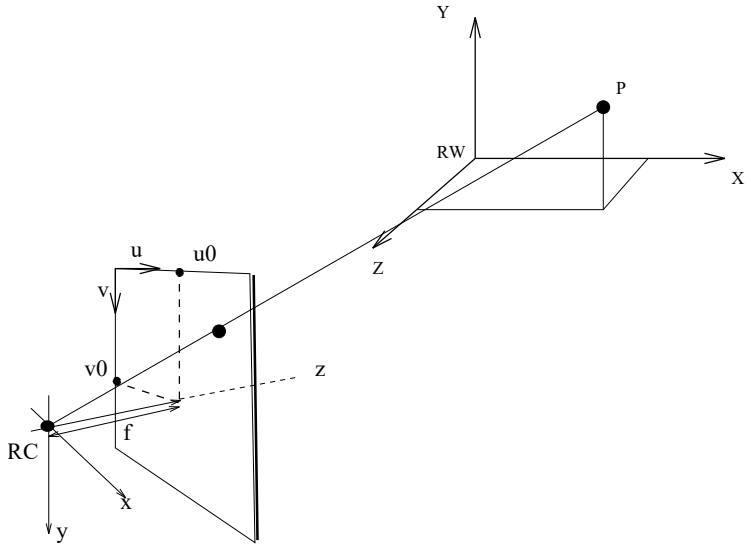


FIG. 5.4 – Projection d'un point $\mathbf{P}$ du monde dans le repère caméra puis dans celui de l'image avec les différents systèmes de coordonnées. Le modèle sténopé est utilisé. Source [67].

En éliminant λ_i dans l'équation 5.13 et en supprimant l'indice i, nous obtenons les *équations de colinéarité* suivantes :

$$\begin{cases} x &= f\frac{r_{11}X+r_{12}Y+r_{13}Z+T_x}{r_{31}X+r_{32}Y+r_{33}Z+T_z} \\ y &= f\frac{r_{21}X+r_{22}Y+r_{23}Z+T_y}{r_{31}X+r_{32}Y+r_{33}Z+T_z} \end{cases}. \tag{5.14}$$

En exprimant (x, y) dans le système de coordonnées pixel de l'image, on en déduit l'équation suivante :

$$\begin{cases} x &= (u + e_x - u_0)dx - do_x \\ y &= (v + e_y - v_0)dy - do_y \end{cases}. \tag{5.15}$$

Les paramètres e_x et e_y sont les erreurs de mesures selon les coordonnées x et y (i.e. les corrections à apporter aux mesures pour qu'il existe une correspondance parfaite avec les données de la fonction de projection). do_x et do_y représentent les composantes de distorsion optique selon x et y. Elles se divisent en deux parties : distorsions *radiales* et *tangentielles*, (i.e. $do_x = do_{xr} + do_{xt}$ et $do_y = do_{yr} + do_{yt}$). Les équations des distorsion sont les suivantes :

$$\begin{cases} do_{xr} &= (u - u_0)dx\left(a_1(r^2 - r_0^2) + a_2(r^4 - r_0^4) + a_3(r^6 - r_0^6)\right) \\ do_{yr} &= (v - v_0)dy\left(a_1(r^2 - r_0^2) + a_2(r^4 - r_0^4) + a_3(r^6 - r_0^6)\right) \end{cases}, \tag{5.16}$$

$$\begin{cases} do_{xt} &= p_1[r^2 + 2(u - u_0)^2dx^2] + 2p_2(u - u_0)dx(v - v_0)dy \\ do_{yt} &= p_2[r^2 + 2(v - v_0)^2dy^2] + 2p_1(u - u_0)dx(v - v_0)dy \end{cases}, \tag{5.17}$$

où :

- a_1, a_2 et a_3 sont les coefficients du polynôme modélisant la distorsion radiale ;
- p_1 et p_2 sont les coefficients du polynôme modélisant la distorsion tangentielle ;

- r_0 est une distance radiale pour laquelle la distorsion est nulle.

En introduisant les équations (5.15), (5.16) et (5.17) dans l'équation (5.14), les équations de colinéarité s'expriment de la manière suivante :

$$\begin{cases} u + e_x &= u_0 + do_{xr} + do_{xt} + f_x \frac{r_{11}X + r_{12}Y + r_{13}Z + T_x}{r_{31}X + r_{32}Y + r_{33}Z + T_z} &= P(\mathbf{\Phi}) \\ v + e_y &= v_0 + (do_{yr} + do_{yt})\frac{f_x}{f_y} + f_y \frac{r_{21}X + r_{22}Y + r_{23}Z + T_y}{r_{31}X + r_{32}Y + r_{33}Z + T_z} &= Q(\mathbf{\Phi}) \end{cases}. \tag{5.18}$$

Le vecteur $\mathbf{\Phi}$ contient 15 paramètres :

$$\mathbf{\Phi} = [u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, T_x, T_y, T_z, \alpha, \beta, \gamma]^t. \tag{5.19}$$

En exprimant les équations de colinéarité (éq. (5.18)) de la manière suivante :

$$\begin{pmatrix} e_x \\ e_y \end{pmatrix} = \begin{pmatrix} P(\mathbf{\Phi}) - u \\ Q(\mathbf{\Phi}) - x \end{pmatrix} = V(\mathbf{\Phi}), \tag{5.20}$$

le problème revient à déterminer la valeur des paramètres $\mathbf{\Phi}$ qui minimisent l'erreur :

$$S = \sum_{i=1}^{n} (e_{x_i}^2 + e_{y_i}^2). \tag{5.21}$$

Pour la résolution, on utilise n points d'une mire étalon dont les distances entre eux sont connues avec une précision très élevée. Pour n points 3D et leurs projetés correspondant dans l'image, il y a donc $2n$ équations (de colinéarité). Comme l'équation (5.20) est non-linéaire, on utilisera une approximation linéaire suivie d'une méthode d'optimisation non-linéaire pour estimer les valeurs des paramètres.

Afin de réduire les erreurs de mesure, l'optimisation sera faite sur plusieurs images I_j, où $j \in [\![1 \dots m]\!]$, capturées par la même caméra mais avec des positions spatiales (i.e. rotations et translations) différentes, $\{T_x^j, T_y^j, T_z^j, \alpha^j, \beta^j, \gamma^j\}_{j \in [\![1 \dots m]\!]}$. Les 9 paramètres intrinsèques $[u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y]$ de la caméra sont les mêmes pour toutes les images. Le calibrage estime les paramètres du vecteur :

$$\begin{aligned} \mathbf{\Phi}_{9+6m} = [&u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, \\ &T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \cdots, T_x^m, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m\Big]^t. \end{aligned} \tag{5.22}$$

Comme il est coûteux d'obtenir une mire d'étalonnage de qualité, il est intéressant, lors du calibrage, d'estimer les positions 3D, (X_i, Y_i, Z_i), des n points de la mire se projetant en coordonnées 2D (u_i, v_i) sur l'image. C'est ce que l'on appelle l'*ajustement de faisceaux* (en anglais *bundle adjustment*) et qui permet de calibrer la caméra avec une mire moins précise (et donc moins onéreuse). Le vecteur de paramètres à estimer sera donc :

$$\begin{aligned} \mathbf{\Phi}_{9+6m+3n} = [&u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, \\ &T_x^1, T_y^1, T_z^1, \alpha^1, \beta^a, \gamma^1, \cdots, T_x^m, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m, \end{aligned} \tag{5.23}$$

$$X_1, Y_1, Z_1, \cdots, X_n, Y_n, Z_n]^t. \tag{5.24}$$

On cherche les paramètres du vecteur $\mathbf{\Phi}_{9+6m+3n}$ qui minimisent l'erreur :

$$S = \sum_{j=1}^{m} \sum_{i=1}^{n} (e_{x_{ij}}^2 + e_{y_{ij}}^2). \tag{5.25}$$

Le nombre d'équations est égale à $2nm$ et le nombre d'inconnues à $9 + 6m + 3n$. Pour une mire de $n = 11$ points, un minimum de $m = 3$ images est nécessaire pour surdéterminer le système d'équations.

Comme la matrice des paramètres intrinsèques est toujours déterminée à un facteur d'échelle près, on impose la contrainte $dx = 1$ afin de fixer l'ensemble des paramètres en une unité arbitraire de pixels.

Pour les paramètres extrinsèques, la géométrie du système est aussi fixée à un facteur d'échelle près, puisqu'il est toujours possible de trouver une mire plus volumineuse observée de plus loin et qui donnerait la même image. Pour l'étalonnage d'une caméra, cette perte de « métrique » (i.e. de distance de référence) n'a que peu d'importance puisque ce sont les paramètres extrinsèques qui ont un intérêt. Néanmoins, deux points de la mire peuvent ne pas être reconstruits afin d'imposer la « métrique » du système. Afin de fixer la géométrie extrinsèque du système, une des coordonnées d'un point quelconque parmi n restera figée.

5.1.3 Calibrage d'un système de stéréovision par appareil photographique

Comme illustré sur la figure 5.5, la même caméra (immobile) capture plusieurs images d'une mire dite « en damier » avec des orientations différentes.

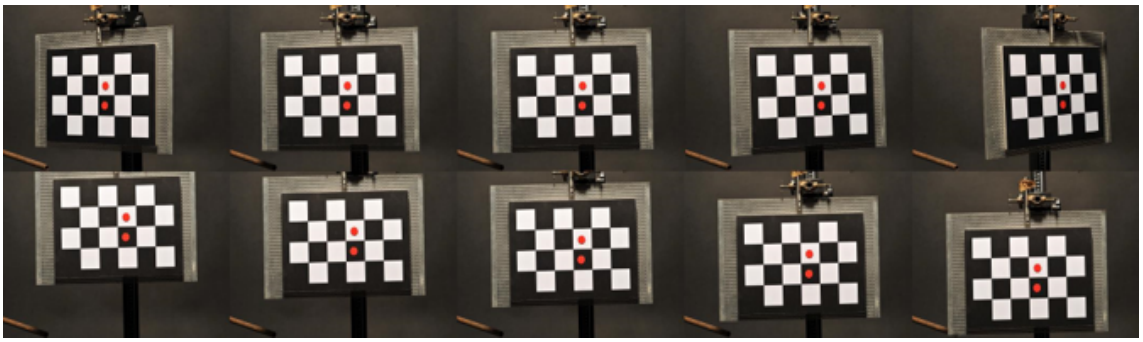


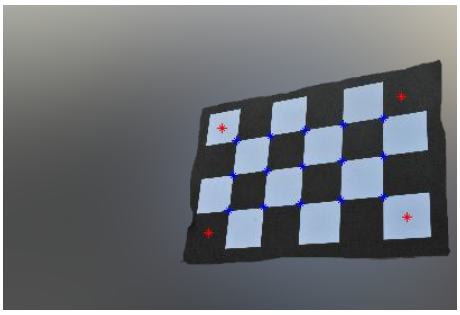
FIG. 5.5 – Acquisitions de différentes vues d'une mire damier. Source [312].

On connaît (de manière assez approximative) la longueur du côté de chaque carré de la mire qui a été réalisée par impression. Il ne s'agit donc pas d'une mire de haute précision. Pour chaque image, on déterminera les 6 paramètres extrinsèques (i.e. rotation et translation) et les 9 paramètres intrinsèques. En combinant l'ensemble des images, les positions 3D précises des points de la mire seront aussi estimées. Dans un premier temps, les coordonnées des points d'intersection de la mire sont extraites de manière approximative dans les images (Fig. 5.6) grâce à des opérations morphologiques. Puis, un ajustement fin est effectué par ajustement de quatre droites sur la norme du gradient de l'image (Fig. 5.7). Le centre du parallélogramme correspond au coordonnées (u_i, v_i) du point d'intersection avec une précision sub-pixelique (Fig. 5.7f). Ce modèle permet d'améliorer la précision du détecteur.

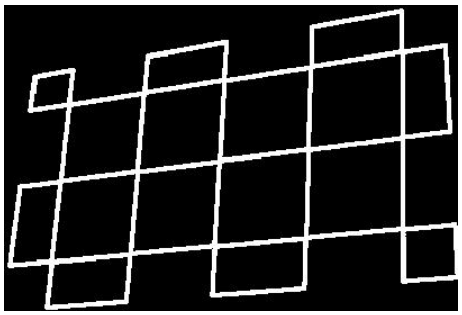
La détection des points 2D dans les images doit être la plus précise possible car elle est déterminée une fois pour toutes (i.e. elle ne variera pas dans l'estimation des paramètres modèle).

Les paramètres du système sont ensuite estimés en deux étapes : une initialisation suivie d'une optimisation non-linéaire. L'initialisation concerne les paramètres suivants.

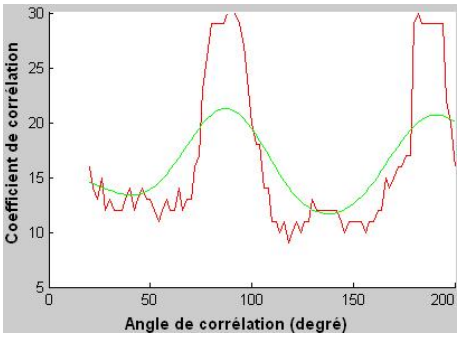
- Les paramètres intrinsèques $[f_x, f_y, u_0, v_0]^t$ peuvent être initialisés directement. u_0 et v_0 correspondent au centre de l'image. La valeur de la focale f peut être récupérée dans les données Exif (*Exchangeable image file format*) incluses avec le fichier de sauvegarde de l'image ou dans la documentation de la caméra. Cette dernière contient les dimensions du capteur ; ce qui permet d'en déduire celles d'un pixel (dx, dy). De manière alternative, la méthode de Zhang [337] permet d'initialiser automatiquement les paramètres intrinsèques grâce à l'analyse des matrices d'homographies (i.e. les matrices carrées de taille 3 non-singulières).
- Les paramètres extrinsèques $\left[T_x^j, T_y^j, T_z^j, \alpha^j, \beta^j, \gamma^j\right]^t_{j\in[\![1\ldots m]\!]}$ sont initialisés par la méthode Zhang [337] ou, si l'on connaît préalablement les paramètres intrinsèques, par celle de DeMenthon [64], [235].



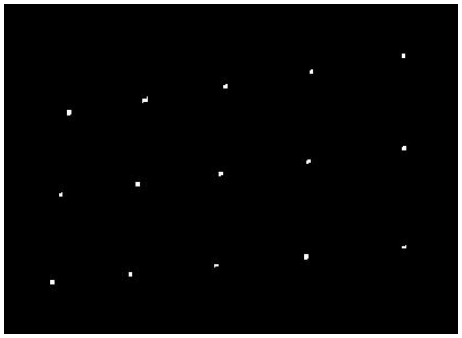
(a) Mire dans une position.



(b) Détection du damier.



(c) Courbe de corrélation directionnelle.



(d) Détection approximative des points d'intersection.

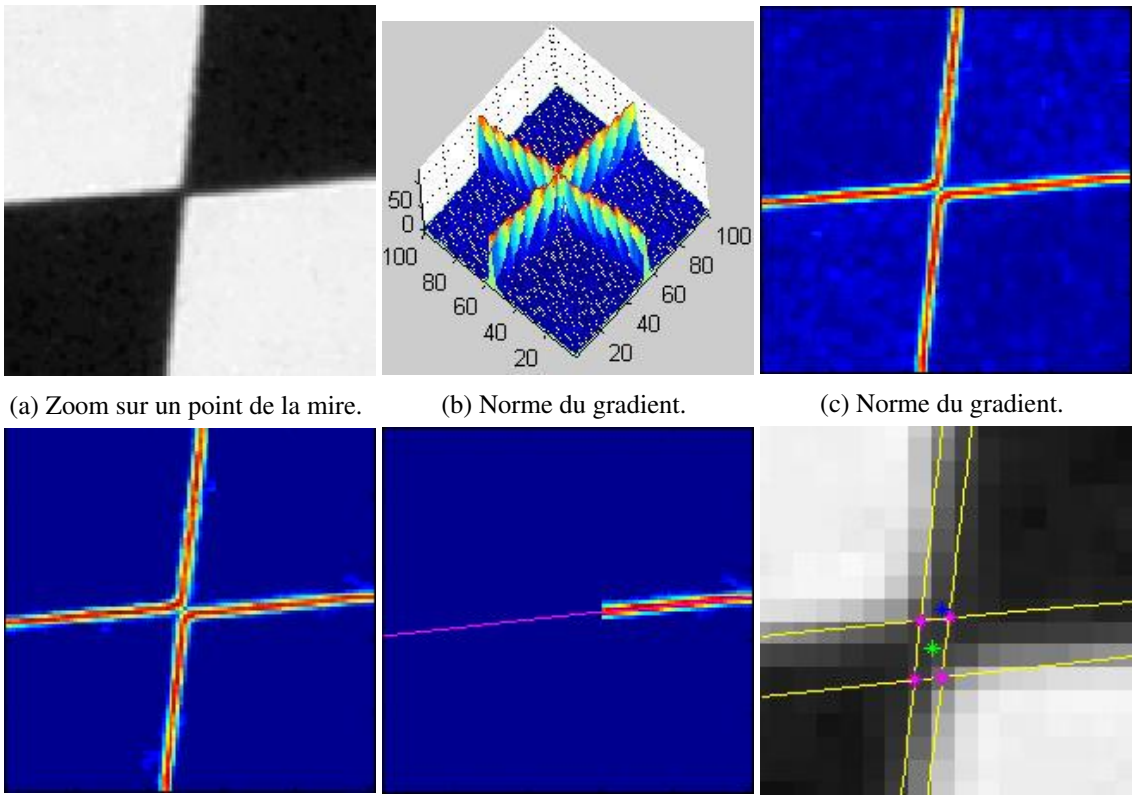
FIG. 5.6 – (a) Image de la mire en damier. (b) Détection des frontières du damier par des opérations morphologiques. (c) Courbes de corrélation directionnelle pour pouvoir faire un filtrage directionnel [284]. (d) Détection approximative des points d'intersection par filtrage directionnel. Source [312].

- Les paramètres de distorsion $[a_1, a_2, a_3, p_1, p_2]^t$ sont en général initialisés à la valeur nulle (lorsque les déformations ne sont pas trop importantes).
- Les positions 3D des points de la mire dans l'espace $[X_i, Y_i, Z_i]^t_{i\in[\![1\ldots n]\!]}$ sont initialisées grâce à une mesure directe sur la mire.

Après l'initialisation des paramètres du vecteur $\mathbf{\Phi}_{9+6m+3n}$ (éq. (5.24)), ceux-ci sont affinés grâce à une méthode d'optimisation non-linéaire. Cette dernière minimise l'erreur S (éq. (5.25)) entre les points 2D de la mire extraits dans l'image et ceux projetés par le modèle. Les algorithmes de *Gauss-Newton* [35] ou de *Levenberg-Marquardt* [35], [193] peuvent être utilisés pour l'optimisation.

Afin d'avoir des dynamiques comparables entre les paramètres lors de l'optimisation non-linéaire, ceux-ci sont normalisés de la manière suivante.

- Les paramètres intrinsèques $[f_x, f_y, u_0, v_0]^t$ sont divisés par leur valeurs initiales afin d'avoir un amplitude proche de 1.
- Les angles de rotation $[\alpha^j, \beta^j, \gamma^j]^t_{j\in[\![1\ldots m]\!]}$ sont divisés par 180° afin que leurs valeurs varient dans l'intervalle $[-1; 1]$.
- Les translations $\left[T^j_x, T^j_y, T^j_z\right]^t_{j\in[\![1\ldots m]\!]}$ sont divisées par $\max_j \{|T^j_x|\}$, $\max_j \{|T^j_y|\}$ et $\max_j \{|T^j_z|\}$.
- Les paramètres de distorsions $[a_1, a_2, a_3, p_1, p_2]^t$ sont divisés par 1×10^{-6} car leurs valeurs sont très faibles.
- Les coordonnées 3D de la mire $[X_i, Y_i, Z_i]^t_{i\in[\![1\ldots n]\!]}$ sont normalisées par 1×10^{-3} dans cet exemple.



(a) Zoom sur un point de la mire. (b) Norme du gradient. (c) Norme du gradient.

(d) Suppression du bruit de fond. (e) Ajustement d'une droite. (f) Modèle de point.

FIG. 5.7 – (a) Zoom sur un point de la mire en damier. (b) et (c) Norme du gradient. (d) Suppression du bruit de fond par seuillage. Le gradient apparaît vu de dessus comme deux hyperboles (e) Ajustement d'une droite sur un des « bras » du gradient. (f) Ajustement du modèle de points. Source [312].

Afin de garantir la convergence de l'algorithme d'optimisation, les paramètres sont optimisés itérativement et plusieurs fois selon les catégories suivantes :

i) les paramètres intrinsèques $[f_x, f_y, u_0, v_0]^t$;

ii) les angles de rotation $\left[\alpha^j, \beta^j, \gamma^j\right]^t_{j\in[\![1\ldots m]\!]}$;

iii) les paramètres de distorsions $[a_1, a_2, a_3, p_1, p_2]^t$;

iv) les coordonnées 3D de la mire $[X_i, Y_i, Z_i]^t_{i\in[\![1\ldots n]\!]}$.

Ensuite une optimisation finale est réalisée sur le vecteur $\mathbf{\Phi}_{9+6m+3n}$ (éq. (5.24)) de tous les paramètres.

5.1.4 Reconstruction 3D

Afin de reconstruire un point 3D il est nécessaire d'avoir deux images du même objet acquises dans des poses différentes (Fig. 5.8).

La reconstruction 3D nécessite un appariement des points dans les images. Il existe différents algorithmes pour effectuer cette mise en correspondance tels que le *SIFT* (*Scale-Invariant Feature Transform*) [166], *DAISY* [301], la corrélation sur un mouchetis [274] ou plus récemment l'appariement par apprentissage profond : *DeepMatching* [256]. Lorsque l'on connaît la transformation, $[\mathbf{RT}]$, composée d'une rotation et d'une translation, entre les deux poses de la caméra, on peut appliquer la contrainte épipolaire : un point d'une image est transformé en une ligne sur l'autre image (Fig. 5.8). Le champ de recherche du point correspondant peut donc être restreint à cette ligne.

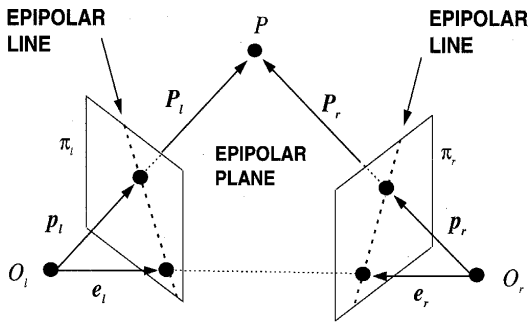


FIG. 5.8 – Principe de reconstruction 3D par stéréovision et géométrie épipolaire. $\mathbf{O}_l$ et $\mathbf{O}_r$ sont les centres optiques des caméras gauche et droite, respectivement. $\mathbf{P}_l = [X_l, Y_l, Z_l]^t$ et $\mathbf{P}_r = [X_r, Y_r, Z_r]^t$ sont les coordonnées du même point $\mathbf{P}$ dans les repères des caméras gauche et droite. $\mathbf{p}_l$ et $\mathbf{p}_r$ sont les projetés du point $\mathbf{P}$ dans les plans gauche π_l et droit π_r de la caméra. Source [302].

Pour effectuer la reconstruction, la méthode de Trucco est utilisée [302, chap. 7]. Comme illustré sur la figure 5.8, le point $\mathbf{P}$ est projeté en une paire de points $\mathbf{p}_l$ et $\mathbf{p}_r$ dans les plans π_l et π_r des caméras gauche (*left*) et droite (*right*), respectivement. $\mathbf{O}_l$ et $\mathbf{O}_r$ sont les centres optiques des deux caméras gauche et droite. Le point $\mathbf{P}$ est à l'intersection des deux rayons de $\mathbf{O}_l$ à $\mathbf{p}_l$ et de $\mathbf{O}_r$ à $\mathbf{p}_r$. Dans un cas idéal, les deux rayons sont exactement connus et s'intersectent dans l'espace. Cependant, comme les paramètres du système et les point projetés $\mathbf{p}_l$ et $\mathbf{p}_r$ ne sont qu'approximativement connus, *les deux rayons ne s'intersectent pas* dans l'espace. Leur intersection peut être estimée comme étant le point $\mathbf{P}'$ de distance minimum entre les deux rayons, comme illustré sur la figure 5.9.

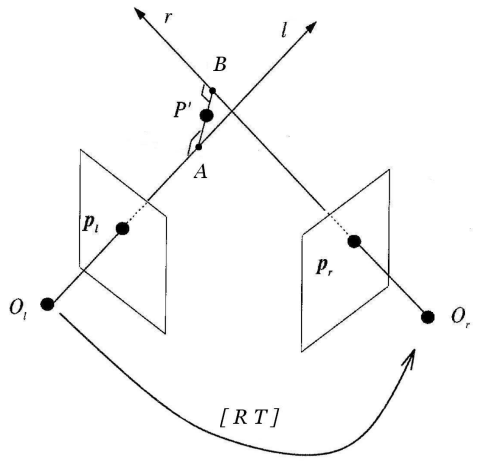


FIG. 5.9 – Reconstruction 3D par triangulation avec des rayons qui ne s'intersectent pas. Source [302].

Présentons maintenant le détail de la méthode de calcul par triangulation. Lors du calibrage du système de stéréovision à deux caméras, les matrices de rotations ($\mathbf{R}_l$ et $\mathbf{R}_r$) et de translations ($\mathbf{T}_l$ et $\mathbf{T}_r$) de chacune des caméras par rapport à la mire sont déterminées. On en déduit la transformation $[\mathbf{RT}]$ entre les deux caméras où la matrice de rotation $\mathbf{R}$ et le vecteur de translation $\mathbf{T}$ sont définis par :

$$\mathbf{R} = \mathbf{R}_r \mathbf{R}_l^t \,, \tag{5.26}$$

$$\mathbf{T} = \mathbf{T}_l - \mathbf{R}^t \mathbf{T}_r \,. \tag{5.27}$$

Soit $a\mathbf{p}_l$ ($a \in \mathbb{R}$) le rayon l passant par $\mathbf{O}_l$ et le projeté $\mathbf{p}_l$. Lorsque $a = 0$, nous avons $a\mathbf{p}_l = \mathbf{O}_l$ et

lorsque $a = 1$, nous avons $a\mathbf{p}_l = \mathbf{p}_l$. Soit $\mathbf{T}+b\mathbf{R}^t\mathbf{p}_r$ ($b \in \mathbb{R}$) le rayon r passant par $\mathbf{O}_r$ et le projeté $\mathbf{p}_r$ exprimé dans le repère du plan caméra gauche. Lorsque $b = 0$, nous avons $\mathbf{T}+b\mathbf{R}^t\mathbf{p}_r = \mathbf{T}$ (i.e. le point $\mathbf{O}_r$ exprimé dans le repère de la caméra gauche) et lorsque $b = 1$, nous avons $\mathbf{T} + \mathbf{R}^t\mathbf{p}_r$ (i.e. le point $\mathbf{p}_r$ exprimé dans le repère de la caméra gauche). Soit $\mathbf{w}$ le vecteur orthogonal aux deux rayons l et r. Soient $\mathbf{A}$ et $\mathbf{B}$ les point d'intersection de la droite portée par $\mathbf{w}$ avec les rayons l et r (Fig. 5.9). Notre problème revient à trouver le point milieu, $\mathbf{P}'$, du segment $[\mathbf{AB}]$. Cherchons les coefficients a_0 et b_0 tels que $a_0\mathbf{p}_l = \mathbf{A}$ et $\mathbf{T} + b_0\mathbf{R}^t\mathbf{p}_r = \mathbf{B}$. Le vecteur $\mathbf{w}$ peut s'exprimer comme $\mathbf{w} = \mathbf{p}_l \times \mathbf{R}^t\mathbf{p}_r$, où $\times$ est le produit vectoriel. Nous avons donc :

$$\begin{aligned} &\mathbf{A} + c\mathbf{w} = \mathbf{B} \\ \Longleftrightarrow\ & a_0\mathbf{p}_l - b_0\mathbf{R}^t\mathbf{p}_r + c_0(\mathbf{p}_l \times \mathbf{R}^t\mathbf{p}_r) = \mathbf{T}. \end{aligned} \tag{5.28}$$

Comme cette équation est valable pour les 3 composantes des points $\mathbf{p}_l$ et $\mathbf{p}_r$, il y a une unique solution aux trois inconnues a_0, b_0 et c_0. On en déduit la position du point $\mathbf{P}'$ comme étant le milieu du segment $[\mathbf{AB}]$, $\mathbf{P}' = a_0\mathbf{p}_l + \frac{c_0}{2}(\mathbf{p}_l \times \mathbf{R}^t\mathbf{p}_r)$.

La précision de reconstruction peut être estimée par la longueur du segment $[\mathbf{AB}]$. Plus la longueur du segment est courte, plus la reconstruction sera précise. La méthode de reconstruction peut aussi être étendue à plus de deux caméras. Une fois la reconstruction terminée pour l'ensemble de points, on obtient un nuage de points sur lequel un maillage peut être appliqué.

Thibaud Vilboux a ensuite estimé la précision de reconstruction du système de stéréovision sur une mire (Fig. 5.10c). Il a utilisé une mire en damier (Fig. 5.10a) pour calibrer les caméras du système de stéréovision (Fig. 5.10b). Puis, il a reconstruit une mire tridimensionnelle en fibres de carbone et la mire en damier. Sur les images de la mire en damier, il a utilisé un détecteur sub-pixelique de points, alors que pour les images de la mire en carbone le choix des points a été fait manuellement (i.e. de manière beaucoup moins précise).

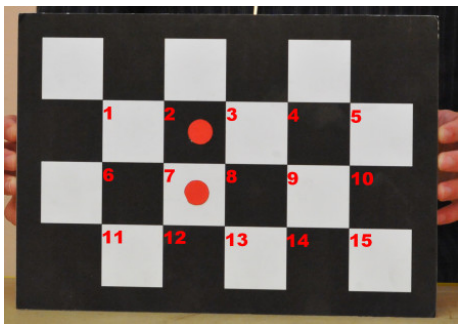
La figure 5.11 illustre 6 paires d'images acquises pour une pose de chacune des deux caméras du système de stéréovision. Les images montrent la mire en fibres de carbone selon 6 orientations différentes. Sur les images de cette mire, 12 points sont sélectionnés manuellement (Fig. 5.10b) pour faire la reconstruction 3D. La figure 5.12 montre la reconstruction de la mire selon les 6 orientations.

Thibaud Vilboux a pu estimer une précision de reconstruction en faisant la moyenne des longueurs des segments $[\mathbf{AB}]$ normaux aux rayons pour chacun des points reconstruits dans les 6 orientations de la mire (Tab. 5.1). Deux distances entre la caméra et la mire ont été considérées : 1,5 m et 2,5 m. On constate qu'avec une distance plus faible, la précision de reconstruction est meilleure. En effet, plus la distance caméra/mire est faible, plus le nombre de pixels décrivant la mire est élevé. L'estimation de la position des points sur les images est donc plus précise à plus faible distance.

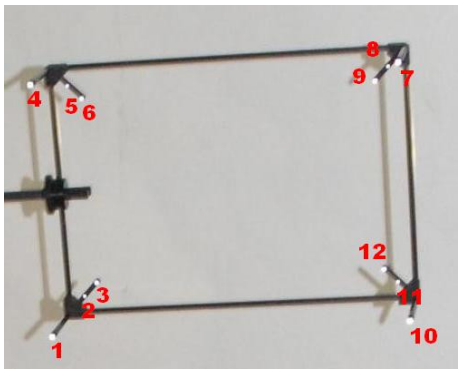
TABLE 5.1 – Moyenne des précisions de reconstruction de chaque point des mires acquises dans différentes positions et pour deux distances caméra/objet différentes. Source [312].

	Acquisition à 1,5 m	Acquisition à 2,5 m
Mire en carbone	0,218 mm	0,316 mm
Mire damier	0,065 mm	0,084 mm

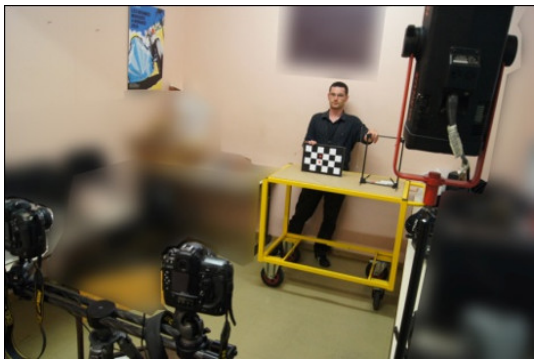
Afin de reconstruire un point en 3D, il est aussi possible de le faire par optimisation non-linéaire. Grâce à l'équation (5.12), p. 190, le vecteur $\mathbf{\Gamma}_i = [X_i, Y_i, Z_i]^t$, représentant les coordonnées des points à reconstruire, est optimisé en fixant les valeurs des paramètres des matrices intrinsèque $\mathbf{M}_i$ et extrinsèque $\mathbf{M}_e$ qui, elles, ont été déterminées lors du calibrage. En pratique, cette méthode d'estimation s'avère être très efficace .



(a) Mire en damier.



(b) Mire en fibres de carbone.



(c) Système de stéréovision.

FIG. 5.10 – (a) Mire en forme damier utilisée pour le calibrage puis pour la reconstruction. (b) Mire en fibre de carbone utilisée pour la reconstruction. (c) Photo du montage. Source [312].

(a) Pose 1, ori 1. (b) Pose 1, ori 2. (c) Pose 1, ori 3. (d) Pose 1, ori 4. (e) Pose 1, ori 5. (f) Pose 1, ori 6.

(g) Pose 2, ori 1. (h) Pose 2, ori 2. (i) Pose 2, ori 3. (j) Pose 2, ori 4. (k) Pose 2, ori 5. (l) Pose 2, ori 6.

FIG. 5.11 – (a-f) Poses n° 1 de la caméra pour 6 orientations (ori) différentes de la mire. (g-l) Poses n° 2 de la caméra pour 6 orientations (ori) différentes de la mire. Source [312].

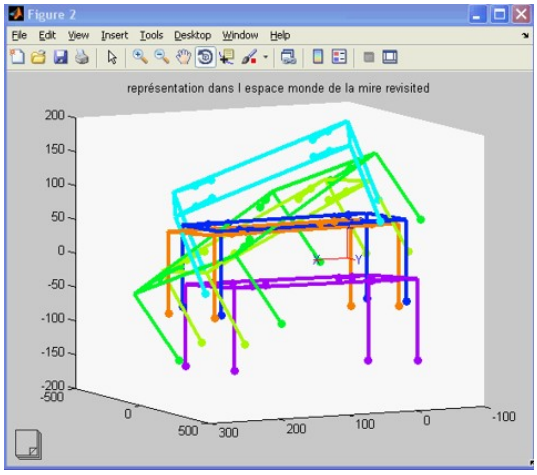


FIG. 5.12 – Reconstruction de la mire en carbone selon 6 orientations. Source [312].

5.2 Système de stéréovision par radiographie

Cette section n'est pas présentée dans le mémoire en ligne.

5.3 Conclusion du chapitre

Dans ce chapitre, j'ai présenté mes travaux sur la reconstruction 3D par stéréovision dans le domaine du visible, par photographie, et dans celui des rayons X, par radiographie. Le principe général consiste à calibrer le système de stéréovision grâce à une mire afin d'obtenir les paramètres de projection d'un objet sur l'image. Dans le domaine du visible, le stage de Thibaud Vilboux [312] et mes travaux ont permis de comprendre et programmer en détail la chaîne complète de reconstruction 3D. Puis, j'ai développé cette chaîne pour un système de stéréovision par rayons X. J'ai conçu une méthode pour estimer les paramètres du modèle de projection dans le domaine des rayons X. J'ai effectué la reconstruction 3D d'une mire et des nappes métalliques d'un pneumatique après avoir extrait ces dernières par filtrage dans les radiographies. La difficulté a consisté à enchaîner les différentes étapes depuis l'acquisition, l'extraction des objets sur les images, le calibrage et la reconstruction tout en limitant l'erreur de reconstruction.

Des premiers résultats très encourageants ont été présentés afin de valider le principe. Il conviendrait de les affiner en vue d'améliorer la précision de reconstruction.

Chapitre 6

Recalage d'images pour les pneumatiques et pour le fond d'œil

En traitement d'image, dès que plusieurs acquisitions du même objet sous des vues différentes sont disponibles, il est souvent intéressant de les superposer entre elles. Cependant, à cause de différents phénomènes de déformations entre les vues, la superposition directe (i.e. rigide qui comprend une rotation, une translation et une mise à l'échelle) n'est en général pas possible. Il est alors nécessaire de déformer une (ou les deux) image(s) afin de la (ou les) recaler l'une sur l'autre. Dans ce chapitre, nous verrons deux problématiques :

i) le recalage entre une épure (i.e. un plan) d'une coupe radiale d'un pneumatique et une image de cette coupe à l'aide de Formes Libres de Déformations (en anglais *Free Form Deformation*) dont la définition repose sur l'utilisation de B-splines [264] ;

ii) le recalage entre deux photographies du fond d'œil capturées à un an d'intervalle grâce à un modèle de déformation [210], [234].

6.1 Recalage d'images pour les coupes de pneumatiques

Cette section n'est pas présentée dans le mémoire en ligne.

6.2 Recalage d'images du fond d'œil dans les grandes banques de données de santé publique

Afin de superposer les images du fond d'œil acquises avec une année d'intervalle pour le suivi de la rétinopathie diabétique, j'ai développé un modèle de recalage adapté à la problématique. Il se compose : d'une correction des distorsions radiales sur chacune des deux images et d'une transformation affine (i.e. une translation, une rotation et une mise à l'échelle anisotrope). Les paramètres de ce modèle ont été estimés en minimisant l'erreur de recalage de deux manières :

i) d'abord sur des points caractéristiques extraits et appariés par la méthode SIFT (*scale-invariant feature transform*) [234] ;

ii) puis, sur la ligne centrale des vaisseaux après une initialisation avec les points SIFT [210].

6.2.1 Ajustement d'un nouveau modèle de recalage sur les points SIFT

Mes travaux présentés dans cette partie ont été publiés dans l'article de journal [234]. Ils ont été réalisés en collaboration avec l'équipe du Professeur David Owens, diabétologue de *Swansea University* et le centre national de dépistage *Diabetic Eye Screening Wales*, de Cardiff au Pays de Galle, Royaume-Uni.

Présentation du problème

L'existence de programmes de dépistage de la Rétinopathie Diabétique (RD) depuis de nombreuses années, permet maintenant d'envisager une analyse longitudinale (i.e. temporelle) des

images du fond d'œil d'un même patient. Cependant, afin de comparer de manière précise les évolutions de la RD au cours du temps, les images doivent pouvoir être parfaitement superposées.

La superposition directe de deux images du même patient ne donne jamais de bons résultats (cf. Fig. 6.1). En effet, pour deux examens photographiques de l'œil, le patient n'est jamais exactement dans la même position et, en outre, la caméra peut avoir changé. Donc, la méthode de superposition doit tenir compte des différentes causes de la déformation comme :

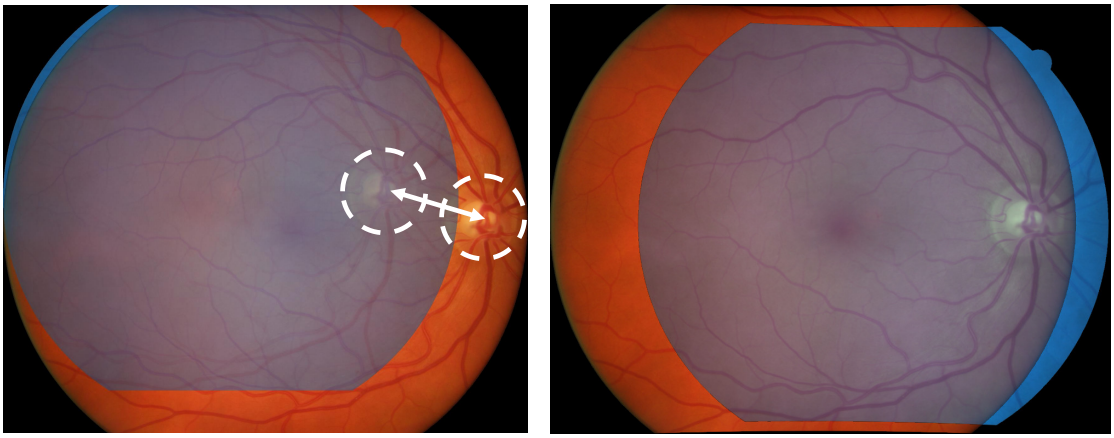
i) la position du patient : en prenant en compte la rotation, la translation et la mise à l'échelle ;

ii) le changement de la caméra : en utilisant une mise à l'échelle et une correction radiale ;

iii) la projection d'une scène 3D assimilée à une sphère (i.e. la rétine de l'œil) sur le plan (2D) du capteur de la caméra : en utilisant une correction radiale ;

iv) la déformation radiale due à l'objectif de la caméra : en utilisant une correction radiale ;

v) la variabilité couleur entre des images due à l'intensité lumineuse et au capteur.



(a) Superposition directe. (b) Superposition avec modèle.

FIG. 6.1 – Superposition d'images du fond d'œil acquises avec des caméras et des résolutions d'images différentes. (a) Superposition directe et infructueuse due à l'absence de modèle adapté. (b) Superposition correcte grâce à un modèle. Afin de mieux visualiser les résultats, une des images du fond d'œil est représentée avec une fausse couleur bleutée.

Afin d'effectuer une superposition – aussi appelée recalage – deux étapes sont considérées comme nécessaires : a) un modèle de déformation et b) un critère de mise en correspondance pour ajuster le modèle. Le recalage d'images médicales a été un champ très actif de recherche durant les dernières décennies, spécialement dans le domaine de la radiologie. Pour une vue d'ensemble des méthodes de recalage d'image médicales se référer e.g. à [168], [237], [311]. Pluim et al. [244] ont aussi écrit un rapport à propos du recalage d'images médicales à partir de l'information-mutuelle. Sotiras et al. [285] ont écrit à propos du recalage déformable. Dans cette sous-section, nous nous concentrerons sur le recalage des images du fond d'œil.

Plusieurs modèles de superposition existent entre des paires d'images du fond d'œil. Les premières méthodes dépendaient des images à la fluorescéine et utilisaient une composition de translation, rotation et mise à l'échelle – i.e. un modèle de transformation affine [153], [332]. Les bifurcations des vaisseaux ont été employées pour mettre en correspondance les points et ajuster le modèle. Un autre critère d'appariement consiste à minimiser les différences d'intensité des images [6], [52], [179], [259]. La mise en correspondance par des points saillants grâce au détecteur SURF – *Speeded Up Robust Features* – a été utilisée dans [46].

D'autres méthodes utilisent une similarité (i.e. une rotation et une translation) [165], [179] et un modèle élastique de déformation [78], [328]. Dans [50], [88], de nouveaux descripteurs PIIFD

(*Partial Intensity Invariant Feature Descriptor*) ont été introduits pour la superposition d'images multimodales d'auto-fluorescence, infrarouges et anérythres (« sans rouge », i.e. avec un filtre centré sur la longueur d'onde du vert qui permet de supprimer les couleurs rouges [175], [271]). Cette méthode utilise un modèle de déformation affine et quadratique. Elle s'avère très efficace pour le recalage d'images multimodales, mais n'a pas été conçue et testée pour la superposition des images couleur du fond d'œil.

Plus récemment, il a été montré qu'un modèle quadratique donne de meilleurs résultats de recalage [6], [42], [48], [266], [291]. La difficulté inhérente à ces modèles est d'estimer leurs paramètres. Pour dépasser cette limite, un modèle de distorsion radiale a été introduit par Lee et al. [159] et comparé aux méthodes précédentes dans [161]. Il consiste à ajouter un modèle radial à une transformation affine dans le but de corriger les effets d'une distorsion radiale due à la géométrie de la caméra et de l'œil. La superposition des images du fond d'œil a aussi été effectuée dans l'espace tridimensionnel en utilisant un modèle de surface presque plane (en anglais *near planar surface*) [163] ou un modèle d'ellipsoïde [115].

Cependant, pour toutes ces méthodes la superposition des images du fond d'œil est effectuée avec des images couleur acquises par la même caméra durant le même examen. Superposer des images couleur acquises par différentes caméras avec au moins un an d'intervalle et enregistrées dans de grandes banques de dépistages reste encore un défi. Deux nouveaux problèmes apparaissent :

i) la distorsion radiale doit être corrigée pour les deux lentilles différentes, parce que la caméra a changé et

ii) la couleur de l'œil peut avoir changé à cause d'une source de lumière différente ou d'une évolution de l'anatomie des yeux (e.g. enlèvement de la cataracte, évolution de la rétinopathie à un stade plus sévère, etc.).

Actuellement, pour les images couleur du fond d'œil, aucune méthode existante n'a été conçue pour résoudre ce problème de changement : des lentilles de la caméra, de la caméra et de la lumière pour les examens rétiniens utilisés dans le dépistage de la RD (i.e. à un intervalle annuel). En outre, dans les méthodes existantes, même lorsqu'une distorsion radiale est corrigée par une correction de distorsion radiale ou par un modèle polynomial de second ordre, le critère de recalage nécessite une intensité (i.e. une couleur) similaire et uniforme sur les images, particulièrement si l'extraction de caractéristiques anatomiques comme les vaisseaux s'avère nécessaire. Ces méthodes fonctionnent dans les conditions de laboratoire (plusieurs acquisitions le même jour avec le même équipement) mais échouent lorsqu'elles sont utilisées sur des banques de données de santé publique. Ces banques ont été constituées durant de nombreuses années, grâce à un dépistage en routine de patients avec du diabète.

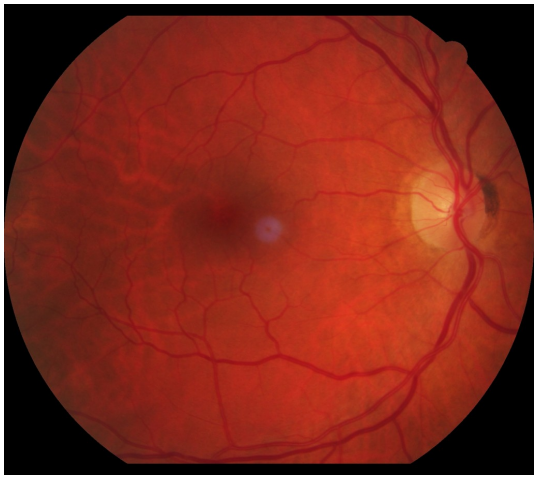
Dans l'article [234], notre contribution a été de répondre à ce défi en présentant une méthode de superposition conçue pour le dépistage longitudinal dans les grandes banques d'images de santé publique. Après avoir pris en compte la variation de couleur, nous introduisons un nouveau modèle de recalage d'image composé de deux distorsions radiales et d'une homographie. Surtout, ce modèle est inversible ce qui est utile pour la déformation de l'image (en anglais *image warping*) et le rend adapté pour du calcul rapide ; un facteur clef pour une analyse efficace des grandes banques de données. Nous introduisons aussi des équations linéaires pour estimer les paramètres du modèle d'une manière simple et rapide. Finalement, nous créons un cadre de travail complet pour la superposition robuste d'images. Le critère de recalage utilise une méthode habituelle d'extraction de point saillants, qui ne repose pas sur l'extraction de caractéristiques anatomiques.

Méthode

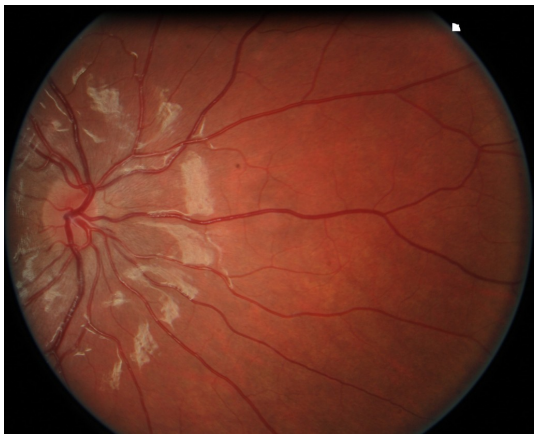
Pour le dépistage de la RD, au cours d'un examen photographique de l'œil, les deux images du fond d'œil (i.e. de la rétine) des deux yeux sont acquises dans deux positions : i.e. un champ « nasal » et « maculaire » de 45° (Fig. 6.2). Comme les patients avec du diabète sont dépistés

annuellement, cela donne une série d'images disponibles pour l'analyse longitudinale. C'est un problème différent des grandes mosaïques d'images du fond d'œil acquises durant le même examen avec la même caméra [48]. Par conséquent, le but de cette étude était de développer une méthode pour superposer les images dans les mêmes positions, même si elles ont été capturées au cours de deux examens différents ou souvent avec des caméras et des résolutions différentes.

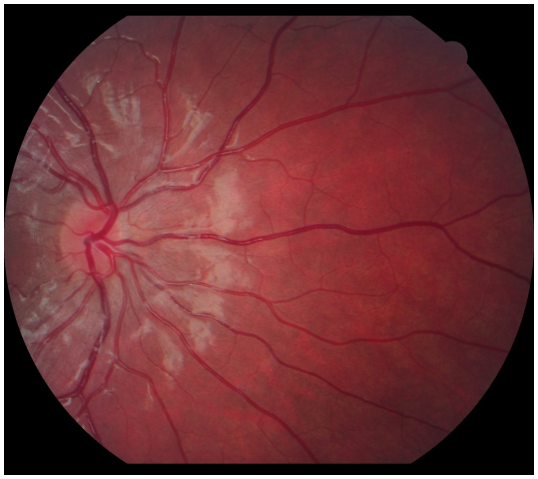
(a) Patient n° 1, champ maculaire.



(b) Patient n° 1, champ maculaire.



(c) Patient n° 2, champ nasal.



(d) Patient n° 2, champ nasal.

FIG. 6.2 – (a) et (b) Images du fond d'œil « maculaires » acquises au cours de deux examens à un an d'intervalle sur un patient (n° 1). (c) et (d) Images du fond d'œil « nasales » acquises au cours de deux examens à un an d'intervalle sur un autre patient (n° 2). Pour les deux patients : différentes caméras, résolutions et conditions d'éclairement ont été utilisées.

Une description schématique de l'étude est représentée sur la figure 6.3. Les différentes étapes ont été conçues pour fournir des solutions efficaces à la superposition d'images acquises pour du dépistage pratique. Entre deux examens, la caméra et les conditions d'éclairement peuvent avoir changé, ce qui produit des différences dans les couleurs et nécessite une étape de pré-traitement. Celle-ci consiste à (1) normaliser la couleur de l'image du fond d'œil par la méthode que j'ai introduite dans [233]. En outre, la position, la résolution et la distorsion radiale entre des images peuvent être différentes, ce qui requiert une transformation de l'image estimée grâce à (2) l'extraction de points caractéristiques dans des paires d'images, (3) une procédure de mise en correspondance, (4) l'utilisation d'un modèle corrigeant la distorsion radiale des deux images et (5) l'estimation des paramètres du modèle par une optimisation robuste. Finalement, la méthode est (6) validée sur un montage simulé et les superposition d'images de la banque de données sont vérifiées

et (7) comparées aux méthodes de l'état de l'art.

Les paires d'images du même œil dans la même position, « nasale » ou « maculaire », sont recalées indépendamment les unes des autres.

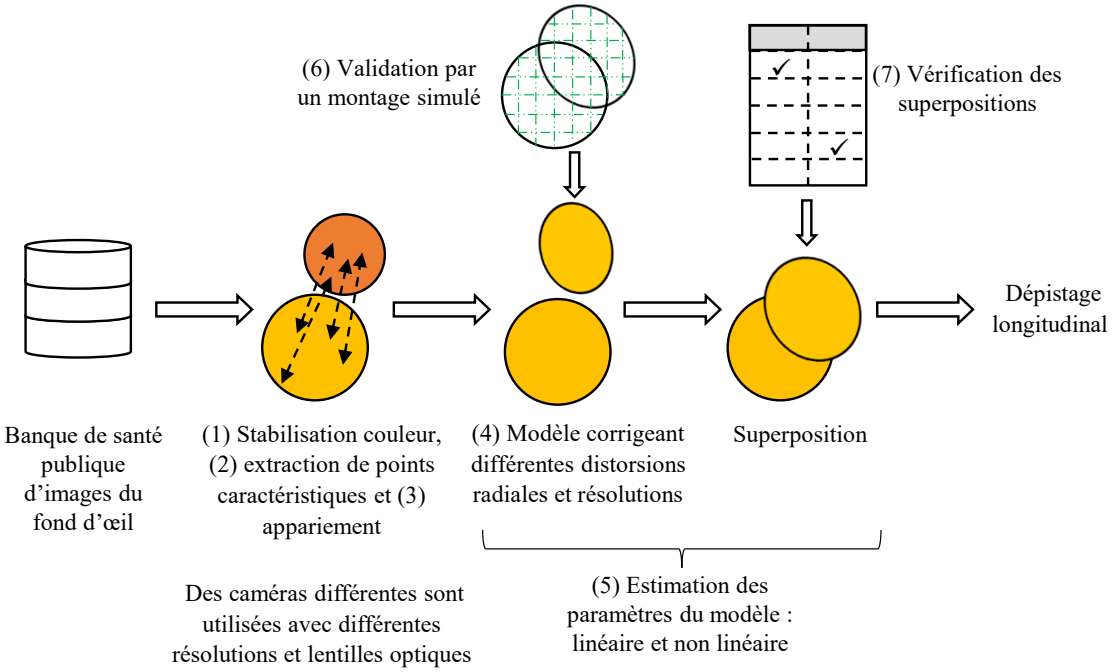
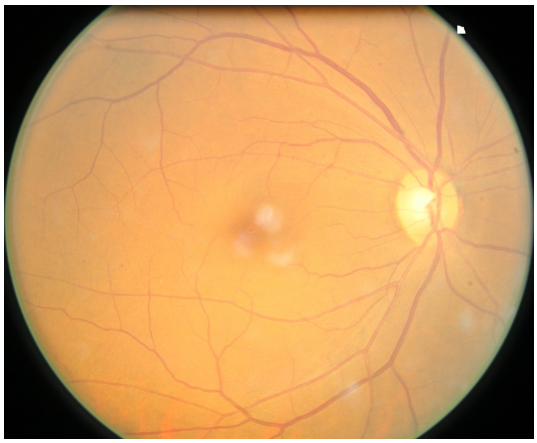


FIG. 6.3 – Cadre de travail de la superposition d'images du fond d'œil pour du dépistage longitudinal. Après la stabilisation couleur (pré-traitement) et l'extraction des points caractéristiques, les points sont appariés et les paramètres du modèle sont estimés. Les différences sont corrigées en ce qui concerne la distorsion radiale et de résolutions entre les images.

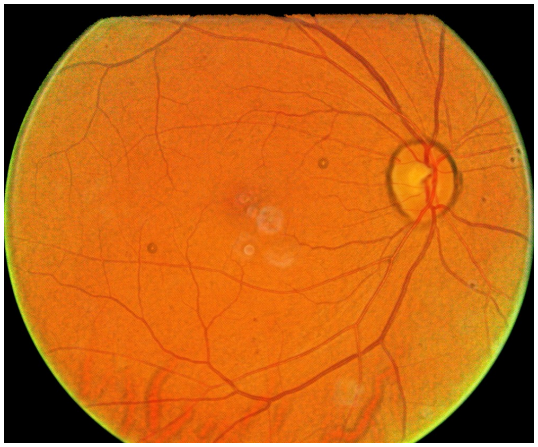
Pré-traitement La clarté (ou luminosité) des images du fond d'œil est non-uniforme pour diverses raisons : la maladie comme la cataracte, le mouvement du patient, les conditions d'acquisitions et les différences d'absorptions de la lumière dans l'œil [313]. Certaines parties des images sont claires tandis que d'autres apparaissent sombres. De plus, le changement possible de la caméra du fond d'œil entre deux examens séparés peut contribuer à un changement de couleur entre deux images du même œil (Fig. 6.4a et 6.4b). Nous avons utilisé une méthode [233] pour corriger les variations de contraste couleur entre les images. La méthode s'appuie sur un modèle couleur cohérent avec les principes physiques de formation de l'image. Les contrastes des éléments sombres et clairs sont ajustés d'une manière qui donne un aspect couleur similaire aux lésions telles que les micro-anévrismes ou aux structures anatomiques telles que des vaisseaux ou des veines. Les résultats peuvent être vus sur les figures 6.4c et 6.4d. Notre méthode couleur va plus loin que les méthodes précédentes existantes à niveaux de gris qui utilisent un ajustement polynomial [314], la morphologie mathématique [336] ou l'égalisation d'histogramme [340].



(a) Image initiale, année t.

(b) Image initiale, année $t + 1{,}5$.

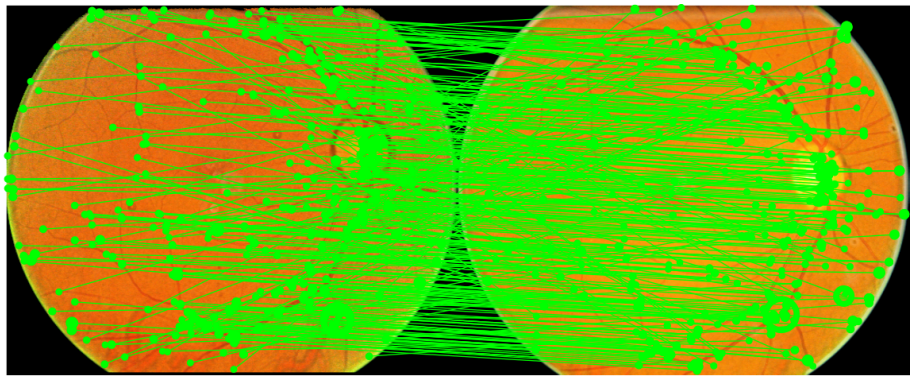
(c) Image améliorée, année t.



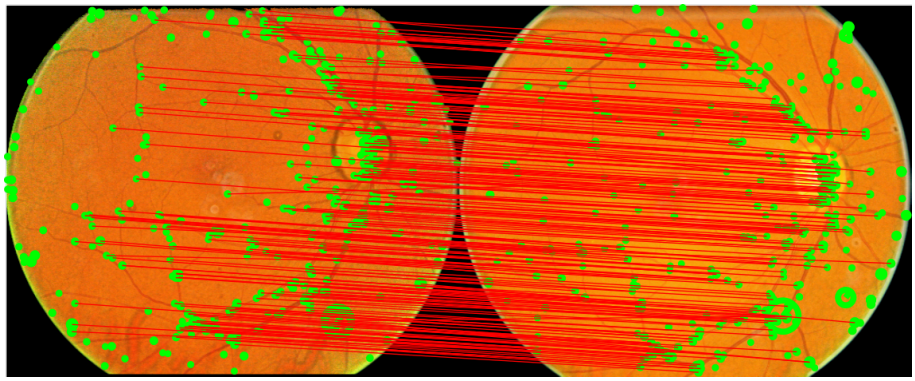
(d) Image améliorée, année $t + 1{,}5$.

FIG. 6.4 – (a), (b) Images du fond d'œil du même patient acquises avec la même caméra durant deux examens avec 1,5 années d'intervalle. Ces images de mauvaise qualité ont été acquises dans des conditions difficiles. (c), (d) Stabilisation couleur des images du fond d'œil (a) et (b).

Extraction des points caractéristiques Le pré-traitement est suivi par l'extraction de plusieurs points saillants (Fig. 6.5) en utilisant l'algorithme *scale-invariant feature transform* – SIFT – [166], [308]. Ce dernier a été conçu pour être robuste aux variations de l'angle d'observation et de l'éclairement. Brièvement, il consiste à extraire les points clefs grâce à une analyse multi-échelles. Puis, une série de descripteurs est calculée pour chaque point saillant. Ces descripteurs sont utilisés pour la mise en correspondance des points.



(a) Mise en correspondance initiale.



(b) Mise en correspondance après simplification.

FIG. 6.5 – Extraction de points et mise en correspondance entre les images. (a) Mise en correspondance initiale (flèches vertes). (b) Mise en correspondance après simplification (flèches rouges).

Mise en correspondance des points avec une estimation préliminaire de la déformation de l'image La mise en correspondance des points est une étape nécessaire avant l'estimation du modèle de déformation. Avec la méthode SIFT qui extrait les points caractéristiques, Lowe [166] a proposé une méthode d'appariement. Cependant, elle s'est avérée n'être pas assez robuste pour estimer le modèle de déformation dans la banque de données. Donc, une procédure en trois étapes a été proposée :

(a) Une première mise en correspondance par la méthode de Lowe suivie par une sélection affinée des vecteurs de correspondance selon leur taille et leur orientation (cf. ci-après).

(b) Une estimation préliminaire de la déformation de l'image par une homographie. La position des points caractéristiques est alors transformée par cette homographie (cf. ci-après).

(c) L'étape (a) est exécutée une deuxième fois en utilisant les points transformés.

Sélection des vecteurs de correspondance Tout d'abord, les régions d'intérêt des deux images sont extraites et jointes ensemble après changement de taille et rajout de zéros (*zero padding*) pour avoir la même hauteur (Fig. 6.5). Comme la déformation de l'image est proche de celle d'une similarité (i.e. rotation, translation et mise à l'échelle isotrope), les rapports de longueurs et les angles sont invariants par similarité [107]. Cette propriété est utilisée pour enlever les vecteurs de correspondance entre les points appariés de manière incohérente. Avec les deux images jointes ensemble, une sélection en deux étapes est effectuée sur les longueurs l et les orientations θ des vecteurs de correspondance.

(i) Seuls sont conservés les vecteurs v dont la longueur l_v et l'orientation θ_v sont dans l'intervalle $\{|l_v - E\{l_v\}| \leq \sigma\{l_v\}$ et $|\theta_v - E\{\theta_v\}| \leq 5°\}$. $E\{.\}$ représente l'espérance et $\sigma\{.\}$

l'écart-type d'une variable aléatoire.

(ii) Parmi les vecteurs sélectionnés $\tilde{v}$, seuls sont conservés les vecteurs dont la longueur $l_{\tilde{v}}$ et l'orientation $\theta_{\tilde{v}}$ sont dans l'intervalle $\{|l_{\tilde{v}} - E\{l_{\tilde{v}}\}| \leq \max(3\sigma\{l_{\tilde{v}}\}, 5\ \% \cdot Height)$ et $|\theta_{\tilde{v}} - E\{\theta_{\tilde{v}}\}| \leq \max(5°, \sigma\{\theta_{\tilde{v}}\})\}$. $Height$ représente le nombre de lignes dans l'image (i.e. la hauteur).

Remarque. La valeur de 5° qui correspond à une différence par rapport à une moyenne, a été choisie empiriquement et fonctionne parfaitement pour toutes les images traitées.

Estimation d'une homographie L'homographie affine (ou transformation affine) H est définie comme :

$$H = \begin{bmatrix} A & T \\ 0^t & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{6.1}$$

où $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}$, est une matrice non singulière (i.e. avec un déterminant non nul) qui représente les applications linéaires. $T = [t_x, t_y]^t \in \mathbb{R}^2$ est une translation.

Les paramètres d'une homographie sont estimés de manière robuste par maximisation de la vraisemblance avec l'algorithme de *random sample consensus* (RANSAC) qui est appliqué sur l'algorithme de *linear normalised DLT* (*direct linear transform*) [105].

Cela correspond à l'étape d'initialisation du *gold standard algorithm* de Hartley et Zisserman [105]. Cet algorithme d'initialisation estime environ 100 homographies sur quatre paires de points sélectionnées aléatoirement. Puis, l'homographie avec l'erreur minimale, lorsque les points appariés sont transformés, est sélectionnée. Cela permet d'enlever les appariements incorrects. Cependant, comme il existe aussi une distorsion radiale dans l'image, la déformation n'est pas entièrement modélisée par une homographie. Donc, certaines homographies doivent être écartées. En particulier, celles présentant une différence relative entre les facteurs d'échelles sur les axes x et y supérieure à 1 %, car la résolution du capteur est quasiment la même en x et en y. Dans ce but, plusieurs estimations (jusqu'à 50) avec l'algorithme d'initialisation sont effectuées si la différence relative entre les facteurs d'échelles est plus grande que 1 %. Si la valeur de 1 % n'est jamais atteinte, alors l'homographie avec la plus petite différence relative entre les facteurs d'échelles est conservée. Finalement, une estimation non-linéaire est effectuée avec l'algorithme de Levenberg-Marquardt [35], [193]. Il sert de bonne initialisation à l'estimation complète du modèle de déformation.

Modèle de déformation Le modèle de déformation assure une superposition correcte entre les images. Plusieurs déformations sont prises en compte :

(i) la différence de positions de l'œil entre une paire d'images sera corrigée par une transformation affine (i.e. une homographie) et

(ii) les déformations radiales dues à la projection de l'œil sur la caméra et dues à l'optique de la caméra [104] seront corrigées en utilisant une transformation radiale. Dans ce but, Lee et al. [159], [160] ont proposé un modèle couplant une unique transformation radiale pour les deux images et une homographie. Dans [161], ils ont effectué la comparaison avec deux autres modèles du second ordre.

Dans [234], j'ai étendu leur approche en définissant un modèle avec une homographie H et deux distorsions radiales, i.e. une pour chaque image. En effet, la caméra peut avoir changé entre deux examens de dépistage sur un nombre important de patients.

La distorsion radiale due à la surface sphérique de l'œil et à la distorsion radiale de la caméra a été modélisée grâce à un *modèle par division* [79], comme il s'ensuit :

$$\overline{P}^d = \left(1 + k(r^d)^2\right) \overline{P}^u, \tag{6.2}$$

où :

- $P^d \in \mathbb{R}^2$ sont les coordonnées distordues dans l'image originale (i.e. distordue) ;
- $P^u \in \mathbb{R}^2$ sont les coordonnées non-distordues dans l'image non-distordue ;
- $\overline{P}^d \in \mathbb{R}^2$ sont les coordonnées distordues centrées sur le centre C de l'image : $\overline{P}^d = P^d - C$;
- $\overline{P}^u \in \mathbb{R}^2$ sont les coordonnées non-distordues centrées sur le centre C de l'image : $\overline{P}^u = P^u - C$;
- $r^d = \left\| P^d - C \right\|_2 = \left\| \overline{P}^d \right\|_2 \in \mathbb{R}$, est la distance des coordonnées déformées P^d depuis le centre optique C (i.e. le centre de l'image) ;
- $k \in \mathbb{R}$ est le paramètre de distorsion. Sa version normalisée, $\widetilde{k} = k(1 + \|C^2\|_2)$, est dans l'intervalle $[-0{,}2; 0{,}2]$.

Ce modèle a été nommé *modèle par division* car les coordonnées distordues sont divisées par la distorsion radiale $\overline{P}^d/(1 + k(r^d)^2) = \overline{P}^u$. L'image distordue correspond à l'image originale et l'image non-distordue est l'image après correction de la distorsion radiale.

Soient P_1^d et P_2^d, les coordonnées des points dans les images originales (i.e. déformés) n°s 1 et 2. Soient k_1, k_2, les paramètres de distorsion et C_1, C_2, les centres des images. L'application transformant les points de l'image n° 1 en ceux de l'image n° 2 est définie (en coordonnées homogènes) comme il s'ensuit :

$$\begin{bmatrix} \frac{\overline{P}_2^d}{1+k_2(r_2^d)^2} + C_2 \\ 1 \end{bmatrix} = H \begin{bmatrix} \frac{\overline{P}_1^d}{1+k_1(r_1^d)^2} + C_1 \\ 1 \end{bmatrix} \text{ ou} \tag{6.3}$$

$$u_{k_2}(P_2^u) = H\left[u_{k_1}(P_1^u)\right]. \tag{6.4}$$

La distorsion radiale u_k est définie par $u_k(P^u) = [P^u, 1]^t = \left[\overline{P}^u + C, 1\right]^t \in \mathbb{R}^3$. Si la même caméra est utilisée pour faire l'acquisition des deux images, les paramètres de distorsion sont égaux, $k_1 = k_2 = k$, et le modèle correspond à celui de Lee et al. [159]. Cependant, dans notre approche, le modèle est estimé après avoir extrait et mis en correspondance les points dans la paire d'images originales (cible et référence). Donc, la correction de la distorsion radiale s'effectue après la détection des caractéristiques des points de correspondance dans l'image originale (i.e. distordue). Elle a été programmée de la manière résultante.

Après l'appariement des points, les paramètres du modèle sont estimés en minimisant l'erreur de déformation err entre les points non-distordus dans l'image qui ont été estimés par le modèle (cf. éq. (6.4)), $[\widehat{\overline{P}_2^u}, 1]^t = u_{k_2}(\widehat{P_2^u}) - [C_2, 1]^t$, et les points non-distordus extraits de l'image n° 2, $\overline{P}_2^u$. L'erreur de déformation est égale à $err = \left\| \widehat{\overline{P}_2^u} - \overline{P}_2^u \right\|_2$.

Estimation des paramètres du modèle Les paramètres du modèle sont estimés dans la figure 6.6 par une méthode différente de celle de Lee et al. [159] et d'Abramoff et al. [1]. En particulier, lors de l'étape d'initialisation, le type de modèle est sélectionné : (a) soit un modèle simple de distorsion, si les images sont acquises par la même caméra ; (b) soit un modèle à deux distorsions, si deux caméras sont utilisées. De plus, la distorsion radiale est estimée après une homographie, sans avoir besoin d'initialisation préliminaire par une calibration de la caméra [106]. En outre, nous avons introduit des estimateurs linéaires à chaque étape du processus.

Le processus d'estimation est le suivant (Fig. 6.6). L'homographie et les distorsions radiales sont initialisées en utilisant des estimateurs linéaires. Puis, une estimation itérative est effectuée indépendamment pour l'homographie et les distorsions radiales jusqu'à ce que l'erreur de déformation ait convergé. Afin d'assurer la convergence, à chaque étape, un estimateur linéaire est utilisé pour l'initialisation d'un optimiseur non-linéaire (itératif) comme l'algorithme de Levenberg-

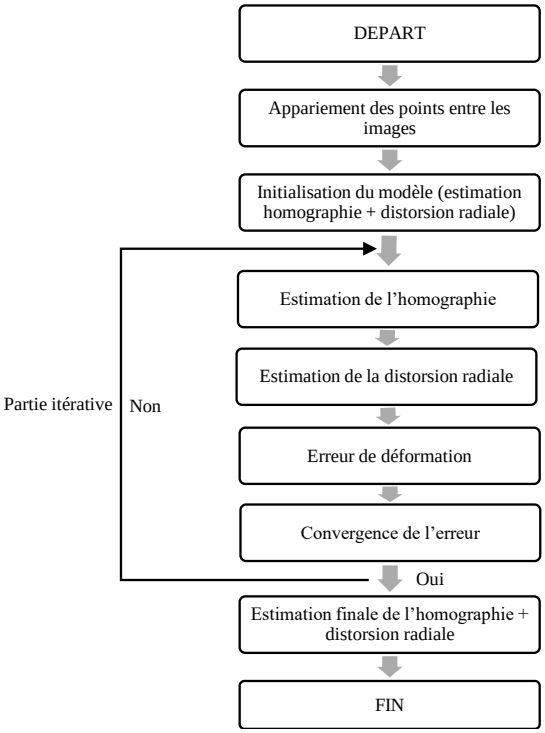


FIG. 6.6 – Schéma de la méthode de superposition à partir d'une homographie et d'un processus de distorsion radiale.

Marquardt [35], [193]. En effet, sans initialisation, l'estimateur non-linéaire peut converger jusqu'à un minimum local de l'erreur [105, p. 110]. Le critère de convergence est défini comme :

$$convergence = [err < \varepsilon] \text{ et } \left[\frac{err_n - err_{n-1}}{err_{n-1}} < tol\right] \text{ et } \left[\tilde{k} \text{ ou } \{\tilde{k}_1 \text{ et } \tilde{k}_2\} \in [-0{,}2; 0{,}2]\right] \text{ et } [n < MaxIter], \quad (6.5)$$

où err est l'erreur moyenne à l'itération n. $\varepsilon = 0{,}01$ est la tolérance sur l'erreur. $tol = 0{,}01$ est une tolérance sur l'erreur relative entre les itérations. $MaxIter = 100$ est le nombre maximal d'itérations. Pour chaque itération, les paramètres de distorsions radiaux normalisés $\tilde{k}$ doivent être dans l'intervalle $[-0{,}2; 0{,}2]$. Sinon, l'algorithme s'arrête et l'estimation du modèle avec la plus petite erreur est sélectionnée.

L'optimisation finale s'effectue conjointement sur l'homographie et la distorsion radiale avec un estimateur linéaire suivi par un optimiseur non-linéaire comme le Levenberg-Marquardt pour une seule distorsion ou la *trust region method* [35], [192] avec les bornes $[-0{,}2; 0{,}2]$ pour les deux paramètres de distorsions radiaux.

Remarque. Même si des équations linéaires sont introduites pour estimer les paramètres du modèle, un estimateur non-linéaire final est utilisé pour deux raisons. Premièrement, une estimation du modèle est effectuée en utilisant la paramétrisation (non-linéaire) de la matrice A en un ensemble de deux angles et deux facteurs d'échelles [107, p. 40], qui ont des interprétations physiques directes. Deuxièmement, la position des points de correspondance peut être réestimée conjointement avec les paramètres du modèle en utilisant l'erreur de « *Sampson* » ou l'erreur de l'algorithme « *gold standard* » [105, p. 114]. Dans cette section, la position des points de correspondance n'a pas été réestimée. Cela fera l'objet de futurs travaux.

Estimation d'un modèle avec deux distorsions radiales Lorsque qu'une caméra différente est utilisée pour chacune des images d'une paire, deux distorsions radiales sont estimées,

k_1 et k_2. Définissons des estimateurs linéaires pour initialiser les estimateurs non-linéaires des paramètres du modèle.

Estimateurs linéaires des paramètres de distorsions radiales k_1 et k_2 D'après les équations (6.1) et (6.3), on déduit que :

$$k_1k_2\left[\left(r_1^d r_2^d\right)^2 \cdot D\right] + k_1\left[r_1^{d^2} \cdot D + r_1^{d^2} \cdot \overline{P}_2^d\right] + k_2\left[r_2^{d^2} \cdot D - r_2^{d^2} \cdot A\overline{P}_1^d\right] = \\ -\left[\overline{P}_2^d + D - A\overline{P}_1^d\right], \quad (6.6)$$

où $D = C_2 - AC_1 - T$. L'équation (6.6) est quadratique par rapport à ses inconnues k_1 et k_2 : $k_1k_2\left[M_{12}\right] + k_1\left[M_1\right] + k_2\left[M_2\right] + \left[M_0\right] = 0$. Les variables M_i, avec $i = \{0; 1; 2; 12\}$, ont deux composantes $M_i = [M_i^x, M_i^y]^t \in \mathbb{R}^2$. Pour chaque composante, l'équation est vérifiée pour chacune des N paires de points appariés. Ceci conduit à deux systèmes d'équations quadratiques. L'ensemble des équations correspond à l'intersection de deux coniques, qui a une unique paire de solutions, selon le théorème de Bernstein pour les polynôme bivariés [295, p. 37]. Ce résultat peut être expliqué d'une manière plus intuitive ; entre une paire de points appariés, il existe une paire de solutions par équations (i.e. deux paires). Comme la solution doit être valide pour les deux systèmes d'équations, il y a une unique paire de solutions $\{k_1, k_2\}$ valide pour tous les points appariés. Les autres solutions sont seulement valides pour une seule paire de points correspondants. La paire $\{k_1, k_2\}$ qui est solution unique, satisfait conjointement les deux systèmes d'équations et peut être obtenue par élimination du terme de second ordre :

$$\left.\begin{array}{l} k_1k_2 + \frac{M_1^x}{M_{12}^x}k_1 + \frac{M_2^x}{M_{12}^x}k_2 + \frac{M_0^x}{M_{12}^x} = 0 \\ k_1k_2 + \frac{M_1^y}{M_{12}^y}k_1 + \frac{M_2^y}{M_{12}^y}k_2 + \frac{M_0^y}{M_{12}^y} = 0 \end{array}\right\} \Rightarrow k_1\left[\frac{M_1^x}{M_{12}^x} - \frac{M_1^y}{M_{12}^y}\right] + k_2\left[\frac{M_2^x}{M_{12}^x} - \frac{M_2^y}{M_{12}^y}\right] = \\ -\left[\frac{M_0^x}{M_{12}^x} - \frac{M_0^y}{M_{12}^y}\right]. \quad (6.7)$$

La paire $\{k_1, k_2\}$, solution unique de l'équation (6.7), est déterminée par la méthode des moindres carrés. Lorsque la caméra est la même pour les deux images, seule une distorsion radiale, $k = k_1 = k_2$, a besoin d'être estimée dans l'équation (6.7).

Estimateurs linéaires des paramètres de distorsions radiales k_1 et k_2 et de l'homographie D'après les équations (6.1) et (6.3), on déduit que :

$$k_1k_2D\left[\left(r_1^d r_2^d\right)^2\right] + k_1D\left[r_1^{d^2}\right] + k_1\left[r_1^{d^2}.\overline{P}_2^d\right] + k_2D\left[r_2^{d^2}\right] - k_2A\left[r_2^{d^2} \cdot \overline{P}_1^d\right] - \\ A\left[\overline{P}_1^d\right] + D\left[1\right] = -\left[\overline{P}_2^d\right]. \quad (6.8)$$

L'équation (6.8) est linéaire par rapport à ses variables entre crochets. Elle est vérifiée pour chaque composante (x et y) des vecteurs $\overline{P}_1^d = \left[\overline{P}_1^{d,x}, \overline{P}_1^{d,y}\right]^t$, $\overline{P}_2^d = \left[\overline{P}_2^{d,x}, \overline{P}_2^{d,y}\right]^t$ et $D = [D^x, D^y]^t$, conduisant à un système de deux équations :

$$\left\{\begin{array}{c} k_1k_2D^x\left[\left(r_1^d r_2^d\right)^2\right] + k_1D^x\left[r_1^{d^2}\right] + k_1\left[r_1^{d^2}\overline{P}_2^{d,x}\right] + k_2D^x\left[r_2^{d^2}\right] - \\ k_2a_{11}\left[r_2^{d^2}\overline{P}_1^{d,x}\right] - k_2a_{12}\left[r_2^{d^2}\overline{P}_1^{d,y}\right] - a_{11}\left[\overline{P}_1^{d,x}\right] - a_{12}\left[\overline{P}_1^{d,y}\right] + D^x\left[1\right] = -\left[\overline{P}_2^{d,x}\right] \\ k_1k_2D^y\left[\left(r_1^d r_2^d\right)^2\right] + k_1D^y\left[r_1^{d^2}\right] + k_1\left[r_1^{d^2}\overline{P}_2^{d,y}\right] + k_2D^y\left[r_2^{d^2}\right] - \\ k_2a_{21}\left[r_2^{d^2}\overline{P}_1^{d,x}\right] - k_2a_{22}\left[r_2^{d^2}\overline{P}_1^{d,y}\right] - a_{21}\left[\overline{P}_1^{d,x}\right] - a_{22}\left[\overline{P}_1^{d,y}\right] + D^y\left[1\right] = -\left[\overline{P}_2^{d,y}\right] \end{array}\right. . \quad (6.9)$$

Les 20 variables – entre crochets – correspondent aux données. Les 18 paramètres du système d'équations sont estimés en utilisant la méthode des moindres carrés appliquée à l'ensemble des deux équations. Comme il y a de multiples relations entre les 18 estimateurs de paramètres, seuls les 8 paramètres du modèle — k_1, k_2 et H (par l'intermédiaire de la matrice $A = [a_{ij}]_{\{i,j\in\{1;2\}\}}$ et du vecteur D) — sont inconnus. Une solution admissible pour k_1, k_2 et H est déduite des 18 paramètres de l'estimateur linéaire. La valeur de k_1 est directement obtenue (3e terme de la 1re équation) ainsi que celles de D^x et D^y (9e termes des deux équations) et a_{ij} (7e et 8e termes des deux équations). La valeur de k_2 est déduite, e.g. à partir du 4e terme de la 1re équation. Puis, cette solution admissible est introduite comme initialisation de l'estimation non-linéaire.

Déformation de l'image Afin d'analyser une grande banque de données, un algorithme rapide de déformation d'image est nécessaire. La déformation directe (*forward warping*) prend du temps et nous utilisons donc la déformation inverse (*inverse warping*). Cependant, le modèle de recalage doit être inversible [319]. La distorsion radiale est modélisée dans l'équation (6.3) grâce à un modèle par division [79]. Wonpil [321] et Park et al. [240] ont calculé une transformation approchée pour une méthode distorsion usuelle. Ici, l'inversion exacte du modèle par division est déterminée analytiquement.

Etant donnée $r^u = \|P^u - C\| = \overline{P}^u \in \mathbb{R}$, la distance des coordonnées non-distordues P^u depuis le centre optique C, d'après l'équation (6.2), on obtient :

$$r^d = \left(1 + k\left(r^d\right)^2\right) r^u. \tag{6.10}$$

Des équations (6.2) et (6.10), on déduit que :

$$\overline{P}^u = \frac{r^u}{r^d}\overline{P}^d = W(\overline{P}^d). \tag{6.11}$$

Afin d'utiliser la déformation inverse, il est nécessaire de déterminer l'application W^{-1} transformant les points non-distordus $\overline{P}^u$ en des points distordus $\overline{P}^d$. D'après l'équation (6.11), il est équivalent de déterminer r^u connaissant r^d.

L'équation (6.10) est équivalente à $kr^u{r^d}^2 - r^d + r^u = 0$, une équation du second ordre en r^d. Dans la démonstration ci-après, nous montrons que son discriminant, $\Delta = 1 - 4kr^{u2}$, est strictement positif et que la transformation inverse W^{-1} correspond à la racine :

$$r^d = W^{-1}(r^u) = \frac{1 - \sqrt{1 - 4k(r^u)^2}}{2kr^u}. \tag{6.12}$$

Démonstration. Démontrons (i) la positivité du discriminant et (ii) le choix de la racine retenue. Nous avons :

$$\left.\begin{array}{r} r^u = \|P^u - C\| \leq \|C\| \\ k = \frac{|\tilde{k}|}{1+\|C\|^2} \leq \frac{2\times 10^{-1}}{1+\|C\|^2} \end{array}\right\} \Longrightarrow |k|\, r^u = \frac{\left|\tilde{k}\right|}{1+\|C\|^2}\|P^u - C\| \leq 2\times 10^{-1}\frac{\|C\|}{1+\|C\|^2} \leq \frac{2\times 10^{-1}}{\|C\|}.$$

Ceci implique que : $\frac{1}{2|k|r^u} \geq \frac{\|C\|}{4\times 10^{-1}}$ et $4\,|k|\,r^{u2} \leq 8\times 10^{-1} < 1$. Donc, nous avons : $\Delta = 1 - 4kr^{u2} > 0$, (i) Q.E.D.

Lorsque le discriminant est positif, il existe deux racines :

a) $\frac{1+\sqrt{\Delta}}{2kr^u} \geq \frac{\|C\|}{4\times 10^{-1}} = 2{,}5\,\|C\|$, qui correspond à une solution où les positions des pixels sont rejetées en dehors d'un certain rayon, $2{,}5\,\|C\|$. Cette racine correspond à une solution aberrante.

b) $\frac{1-\sqrt{\Delta}}{2kr^u} \leq \frac{C}{4\times 10^{-1}} = 2{,}5\,\|C\|$, qui correspond à une solution où les positions des pixels sont à l'intérieur d'un certain rayon, $2{,}5\,\|C\|$. C'est la solution acceptable, (ii) Q.E.D.

□

Donc, la transformation W est inversible (eq. (6.12)). Une méthode de déformation d'image inversible comparée à une méthode non-inversible réduit drastiquement le temps de calcul [319]. Nous avons expérimenté une réduction d'environ 10 min à quelques secondes (environ 5 s) sur un ordinateur standard en utilisant *MATLAB*® (processeur Intel® Core™ i7 CPU 4702HQ ; 2,20 GHz ; 4 cœurs ; 8 threads avec 16 Go de RAM).

Validation expérimentale

Validation par un montage simulé Un montage simulé a été créé par une superposition de deux images du fond d'œil et une déformation selon la méthode présentée par Lee et al. [161]. De vraies images recalées ont été sélectionnées avec un pourcentage de recouvrement de 80 % correspondant au cas le plus commun dans les banques de données longitudinales. Aucune modification de couleur n'a été faite dans les images. Après avoir ajouté des points de repères (i.e. la vérité terrain), les images ont été découpées et déformées selon le modèle de Lee et al. [161]. Une transformation affine a été utilisée, avec une rotation et une mise à l'échelle anisotrope. Puis, l'image a été modifiée par une distorsion projective. Le rayon de la boule de l'œil est égal au rapport entre le rayon du disque du champ de vision de l'image et l'angle d'observation de la caméra (45°). Puis, les images sont recalées et l'erreur est mesurée entre les points de repères recalés et leurs vraies positions (i.e. la vérité terrain).

La qualité de la superposition de l'image a été évaluée grâce au montage simulé pour trois approches :

(i) notre méthode avec une seule distorsion radiale ;

(ii) notre méthode avec deux distorsions radiales ;

(iii) une méthode de l'état de l'art « gdbicp » [325] qui utilise un modèle quadratique.

La vérité terrain du montage simulé permet une comparaison entre notre modèle et le modèle quadratique « gdbicp ». La durée de chaque méthode de superposition est aussi mesurée.

Validation avec une banque de données de santé publique Notre méthode doit être visuellement évaluée et comparée à d'autres de l'état de l'art dans des banques de données de santé publique utilisées pour le dépistage de la rétinopathie. Afin d'évaluer l'évolution de la rétinopathie diabétique dans le monde, il existe des programme de dépistage. Des essais ont été effectués sur une banque de données de 69 patients sélectionnés aléatoirement et provenant du programme *Diabetic Eye Screening Wales* (DESW), au Pays de Galle, Royaume-Uni [298]. Tous les patients ont du diabète et différents stades de sévérités, y compris aucun signe de rétinopathie ou de maculopathie. Tous les patients sélectionnés ont été dépistés annuellement pendant plusieurs années, en moyenne sept ans. Pour chacun d'eux, nous avons utilisé deux séries de deux examens avec un intervalle de dépistage d'environ un an entre les deux événements d'examens. Pour chaque événement d'examen, quatre images sont disponibles avec deux positions (nasale et maculaire) par œil. Il y avait deux séries d'images acquises durant des années différentes : la première série se composait de 271 paires d'images de qualité suffisante (dont 63 % étaient acquises avec des caméras différentes) et la deuxième série contenait 268 paires (dont 9 % étaient acquises avec des caméras différentes). Dans la première série, 11 paires ont une faible aire de recouvrement (environ 30 % de la surface de la paire recalée). Pour chaque position, nous avons effectué la superposition des images entre deux examens différents. Toutes les photographies rétiniennes avaient une haute qualité selon un protocole incluant une dilatation de la pupille.

En outre, afin de montrer que notre méthode n'est pas seulement utile sur des images de haute qualité acquises après dilatation des pupilles, mais aussi sur des images de faible qualité sans dilatation des pupilles avant la photographie, un deuxième test de validation est effectué pour superposer les images de faible/mauvaise qualité pour cinq patients avec du diabète. La qualité de la superposition d'images est vérifiée pour les 20 paires d'images (i.e. quatre par patient). Les

conditions d'acquisition étaient significativement plus difficiles comparativement aux images de haute qualité et la qualité des images était hétérogène, en partie à cause de la dilatation de la pupille avant la photographie. Dans cette banque de données, les patients avec du diabète ont deux examens rétiniens avec au moins une année d'intervalle et en utilisant la même caméra.

Comme il n'y a pas de recalage de référence pour les banques de données (i.e. pas de vérité terrain), pour toutes les paires d'images la superposition a été vérifiée par le même expert. L'aire entière de recouvrement de chaque paire a été vérifiée avec soin. La classification a été faite en deux catégories :

(1) pas de différence visible (i.e. correcte) et

(2) différence visible (i.e. incorrecte). La catégorie « incorrecte » inclut trois sous-catégories : (i) différences de la taille d'un vaisseau de petit diamètre ; (ii) différences de la taille d'un vaisseau de gros diamètre ou (iii) même plus large. Dans cette section, les trois sous-catégories ont été regroupées en une seule catégorie « incorrecte ».

Pour chaque ensemble d'environ 270 images, la durée d'une vérification pour un modèle est plus longue qu'une heure et demie. Pour chaque méthode et chaque paire de chaque ensemble, une première vérification a été effectuée et une seconde vérification a été faite par la même personne. Pour presque toutes les paires, les deux résultats étaient en accord entre les deux vérifications. Dans les rares cas de divergence, la décision a toujours été prise de la manière la plus restrictive (i.e. « incorrecte »). Par conséquent, la durée totale du processus a été de plus de 30 heures pour les cinq méthodes testées ci-après.

Une évaluation visuelle a été effectuée sur les deux séries d'images de haute qualité et sur la série d'images de faible qualité afin de comparer notre approche à deux autres de l'état de l'art qui sont bien établies et dont le code est disponible publiquement : « gdbicp » [291], [325] ou que nous avons reprogrammées (Lee et al. [159]). En particulier, nous avons testé les cinq méthodes suivantes (cf. table 6.3) :

(i) notre méthode sur les images normalisées avec un choix automatique d'une ou de deux distorsions selon le nombre de caméras utilisées ;

(ii) notre modèle avec une seule distorsion pour toutes les paires d'images originales sans normalisation couleur, afin d'être similaire à la méthode de l'état de l'art de Lee et al. [159] ;

(iii) la méthode de l'état de l'art « gdbicp » avec un modèle quadratique [291] sur les images originales ;

(iv) « gdbicp » avec un modèle radial sur les images originales [325] ;

(v) et « gdbicp » avec un modèle quadratique [291] sur les images normalisées.

Remarque. Pour faciliter la comparaison entre (i) notre modèle et (ii) celui similaire de Lee et al. [159], nous avons effectué l'ajustement des deux modèles avec notre méthode, en utilisant les points caractéristiques extraits par le SIFT, tandis que Lee et al [159] utilisent la ligne centrale des vaisseaux et les centres des embranchements pour leur modèle.

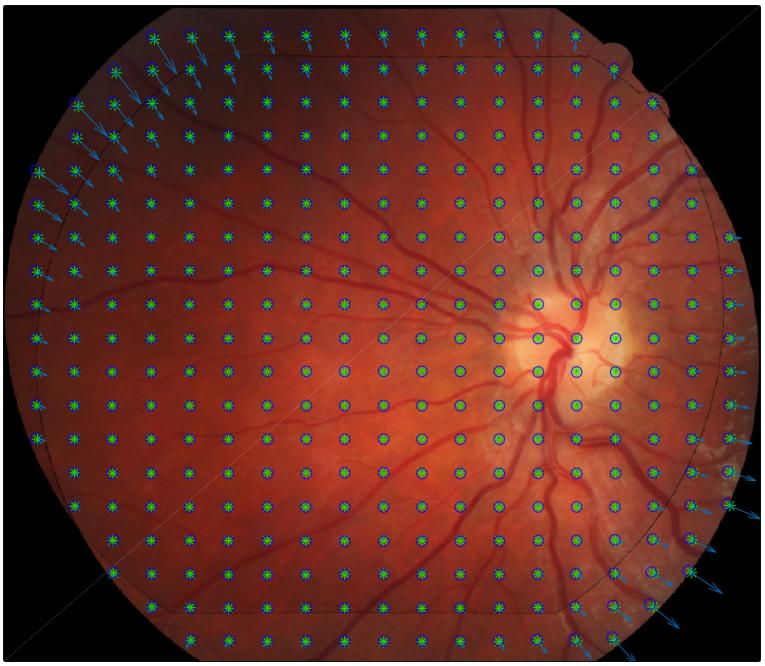


FIG. 6.7 – Validation du modèle de superposition en recalant une paire d'images précédemment déformées. Les étoiles vertes correspondent aux points de l'image de référence et les points bleus à ceux de l'image courante. Les flèches montrent l'erreur de recalage entre les deux images.

Résultats

Validation par un montage simulé En utilisant notre méthode, avec une seule distorsion, l'erreur moyenne de recalage est de 0,86 pixel (écart-type de 1,75 pixels) dans des images de taille $1\,568 \times 2\,352$ pixels (Fig. 6.7), avec des vaisseaux d'un diamètre maximum de 30 pixels. L'erreur relative par rapport à la diagonale de l'image est de 0,03 %, et par rapport aux vaisseaux de 2,9 %. Avec deux distorsions, l'erreur moyenne de recalage est de 0,92 pixels (écart-type de 1,94 pixels) et l'erreur relative est de 0,03 % par rapport à l'image et 3,1 % par rapport aux vaisseaux (voir table 6.1). Sur la figure 6.7, l'erreur est principalement située sur la partie externe de la paire superposée d'images.

Notre méthode est comparée à une autre méthode de l'état de l'art « gdbicp » qui superpose les images du fond d'œil avec un modèle quadratique [291]. En utilisant le montage simulé, l'erreur moyenne de recalage est de 2,44 pixels (écart-type de 1,64 pixels). Pour rappel, le modèle de déformation proposé dans cette section avec une seule distorsion radiale correspond au modèle de Lee et al. [159], cependant, avec une méthode différente pour estimer les paramètres.

Comme test de vitesse, la durée des différentes méthodes de superposition est mesurée avec les images du montage simulé (voir table 6.2). En 2016, les expériences ont été effectuées sur un ordinateur récent (processeur Intel® Core™ i7 CPU 4702HQ ; 2,20 GHz ; 4 cœurs ; 8 threads avec 16 Go de RAM). La méthode inverse de déformation avec une ou deux distorsions dure approximativement 5,5 s en utilisant *MATLAB*®.

TABLE 6.1 – Erreur de superposition pour différentes méthodes de recalage avec le montage simulé.

Modèle	Moyenne	Ecart-type	Erreur relative moyenne par rapport à l'image	Erreur relative moyenne par rapport aux vaisseaux
• **Homographie et une distorsion radiale** (images normalisées) (**notre méthode**)	**0,86 pixels**	**1,75 pixels**	**0,03 %**	**2,9 %**
• **Homographie et deux distorsions radiales** (images normalisées) (**notre méthode**)	**0,92 pixels**	**1,94 pixels**	**0,03 %**	**3,1 %**
• « gdbicp » quadratique (images originales)	2,44 pixels	1,64 pixels	0,08 %	8,1 %

TABLE 6.2 – Durée de chaque méthode de superposition.

Modèle	Durée
Homographie et une distorsion radiale (images normalisées) (**notre méthode**)	**64,5 s**
Homographie et deux distorsions radiales (images normalisées) (**notre méthode**)	71,7 s
« gdbicp » quadratique (images originales)	169 s

Validation sur une banque de données de santé publique et comparaison à d'autres méthodes de l'état de l'art Une évaluation visuelle est essentielle sur les banques de données de santé publique pour les cinq méthodes précédemment décrites. Des résultats préliminaires ont été présentés dans [232].

TABLE 6.3 – Pourcentages de superpositions correctes de paires d'image du fond d'œil dans des banques de données de santé publique pour différentes méthodes et modèles.

Modèle	1re série – 271 paires (63 % caméras différentes) (images haute qualité)	2e série – 268 paires (9 % caméras différentes (images haute qualité)	3e série – 20 paires caméras similaires (images faible qualité)
(i) **Homographie et une ou deux distorsions radiales** (images normalisées) (**notre méthode**)	**92 %**	**98 %**	**100 %**
(ii) Homographie et une distorsion radiale (images originales) (similaire à Lee et al. [159])	88 %	95 %	95 %
(iii) « gdbicp » quadratique (images originales) (Yang et al. [325])	74 %	58 %	40 %
(iv) « gdbicp » radial (images originales)	48 %	33 %	25 %
(v) « gdbicp » quadratique (images normalisées)	88 %	87 %	100 %

(i) Tout d'abord, notre méthode de superposition est appliquée sur des images normalisées (pour les contrastes couleur). Si la même caméra a capturé les deux images d'une paire, notre modèle est utilisé avec une seule distorsion radiale. Si deux caméras ont été utilisées, le modèle avec deux distorsions radiales est appliqué. Sur les figures 6.1 et 6.8, quatre exemples de superpositions avec le modèle radial de distorsion sont montrés. On peut remarquer la bonne qualité de la superposition. Dans les figures 6.8a et 6.8b, deux distorsions radiales ont été utilisées pour corriger la déformation des deux caméras différentes, tandis que dans la figure 6.8c une seule distorsion corrige la déformation d'une seule caméra. Les résultats de validation pour la banque de données complète sont présentés dans la table 6.3. Dans la

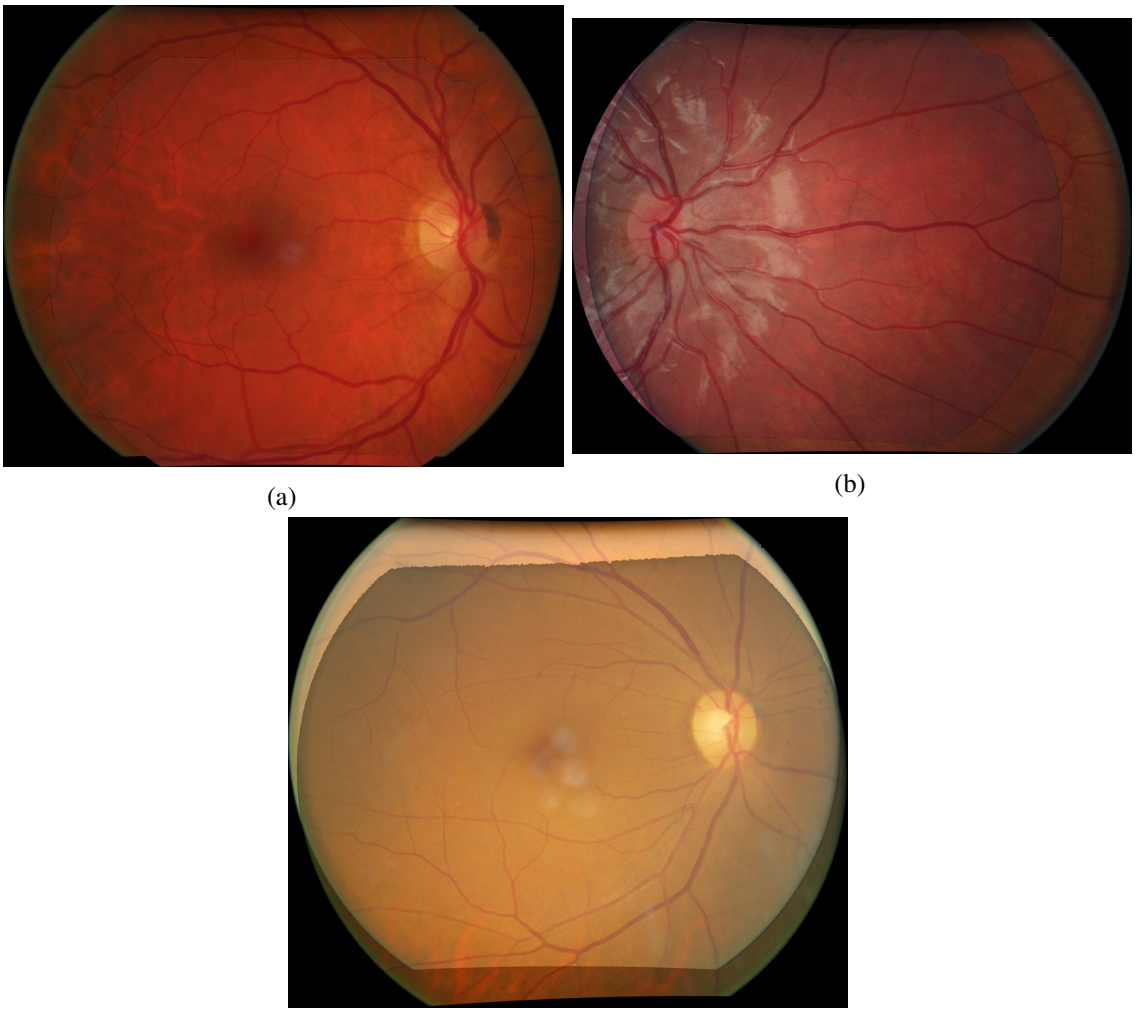


FIG. 6.8 – Superposition de trois paires d'images du fond d'œil avec correction d'une ou deux distorsions radiales. (a) Superposition de la paire des figures 6.2a et 6.2b, avec deux distorsions radiales $\tilde{k}_1 = -0{,}054\,6$; $\tilde{k}_2 = -0{,}064\,6$. (b) Superposition de la paire des figures 6.2c et 6.2d, avec deux distorsions radiales $\tilde{k}_1 = -0{,}069\,5$; $\tilde{k}_2 = -0{,}053\,6$. (c) Superposition de la paire des figures 6.2a et 6.2b, avec une seule distorsion radiale $\tilde{k} = -0{,}081\,9$.

première série de 271 paires, deux paires ont de petites différences dans la partie externe de la superposition. Ces différences sont de la taille du diamètre d'un vaisseau (Fig. 6.10). Lorsque le pourcentage de surface de superposition est faible (autour de 30 %) comparé à la surface de l'image superposée, les superpositions sont correctes dans 92 % des paires et 96 %, si l'on considère les paires dans la même position. Dans la deuxième série de 268 paires d'images acquises quelques années après, la superposition est correcte pour 98 % des paires. Parmi les paires incorrectement superposées, la partie externe présentait des différences tandis que la partie centrale était parfaitement superposée. Pour les deux séries et pour les paires avec un recouvrement suffisant (supérieur à 50 %), aucune différence visible n'a été perçue entre elles.

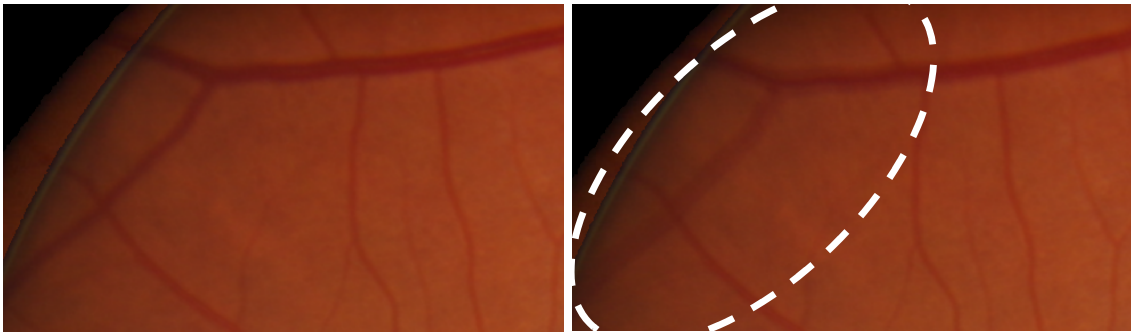
(ii) Notre approche est comparée à celle similaire de Lee et al. [159] (table 6.3). Dans la première séries avec 271 paires (respectivement la seconde série avec 268 paires), la méthode similaire à celle de Lee et al. [159] trouve une solution de superposition pour 88 % des paires (respectivement 95 %), tandis que la nôtre donne un résultat correct pour 92 % des paires

(respectivement 98 %). Dans la figure 6.9, avec leur approche, une différence visible apparaît près d'un vaisseau situé sur la partie externe de l'image, où l'effet de la déformation radiale est plus important, alors qu'il n'y a pas de différence notable avec notre méthode.

(iii) Notre méthode est comparée à la méthode de l'état de l'art « gdbicp » [291], [325], qui superpose des images avec un modèle quadratique (table 6.3). Dans la première série avec 271 paires (respectivement la seconde série avec 268 paires), leur méthode a trouvé une solution de superposition pour 74 % des paires (resp. 58 %), tandis que notre méthode donne un résultat correct pour 92 % des paires (resp. 98 %).

(iv) En outre, lorsque l'on utilise la méthode « gdbicp » avec une homographie et deux distorsions radiales par un modèle multiplicatif – qui n'est donc pas inversible – peu de paires sont recalées (48 % dans la première série et 33 % dans la deuxième série). La méthode « gdbicp » avec un modèle quadratique donne de meilleurs résultats pour des images du fond d'œil que « gdbicp » avec des distorsions radiales, parce que « gdbicp » avec un modèle quadratique a été spécifiquement conçue pour superposer les images du fond d'œil dans [291].

Afin d'illustrer la robustesse de notre méthode, la superposition est effectuée sur la banque de données de cinq patients avec 20 paires d'images de faible qualité. Dans chacun des cas, il n'y avait pas de différence notable dans la superposition (i) avec notre méthode, même sur le bord des images. (ii) La méthode similaire à celle de Lee et al. [159] donne une superposition correcte pour 95 % des paires, (iii) « gdbicp » avec un modèle quadratique 40 % et (iv) « gdbicp » avec une homographie et deux distorsions radiales, 25 %. Dans cette seconde validation, notre méthode donne de meilleurs résultats que les autres.



(a) (b)

FIG. 6.9 – (a) Superposition correcte avec notre méthode fondée sur une homographie et deux distorsions radiales sur des images normalisées ($k_1 = -0{,}061\,6$ et $k_2 = -0{,}068\,5$). (b) Superposition incorrecte avec des différences visibles (vaisseaux flous) en utilisant la méthode de Lee et al. [159] sur les images initiales ($k = -0{,}033\,4$). Les différences peuvent être plus visibles dans la version électronique.

Interprétation et discussion

La précision des méthodes de superposition a été validée sur un montage simulé (table 6.1). L'erreur de superposition est plus faible pour notre modèle qui comprend une ou deux distorsions que pour un modèle quadratique de superposition tel que celui utilisé dans « gdbicp » quadratique [6], [50], [291], [325]. Cela confirme le résultat de Lee et al. [161] qui a précédemment montré qu'un modèle quadratique a une erreur plus forte qu'une homographie et un modèle de distorsion radial. De tels résultats démontrent que notre méthode produit une superposition sans différence visible par un observateur humain. Donc, cette approche est adaptée pour effectuer l'analyse de grandes banques de données de santé publique.

En matière de vitesse, notre méthode est 2,4 fois plus rapide que « gdbicp » quadratique (table 6.2), même en utilisant un langage informatique plus lent (*MATLAB*® face à un langage

compilé). Les résultats de durée montrent que notre méthode est efficace d'un point de vue calculatoire lorsque l'on analyse des grandes banques de données. Les bons résultats de vitesse sont dus aux estimateurs linéaires des paramètres du modèle et l'utilisation d'un modèle inversible de déformation. Dans le futur, nous étudierons une implémentation plus efficace en utilisant un langage compilé, qui est plus rapide.

Afin d'évaluer l'efficacité à des fins de santé publique, notre méthode de superposition a été validée sur des banques de données de santé publique avec des images de haute qualité pour 69 patients (deux séries comprenant 271 paires et 268 paires différentes) et cinq patients (20 paires) avec des images de faible qualité. Dans la première série, la majorité des paires a été acquise avec deux caméras, tandis que dans la seconde série la majorité a été acquise avec la même caméra. Quel que soit le nombre de caméras, il n'y pas de différences visibles dans les images superposées si le recouvrement est suffisant (plus de 50 % environ) (table 6.3-(i)). Pour des paires avec une faible aire de recouvrement, nous avons développé un autre algorithme qui utilise, en plus des points de correspondance, la distance entre les vaisseaux rétiniens. Cet algorithme est similaire à celui décrit par Can et al. [42] et Lee et al. (2010) [161], mais possède néanmoins deux corrections de distorsion radiales. Ces résultats seront présentés dans la sous-section 6.2.2, p. 222.

Cependant, l'intérêt de la superposition est de comparer les évolutions dans des banques de données de santé publique sur plusieurs années, ce qui est seulement utile lorsque le recouvrement est suffisamment large. Cela signifie que notre méthode, efficace lorsque le recouvrement est suffisant, est adaptée pour être appliquée à l'analyse des grandes banques de données.

(v) Finalement, afin de montrer l'influence de la normalisation (ou stabilisation) couleur sur la superposition, la méthode « gdbicp » avec un modèle quadratique est appliquée aux images avec une stabilisation couleur (i.e. normalisées). Les résultats obtenus avec la stabilisation couleur sont bien meilleurs que sans : dans la première série avec 271 paires, 88 % contre 48 %, dans la deuxième série avec 268 paires, 87 % contre 33 % et dans la troisième série avec 20 paires de faible qualité, 100 % contre 25 %.

Comparée à l'autre méthode de l'état de l'art de Lee et al. [159] composée d'une homographie et d'une seule distorsion (ii), pour toutes les trois séries d'images, notre méthode (i) améliore systématiquement les résultats obtenus (table 6.3). Ce résultat est dû à l'utilisation de deux distorsions radiales au lieu d'une, lorsque des images de différentes tailles sont acquises par différentes caméras. Cela montre l'importance d'utiliser un modèle adapté aux types de caméras employées pour acquérir une paire d'images. Donc, notre modèle va plus loin que le précédent (de Lee et al. [159]) qui ne prenait pas en compte le changement de caméra. Nous soulignons que dans [159], [161], les auteurs étaient intéressés par de grandes représentations sous formes de mosaïques avec des images acquises en utilisant la même caméra au cours du même examen du patient, et donc ils utilisaient une banque de données correspondante (i.e. des images de couleur similaire et le même paramètre de distorsion). Notre banque de données se compose d'images de patients acquises avec au moins une année d'intervalle par des caméras différentes et elles peuvent être de faible qualité.

Comparée à la méthode « gdbicp » quadratique [325] (iii), qui utilise un modèle quadratique, notre méthode (i) recale mieux les images du fond d'œil. Notre approche superpose toujours mieux les images que la méthode « gdbicp » quadratique avec des images normalisées (v). Cependant, lorsque l'aire de recouvrement de l'image est faible (i.e. inférieure à 50 %), l'algorithme « gdbicp » quadratique (v) recale mieux que notre méthode. Ceci est dû à la détection des vaisseaux comme repères, ce qui rend l'ajustement du modèle plus efficace avec un faible recouvrement. Cependant, dans cette configuration des images, la solution de superposition donnée par le « gdbicp » quadratique ((iii), (v)) présente souvent une forte déformation de cisaillement (i.e. une mise à l'échelle anisotrope) qui n'a pas de justification physique. Cela signifie que le modèle quadratique n'est pas bien adapté lorsque la surface de recouvrement est faible. Cela confirme les résultats obtenus par Lee et al. [161] qui ont montré que les modèles quadratiques introduisent plus d'erreurs que les modèles fondés sur une homographie et une distorsion radiale. En effet, avec 12 paramètres à ajuster (contre 8 paramètres pour notre modèle – 6 pour l'homographie et 2 pour les distor-

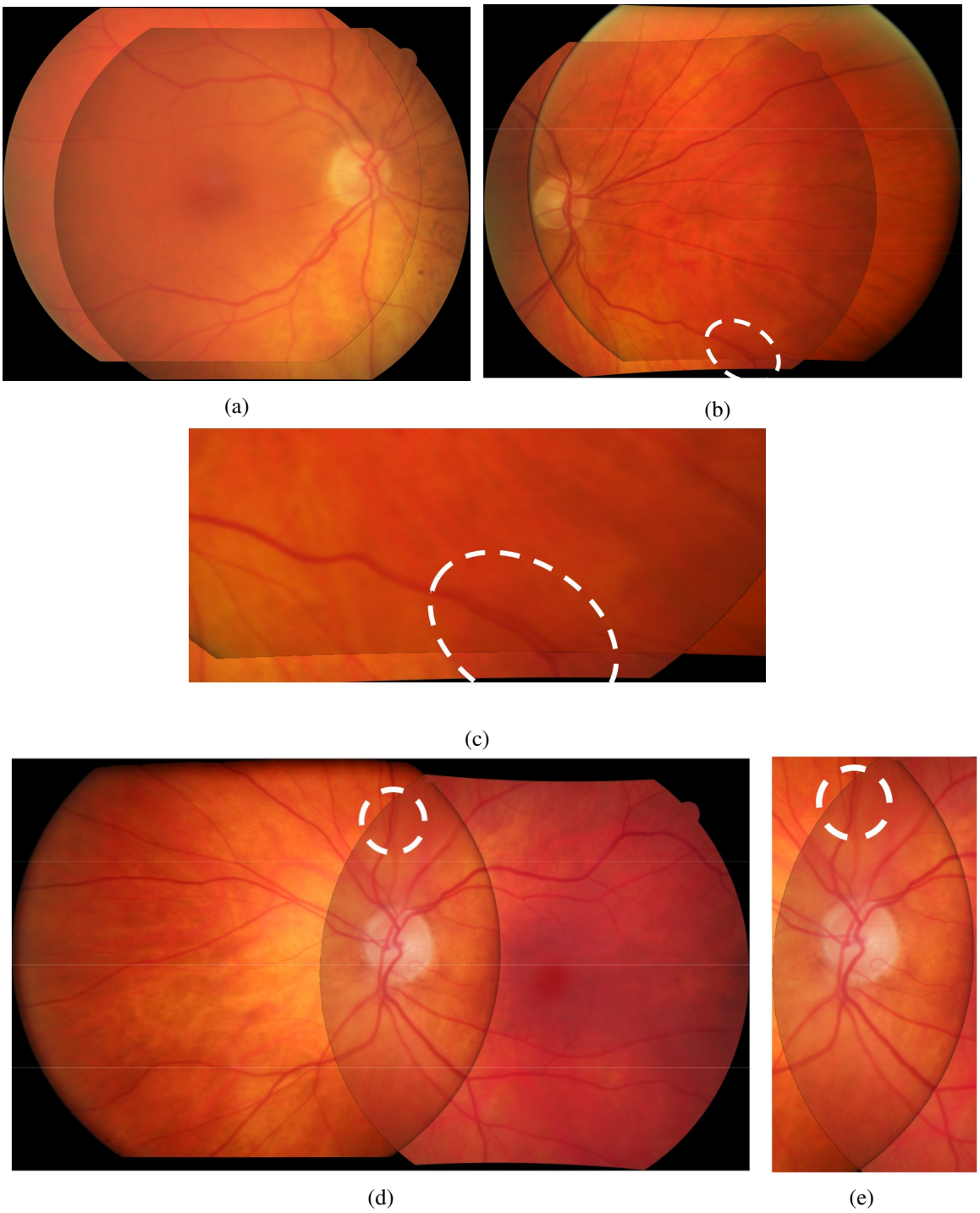


FIG. 6.10 – Exemples de superposition avec notre méthode. (a) Bonne superposition. (b) Superposition avec de petites différences (dans l'ellipse blanche) et son agrandissement (c) et (d) Superposition avec de petites différences (dans le cercle blanc) et son agrandissement (e). Il y a seulement un faible recouvrement dans la paire de (d) et (e).

sions radiales) les modèles quadratiques ont un plus grand nombre de degrés de liberté, ce qui est une source d'erreur supplémentaire. Donc, notre modèle va plus loin que les approches précédentes avec des modèles quadratiques [6], [42], [50], [161]. Quand la surface de recouvrement est faible (inférieure à 50 %), la méthode « gdbicp » quadratique avec des images normalisées (v) est complémentaire à notre méthode (i), à cause de la détection des vaisseaux comme repère pour le recalage.

Pour résumer la comparaison, notre méthode (i) est celle qui superpose le mieux les images de notre banque de données de santé publique (table 6.3). Le second modèle est celui de Lee et al. [159] (ii), dont les résultats sont systématiquement améliorés par notre méthode. La troisième méthode est le « gdbicp » quadratique [325] (iii) qui est amélioré quand les images sont normalisées (v). La méthode « gdbicp » quadratique, et aussi celle décrite par Lee et al. [159] reposent sur la détection des vaisseaux. Donc, lorsqu'elles sont utilisées sur des images avec un faible recouvrement, la détection des vaisseaux les rend plus efficaces et complémentaires à notre méthode. Cette amélioration sera présentée pour notre approche dans la sous-section 6.2.2, p. 222.

Les résultats obtenus avec le montage de validation et avec les banques de données de santé publique démontrent que notre modèle corrige mieux les erreurs provenant : des différentes positions du patient lors de l'acquisition d'image, du changement de caméra employée (résolutions et lentilles optiques), de la projection d'une scène 3D sur un plan et de la variabilité des couleurs entre les images.

En outre, l'étape de stabilisation couleur est utile lorsque les images présentent de fortes variations de contrastes. Les résultats obtenus pour l'algorithme « gdbicp » quadratique avec la stabilisation couleur (v), qui sont meilleurs avec que sans (iii), montrent l'importance d'utiliser des images dont la couleur est normalisée dans des banques de données de santé publique (table 6.3). Dans notre méthode (i), les points SIFT sont extraits et appariés sur les images dont le contraste couleur est normalisé avant d'ajuster le modèle de recalage. Les expériences ont montré que l'utilisation d'images normalisées donne des résultats plus robustes quand on extrait et apparie les points SIFT, même si ces points sont connus pour être robustes aux variations couleur [166], [308]. Surtout, même lorsque l'on utilise des images normalisées avec « gdbicp », notre méthode reste meilleure.

En outre, notre méthode a été testée pour des images acquises avec un champ de vision (*Field Of View* – FOV) de 45° qui sont utilisées pour du dépistage de la rétinopathie diabétique par des photographie du fond d'œil à champ simple (*single-field fundus photography*) [318]. De plus, notre approche pourrait être utile pour la détection automatique des patients à adresser à un ophtalmologiste pour cause de rétinopathie diabétique [2], [63], [80], [253].

Malgré les bons résultats obtenus, notre méthode a des limites. Pour des images avec un faible recouvrement (e.g. 30 % de la surface de l'image mosaïque), la superposition peut présenter de faibles différences sur la partie externe. Afin de résoudre ce problème, nous avons développé un autre algorithme qui utilise, en plus des points appariés, une distance entre des vaisseaux. Il sera présenté dans la sous-section 6.2.2, p. 222.

Conclusions de l'étude

Nous avons donc introduit une nouvelle méthode pour répondre aux principaux défis de la superposition des images du fond d'œil provenant de grandes banques de données de santé publique. En plus des méthodes précédemment existantes, la nôtre a été conçue pour gérer les changements de caméras, de lentilles, de résolution d'image et de fortes variations de couleur entre deux examens rétiniens. La méthode présentée consiste à ajuster un modèle de recalage composé d'une homographie et d'une ou deux distorsions radiales sur des points saillants extraits dans des images après stabilisation couleur. Le choix entre le nombre de distorsions radiales est fait automatiquement selon le nombre de caméras utilisées pour acquérir les paires d'images. Puis, la déformation de l'image est effectuée en utilisant un modèle par division, qui est inversible ; ce qui la rend rapide à calculer. Pour estimer les paramètres du modèle de manière rapide, des équations linéaires ont été introduites, suivies par des estimateurs non-linéaires. Surtout, la méthode est facile à utiliser et ne requiert pas d'extraire des caractéristiques intrinsèques comme des vaisseaux ou leurs embranchements.

L'efficacité de notre méthode a été validée sur un montage simulé avec une erreur de superposition qui est inférieure à celle d'un modèle quadratique. Son efficacité a aussi été démontrée sur des banques de données de santé publique avec des images de haute et de faible qualité, causées par

des différences dans les conditions d'acquisition, et sur des images acquises avec la même caméra ou des caméras différentes. Quoiqu'il en soit, les résultats montrent qu'il n'y a pas de différence visible entre les images provenant de deux examens avec l'œil dans la même position (nasale ou maculaire). La superposition est correcte dans plus de 92 % des cas (96 % avec une aire de recouvrement supérieure à 50 %). Dans ces banques de données, notre méthode superpose mieux les images que deux autres de l'état de l'art : celle de Lee et al. [159] et « gdbicp » de Stewart et al. [291], [325].

6.2.2 Ajustement d'un modèle de recalage sur les vaisseaux

Les travaux de cette partie sont extraits de l'article de conférence [210]. Ils ont été réalisés en collaboration avec l'équipe du Professeur David Owens, diabétologue de *Swansea University* et le centre national de dépistage *Diabetic Eye Screening Wales*, de Cardiff au Pays de Galle, Royaume-Uni.

Motivation

Dans la sous-section 6.2.1, p. 201, a été présentée une méthode de superposition des images du fond d'œil dans les grandes banques de données de santé publique, en vue de l'étude longitudinale de la rétinopathie diabétique. Notre approche comportait deux étapes : un pré-traitement destiné à corriger les variations de contraste et un modèle de recalage composé d'une affinité et de deux corrections de distorsion radiales. Les paramètres du modèle étaient estimés avec des points caractéristiques extraits par la *scale-invariant feature transform* (SIFT) [166]. Cette estimation s'avère généralement suffisante dans de nombreuses images. Cependant, certaines images peuvent présenter des écarts de recalage visibles sur leur partie externe, en particulier lorsque la surface de recouvrement est faible (inférieure à 50 %). Le but de cette étude est de résoudre ce problème en utilisant les vaisseaux, en plus des points SIFT, pour ajuster le modèle. Des équations explicites des frontières de la surface de recouvrement seront données afin de sélectionner efficacement les vaisseaux dans cette zone. L'étude est organisée en deux parties. Premièrement, la méthode améliorée sera présentée. Deuxièmement, elle sera comparée à une autre récente dans l'état de l'art, à savoir « REMPE » [116]. Un ensemble de données de santé publique sera utilisé avec 271 paires acquises principalement avec différentes caméras (cf. sous-sous-section « Résultats », p. 215) [234]. Les résultats obtenus précédemment dans cet ensemble de données, seront rappelés pour trois méthodes de l'état de l'art. Une deuxième comparaison sera effectuée dans l'ensemble de données disponible publiquement « FIRE » [118] et associé à la méthode « REMPE » [116].

Méthode

Ajustement du modèle Le modèle est celui présenté dans l'équation (6.3), p. 209 :

$$u_{k_2}(P_2^u) = H\left[u_{k_1}(P_1^u)\right].$$

Il se compose d'une homographie $H \in \mathbb{R}^{3\times 3}$ et de deux distorsions radiales $u_{k_1}(P_1^u)$ et $u_{k_2}(P_2^u) = [P_2^u, 1]^t \in \mathbb{R}^3$. Les points P_1^u et $P_2^u \in \mathbb{R}^2$ sont les coordonnées non-distordues dans les images non-distordue n°s 1 et 2. Elles sont reliées aux coordonnées des points distordus P^d dans les images initiales par la relation $P^u - C = (P^d - C)/(1 + k\|P^d - C\|^2)$, où $C \in \mathbb{R}^2$ représente le centre de l'image. Si la même caméra est utilisée pour capturer les deux images, les paramètres réels de distorsion sont égaux à $k_1 = k_2 = k$.

Le modèle est ajusté par une approche en plusieurs étapes (Fig. 6.11) décrites ci-après.

(1) Comme de nombreuses images dans les banques de données de santé publique présentent une clarté non-uniforme, un pré-traitement corrige les variations locales de contraste couleur.

(2) Des points caractéristiques sont extraits en utilisant l'algorithme SIFT [166] et mis en correspondance entre les images selon la méthode présentée dans le paraphe « Mise en correspondance des points avec une estimation préliminaire de la déformation de l'image », p. 207 [234].

(3) Les points appariés servent à initialiser les paramètres du modèle et le nombre de distorsions – 1 ou 2 – est automatiquement choisi.

(4) Une estimation itérative est effectuée sur les paramètres jusqu'à la convergence de l'erreur de recalage. Des estimateurs linéaires sont utilisés pour initialiser les optimiseurs non-linéaires [234].

(5) Un optimiseur non-linéaire affine l'estimation des paramètres du modèle [234]. Dans cette sous-section, aux étapes (4) et (5), nous remplacerons l'erreur entre les points SIFT, que nous avons utilisée précédemment, dans [234] par une erreur entre les vaisseaux.

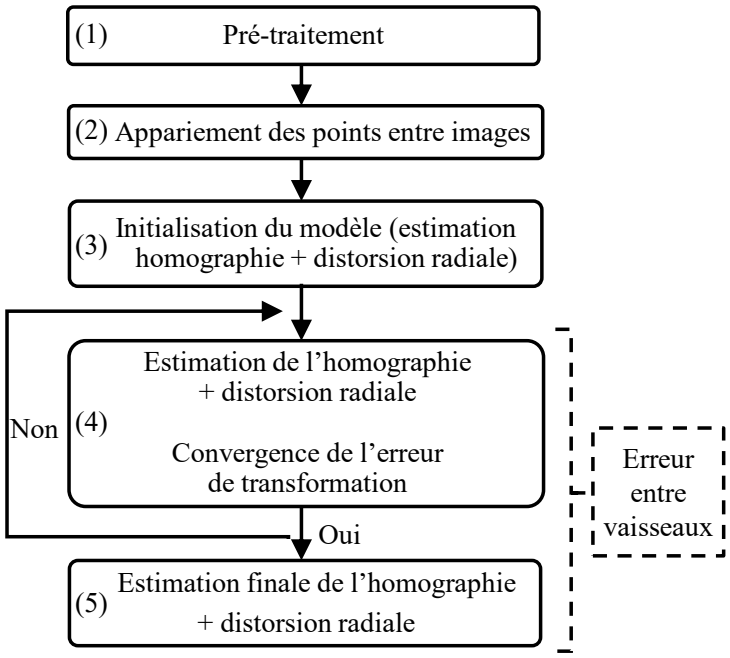


FIG. 6.11 – Organigramme de la méthode. Une accolade en pointillés indique l'amélioration apportée grâce à la minimisation de l'erreur entre les vaisseaux.

Un critère d'erreur entre vaisseaux Les points SIFT sont utilisés pour l'étape d'initialisation (3) (Fig. 6.11). Cependant, durant les étapes (4) et (5), une erreur entre les vaisseaux les plus proches est minimisée afin d'assurer une meilleure superposition sur l'ensemble de la surface de recouvrement. Dans le plan, l'erreur est mesurée entre les lignes centrales des vaisseaux les plus proches extraites par la méthode de Staal et al. [287]. Soit $\mathcal{S}$ la courbe (de la ligne centrale) d'un vaisseau source et $\mathcal{R}$ la courbe du vaisseau de référence correspondant. L'erreur « courbe-à-courbe » $d(\mathcal{S},\mathcal{R})$ est la somme des carrés des distances entre tous les points s de $\mathcal{S}$ à la courbe $\mathcal{R}$, $d(\mathcal{S},\mathcal{R}) = \sum_{\mathrm{s}\in\mathcal{S}} d^2(\mathrm{s},\mathcal{R})$ [40]. Le carré de la distance point-courbe $d^2(\mathrm{s},\mathcal{R})$ est défini pour chaque point $\mathrm{s} \in \mathcal{S}$ comme le carré de la distance euclidienne à ses *points les plus proches* r^* sur $\mathcal{R}$, $d^2(\mathrm{s},\mathcal{R}) = \min_{\mathrm{r}\in\mathcal{R}} \|\mathrm{r}-\mathrm{s}\|_2^2 = \|\mathrm{r}^*-\mathrm{s}\|_2^2$. Une approximation du second ordre du carré de la distance point-courbe, a été définie dans [247] pour les courbes discrètes, comme :

$$d^2(\mathrm{s},\mathcal{R}) \approx \frac{d}{d-\rho}[(\mathrm{s}-\mathrm{r}^*)\cdot \mathrm{t}(\mathrm{r}^*)]^2 + [(\mathrm{s}-\mathrm{r}^*)\cdot \mathrm{n}(\mathrm{r}^*)]^2,$$

où $\mathrm{t}(\mathrm{r}^*)$ et $\mathrm{n}(\mathrm{r}^*)$ sont les vecteurs unitaires tangents et normaux à la courbe définis au point r^*. « $\cdot$ » est le produit scalaire. ρ est le rayon de courbure (signé) au point r^*. d est la distance signée

au point le plus proche r^* définie par $d = \|\mathrm{s} - \mathrm{r}^*\|_2$, lorsque r^* et $\mathrm{n}(\mathrm{r}^*)$ sont du même côté de la courbe, et par $d = -\|\mathrm{s} - \mathrm{r}^*\|_2$, sinon. La distance signée d trouve le point $\mathrm{r}^* \in \mathcal{R}$ le plus proche dans la courbe de référence, pour chaque point $\mathrm{s} \in \mathcal{S}$ de la courbe source. Afin d'éviter les problèmes de distances d anormalement élevées, nous sélectionnons seulement les points dont la distance signée d est telle que $\left(1 - (d/(a\sigma(d))^2)\right)^2 > 0$; ce qui correspond à la fonction de poids de « Beaton-Tukey », avec $a = 4{,}685$ [42], [120] et $\sigma(d)$ l'écart-type de d pour tout $\mathrm{s} \in \mathcal{S}$. En pratique, les rayons de courbures sont calculés une fois pour toutes les itérations de l'étape (4). Le carré de la distance $d^2(\mathrm{s}, \mathcal{R})$ est estimé entre le point s et son plus proche vaisseau $\mathcal{R}$.

Equations de la zone de recouvrement pour sélectionner les vaisseaux La connaissance des équations des frontières de la zone de recouvrement permet de sélectionner uniquement les vaisseaux présents dans cette surface. Dans chaque image n° $i \in \{1; 2\}$, un cercle est ajusté sur les frontières du champ de vision. Son centre correspond à celui de l'image $C_i = [x_i, y_i]^t \in \mathbb{R}^2$ et son rayon est noté r_i. L'équation du cercle de l'image n° i peut s'exprimer comme $(x - x_i)^2 + (y - y_i)^2 = r_i^2$ ou sous forme matricielle par $X^t \Gamma_i X = 0$. $X = [x, y, 1]^t$ représente la variable de l'équation et la matrice du cercle Γ_i est égale à :

$$\Gamma_i = \begin{bmatrix} 1 & 0 & -x_i \\ 0 & 1 & -y_i \\ -x_i & -y_i & -r_i^2 + x_i^2 + y_i^2 \end{bmatrix}. \tag{6.13}$$

La distorsion radiale transforme le disque de rayon r_i en un disque non-distordu de rayon $r_i^u = r_i/(1+k_i r_i^2)$, dont l'équation est $X^t \Gamma_i^u X = 0$. La matrice Γ_i^u à la même expression que la matrice Γ_i excepté le rayon r_i qui est remplacé par r_i^u. A cause de la transformation par homographie $X' = \mathrm{H}X$, l'équation du disque non-distordu n° 1 devient [106] :

$$X^t \Gamma_1^u X = X'^t (\mathrm{H}^{-1})^t \Gamma_1^u (\mathrm{H}^{-1}) X' = X'^t \Gamma_1^r X', \tag{6.14}$$

où $\Gamma_1^r = (\mathrm{H}^{-1})^t \Gamma_1^u (\mathrm{H}^{-1})$. Γ_1^r est la matrice d'une conique [106] et dans notre cas d'une ellipse. En effet, le déterminant du bloc supérieur gauche 2×2 de la matrice est strictement positif $det(\Gamma_1^r(1; 2-1; 2)) = (det(A))^2 > 0$ [238, chap. 7.5]. La surface de recouvrement entre le cercle Γ_2^u et l'ellipse Γ_1^r est alors définie par les points à l'intérieur du cercle et de l'ellipse :

$$\begin{cases} X'^t \Gamma_1^r X' & \leq & 0 \\ X'^t \Gamma_2^u X' & \leq & 0 \end{cases}. \tag{6.15}$$

A chaque itération de l'estimation du modèle, ce système d'équation permet de sélectionner uniquement les vaisseaux à l'intérieur de la surface de recouvrement.

Expériences et résultats

Des expériences ont été menées afin de comparer la méthode actuelle à d'autres. Deux ensembles de données ont été utilisés : un ensemble de santé publique [234] et un autre disponible publiquement.

Expériences dans un ensemble de données de santé publique Ce jeu de données de santé publique se compose de 69 patients sélectionnés aléatoirement et provenant du programme *Diabetic Eye Screening Wales* (DESW), au Royaume-Uni. Tous les patients avaient du diabète et différents stades de sévérité de rétinopathie ou de maculopathie. Chacun d'eux a été dépisté annuellement durant plusieurs années et 4 images étaient disponibles par dépistage. Nous avons sélectionné une série de 271 paires d'images :

(1) de qualité suffisante ;

(2) avec un intervalle de dépistage d'environ une année entre les événements d'examens [234] ;

(3) capturées lorsque le service de dépistage renouvelait ses caméras du fond d'œil. 63 % des paires ont été capturées avec des caméras différentes – i.e. des résolutions et distorsions différentes.

10 paires avaient une faible surface de recouvrement d'environ 30 % de l'aire de l'image superposée. Toutes les photographies rétiniennes étaient de haute qualité et étaient capturées selon un protocole incluant une dilatation de la pupille. Afin de valider la qualité de superposition, toutes les surfaces de recouvrement recalées ont été vérifiées avec soin par un expert selon la classification visuelle présentée dans [234] (cf. sous-paragraphe « Validation avec une banque de données de santé publique », p. 213). Deux catégories ont été considérées :

(a) pas de différence visible (i.e. *correcte*) et

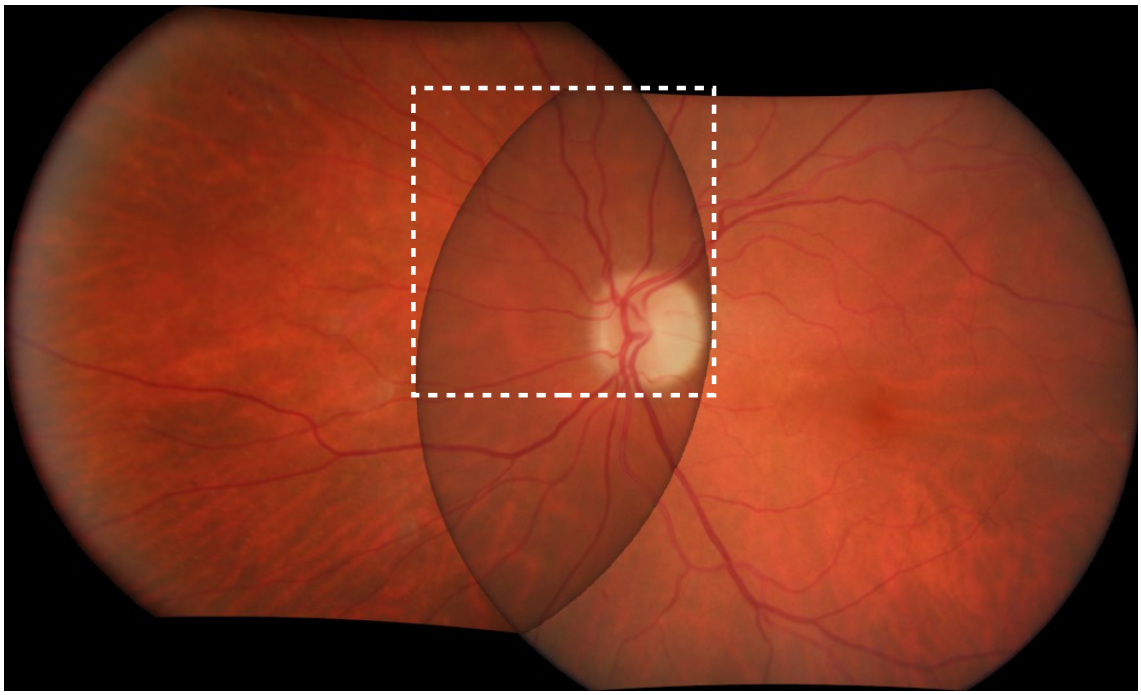
(b) différence visible (i.e. incorrecte) avec trois sous-catégories : (b.1) différences de la taille d'un vaisseau de petit diamètre, (b.2) différences de la taille d'un vaisseau de gros diamètre ou (b.3) même plus large. Les trois sous-catégories ont été regroupées en une seule *incorrecte*.

En utilisant cette échelle visuelle, nous avons évalué (i) la méthode actuelle et nous l'avons comparée aux résultats obtenus dans [234] dans le même jeu de données pour trois autres méthodes : (ii) la précédente [234], (iii) la méthode de Lee et al [159] et (iv) « gdbicp » quadratique [325]. Nous avons ajouté une comparaison avec une méthode plus récente (v) « REMPE » [116] qui repose sur une hypothèse d'œil de forme sphérique. Les paramètres de référence ont été utilisés pour cette méthode.

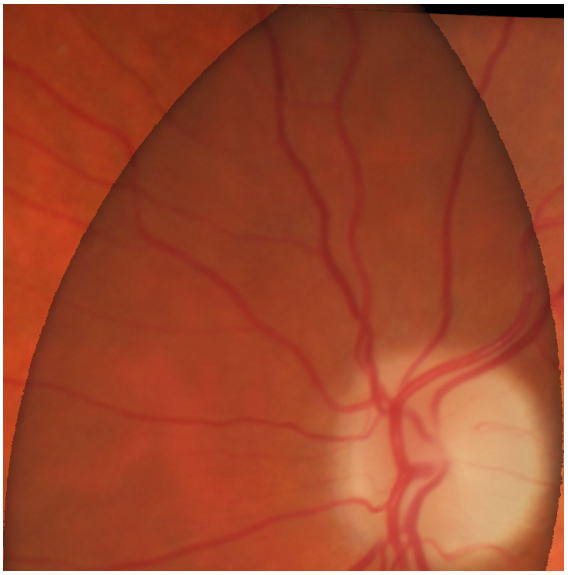
Résultats dans l'ensemble de données de santé publique Dans la table 6.4 ; 96 % des paires sont correctement superposées (i) avec la méthode actuelle. Cette méthode recale mieux les paires que : (ii) la précédente [234], (iii) celle de Lee et al. [159], (v) « REMPE » et (iv) « gdbicp » quadratique [325]. La figure 6.12 illustre la méthode actuelle (i) qui superpose avec succès une paire avec un faible recouvrement, tandis que la précédente (ii) échoue. Notre méthode est donc meilleure que les autres (ii), (iii), (iv) et (v), dans cet ensemble de données de santé publique où 63 % des paires ont été capturées avec des caméras différentes.

	Méthode	Superpositions correctes
(i)	**Modèle actuel (vaisseaux)**	**96 %**
(ii)	Modèle actuel (points SIFT) [234]	92 %
(iii)	Lee et al. [159]	88 %
(v)	« REMPE » [116]	75 %
(iv)	« gdbicp » quadratique [325]	74 %

TABLE 6.4 – Pourcentage décroissant des superpositions correctes pour cinq méthodes dans un jeu de données de santé publique.



(a) Méthode actuelle (i).



(b) Agrandissement sur (a).

(c) Méthode précédente (ii).

FIG. 6.12 – Superposition d'une paire d'images avec un faible recouvrement. (a) Recalage correct avec la méthode actuelle (i). (b) Agrandissement sur (a). (c) Recalage incorrect avec la méthode précédente (ii).

Expériences dans un ensemble de données disponible publiquement : FIRE Nous avons aussi comparé (i) la méthode actuelle à (iv) « REMPE » en utilisant l'ensemble de données FIRE qui inclut une vérité terrain [118]. Il se compose de 129 images formant 134 paires et divisées en 3 catégories :

(1) la catégorie $\mathcal{S}$ qui contient 71 paires, avec un fort recouvrement et pas de différence anatomique ;

(2) la catégorie $\mathcal{P}$ qui contient 49 paires, avec un faible recouvrement et pas de différence anatomique ;

(3) la catégorie $\mathcal{A}$ qui contient 14 paires, avec un fort recouvrement et de grands changements anatomiques.

La vérité terrain a été créée par des experts [118] qui ont sélectionnés 10 points correspondants dans chaque paire d'images. Pour les deux méthodes (i) et (v), nous avons recalé ces points et nous avons calculé la moyenne des distances euclidiennes entre chaque image de la paire.

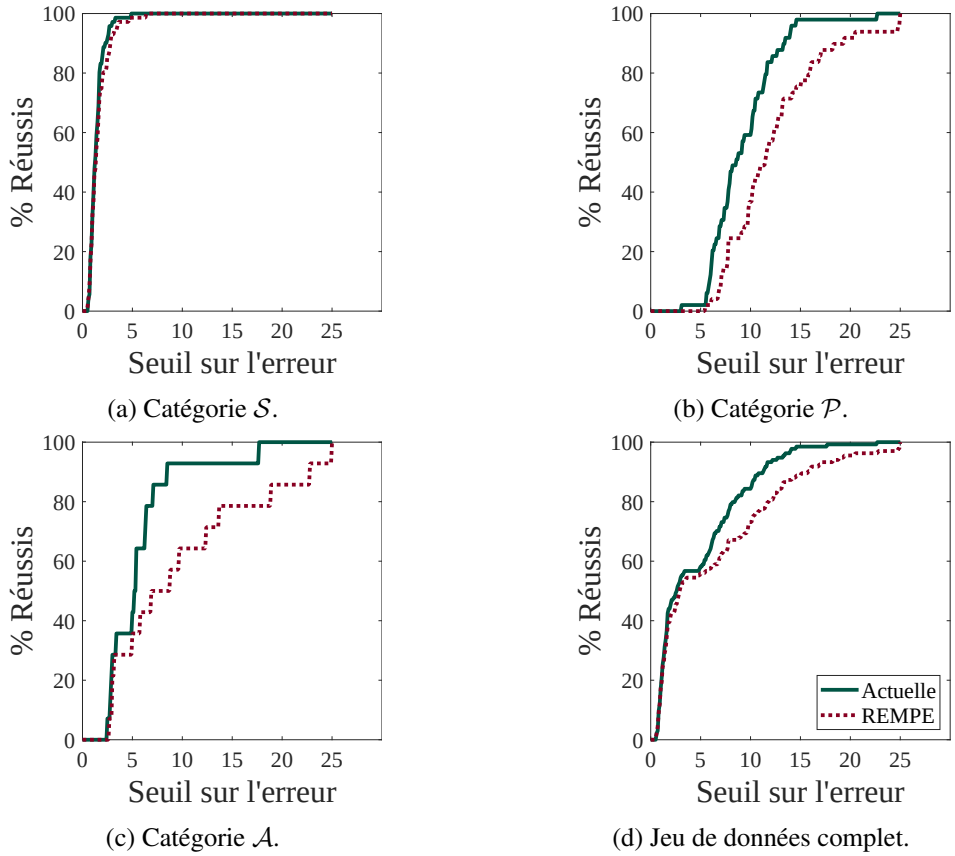
Résultats dans l'ensemble de données disponible publiquement : FIRE L'erreur de recalage est donnée entre les graphiques de la figure 6.13 selon l'approche de Hernandez-Matas et al. [116], [118]. L'axe horizontal représente le seuil sur l'erreur moyenne en dessous duquel un recalage est considéré comme réussi. Un pas de 0,1 pixel est utilisé pour le seuil sur l'erreur. L'axe vertical est le pourcentage de recalages réussis. L'aire sous la courbe (*Area Under Curve* – AUC) est aussi calculée pour chaque courbe. Plus la valeur de la courbe de recalage (ou l'AUC) est élevée, meilleur est le recalage. Les courbes de recalage (Fig. 6.13), l'aire sous la courbe (AUC, table 6.5) et l'erreur moyenne (table 6.6) sont données pour chaque catégorie $\mathcal{S}$, $\mathcal{P}$ ou $\mathcal{A}$ et pour l'ensemble du jeu de données. Dans chacun des cas, la courbe de recalage et l'aire sous la courbe (AUC) de (i) la méthode actuelle sont supérieure à celles de (v) « REMPE » (table 6.5). Dans chacun des cas, l'erreur moyenne et l'écart-type de (i) la méthode actuelle sont inférieurs à ceux de (v) « REMPE » (table 6.6). La méthode actuelle est donc meilleure que « REMPE » dans toutes les catégories du jeu de données FIRE et particulièrement pour les catégories $\mathcal{P}$ (avec un faible recouvrement) et $\mathcal{A}$ (avec de grandes différences anatomiques). Cependant, pour ces deux catégories, l'erreur moyenne reste élevée, 9,25 et 5,81 pixels, respectivement (table 6.6), pour des images de taille $2\,912^2$ pixels. La superposition peut encore être améliorée dans ces deux catégories, en particulier pour les images avec un faible recouvrement (moins de 50 % de la surface de l'image recalée).

	Méthode	$\mathcal{S}$	$\mathcal{P}$	$\mathcal{A}$	FIRE
(i)	**Méthode actuelle**	**0,942**	**0,632**	**0,768**	**0,810**
(v)	« REMPE »	0,935	0,511	0,599	0,745

TABLE 6.5 – Aire sous la courbe (AUC) de la méthode actuelle et de « REMPE » pour les catégories $\mathcal{S}$, $\mathcal{P}$ et $\mathcal{A}$ et l'ensemble du jeu de données FIRE.

	Méthode	$\mathcal{S}$	$\mathcal{P}$	$\mathcal{A}$	FIRE
(i)	**Méthode actuelle**	**1,46 (1,12)**	**9,25 (10,00)**	**5,81 (7,21)**	**4,76 (7,47)**
(ii)	« REMPE »	1,63 (1,57)	12,64 (15,19)	14,05 (25,73)	6,96 (13,65)

TABLE 6.6 – Erreur moyenne (et son écart-type) de la méthode actuelle et de « REMPE » pour les catégories $\mathcal{S}$, $\mathcal{P}$, $\mathcal{A}$ et l'ensemble du jeu de données FIRE. Les unités sont en pixels.



(a) Catégorie $\mathcal{S}$.

(b) Catégorie $\mathcal{P}$.

(c) Catégorie $\mathcal{A}$.

(d) Jeu de données complet.

FIG. 6.13 – Courbes de recalage des différentes catégories (a) $\mathcal{S}$, (b) $\mathcal{P}$, (c) $\mathcal{A}$ et (d) pour le jeu de données complet FIRE. Le seuil sur l'erreur est la valeur en dessous de laquelle un recalage est considéré comme réussi. L'axe vertical représente le pourcentage de recalages réussis.

Conclusions de l'étude

Nous avons introduit un nouveau critère d'erreur pour estimer notre modèle de recalage composé d'une affinité (i.e. rotation, translation et mise à l'échelle anisotrope) et de deux distorsions radiales. Ce critère repose sur une erreur entre les vaisseaux qui sont sélectionnés par les équations de la frontière de la surface de recouvrement, i.e. l'intersection entre un disque et une surface d'ellipse. Ces équations sont déduites de celles du modèle. Des expériences ont montré que notre méthode superpose avec succès 96 % des paires d'un ensemble de données de santé publique dont les paires d'images sont surtout acquises avec des caméras différentes. Ce résultat est meilleur que celui de notre méthode précédente [234] et que ceux de trois autres de l'état de l'art [116], [159], [325]. Dans le jeu de données publiquement disponible, « FIRE » [118], notre approche superpose toujours mieux les images que la méthode de l'état de l'art « REMPE » [116] même si toutes les paires ne sont pas parfaitement superposées. Quoiqu'il en soit, les résultats montrent que notre méthode est efficace pour les images de banques de données de santé publique qui sont employées pour le dépistage de la rétinopathie diabétique et qui présentent de fortes variations de contrastes et des distorsions radiales.

6.3 Conclusion du chapitre

Dans ce chapitre, différentes méthodes de recalage ont été présentées, tant dans le domaine des pneumatiques que dans celui des images du fond d'œil. La superposition entre une image d'une coupe radiale de pneumatique et son épure (i.e. son plan), ne peut être réalisée de manière

rigide puisque le pneumatique est un objet déformable. Le stage de Florian Rousset, que j'ai supervisé, a permis de tester sur le pneumatique l'approche de recalage souple [265] proposée par Rouhanni et Sappa [263], [264] dans un autre domaine applicatif. Celle-ci repose sur un ajustement d'une forme de référence sur une autre grâce à des splines cubiques réparties sur une grille de déformation. L'ajustement d'une forme sur l'autre revient à la résolution d'un système linéaire ce qui le rend rapide à calculer et est donc très utile pour recaler de grandes images. Les premiers résultats obtenus ont été concluants sur le pneumatique.

Pour les images du fond d'œil, j'ai conçu un modèle de recalage pour l'étude longitudinale (i.e. temporelle) de la rétinopathie diabétique chez les patients. La difficulté consistait à superposer des images provenant de banques de données de santé publiques utilisées pour le dépistage de la rétinopathie diabétique. Les images avaient été acquises dans des conditions hétérogènes et avec une année d'intervalle. Ces hétérogénéités peuvent être causées par des changements d'éclairement, un changement éventuel de caméra, un mouvement de l'œil du patient, une évolution de la sévérité de la rétinopathie, une opération de la cataracte, etc. Cela a pour conséquence, des variations de contraste et des déformations dans les images notamment causées par la projection d'une surface non plane, i.e. l'œil, sur le plan du capteur de la caméra. Afin de superposer les images de la rétine, j'ai introduit un nouveau modèle de recalage composé d'une homographie affine (i.e. une rotation, translation et mise à l'échelle anisotrope) et de la correction d'une ou deux distorsions radiales, selon le remplacement ou non de la caméra entre les images de la paire. Pour ajuster le modèle, j'ai proposé des estimateurs linéaires des paramètres suivis d'optimiseurs non linéaires. Deux versions de l'erreur de recalage ont été présentées :

(i) soit une erreur euclidienne entre des points caractéristiques [234] extraits par la méthode SIFT *scale-invariant feature transform* (SIFT) [166] dans les images,

(ii) soit une erreur entre des courbes [40], [247] : les lignes centrales des vaisseaux de l'image [210] après initialisation sur les points SIFT.

Chacune des versions a donné lieu à publication. La première version a été publiée comme article de journal [234]. Elle a été réalisée avec l'équipe du Professeur David Owens, diabétologue à *Swansea University* et le centre national de dépistage *Diabetic Eye Screening Wales*, de Cardiff au Pays de Galle, Royaume-Uni. La deuxième version a donné lieu, pour l'instant, à un article de la conférence *IEEE International Symposium on Biomedical Imaging 2020* [210]. La deuxième version améliore la première, en particulier lorsque la surface de recouvrement entre les deux images est faible (par rapport à la surface de l'image recalée). Comparées à d'autres méthodes de l'état de l'art, les deux versions de la nôtre sont meilleures dans les banques de données de santé publique. Afin d'améliorer encore le recalage entre les images du fond d'œil, il pourrait être intéressant de travailler sur un modèle tridimensionnel du fond d'œil comme proposé dans [117]. Il est cependant important de na pas trop complexifier le modèle afin de ne pas accroître l'erreur de recalage.

En outre, la méthode d'amélioration des contrastes couleur du fond d'œil que j'ai proposée, améliore la qualité de recalage surtout pour des images provenant de grandes banques de données de santé publique.

Chapitre 7

Conclusion et perspectives de recherche

7.1 Conclusion

Ce mémoire présente mes travaux sur l'analyse morphologique et logarithmique de grandes banques d'images à la fois industrielles et médicales. Ma démarche a consisté à observer des phénomènes, à formuler une problématique de recherche puis à proposer un cadre théorique permettant de lever des verrous technologiques. Cette approche peut s'expliquer par mon parcours tant industriel qu'académique puisqu'après mon doctorat au Centre de Morphologie Mathématique de l'Ecole des Mines de Paris, j'ai choisi d'avoir une expérience significative dans la recherche industrielle avant de revenir vers la recherche académique. Ces travaux ont été menés sur douze années : 6 ans comme chercheur au Centre de Recherche mondial de Michelin (3500 chercheurs) à Clermont-Ferrand et 6 ans comme *Research Director* à l'*International Prevention Research Institute* (IPRI) à Lyon. Depuis 2017, je suis aussi *Visiting Researcher* à *University of Strathclyde* à Glasgow en Ecosse, Royaume-Uni. Dans ce mémoire, mes travaux sont décrits comme il s'ensuit.

Après un rappel du contexte de recherche dans le chapitre 2, j'ai présenté mes travaux en segmentation morphologique et en apprentissage statistique dans les grandes banques d'images au chapitre 3. Ceux-ci ont porté sur l'analyse automatique de la surface des pneumatiques pour le contrôle industriel. Des images de très grande taille et très nombreuses ont été segmentées par Morphologie Mathématique afin d'extraire des anomalies en relief ou des anomalies dans des textures. J'ai créé moi-même ou grâce à des collaborations avec l'équipe du Professeur Dominique Jeulin, de l'Ecole des Mines de Paris, de nouvelles méthodes de segmentation et d'analyse de texture pour les pneumatiques qui ont permis l'obtention de 8 brevets dans plusieurs pays ou entités géopolitiques (Union européenne, Etats-Unis, Japon, Chine). De plus, j'ai travaillé à l'accélération d'algorithmes morphologiques pour le contrôle des pneumatiques en collaboration avec l'Ecole des Mines de Paris ou bien en supervisant un jeune ingénieur de recherche, Pierre Guilbert, durant une année à l'IPRI. Puis, j'ai encadré deux stagiaires, chacun durant six mois. La première, Sabina Stefan, diplômée d'un *Bachelor* (Bac +4) de *University of Cape Town* en Afrique du Sud, a testé des méthodes d'apprentissage statistique pour évaluer la qualité des images du fond d'œil dans les grandes banques de données de santé publique servant au dépistage de la rétinopathie diabétique. Le deuxième, Lucas Arnaldo Alves, étudiant brésilien en échange à Télécom Saint-Etienne, a effectué son stage de fin d'études d'ingénieur et de master sur la classification automatique des images du fond d'œil par apprentissage statistique.

Dans le chapitre 4, j'ai étudié l'analyse des images faiblement contrastées grâce au Traitement d'Image Logarithmique (*Logarithmic Image Processing* – LIP) et aux métriques fonctionnelles d'Asplund. J'ai proposé un algorithme d'amélioration des contrastes couleur locaux des images du fond d'œil. Cet algorithme a été testé sur plusieurs milliers d'images provenant de grandes banques de données de santé publique. Il permet de rendre le diagnostic plus rapide et plus évident pour le médecin. Pour cette étude, j'ai obtenu deux financements de la Banque Publique d'Investissement (BPI-France) et un projet européen (Eye-Light) du programme Horizon 2020, afin de développer le système et de planifier sa validation clinique. Celle-ci a été organisée avec le Professeur Chris Robertson, épidémiologiste pour la santé publique au Département de Mathématiques et Statistiques de *University of Strathclyde*. Ces travaux ont d'abord été réalisés avec l'équipe du Professeur David Owens, diabétologue à *Swansea University*, avec le soutien du centre national de

dépistage *Diabetic Eye Screening Wales*, de Cardiff au Pays de Galle, Royaume-Uni. La deuxième partie de l'étude a été réalisée avec l'équipe du Professeur Laurent Kodjikian, ophtalmologiste aux Hospices Civils de Lyon. En outre, avec le Professeur Michel Jourlin, de l'Université Jean Monnet à Saint-Etienne, j'ai étudié les métriques d'Asplund fonctionnelles définies dans le cadre du LIP ; ce qui leur donne de bonnes propriétés optiques et notamment d'être robustes aux variations d'éclairement. Elles sont donc bien adaptées à l'analyse des grandes banques d'images acquises dans des conditions d'éclairement hétérogènes, notamment en santé publique ou dans l'industrie. Ces métriques sont très utiles pour la reconnaissance de forme car elles permettent de calculer une carte de distances entre une fonction de référence et une image. J'ai établi le lien de ces cartes de distances d'Asplund avec la Morphologie Mathématique, ce qui les place dans un cadre bien connu. J'ai aussi étendu ces métriques aux images couleur et multivariées. Le stage de fin d'études d'ingénieur et de Master de Benoît Morel, étudiant à CPE-Lyon, a permis d'illustrer leur intérêt sur des radiographies de pneumatiques. Puis, j'ai établi le cadre de la Morphologie Mathématique Logarithmique en définissant les opérations fondamentales d'érosion et de dilatation grâce au modèle LIP, ce qui les rend bien adaptées aux images. En effet, la somme de deux images n'est pas linéaire mais est de type logarithmique et de plus les images ont une valeur maximale indépassable qui est prise en compte dans la définition du modèle LIP. Grâce au nouveau cadre de la Morphologie Mathématique Logarithmique, des opérateurs adaptatifs ou robustes aux variations d'éclairement peuvent être créés. J'ai aussi montré que les cartes de distances d'Asplund pour la loi LIP-additive étendent le gradient LIP-morphologique aux fonctions structurantes (i.e. aux éléments structurants non-plats). Elles correspondent, en effet, à une différence-LIP entre une dilatation et une érosion logarithmiques. Les cartes de distances d'Asplund LIP-multiplicatives, correspondent, quant à elles, à une différence entre une dilatation et une érosion fonctionnelles. Avec le Professeur Michel Jourlin, j'ai défini de nouveaux critères d'homogénéité de région fondés sur des contrastes logarithmiques. Lorsqu'ils sont intégrés dans une méthode de type croissance de région, les segmentations obtenues s'avèrent robustes aux variations importantes d'éclairement. En lien avec l'équipe du Professeur Laurent Kodjikian, ophtalmologiste aux Hospices Civils de Lyon, j'ai proposé un nouvel algorithme de segmentation des vaisseaux du fond d'œil à partir d'opérateurs de Morphologie Mathématique Logarithmique. Avec le Professeur Michel Jourlin, j'ai étudié les méthodes de percolation logarithmique à niveaux de gris pour le suivi de fils métalliques ou de frontières des composants des pneumatiques. J'ai aussi étendu les méthodes de percolation grâce à des contrastes couleur.

Dans le chapitre 5, j'ai exposé mes travaux en stéréovision pour l'analyse des performances des pneumatiques. J'ai développé une chaîne complète d'acquisition tri-dimensionnelle allant de la capture d'images, au calibrage des caméras et à la reconstruction 3D. Le stage de fin d'étude de Thibaud Vilboux, élève-ingénieur à Télécom Saint-Etienne, a permis d'étudier le calibrage des caméras dans le domaine du visible. Des discussions avec le Professeur Jean-Marc Lavest, de l'Université Blaise Pascal à Clermont-Ferrand, et mes travaux ont permis de concevoir et de réaliser un système complet de reconstruction 3D dans le domaine de la lumière visible, par stéréovision photographique, et dans le domaines des rayons X, par stéréovision radiographique. J'ai proposé une méthode de calibrage pour un système de stéréovision par rayons X, incluant des estimateurs initiaux des paramètres du système qui sont ensuite affinés par optimisation.

Le chapitre 6 présente mes travaux en recalage d'images sur les pneumatiques puis sur la rétine. Afin de superposer une image d'une coupe radiale de pneumatique avec son épure (i.e. son plan), le stage de fin d'étude et de Master de Florian Rousset, élève-ingénieur à CPE-Lyon, a permis d'étudier une méthode de recalage déformable par splines cubiques. Celle-ci appartient à la catégorie des formes libres de déformation. Ses paramètres de déformation présentent l'avantage de pouvoir être estimés de manière linéaire. Dans les banques de données de santé publique, à des fins d'analyse longitudinale de la rétinopathie diabétique, j'ai conçu une nouvelle méthode de recalage des images du fond d'œil. Celle-ci repose sur une transformation par homographie affine (i.e. rotation, translation et mise à l'échelle anisotrope) ainsi que sur une ou deux corrections

de distorsions radiales. Cette méthode entièrement automatique donne de meilleurs résultats que d'autres de l'état de l'art (96 % de superpositions correctes dans une série de 271 paires d'images acquises à un an d'intervalle et principalement avec des caméras différentes entre les deux photographies rétiniennes). Ces travaux ont été réalisés avec l'équipe du Professeur David Owens, diabétologue à *Swansea University*, avec le soutien du centre national de dépistage *Diabetic Eye Screening Wales*, de Cardiff au Pays de Galle, Royaume-Uni.

7.2 Perspectives

Concernant les perspectives de mes travaux, outre les améliorations de points particuliers sur chacune des méthodes, deux axes peuvent être dessinés.

i) Le premier concerne le développement de méthodes morphologiques et logarithmiques.

ii) Le second vise à combiner la Morphologie Mathématique avec l'apprentissage statistique.

7.2.1 Méthodes morphologiques et logarithmiques

Le premier axe consiste notamment en l'étude des propriétés du modèle LIP et de la Morphologie Mathématique à travers le cadre de la Morphologie Mathématique Logarithmique, que j'ai développé. Grâce à ce nouveau cadre, des opérateurs robustes aux variations d'éclairement modélisées par une addition LIP et adaptés aux images vont pouvoir être définis. Jusqu'à présent, j'ai étudié deux opérateurs présentant cette propriété de robustesse : à savoir la carte des distances d'Asplund LIP-additives et la différence-LIP d'érosions logarithmiques. Dans le futur, je souhaite définir d'autres opérateurs robustes à ces variations ; e.g. des différences-LIP de dilatations ou d'ouvertures et fermeture logarithmiques. La dilatation logarithmique géodésique par des fonctions structurantes pourrait aussi être étudiée ; ce qui permettrait l'introduction d'algorithmes de Lignes de Partages des Eaux robustes aux variations d'éclairement. La morphologie mathématique logarithmique LIP-multiplicative pourrait aussi être définie et étudiée avec la loi LIP-multiplicative. Ceci ouvrira la voie à la définition d'opérateurs morphologiques robustes aux variations de contraste causées par des changements d'opacité de l'objet et qui constitue une difficulté d'analyse importante dans les grandes banques d'images acquises en lumière transmise (radiographie, tomographie, spectrophotométrie, etc.).

En outre, la Morphologie Mathématique Logarithmique définit des opérateurs analysant une fonction à partir d'une fonction structurante (i.e. non plate). Bien que la Morphologie Mathématique fonctionnelle permette théoriquement l'étude d'une fonction par une autre fonction, en pratique la fonction structurante est presque toujours plate. La Morphologie Mathématique Logarithmique permet également d'envisager de définir des opérateurs par double palpation de la fonction avec deux sondes différentes l'une de l'autre et adaptées à la forme à détecter. Ce cadre de travail ouvre donc un nouveau champ exploratoire.

Quant aux contrastes logarithmiques, ils pourraient être étudiés sur l'*arbre des composantes* des ensembles de niveaux de gris de l'image [242], [267]. Ces arbres sont théoriquement invariants aux changements de contraste causés par une fonction continue et croissante appliquée à l'image [191]. Cependant, lorsque le contraste maximal de l'image présente une dynamique très faible, ces méthodes sont limitées. Dans ce cas, les contrastes logarithmiques pourraient étendre la robustesse de ces approches aux faibles contrastes. Une autre application des contrastes logarithmiques concerne la percolation sur des images présentant des variations importantes de contrastes en vue de l'extraction de structures tubulaires. En percolation seule une information locale ou régionale est utilisée, contrairement à une méthode globale tel que le calcul des chemins et distances géodésiques. C'est donc une méthode prometteuse pour réduire les temps de traitement dans les grandes banques d'images.

En ce qui concerne les métriques fonctionnelles d'Asplund, comme expliqué par Michel Jourlin dans son ouvrage [140], il serait intéressant de les étudier dans le cadre de la théorie des jauges.

Le lecteur intéressé par un exposé plus détaillé de ce point pourra de reporter à la référence [140, section 7.3]. Seule une présentation très succincte en est ici donnée. Soit E un espace vectoriel réel et A un sous-ensemble de E. A est dit *équilibré* si $\forall \lambda \in \mathbb{R}, |\lambda| \leq 1 \implies \lambda A \subset A$. A est dit *absorbant* si $\forall x \in E, \exists \alpha_x > 0, \forall \lambda : |\lambda| \geq \alpha_x \implies x \in \lambda A$. Un ensemble équilibré A contient nécessairement l'origine 0_E de l'espace vectoriel E et A est symétrique par rapport à 0_E. De plus, 0_E est toujours un élément d'un ensemble absorbant A. Une *jauge* peut maintenant être définie.

Etant donné un sous-ensemble A d'un espace vectoriel E, on appelle *jauge* de A l'application j_A définie sur E et à valeurs dans $[0; +\infty[$ par :

$$x \to j_A(x) = \begin{cases} \inf\{\alpha > 0 \mid x \in \alpha A\} & \text{s'il existe un } \alpha \text{ tel que } x \in \alpha A, \\ +\infty & \text{si pour tout } \alpha > 0,\ x \notin \alpha A. \end{cases}$$

Dans le cas du sous-espace $([0; M[^D, \triangle, \triangle)$ de l'espace vectoriel $([-\infty; M[^D, \triangle, \triangle)$ le sous-graphe de toute image $g > 0$, $sg(g)$, est un voisinage de 0_F, i.e. la fonction nulle. Pour une image $f \in [0; M[^D$, l'homothétique $\lambda_0 \triangle g$ qui absorbe f pour la première fois est défini par :

$$\lambda_0 = \inf\{\lambda, \forall x \in D, \lambda \triangle g(x) \geq f(x)\}.$$

Il correspond à la jauge $j_{sg(g)}(sg(f))$, où $sg(g)$ est le sous-graphe de g dans le sous-espace $[0; M[^D$. λ_0 correspond à la définition de la sonde supérieure $\lambda_0 \triangle g$ de l'image f utilisée dans le cas de la métrique d'Asplund LIP-multiplicative. Il y a donc un lien à étudier entre cette métrique et la théorie des jauges.

De plus, l'extension des cartes de distances d'Asplund binaires et fonctionnelles aux opérations morphologiques sur des graphes avec une maille irrégulière ou sur des nuages de points constituerait un nouveau champ exploratoire. Il sera en effet nécessaire de définir un voisinage sur ces graphes [34], [200] ou nuages de points [9], [75], ce qui n'est pas toujours trivial.

En outre, des métriques fonctionnelles d'Asplund avec une variation temporelle de la forme de référence pourront être définies. Pour les métriques fonctionnelles qui sont robustes aux variations d'intensité dans l'image, une des difficultés sera d'ajouter une invariance au changement d'échelle de la forme comme pour les métriques d'Asplund entre ensembles. Les nouvelles métriques d'Asplund spatio-temporelles serviront à l'analyse de formes changeant au cours du temps. En effet, avec l'utilisation de plus en plus massive de la photographie numérique, de grandes banques de données d'images ont été constituées durant ces dernières décennies dans de nombreux domaines : santé, télédétection, industrie, etc. Dans ces banques, plusieurs images de la même scène présentant un phénomène temporel ont été acquises à des périodes différentes et dans des conditions diverses. Il devient maintenant possible d'étudier l'évolution spatio-temporelle de ces phénomènes sur des périodes pouvant aller de quelques secondes à plusieurs années. E.g. dans les images du fond d'œil, il est possible de suivre l'évolution des lésions au cours du temps chez un patient grâce aux images rétiniennes acquises à une fréquence d'une par an [12], [273]. En télédétection par satellite, des séries d'images permettent de suivre l'évolution du bâti au cours du temps [62], [87], [297]. En imagerie biologique par microscopie, il devient possible d'étudier l'évolution de cellules au cours du temps [26] ou des battements cardiaques sur des embryons de poisson [173]. Dans l'industrie, un matériau mis sous contrainte subit aussi une évolution temporelle [274]. Les métriques fonctionnelles d'Asplund avec une variation temporelle de la forme de référence pourront servir à étudier ces phénomènes spatio-temporels.

Concernant le modèle LIP, deux pistes de recherches sont à explorer. La première est la définition d'un modèle LIP pour les images hyperspectrales à partir du modèle LIP couleur. La seconde est la prise en compte de phénomènes de diffusion dans le modèle LIP comme ceux existant lors de la propagation de la lumière au travers du papier [260] ou de milieux biologiques [307] comme l'épiderme. En effet, le modèle LIP, qui repose sur la loi optique des transmittances, considère une propagation rectiligne des ondes. Comme celle-ci ne tient pas compte des phénomènes de diffusion, il serait intéressant d'introduire ces phénomènes dans ce modèle.

7.2.2 Morphologie Mathématique et apprentissage statistique

Le deuxième axe des perspectives concerne la combinaison de la Morphologie Mathématique avec l'apprentissage statistique. Afin de trouver un modèle d'une forme de référence, il pourrait être intéressant d'apprendre cette forme par apprentissage statistique à partir de segmentations manuelles par exemple ou obtenues par une autre méthode de segmentation d'images. Cette forme de référence servira ensuite au calcul d'une carte des distances d'Asplund sur l'image à étudier. Par exemple en ophtalmologie, une forme modélisant des lésions, comme un AMIR (anomalie microvasculaire intra-rétinienne) ou bien un micro-anévrisme, pourrait être apprise puis utilisée pour rechercher d'autres lésions dans les images du fond d'œil.

Afin de faire de l'apprentissage statistique, on pourrait envisager d'avoir des opérateurs avec des invariances comme celles au changement de contraste et d'échelle et de venir apprendre des formes de références afin de retrouver des objets similaires dans les images. Les opérateurs pourraient être composés (i.e. mis en cascade) afin de minimiser l'erreur de classification. Utiliser des opérateurs avec des invariances fortes pourrait permettre de réduire le nombre de couches d'apprentissage (et donc de paramètres à estimer) plutôt que de chercher à apprendre ces invariances grâce à des espaces d'échelles [269] ou à de l'augmentation de données [94].

En outre, dans les grandes banques d'images de phénomènes évoluant avec le temps, se pose la question suivante : comment prédire la progression du phénomène observé sur un objet dans les images ? De manière plus précise, connaissant la classe du phénomène sur l'objet pour les instants $t-n; \dots ; t-2; t-1; t$, est-il possible de prédire sa classe pour l'instant $t+1$? La prédiction devra utiliser la comparaison entre les images du même objet. Des essais de prédiction d'évolutions temporelles ont pu être effectués dans différents domaines [12], [87] en utilisant des descripteurs purement statistiques. Cependant, ceux-ci n'ont pas d'interprétation évidente d'un point de vue théorique [239] ou par un expert du phénomène (e.g. médecin [320], biologiste, etc.). Le suivi et la prédiction de phénomènes spatio-temporels dans des images par apprentissage statistique avec des critères morphologiques ou anatomiques restent un champ de recherche à explorer. *Développer de telles méthodes permettrait de lever un verrou technologique important.* Afin de lever ce verrou, il serait donc utile de développer *des modèles qui pourront être reliés aux caractéristiques morphologiques propres des images du phénomène* telles que des lésions [164], des divisions cellulaires, le suivi d'un phénomène physique ou biologique ou encore des anomalies qui apparaissent. Les métriques d'Asplund fonctionnelles pourraient permettre de relever ce défi.

Afin de suivre et prédire un phénomène avec une signature spatiale évoluant au cours du temps, il serait intéressant de combiner des métriques de type Asplund, invariantes à certaines évolutions spatio-temporelles de la fonction de référence, avec des méthodes d'apprentissage statistiques qui tiennent compte des relations temporelles. Parmi ces approches, on peut citer : les réseaux RNN (*Recurrent Neural Network*), LSTM (*Long Short-Term Memory*), GRU (*Gated Recurrent Units*) ou *Transformer* [306] et leurs adaptations aux images : R-CNN (*Recurrent Convolutionnal Neural Network*) [93], YOLO (*You Only Look Once*) [255] ou ViViT (*A Video Vision Transformer*) [13]. Ces réseaux permettent de tenir compte des dépendances temporelles entre les variables [94]. On pourrait également envisager d'apprendre certains paramètres de déformations de la forme de référence grâce à de l'apprentissage des paramètres des métriques [24], [114].

En outre, à partir de formes détectées dans les images, est-on capable de générer une nouvelle image réaliste contenant ces formes ? Les paramètres spatio-temporels appris à partir des données lors du suivi de l'évolution du phénomène pourront servir dans la simulation d'images réalistes de stades ultérieurs du phénomène. Les distributions statistiques de paramètres permettant de générer des images du phénomène seront ainsi apprises sur les données. Ces images simulées permettront la validation du modèle lors de la prédiction. Pour évaluer la réalité de la prédiction du point de vue des caractéristiques propres du phénomène, on s'appuiera sur les travaux récents [239], [334] qui permettent l'interprétation des décisions prises par l'apprentissage profond. Comme les métriques fonctionnelles d'Asplund reposent sur le modèle *Logarithmic Image Processing* qui lui-même dépend de la loi optique des transmittances, des connaissances physiques seront introduites

dans l'apprentissage. Les applications suivantes pourront être étudiées : en santé pour la prédiction de l'évolution de la rétinopathie diabétique [97], [98], [234], [252] ou en télédétection pour suivre l'évolution du bâti au cours du temps [62].

Ces approches compléteront les méthodes existantes en géométrie stochastique [43], [44], [132], [133]. Au niveau applicatif, la prédiction de l'évolution d'un phénomène temporel dans les images à partir de photographies des états précédents de ce phénomène est un champ qui reste à explorer. En particulier, il serait intéressant de prédire l'évolution de la rétinopathie diabétique dans les images rétiniennes d'un patient connaissant ses examens des années précédentes. Il pourrait être possible de déterminer en quelle année t administrer un traitement au patient avant qu'il ne développe en année $t + 1$ une forme trop sévère. De plus, est-il possible de générer une image rétinienne réaliste indiquant l'évolution de la rétinopathie ? Peu de travaux se sont encore penchés sur le sujet [12], [58], [273]. Les méthodes développées pourraient apporter un bénéfice à la santé des patients en facilitant le diagnostic et le traitement de la maladie. Les résultats pourront ensuite être utilisés dans un projet industriel ou européen de transfert technologique vers la clinique. En outre, les méthodes développées en amélioration de contraste, segmentation et recalage pourront être appliquées à d'autres domaines en imagerie biologique, médicale ou industrielle.

Annexe : Enseignement et encadrement de projets de recherche

1 Enseignement

Depuis la fin de mes études d'ingénieur, j'ai assuré un certain nombre d'interventions en formation dans des établissements d'enseignement supérieur. Toutes mes interventions concernent les domaines des sciences de l'ingénieur et majoritairement le Traitement d'Image, les Mathématiques et les Sciences Physiques. Elles s'adressaient à des publics de niveau Licence, Master, Doctorat et des chercheurs. J'ai dispensé une bonne partie de mes cours et séminaires en langue anglaise. La liste de mon activité d'enseignement et des séminaires que j'ai donnés est présentée dans les tables 1 et 2 ci-dessous.

TABLE 1 – Activité d'enseignement. CM : Cours Magistral ; TD : Travaux Dirigés ; TP : Travaux Pratiques ; L,M,D : Licence, Master, Doctorat ; CDI : Contrat à Durée Indéterminée.

Statut	**Année**	**Durée**	**Etablissement**	**Public**	**Niveau**	**Domaine, matières responsabilités**
Intervenant	2019 Séminaire (CM)	1 h	Labo. MatéiS INSA Lyon	Doctorants, chercheurs	D	Traitement d'images acquises en milieu perturbé. Création du support.
CDI	2018 Séminaire (CM)	1 h	IPRI	Etudiants en bachelor de *University of North Florida USA*	L3 M1	*Diabetes and Retinopathy. Image analysis of the fundus of the eye.* Création du support.
Intervenant	2017 Séminaire (CM)	2 h	*University of Strathclyde* Royaume-Uni	Pr. et chercheurs		*Image analysis of eye fundus images for diabetic retinopathy screening.* Création du support.
CDI	2017 Séminaire (CM)	2 h	IPRI	Pr. et chercheurs		*Image analysis of eye fundus images for diabetic retinopathy screening.* Création du support.
Intervenant	2016 Séminaire (CM)	2 h	*Diabetic Eye Screening Wales* Cardiff Royaume-Uni	Techniciens et chercheurs		*Image analysis of eye fundus images for diabetic retinopathy screening.* Création du support.
Intervenant	2015 Séminaire (CM)	2 h	*Mac Master University* Hamilton, Ontario, Canada	Pr. et chercheurs	D	*Image analysis of eye fundus images for diabetic retinopathy screening.* Création du support.
CDI	2015 Séminaire (CM)	2 h	IPRI	Pr. et chercheurs		*Contrast enhancement of eye fundus images for diabetic retinopathy screening.* Création du support.

Suite dans la table 2 à la page suivante.

TABLE 2 – *Suite de la table 1 à la page précédente.*

Vacataire	2014 (CM)	4 h	CPE-Lyon	Cycle ingénieur Master 2	M2	Méthodes statistiques d'analyse de données, de classification et d'apprentissage pour images multivariées. Création du support.
CDI	2011-13 CM	8 h	Centre Recherche Michelin	Formation continue Ingénieurs et chercheurs		*Hyperspectral segmentation by Mathematical Morphology. SVM classification for hyperspectral images.* Création du support.
CDI	2012 CM et TD	8 h	Centre Recherche Michelin	Formation continue Ingénieurs et chercheurs		Formation traitement du signal (analyse spectrale et filtrage) pour l'Ecole de la Physique de Michelin.
CDI	2010 Tutoriel	30 h	Centre Recherche Michelin	Formation continue Chercheurs		2 tutoriels sur le traitement d'images et la morphologie mathématique 20 h (en anglais) et 10 h.
CDI	2007 Séminaire (CM)	2 h	*University of Heidelberg*, Allemagne	Equipe du Pr. Fred Hamprecht Doctorants et chercheurs	D	*Segmentation of hyperspectral images by mathematical morphology.* Création du support.
Intervenant	2007-8 4 séminaires (CM)	12 h	Hôpital Georges Pompidou, Paris	Equipe du Pr. C.A. Cuénod Doctorants et chercheurs	D	Segmentation d'images DCE-MRI du petit animal par Morphologie Mathématique hyperspectrale. Création du support.
Monitorat	2005-8 TD	48 h	Ecole des Mines de Paris	Cycle ingénieur Master 2	M2	Image : Morphologie Mathématique, Modèles aléatoires. Physique des milieux aléatoires. Dispensés pour la majorité en anglais.
Monitorat	2007-8 CM et TD	12 h	Ecole des Mines de Paris	Formation continue Ingénieurs et chercheurs		Géodésie et segmentation d'images hyperspectrales par Morphologie Mathématique.
Vacataire	2005-7 TD	44 h	CPE-Lyon	Cycle ingénieur L3, M1	L3 M1	Probabilités, Analyse Numérique pour l'image.
Assistant	2005 TP	16 h	CPE-Lyon	Cycle ingénieur Master M2	M2	Image : reproduction couleur (en anglais et français).
Vacataire	2004-8 Interrogations (colles)	122 h	Classes préparatoires à CPE-Lyon	Math Sup (L1) Math Spé (L2)	L1 L2	Physique.
	TOTAL	**318 h**				

2 Encadrements de projets de recherche, bourses et autres activités de recherche

2.1 Projets de recherche ayant donné lieu à des transferts

- A l'IPRI, entre 2014 et 2020, j'ai développé un projet de recherche en analyse d'images médicales qui a conduit à un transfert vers la clinique. J'ai créé un système d'amélioration des

images du fond d'œil pour un meilleur diagnostic de la rétinopathie diabétique. J'ai obtenu trois financements (cf. sous-section 2.2 des annexes) qui comprenaient : un projet Européen, *Eye-Light*, (50 k€) et deux financements de BPI France (5 k€ et 18 k€). L'objectif était d'étudier la faisabilité d'un service en ligne d'amélioration des images du fond d'œil, certifié comme un dispositif médical au sens de la réglementation européenne. Une étude clinique a été planifiée pour valider le service d'un point de vue clinique. Le Professeur Chris Robertson, de *University of Strathclyde*, épidémiologiste pour la santé publique, a contribué au plan d'étude statistique. Il était attendu que le service rende le diagnostic de la rétinopathie meilleur et plus rapide. Un *business plan* a aussi été préparé. Le projet a permis la publication d'un article de journal [234], de deux articles de conférences [210], [211] et de deux résumés de conférences médicales [232], [233].

- Au centre Mondial de Recherche de Michelin, de 2011 à 2014, pendant trois ans à temps plein, j'ai organisé et géré un projet de recherche sur la reconstruction 3D des pneumatiques par acquisitions multimodales (domaine du visible et des rayons X) pour étudier les performances des pneus. J'ai conçu et réalisé un système de reconstruction 3D par stéréovision pour analyser la surface et les composants des pneumatiques (cf. chapitre 5). Cette approche innovante a permis d'avoir accès à l'information à l'intérieur du pneumatique en statique ou en dynamique. Les images étaient capturées sous différentes vues afin de reconstruire la forme 3D interne des objets. En combinant de la stéréovision par photographie et par radiographie, on obtenait des images multimodales. J'ai développé le système complet : l'acquisition des images, le calibrage du dispositif et les algorithmes de reconstruction 3D. J'ai aussi travaillé sur la segmentation des coupes de pneumatiques. J'ai développé des collaborations avec l'Ecole des Mines de Paris, le Professeur Michel Jourlin de l'Université Jean Monnet (Saint-Etienne) et une Petite et Moyenne Entreprise (PME) innovante. Le montant des collaborations externe représentait des montants conséquents.
- Au centre Mondial de Recherche de Michelin, de 2008 à 2010 pendant trois ans à temps plein, j'ai organisé et géré un projet de recherche sur le contrôle industriel des pneumatiques par traitement d'image. Le système utilisait la Morphologie Mathématique mais aussi l'analyse de texture et l'intelligence artificielle pour détecter les anormalités à la surface des pneumatiques (cf. chapitre 3). Les anomalies étaient extraites en construisant des séries d'images multivariées comme la réponse à des séries de filtres appliquées à des images multimodales. La réduction de données et la classification par apprentissage statistique étaient utilisées pour extraire les objets d'intérêt en temps réel. La difficulté était de segmenter des objets texturés avec un faible contraste dans l'environnement très contraint du pneumatique. Il était également important de développer des algorithmes de segmentation rapides pour traiter efficacement ces grandes banques d'images industrielles. Le système a été testé et utilisé dans l'industrie. J'ai dû définir les objectifs, développer le réseau de collaborations scientifiques, négocier les contrats de collaboration, et gérer scientifiquement le projet. Les partenaires se composaient de l'Ecole des Mines de Paris, de PME innovantes et d'autres centres de recherche. Plusieurs chercheurs et post-doctorants de différents laboratoires ont travaillé sur le projet. Le montant des collaborations externe représentait des montants conséquents. Le projet a permis l'obtention de huit brevets [195]-[197], [213]-[215], [225], [226] obtenus dans plusieurs pays ou entités géopolitiques (Union Européenne, Etats-Unis, Japon, Chine).

2.2 Financements pour les projets

1) En 2016, j'ai obtenu et coordonné le projet européen *Eye-Light*. Il s'agissait d'un projet *SME-Instrument* du programme *Horizon2020* de l'Union Européenne. Le budget obtenu était de 50 k€ pour une durée de six mois. Le numéro de convention de subvention est le 717108 et le lien internet est le suivant `https://cordis.europa.eu/project/id/717108/fr`. Après avoir développé une nouvelle méthode d'amélioration des images du fond d'œil, l'ob-

jectif du projet était d'effectuer une étude de faisabilité pour un service d'amélioration d'image pour le diagnostic de la rétinopathie diabétique. Une attention particulière était portée aux pays en voie de développement.

Pour développer ce projet, j'ai aussi obtenu deux financements de BPI France (Banque Publique d'Investissement) :

2) en 2016, une bourse de 18 k€ (Aide au Partenariat Technologique) pour préparer l'implémentation du projet *Eye-Light* après la phase 1 ;
3) en 2015, une bourse de 5 k€ (Aide aux Programmes Européens) pour préparer la soumission du projet *Eye-Light*.

J'ai aussi soumis d'autres demandes de projets européens qui n'ont pas été retenues.

- En 2017, j'ai déposé une demande pour la phase 2 du projet *Eye-Light* (*SME-Instrument*, *Horizon2020*). L'objectif était de certifier le service comme un *dispositif médical* d'amélioration des images du fond d'œil. Une étude clinique de validation était prévue. Cette demande impliquait une collaboration avec Diabetic Eye Screening Wales, et University of Swansea, Pays de Galles, Royaume-Uni. Le budget demandé était de 4,8 M€ pour une durée 36 mois.
- En 2020, avec mon collègue le Professeur Alberto d'Onofrio, nous avons déposé un projet Marie-Sklodowska Curie *European Training Network* pour un réseau de doctorat. La demande de projet s'intitulait PHIRST : *Public Health Innovative Research On Space and Time*. Avec mon collègue, nous avons coordonné la soumission du projet. Les principaux participants étaient : *University of Strathclyde* (Royaume-Uni), IPRI (France), *University of Pisa* (Italie), *Istituto Superiore Di Sanita* (Italie), *University of Uppsala* (Suède), *University of Utrecht* (Pays-bas), *Foundation for Research and Technology Hellas* - FORTH (Grèce) et *University of Porto* (Portugal). Le budget demandé était de 3,8 M€ pour une durée de 48 mois.

2.3 Encadrement de jeunes chercheurs et d'étudiants

N'ayant pas été dans le milieu universitaire français après mon doctorat, mes missions ne comprenaient pas le co-encadrement de thèse. Par contre, au cours de mes travaux de recherche chez Michelin puis chez IPRI, j'ai pu superviser un jeune ingénieur sur un projet de recherche pendant une année et 6 étudiants en stage. Voici la liste détaillée de mes encadrements.

1) De mi-2015 à mi-2016, durant un an à l'IPRI, j'ai encadré un jeune ingénieur CPE-Lyon, Pierre Guilbert, en CDI sur un projet de recherche intitulé : « Implémentation d'algorithmes rapides en morphologie mathématique ».
2) Entre janvier et juillet et 2016 à l'IPRI, j'ai encadré pour un stage de 6 mois, Sabina Stefan, diplômée d'un Bachelor de 4 ans de *University of Cape-Town*, Afrique du Sud. Le sujet était : « *Image quality assessment and cataract detection in eye fundus images* ». Elle est ensuite partie comme étudiante de doctorat à *Brown University*, Providence, Rhode Island, Etats-Unis.

   J'ai encadré les étudiants suivants en stage de fin de Master 2 ou d'école d'ingénieur.
3) De février à août 2019, durant 6 mois à l'IPRI, Lucas Arnaldo Alves a effectué son stage de fin d'études pour Télécom Saint-Etienne et pour le Master Recherche Optique Image Vision, de l'Université Jean Monnet, Saint-Etienne. Le titre était « Deep learning pour la classification des images du fond d'œil » [14].
4) En 2014, au Centre de Recherche de Michelin, Florian Rousset a effectué pendant 6 mois son stage d'élève-ingénieur pour CPE-Lyon et de Master 2 Recherche Optique Image Vision, de l'Université Jean Monnet, Saint-Etienne. Je l'ai supervisé de début mars à fin avril 2014, jusqu'à mon changement de poste de Michelin à l'IPRI. Florian a ensuite poursuivi son stage jusqu'à son terme, en août 2014. Son sujet était « Recalage d'images de pneumatiques déformés » [265].

5) D'avril à septembre 2013, au Centre de Recherche de Michelin, j'ai supervisé Benoît Morel durant 6 mois pour son stage de fin d'étude de CPE-Lyon et du Master Recherche Electronique Electrotechnique Automatique Procédés, Systèmes et Images de l'INSA de Lyon. Son sujet était « Rehaussement de contraste en imagerie Rayons-X des pneumatiques » [198].
6) De mars à septembre 2011, au Centre de Recherche de Michelin, j'ai co-encadré Thibaud Vilboux pour son stage de fin d'étude de Télécom Saint-Etienne. Le sujet était « Stéréovision » avec une application au domaine des pneumatiques [312].
7) En 2012, j'ai participé à la supervision de Cédric Ruivo pour un stage d'un an en 4ème année de l'école d'ingénieur CPE-Lyon. Le sujet était « Estimation de la perte de gomme de pneus en utilisant des images 3D ».

2.4 Membre du comité de sélection du prix de la meilleure thèse lors d'une conférence internationale

En 2019, j'ai fait partie du comité de sélection du prix de la meilleure thèse de l'*International Society of Stereology*. Ce prix a été remis lors du *15th International Congress for Stereology and Image Analysis* qui s'est déroulé à Aarhus, au Danemark, du 27 au 30 mai 2019. A cette occasion, j'ai eu deux thèses à évaluer.

2.5 Prix

En 2007, j'ai reçu le prix du « *Young Stereologist* » remis par l'*International Society of Stereology* lors du *12th International Congress of Stereology* qui s'est déroulé à Saint-Etienne en France du 30 août au 7 septembre 2007.

2.6 Relecteur dans des revues à comité de lecture

Au 8 mars 2021, j'avais effectué 40 relectures dans des revues scientifiques internationales à comité de lecture. La liste des journaux est donnée ci-après. Elle est aussi accessible à l'adresse internet suivante `https://publons.com/a/1485666/`.

- *Physics in Medicine and Biology*;
- *Pattern Recognition Letters*;
- *International Journal of Remote Sensing*;
- *Image Analysis and Stereology*;
- *Journal of Optics*;
- *Measurement Science and Technology*;
- *Journal of Modern Optics*;
- *Sensors*;
- *Mathematical Morphology - Theory and Applications*;
- *Biomedical Physics and Engineering Express*.

2.7 Relecteur pour des conférences internationales

- DGMM 2021 - *IAPR International Conference on Discrete Geometry and Mathematical Morphology*, 24-27 mai 2021, *Uppsala University*, Suède – 2 articles relus.
- IEEE ISBI 2021 - *International Symposium on Biomedical Imaging*, 13-16 avril 2021, conférence virtuelle – 6 articles relus.
- ICIAR 2020 - *17th International Conference on Image Analysis and Recognition*, 24-26 juin 2020, Conférence virtuelle à cause de la Covid-19, Portugal – 3 articles relus.

- ISMM 2019 - *International Symposium on Mathematical Morphology*, Saarbrück, Allemagne, 8-10 juillet 2019 – 3 articles relus.

2.8 Membre du comité de programme de conférences internationales

- ICIAR 2020 - *17th International Conference on Image Analysis and Recognition*, 24-26 juin 2020, Conférence virtuelle à cause de la Covid-19, Portugal.
- ISMM 2019 - *International Symposium on Mathematical Morphology*, Saarbrück, Allemagne, 8-10 juillet 2019.

2.9 Président de session dans des conférences

J'ai été **président de session dans une conférence internationale** :

- CIARP 2016 - *21st IberoAmerican Congress on Pattern Recognition (IAPR conference)*, Lima, Pérou, 8-11 novembre 2016.

J'ai été **président de session dans une conférence nationale** :

- ISS 2020 - 43e réunion de l'*International Society of Stereology* France, Ecole des Mines de Paris, France, 6 février 2020.

2.10 Membre de comités techniques

Je suis membre des comités techniques suivants :

- *IAPR (International Association for Pattern Recognition) - Technical Committee 18 : Discrete Geometry and Mathematical Morphology* ;
- GdR IM : Groupe de recherche en Informatique Mathématique du Centre National de la Recherche Scientifique (CNRS) ;
- GT GDMM groupe de travail en géométrie discrète et morphologie mathématique des GdR IM et IGRV (Informatique Géométrique et Graphique, Réalité Virtuelle et Visualisation).

2.11 Membre de sociétés scientifiques

Je suis membre des sociétés scientifiques suivantes :

- *IEEE (Institute of Electrical and Electronics Engineers)* ;
- *IEEE Signal Processing Society* ;
- *IEEE Computer Society* ;
- *IAPR (International Association for Pattern Recognition)* et AFRIF (Association Française pour la Reconnaissance et l'Interprétation des Formes) ;
- *ISSIA : International Society for Stereology & Image Analysis*.

2.12 Développement de logiciels

J'ai développé les logiciels listés ci-dessous :

- Programmation d'une boîte à outils *MATLAB*® de traitement d'image avec des fonctionnalités dans les sous-domaines suivants :
  - Morphologie Mathématique ;
  - *Logarithmic Image Processing* pour des images à niveaux de gris et couleur ;
  - métriques fonctionnelles d'Asplund pour des images à niveaux de gris, couleur et hyperspectrales ;

  - reconstruction 3D par stéréovision ;
  - calibrage de caméras.

  Création de la documentation. Programmation des tests unitaires, d'intégration et de validation. Utilisation de cette boîte à outils dans toutes mes publications depuis 2014.

- Programmation en C++ et OpenMP de méthodes de morphologie mathématique, des métriques d'Asplund présentées dans l'article [209] et de seuillage de Köhler [207]. Encadrement de Pierre Guilbert (ingénieur-chercheur) pour programmer :
  - des opérations morphologiques de dilatation et d'érosions rapides avec des éléments structurants en formes de disque, ligne, rectangle, etc. ;
  - la méthode de seuillage de Köhler ;
  - la documentation avec Doxygen ;
  - les tests unitaires, d'intégration et de validation.
- Encadrement de Lucas Arnaldo-Alves, stagiaire ingénieur/M2, pour la programmation en Python avec *TensorFlow* et *Keras* d'une méthode d'apprentissage profond pour la classification automatique des images du fond d'œil des patients diabétiques.
- Programmation *MATLAB*® d'une méthode de recalage d'images du fond d'œil :
  - création de la méthode ;
  - programmation des différentes versions ;
  - publications des résultats dans les articles [210], [234].
- Programmation *MATLAB*® d'une méthode d'amélioration d'images du fond d'œil :
  - création de l'algorithme ;
  - dépôt de la méthode à l'INPI (Institut National de la Propriété Intellectuelle) en enveloppe Soleau ;
  - programmation et test des différentes versions ;
  - publications des résultats dans les articles [232]-[234].
- Programmation *MATLAB*® d'une méthode de détection des vaisseaux dans les images du fond d'œil :
  - création de la méthode ;
  - programmation des différentes versions ;
  - publications des résultats dans l'article [211].
- Programmation *MATLAB*® des tests unitaires et d'intégration pour une boîte à outils Michelin sur la simulation de pneumatiques.
- Programmation *MATLAB*® et Python des méthodes de segmentations d'objets en surface des pneumatiques. Dépôt des brevets [214], [215] pour protéger la méthode.
- Organisation d'une collaboration avec l'Ecole des Mines de Paris pour une méthode et un programme en C++ et Python sur l'analyse de texture et la segmentation rapide d'anomalies dans les images de pneumatiques. L'objectif était de créer une méthode automatique de contrôle industriel. Développement et validation de la méthode en collaboration avec l'Ecole des Mines. Dépôt des brevets [213], [225], [226] pour protéger la méthode.
- Programmation en R et Python d'une méthode d'analyse morphologique des images hyperspectrales durant ma thèse. Création d'un module de Travaux Pratiques (TP) avec ces programmes pour de la formation continue à destination des industriels.
- Programmation MySQL pour créer et gérer des bases de données d'images industrielles (chez Michelin).

3 Responsabilités collectives et activités administratives

Depuis ma thèse, j'ai eu différentes responsabilités collectives et activités administratives qui sont listées ci-après de manière succincte.

- A l'*International Prevention Research Institute*, j'ai été responsable de l'axe de recherche analyse d'images médicales, durant 10 % de mon temps.
- Pour l'IPRI et *University of Strathclyde*, j'ai soumis et coordonné des projets nationaux et européens, durant 30 % de mon temps. J'ai obtenu un projet européen (*Eye-Light* – 50 k€) et deux financements de BPI France (5 k€ et 18 k€).
- J'ai participé au recrutement de stagiaires et de jeunes chercheurs à l'IPRI durant 2 % de mon temps.
- J'ai été responsable de la définition des axes de recherche en traitement d'images pour le groupe Michelin en lien avec les responsables Michelin des collaborations scientifiques, durant 5 % de mon temps.
- J'ai apporté mon expertise scientifique et technique sur plusieurs projets pour le groupe Michelin, durant 20 % de mon temps.
- Au centre de recherche de Michelin, j'ai été responsable de plusieurs collaborations scientifiques externes avec des partenaires académiques, durant 20 % de mon temps.
- Pour Michelin, j'ai été formateur en traitement d'image et de signal, durant 2 % de mon temps.
- Au centre de recherche de Michelin, j'ai été responsable de projets de recherche et développement, durant 5 % de mon temps.

Liste de publications

Ma liste de publications est disponible sur les sites internet suivants :

- HAL : https://cv.archives-ouvertes.fr/guillaume-noyel;
- Orcid : https://orcid.org/0000-0002-7374-548X;
- Google Scholar : https://scholar.google.fr/citations?user=8RJLn60AAAAJ&hl=en.

1 Revues internationales à comité de lecture

[R1] G. NOYEL et M. JOURLIN, « Functional Asplund Metrics for Pattern Matching, Robust to Variable Lighting Conditions », *Image Analysis & Stereology*, t. 39, n° 2, p. 53-71, 2020, ISSN : 1854-5165. DOI : 10.5566/ias.2292.

[R2] ——, « Region homogeneity in the logarithmic image processing framework : application to region growing algorithms », *Image Analysis & Stereology*, t. 38, n° 1, p. 43-52, 2019, ISSN : 1854-5165. DOI : 10.5566/ias.2038.

[R3] G. NOYEL, R. THOMAS, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Superimposition of eye fundus images for longitudinal analysis from large public health databases », *Biomedical Physics & Engineering Express*, t. 3, n° 4, p. 045 015, juill. 2017. DOI : 10.1088/2057-1976/aa7d16.

[R4] G. NOYEL, J. ANGULO, D. JEULIN, D. BALVAY et C.-A. CUENOD, « Multivariate mathematical morphology for DCE-MRI image analysis in angiogenesis studies », *Image Analysis & Stereology*, t. 34, n° 1, p. 1-25, 2015, ISSN : 1854-5165. DOI : 10.5566/ias.1109.

[R5] G. NOYEL, J. ANGULO et D. JEULIN, « Fast computation of all pairs of geodesic distances », *Image Analysis & Stereology*, t. 30, n° 2, p. 101-109, 2011, ISSN : 1854-5165. DOI : 10.5566/ias.v30.p101-109.

[R6] ——, « A new spatio-spectral morphological segmentation for multi-spectral remote-sensing images », *International Journal of Remote Sensing*, t. 31, n° 22, p. 5895-5920, 2010. DOI : 10.1080/01431161.2010.512314.

[R7] ——, « Morphological segmentation of hyperspectral images », *Image Analysis & Stereology*, t. 26, n° 3, p. 101-109, 2007, ISSN : 1854-5165. DOI : 10.5566/ias.v26.p101-109.

2 Chapitre d'ouvrage

Ouvrage présentant les meilleures publications de congrès internationaux.

[R8] G. NOYEL, J. ANGULO et D. JEULIN, « Regionalized Random Germs by a Classification for Probabilistic Watershed Application : Angiogenesis Imaging Segmentation », in *Progress in Industrial Mathematics at ECMI 2008*, A. D. FITT, J. NORBURY, H. OCKENDON et E. WILSON, éd., sér. Mathematics in Industry. Berlin, Heidelberg : Springer, 2010, p. 211-216, ISBN : 978-3-642-12110-4. DOI : 10.1007/978-3-642-12110-4_29.

3 Brevets

8 brevets publiés et acceptés sur l'analyse de texture pour le traitement d'image du pneumatique (morphologie mathématique, analyse de données, apprentissage statistique).

[R9] G. NOYEL, D. JEULIN, E. PARRA-DENIS et M. BILODEAU, « Improved method of checking the appearance of the surface of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2013045593 (A1), aussi publié comme [**USA**] US9189841 (B2), [**France**] FR2980735 (A1), FR2980735 (B1), [**Europe**] EP2761587 (A1), [**Chine**] CN103843034 (A), avr. 2013. En ligne : `https://patentscope.wipo.int/search/en/WO2013045593`.

[R10] ——, « Method of fast analysis of the relief elements featuring on the internal surface of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2013045594 (A1), aussi publié comme [**USA**] US9196034 (B2), [**France**] FR2980896 (A1), FR2980896 (B1), [**Europe**] EP2761588 (A1), EP2761588 (B1), [**Chine**] CN103858147 (A), CN103858147 (B), avr. 2013. En ligne : `https://patentscope.wipo.int/search/en/WO2013045594`.

[R11] G. NOYEL et al., « Method for determining elements in relief on the surface of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2012156260 (A1), aussi publié comme [**USA**] US2015043780 (A1), US9002130 (B2), [**France**] FR2975523 (A1), FR2975523 (B1), [**Europe**] EP2710555 (A1), EP2710555 (B1), [**Chine**] CN103649995 (A), CN103649995 (B), nov. 2012. En ligne : `https://patentscope.wipo.int/search/en/WO2012156260`.

[R12] ——, « Method for determining markings in relief on the outer surface of the sidewall of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2012156262 (A1), aussi publié comme [**USA**] US2014219541 (A1), US9025853 (B2), [**France**] FR2975524 (A1), FR2975524 (B1), [**Europe**] EP2710552 (A1), EP2710552 (B1), [**Chine**] CN103649991 (A), CN103649991 (B), nov. 2012. En ligne : `https://patentscope.wipo.int/search/en/WO2012156262`.

[R13] C. MOREAU, G. NOYEL, Y. SHEN et J.-P. ZANELLA, « Analysis of the digital image of the internal surface of a tyre and processing of false measurement points », [Org. Mondiale Propriété Intellectuelle] WO2012143199 (A1), aussi publié comme [**USA**] US2014270390 (A1); US9230337 (B2); [**France**] FR2974220 (A1); FR2974220 (B1); EP2700052 (A1); [**Europe**] EP2700052 (B1); [**Japon**] JP2014513832 (A); JP5876139 (B2); [**Chine**] CN103493096 (A); CN103493096 (B); [**Brésil**] BR112013024108 (A2), oct. 2012. En ligne : `https://patentscope.wipo.int/search/en/WO2012143199`.

[R14] ——, « Analysis of the digital image of the external surface of a tyre and processing of false measurement points », [Org. Mondiale Propriété Intellectuelle] WO2012143198 (A1), aussi publié comme [**USA**] US2014307941 (A1); US9230318 (B2); [**France**] FR2974219(A1); [**Europe**] EP2700050 (A1); [**Japon**] JP2014517265 (A); [**Chine**] CN103649989 (A); CN103649989 (B); [**Brésil**] BR112013024106 (A2), oct. 2012. En ligne : `https://patentscope.wipo.int/search/en/WO2012143198`.

[R15] ——, « Analysis of the digital image of the surface of a tyre and processing of non-measurement points », [Org. Mondiale Propriété Intellectuelle] WO2012143197 (A1), aussi publié comme [**USA**] US2014185883 (A1); US9224198 (B2); [**France**] FR2974218 (A1); [**Europe**] EP2700051 (A1); [**Japon**] JP2014513793 (A); [**Chine**] CN103493095 (A); [**Brésil**] BR112013023977 (A2), oct. 2012. En ligne : `https://patentscope.wipo.int/search/en/WO2012143197`.

[R16] G. NOYEL, « Method of Monitoring the Appearance of the Surface of a Tire », [Org. Mondiale Propriété Intellectuelle] WO2011131410 (A1), aussi publié comme [**USA**] US2013129182 (A1), US9002093 (B2), [**France**] FR2959046 (A1), FR2959046 (B1), [**Europe**] EP2561479 (A1), [**Japon**] JP2013532315 (A), JP5779232 (B2), [**Chine**] CN102844791 (A), CN102844791 (B), [**Brésil**] BR112012025402 (A2), oct. 2011. En ligne : https://patentscope.wipo.int/search/en/WO2011131410.

4 Conférences avec actes

Communications dans des conférences internationales à comité de lecture et avec actes.

[R17] G. NOYEL, C. VARTIN, P. BOYLE et L. KODJIKIAN, « Retinal Vessel Segmentation by Probing Adaptive to Lighting Variations », in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Iowa City, United States : IEEE, avr. 2020, p. 1246-1249. DOI : 10.1109/ISBI45749.2020.9098332.

[R18] G. NOYEL, R. THOMAS, S. ILES, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Registration Of Retinal Images From Public Health By Minimising An Error Between Vessels Using An Affine Model With Radial Distortions », in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Venice, Italy : IEEE, avr. 2019, p. 561-564. DOI : 10.1109/ISBI.2019.8759415.

[R19] G. NOYEL, « A Link Between the Multiplicative and Additive Functional Asplund's Metrics », in *14th International Symposium on Mathematical Morphology (ISMM 2019)*, B. BURGETH, A. KLEEFELD, B. NAEGEL, N. PASSAT et B. PERRET, éd., Saarland University, Saarbrücken, Germany : Springer International Publishing, juill. 2019, p. 41-53, ISBN : 978-3-030-20867-7. DOI : 10.1007/978-3-030-20867-7_4.

[R20] ——, « Logarithmic Mathematical Morphology : A New Framework Adaptive to Illumination Changes », in *23rd Iberoamerican Congress on Pattern Recognition (CIARP 2018)*, R. VERA-RODRIGUEZ, J. FIERREZ et A. MORALES, éd., sér. Lecture Notes in Computer Science, t. 11401, Madrid, Spain : Springer International Publishing, nov. 2019, p. 453-461, ISBN : 978-3-030-13469-3. DOI : 10.1007/978-3-030-13469-3_53.

[R21] M. JOURLIN et G. NOYEL, « Homogeneity of a region in the logarithmic image processing framework : application to region growing algorithms », in *Physics and Mechanics of Random Structures : from Morphology to Material Properties*, F. WILLOT et S. FOREST, éd., sér. Physics and Mechanics of Random Structures : from Morphology to Material Properties, Ile d'Oléron, France : Presse des Mines, juin 2018, p. 5-7. En ligne : https://hal.archives-ouvertes.fr/hal-01822522.

[R22] G. NOYEL, « Speeding up the Köhler's method of contrast thresholding », in *2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, éd., IEEE, Beijing, China : IEEE, sept. 2017, p. 320-324. DOI : 10.1109/ICIP.2017.8296295.

[R23] G. NOYEL et M. JOURLIN, « Double-Sided Probing by Map of Asplund's Distances Using Logarithmic Image Processing in the Framework of Mathematical Morphology », in *13th International Symposium on Mathematical Morphology (ISMM 2017)*, J. ANGULO, S. VELASCO-FORERO et F. MEYER, éd., sér. Lecture Notes in Computer Science, t. 10225, Fontainebleau, France : Springer International Publishing, mai 2017, p. 408-420, ISBN : 978-3-319-57240-6. DOI : 10.1007/978-3-319-57240-6_33.

[R24] ——, « Spatio-Colour Asplund's Metric and Logarithmic Image Processing for Colour Images (LIPC) », in *21st Iberoamerican Congress on Pattern Recognition (CIARP 2016)*, C. BELTRÁN-CASTAÑÓN, I. NYSTRÖM et F. FAMILI, éd., sér. Lecture Notes in Computer Science, t. 10125, Lima, Peru : Springer International Publishing, nov. 2017, p. 36-43, ISBN : 978-3-319-52277-7. DOI : 10.1007/978-3-319-52277-7_5.

[R25] G. NOYEL et M. JOURLIN, « Asplund's metric defined in the logarithmic image processing (LIP) framework for colour and multivariate images », in *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, éd., Québec City, Canada : IEEE, sept. 2015, p. 3921-3925. DOI : 10.1109/ICIP.2015.7351540.

[R26] G. NOYEL, J. ANGULO et D. JEULIN, « Classification-driven stochastic watershed. Application to multispectral segmentation », in *4th European Conference on Colour in Graphics, Imaging, and Vision and 10th International Symposium on Multispectral Colour Science (CGIV 2008/MCS 2008)*, Terrassa - Barcelona, Spain : Society for Imaging Science and Technology, juin 2008, p. 471-476. En ligne : www.ingentaconnect.com/content/ist/cgiv/2008/00002008/00000001/art00101.

[R27] G. NOYEL, J. ANGULO, D. JEULIN, D. BALVAY et C. CUENOD, « Filtering, segmentation and region classification by hyperspectral mathematical morphology of DCE-MRI series for angiogenesis imaging », in *2008 5th IEEE International Symposium on Biomedical Imaging (ISBI) : From Nano to Macro*, Paris, France : IEEE, mai 2008, p. 1517-1520. DOI : 10.1109/ISBI.2008.4541297.

[R28] G. NOYEL, J. ANGULO et D. JEULIN, « On distances, paths and connections for hyperspectral image segmentation », in *8th International Symposium on Mathematical Morphology (ISMM 2007)*, G. J. F. BANON, J. BARRERA, U. d. M. BRAGA-NETO et N. S. T. HIRATA, éd., ISBN 978-85-17-00035-5, t. 1, Rio de Janeiro, Brazil : Instituto Nacional de Pesquisas Espaciais (INPE), oct. 2007, p. 399-410. En ligne : http://urlib.net/dpi.inpe.br/ismm@80/2007/03.05.11.23.

[R29] G. NOYEL, J. ANGULO et D. JEULIN, « Random Germs and Stochastic Watershed for Unsupervised Multispectral Image Segmentation », in *11th International Conference Knowledge-Based Intelligent Information and Engineering Systems (KES 2007)*, B. APOLLONI, R. J. HOWLETT et L. JAIN, éd., sér. Lecture Notes in Computer Science, t. 4694, Vietri sul Mare, Salerno, Italy : Springer, sept. 2007, p. 17-24, ISBN : 978-3-540-74829-8. DOI : 10.1007/978-3-540-74829-8_3.

[R30] G. NOYEL, J. ANGULO et D. JEULIN, « Morphological segmentation of hyperspectral images - ICSXII », in *12th International Congress for Stereology (ICS XII)*, Université Jean Monnet, Saint-Etienne, France, août 2007. En ligne : https://hal.archives-ouvertes.fr/hal-02958884.

5 Conférences avec résumé

Résumés et posters dans des conférences internationales avec comité de lecture.

[R31] G. NOYEL, M. JOURLIN, M. SMANS, R. THOMAS, S. ILES, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Enhancing Eye Fundus Images for Diabetic Retinopathy Screening », in *American Diabetes Association 77th Scientific Sessions - ADA2017*, American Diabetes Association, San Diego, CA, United States, juin 2017. En ligne : https://hal.archives-ouvertes.fr/hal-03162202.

[R32] G. NOYEL, M. JOURLIN, R. THOMAS, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Contrast enhancement of eye fundus images », in *IDF2015 - World Diabetes Congress*, International Diabetes Federation, Vancouver, Canada, nov. 2015. En ligne : https://hal.archives-ouvertes.fr/hal-03162208.

6 Conférences sans actes

Communications dans des conférences internationales à comité de lecture sans actes.

[R33] G. NOYEL et M. JOURLIN, « A simple expression for the map of Asplund's distances with the multiplicative Logarithmic Image Processing (LIP) law », in *12th European Congress for Stereology and Image Analysis 2017 (ECSIA 2017)*, University of Kaiserslautern and Fraunhofer-Institut für Techno- und Wirtschaftsmathematik (ITWM), Kaiserslautern, Germany, sept. 2017. En ligne : `https://arxiv.org/abs/1708.08992`.

[R34] G. NOYEL, J. ANGULO, D. JEULIN, D. BALVAY et C.-A. CUÉNOD, « Hyperspectral mathematical morphology applied to DCE-MRI series for angiogenesis imaging », in *15th European Conference on Mathematics for Industry (ECMI'2008)*, University College London, United Kingdom, juill. 2008. En ligne : `https://hal-mines-paristech.archives-ouvertes.fr/hal-00904082`.

7 Thèse

[R35] G. NOYEL, « Filtrage, réduction de dimension, classification et segmentation morphologique hyperspectrale », Sous la direction de Dominique Jeulin et Jesus Angulo, thèse de doct., École Nationale Supérieure des Mines de Paris, Fontainebleau, France, sept. 2008. En ligne : `https://pastel.archives-ouvertes.fr/pastel-00004473`.

8 Conférences nationales

Communications dans des conférences nationales sans actes.

[R36] G. NOYEL, « Logarithmic Mathematical Morphology : a new framework robust to illumination changes », in *43e réunion de l'International Society of Stereology - France*, Ecole des Mines de Paris, France, fév. 2020.

[R37] ——, « Linking the map of Asplund's distances with Mathematical Morphology », in *42e réunion de l'International Society of Stereology - France*, Ecole des Mines de Paris, France, fév. 2019.

[R38] ——, « Aspect des pneumatiques », in *Conférence Sensomines*, Ecole des Mines de Paris, France, fév. 2010.

[R39] G. NOYEL, J. ANGULO, D. JEULIN, D. BALVAY et C.-A. CUENOD, « Filtrage, segmentation et classification en régions, par morphologie mathématique hyperspectrale, de séries DCE-MRI en imagerie angiogénique », in *31e réunion de l'International Society of Stereology - France*, Ecole des Mines de Paris, France, fév. 2008.

[R40] G. NOYEL, J. ANGULO et D. JEULIN, « Segmentation d'images hyperspectrales par morphologie mathématique », in *GDR–ISIS Analyse d'Images Multivariées*, Ecole Nationale Supérieure des Télécommunications de Paris, France, nov. 2007.

[R41] ——, « Segmentation d'images hyperspectrales par morphologie mathématique », in *30e réunion de l'International Society of Stereology - France*, Ecole des Mines de Paris, Paris, France, fév. 2007.

[R42] ——, « Segmentation d'images hyperspectrales par morphologie mathématique », in *Ecole d'Hiver d'Imagerie Numérique Couleur*, Poitiers, France, jan. 2007.

Bibliographie

[1] M. ABRAMOFF, M. NIEMEIJER, S. LEE et J. REINHARDT, « Optimal registration of multiple deformed images using a physical model of the imaging distortion », US8194936 (B2), juin 2012. En ligne : www.google.fr/patents/US8194936.

[2] M. D. ABRÀMOFF, J. C. FOLK, D. P. HAN, J. D. WALKER, D. F. WILLIAMS, S. R. RUSSELL, P. MASSIN, B. COCHENER, P. GAIN, L. TANG, M. LAMARD, D. C. MOGA, G. QUELLEC et M. NIEMEIJER, « Automated Analysis of Retinal Images for Detection of Referable Diabetic Retinopathy », *JAMA Ophthalmology*, t. 131, n° 3, p. 351-357, mars 2013, ISSN : 2168-6165. DOI : 10.1001/jamaophthalmol.2013.1743.

[3] M. D. ABRÀMOFF, P. T. LAVIN, M. BIRCH, N. SHAH et J. C. FOLK, « Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices », *Nature Partner Journals Digital Medicine*, t. 1, n° 1, p. 39, déc. 2018, ISSN : 2398-6352. DOI : 10.1038/s41746-018-0040-6.

[4] M. D. ABRÀMOFF, J. M. REINHARDT, S. R. RUSSELL, J. C. FOLK, V. B. MAHAJAN, M. NIEMEIJER et G. QUELLEC, « Automated Early Detection of Diabetic Retinopathy », *Ophthalmology*, t. 117, n° 6, p. 1147-1154, juin 2010, ISSN : 01616420. DOI : 10.1016/j.ophtha.2010.03.046.

[5] M. D. ABRÀMOFF, Y. LOU, A. ERGINAY, W. CLARIDA, R. AMELON, J. C. FOLK et M. NIEMEIJER, « Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning », *Investigative Ophthalmology & Visual Science*, t. 57, n° 13, p. 5200-5206, oct. 2016, ISSN : 1552-5783. DOI : 10.1167/iovs.16-19964.

[6] K. M. ADAL, R. M. ENSING, R. COUVERT, P. van ETTEN, J. P. MARTINEZ, K. A. VERMEER et L. J. van VLIET, « A Hierarchical Coarse-to-Fine Approach for Fundus Image Registration », in *Biomedical Image Registration*, S. OURSELIN et M. MODAT, éd., Cham : Springer International Publishing, 2014, p. 93-102, ISBN : 978-3-319-08554-8. DOI : 10.1007/978-3-319-08554-8_10.

[7] R. ADAMS, « Radial Decomposition of Disks and Spheres », *CVGIP : Graphical Models and Image Processing*, t. 55, n° 5, p. 325-332, 1993, ISSN : 1049-9652. DOI : 10.1006/cgip.1993.1024.

[8] J. ANGULO et D. JEULIN, « Stochastic watershed segmentation », in *8th International Symposium on Mathematical Morphology*, G. BANON et al., éd., t. 1, Instituto Nacional de Pesquisas Espaciais (INPE), 2007, p. 265-276. En ligne : http://urlib.net/dpi.inpe.br/ismm@80/2007/07.16.14.39.

[9] J. ANGULO, « Morphological Scale-Space Operators for Images Supported on Point Clouds », in *Scale Space and Variational Methods in Computer Vision*, J.-F. AUJOL, M. NIKOLOVA et N. PAPADAKIS, éd., Cham : Springer International Publishing, 2015, p. 78-89, ISBN : 978-3-319-18461-6. DOI : 10.1007/978-3-319-18461-6_7.

[10] J. ANGULO, « Morphologie mathématique et indexation d'images couleur : application à la microscopie en biomédecine », Theses, École Nationale Supérieure des Mines de Paris, déc. 2003. En ligne : https://pastel.archives-ouvertes.fr/tel-00007524.

[11] J. ANGULO et S. VELASCO-FORERO, « Semi-supervised hyperspectral image segmentation using regionalized stochastic watershed », in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVI*, S. S. SHEN et P. E. LEWIS, éd., International Society for Optics et Photonics, t. 7695, SPIE, 2010, p. 483-494. DOI : 10.1117/12.850187.

[12] F. ARCADU, F. BENMANSOUR, A. MAUNZ, J. WILLIS, Z. HASKOVA et M. PRUNOTTO, « Deep learning algorithm predicts diabetic retinopathy progression in individual patients », *Nature Partner Journals Digital Medicine*, t. 2, n° 1, p. 92, déc. 2019, ISSN : 2398-6352. DOI : 10.1038/s41746-019-0172-3.

[13] A. ARNAB, M. DEHGHANI, G. HEIGOLD, C. SUN, M. LUCIC et C. SCHMID, « ViViT : A Video Vision Transformer », *CoRR*, t. abs/2103.15691, 2021. arXiv : 2103.15691. En ligne : https://arxiv.org/abs/2103.15691.

[14] L. ARNALDO ALVES, « Deep learning pour la classification des images du fond d'œil », Rapport de Master, Télécom Saint-Etienne et Université Jean Monnet, Saint-Etienne, Master Recherche Optique Image Vision, France, 2019.

[15] E. ASPLUND, « Comparison Between Plane Symmetric Convex Bodies and Parallelograms. », *Mathematica Scandinavica*, t. 8, p. 171-180, déc. 1960, ISSN : 1903-1807. DOI : 10.7146/math.scand.a-10606.

[16] G. AZZOPARDI, N. STRISCIUGLIO, M. VENTO et N. PETKOV, « Trainable COSFIRE filters for vessel delineation with application to retinal images », *Medical Image Analysis*, t. 19, n° 1, p. 46-57, 2015, ISSN : 1361-8415. DOI : 10.1016/j.media.2014.08.002.

[17] F. V. BABALIEVSKI, « Percolation conductivity of Penrose tiling by the transfer-matrix Monte Carlo method », *Physica A : Statistical Mechanics and its Applications*, t. 182, n° 3, p. 325-330, 1992, ISSN : 0378-4371. DOI : 10.1016/0378-4371(92)90346-R.

[18] G. BANON et S. FARIA, « Morphological approach for template matching », in *Proc. X Brazilian Symp. on Comput. Graphics and Image Process.*, IEEE Comput. Soc, oct. 1997, p. 171-178, ISBN : 0-8186-8102-0. DOI : 10.1109/SIGRA.1997.625169.

[19] G. J. F. BANON et J. BARRERA, « Decomposition of mappings between complete lattices by mathematical morphology, part I : General lattices », *Signal Process.*, t. 30, n° 3, p. 299-327, fév. 1993, ISSN : 0165-1684. DOI : 10.1016/0165-1684(93)90015-3.

[20] C. BARAT, C. DUCOTTET et M. JOURLIN, « Pattern matching using morphological probing », in *Proceedings 2003 IEEE International Conference on Image Processing*, t. 1, 2003, p. I-369. DOI : 10.1109/ICIP.2003.1246975.

[21] C. BARAT, C. DUCOTTET et M. JOURLIN, « Virtual double-sided image probing : A unifying framework for non-linear grayscale pattern matching », *Pattern Recognition*, t. 43, n° 10, p. 3433-3447, 2010, ISSN : 0031-3203. DOI : 10.1016/j.patcog.2010.04.020.

[22] V. BARNETT, « The Ordering of Multivariate Data », *Journal of the Royal Statistical Society. Series A (General)*, t. 139, n° 3, p. 318-355, 1976, ISSN : 00359238. DOI : 10.2307/2344839.

[23] P. BAWANKAR, N. SHANBHAG, S. S. K., B. DHAWAN, A. PALSULE, D. KUMAR, S. CHANDEL et S. SOOD, « Sensitivity and specificity of automated analysis of single-field non-mydriatic fundus photographs by Bosch DR Algorithm—Comparison with mydriatic fundus photography (ETDRS) for screening in undiagnosed diabetic retinopathy », *PLOS ONE*, t. 12, n° 12, p. 1-13, déc. 2017. DOI : 10.1371/journal.pone.0189854.

[24] A. BELLET, A. HABRARD et M. SEBBAN, *Metric Learning*, 1. jan. 2015, t. 9, p. 1-151. DOI : 10.2200/S00626ED1V01Y201501AIM030.

[25] J. BENZÉCRI, *L'Analyse Des Données, L'Analyse des Correspondances*. Paris : Dunod, 1973, t. 2, ISBN : 204007225X.

[26] A. BERDEU, B. LAPERROUSAZ, T. BORDY, O. MANDULA, S. MORALES, X. GIDROL, N. PICOLLET-D'HAHAN et C. ALLIER, « Lens-free microscopy for 3D + time acquisitions of 3D cell culture », *Scientific Reports*, t. 8, n° 1, p. 16 135, déc. 2018, ISSN : 2045-2322. DOI : 10.1038/s41598-018-34253-6.

[27] S. BEUCHER et C. LANTUÉJOUL, « Use of watersheds in contour detection », in *Proc. Int. Workshop on image processing, real-time edge and motion detection-estimation*, 1979, p. 17-21.

[28] S. BEUCHER, « Image segmentation and mathematical morphology », Thèse, Ecole Nationale Supérieure Mines Paris, France, juin 1990. En ligne : https://pastel.archives-ouvertes.fr/tel-00108290.

[29] ——, « Transformations Résiduelles en Morphologie Numérique », Centre de Morphologie Mathématique, École des Mines de Paris, rapp. tech., 2003, p. 32. En ligne : https://people.cmm.minesparis.psl.eu/users/beucher/publi/opresid.pdf.

[30] ——, « Numerical Residues », in *Mathematical Morphology : 40 Years On*, C. RONSE, L. NAJMAN et E. DECENCIÈRE, éd., Dordrecht : Springer Netherlands, 2005, p. 23-32, ISBN : 978-1-4020-3443-5. DOI : 10.1007/1-4020-3443-1_3.

[31] ——, *Fast implementation of large erosions and dilations in Mamba*, Algorithmic description of the implementation of fast morphological operators with large structuring elements., oct. 2010. En ligne : https://hal.archives-ouvertes.fr/hal-01403935.

[32] S. BEUCHER et F. MEYER, « The morphological approach to segmentation : The watershed transformation », in *Mathematical Morphology in Image Processing*, E. DOUGHERTY, éd., sér. Optical Engineering, New York, NY, USA : Marcel Dekker, 1993, chap. 12, p. 433-481, ISBN : 9780824787240. DOI : 10.1201/9781482277234-12.

[33] M. BILODEAU, E. PARRA-DENIS et D. JEULIN, « New Operator Based on a Multi Support Point Algorithm Applied to Feature Extraction », in *International Congress for Stereology*, Beijing, China, oct. 2011. En ligne : https://hal.archives-ouvertes.fr/hal-00880286.

[34] S. BLUSSEAU, S. VELASCO-FORERO, J. ANGULO et I. BLOCH, « Tropical and Morphological Operators for Signals on Graphs », in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, p. 1198-1202. DOI : 10.1109/ICIP.2018.8451395.

[35] J. F. BONNANS, J. C. GILBERT, C. LEMARÉCHAL et C. A. SAGASTIZABAL, *Numerical Optimization*. Berlin, Heidelberg : Springer, 2006, ISBN : 978-3-540-35445-1. DOI : 10.1007/978-3-540-35447-5.

[36] N. BOURBAKI, *Intégration : Chapitres 1 à 4*. Springer Berlin Heidelberg, 2007, ISBN : 9783540353294. DOI : 10.1007/978-3-540-35329-4.

[37] J. BRAILEAN, B. SULLIVAN, C. CHEN et M. GIGER, « Evaluating the EM algorithm for image processing using a human visual fidelity criterion », in *IEEE Int. Conf. Acoustics, Speech, Signal Process.*, t. 4, avr. 1991, p. 2957-2960. DOI : 10.1109/ICASSP.1991.151023.

[38] J. E. BRESENHAM, « Algorithm for computer control of a digital plotter », *IBM Systems Journal*, t. 4, n° 1, p. 25-30, 1965. DOI : 10.1147/sj.41.0025.

[39] S. R. BROADBENT et J. M. HAMMERSLEY, « Percolation processes : I. Crystals and mazes », *Mathematical Proceedings of the Cambridge Philosophical Society*, t. 53, n° 3, p. 629-641, 1957. DOI : 10.1017/S0305004100032680.

[40] A. BRONSTEIN, M. BRONSTEIN et R. KIMMEL, « In the Rigid Kingdom », in *Numerical Geometry of Non-Rigid Shapes*. New York, NY : Springer, 2009, chap. 6, p. 119-135, ISBN : 978-0-387-73301-2. DOI : 10.1007/978-0-387-73301-2_6.

[41] H. CAI, Z. YANG, X. CAO, W. XIA et X. XU, « A New Iterative Triclass Thresholding Technique in Image Segmentation », *IEEE Trans. Image Process.*, t. 23, n° 3, p. 1038-1046, mars 2014, ISSN : 1057-7149. DOI : 10.1109/TIP.2014.2298981.

[42] A. CAN, C. V. STEWART, B. ROYSAM et H. L. TANENBAUM, « A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 24, n° 3, p. 347-364, 2002. DOI : 10.1109/34.990136.

[43] V. CAPASSO, *An Introduction to Random Currents and Their Applications*, sér. SpringerBriefs in Mathematics. Cham : Springer International Publishing, 2018, ISBN : 978-3-319-94576-7. DOI : 10.1007/978-3-319-94577-4.

[44] V. CAPASSO, A. MICHELETTI et D. MORALE, « Stochastic geometric models, and related statistical issues in tumour-induced angiogenesis », *Mathematical Biosciences*, t. 214, n° 1, p. 20-31, 2008, BICOMP 2008, ISSN : 0025-5564. DOI : 10.1016/j.mbs.2008.03.009.

[45] M. CARRE et M. JOURLIN, « LIP operators : Simulating exposure variations to perform algorithms independent of lighting conditions », in *2014 Int. Conf. on Multimedia Computing and Syst. (ICMCS)*, IEEE, avr. 2014, p. 122-126. DOI : 10.1109/ICMCS.2014.6911247.

[46] P. C. CATTIN, H. BAY, L. VAN GOOL et G. SZÉKELY, « Retina Mosaicing Using Local Features », in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, R. LARSEN, M. NIELSEN et J. SPORRING, éd., Berlin, Heidelberg : Springer, 2006, p. 185-192, ISBN : 978-3-540-44728-3. DOI : 10.1007/11866763_23.

[47] J. CHANIOT, M. MOREAUD, L. SORBIER, D. JEULIN, J.-M. BECKER et T. FOURNEL, « Heterogeneity Assessment Based on Average Variations of Morphological Tortuosity for Complex Porous Structures Characterization », *Image Analysis & Stereology*, t. 39, n° 2, p. 111-128, 2020, ISSN : 1854-5165. DOI : 10.5566/ias.2370.

[48] T. CHANWIMALUANG, G. FAN et S. R. FRANSEN, « Hybrid retinal image registration », *IEEE Trans. Inf. Technol. Biomed.*, t. 10, n° 1, p. 129-142, 2006. DOI : 10.1109/TITB.2005.856859.

[49] B. CHAPMAN, G. JOST et R. VAN DER PAS, *Using OpenMP : Portable Shared Memory Parallel Programming (Scientific and Engineering Computation)*. The MIT Press, 2007, ISBN : 9780262533027. En ligne : https://ieeexplore.ieee.org/servlet/opac?bknumber=6267237.

[50] J. CHEN, J. TIAN, N. LEE, J. ZHENG, R. T. SMITH et A. F. LAINE, « A Partial Intensity Invariant Feature Descriptor for Multimodal Retinal Image Registration », *IEEE Trans. Biomed. Eng.*, t. 57, n° 7, p. 1707-1718, 2010. DOI : 10.1109/TBME.2010.2042169.

[51] W. CHEN, M. J. ER et S. WU, « Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain », *IEEE Trans. Syst., Man, Cybern. B*, t. 36, n° 2, p. 458-466, avr. 2006, ISSN : 1083-4419. DOI : 10.1109/TSMCB.2005.857353.

[52] A. V. CIDECIYAN, « Registration of ocular fundus images : an algorithm using cross-correlation of triple invariant image descriptors », *IEEE Eng. Med. Biol. Mag.*, t. 14, n° 1, p. 52-58, 1995. DOI : 10.1109/51.340749.

[53] « CIE illuminant D65 », Colour & Vision Research Laboratory, Institute of Ophthalmology, University College London. (2006), En ligne : http://cvrl.ucl.ac.uk/ (visité le 04/12/2020).

[54] « CIE S 014-2E :2006 : Colorimetry - Part 2 : CIE Standard Illuminants for Colorimetry », CIE Standard, 2006. En ligne : http://cie.co.at/publications/colorimetry-part-2-cie-standard-illuminants-colorimetry-0.

[55] P. COCKSHOTT et K. RENFREW, *SIMD Programming Manual for Linux and Windows*. London : Springer, 2004, ISBN : 978-1-84996-920-8. DOI : 10.1007/978-1-4471-3862-4.

[56] A. CORD, F. BACH et D. JEULIN, « Texture classification by statistical learning from morphological image processing : application to metallic surfaces », *J. Microsc. (Oxford, U.K.)*, t. 239, nº 2, p. 159-166, 2010. DOI : 10.1111/j.1365-2818.2010.03365.x.

[57] A. CORD, D. JEULIN et F. BACH, « Segmentation of random textures by morphological and linear operators », in *Proceedings of ISMM 2007. 8th International Symposium on Mathematical Morphology*, G. J. F. BANON, J. BARRERA, U. d. M. BRAGA-NETO et N. S. T. HIRATA, éd., t. 1, São José dos Campos : Instituto Nacional de Pesquisas Espaciais (INPE), 2007, p. 387-398, ISBN : 978-85-17-00035-5. En ligne : http://urlib.net/rep/dpi.inpe.br/ismm@80/2007/03.26.17.28.

[58] P. COSTA, A. GALDRAN, M. I. MEYER, M. NIEMEIJER, M. ABRÀMOFF, A. M. MENDONÇA et A. CAMPILHO, « End-to-End Adversarial Retinal Image Synthesis », *IEEE Trans. Med. Imag.*, t. 37, nº 3, p. 781-791, 2018. DOI : 10.1109/TMI.2017.2759102.

[59] M. COSTER et J. CHERMANT, *Précis d'analyse d'images*, sér. CNRS plus. Presses du CNRS, 1989, ISBN : 9782876820203. En ligne : https://books.google.fr/books?id=5ufqAAAACAAJ.

[60] J. COUSTY, G. BERTRAND, L. NAJMAN et M. COUPRIE, « Watershed Cuts : Minimum Spanning Forests and the Drop of Water Principle », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 31, nº 8, p. 1362-1374, 2009. DOI : 10.1109/TPAMI.2008.173.

[61] G. CYBENKO, « Approximation by superpositions of a sigmoidal function », *Mathematics of Control, Signals, and Systems*, t. 2, nº 4, p. 303-314, déc. 1989, ISSN : 09324194. DOI : 10.1007/BF02551274.

[62] M. DAS et S. K. GHOSH, « Deep-STEP : A Deep Learning Approach for Spatiotemporal Prediction of Remote Sensing Data », *IEEE Geosci. Remote Sens. Lett.*, t. 13, nº 12, p. 1984-1988, 2016. DOI : 10.1109/LGRS.2016.2619984.

[63] E. DECENCIÈRE, G. CAZUGUEL, X. ZHANG, G. THIBAULT, J.-C. KLEIN, F. MEYER, B. MARCOTEGUI, G. QUELLEC, M. LAMARD, R. DANNO, D. ELIE, P. MASSIN, Z. VIKTOR, A. ERGINAY, B. LAŸ et A. CHABOUIS, « TeleOphta : Machine learning and image processing methods for teleophthalmology », *IRBM*, t. 34, nº 2, p. 196-203, 2013, Special issue : ANR TECSAN : Technologies for Health and Autonomy, ISSN : 1959-0318. DOI : 10.1016/j.irbm.2013.01.010.

[64] D. DEMENTHON et L. DAVIS, « Model-based object pose in 25 lines of code », *International Journal of Computer Vision*, t. 15, p. 123-141, 2005.

[65] V. DESHAYES, P. GUILBERT et M. JOURLIN, « How simulating exposure time variations in the LIP Model. Application : moving objects acquisition. », in *Acta Stereol., Proc. 14th International Conference on Stereology and Image Analysis (ICSIA)*, juill. 2015. En ligne : https://popups.uliege.be:443/0351-580X/index.php?id=3663.

[66] M. M. DEZA et E. DEZA, *Encyclopedia of Distances*. Berlin, Heidelberg : Springer, 2016, ISBN : 978-3-662-52843-3. DOI : 10.1007/978-3-662-52844-0.

[67] M. DHOME et J.-M. LAVEST, *Etalonnage et Localisation*, Cours de Master, Université Blaise Pascal, Clermont-Ferrand, 2004.

[68] « Diabetic Retinopathy Detection. Identify signs of diabetic retinopathy in eye images », Kaggle Diabetic Retinopathy competition. (2015), En ligne : www.kaggle.com/c/diabetic-retinopathy-detection.

[69] E. DIDAY, « Une nouvelle méthode en classification automatique et reconnaissance des formes la méthode des nuées dynamiques », *Revue de statistique appliquée*, t. 19, n° 2, p. 19-33, 1979. En ligne : www.numdam.org/item/RSA_1971__19_2_19_0.

[70] E. R. DOUGHERTY, « Application of the Hausdorff metric in gray-scale mathematical morphology via truncated umbrae », *J. Visual Commun. Image Representation*, t. 2, n° 2, p. 177-187, juin 1991, ISSN : 10473203. DOI : 10.1016/1047-3203(91)90007-3.

[71] R. O. DUDA, P. E. HART et D. G. STORK, *Pattern Classification*, 2e éd. New York : John Wiley & Sons, 2000, ISBN : 9780471056690. En ligne : https://books.google.fr/books?id=YoxQAAAAMAAJ.

[72] EARLY TREATMENT DIABETIC RETINOPATHY STUDY RESEARCH GROUP, « Early Photocoagulation for Diabetic Retinopathy : ETDRS Report Number 9 », *Ophthalmology*, t. 98, n° 5, Supplement, p. 766-785, 1991, ISSN : 0161-6420. DOI : 10.1016/S0161-6420(13)38011-7.

[73] ——, « Grading Diabetic Retinopathy from Stereoscopic Color Fundus Photographs—An Extension of the Modified Airlie House Classification : ETDRS Report Number 10 », *Ophthalmology*, t. 98, n° 5, Supplement, p. 786-806, 1991, ISSN : 0161-6420. DOI : 10.1016/S0161-6420(13)38012-9.

[74] M. ELAD, R. KIMMEL, D. SHAKED et R. KESHET, « Reduced complexity Retinex algorithm via the variational approach », *Journal of Visual Communication and Image Representation*, t. 14, n° 4, p. 369-388, 2003, ISSN : 1047-3203. DOI : 10.1016/S1047-3203(03)00045-2.

[75] A. ELMOATAZ, F. LOZES et H. TALBOT, « Morphological PDEs on Graphs for Image Processing on Surfaces and Point Clouds », *ISPRS International Journal of Geo-Information*, t. 5, n° 11, 2016, ISSN : 2220-9964. DOI : 10.3390/ijgi5110213.

[76] M. FAESSEL et M. BILODEAU, « SMIL : Simple Morphological Image Library », in *Séminaire Performance et Généricité, LRDE*, mars 2013. En ligne : http://cmm.ensmp.fr/~faessel/documents/2013_03_SMIL_LRDE.pdf.

[77] J. FALCOU, J.-T. LAPRESTÉ, T. CHATEAU et J. SEROT, « NT2 : Une bibliothèque haute-performance pour la vision artificielle », in *ORASIS 2007, Congrès Jeunes Chercheurs en Vision par Ordinateur*, Obernai, France, 2007, cdrom. En ligne : https://hal.archives-ouvertes.fr/hal-00167387.

[78] B. FANG et Y. Y. TANG, « Elastic registration for retinal images based on reconstructed vascular trees », *IEEE Trans. Biomed. Eng.*, t. 53, n° 6, p. 1183-1187, 2006. DOI : 10.1109/TBME.2005.863927.

[79] A. W. FITZGIBBON, « Simultaneous linear estimation of multiple view geometry and lens distortion », in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, t. 1, 2001, p. I-I. DOI : 10.1109/CVPR.2001.990465.

[80] A. D. FLEMING, K. A. GOATMAN, S. PHILIP, G. J. PRESCOTT, P. F. SHARP et J. A. OLSON, « Automated grading for diabetic retinopathy : a large-scale audit using arbitration by clinical experts », *British Journal of Ophthalmology*, t. 94, n° 12, p. 1606-1610, déc. 2010, ISSN : 0007-1161. DOI : 10.1136/bjo.2009.176784.

[81] G. L. FORESTI, C. MICHELONI, L. SNIDARO, P. REMAGNINO et T. ELLIS, « Active video-based surveillance system : the low-level image and video processing techniques needed for implementation », *IEEE Signal Process. Mag.*, t. 22, n° 2, p. 25-37, mars 2005, ISSN : 1053-5888. DOI : 10.1109/MSP.2005.1406473.

[82] W. T. FREEMAN et E. H. ADELSON, « The design and use of steerable filters », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 13, n° 9, p. 891-906, 1991. DOI : 10.1109/34.93808.

[83] R. GARGEYA et T. LENG, « Automated Identification of Diabetic Retinopathy Using Deep Learning », *Ophthalmology*, t. 124, n° 7, p. 962-969, juill. 2017, ISSN : 01616420. DOI : 10.1016/j.ophtha.2017.02.008.

[84] GDXRAY, *Database of X-ray images*, http://dmery.ing.puc.cl/index.php/material/gdxray/, Ensemble d'images de bagages n° B0063, image n° 3 ; 10 et 11. Accédé le 16 Octobre 2018, 2015.

[85] A. GEIGER, M. LAUER, C. WOJEK, C. STILLER et R. URTASUN, « 3D Traffic Scene Understanding From Movable Platforms », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 36, n° 5, p. 1012-1025, mai 2014, ISSN : 0162-8828. DOI : 10.1109/TPAMI.2013.185.

[86] T. GÉRAUD, H. TALBOT et M. VAN DROOGENBROECK, « Algorithms for Mathematical Morphology », in *Mathematical Morphology*. John Wiley & Sons, Ltd, 2013, chap. 12, p. 323-353, ISBN : 9781118600788. DOI : 10.1002/9781118600788.ch12.

[87] F. GERBER, R. DE JONG, M. E. SCHAEPMAN, G. SCHAEPMAN-STRUB et R. FURRER, « Predicting Missing Values in Spatio-Temporal Remote Sensing Data », *IEEE Trans. Geosci. Remote Sens.*, t. 56, n° 5, p. 2841-2853, 2018. DOI : 10.1109/TGRS.2017.2785240.

[88] Z. GHASSABI, J. SHANBEHZADEH, A. SEDAGHAT et E. FATEMIZADEH, « An efficient approach for robust multimodal retinal image registration based on UR-SIFT features and PIIFD descriptors », *EURASIP Journal on Image and Video Processing*, t. 2013, n° 1, p. 25, déc. 2013, ISSN : 1687-5281. DOI : 10.1186/1687-5281-2013-25.

[89] J. GIL et M. WERMAN, « Computing 2-D min, median, and max filters », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 15, n° 5, p. 504-507, 1993. DOI : 10.1109/34.211471.

[90] J. Y. GIL et R. KIMMEL, « Efficient dilation, erosion, opening, and closing algorithms », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 24, n° 12, p. 1606-1617, 2002. DOI : 10.1109/TPAMI.2002.1114852.

[91] L. GILLIBERT, C. PEYREGA, D. JEULIN, V. GUIPONT et M. JEANDIN, « 3D multiscale segmentation and morphological analysis of x-ray microtomography from cold-sprayed coatings », *Journal of Microscopy*, t. 248, n° 2, p. 187-199, 2012. DOI : 10.1111/j.1365-2818.2012.03655.x.

[92] L. GILLIBERT et D. JEULIN, « Stochastic Multiscale Segmentation Constrained by Image Content », in *Mathematical Morphology and Its Applications to Image and Signal Processing*, P. SOILLE, M. PESARESI et G. K. OUZOUNIS, éd., Springer Berlin Heidelberg, 2011, p. 132-142, ISBN : 978-3-642-21569-8. DOI : 10.1007/978-3-642-21569-8_12.

[93] R. GIRSHICK, J. DONAHUE, T. DARRELL et J. MALIK, « Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation », in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, p. 580-587. DOI : 10.1109/CVPR.2014.81.

[94] I. GOODFELLOW, Y. BENGIO et A. COURVILLE, *Deep Learning*. MIT Press, 2016, www.deeplearningbook.org.

[95] I. GOODFELLOW, D. WARDE-FARLEY, M. MIRZA, A. COURVILLE et Y. BENGIO, « Maxout Networks », in *Proceedings of the 30th International Conference on Machine Learning*, S. DASGUPTA et D. MCALLESTER, éd., sér. Proceedings of Machine Learning Research, t. 28, Atlanta, Georgia, USA : PMLR, juin 2013, p. 1319-1327. En ligne : http://proceedings.mlr.press/v28/goodfellow13.html.

[96] B. GRÜNBAUM, « Measures of symmetry for convex sets », in *Convexity, Proceedings of Symposia in Pure Mathematics*, t. 7, 1963, p. 233-270. DOI : 10.1090/pspum/007.

[97] A. GRZYBOWSKI, P. BRONA, G. LIM, P. RUAMVIBOONSUK, G. S. W. TAN, M. ABRAMOFF et D. S. W. TING, « Artificial intelligence for diabetic retinopathy screening : a review », *Eye*, t. 34, nº 3, p. 451-460, mars 2020. DOI : 10.1038/s41433-019-0566-0.

[98] V. GULSHAN, L. PENG, M. CORAM, M. C. STUMPE, D. WU, A. NARAYANASWAMY, S. VENUGOPALAN, K. WIDNER, T. MADAMS, J. CUADROS, R. KIM, R. RAMAN, P. C. NELSON, J. L. MEGA et D. R. WEBSTER, « Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs », *JAMA*, t. 316, nº 22, p. 2402-2410, déc. 2016, ISSN : 0098-7484. DOI : 10.1001/jama.2016.17216.

[99] B. M. ter HAAR ROMENY, *Front-End Vision and Multi-Scale Image Analysis : Multi-Scale Computer Vision Theory and Applications, written in Mathematica*. Dordrecht : Springer Netherlands, 2003, ISBN : 978-1-4020-8840-7. DOI : 10.1007/978-1-4020-8840-7.

[100] H. HADWIGER, *Vorlesungen Über Inhalt, Oberfläche und Isoperimetrie*. Berlin, Heidelberg : Springer, 1957, ISBN : 978-3-642-94703-2. DOI : 10.1007/978-3-642-94702-5.

[101] ——, « Minkowskische Addition und Subtraktion beliebiger Punktmengen und die Theoreme von Erhard Schmidt », *Mathematische Zeitschrift*, t. 53, nº 3, p. 210-218, juin 1950, ISSN : 00255874. DOI : 10.1007/BF01175656.

[102] S. HARDING, R. GREENWOOD, S. ALDINGTON, J. GIBSON, D. OWENS, R. TAYLOR, E. KOHNER, P. SCANLON, G. LEESE et THE DIABETIC RETINOPATHY GRADING AND DISEASE MANAGEMENT WORKING PARTY, « Grading and disease management in national screening for diabetic retinopathy in England and Wales », *Diabetic Medicine*, t. 20, nº 12, p. 965-971, 2003. DOI : 10.1111/j.1464-5491.2003.01077.x.

[103] J. A. HARTIGAN et M. A. WONG, « A K-Means Clustering Algorithm », *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, t. 28, nº 1, p. 100-108, 1979. DOI : 10.2307/2346830.

[104] R. HARTLEY et A. ZISSERMAN, « Computation of the Camera Matrix P », in *Multiple View Geometry in Computer Vision*, 2e éd. Cambridge University Press, 2004, p. 178-194. DOI : 10.1017/CBO9780511811685.011.

[105] ——, « Estimation – 2D Projective Transformations », in *Multiple View Geometry in Computer Vision*, 2e éd. Cambridge University Press, 2004, p. 87-131. DOI : 10.1017/CBO9780511811685.007.

[106] ——, *Multiple View Geometry in Computer Vision*, 2e éd. Cambridge University Press, 2004. DOI : 10.1017/CBO9780511811685.

[107] ——, « Projective Geometry and Transformations of 2D », in *Multiple View Geometry in Computer Vision*, 2e éd. Cambridge University Press, 2004, p. 25-64. DOI : 10.1017/CBO9780511811685.005.

[108] T. HASTIE, R. TIBSHIRANI et J. FRIEDMAN, *The Elements of Statistical Learning*, Second Edition, sér. Springer Series in Statistics. New York, NY : Springer, 2009, ISBN : 978-0-387-84857-0. DOI : 10.1007/978-0-387-84858-7.

[109] N. HAUTIÈRE, R. LABAYRADE et D. AUBERT, « Real-time disparity contrast combination for onboard estimation of the visibility distance », *IEEE Trans. Intell. Transp. Syst.*, t. 7, n° 2, p. 201-212, juin 2006, ISSN : 1524-9050. DOI : 10.1109/TITS.2006.874682.

[110] N. HAUTIÈRE et M. JOURLIN, « Measurement of local contrast in images, Application to the measurement of visibility distance through use of an onboard camera », *Traitement du Signal*, t. 23, n° 2, p. 145-158, 2006, ISSN : 0765-0019. En ligne : http://hdl.handle.net/2042/5826.

[111] H. J. A. M. HEIJMANS, « Theoretical aspects of gray-level morphology », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 13, n° 6, p. 568-582, juin 1991, ISSN : 1939-3539. DOI : 10.1109/34.87343.

[112] H. HEIJMANS et C. RONSE, « The algebraic basis of mathematical morphology I. Dilations and erosions », *Comput. Vision Graphics and Image Process.*, t. 50, n° 3, p. 245-295, juin 1990, ISSN : 0734-189X. DOI : 10.1016/0734-189X(90)90148-O.

[113] H. HEIJMANS, *Morphological image operators*, sér. Advances in Imaging and Electron Physics. San Diego, CA, USA : Academic, 1994, t. 216, ISBN : 9780128210048. DOI : 10.1016/s1076-5670%2820%29x0005-x.

[114] M. HEITZ, N. BONNEEL, D. COEURJOLLY, M. CUTURI et G. PEYRÉ, « Ground Metric Learning on Graphs », *Journal of Mathematical Imaging and Vision*, t. 63, n° 1, p. 89-107, jan. 2021, ISSN : 0924-9907. DOI : 10.1007/s10851-020-00996-z.

[115] C. HERNANDEZ-MATAS, X. ZABULIS et A. A. ARGYROS, « Retinal image registration through simultaneous camera pose and eye shape estimation », in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, p. 3247-3251. DOI : 10.1109/EMBC.2016.7591421.

[116] ——, « An experimental evaluation of the accuracy of keypoints-based retinal image registration », in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, p. 377-381. DOI : 10.1109/EMBC.2017.8036841.

[117] C. HERNANDEZ-MATAS, X. ZABULIS et A. A. ARGYROS, « REMPE : Registration of Retinal Images Through Eye Modelling and Pose Estimation », *IEEE J. Biomed. Health Informat.*, t. 24, n° 12, p. 3362-3373, 2020. DOI : 10.1109/JBHI.2020.2984483.

[118] C. HERNANDEZ-MATAS, X. ZABULIS, A. TRIANTAFYLLOU, P. ANYFANTI, S. DOUMA et A. A. ARGYROS, « FIRE : Fundus Image Registration Dataset », *Journal for Modeling in Ophthalmology*, t. 1, n° 4, p. 16-28, juill. 2017. En ligne : www.modeling-ophthalmology.com/index.php/JMO/article/view/42.

[119] D. HILBERT, « Ueber die vollen Invariantensysteme », *Mathematische Annalen*, t. 42, n° 3, p. 313-373, sept. 1893, ISSN : 00255831. DOI : 10.1007/BF01444162.

[120] P. W. HOLLAND et R. E. WELSCH, « Robust regression using iteratively reweighted least-squares », *Communications in Statistics - Theory and Methods*, t. 6, n° 9, p. 813-827, 1977. DOI : 10.1080/03610927708827533.

[121] H. HOTELLING, « Analysis of a complex of statistical variables into principal components », *Journal of Educational Psychology*, t. 24, p. 498-520, 1933. DOI : 10.1037/h0071325.

[122] K. HU, Z. ZHANG, X. NIU, Y. ZHANG, C. CAO, F. XIAO et X. GAO, « Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function », *Neurocomputing*, t. 309, p. 179-191, 2018, ISSN : 0925-2312. DOI : 10.1016/j.neucom.2018.05.011.

[123] I. HUSSAIN et T. R. REED, « A bond percolation-based model for image segmentation », *IEEE Trans. Image Process.*, t. 6, n° 12, p. 1698-1704, déc. 1997, ISSN : 1057-7149. DOI : 10.1109/83.650123.

[124] D. P. HUTTENLOCHER, G. A. KLANDERMAN et W. J. RUCKLIDGE, « Comparing images using the Hausdorff distance », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 15, n° 9, p. 850-863, sept. 1993, ISSN : 0162-8828. DOI : 10.1109/34.232073.

[125] M. INAM UL HAQ, « Texture analysis in the Logarithmic Image Processing (LIP) framework », thèse de doct., Université Jean Monnet - Saint-Etienne, juin 2013. En ligne : https://tel.archives-ouvertes.fr/tel-00998492.

[126] INTERNATIONAL DIABETES FEDERATION, *IDF Diabetes Atlas*, 9e éd. Bruxelles, Belgique, 2019. En ligne : https://diabetesatlas.org.

[127] INTERNATIONAL DIABETES FEDERATION et THE FRED HOLLOWS FOUNDATION, *Diabetes Eye Health. A guide for health professionals*. Bruxelles, Belgique, 2015. En ligne : https://idf.org/our-activities/care-prevention/eye-health/eye-health-guide.html.

[128] D. JEULIN, « Random texture models for material structures », *Statistics and Computing*, t. 10, n° 2, p. 121-132, 2000, ISSN : 09603174. DOI : 10.1023/A:1008942325749.

[129] ——, « Remarques sur la segmentation probabiliste », Ecole des Mines de Paris, Note interne N-10/08/MM, 2008.

[130] ——, « Extraction de bosses sur des profils », Ecole des Mines de Paris, Note interne N-09/10/MM, juin 2010.

[131] ——, « Morphological probabilistic hierarchies for texture segmentation », *Mathematical Morphology - Theory and Applications*, t. 1, n° 1, p. 216-234, 2016. DOI : doi:10.1515/mathm-2016-0012.

[132] ——, « Some dense random packings generated by the dead leaves model », *Image Analysis & Stereology*, t. 38, n° 1, p. 3-13, 2019, ISSN : 1854-5165. DOI : 10.5566/ias.2081.

[133] ——, « Morphological Models », in *Encyclopedia of Continuum Mechanics*, H. ALTENBACH et A. ÖCHSNER, éd. Berlin, Heidelberg : Springer, 2020, p. 1754-1764, ISBN : 978-3-662-55771-6. DOI : 10.1007/978-3-662-55771-6_70.

[134] R. JONES et P. SOILLE, « Periodic lines : Definition, cascades, and application to granulometries », *Pattern Recognition Letters*, t. 17, n° 10, p. 1057-1063, 1996, ISSN : 0167-8655. DOI : 10.1016/0167-8655(96)00066-9.

[135] M. JOURLIN et N. MONTARD, « A logarithmic version of the top-hat transform in connection with the Asplund distance », *Acta Stereologica*, t. 16, p. 201-208, 1997.

[136] M. JOURLIN, « Chapter Five - Ability of the LIP Model to Simulate Variable Acquisition Conditions », in *Logarithmic Image Processing : Theory and Applications*, sér. Advances in Imaging and Electron Physics, M. JOURLIN, éd., t. 195, Elsevier, 2016, p. 165-184. DOI : 10.1016/bs.aiep.2016.04.005.

[137] ——, « Chapter Four - Dynamic Range Expansion, Night Vision. Stabilization, Centering. Industrial and Biomedical Applications », in *Logarithmic Image Processing : Theory and Applications*, sér. Adv. Imag. Electron Phys. M. JOURLIN, éd., t. 195, Elsevier, 2016, p. 115-163. DOI : 10.1016/bs.aiep.2016.04.004.

[138] ——, « Chapter One - Gray-Level LIP Model. Notations, Recalls, and First Applications », in *Logarithmic Image Processing : Theory and Applications*, sér. Adv. Imag. Electron Phys. M. JOURLIN, éd., t. 195, Elsevier, 2016, p. 1-26. DOI : 10.1016/bs.aiep.2016.04.001.

[139] ——, « Chapter Six - Transfer of Classical Tools to the LIP Context », in *Logarithmic Image Processing : Theory and Applications*, sér. Advances in Imaging and Electron Physics, M. JOURLIN, éd., t. 195, Elsevier, 2016, p. 185-244. DOI : 10.1016/bs.aiep.2016.04.006.

[140] ——, « Chapter Three - Metrics Based on Logarithmic Laws », in *Logarithmic Image Processing : Theory and Applications*, sér. Adv. Imag. Electron Phys. M. JOURLIN, éd., t. 195, Elsevier, 2016, p. 61-113. DOI : http://dx.doi.org/10.1016/bs.aiep.2016.04.003.

[141] ——, « Chapter Two - Various Contrast Concepts », in *Logarithmic Image Processing : Theory and Applications*, sér. Adv. Imag. Electron Phys. M. JOURLIN, éd., t. 195, Elsevier, 2016, p. 27-60. DOI : 10.1016/bs.aiep.2016.04.002.

[142] M. JOURLIN, E. COUKA, B. ABDALLAH, J. CORVO et J. BREUGNOT, « Asplünd's metric defined in the Logarithmic Image Processing (LIP) framework : A new way to perform double-sided image probing for non-linear grayscale pattern matching », *Pattern Recognit.*, t. 47, n° 9, p. 2908-2924, sept. 2014, ISSN : 0031-3203. DOI : 10.1016/j.patcog.2014.03.031.

[143] M. JOURLIN et J. PINOLI, « Logarithmic image processing : The mathematical and physical framework for the representation and processing of transmitted images », in sér. Adv. Imag. Electron Phys. P. W. HAWKES, éd., t. 115, Elsevier, 2001, p. 129-196. DOI : 10.1016/S1076-5670(01)80095-1.

[144] M. JOURLIN, *Logarithmic Image Processing : Theory and Applications*, sér. Adv. Imag. Electron Phys. New York, NY, USA : Elsevier, 2016, t. 195, ISBN : 9780128052297. En ligne : www.sciencedirect.com/science/bookseries/10765670/195.

[145] M. JOURLIN, J. BREUGNOT, F. ITTHIRAD, M. BOUABDELLAH et B. CLOSS, « Chapter 2 - Logarithmic Image Processing for Color Images », in *Adv. Imag. Electron Phys.* Sér. Adv. Imag. Electron Phys. P. W. HAWKES, éd., t. 168, Elsevier, 2011, p. 65-107. DOI : 10.1016/B978-0-12-385983-9.00002-8.

[146] M. JOURLIN, M. CARRÉ, J. BREUGNOT et M. BOUABDELLAH, « Chapter 7 - Logarithmic Image Processing : Additive Contrast, Multiplicative Contrast, and Associated Metrics », in sér. Adv. Imag. Electron Phys. T. 171, Elsevier, 2012, p. 357-406. DOI : 10.1016/B978-0-12-394297-5.00007-6.

[147] M. JOURLIN et G. NOYEL, « Homogeneity of a region in the logarithmic image processing framework : application to region growing algorithms », in *Physics and Mechanics of Random Structures : from Morphology to Material Properties*, F. WILLOT et S. FOREST, éd., sér. Physics and Mechanics of Random Structures : from Morphology to Material Properties, Ile d'Oléron, France : Presse des Mines, juin 2018, p. 5-7. En ligne : https://hal.archives-ouvertes.fr/hal-01822522.

[148] M. JOURLIN et J.-C. PINOLI, « Image dynamic range enhancement and stabilization in the context of the logarithmic image processing model », *Signal Process.*, t. 41, n° 2, p. 225-237, jan. 1995, ISSN : 0165-1684. DOI : 10.1016/0165-1684(94)00102-6.

[149] M. JOURLIN, J.-C. PINOLI et R. ZEBOUDJ, « Contrast definition and contour detection for logarithmic images », *Journal of Microscopy*, t. 156, n° 1, p. 33-40, 1989. DOI : 10.1111/j.1365-2818.1989.tb02904.x.

[150] M. JOURLIN et J.-C. PINOLI, « A model for logarithmic image-processing », *J. Microsc.*, t. 149, n° 1, p. 21-35, jan. 1988. DOI : 10.1111/j.1365-2818.1988.tb04559.x.

[151] T. KAUPPI, V. KALESNYKIENE, J.-K. KAMARAINEN, L. LENSU, I. SORRI, A. RANINEN, R. VOUTILAINEN, H. UUSITALO, H. KALVIAINEN et J. PIETILA, « The DIARETDB1 diabetic retinopathy database and evaluation protocol », in *Proc BMVC*, 2007, p. 15.1-15.10, ISBN : 1-901725-34-0. DOI : 10.5244/C.21.15.

[152] M. KHOSRAVI et R. SCHAFER, « Template matching based on a grayscale hit-or-miss transform », *IEEE Trans. Image Process.*, t. 5, n° 6, p. 1060-1066, juin 1996, ISSN : 10577149. DOI : 10.1109/83.503921.

[153] J. KLEIN, « A registration algorithm of eye fundus images using a Bayesian Hough transform », *IET Conference Proceedings*, p. 479-483, jan. 1999. DOI : 10.1049/cp:19990368.

[154] R. KLEIN, B. E. K. KLEIN, S. E. MOSS, M. D. DAVIS et D. L. DEMETS, « The Wisconsin Epidemiologic Study of Diabetic Retinopathy : IX. Four-Year Incidence and Progression of Diabetic Retinopathy When Age at Diagnosis Is Less Than 30 Years », *Archives of Ophthalmology*, t. 107, n° 2, p. 237-243, fév. 1989, ISSN : 0003-9950. DOI : 10.1001/archopht.1989.01070010243030.

[155] ——, « The Wisconsin Epidemiologic Study of Diabetic Retinopathy : X. Four-Year Incidence and Progression of Diabetic Retinopathy When Age at Diagnosis Is 30 Years or More », *Archives of Ophthalmology*, t. 107, n° 2, p. 244-249, fév. 1989, ISSN : 0003-9950. DOI : 10.1001/archopht.1989.01070010250031.

[156] R. KOHLER, « A segmentation system based on thresholding », *Computer Graphics and Image Processing*, t. 15, n° 4, p. 319-338, 1981, ISSN : 0146-664X. DOI : http://dx.doi.org/10.1016/S0146-664X(81)80015-9.

[157] N. LARSEN, J. GODT, M. GRUNKIN, H. LUND-ANDERSEN et M. LARSEN, « Automated detection of diabetic retinopathy in a fundus photographic screening population », *Investigative Ophthalmology and Visual Science*, t. 44, n° 2, p. 767-771, fév. 2003, ISSN : 01460404. DOI : 10.1167/iovs.02-0417.

[158] Y. LECUN, Y. BENGIO et G. HINTON, « Deep learning », *Nature*, t. 521, n° 7553, p. 436-444, mai 2015, ISSN : 0028-0836. DOI : 10.1038/nature14539.

[159] S. LEE, M. D. ABRÀMOFF et J. M. REINHARDT, « Feature-based pairwise retinal image registration by radial distortion correction », in *Medical Imaging 2007 : Image Processing*, J. P. W. PLUIM et J. M. REINHARDT, éd., International Society for Optics et Photonics, t. 6512, SPIE, 2007, p. 701-710. DOI : 10.1117/12.710676.

[160] ——, « Retinal image mosaicing using the radial distortion correction model », in *Medical Imaging 2008 : Image Processing*, J. M. REINHARDT et J. P. W. PLUIM, éd., International Society for Optics et Photonics, t. 6914, SPIE, 2008, p. 1040-1048. DOI : 10.1117/12.773161.

[161] S. LEE, J. M. REINHARDT, P. C. CATTIN et M. D. ABRÀMOFF, « Objective and expert-independent validation of retinal image registration algorithms by a projective imaging distortion model », *Medical Image Analysis*, t. 14, n° 4, p. 539-549, 2010, ISSN : 1361-8415. DOI : 10.1016/j.media.2010.04.001.

[162] J. LI et B.-L. LU, « An adaptive image Euclidean distance », *Pattern Recognit.*, t. 42, n° 3, p. 349-357, mars 2009, ISSN : 0031-3203. DOI : 10.1016/j.patcog.2008.07.017.

[163] Y. LIN et G. MEDIONI, « Retinal image registration from 2D to 3D », in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008, p. 1-8. DOI : 10.1109/CVPR.2008.4587705.

[164] L. LIND, J. KULLBERG, H. AHLSTRÖM, K. MICHAËLSSON et R. STRAND, « Proof of principle study of a detailed whole-body image analysis technique, “Imiomics”, regarding adipose and lean tissue distribution », *Scientific Reports*, t. 9, n° 1, p. 7388, déc. 2019, ISSN : 2045-2322. DOI : 10.1038/s41598-019-43690-w.

[165] D. LLORET, J. SERRAT, A. M. LOPEZ, A. SOLER et J. J. VILLANEUVA, « Retinal image registration using creases as anatomical landmarks », in *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, t. 3, 2000, 203-206 vol.3. DOI : 10.1109/ICPR.2000.903521.

[166] D. G. LOWE, « Distinctive Image Features from Scale-Invariant Keypoints », *Int. J. Comput. Vision*, t. 60, n° 2, p. 91-110, nov. 2004, ISSN : 1573-1405. DOI : 10.1023/B:VISI.0000029664.99615.94.

[167] L. LU, Y. ZHENG, G. CARNEIRO et L. YANG, éd., *Deep Learning and Convolutional Neural Networks for Medical Image Computing - Precision Medicine, High Performance and Large-Scale Datasets*, sér. Advances in Computer Vision and Pattern Recognition. Springer, 2017, ISBN : 978-3-319-42998-4. DOI : 10.1007/978-3-319-42999-1.

[168] J. MAINTZ et M. A. VIERGEVER, « A survey of medical image registration », *Medical Image Analysis*, t. 2, n° 1, p. 1-36, 1998, ISSN : 1361-8415. DOI : 10.1016/S1361-8415(01)80026-8.

[169] J. MALIK, S. BELONGIE, T. LEUNG et J. SHI, « Contour and texture analysis for image segmentation », *International Journal of Computer Vision*, t. 43, n° 1, p. 7-27, juin 2001, ISSN : 09205691. DOI : 10.1023/A:1011174803800.

[170] F. MALMBERG et C. L. LUENGO HENDRIKS, « An efficient algorithm for exact evaluation of stochastic watersheds », *Pattern Recognition Letters*, t. 47, p. 80-84, 2014, Advances in Mathematical Morphology, ISSN : 0167-8655. DOI : 10.1016/j.patrec.2014.03.016.

[171] P. MARAGOS, « A representation theory for morphological image and signal processing », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 11, n° 6, p. 586-599, juin 1989. DOI : 10.1109/34.24793.

[172] P. MARAGOS et R. SCHAFER, « Morphological filters–Part II : Their relations to median, order-statistic, and stack filters », *IEEE Trans. Acoust., Speech, Signal Process.*, t. 35, n° 8, p. 1170-1184, août 1987, ISSN : 0096-3518. DOI : 10.1109/TASSP.1987.1165254.

[173] O. MARIANI, A. ERNST, N. MERCADER et M. LIEBLING, « Reconstruction of Image Sequences From Ungated and Scanning-Aberrated Laser Scanning Microscopy Images of the Beating Heart », *IEEE Transactions on Computational Imaging*, t. 6, p. 385-395, 2020. DOI : 10.1109/TCI.2019.2948772.

[174] D. MARR et T. POGGIO, « A Computational Theory of Human Stereo Vision », *Proceedings of the Royal Society of London. Series B, Biological Sciences*, t. 204, n° 1156, p. 301-328, 1979, ISSN : 00804649. En ligne : www.jstor.org/stable/77509.

[175] P. MASSIN et A. ERGINAY, *Rétinopathie diabétique*. Elsevier Health Sciences France, 2012, ISBN : 9782294716812. En ligne : https://books.google.fr/books?id=U6nEbTXZkL0C.

[176] G. MATHERON, *Eléments pour une théorie des milieux poreux*. Paris, Fr. : Masson, 1967. En ligne : http://cg.ensmp.fr/bibliotheque/public/MATHERON_Ouvrage_00101.pdf.

[177] ——, *Random sets and integral geometry*, sér. Wiley series in probability and mathematical statistics. New York, NY, USA : Wiley, 1975, ISBN : 9780471576211. En ligne : https://books.google.fr/books?id=bgzvAAAAMAAJ.

[178] G. MATHERON, « Filters and Lattices », in *Image analysis and Mathematical Morphology : Theoretical advances*, J. SERRA, éd. Academic Press, 1988, t. 2, chap. 6, p. 115-140, ISBN : 9780126372410. En ligne : https://books.google.fr/books?id=BpdTAAAAYAAJ.

[179] G. K. MATSOPOULOS, N. A. MOURAVLIANSKY, K. K. DELIBASIS et K. S. NIKITA, « Automatic retinal image registration scheme using global optimization techniques », *IEEE Trans. Inf. Technol. Biomed.*, t. 3, n° 1, p. 47-60, 1999. DOI : 10.1109/4233.748975.

[180] F. MAYET et M. JOURLIN, « Classification d'images par analyse fractale et multifractale », www.worldcat.org/oclc/492527269, thèse de doct., Université Jean Monnet, Saint-Etienne, France, 2001.

[181] F. MAYET, J.-C. PINOLI et M. JOURLIN, « Physical justifications and applications of the LIP model for the processing of transmitted light images », *Trait Signal*, t. 13, n° 3, p. 251-262, 1996, ISSN : 0765-0019. En ligne : http://hdl.handle.net/2042/1959.

[182] A. M. MENDONCA et A. CAMPILHO, « Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction », *IEEE Trans. Med. Imag.*, t. 25, n° 9, p. 1200-1213, 2006. DOI : 10.1109/TMI.2006.879955.

[183] O. MERVEILLE, H. TALBOT, L. NAJMAN et N. PASSAT, « Curvilinear Structure Analysis by Ranking the Orientation Responses of Path Operators », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 40, n° 2, p. 304-317, fév. 2018. DOI : 10.1109/TPAMI.2017.2672972.

[184] D. MERY, V. RIFFO, U. ZSCHERPEL, G. MONDRAGÓN, I. LILLO, I. ZUCCAR, H. LOBEL et M. CARRASCO, « GDXray : The Database of X-ray Images for Nondestructive Testing », *Journal of Nondestructive Evaluation*, t. 34, n° 4, p. 42, nov. 2015, ISSN : 1573-4862. DOI : 10.1007/s10921-015-0315-7.

[185] C. METZ, « ROC Methodology in Radiologic Imaging », *Investigative Radiology*, t. 21, p. 720-733, 1986. DOI : 10.1097/00004424-198609000-00009.

[186] F. MEYER, « Iterative image transformations for an automatic screening of cervical smears. », *J. Histochemistry & Cytochemistry*, t. 27, n° 1, p. 128-135, 1979. DOI : 10.1177/27.1.438499.

[187] ——, *Topographical Tools for Filtering and Segmentation 1 : Watersheds on Node- or Edge-weighted Graphs*. Hoboken, NJ, USA : Wiley, 2019, ISBN : 9781119579519. DOI : 10.1002/9781119579519.

[188] ——, *Topographical Tools for Filtering and Segmentation 2 : Flooding and Marker-based Segmentation on Node- or Edge-weighted Graphs*. Hoboken, NJ, USA : Wiley, 2019, ISBN : 9781119575139. DOI : 10.1002/9781119575139.

[189] F. MEYER et J. STAWIASKI, « A Stochastic Evaluation of the Contour Strength », in *Pattern Recognition*, M. GOESELE, S. ROTH, A. KUIJPER, B. SCHIELE et K. SCHINDLER, éd., Springer Berlin Heidelberg, 2010, p. 513-522, ISBN : 978-3-642-15986-2. DOI : 10.1007/978-3-642-15986-2_52.

[190] H. MINKOWSKI, « Volumen und Oberfläche », *Mathematische Annalen*, t. 57, n° 4, p. 447-495, déc. 1903, ISSN : 1432-1807. DOI : 10.1007/BF01445180.

[191] P. MONASSE et F. GUICHARD, « Fast computation of a contrast-invariant image representation », *IEEE Trans. Image Process.*, t. 9, n° 5, p. 860-872, mai 2000, ISSN : 1057-7149. DOI : 10.1109/83.841532.

[192] J. J. MORÉ, « Recent Developments in Algorithms and Software for Trust Region Methods », in *Mathematical Programming The State of the Art : Bonn 1982*, A. BACHEM, B. KORTE et M. GRÖTSCHEL, éd. Berlin, Heidelberg : Springer, 1983, p. 258-287, ISBN : 978-3-642-68874-4. DOI : 10.1007/978-3-642-68874-4_11.

[193] J. J. MORÉ, « The Levenberg-Marquardt algorithm : Implementation and theory », in *Numerical Analysis*, G. A. WATSON, éd., Berlin, Heidelberg : Springer, 1978, p. 105-116, ISBN : 978-3-540-35972-2.

[194] C. MOREAU, G. NOYEL, Y. SHEN et J.-P. ZANELLA, « Analysis of the digital image of the external surface of a tyre and processing of false measurement points », WO2012143198 (A1), Aussi publié comme US2014307941 (A1) ; US9230318 (B2) ; FR2974219(A1) ; EP2700050 (A1) ; JP2014517265 (A) ; CN103649989 (A) ; CN103649989 (B) ; BR112013024106 (A2), oct. 2012. En ligne : https : / / patentscope.wipo.int/search/en/WO2012143198.

[195] ——, « Analysis of the digital image of the external surface of a tyre and processing of false measurement points », [Org. Mondiale Propriété Intellectuelle] WO2012143198 (A1), aussi publié comme [**USA**] US2014307941 (A1) ; US9230318 (B2) ; [**France**] FR2974219(A1) ; [**Europe**] EP2700050 (A1) ; [**Japon**] JP2014517265 (A) ; [**Chine**] CN103649989 (A) ; CN103649989 (B) ; [**Brésil**] BR112013024106 (A2), oct. 2012. En ligne : https://patentscope.wipo.int/search/en/WO2012143198.

[196] ——, « Analysis of the digital image of the internal surface of a tyre and processing of false measurement points », [Org. Mondiale Propriété Intellectuelle] WO2012143199 (A1), aussi publié comme [**USA**] US2014270390 (A1) ; US9230337 (B2) ; [**France**] FR2974220 (A1) ; FR2974220 (B1) ; EP2700052 (A1) ; [**Europe**] EP2700052 (B1) ; [**Japon**] JP2014513832 (A) ; JP5876139 (B2) ; [**Chine**] CN103493096 (A) ; CN103493096 (B) ; [**Brésil**] BR112013024108 (A2), oct. 2012. En ligne : https : //patentscope.wipo.int/search/en/WO2012143199.

[197] ——, « Analysis of the digital image of the surface of a tyre and processing of non-measurement points », [Org. Mondiale Propriété Intellectuelle] WO2012143197 (A1), aussi publié comme [**USA**] US2014185883 (A1) ; US9224198 (B2) ; [**France**] FR2974218 (A1) ; [**Europe**] EP2700051 (A1) ; [**Japon**] JP2014513793 (A) ; [**Chine**] CN103493095 (A) ; [**Brésil**] BR112013023977 (A2), oct. 2012. En ligne : https : / / patentscope.wipo.int/search/en/WO2012143197.

[198] B. MOREL, « Rehaussement de contraste en imagerie Rayons-X des pneumatiques », Rapport de Master, CPE-Lyon et Master Recherche Electronique Electrotechnique Automatique Procédés, Systèmes et Images, INSA de Lyon, France, 2013.

[199] B. NAEGEL et N. PASSAT, « Interactive Segmentation Based on Component-trees », *Image Processing On Line*, t. 4, p. 89-97, 2014. DOI : 10.5201/ipol.2014.71.

[200] L. NAJMAN et J. COUSTY, « A graph-based mathematical morphology reader », *Pattern Recognition Letters*, t. 47, p. 3-17, 2014, Advances in Mathematical Morphology, ISSN : 0167-8655. DOI : 10.1016/j.patrec.2014.05.007.

[201] L. NAJMAN et H. TALBOT, *Mathematical Morphology : From Theory to Applications*, 1re éd. Hoboken, NJ, USA : Wiley, 2013, ISBN : 9781118600788. DOI : 10.1002/9781118600788.

[202] L. NAVARRO, G. COURBEBAISSE et M. JOURLIN, « Chapter Two - Logarithmic Wavelets », in sér. Adv. Imag. Electron Phys. P. W. HAWKES, éd., t. 183, Elsevier, 2014, p. 41-98. DOI : 10.1016/B978-0-12-800265-0.00002-3.

[203] L. NAVARRO, G. DENG et G. COURBEBAISSE, « The symmetric logarithmic image processing model », *Digital Signal Process.*, t. 23, no 5, p. 1337-1343, sept. 2013, ISSN : 1051-2004. DOI : 10.1016/j.dsp.2013.07.001.

[204] M. NIELSEN, *Neural Networks and Deep Learning*. Determination Press, 2015. En ligne : http://neuralnetworksanddeeplearning.com.

[205] M. NIEMEIJER, « Automatic Detection of Diabetic Retinopathy in Digital Fundus Photographs », thèse de doct., Utrecht University, The Netherlands, 2006. En ligne : `http://dspace.library.uu.nl/handle/1874/18680`.

[206] M. NIEMEIJER, M. D. ABRÀMOFF et B. VAN GINNEKEN, « Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening », *Medical Image Analysis*, t. 10, n° 6, p. 888-898, 2006, ISSN : 1361-8415. DOI : `10.1016/j.media.2006.09.006`.

[207] G. NOYEL, « Speeding up the Köhler's method of contrast thresholding », in *2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, éd., IEEE, Beijing, China : IEEE, sept. 2017, p. 320-324. DOI : `10.1109/ICIP.2017.8296295`.

[208] G. NOYEL, J. ANGULO, D. JEULIN, D. BALVAY et C. CUENOD, « Filtering, segmentation and region classification by hyperspectral mathematical morphology of DCE-MRI series for angiogenesis imaging », in *2008 5th IEEE International Symposium on Biomedical Imaging (ISBI) : From Nano to Macro*, Paris, France : IEEE, mai 2008, p. 1517-1520. DOI : `10.1109/ISBI.2008.4541297`.

[209] G. NOYEL et M. JOURLIN, « Asplund's metric defined in the logarithmic image processing (LIP) framework for colour and multivariate images », in *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, éd., Québec City, Canada : IEEE, sept. 2015, p. 3921-3925. DOI : `10.1109/ICIP.2015.7351540`.

[210] G. NOYEL, R. THOMAS, S. ILES, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Registration Of Retinal Images From Public Health By Minimising An Error Between Vessels Using An Affine Model With Radial Distortions », in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Venice, Italy : IEEE, avr. 2019, p. 561-564. DOI : `10.1109/ISBI.2019.8759415`.

[211] G. NOYEL, C. VARTIN, P. BOYLE et L. KODJIKIAN, « Retinal Vessel Segmentation by Probing Adaptive to Lighting Variations », in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Iowa City, United States : IEEE, avr. 2020, p. 1246-1249. DOI : `10.1109/ISBI45749.2020.9098332`.

[212] G. NOYEL, « Filtrage, réduction de dimension, classification et segmentation morphologique hyperspectrale », Sous la direction de Dominique Jeulin et Jesus Angulo, thèse de doct., École Nationale Supérieure des Mines de Paris, Fontainebleau, France, sept. 2008. En ligne : `https://pastel.archives-ouvertes.fr/pastel-00004473`.

[213] ——, « Method of Monitoring the Appearance of the Surface of a Tire », [Org. Mondiale Propriété Intellectuelle] WO2011131410 (A1), aussi publié comme [**USA**] US2013129182 (A1), US9002093 (B2), [**France**] FR2959046 (A1), FR2959046 (B1), [**Europe**] EP2561479 (A1), [**Japon**] JP2013532315 (A), JP5779232 (B2), [**Chine**] CN102844791 (A), CN102844791 (B), [**Brésil**] BR112012025402 (A2), oct. 2011. En ligne : `https://patentscope.wipo.int/search/en/WO2011131410`.

[214] G. NOYEL et al., « Method for determining elements in relief on the surface of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2012156260 (A1), aussi publié comme [**USA**] US2015043780 (A1), US9002130 (B2), [**France**] FR2975523 (A1), FR2975523 (B1), [**Europe**] EP2710555 (A1), EP2710555 (B1), [**Chine**] CN103649995 (A), CN103649995 (B), nov. 2012. En ligne : `https://patentscope.wipo.int/search/en/WO2012156260`.

[215] ——, « Method for determining markings in relief on the outer surface of the sidewall of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2012156262 (A1), aussi publié comme [**USA**] US2014219541 (A1), US9025853 (B2), [**France**] FR2975524 (A1), FR2975524 (B1), [**Europe**] EP2710552 (A1), EP2710552 (B1), [**Chine**] CN103649991 (A), CN103649991 (B), nov. 2012. En ligne : https://patentscope.wipo.int/search/en/WO2012156262.

[216] G. NOYEL, « A Link Between the Multiplicative and Additive Functional Asplund's Metrics », in *14th International Symposium on Mathematical Morphology (ISMM 2019)*, B. BURGETH, A. KLEEFELD, B. NAEGEL, N. PASSAT et B. PERRET, éd., Saarland University, Saarbrücken, Germany : Springer International Publishing, juill. 2019, p. 41-53, ISBN : 978-3-030-20867-7. DOI : 10.1007/978-3-030-20867-7_4.

[217] ——, « Logarithmic Mathematical Morphology : A New Framework Adaptive to Illumination Changes », in *23rd Iberoamerican Congress on Pattern Recognition (CIARP 2018)*, R. VERA-RODRIGUEZ, J. FIERREZ et A. MORALES, éd., sér. Lecture Notes in Computer Science, t. 11401, Madrid, Spain : Springer International Publishing, nov. 2019, p. 453-461, ISBN : 978-3-030-13469-3. DOI : 10.1007/978-3-030-13469-3_53.

[218] G. NOYEL, J. ANGULO et D. JEULIN, « Classification-driven stochastic watershed. Application to multispectral segmentation », in *4th European Conference on Colour in Graphics, Imaging, and Vision and 10th International Symposium on Multispectral Colour Science (CGIV 2008/MCS 2008)*, Terrassa - Barcelona, Spain : Society for Imaging Science and Technology, juin 2008, p. 471-476. En ligne : www.ingentaconnect.com/content/ist/cgiv/2008/00002008/00000001/art00101.

[219] G. NOYEL, J. ANGULO, D. JEULIN, D. BALVAY et C.-A. CUENOD, « Multivariate mathematical morphology for DCE-MRI image analysis in angiogenesis studies », *Image Analysis & Stereology*, t. 34, n° 1, p. 1-25, 2015, ISSN : 1854-5165. DOI : 10.5566/ias.1109.

[220] G. NOYEL, J. ANGULO et D. JEULIN, « Morphological segmentation of hyperspectral images », *Image Analysis & Stereology*, t. 26, n° 3, p. 101-109, 2007, ISSN : 1854-5165. DOI : 10.5566/ias.v26.p101-109.

[221] ——, « Random Germs and Stochastic Watershed for Unsupervised Multispectral Image Segmentation », in *11th International Conference Knowledge-Based Intelligent Information and Engineering Systems (KES 2007)*, B. APOLLONI, R. J. HOWLETT et L. JAIN, éd., sér. Lecture Notes in Computer Science, t. 4694, Vietri sul Mare, Salerno, Italy : Springer, sept. 2007, p. 17-24, ISBN : 978-3-540-74829-8. DOI : 10.1007/978-3-540-74829-8_3.

[222] ——, « A new spatio-spectral morphological segmentation for multi-spectral remote-sensing images », *International Journal of Remote Sensing*, t. 31, n° 22, p. 5895-5920, 2010. DOI : 10.1080/01431161.2010.512314.

[223] ——, « Regionalized Random Germs by a Classification for Probabilistic Watershed Application : Angiogenesis Imaging Segmentation », in *Progress in Industrial Mathematics at ECMI 2008*, A. D. FITT, J. NORBURY, H. OCKENDON et E. WILSON, éd., sér. Mathematics in Industry. Berlin, Heidelberg : Springer, 2010, p. 211-216, ISBN : 978-3-642-12110-4. DOI : 10.1007/978-3-642-12110-4_29.

[224] ——, « Fast computation of all pairs of geodesic distances », *Image Analysis & Stereology*, t. 30, n° 2, p. 101-109, 2011, ISSN : 1854-5165. DOI : 10.5566/ias.v30.p101-109.

[225] G. NOYEL, D. JEULIN, E. PARRA-DENIS et M. BILODEAU, « Improved method of checking the appearance of the surface of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2013045593 (A1), aussi publié comme [**USA**] US9189841 (B2), [**France**] FR2980735 (A1), FR2980735 (B1), [**Europe**] EP2761587 (A1), [**Chine**] CN103843034 (A), avr. 2013. En ligne : https://patentscope.wipo.int/search/en/WO2013045593.

[226] ——, « Method of fast analysis of the relief elements featuring on the internal surface of a tyre », [Org. Mondiale Propriété Intellectuelle] WO2013045594 (A1), aussi publié comme [**USA**] US9196034 (B2), [**France**] FR2980896 (A1), FR2980896 (B1), [**Europe**] EP2761588 (A1), EP2761588 (B1), [**Chine**] CN103858147 (A), CN103858147 (B), avr. 2013. En ligne : https://patentscope.wipo.int/search/en/WO2013045594.

[227] G. NOYEL et M. JOURLIN, « A simple expression for the map of Asplund's distances with the multiplicative Logarithmic Image Processing (LIP) law », in *12th European Congress for Stereology and Image Analysis 2017 (ECSIA 2017)*, University of Kaiserslautern and Fraunhofer-Institut für Techno- und Wirtschaftsmathematik (ITWM), Kaiserslautern, Germany, sept. 2017. En ligne : https://arxiv.org/abs/1708.08992.

[228] ——, « Double-Sided Probing by Map of Asplund's Distances Using Logarithmic Image Processing in the Framework of Mathematical Morphology », in *13th International Symposium on Mathematical Morphology (ISMM 2017)*, J. ANGULO, S. VELASCO-FORERO et F. MEYER, éd., sér. Lecture Notes in Computer Science, t. 10225, Fontainebleau, France : Springer International Publishing, mai 2017, p. 408-420, ISBN : 978-3-319-57240-6. DOI : 10.1007/978-3-319-57240-6_33.

[229] ——, « Spatio-Colour Asplund's Metric and Logarithmic Image Processing for Colour Images (LIPC) », in *21st Iberoamerican Congress on Pattern Recognition (CIARP 2016)*, C. BELTRÁN-CASTAÑÓN, I. NYSTRÖM et F. FAMILI, éd., sér. Lecture Notes in Computer Science, t. 10125, Lima, Peru : Springer International Publishing, nov. 2017, p. 36-43, ISBN : 978-3-319-52277-7. DOI : 10.1007/978-3-319-52277-7_5.

[230] ——, « Region homogeneity in the logarithmic image processing framework : application to region growing algorithms », *Image Analysis & Stereology*, t. 38, n° 1, p. 43-52, 2019, ISSN : 1854-5165. DOI : 10.5566/ias.2038.

[231] ——, « Functional Asplund Metrics for Pattern Matching, Robust to Variable Lighting Conditions », *Image Analysis & Stereology*, t. 39, n° 2, p. 53-71, 2020, ISSN : 1854-5165. DOI : 10.5566/ias.2292.

[232] G. NOYEL, M. JOURLIN, M. SMANS, R. THOMAS, S. ILES, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Enhancing Eye Fundus Images for Diabetic Retinopathy Screening », in *American Diabetes Association 77th Scientific Sessions - ADA2017*, American Diabetes Association, San Diego, CA, United States, juin 2017. En ligne : https://hal.archives-ouvertes.fr/hal-03162202.

[233] G. NOYEL, M. JOURLIN, R. THOMAS, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Contrast enhancement of eye fundus images », in *IDF2015 - World Diabetes Congress*, International Diabetes Federation, Vancouver, Canada, nov. 2015. En ligne : https://hal.archives-ouvertes.fr/hal-03162208.

[234] G. NOYEL, R. THOMAS, G. BHAKTA, A. CROWDER, D. OWENS et P. BOYLE, « Superimposition of eye fundus images for longitudinal analysis from large public health databases », *Biomedical Physics & Engineering Express*, t. 3, n° 4, p. 045 015, juill. 2017. DOI : 10.1088/2057-1976/aa7d16.

[235] D. OBERKAMPF, D. F. DEMENTHON et L. S. DAVIS, « Iterative Pose Estimation Using Coplanar Feature Points », *Computer Vision and Image Understanding*, t. 63, n° 3, p. 495-511, 1996, ISSN : 1077-3142. DOI : 10.1006/cviu.1996.0037.

[236] F. ODONE, E. TRUCCO et A. VERRI, « General Purpose Matching of Grey Level Arbitrary Images », in *Visual Form 2001*, Berlin, Heidelberg : Springer, 2001, p. 573-582, ISBN : 978-3-540-45129-7. DOI : 10.1007/3-540-45129-3_53.

[237] F. P. OLIVEIRA et J. M. R. TAVARES, « Medical image registration : a review », *Computer Methods in Biomechanics and Biomedical Engineering*, t. 17, n° 2, p. 73-93, 2014, PMID : 22435355. DOI : 10.1080/10255842.2012.670855.

[238] *Opera Magistris*, 3e éd. www.sciences.ch, 2018, ISBN : 978-2-8399-0932-7. En ligne : https://archive.org/details/OperaMagistris.

[239] V. PAPYAN, Y. ROMANO, J. SULAM et M. ELAD, « Theoretical Foundations of Deep Learning via Sparse Representations : A Multilayer Sparse Model and Its Connection to Convolutional Neural Networks », *IEEE Signal Process. Mag.*, t. 35, n° 4, p. 72-89, 2018. DOI : 10.1109/MSP.2018.2820224.

[240] J. PARK, S. BYUN et B. LEE, « Lens distortion correction using ideal image coordinates », *IEEE Trans. Consum. Electron.*, t. 55, n° 3, p. 987-991, 2009. DOI : 10.1109/TCE.2009.5278053.

[241] E. PARRA-DENIS, M. BILODEAU et D. JEULIN, « Multistep Detection of Oriented Structure in Complex Textures », in *International Congress for Stereology*, Beijing, China, oct. 2011. En ligne : https://hal.archives-ouvertes.fr/hal-00880311.

[242] N. PASSAT, B. NAEGEL, F. ROUSSEAU, M. KOOB et J.-L. DIETEMANN, « Interactive segmentation based on component-trees », *Pattern Recogn*, t. 44, n° 10, p. 2539-2554, 2011.

[243] C. PEYREGA et D. JEULIN, « Estimation of tortuosity and reconstruction of geodesic paths in 3D », *Image Analysis & Stereology*, t. 32, n° 1, p. 27-43, 2013, ISSN : 1854-5165. DOI : 10.5566/ias.v32.p27-43.

[244] J. P. W. PLUIM, J. B. A. MAINTZ et M. A. VIERGEVER, « Mutual-information-based registration of medical images : a survey », *IEEE Trans. Med. Imag.*, t. 22, n° 8, p. 986-1004, 2003. DOI : 10.1109/TMI.2003.815867.

[245] P. PORWAL, S. PACHADE, R. KAMBLE, M. KOKARE, G. DESHMUKH, V. SAHASRABUDDHE et F. MERIAUDEAU. « Indian Diabetic Retinopathy Image Dataset (IDRiD) », IEEE Dataport. (2018), En ligne : https://dx.doi.org/10.21227/H25W98.

[246] ——, « Indian Diabetic Retinopathy Image Dataset (IDRiD) : A Database for Diabetic Retinopathy Screening Research », *Data*, t. 3, n° 3, p. 25, juill. 2018, ISSN : 2306-5729. DOI : 10.3390/data3030025.

[247] H. POTTMANN et M. HOFER, « Geometry of the Squared Distance Function to Curves and Surfaces », in *Visualization and Mathematics III*, H.-C. HEGE et K. POLTHIER, éd., Berlin, Heidelberg : Springer, 2003, p. 221-242, ISBN : 978-3-662-05105-4. DOI : 10.1007/978-3-662-05105-4_12.

[248] P. PUDIL, J. NOVOVIČOVÁ et J. KITTLER, « Floating search methods in feature selection », *Pattern Recognition Letters*, t. 15, n° 11, p. 1119-1125, 1994, ISSN : 0167-8655. DOI : 10.1016/0167-8655(94)90127-9.

[249] T. PUNTOUS, J. RECH, D. DELAFOSSE et M. JOURLIN, « Perceptual percolation based on logarithmic image processing model », in *Proceedings 13th International Congress for Stereology (ICS 13)*, Beijing, China : International Society for Stereology, 2011. En ligne : http://cstm.cnki.net/stmt/TitleBrowse/KnowledgeNet/EGTA201110001186?db=STMI8319.

[250] T. PUNTOUS, « Perception et détection automatique des rayures de polissage », www.worldcat.org/oclc/919252120, thèse de doct., Université Jean Monnet, Saint-Etienne, France, 2013.

[251] J. QIU, C. CHEN, S. LIU, H.-Y. ZHANG et B. ZENG, « SlimConv : Reducing Channel Redundancy in Convolutional Neural Networks by Features Recombining », *IEEE Trans. Image Process.*, t. 30, p. 6434-6445, 2021. DOI : 10.1109/TIP.2021.3093795.

[252] G. QUELLEC, K. CHARRIÈRE, Y. BOUDI, B. COCHENER et M. LAMARD, « Deep image mining for diabetic retinopathy screening », *Medical Image Analysis*, t. 39, p. 178-193, 2017, ISSN : 1361-8415. DOI : 10.1016/j.media.2017.04.012.

[253] G. QUELLEC, M. LAMARD, A. ERGINAY, A. CHABOUIS, P. MASSIN, B. COCHENER et G. CAZUGUEL, « Automatic detection of referral patients due to retinal pathologies through data mining », *Medical Image Analysis*, t. 29, p. 47-64, 2016, ISSN : 1361-8415. DOI : 10.1016/j.media.2015.12.006.

[254] N. P. RAMAIAH, E. P. IJJINA et C. K. MOHAN, « Illumination invariant face recognition using convolutional neural networks », in *2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, fév. 2015, p. 1-4. DOI : 10.1109/SPICES.2015.7091490.

[255] J. REDMON, S. DIVVALA, R. GIRSHICK et A. FARHADI, « You Only Look Once : Unified, Real-Time Object Detection », in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, p. 779-788. DOI : 10.1109/CVPR.2016.91.

[256] J. REVAUD, P. WEINZAEPFEL, Z. HARCHAOUI et C. SCHMID, « DeepMatching : Hierarchical Deformable Dense Matching », *International Journal of Computer Vision*, t. 120, n° 3, p. 300-323, déc. 2016, ISSN : 0920-5691. DOI : 10.1007/s11263-016-0908-3.

[257] C. REVOL et M. JOURLIN, « A new minimum variance region growing algorithm for image segmentation », *Pattern Recognition Letters*, t. 18, n° 3, p. 249-258, 1997, ISSN : 0167-8655. DOI : 10.1016/S0167-8655(97)00012-3.

[258] L. RIBEIRO, C. M. OLIVEIRA, C. NEVES, J. D. RAMOS, H. FERREIRA et J. CUNHA-VAZ, « Screening for Diabetic Retinopathy in the Central Region of Portugal. Added Value of Automated 'Disease/No Disease' Grading », *Ophthalmologica*, t. 233, n° 2, p. 96-103, 2015, ISSN : 0030-3755. DOI : 10.1159/000368426.

[259] N. RITTER, R. OWENS, J. COOPER, R. H. EIKELBOOM et P. P. VAN SAARLOOS, « Registration of stereo and temporal images of the retina », *IEEE Trans. Med. Imag.*, t. 18, n° 5, p. 404-418, 1999. DOI : 10.1109/42.774168.

[260] G. ROGERS, O. CORBLET, T. FOURNEL et M. HEBERT, « Measurement of the diffusion of light within paper », *J. Opt. Soc. Am. A*, t. 36, n° 4, p. 636-640, avr. 2019. DOI : 10.1364/JOSAA.36.000636.

[261] C. RONSE et H. HEIJMANS, « The algebraic basis of mathematical morphology : II. Openings and closings », *CVGIP : Image Understanding*, t. 54, n° 1, p. 74-97, 1991, ISSN : 1049-9660. DOI : 10.1016/1049-9660(91)90076-2.

[262] F. ROSENBLATT, « The perceptron : A probabilistic model for information storage and organization in the brain. », *Psychological Review*, t. 65, n° 6, p. 386-408, nov. 1958, ISSN : 1939-1471. DOI : 10.1037/h0042519.

[263] M. ROUHANI, « Shape representation and registration using implicit functions », Director : Angel D. Sappa, thèse de doct., Centre de Visió per Computador (Bellaterra, Catalunya), Universitat Autònoma de Barcelona, Barcelona, Spain, 2012. En ligne : https://ddd.uab.cat/record/117856.

[264] M. ROUHANI et A. D. SAPPA, « Non-rigid Shape Registration : A Single Linear Least Squares Framework », in *Computer Vision – ECCV 2012*, A. FITZGIBBON, S. LAZEBNIK, P. PERONA, Y. SATO et C. SCHMID, éd., Berlin, Heidelberg : Springer, 2012, p. 264-277, ISBN : 978-3-642-33786-4. DOI : 10.1007/978-3-642-33786-4_20.

[265] F. ROUSSET, « Recalage d'images de pneumatiques déformés », Rapport de Master, CPE-Lyon et Université Jean Monnet, Saint-Etienne, Master Recherche Optique Image Vision, France, 2014.

[266] N. RYAN, C. HENEGHAN et P. DE CHAZAL, « Registration of digital retinal images using landmark correspondence by expectation maximization », *Image and Vision Computing*, t. 22, n° 11, p. 883-898, 2004, ISSN : 0262-8856. DOI : 10.1016/j.imavis.2004.04.004.

[267] P. SALEMBIER, A. OLIVERAS et L. GARRIDO, « Antiextensive connected operators for image and sequence processing », *IEEE Trans. Image Process.*, t. 7, n° 4, p. 555-570, avr. 1998, ISSN : 1057-7149. DOI : 10.1109/83.663500.

[268] P. SALEMBIER, « Comparison of some morphological segmentation algorithms based on contrast enhancement. Application to automatic defect detection », in *Proceedings of 5th European signal processing conference*, L. TORRES, E. MASGRAU et M. LAGUNAS, éd., Barcelona, Spain : Elsevier, 1990, p. 833-836. En ligne : www.eurasip.org/Proceedings/Eusipco/Eusipco1990/EUSIPCO90%7B%5C_%7DVolI.pdf.

[269] M. SANGALLI, S. BLUSSEAU, S. VELASCO-FORERO et J. ANGULO, « Scale Equivariant Neural Networks with Morphological Scale-Spaces », in *Discrete Geometry and Mathematical Morphology*, J. LINDBLAD, F. MALMBERG et N. SLADOJE, éd., Cham : Springer International Publishing, 2021, p. 483-495, ISBN : 978-3-030-76657-3. DOI : 10.1007/978-3-030-76657-3_35.

[270] D. SANTANA MAIA, J. COUSTY, L. NAJMAN et B. PERRET, « On the Probabilities of Hierarchical Watersheds », in *Mathematical Morphology and Its Applications to Signal and Image Processing*, B. BURGETH, A. KLEEFELD, B. NAEGEL, N. PASSAT et B. PERRET, éd., Cham : Springer International Publishing, 2019, p. 137-149, ISBN : 978-3-030-20867-7. DOI : 10.1007/978-3-030-20867-7_11.

[271] P. H. SCANLON, C. P. WILKINSON, S. J. ALDINGTON et D. R. MATTHEWS, *A practical manual of diabetic retinopathy management*. Wiley Online Library, 2009, ISBN : 9781405170352. DOI : 10.1002/9781444308174.

[272] J. SCHANDA, « CIE Colorimetry », in *Colorimetry : Understanding the CIE System*. John Wiley & Sons, Ltd, 2007, chap. 3, p. 25-78, ISBN : 9780470175637. DOI : 10.1002/9780470175637.ch3.

[273] U. SCHMIDT-ERFURTH, A. SADEGHIPOUR, B. S. GERENDAS, S. M. WALDSTEIN et H. BOGUNOVIĆ, « Artificial intelligence in retina », *Progress in Retinal and Eye Research*, t. 67, p. 1-29, 2018, ISSN : 1350-9462. DOI : 10.1016/j.preteyeres.2018.07.004.

[274] H. SCHREIER, J.-J. ORTEU et M. A. SUTTON, *Image Correlation for Shape, Motion and Deformation Measurements*. Boston, MA : Springer US, 2009, ISBN : 978-0-387-78746-6. DOI : 10.1007/978-0-387-78747-3.

[275] J. SERRA, *Image Analysis and Mathematical Morphology*. New York : Academic Press, 1982, t. 1, ISBN : 9780126372403. En ligne : https://books.google.fr/books?id=RQIUAQAAIAAJ.

[276] ——, *Alternating Sequential Filters*, J. SERRA, éd. Academic Press, 1988, t. 2, chap. 10, p. 203-216, ISBN : 9780126372410. En ligne : https://books.google.fr/books?id=BpdTAAAAYAAJ.

[277] J. SERRA, éd., *Image analysis and Mathematical Morphology : Theoretical advances*. Academic Press, 1988, t. 2, ISBN : 9780126372410. En ligne : https://books.google.fr/books?id=BpdTAAAAYAAJ.

[278] J. A. SETHIAN, « A fast marching level set method for monotonically advancing fronts », *Proceedings of the National Academy of Sciences*, t. 93, n° 4, p. 1591-1595, 1996, ISSN : 0027-8424. DOI : 10.1073/pnas.93.4.1591.

[279] ——, *Level Set Methods and Fast Marching Methods : Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, sér. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 1999, ISBN : 9780521645577. En ligne : https://books.google.fr/books?id=ErpOoynE4dIC.

[280] J. H. SHAH, M. SHARIF, M. RAZA, M. MURTAZA et SAEED-UR-REHMAN, « Robust Face Recognition Technique under Varying Illumination », *Journal of Applied Research and Technology*, t. 13, n° 1, p. 97-105, 2015, ISSN : 1665-6423. DOI : 10.1016/S1665-6423(15)30008-0.

[281] K. SIMONYAN, A. VEDALDI et A. ZISSERMAN, « Deep inside convolutional networks : visualising image classification models and saliency maps », ICLR, 2014, p. 1-8. En ligne : https://arxiv.org/abs/1312.6034.

[282] P. SOILLE, E. J. BREEN et R. JONES, « Recursive implementation of erosions and dilations along discrete lines at arbitrary angles », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 18, n° 5, p. 562-567, 1996. DOI : 10.1109/34.494646.

[283] P. SOILLE, « Generalized geodesy via geodesic time », *Pattern Recognition Letters*, t. 15, n° 12, p. 1235-1240, 1994, ISSN : 0167-8655. DOI : 10.1016/0167-8655(94)90113-9.

[284] ——, *Morphological Image Analysis : Principles and Applications*, 2e éd. New York : Springer-Verlag, 2003, ISBN : 3540429883. En ligne : www.springer.com/fr/book/9783540429883.

[285] A. SOTIRAS, C. DAVATZIKOS et N. PARAGIOS, « Deformable Medical Image Registration : A Survey », *IEEE Trans. Med. Imag.*, t. 32, n° 7, p. 1153-1190, 2013. DOI : 10.1109/TMI.2013.2265603.

[286] N. SRIVASTAVA, G. HINTON, A. KRIZHEVSKY, I. SUTSKEVER et R. SALAKHUTDINOV, « Dropout : A Simple Way to Prevent Neural Networks from Overfitting », *Journal of Machine Learning Research*, t. 15, n° 56, p. 1929-1958, 2014. En ligne : http://jmlr.org/papers/v15/srivastava14a.html.

[287] J. STAAL, M. D. ABRAMOFF, M. NIEMEIJER, M. A. VIERGEVER et B. VAN GINNEKEN, « Ridge-based vessel segmentation in color images of the retina », *IEEE Trans. Med. Imag.*, t. 23, n° 4, p. 501-509, avr. 2004. DOI : 10.1109/TMI.2004.825627.

[288] S. R. STERNBERG, « Parallel architectures for image processing », in *COMPSAC 79. Proc. Comput. Softw. and The IEEE Comput. Soc.'s Third Int. Appl. Conf.*, nov. 1979, p. 712-717. DOI : 10.1109/CMPSAC.1979.762586.

[289] ——, « Computer Architectures Specialized for Mathematical Morphology », in *Workshop on Algorithmically Specialized Computer Organizations*, Purdue University, West Lafayette, IN, USA, 1982.

[290] S. R. STERNBERG, « Grayscale morphology », *Computer Vision, Graphics, and Image Processing*, t. 35, n° 3, p. 333-355, 1986, ISSN : 0734-189X. DOI : 10.1016/0734-189X(86)90004-6.

[291] C. V. STEWART, CHIA-LING TSAI et B. ROYSAM, « The dual-bootstrap iterative closest point algorithm with application to retinal image registration », *IEEE Trans. Med. Imag.*, t. 22, n° 11, p. 1379-1394, 2003. DOI : 10.1109/TMI.2003.819276.

[292] W. STILES et J. BURCH, « N.P.L. Colour-matching Investigation : Final Report (1958) », *Optica Acta : International Journal of Optics*, t. 6, n° 1, p. 1-26, 1959. DOI : 10.1080/713826267.

[293] « Stiles & Burch individual 10-degree RGB Colour Matching Functions », Colour & Vision Research Laboratory, Institute of Ophthalmology, University College London. (1959), En ligne : http://cvrl.ucl.ac.uk/ (visité le 04/12/2020).

[294] P. STOCK, « Efficiency and Redundancy in Deep Learning Models : Theoretical Considerations and Practical Applications », thèse de doct., Université de Lyon, avr. 2021. En ligne : https://tel.archives-ouvertes.fr/tel-03208517.

[295] B. STURMFELS, *Solving Systems of Polynomial Equations*, sér. CBMS Regional Conference Series in Mathematics. Providence, Rhode Island, USA : American Mathematical Society, sept. 2002, t. 97, ISBN : 9780821832516. DOI : 10.1090/cbms/097.

[296] J. Z. SUN, G. I. WANG, V. K. GOYAL et L. R. VARSHNEY, « A framework for Bayesian optimality of psychophysical laws », *Journal of Mathematical Psychology*, t. 56, n° 6, p. 495-501, 2012, ISSN : 0022-2496. DOI : 10.1016/j.jmp.2012.08.002.

[297] H. TAUBENBÖCK, M. WIESNER, A. FELBIER, M. MARCONCINI, T. ESCH et S. DECH, « New dimensions of urban landscapes : The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data », *Applied Geography*, t. 47, p. 137-153, 2014, ISSN : 0143-6228. DOI : 10.1016/j.apgeog.2013.12.002.

[298] R. L. THOMAS, F. DUNSTAN, S. D. LUZIO, S. ROY CHOWDURY, S. L. HALE, R. V. NORTH, R. L. GIBBINS et D. R. OWENS, « Incidence of diabetic retinopathy in people with type 2 diabetes mellitus attending the Diabetic Retinopathy Screening Service for Wales : retrospective analysis », *BMJ*, t. 344, 2012, ISSN : 0959-8138. DOI : 10.1136/bmj.e874.

[299] B. THOMEE, D. A. SHAMMA, G. FRIEDLAND, B. ELIZALDE, K. NI, D. POLAND, D. BORTH et L.-J. LI, « YFCC100M : The New Data in Multimedia Research », *Commun. ACM*, t. 59, n° 2, p. 64-73, jan. 2016, ISSN : 0001-0782. DOI : 10.1145/2812802.

[300] D. S. W. TING, C. Y.-L. CHEUNG, G. LIM, G. S. W. TAN, N. D. QUANG, A. GAN, H. HAMZAH, R. GARCIA-FRANCO, I. Y. SAN YEO, S. Y. LEE, E. Y. M. WONG, C. SABANAYAGAM, M. BASKARAN, F. IBRAHIM, N. C. TAN, E. A. FINKELSTEIN, E. L. LAMOUREUX, I. Y. WONG, N. M. BRESSLER, S. SIVAPRASAD, R. VARMA, J. B. JONAS, M. G. HE, C.-Y. CHENG, G. C. M. CHEUNG, T. AUNG, W. HSU, M. L. LEE et T. Y. WONG, « Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes », *JAMA*, t. 318, n° 22, p. 2211-2223, déc. 2017, ISSN : 0098-7484. DOI : 10.1001/jama.2017.18152.

[301] E. TOLA, V. LEPETIT et P. FUA, « DAISY : An Efficient Dense Descriptor Applied to Wide-Baseline Stereo », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 32, n° 5, p. 815-830, 2010. DOI : 10.1109/TPAMI.2009.77.

[302] E. TRUCCO et A. VERRI, *Introductory techniques for 3-D computer vision*. Prentice Hall, 1998, ISBN : 978-0-13-261108-4. En ligne : https://books.google.fr/books?id=jfAeAQAAIAAJ.

[303] A. TUFAIL, V. V. KAPETANAKIS, S. SALAS-VEGA, C. EGAN, C. RUDISILL, C. G. OWEN, A. LEE, V. LOUW, J. ANDERSON, G. LIEW, L. BOLTER, C. BAILEY, S. SADDA, P. TAYLOR et A. R. RUDNICKA, « An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness », *Health Technology Assessment*, t. 20, n° 92, p. 1-72, déc. 2016, ISSN : 1366-5278. DOI : 10.3310/hta20920.

[304] M. VAN HERK, « A fast algorithm for local minimum and maximum filters on rectangular and octagonal kernels », *Pattern Recognition Letters*, t. 13, n° 7, p. 517-521, 1992, ISSN : 0167-8655. DOI : 10.1016/0167-8655(92)90069-C.

[305] L. R. VARSHNEY et J. Z. SUN, « Why do we perceive logarithmically ? », *Significance*, t. 10, n° 1, p. 28-31, 2013. DOI : 10.1111/j.1740-9713.2013.00636.x.

[306] A. VASWANI, N. SHAZEER, N. PARMAR, J. USZKOREIT, L. JONES, A. N. GOMEZ, Ł. KAISER et I. POLOSUKHIN, « Attention is All You Need », in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, sér. NIPS'17, Long Beach, California, USA : Curran Associates Inc., 2017, p. 6000-6010, ISBN : 9781510860964.

[307] F. VAUDELLE, J.-P. L'HUILLIER et M. L. ASKOURA, « Light source distribution and scattering phase function influence light transport in diffuse multi-layered media », *Optics Communications*, t. 392, p. 268-281, 2017, ISSN : 0030-4018. DOI : https://doi.org/10.1016/j.optcom.2017.02.001.

[308] A. VEDALDI et B. FULKERSON, *VLFeat : An Open and Portable Library of Computer Vision Algorithms*, www.vlfeat.org/, 2008.

[309] R. VERDÚ-MONEDERO, J. ANGULO et J. SERRA, « Anisotropic Morphological Filters With Spatially-Variant Structuring Elements Based on Image-Dependent Gradient Fields », *IEEE Trans. Image Process.*, t. 20, n° 1, p. 200-212, jan. 2011. DOI : 10.1109/TIP.2010.2056377.

[310] F. VIÉNOT et P. WALRAVEN, « Color-Matching Functions : Physiological Basis », in *Colorimetry : Understanding the CIE System*. John Wiley & Sons, Ltd, 2007, chap. 9, p. 219-243, ISBN : 9780470175637. DOI : 10.1002/9780470175637.ch9.

[311] M. A. VIERGEVER, J. A. MAINTZ, S. KLEIN, K. MURPHY, M. STARING et J. P. PLUIM, « A survey of medical image registration – under review », *Medical Image Analysis*, t. 33, p. 140-144, 2016, 20th anniversary of the Medical Image Analysis journal (MedIA), ISSN : 1361-8415. DOI : 10.1016/j.media.2016.06.030.

[312] T. VILBOUX, « La Stéréovision », Rapport de Master, Télécom Saint-Etienne, Université Jean Monnet, France, 2011.

[313] T. WALTER et J.-C. KLEIN, « Automatic Analysis of Color Fundus Photographs and Its Application to the Diagnosis of Diabetic Retinopathy », in *Handbook of Biomedical Image Analysis*, J. S. SURI, D. L. WILSON et S. LAXMINARAYAN, éd. Boston, MA : Springer US, 2005, t. II, p. 315-368, ISBN : 978-0-306-48606-7. DOI : 10.1007/0-306-48606-7_7.

[314] T. WALTER, P. MASSIN, A. ERGINAY, R. ORDONEZ, C. JEULIN et J.-C. KLEIN, « Automatic detection of microaneurysms in color fundus images », *Medical Image Analysis*, t. 11, n° 6, p. 555-566, 2007, ISSN : 1361-8415. DOI : 10.1016/j.media.2007.05.001.

[315] H. WANG, S. Z. LI et Y. WANG, « Face recognition under varying lighting conditions using self quotient image », in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings*, mai 2004, p. 819-824. DOI : 10.1109/AFGR.2004.1301635.

[316] L. WANG, Y. ZHANG et J. FENG, « On the Euclidean distance of images », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 27, n° 8, p. 1334-1339, août 2005, ISSN : 0162-8828. DOI : 10.1109/TPAMI.2005.165.

[317] C. WILKINSON, F. L. FERRIS, R. E. KLEIN, P. P. LEE, C. D. AGARDH, M. DAVIS, D. DILLS, A. KAMPIK, R. PARARAJASEGARAM et J. T. VERDAGUER, « Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales », eng, *Ophthalmology*, t. 110, n° 9, p. 1677-1682, sept. 2003, ISSN : 01616420. DOI : 10.1016/S0161-6420(03)00475-5.

[318] G. A. WILLIAMS, I. U. SCOTT, J. A. HALLER, A. M. MAGUIRE, D. MARCUS et H. MCDONALD, « Single-field fundus photography for diabetic retinopathy screening », eng, *Ophthalmology*, t. 111, n° 5, p. 1055-1062, mai 2004, ISSN : 01616420. DOI : 10.1016/j.ophtha.2004.02.004.

[319] G. WOLBERG, *Digital Image Warping*, 1st, sér. IEEE Computer Society Press monograph. Washington, DC, USA : Wiley, 1990, ISBN : 978-0-818-68944-4. En ligne : www.wiley.com/en-gb/Digital+Image+Warping-p-9780818689444.

[320] T. Y. WONG et N. M. BRESSLER, « Artificial Intelligence With Deep Learning Technology Looks Into Diabetic Retinopathy Screening », *JAMA*, t. 316, n° 22, p. 2366-2367, déc. 2016, ISSN : 0098-7484. DOI : 10.1001/jama.2016.17563.

[321] WONPIL YU, « An embedded camera lens distortion correction method for mobile computing applications », *IEEE Trans. Consum. Electron.*, t. 49, n° 4, p. 894-901, 2003. DOI : 10.1109/TCE.2003.1261171.

[322] Y. XU, T. GÉRAUD et L. NAJMAN, « Connected Filtering on Tree-Based Shape-Spaces », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 38, n° 6, p. 1126-1140, juin 2016, ISSN : 0162-8828. DOI : 10.1109/TPAMI.2015.2441070.

[323] T. YAMAGUCHI et S. HASHIMOTO, « Fast crack detection method for large-size concrete surface images using percolation-based image processing », *Machine Vision and Applications*, t. 21, n° 5, p. 797-809, août 2010, ISSN : 1432-1769. DOI : 10.1007/s00138-009-0189-8.

[324] T. YAMAGUCHI, S. NAKAMURA et S. HASHIMOTO, « An efficient crack detection method using percolation-based image processing », in *2008 3rd IEEE Conference on Industrial Electronics and Applications*, juin 2008, p. 1875-1880. DOI : 10.1109/ICIEA.2008.4582845.

[325] G. YANG et al., « Registration of Challenging Image Pairs : Initialization, Estimation, and Decision », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 29, n° 11, p. 1973-1989, nov. 2007, ISSN : 0162-8828. DOI : 10.1109/TPAMI.2007.1116.

[326] YFCC100M, *Stripes image from the Yahoo Flickr Creative Commons 100 Million YFCC100M dataset*, www.flickr.com/photos/34994230@N00/356103396/, Licence CC-BY-NC-SA 2.0, 2007.

[327] ——, *Image of a butterfly from the Yahoo Flickr Creative Commons 100 Million (YFCC100M) dataset*, www.flickr.com/photos/45563311@N04/4350683057/, Licence CC BY-NC-SA 2.0, 2010.

[328] X. YOU, B. FANG, Z. HE et Y. Y. TANG, « A Global-to-Local Matching Strategy for Registering Retinal Fundus Images », in *Pattern Recognition and Image Analysis*, J. S. MARQUES, N. PÉREZ DE LA BLANCA et P. PINA, éd., Berlin, Heidelberg : Springer, 2005, p. 259-267, ISBN : 978-3-540-32237-5. DOI : 10.1007/11492429_32.

[329] H. YU et J. FAN, « A novel segmentation method for uneven lighting image with noise injection based on non-local spatial information and intuitionistic fuzzy entropy », *EURASIP Journal on Advances in Signal Processing*, t. 2017, n° 1, p. 74, oct. 2017, ISSN : 1687-6180. DOI : 10.1186/s13634-017-0509-5.

[330] E. ZAHARESCU, « Morphological Enhancement of Medical Images in a Logarithmic Image Environment », in *2007 15th International Conference on Digital Signal Processing*, 2007, p. 171-174. DOI : 10.1109/ICDSP.2007.4288546.

[331] F. ZANA et J.-C. KLEIN, « Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation », *IEEE Trans. Image Process.*, t. 10, n° 7, p. 1010-1019, 2001. DOI : 10.1109/83.931095.

[332] F. ZANA et J. C. KLEIN, « A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform », *IEEE Trans. Med. Imag.*, t. 18, n° 5, p. 419-428, 1999. DOI : 10.1109/42.774169.

[333] R. ZÉBOUDJ et M. JOURLIN, « Filtrage, seuillage automatique, contraste et contours du pré-traitement à l'analyse d'image », Thèse de doctorat Sciences. Informatique, thèse de doct., Université de Saint-Etienne, France, 1988, 152 p. En ligne : www.sudoc.fr/043729355.

[334] Q. ZHANG, X. WANG, Y. N. WU, H. ZHOU et S.-C. ZHU, « Interpretable CNNs for Object Classification », *IEEE Trans. Pattern Anal. Mach. Intell.*, p. 1-1, 2020. DOI : 10.1109/TPAMI.2020.2982882.

[335] W. ZHANG, X. ZHAO, J. MORVAN et L. CHEN, « Improving Shadow Suppression for Illumination Robust Face Recognition », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 41, n° 3, p. 611-624, mars 2019, ISSN : 0162-8828. DOI : 10.1109/TPAMI.2018.2803179.

[336] X. ZHANG, G. THIBAULT, E. DECENCIÈRE, B. MARCOTEGUI, B. LAŸ, R. DANNO, G. CAZUGUEL, G. QUELLEC, M. LAMARD, P. MASSIN, A. CHABOUIS, Z. VICTOR et A. ERGINAY, « Exudate detection in color retinal images for mass screening of diabetic retinopathy », *Medical Image Analysis*, t. 18, n° 7, p. 1026-1043, 2014, ISSN : 1361-8415. DOI : 10.1016/j.media.2014.05.004.

[337] Z. ZHANG, « A flexible new technique for camera calibration », *IEEE Trans. Pattern Anal. Mach. Intell.*, t. 22, n° 11, p. 1330-1334, 2000. DOI : 10.1109/34.888718.

[338] Y. ZHAO, L. RADA, K. CHEN, S. P. HARDING et Y. ZHENG, « Automated Vessel Segmentation Using Infinite Perimeter Active Contour Model with Hybrid Region Information with Application to Retinal Images », *IEEE Trans. Med. Imag.*, t. 34, n° 9, p. 1797-1807, sept. 2015, ISSN : 1558-254X. DOI : 10.1109/TMI.2015.2409024.

[339] C. ZHU, B. ZOU, R. ZHAO, J. CUI, X. DUAN, Z. CHEN et Y. LIANG, « Retinal vessel segmentation in colour fundus images using Extreme Learning Machine », *Comput Med Imag Grap*, t. 55, p. 68-77, 2017, ISSN : 0895-6111. DOI : 10.1016/j.compmedimag.2016.05.004.

[340] K. ZUIDERVELD, « VIII.5. - Contrast Limited Adaptive Histogram Equalization », in *Graphics Gems*, P. S. HECKBERT, éd., Academic Press, 1994, p. 474-485, ISBN : 978-0-12-336156-1. DOI : 10.1016/B978-0-12-336156-1.50061-6.