



HAL
open science

Analyse automatique de la qualité des images issues de détecteurs plats à rayons X

Yann Le Meur

► **To cite this version:**

Yann Le Meur. Analyse automatique de la qualité des images issues de détecteurs plats à rayons X. Interface homme-machine [cs.HC]. Institut National Polytechnique de Grenoble - INPG, 2009. Français. NNT: . tel-00424459

HAL Id: tel-00424459

<https://theses.hal.science/tel-00424459>

Submitted on 15 Oct 2009

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.

INSTITUT POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

THÈSE

pour obtenir le grade de

DOCTEUR DE L'Institut polytechnique de Grenoble

Spécialité : « Signal Image Parole et Télécom »

préparée au Laboratoire Grenoble Image Signal Parole Automatique (GIPSA)

dans le cadre de l'École Doctorale « **Electronique , Electrotechnique, Automatique, Traitement du Signal** »

présentée et soutenue publiquement

par

Yann Le Meur

le 7 avril 2009

Titre :

**ANALYSE AUTOMATIQUE DE LA QUALITÉ DES IMAGES ISSUES DE DÉTECTEURS PLATS
À RAYONS X**

Directeur de thèse : Jocelyn CHANUSSOT

JURY

Pr.	Philippe Bolon	Président
Pr.	Philippe Réfrégier	Rapporteur
Pr.	Fabrice Mériaudeau	Rapporteur
Pr.	Jocelyn Chanussot	Directeur de thèse
M.	Jean-Michel Vignolle	Co-encadrant
Dr.	Jean-Luc Starck	Examineur
Pr.	Lionel Moisan	Examineur

Remerciements

J'ai effectué ma thèse entre novembre 2005 et novembre 2008 en passant la grande majorité de mon temps dans l'entreprise Trixell, effectuée à Moirans. Je tiens néanmoins à remercier les personnes du Gipsa-lab avec qui j'ai passé quelques moments, même si ils furent, je le concède, peu fréquents.

Mes premiers remerciements vont à mon directeur de thèse, le Pr. Jocelyn Chanussot. Je le remercie avant tout pour ses grandes qualités humaines dont il a fait preuve dans les moments difficiles. Il fut comme un phare dans la tempête, non pas grâce aux couleurs chatoyantes qu'il aime porter, mais plutôt grâce à ses conseils, à son esprit de synthèse et à sa capacité à remotiver ses doctorants aux moments opportuns. Je le remercie également pour m'avoir fait confiance dans ce travail de thèse. Pour ceci et pour avoir su, entre autres, parfaitement me guider pendant ces trois ans, je lui suis très reconnaissant.

Ce travail de thèse n'existerait pas sans l'apport des idées de Jean-Michel Vignolle. Son expérience et ses qualités techniques ont permis, au minimum d'initier, et parfois plus, une bonne partie des développements contenus dans ce manuscrit. Pour ceci et pour la confiance qu'il a placé en moi à mon arrivée à Trixell, je le remercie.

J'adresse également mes remerciements les plus chaleureux à mes collègues et amis de Trixell, en particulier les experts traitement et correction d'images que sont Cyril Marmajou et Benoit Candiard. Merci notamment pour votre accueil, votre bonne humeur, vos conseils précieux. Ils m'ont fait apprécier chaque journée passée à Trixell. Je remercie également les autres participants du groupe image, notamment Jérôme Moré pour son amitié et les nombreuses heures que nous avons passées à discuter de sujets variés (autour d'une pinte ou d'un café), Maude Laborie pour la fraîcheur et la bonne humeur qu'elle a apporté pendant plus d'un an au sein du Groupe Image, et Alliaume Rico pour sa sympathie, ses blagues, ses fonctions cachées dans les logiciels et sa coupe de cheveux.

Enfin je remercie ma mère et mon frère pour leur soutien et pour m'avoir supporté quelques week end pendant ces trois années. Un grand merci également à Laurence et Philippe qui ont su être là pour moi quand il le fallait, et pour les nombreux bons moments passés avec eux (et on ne va pas s'arrêter en si bon chemin !).

Pour finir, avant d'écrire ces remerciements, ma plus grande fierté, la plus belle chose qui me soit arrivée pendant ces trois ans, n'apparaissait pas dans ce manuscrit. Je répare donc ce paradoxe : merci ma chérie pour tout ce que tu m'as donné et pour tout ce que tu continues à me donner.

Table des matières

Introduction	3
I L'analyse de la qualité image dans un contexte industriel	5
1 Principe technologique des détecteurs plats numériques pour radiographie	7
1.1 La radiographie	9
1.1.1 Le rayonnement X	9
1.1.2 La radiographie numérique	10
1.1.3 Impacts de la source de rayons X dans les images acquises par des détecteurs numériques	11
1.2 La technologie des détecteurs plats	12
1.2.1 La couche scintillatrice	12
1.2.2 La dalle de photodiodes et l'électronique de lecture	13
1.3 Les applications radiographiques	14
1.4 La chaîne de correction d'images	18
1.5 Conclusions	18
2 Chaîne d'analyse automatique de la qualité image	21
2.1 La chaîne complète	23
2.2 Pré-traitement des images	26
2.2.1 Objectifs et enjeux	26
2.2.2 Méthode retenue	27
2.2.3 Limitations de la méthode de prétraitement	29
2.2.4 Conclusions et remarques sur le prétraitement	30
2.3 Détection des zones défectueuses	30
2.3.1 Objectifs et enjeux	30
2.3.2 Caractérisation des défauts à détecter et présentation des méthodes de détection	30
2.3.3 Choix des seuils de détection	34
2.3.4 Limitations de la méthode de détection	35
2.3.5 Conclusions et remarques sur la détection	35
2.4 Extraction des zones défectueuses	36
2.4.1 Objectifs et enjeux	36
2.4.2 Calcul adapté de la zone extraite	36
2.4.3 Conclusions et remarques sur l'extraction	38
2.5 Localisation des pixels défectueux	39
2.5.1 Objectifs et enjeux	39
2.5.2 Méthode proposée	39
2.5.3 Limitation de la méthode de localisation	39
2.5.4 Conclusion et remarques sur la localisation	40
2.6 Débruitage de la zone défectueuse	40

2.6.1	Objectifs et enjeux	40
2.6.2	Quelques méthodes existantes	42
2.6.3	Méthode proposée	44
2.6.4	Conclusions sur le débruitage	48
2.7	Introduction du défaut dans un contexte d'observation	49
2.7.1	Objectifs et enjeux	49
2.7.2	Calcul du contexte	49
2.7.3	Introduction du défaut dans le contexte	50
2.8	Mesure de visibilité du défaut	50
2.8.1	Objectifs et enjeux	50
2.8.2	Méthode proposée	52
2.8.3	Résultats obtenus et commentaires	52
2.9	Conclusions sur la chaîne de traitement	55
II	Détection de pixels défectueux	57
3	Détection automatique de défauts : objectifs et évaluation des performances	59
3.1	Objectifs de la détection	61
3.2	Méthode d'évaluation des performances de détection : introduction	61
3.3	La tâche d'inspection	62
3.4	L'analyse ROC	63
3.4.1	Définitions et courbes ROC	63
3.5	La comparaison des masques de détection	65
3.5.1	Comparaison des masques pixel à pixel	67
3.5.2	Comparaison des masques par la méthode de Theiler	69
3.5.3	Méthode proposée : comparaison souple	71
3.6	Exemples d'applications de la méthode souple	76
3.6.1	Mesure de performances d'algorithmes de détection par courbes ROC	76
3.6.2	Calibration automatique des algorithmes de détection	79
3.7	Commentaires et extensions possibles de la méthode	80
3.8	Conclusion	81
4	Méthodes de détection de défauts	83
4.1	Objectifs de la détection de défauts	85
4.1.1	Données de test	86
4.2	Introduction aux méthodes de détection	87
4.2.1	Les méthodes classiques de détection ou de segmentation	87
4.3	La méthode de base : seuillage direct de l'image	92
4.4	Méthodes à base de statistiques locales : moyenne, variance, skewness et kurtosis	92
4.4.1	Description de l'algorithme	94
4.4.2	Résultats sur défauts synthétiques	95
4.4.3	Conclusions sur les méthodes à base de statistiques locales	100
4.5	Détection par décomposition en ondelettes	100
4.5.1	Principe	100
4.5.2	Description de l'algorithme	101
4.5.3	Résultats sur défauts synthétiques	104
4.5.4	Conclusion sur la méthode de détection par décomposition en ondelettes	105
4.6	Détection par filtrage directionnel	106
4.6.1	Principe	106
4.6.2	Description de l'algorithme	106
4.6.3	Résultats sur défauts synthétiques	110

4.6.4	Conclusions sur la détection par filtrage directionnel	111
4.7	Détection par mesure d'anisotropie	112
4.7.1	Principe	112
4.7.2	Description de l'algorithme	112
4.7.3	Résultats sur défauts synthétiques	113
4.7.4	Conclusion sur la méthode de détection par calcul d'anisotropie	114
4.8	Détection par moyenne statistique pondérée	116
4.8.1	Description de l'algorithme	116
4.8.2	Résultats sur défauts synthétiques	117
4.8.3	Conclusions sur l'algorithme de détection par moyenne pondérée	117
4.9	Détection par snakes statistiques et rapports de vraisemblance	119
4.9.1	Snakes statistiques et rapport de vraisemblance : principes	119
4.9.2	Résultats sur défauts synthétiques	121
4.9.3	Conclusions sur la méthode de détection par snakes statistiques et MLRT	121
4.10	Synthèse	122
4.11	Conclusions et perspectives	124
III Mesure de visibilité des défauts dans une image		125
5	Mesure objective de visibilité de défauts	127
5.1	But d'une mesure objective de visibilité	129
5.2	Mesures objectives : éléments bibliographiques	130
5.2.1	Mesures classiques d'artefacts	130
5.2.2	Les mesures statistiques	133
5.2.3	Les mesures de type «Model Observer»	135
5.3	Application de mesures de visibilité sur les défauts synthétiques	138
5.3.1	Mesure de référence : énergie avec normalisation locale	139
5.3.2	Mesure de l'Hotelling Observer (HO)	140
5.3.3	L'Hotelling Observer non blanchisseur avec filtre de l'œil (NPWE)	140
5.4	Conclusions	143
Conclusions et perspectives		145
IV Annexes		151
A	Calculs des seuils théoriques pour un taux de fausse alarme fixé	153
A.1	Détermination de t_p	155
A.2	Détermination de t_m	156
A.3	Détermination de t_s	158
B	Etude psychovisuelle pour la fixation des seuils t_p, t_m et t_s	161
B.1	Choix du seuil t_p	163
B.2	Choix des seuils t_m et t_s	163
Bibliographie		169

Introduction

Introduction

Les énormes progrès de l'industrie micro-électronique depuis une vingtaine d'années ont permis le développement rapide des capteurs numériques. Aujourd'hui, ces capteurs sont présents dans de nombreux domaines d'application (vidéo-surveillance et sécurité, médical, grand public) et leur capacité en termes de résolution est en augmentation constante. Les capteurs récents affichent généralement des définitions supérieures à dix millions de pixels. Cette grande taille entraîne statistiquement une augmentation des risques de voir apparaître un défaut sur l'image acquise. Les industriels doivent donc améliorer dans le même temps leurs procédés de fabrication afin de conserver une bonne fiabilité de leurs capteurs, et donc une viabilité économique.

Le domaine de l'imagerie médicale a également profité de la numérisation des capteurs. En imagerie X, les systèmes d'acquisition à film photosensibles sont progressivement remplacés par des systèmes à capteurs numériques CCD ou CMOS, fournissant des images haute définition. Trixell, fabriquant de détecteurs plats numériques pour radiographie, dispose de détecteurs atteignant des définitions de l'ordre de 3000×3000 pixels. Le domaine médical impose que l'image en sortie du détecteur ne contienne aucun défaut, et soit d'une qualité suffisante pour permettre aux médecins et aux radiologues d'effectuer un diagnostic sûr. Cette forte contrainte de qualité impose à Trixell un contrôle drastique des images produites par ses détecteurs.

Cette thèse s'inscrit dans le contexte industriel de contrôle de la qualité des images issues de détecteurs radiologiques plats. En particulier, nous avons cherché à analyser le contenu des images à la recherche de structures se différenciant du bruit de fond des images plates (images acquises à l'aide d'une dose de rayons X mais en l'absence de sujet, voir figure 0.1), structures que nous appelons défauts.

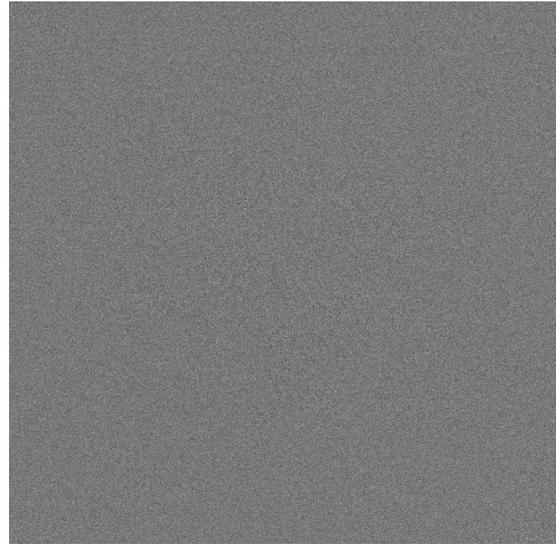
Le contrôle fréquent des images à la recherche de tels défauts génère à Trixell une quantité considérable d'images à analyser. Il existe deux cas de figure différents pouvant amener à une telle analyse d'images :

- contrôle de la qualité au cours de la fabrication : la fabrication d'un détecteur est un processus complexe pouvant durer une semaine. Au cours de ce processus, la qualité du détecteur est analysée, notamment à l'aide d'images de test. La production d'un seul détecteur requiert l'analyse de centaines d'images de test haute résolution. Il faut donc mettre au point des outils efficaces d'analyse des défauts présents dans ces images afin d'automatiser la tâche de contrôle, qui ne peut être faite dans un délai raisonnable, et dans de bonnes conditions d'efficacité, par un opérateur humain.
- Contrôle *a posteriori* : en fonction des retours clients ou de problèmes de qualité image détectés en interne, les experts image de Trixell sont fréquemment amenés à analyser *a posteriori* la base des images à la recherche de défauts particuliers. Depuis sa création, Trixell a produit environ 25 000 détecteurs, tous ayant généré plusieurs centaines d'images. L'enjeu d'une tâche d'analyse *a posteriori* est donc d'être capable d'analyser automatiquement les images, et d'effectuer une synthèse des résultats obtenus. Il est notamment intéressant d'analyser chaque image et d'en extraire les défauts, puis d'affecter une mesure à ces défauts. Cette mesure doit de plus être corrélée à l'impact qu'ont les défauts sur la qualité de l'image.

La quantité d'images à analyser impose des contraintes de performance pour le système de détection de défauts que l'on veut mettre au point : ce système doit avoir un taux de fausse alarme extrêmement faible (une fausse alarme est une détection d'un défaut dans l'image alors qu'il n'y en a pas). En effet, chaque fausse alarme entraîne soit le dérangement inutile d'un opérateur (dans le cas d'un contrôle de qualité au cours de la fabrication), soit de mauvaises analyses (dans le cas du contrôle *a posteriori*).



(a) Image clinique



(b) Image plate

FIGURE 0.1 – Image clinique et image plate d'un détecteur numérique classique (examen radiologique conventionnel)

Cette thèse s'inscrit dans un projet industriel complet d'analyse de la qualité des images issues de détecteurs numériques pour radiographie. Le manuscrit sera organisé en trois parties distinctes. En partant de considérations générales, nous nous attacherons à l'étude de deux processus critiques pour l'analyse de qualité d'une image.

1. La première partie est donc introductive. Elle présentera d'abord le contexte de la thèse (capteurs radiographiques, types d'images, chaîne de correction). Ensuite, nous décrirons la chaîne d'analyse automatique de la qualité des images. Cette chaîne détaille, de l'image d'origine jusqu'aux résultats des défauts quantifiés en fonction de leur visibilité, le processus d'analyse élaboré dans le cadre de la thèse. Ce processus est prévu pour une exploitation industrielle, *i.e.* pour l'analyse de milliers d'images hautes résolutions issues des détecteurs pour radiographie.
2. La seconde partie sera un éclairage sur un des points critiques de la chaîne de traitement : la détection de défauts dans l'image. Plusieurs méthodes de détection y seront présentées, s'appuyant à la fois sur des considérations statistiques et spatiales. Dans le même temps, nous exposerons une méthode de validation de performances de ces algorithmes de détection, adaptée à notre contexte, et ne nécessitant aucun réglage : la méthode du Receiver Operating Characteristic (ROC) souple.
3. Enfin, nous finirons par une étude de mesure objective de visibilité, s'appuyant sur une campagne d'évaluation subjective menée auprès d'experts qualité image à Trixell. Les mesures objectives de qualité sont l'objet de nombreuses études académiques, aucune n'ayant donné de résultats vraiment convaincants dans tous les cas. Notre approche pour cette étude est de s'appuyer sur des données d'évaluation subjective, ainsi que sur la connaissance du type d'artefacts pouvant apparaître sur les images, pour construire une mesure simple de visibilité des défauts.

Première partie

**L'analyse de la qualité image dans
un contexte industriel**

Chapitre 1

Principe technologique des détecteurs plats numériques pour radiographie

Résumé

Dans ce chapitre, le procédé de fabrication du rayonnement X ainsi que la technologie utilisée pour la fabrication des détecteurs numériques sont brièvement exposés. Les impacts de ces éléments sur l'image sont également illustrés. Grâce à la connaissance des constituants d'un détecteur, des corrections sont possibles, et donc appliquées, afin de s'affranchir de phénomènes de dégradation d'image se retrouvant systématiquement sur toute image acquise par le détecteur. On obtient ainsi une image corrigée contenant uniquement du bruit ainsi que d'éventuels défauts dont la détection et la mesure font l'objet de cette thèse.

Sommaire

1.1	La radiographie	9
1.1.1	Le rayonnement X	9
1.1.2	La radiographie numérique	10
1.1.3	Impacts de la source de rayons X dans les images acquises par des détecteurs numériques	11
1.2	La technologie des détecteurs plats	12
1.2.1	La couche scintillatrice	12
1.2.2	La dalle de photodiodes et l'électronique de lecture	13
1.3	Les applications radiographiques	14
1.4	La chaîne de correction d'images	18
1.5	Conclusions	18

Les détecteurs numériques pour radiographie Trixell sont issus d'un processus de fabrication complexe mettant en œuvre aussi bien de l'électronique que de la mécanique et de la chimie. Ce processus, par sa complexité et par sa nécessité d'intervention humaine, n'est pas parfait et un certain nombre de défauts sont présents dans le produit en fin de chaîne. L'expérience permet, à partir d'un défaut constaté sur une image, de retrouver son origine dans le processus de fabrication. Ce raisonnement s'appuie sur une bonne connaissance du procédé de fabrication, et des dérives qu'il peut subir. Un détecteur est globalement constitué de trois éléments critiques : une couche de cristal appelé scintillateur, une couche de constituant électronique sensible à la lumière (la dalle de photodiodes) et l'électronique permettant la lecture et la conversion numérique du signal. Dans ce chapitre, après avoir présenté globalement le principe de la radiographie, une présentation de ces trois couches technologiques sera faite. Elle sera conclue par un tour d'horizon des systèmes de radiographie, et notamment par un détail des différentes classes d'application de chaque système, impliquant chacune des contraintes différentes sur la qualité des images.

1.1 La radiographie

1.1.1 Le rayonnement X

Depuis la découverte du rayonnement X par Wilhelm Röntgen [82] et le premier cliché radiographique (voir figure 1.1) qu'il réalisa en 1895, les outils de diagnostic médicaux se sont multipliés. Cependant, la radiographie reste l'outil de base grâce à sa simplicité, sa polyvalence et son aspect bon marché comparé aux autres techniques.

La radiographie par rayons X consiste en l'exposition d'un patient à des rayons X, ce qui permet d'obtenir par ombre chinoise des clichés de ses os ou de ses organes. Selon la densité de la partie du corps traversée, le rayonnement est plus ou moins absorbé, et on arrive ainsi à voir par transparence les tissus sur lesquels le médecin ou le radiologue cherche des lésions.

Cette capacité à observer toutes les parties du corps fait de la radiographie un outil polyvalent et donc couramment utilisé. Mais son succès s'explique aussi par sa plus grande facilité de mise en œuvre et son ancienneté par rapport aux autres technologies d'imagerie médicale.

Un peu plus de 100 ans après la première radiographie de Röntgen, le principe de détection n'a finalement pas évolué puisqu'il repose toujours sur le film radiographique, sensible aux rayons X comme une pellicule photographique est sensible à la lumière. Pour obtenir un cliché après exposition, il suffit, toujours de façon analogue à la photographie, de développer le film.

Le rayonnement X, dans des applications médicales, est produit dans un tube à rayons X par bombardement d'une anode métallique (en général du tungstène) par un faisceau d'électrons. Lors de ce bombardement, les électrons subissent une décélération au contact de l'anode, produisant ainsi des rayons X.



FIGURE 1.1 – Première radiographie médicale : main d'Anna-Berthe Röntgen, 22 décembre 1895

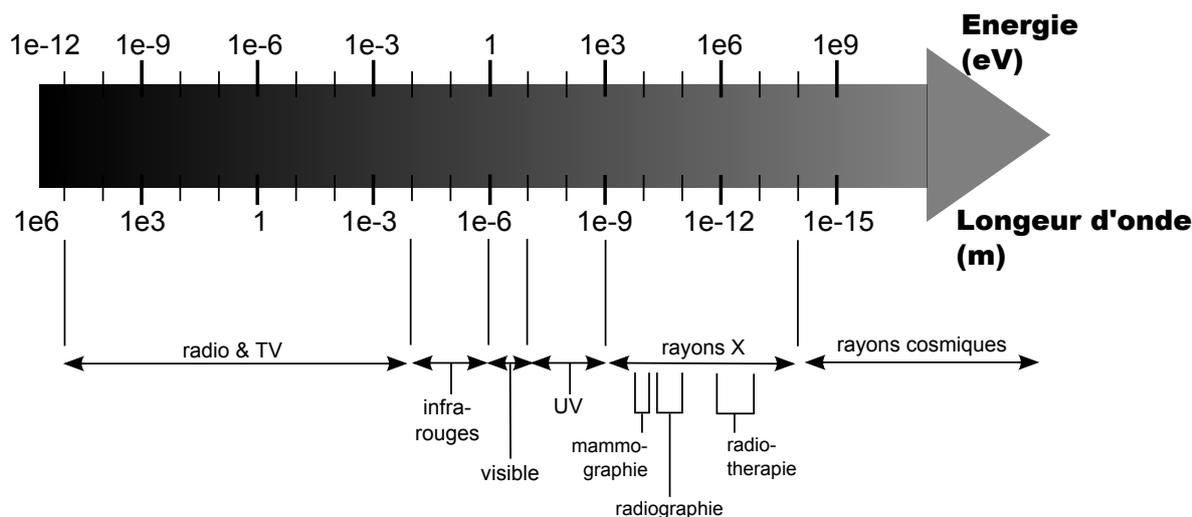


FIGURE 1.2 – Spectre des ondes électromagnétiques : les rayons X sont situés aux niveaux d'énergie élevés du spectre

En imagerie médicale, les rayons X utilisés sont dans la gamme d'énergie 10-100keV , ce qui correspond à une longueur d'onde de $0.1nm$ à $0.01nm$ (voir figure 1.2). Une description complète du procédé de fabrication, des caractéristiques, et de l'utilisation du rayonnement X dans un contexte médical peut être trouvée dans [7].

1.1.2 La radiographie numérique

Depuis quelques années, la radiologie a amorcé son virage vers le tout numérique avec l'avènement des premiers détecteurs plats. La société Trixell joue un rôle moteur dans le développement de cette technologie. Les principaux avantages de ces détecteurs numériques plats, face aux radiographies à base de film, sont :

- une dose d'exposition plus faible du patient aux rayons X, à qualité d'image identique : les détecteurs numériques ont une sensibilité supérieure, ce qui permet de réduire la dose. Ceci est une caractéristique particulièrement stratégique pour les détecteurs numériques, notamment lorsque l'on considère une utilisation dans le cadre d'une intervention chirurgicale, où le patient et le médecin peuvent être exposés à une dose de rayonnement cumulée relativement importante si la sensibilité du matériel d'acquisition est insuffisante,
- une facilité de transmission et de stockage : les clichés radiographiques sont conservés directement sur support informatique, ce qui permet une transmission rapide des données et un stockage plus pratique, et surtout moins polluant, comparé aux traditionnels films,
- les multiples traitements d'image disponibles qui permettent de faciliter le diagnostic : les outils de réglage de contraste permettent au médecin de pouvoir instantanément visualiser plusieurs types d'organes (tissus mous et os), alors qu'avec un film classique, il aurait fallu faire plusieurs clichés en adaptant la dose d'exposition. Le fait de disposer directement de la source numérisée de l'examen a également permis de développer les techniques d'amélioration d'image (débruitage, rehaussement de contours notamment) ou de post-traitement avancé (reconstruction 3D, visualisation couleur par acquisition en «Dual Energy»).

Les multiples avantages des détecteurs numériques leur permettent de prendre petit à petit le dessus vis-à-vis des systèmes à film, qu'ils sont amenés à remplacer dans des contextes où le coût des systèmes n'est pas une contrainte trop forte.

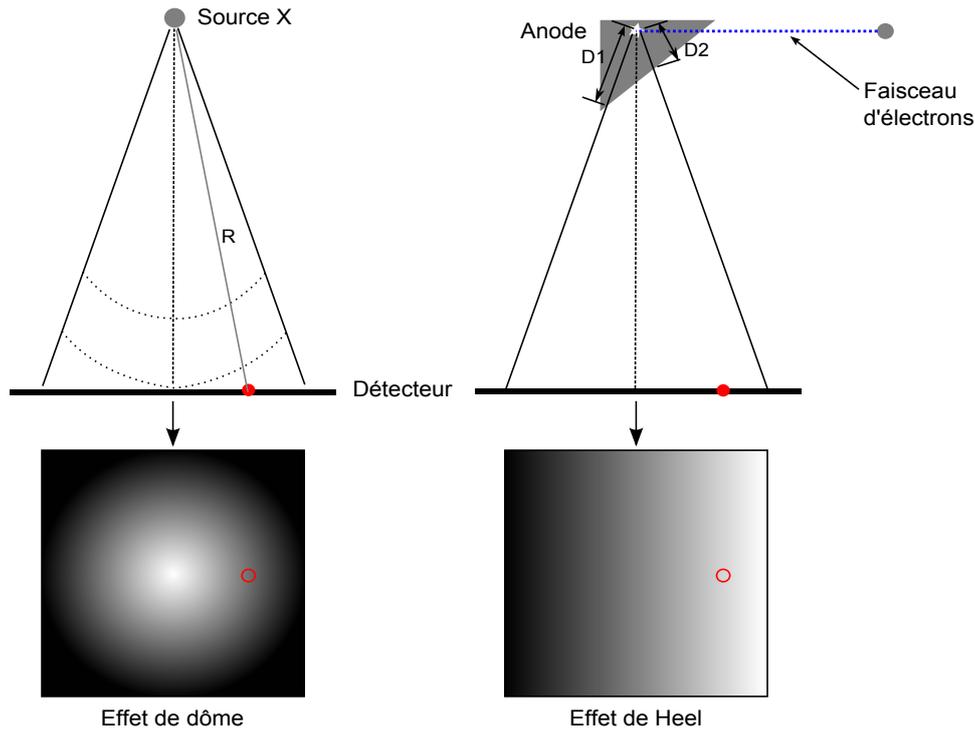


FIGURE 1.3 – Illustration de l’effet de dôme et de l’effet de Heel. L’effet de dôme provoque un dégradé radial car les pixels du détecteur ne sont pas équidistants de la source. L’effet de Heel résulte d’une contrainte de design du tube et provoque un effet de dégradé linéaire : les rayons X produits dans l’anode ne traversent pas une épaisseur de métal équivalente ($D1 > D2$) et sont donc absorbés différemment, d’où le dégradé apparaissant sur le détecteur.

1.1.3 Impacts de la source de rayons X dans les images acquises par des détecteurs numériques

Certaines caractéristiques physiques du tube à rayons X peuvent avoir des effets sur l’image acquise à l’aide du rayonnement X provenant de ce tube. L’impact le plus notable est l’effet conjugué de source ponctuelle de rayonnement et de l’effet de Heel qui entraînent un dégradé de luminosité sur les images acquises à l’aide des détecteurs plats.

Le premier effet, appelé « effet de dôme », est une conséquence de la nature ponctuelle de la source de rayonnement qui n’éclaire pas uniformément la surface plane formée par le détecteur. En effet, le niveau L d’un pixel du détecteur éclairé par une dose Dx de rayons X suit :

$$L \propto \frac{Dx}{R^2} \quad (1.1)$$

avec R la distance du pixel à la source.

Le détecteur étant plan, tous ses pixels ne sont pas équidistants de la source d’où l’effet de dôme observé qui est une décroissance du signal (*i.e.* de la luminosité) sur les bords du détecteur (voir figure 1.3).

Le second effet, appelé effet de Heel, résulte d’une contrainte de design de l’anode du tube à rayons X. Celle-ci doit présenter un léger angle par rapport au faisceau d’électrons venant la frapper. Cet angle permet de diminuer l’échauffement sur la partie illuminée de l’anode : les électrons sont répartis sur une surface plus grande par rapport à une configuration où l’anode serait perpendiculaire au faisceau électronique. La dissipation thermique dans l’anode est rendue ainsi plus efficace. La figure 1.3 illustre l’effet de Heel et le dégradé linéaire ainsi observé sur l’image.

Cet effet du générateur a des répercussions sur les images en sortie des détecteurs plats, comme l’illustre la figure 1.4. On observe sur image plate une fonction d’éclairement qui vient se multiplier

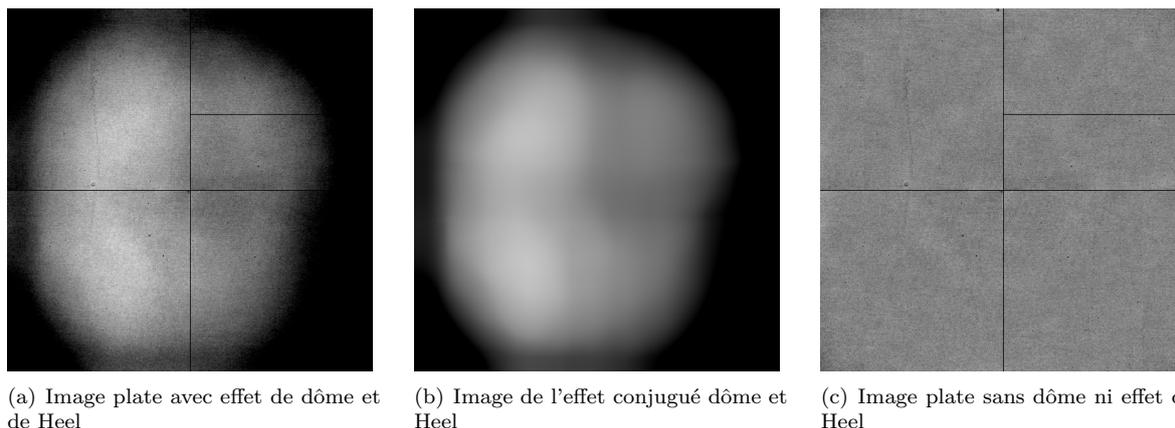


FIGURE 1.4 – Illustration de l'effet de dôme et de l'effet de Heel sur une image en sortie du détecteur : un dôme d'éclairement apparaît clairement sur l'image non corrigée (figure 1.4(a)). L'effet peut facilement être corrigé et isolé (figure 1.4(b)) afin d'obtenir une image d'intensité globalement homogène (figure 1.4(c)).

à l'image d'origine. Cet effet peut facilement être compensé par correction (par exemple, acquisition d'images de référence, et division des images en sortie par cette référence) car l'effet est spatialement stable d'une image à l'autre. On peut également estimer l'effet par filtrage basse-fréquence comme l'illustre la figure 1.4(b). Dans la suite du manuscrit, nous verrons que nous pouvons nous affranchir de cet effet, et nous concentrer sur l'analyse des structures non prévisibles apparaissant dans les images.

D'autres structures connues apparaissent classiquement sur les images issues des détecteurs numériques. En plus des effets d'éclairement causés par la source, que nous venons de voir, les différents constituants du détecteur peuvent être à l'origine de structures, maîtrisées ou non. Nous allons voir dans les prochaines sections les trois principaux constituants d'un détecteur, en indiquant à chaque fois le type de défaut qu'il peut générer.

1.2 La technologie des détecteurs plats

Dans le domaine des détecteurs plats, deux grands principes de détection coexistent :

- la détection directe : les rayons X sont convertis directement en charge électrique au sein d'une substance photoconductrice. Le rayonnement X incident a un effet ionisant sur le matériau (principe inverse de la production de rayons X), généralement du Sélénium Se, et provoque une libération d'électrons. La tension électrique appliquée aux bornes du matériau permet une migration des charges qui sont stockées avant d'être lues et converties en valeur numérique
- la détection indirecte : c'est la technologie adoptée par Trixell. Les rayons X sont transformés en photons visibles au sein du matériau scintillant (oxyde de Gadolinium dopé Terbium ou iodure de Césium dopé Thallium). Les photons visibles sont ensuite convertis en charges électriques puis en signal numérique.

Les détecteurs plats numériques produits par Trixell s'appuient sur la technologie de détection indirecte. Cette technologie a donné naissance à l'un des premiers détecteurs plats en 1995 [67]. La figure 1.5 résume le schéma de détection de la technologie Trixell. Les trois principaux composants y sont illustrés : le scintillateur, la matrice de détection et l'électronique de lecture.

1.2.1 La couche scintillatrice

La couche scintillatrice est utilisée dans la technologie à détection indirecte pour transformer les photons X en photons visibles. Plusieurs technologies existent pour réaliser cette couche, celle retenue

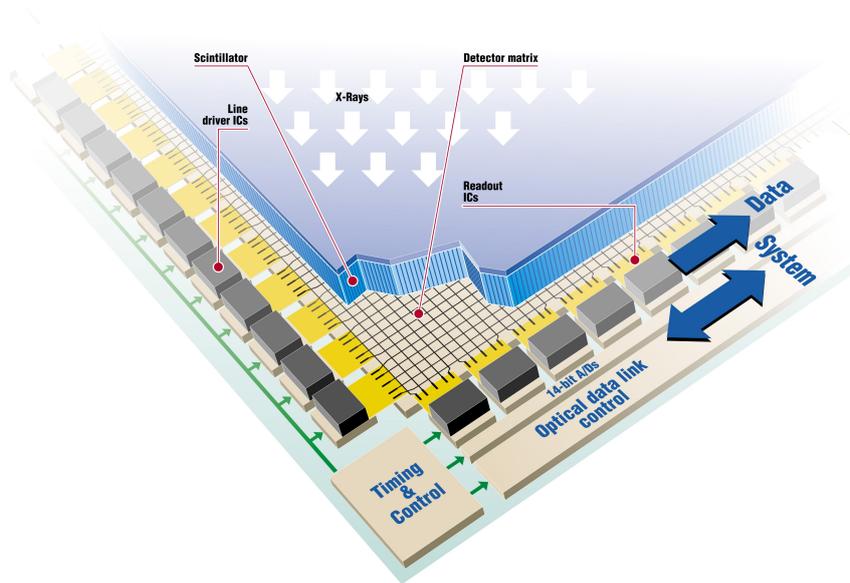


FIGURE 1.5 – Coupe d'un détecteur numérique plat. Les rayons X incidents (partie supérieure) sont transformés en photons visibles par le scintillateur. Au sein de la matrice de photodiodes, les photons visibles sont convertis en charges électriques et lus par l'électronique de lecture.

pour les détecteurs numériques Trixell s'appuie sur un cristal d'iodure de Césium dopé Thallium [31, 19]. L'avantage de cette technologie à base d'iodure de Césium est sa structure en aiguille (voir figure 1.6), qui, comme une fibre optique, permet de guider les photons. Ceci permet d'améliorer la résolution spatiale en limitant la diffusion des photons dans le matériau. Cette structure est obtenue par évaporation : le matériau est chauffé jusqu'à évaporation et se condense sur un substrat qui sert de base de croissance pour les aiguilles. L'effet de scintillation est causé par la nature ionisante du rayonnement X qui vient créer des changements de niveau d'énergie des électrons au sein des atomes du matériau. L'énergie libérée par ce changement de niveau est ainsi rayonnée sous la forme de photons. Le dopage au Thallium du cristal est un élément essentiel qui permet de déplacer la bande d'émission du scintillateur aux alentours de $550nm$, ce qui correspond à la plage de fonctionnement optimal de la matrice de photodiodes (plage d'efficacité maximale). D'autres dopages par des constituants moins nocifs que le Thallium existent, mais la longueur d'onde des photons émis correspond moins bien à la bande de détection des photodiodes, ce qui dégrade l'efficacité de la chaîne d'acquisition.

La plupart des défauts que nous recherchons résultent d'un effet lié au scintillateur. En effet, sa fabrication et son intégration au détecteur sont des opérations complexes et sensibles. De plus, le CsI, même si il dispose de nombreux avantages d'un point de vue optique, est un matériau peu stable dans le temps, et sensible aux variations de température. Les défauts résultant du scintillateur et de son évolution dans le temps sont donc mal déterminés : ils peuvent prendre une forme et un aspect aléatoires.

1.2.2 La dalle de photodiodes et l'électronique de lecture

Comme illustré sur la figure 1.5, le second constituant principal d'un détecteur est la dalle de photodiodes permettant de passer d'un signal lumineux (des photons visibles) à un signal électrique (des charges accumulées dans un condensateur de la photodiode). Chaque photodiode dispose d'un transistor de type TFT qui sert d'interrupteur pour le transfert des charges dans l'électronique de lecture. Les composants de lecture viennent lire les charges selon le principe suivant :

1. choix de la colonne : un signal vient activer l'entrée du composant de lecture sur la colonne à lire.

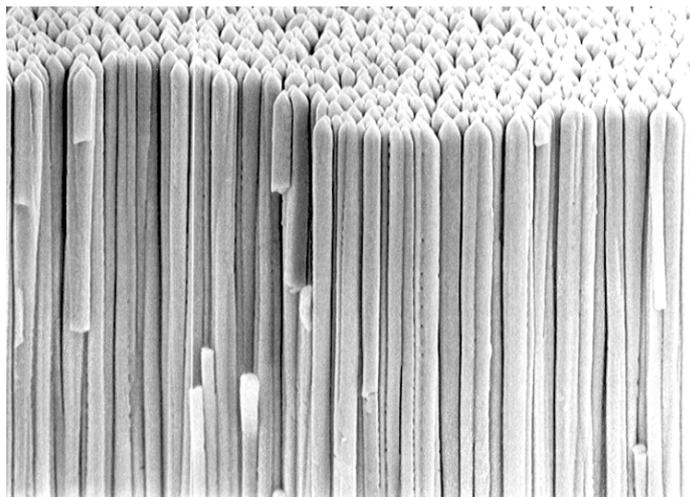


FIGURE 1.6 – Structure en aiguilles du scintillateur vue au microscope électronique : les aiguilles de Iodure de Césium dopée Thallium (CsI :Tl) mesurent en moyenne $6\mu m$ de diamètre et 500 à $600\mu m$ de long

Un même composant est chargé de la lecture d’une centaine de colonnes.

2. lecture des lignes : un séquenceur ouvre successivement les transistors des diodes de chaque ligne, entraînant le transfert des charges qui y étaient stockées.
3. les charges sont numérisées afin d’affecter une valeur en niveau de gris à chaque pixel.

Cette étape peut générer des structures sur les images, qui doivent être corrigées, comme nous verrons dans la prochaine section. En particulier, on peut observer une corrélation en ligne et en colonne du bruit composant une image plate. Cette corrélation spatiale du bruit est caractéristique d’un effet de l’électronique de lecture. La figure 1.7 représente les densités spectrales de bruit d’un détecteur à deux doses différentes. On rappelle que la densité spectrale de puissance $\Gamma(u, v)$ d’une image $f(x, y)$, de transformée de Fourier $F(u, v)$, est définie par :

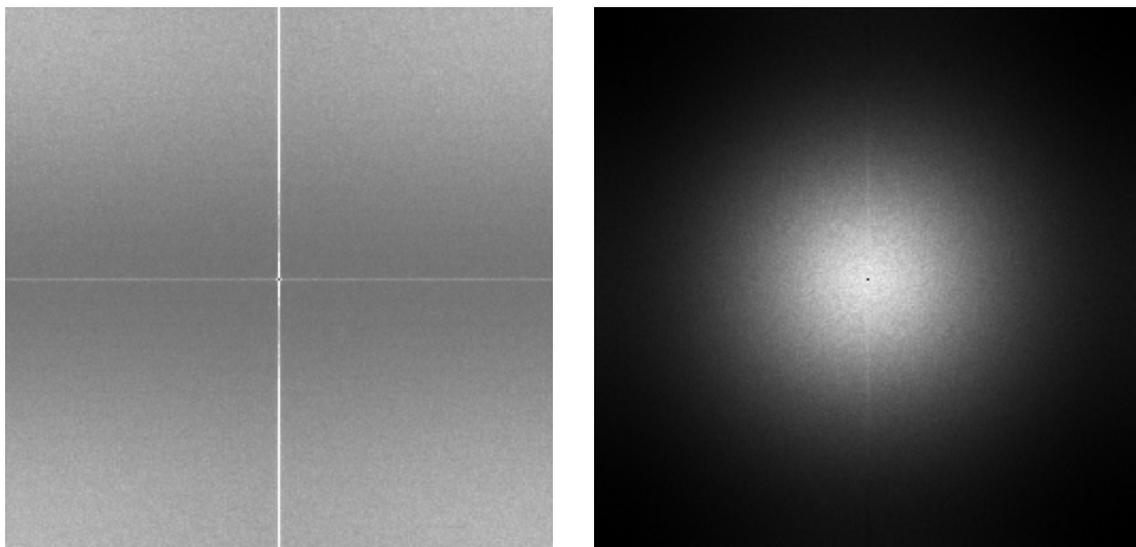
$$\Gamma(u, v) = |F(u, v)|^2 \quad (1.2)$$

Pour déterminer la densité spectrale de bruit dans notre cas, nous l’avons estimée à partir de plusieurs réalisations de ce bruit.

La première densité spectrale (figure 1.7(a)) a été déterminée en l’absence de rayons X incidents, et la seconde pour une dose d’exposition classique (figure 1.7(b)). On voit clairement sur la figure 1.7(a) le bruit de l’électronique de lecture qui présente une corrélation forte en colonne et en ligne (respectivement, ligne horizontale et ligne verticale dans les images de densité spectrale). En présence de rayonnement X, le bruit électronique est couplé à un bruit résultant de l’effet combiné de la source de rayons X et du scintillateur. Le bruit d’une image brute en sortie du détecteur n’est donc pas strictement blanc. Néanmoins, nous verrons dans les chapitres suivants que la chaîne de correction ainsi que le souci de traiter le cas de la détection dans un contexte général nous pousseront à traiter le cas d’un bruit de fond blanc dans les images.

1.3 Les applications radiographiques

Triaxel dispose d’une large gamme de détecteurs radiographiques, dépendant du type d’applications pour lesquelles ils sont destinés. On distingue notamment les détecteurs statiques, utilisés en radiologie conventionnelle et les détecteurs dynamiques (pouvant acquérir plusieurs images en un intervalle de temps réduit) plutôt destinés à de la radiologie interventionnelle. La table 1.1 résume les caractéristiques



(a) Densité Spectrale de Bruit sans X. Réglage du contraste : valeur médiane $\pm 200\%$

(b) Densité Spectrale de Bruit à $1\mu\text{Gy}$. Réglage du contraste : valeur médiane $\pm 90\%$

FIGURE 1.7 – Densités Spectrales de Bruit d'un détecteur numérique



FIGURE 1.8 – Gamme des produits Trixell (de gauche à droite, de l'arrière plan au premier plan) : Pixium RF, Pixium 4600, Pixium Portable, Pixium 4800 et Pixium 4700

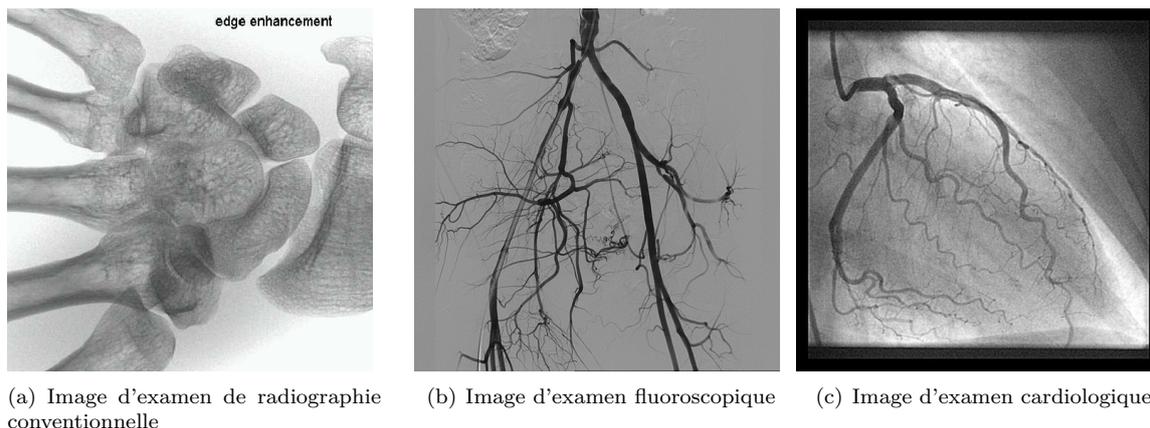


FIGURE 1.9 – Différents types d'examen en radiologie (de gauche à droite) : radiographie d'une main, fluoroscopie pour le suivi de la bifurcation aortique au niveau de l'abdomen, examen dynamique du coeur par injection de produit de contraste

principales des détecteurs Trixell actuellement sur le marché. La gamme est composée (voir figure 1.8) :

- d'un détecteur radiographique, le Pixium 4600 : premier détecteur Pixium mis sur le marché, il est naturellement destiné, grâce à sa grande taille, à la radiographie conventionnelle (radiographie de membres ou de thorax, voir figure 1.9(a)).
- D'un détecteur pour la fluoroscopie, le Pixium 4700 : détecteur dynamique de taille moyenne qui permet de réaliser des examens comme le suivi du flux sanguin dans des artères ou des vaisseaux (après injection d'un produit de contraste, voir figure 1.9(b)).
- D'un détecteur pour cardiologie, le Pixium 4800. Ce détecteur de petite taille (taille adaptée au coeur) permet d'acquérir des images en mode dynamique. Sa rapidité d'acquisition permet son utilisation pour les interventions de chirurgie cardiaque : réparation d'artères autour du coeur : interventions suite à occlusion, anévrisme ou dissection aortique, voir figure 1.9(c).
- D'un détecteur polyvalent, le Pixium RF. Produit faisant partie de la nouvelle gamme, il se décline en deux versions venant remplacer le Pixium 4600 et le Pixium 4700. Sa plus grande taille et son pas d'échantillonnage (pixel pitch) plus faible sont les principales améliorations apportées.
- D'un détecteur portable, le Pixium Portable. Produit utilisé pour la radiographie conventionnelle, comme le Pixium 4600, sa principale caractéristique repose dans son utilisation portable : il transmet ses images par liaison WiFi et dispose d'une poignée de transport et d'une batterie pour fonctionner en mode autonome. Il est principalement destiné aux examens urgents pour lesquels on ne peut pas manipuler ou transporter le patient vers une salle de radiographie (service des urgences par exemple).

Les détecteurs numériques ajoutent à la composante spatiale une composante temporelle. Ceci a des implications dans le cadre du contrôle qualité et la détection de défaut puisque la perception des défauts diffère grandement dans le cas des séquences d'images. La détection de défaut sur image statique et sur images dynamiques étant deux problèmes relativement différents, nous avons décidé de nous placer uniquement dans le cas statique pour cette thèse. On peut *a priori* émettre l'hypothèse que les défauts sont plus difficilement détectables dans le cas statique, car on ne dispose pas de l'information temporelle. Les méthodes de détection sur images statiques peuvent de plus s'appliquer sur images dynamiques (en les appliquant image par image, et en procédant à une fusion de chaque détection obtenue), alors qu'il nous paraît difficile d'appliquer à une seule image les méthodes élaborées sur le cas dynamique. Nous avons donc choisi de restreindre l'étude au cas statique.

	Pixium 4600	Pixium 4700	Pixium 4800
Champs d'application	Radiographie	Fluoroscopie	Cardioscopie
Début de production	2000	2004	2002
Surface imagée	43 × 43cm	40 × 30cm	18 × 18cm
Résolution (en pixels)	3000 × 3000	2480 × 1910	956 × 954
Dynamique des images (bits/pixel)	14	14	14
Pixel pitch (μm)	143	154	184
Fréquence d'acquisition max (en Hz)		60	60
Poids (en kg)	17.2	20	10
	Pixium RF	Pixium Portable	
Champs d'application	Radio/Fluoro	Radiographie	
Début de production	2007	2008	
Surface imagée	43 × 43cm	34 × 43cm	
Résolution (en pixels)	2880 × 2881	2372 × 3000	
Dynamique des images (bits/pixel)	16	16	
Pixel pitch (μm)	148	144	
Fréquence d'acquisition max (en Hz)	30		
Poids (en kg)	25	4.8	

TABLE 1.1 – Caractéristiques des détecteurs Pixium

1.4 La chaîne de correction d'images

Du fait de l'électronique et des effets de dôme ou de l'effet de Heel précédemment décrits, les images en sortie du détecteur présentent des structures caractéristiques se retrouvant de manière systématique. On observe ainsi des bandes verticales sur l'image résultant de la dispersion de gains des composants électroniques. En effet, avec $p_{i,j}$ le niveau d'un pixel dans l'image de sortie, et d la dose de rayonnement X que reçoit ce pixel (qu'on suppose constante sur tout le détecteur), on peut modéliser, en l'absence de bruit, le système par l'équation suivante (voir aussi la figure 1.10(a)) :

$$p_{i,j} = G_{i,j} \times d + O_{i,j} \quad (1.3)$$

La quantité $G_{i,j}$ est le gain du pixel, et $O_{i,j}$ est son offset. Afin de garantir l'homogénéité spatiale des niveaux sur le détecteur, il faut supprimer la dépendance en (i, j) de $p_{i,j}$ dans la formule précédente. Un des buts de la chaîne de correction, en plus de corriger les défauts, est de rétablir l'homogénéité spatiale de l'image en choisissant pour tous les pixels un gain et un offset uniques. Ceci est réalisé en deux temps (voir figure 1.10) :

1. une image de référence $Oref_{i,j}$ acquise sans exposition est retirée à l'image à corriger $I_{i,j}$. Ceci a pour conséquence de créer une image où $\forall(i, j) O_{i,j} = 0$
2. l'image résultante $I_{i,j}^{Oc}$ est divisée par une image de référence de gain $Gref_{i,j}$ de manière à avoir $\forall(i, j) G_{i,j} = 1$

L'image corrigée d'offset et de gain s'écrit donc :

$$I_{i,j}^{OcGc} = \frac{I_{i,j}}{Gref_{i,j}} - Oref_{i,j} \implies \forall(i, j) p_{i,j} \propto d \quad (1.4)$$

Une dernière étape, la correction de défauts, permet de corriger les pixels défectueux (pixel 3 sur la figure 1.10(a)) détectés par une méthode basique de type statistique sur le niveau des pixels. L'image obtenue en fin de chaîne est une image corrigée d'offset, de gain et de défauts.

1.5 Conclusions

La fabrication d'un détecteur est un processus complexe qui ne peut être exempt de défauts et de dérives de procédé. Afin de garantir le principe de «zéro défaut», on doit analyser les images automatiquement de manière à détecter et mesurer chaque défaut. On distingue principalement deux catégories de défauts :

- les dérives dans les structures typiques de l'image. Ce sont des défauts causés par une valeur atypique dans les structures typiques de l'image (structures dues à l'électronique ou au scintillateur) qui peuvent être corrigées par la chaîne de correction. Ces défauts n'évoluant pas, ils sont alors invisibles sur l'image finale entièrement corrigée. Ces défauts de structures connues ne font pas l'objet de cette thèse, puisque des méthodes spécifiques *ad hoc* existent déjà pour les détecter et les mesurer (en général, il s'agit de comparer leur valeur à une valeur spécifiée) .
- les défauts sur image corrigée : des pixels qui ne sont pas détectés comme défectueux par les méthodes actuelles, restant visibles sur l'image corrigée. Cette thèse traite de la détection et de la mesure de ces défauts de structure *a priori* inconnue.

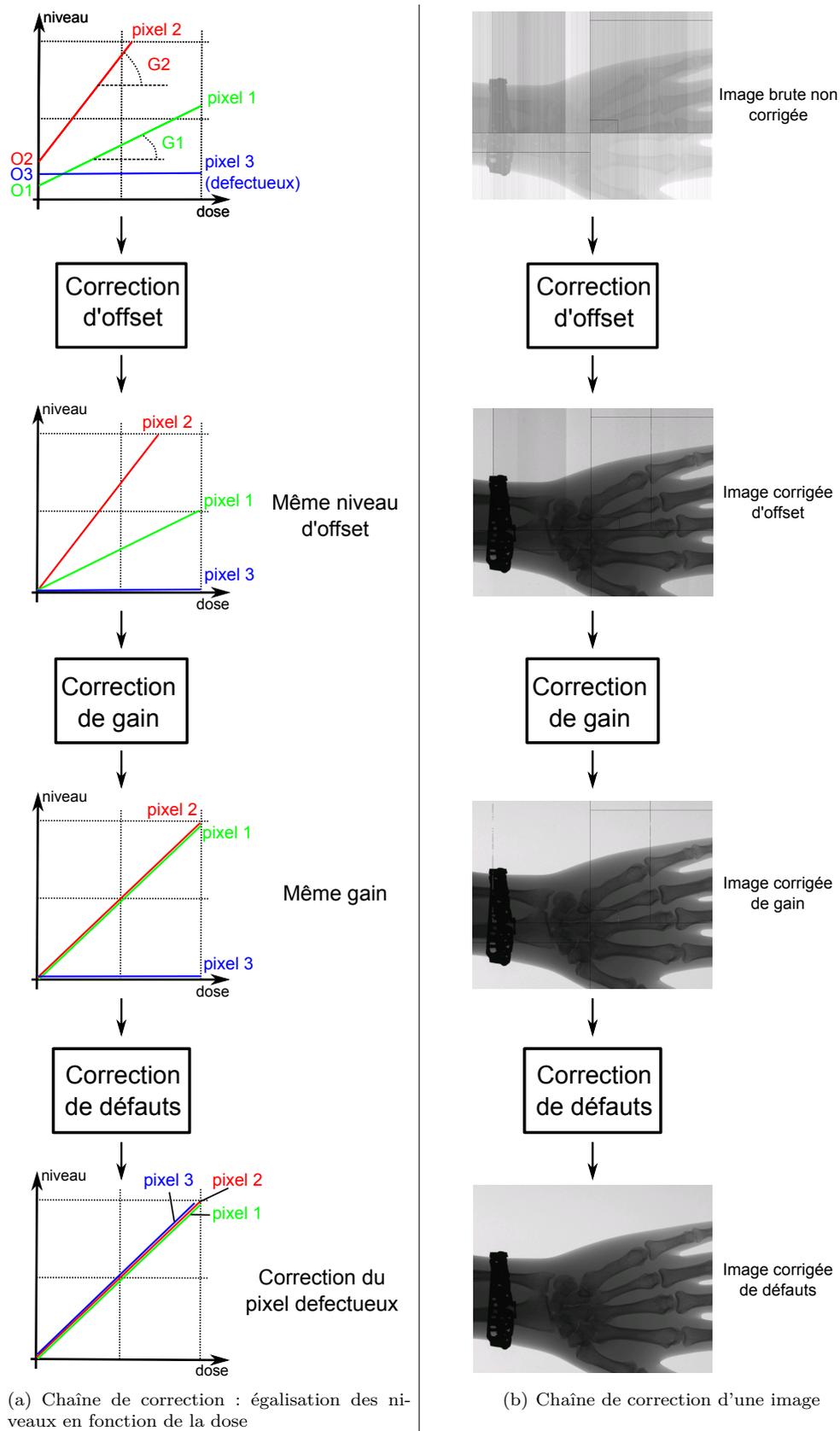


FIGURE 1.10 – Illustration de la chaîne de correction : le principe de la correction consiste à annuler l'offset (ordonnées à l'origine des droites de la figure 1.10(a)) et ramener le gain de tous les pixels (pentes des droites de la figure 1.10(a)) à 1. Ceci revient à corriger les structures observées sur l'image non corrigée de la figure 1.10(b)). La correction de défaut vient corriger tous les pixels qui n'avaient pas un comportement normal (pixel 3 sur la figure 1.10(a)).

Chapitre 2

Chaîne d'analyse automatique de la qualité image

Résumé

Les images issues des détecteurs numériques sont de grande taille - jusqu'à 3000×3000 pixels et peuvent contenir plusieurs défauts de taille et de nature variées. Cette grande taille et le nombre de défauts présents rendent impossible le contrôle et l'analyse de la qualité de ces images par un observateur humain. L'objectif de la chaîne d'analyse automatique de la qualité est donc de disposer d'une procédure automatique permettant de produire une cartographie des défauts présents, et pouvant donner une quantification de l'impact visuel de chaque défaut de manière à produire une appréciation aussi fine que celle qui serait donnée par un observateur humain. Dans ce chapitre, des méthodes permettant de répondre à ces deux objectifs fondamentaux sont décrites et commentées. Des résultats sur des exemples réels viennent illustrer chacune de ces méthodes, et mettre en avant leurs limitations. Ce chapitre présente la chaîne complète, comme elle a été implémentée, à l'aide de traitements simples et robustes. Les procédés permettant la détection et la mesure de visibilité seront traités avec plus de détails dans les chapitres 4 et 5.

Sommaire

2.1	La chaîne complète	23
2.2	Pré-traitement des images	26
2.3	Détection des zones défectueuses	30
2.4	Extraction des zones défectueuses	36
2.5	Localisation des pixels défectueux	39
2.6	Débruitage de la zone défectueuse	40
2.7	Introduction du défaut dans un contexte d'observation	49
2.8	Mesure de visibilité du défaut	50
2.9	Conclusions sur la chaîne de traitement	55

Comme nous l'avons vu dans le chapitre précédent, l'analyse automatique du contenu est un enjeu de grande importance afin de garantir la bonne qualité des images qui seront utilisées par l'utilisateur final, mais également de capitaliser la connaissance des problèmes de qualité d'image afin de tendre vers une amélioration des produits. Dans ce chapitre, nous allons décrire en détails la tâche d'analyse automatique de la qualité d'image, comme elle est vue dans la problématique industrielle servant de cadre à cette thèse. Cette analyse automatique de la qualité des images exploite une image brute, en sortie du détecteur, et exploite toute l'information qui diffère du bruit de fond de cette image. Cette information est exploitée principalement sous deux formes, qui sont deux sous-produits de la chaîne complète d'analyse de la qualité d'une image :

- cartographie de tous les défauts image : chaque défaut dans l'image doit être localisé afin d'être observé, identifié, et archivé. Ce processus nécessite une détection des défauts, et une capture de chaque défaut dans le contexte ayant servi pour sa détection.
- caractérisation et mesure de la visibilité de chaque défaut dans son contexte : à chaque défaut, il faut associer une valeur quantitative permettant de traduire l'impact visuel du défaut dans un contexte particulier, qui peut être différent du contexte dans lequel il a été détecté.

Ce chapitre décrit la chaîne d'analyse complète permettant notamment d'obtenir ces deux informations à partir d'une image Trixell. Dans les paragraphes suivants, nous allons détailler chaque étape de cette chaîne en précisant à chaque fois son but, les enjeux associés à l'étape, et une illustration de réalisation sur un exemple concret d'analyse d'image. Les étapes de localisation des pixels défectueux et de mesure de visibilité ont fait l'objet d'études approfondies et seront donc décrites plus en détails dans les chapitres 4 et 5. Pour les autres étapes de la chaîne, nous proposons une méthode permettant de répondre au problème. D'autres méthodes sont bien entendu possibles, mais la thèse ayant principalement pour objet la détection et la mesure de qualité, nous avons choisi de répondre à certains problèmes de manière pragmatique, en étant conscient que les solutions ainsi proposées pourront être améliorées.

2.1 La chaîne complète

La chaîne d'analyse de la qualité d'une image, construite dans le cadre de la thèse, est présentée sur la fig 2.1.

L'image d'origine présentée sur cette figure est une image non-corrigée présentant des structures connues comme le dôme, l'effet de Heel et des structures corrélées en ligne et en colonne. L'origine de ces structures a été présentée dans le chapitre précédent. Ces structures ont une géométrie commune dans toutes les images à analyser. La première étape de la chaîne de traitement consiste donc à récupérer ces structures et les traiter à part. Cette étape de pré-traitement sera détaillée dans la section 2.2. On extrait trois structures typiques de l'image d'origine : la structure basse fréquence (où on retrouve notamment les effets de dôme et de Heel) et des structures de type colonnes et lignes constantes, dues à l'électronique de lecture. Ces structures sont en pratique corrigées par la chaîne de correction.

Une fois ces structures enlevées de l'image d'origine, nous entrons dans la phase d'analyse des défauts inconnus. La première étape de cette phase est une détection des zones potentiellement défectueuses (voir section 2.3). Cette étape définit des régions relativement étendues (pas de détection au pixel près) qui serviront ensuite à une analyse plus fine des défauts ainsi détectés. Sur la fig 2.1, nous avons ainsi détecté quelques zones défectueuses qui apparaissent en rouge. Nous verrons que la méthode choisie, en première approximation, souffre d'un manque de sensibilité : des défauts bien visibles malgré un rapport signal à bruit éventuellement faible, ne sont pas ou mal détectés. L'amélioration de cette étape a donc fait l'objet d'une étude plus approfondie qui sera exposée dans les chapitres 3 et 4.

Ces régions sont passées en entrée d'un bloc d'extraction de régions d'intérêt (ROI) qui est en charge de la définition d'une zone de capture du défaut (voir section 2.4). La ROI extraite doit à la fois permettre la visualisation du défaut par un observateur, mais aussi définir une zone suffisamment large pour que l'étape suivante de localisation au pixel près soit efficace. En effet, il faut pouvoir estimer des statistiques sur le bruit de fond, la ROI extraite doit donc contenir suffisamment de pixels de bruit pour accomplir cette tâche. Enfin, les ROI extraites doivent être de taille raisonnable pour des raisons pratiques de

stockage des résultats. La fig 2.1 montre un exemple de trois défauts capturés sur l'image d'origine. Nous avons choisi, par souci de clarté de l'illustration, des défauts très visibles. En pratique ces défauts sont parfaitement corrigés par la chaîne de correction. Ils peuvent être dus à des défauts de surface sur le détecteur (capot rayé, particules présentes sur le capot lors de l'acquisition de l'image). Dans tous les cas, ces défauts ne seront évidemment pas présents sur l'image qui sera utilisée par le client.

L'étape suivante consiste en une localisation fine des défauts. Cette étape a un objectif double : confirmer la présence du défaut indiquée par le bloc de détection, et être capable d'identifier le défaut présent au pixel près. Une méthode simple de détection pixellique a été utilisée en première approximation pour définir la chaîne d'analyse (section 2.5). Ainsi, la fig 2.1 présente le résultat de la localisation sur les trois défauts présentés.

L'étape de débruitage (section 2.6) s'appuie ici sur les résultats de la localisation pour fournir une estimation du défaut débruité. L'objectif de cette étape est de s'affranchir du contexte dans lequel a été détecté chaque défaut, de manière à avoir une représentation générique du défaut qui sera fournie aux étapes suivantes. Cette étape s'apparente à une étape de pré-traitement en vue de la caractérisation du défaut. Une méthode de débruitage simple et adaptée à notre problème est proposée dans le cadre de cette étude. On voit ainsi que les trois défauts présentés sont assez bien débruités par cette méthode qui a d'ailleurs éliminé complètement le bruit de fond de l'image d'origine. La quantité considérable de travaux sur le problème de débruitage peut évidemment être mise à profit pour améliorer cette étape. Nous avons donné ici une méthode pragmatique qui répond à l'objectif de l'étape de débruitage tout en étant conscient que cette proposition est largement perfectible.

Les résultats du débruitage sont interprétés par deux étapes aux objectifs différents :

- l'étape de classification a pour but de classer chaque défaut dans une catégorie prédéfinie. Il s'agit d'une étape de caractérisation haut niveau de chaque défaut. Ici aussi, une quantité importante de travaux est disponible. Nous avons choisi de laisser cette étape ouverte. Un nombre très important de techniques de classification a été défini dans la littérature. Pour notre problème, il faudra en choisir une qui permette en particulier de gérer facilement l'apparition de nouvelles classes. Ceci arrive en pratique fréquemment, lors de l'apparition d'un nouveau défaut.
- l'étape d'introduction dans un nouveau contexte : cette étape utilise chaque défaut débruité afin de l'introduire dans un nouveau contexte. Sur la fig 2.1, un nouveau contexte a été défini (il s'agit d'un bruit différent de celui de l'image d'origine) et les trois défauts ont été introduits dans ce nouveau contexte par une opération simple (section 2.7). Cette opération permet de pouvoir facilement changer de contexte sans avoir à refaire intégralement une détection des défauts. Les contextes possibles ici peuvent être des bruits de natures différentes car correspondant à des doses d'acquisition différentes, des images cliniques ou d'autres images techniques utilisées en interne (images de test ou de calibration, avec mires, *etc*).

Enfin, la dernière étape de la chaîne d'analyse permet de quantifier la visibilité de chaque défaut dans le contexte choisi. Une mesure de visibilité est associée à chaque défaut, traduisant l'impact visuel qu'il a sur la qualité de l'image, sachant qu'une image de qualité optimale ne contient que du bruit de fond. Cette étape a également fait l'objet d'une étude plus poussée qui sera développée dans le chapitre 5. Une première méthode simple pour définir cette mesure est donnée en section 2.8. Cette mesure, illustrée sur les exemples de la chaîne d'analyse, répond en première approximation au besoin. En effet, on constate que les trois exemples de défauts sont bien quantifiés par la mesure puisque le classement des défauts par ordre de visibilité croissante donne bien le résultat attendu par l'observation visuelle de ces défauts.

Les méthodes présentées dans la suite ont été testées sur images réelles de taille 3000×3000 pixels. Les tests des méthodes de pré-traitement et de détection ont été conduits sur 6 images de ce type. Par la suite, une seule image a été retenue, et les tests ont alors été menés sur les 1800 défauts identifiés par la méthode de détection sur cette image.

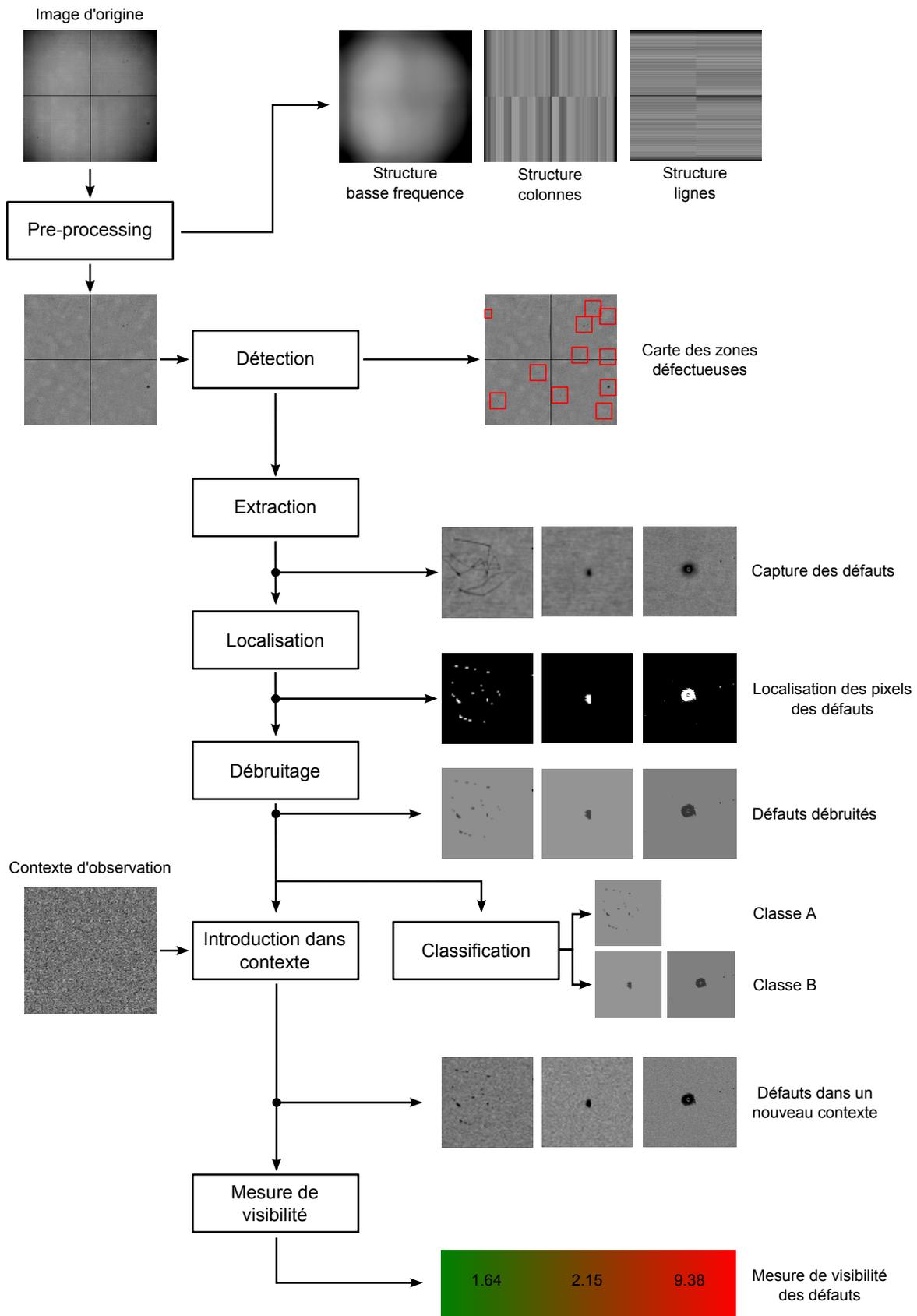


FIGURE 2.1 – Illustration de la chaîne d'analyse de qualité sur un défaut réel. Le but principal de l'analyse de qualité des images est de fournir une capture des défauts et une mesure de leur visibilité. Chaque étape (hormis l'étape de classification) est traitée à l'aide de méthodes qui seront décrites dans les sections 2.2 à 2.8

2.2 Pré-traitement des images

2.2.1 Objectifs et enjeux

Comme exposé dans le chapitre précédent, chaque détecteur numérique produit des images avec des structures caractéristiques, qui sont apparentées à des structures parasites même si elles sont bien maîtrisées et corrigées par la chaîne de correction. En effet, on s'attend idéalement à obtenir une image, sous exposition de rayons X, de niveau totalement constant et proportionnel à la dose reçue par le détecteur. La présence de bruit¹ provoqué par des événements aléatoires au sein de la source ou de l'électronique de lecture est inévitable. Néanmoins, on s'attend à ce que ce bruit ne présente aucune corrélation spatiale forte due à la présence de structures dans l'image. Hors, nous avons vu (fig 1.7) que ce n'est pas le cas. En effet, des structures récurrentes apparaissent. Elles peuvent généralement se décomposer en trois contributions :

- une structure basse fréquence causée par les effets de dôme et de Heel, et éventuellement par la physique du détecteur,
- des effets corrélés en colonne provoqués par la dispersion des composants de l'électronique de lecture et le bruit électronique en général, dans les composants de lecture de chaque colonne,
- des effets corrélés en ligne provoqués par le bruit de lecture des composants d'adressage des pixels en ligne.

D'une manière générale, on peut modéliser l'image brute $o(i, j)$ de dimension $N \times N$ de la manière suivante :

$$\forall (i, j) \in N \times N \quad o(i, j) = e(i, j) \cdot (l(i) \cdot c(j) \cdot p(i, j)) \quad (2.1)$$

où $e(i, j)$ est la composante basse fréquence (dépendance spatiale en ligne et en colonne), $l(i)$ est la composante corrélée en ligne (ne dépend que de l'indice ligne i dans l'image) et $c(j)$ est la composante corrélée en colonne (ne dépend que de l'indice colonne j dans l'image). Dans notre étude, nous nous focalisons sur l'analyse d'images après correction d'offset. Le choix d'un modèle multiplicatif s'impose car les structures observées sont alors dépendantes de la dose reçue, par construction physique du détecteur. Les structures et défauts n'évoluant pas avec la dose ont été supprimés par la correction d'offset et ne feront pas partie de l'analyse qualité².

Le but de l'étape de pré-traitement est d'estimer au mieux $e(i, j)$, $l(i)$ et $c(j)$ et de retrouver ainsi l'image $p(i, j)$ débarrassée des structures récurrentes. Estimer au mieux signifie que ce processus d'estimation et de compensation des structures ne crée aucun artefact dans l'image pré-traitée $p(i, j)$. L'image pré-traitée ne contient alors plus que les artefacts n'entrant pas dans le cadre des trois composantes, ainsi qu'un bruit de fond. Les composantes sont modélisées comme des grandeurs déterministes (non-bruitées).

On prend les hypothèses suivantes :

- les composantes $e(i, j)$, $l(i)$, $c(j)$ et $p(i, j)$ sont des réalisations de trois variables aléatoires indépendantes E , L , C et P
- on suppose les variables aléatoires L , C et P stationnaires au premier ordre, de moyenne respective μ_L , μ_C et μ_P
- on suppose que la variable aléatoire E est stationnaire sur une plage donnée. Ceci implique :

$$\exists (i_0, i_1, j_0, j_1) \text{ tels que } \forall (i, j) \quad i_0 \leq i \leq i_1, j_0 \leq j \leq j_1 :$$

$$\mathbf{E} \{e(i, j)\} = \mu_E \quad (2.2)$$

où l'opérateur $\mathbf{E} \{ \cdot \}$ est l'espérance mathématique.

1. En pratique, ce bruit n'est pas blanc comme nous l'avons vu dans le chapitre précédent. Il est plutôt basse fréquence.

2. En pratique, si on veut étudier ces défauts d'offset, sur image sans rayonnement X, il faut adopter un modèle multiplicatif. De plus, la composante $e(i, j)$ n'est plus présente dans le modèle car elle est uniquement due à la source de rayonnement

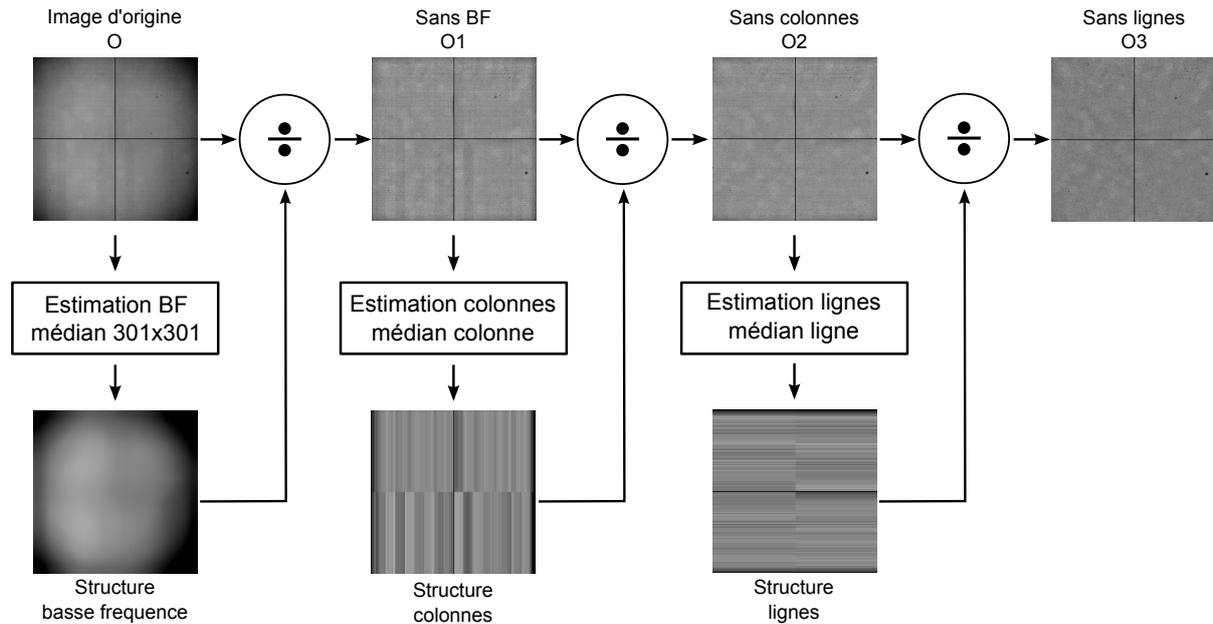


FIGURE 2.2 – Méthode de pré-traitement. Les structures basse-fréquence, colonne et ligne sont estimées successivement par filtrage médian sur un support adapté. Après chaque estimation, la structure est enlevée de l'image par division.

Les hypothèses prises ici vont permettre une estimation simple et séquentielle de chacune des composantes. Le principe général est en effet d'estimer une composante sur un support où cette composante est normalement constante, sachant que sur ce même support on a plusieurs réalisations des autres composantes. Sur ce support, on estime alors à chaque fois la valeur de la composante, aux moyennes près des autres composantes.

Les principales difficultés pour accomplir cette tâche de pré-traitement sont inhérentes au contenu de l'image $p(i, j)$. En effet, cette image contient en général, en plus du bruit de fond, des défauts et d'autres structures non connues *a priori* qui peuvent gêner l'estimation des composantes et créer ainsi des artefacts artificiellement. Si, sur certains supports d'estimation, ces défauts sont suffisamment présents, ils vont venir biaiser l'estimation de la composante ce qui mènera à l'introduction d'un artefact de décomposition, comme nous le verrons en section 2.2.3.

2.2.2 Méthode retenue

La méthode de base pour effectuer cette étape de pré-traitement consiste à estimer successivement chaque composante à l'aide de méthode d'estimation robuste, en adaptant la fenêtre d'estimation à la géométrie de la composante à estimer (figure 2.2). En effet, dans le cadre du modèle défini par l'équation 2.1, on peut réaliser une estimation en trois étapes :

1. calculer $\hat{e}(i, j)$, l'estimation de $e(i, j)$ et l'enlever de l'image initiale en calculant l'image intermédiaire $\widehat{O1}(i, j) = \frac{o(i, j)}{\hat{e}(i, j)}$.
2. calculer $\hat{c}(j)$, et l'enlever de l'image $\widehat{O1}(i, j)$ en calculant l'image intermédiaire $\widehat{O2}(i, j) = \frac{\widehat{O1}(i, j)}{\hat{c}(j)}$.
3. calculer $\hat{l}(i)$, et l'enlever de l'image $\widehat{O2}(i, j)$ en calculant l'image intermédiaire $\widehat{O3}(i, j) = \frac{\widehat{O2}(i, j)}{\hat{l}(i)}$.

On obtient alors l'estimation de l'image de fond : $\widehat{O3}(i, j) = \hat{p}(i, j)$

Les estimations successives des composantes sont effectuées en calculant des calculs de médiane sur des supports adaptés. Ceci permet d'estimer chacune des composantes indépendamment de l'action des autres, et de manière robuste à la présence de défauts inattendus dans la composante de bruit $p(i, j)$.

2.2.2.1 Estimation de la composante basse fréquence

Nous avons pris l'hypothèse que la composante basse fréquence est stationnaire sur un support de taille définie. L'expérience permet d'estimer la taille de ce support. Il s'agit d'une fenêtre carrée de taille 301×301 , sachant que l'image à traiter est de taille 3002×3002 . On appelle cette fenêtre \mathcal{F}_{i_0, j_0}^E centrée en (i_0, j_0) . L'estimation de $e(i, j)$ sur cette fenêtre donne :

$$\text{médiane}_{(i,j) \in \mathcal{F}_{i_0, j_0}^E} (o(i, j)) \propto e(i_0, j_0) \quad (2.3)$$

En effet, les autres composantes, moyennées sur le support choisi, vont tendre vers leur moyenne respective. La composante basse fréquence étant, elle, approximativement constante sur ce support, on va donc estimer sa valeur.

On peut alors estimer la composante basse fréquence sur toute l'image en réalisant l'estimation sur une fenêtre $F_{i,j}$ glissante :

$$\hat{e}(i, j) = \text{médiane}_{F_{i,j}^E} \{o(i, j)\} \propto e(i, j) \quad (2.4)$$

Pour simplifier, on considère que l'estimation de $e(i, j)$ est réalisée à un facteur multiplicatif près et cette estimation est notée $\hat{e}(i, j)$.

On a ainsi l'estimation de la composante basse fréquence, à un facteur multiplicatif près, en effectuant des médianes sur une fenêtre glissante de l'image d'origine. L'utilisation de la médiane pour effectuer les estimations permet d'être plus robuste aux défauts contenus dans les images analysées : leurs valeurs aberrantes peuvent fortement biaiser le résultat du calcul par moyenne empirique, le calcul par médiane permet d'être beaucoup moins sensible à cet effet. L'estimation de la composante basse fréquence revient donc à passer un filtre médian glissant de taille 301×301 sur l'image d'origine. On détermine, à partir de cette estimation, une image corrigée de l'effet basse fréquence par simple division de l'image d'origine par l'estimation de la composante basse fréquence :

$$\widehat{O1}(i, j) = \frac{o(i, j)}{\hat{e}(i, j)} \quad (2.5)$$

Sur l'illustration de la figure 2.2, on voit clairement qu'on a pu corriger l'image d'origine de l'effet de la structure basse fréquence : l'image $O1$ ne possède plus l'effet de dôme (bords de l'image plus sombres) présent sur l'image d'origine.

2.2.2.2 Estimation de la composante colonne

L'estimation de la composante colonne suit le même principe. La seule différence réside dans la fenêtre d'estimation, qui est cette fois-ci une fenêtre de la hauteur totale de la colonne. Dans le cas de l'exemple présenté sur la figure 2.2, on a ainsi estimé les valeurs de chaque colonne en calculant la médiane de toute la colonne. L'image traitée est en fait constituée de quatre quarts différents, par construction du détecteur, sur lesquels on estime les colonnes et les lignes. On a donc utilisé, pour estimer la valeur de la colonne, un médian sur une fenêtre de 1024×1 pixels. Par analogie avec les notations précédentes, nous appelons F_C cette fenêtre. On a alors :

$$\hat{c}(j) = \text{médiane}_{F_j^C} \{ \widehat{O1}(i, j) \} \quad (2.6)$$

Enfin, l'image intermédiaire corrigée de la structure colonne est calculée par :

$$\widehat{O2}(i, j) = \frac{\widehat{O1}(i, j)}{\hat{c}(j)} \quad (2.7)$$

a composante colonne estimée ainsi que l'image corrigée de la composante colonne sont représentées sur la figure 2.2. L'alternance de bandes verticales claires et sombres a bien été estimée dans la composante colonne, et la structure n'est alors plus visible sur l'image débarrassée de cette composante.

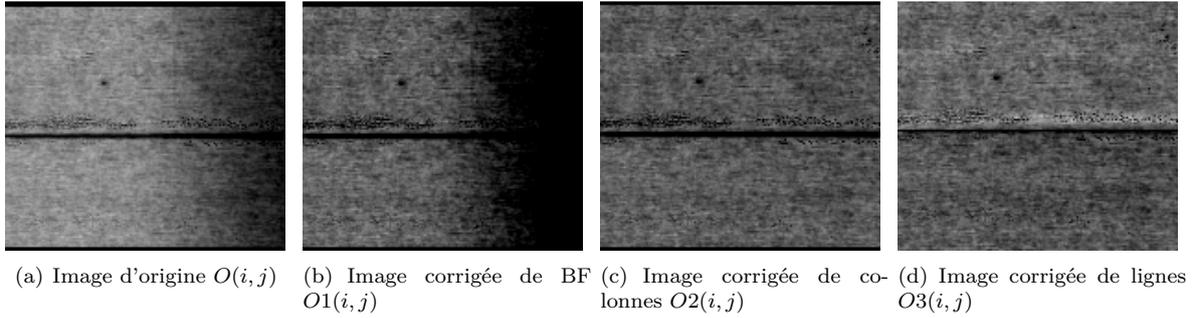


FIGURE 2.3 – Illustration d'un artefact de décomposition : sur l'image en fin de décomposition (image 2.3(d)), une ligne claire apparaît. Cette ligne n'est pas présente sur l'image d'origine (image 2.3(a)). Cette ligne a été introduite lors de la dernière étape de décomposition (composante ligne), du fait de la présence de pixels défectueux sur la longueur d'une ligne.

2.2.2.3 Estimation de la composante ligne

Enfin, l'estimation de la composante ligne s'effectue par calcul des valeurs médianes en ligne sur une fenêtre ligne F_i^L (hauteur 1 pixel et plusieurs pixels de largeur, centrée sur la ligne d'indice i).

Dans notre exemple, de manière analogue à ce qui a été fait pour estimer les valeurs des colonnes, nous calculons les valeurs des lignes par filtrage médian sur un support de taille 1×1024 . La composante lignes estimée, et l'image débarrassée de cette composante sont alors :

$$\hat{l}(i) = \text{médiane}_{F_i^L} \left\{ \widehat{O2}(i, j) \right\} \quad (2.8)$$

$$\widehat{O3}(i, j) = \frac{\widehat{O2}(i, j)}{\hat{l}(i)} \quad (2.9)$$

On obtient finalement une image débarrassée de toute structure basse fréquence, colonne et ligne. Cette image contient uniquement du bruit, et des défauts qui diffèrent des structures classiques apparaissant sur toutes les images.

2.2.3 Limitations de la méthode de prétraitement

La méthode retenue présente certaines limitations, principalement causées par la stratégie d'estimation adoptée. En effet, même si les estimations de l'effet de dôme, des valeurs des lignes et des colonnes, sont réalisées avec le médian, en présence d'une forte densité de pixels défectueux, la robustesse de l'estimateur est insuffisante. La conséquence directe est la mauvaise estimation de la composante, ce qui produit un artefact de décomposition lorsqu'on retire cette composante à l'image traitée.

Cet effet est illustré par la figure 2.3. On voit clairement apparaître sur l'image 2.3(d) un artefact de décomposition, sous la forme d'une trace verticale claire. Cette trace n'est pas présente sur l'image d'origine. La cause de cet artefact est simple : plusieurs pixels défectueux sont présents et distribués horizontalement. Les estimations des lignes sont ainsi perturbées par la présence de ces défauts. Dans ce cas, le médian n'est pas suffisamment robuste, du fait de la grande densité de défauts. Ce problème est visible dans certains cas, mais il reste néanmoins marginal car il s'observe sur des images fortement dégradées par des défauts de très grande taille. La manifestation de ce problème reste en générale assez faible, et limité spatialement. Enfin, on l'observe uniquement en présence de défauts importants dans l'image, donc les artefacts générés par ce léger défaut de la méthode ne seront pas des fausses alarmes, mais plutôt une légère modification artificielle de l'image aux alentours des défauts. En pratique, cette limitation n'est pas réellement gênante et impacte principalement la visualisation des résultats.

Un autre problème est la trop grande rigidité du modèle adopté. En particulier, nous avons modélisé les composantes lignes et colonnes par des constantes. En pratique, on observe une variation de la valeur moyenne le long d'une même ligne ou d'une même colonne. Ainsi, l'estimation faite de la valeur de cette ligne ou colonne peut être perturbée par cette variation le long de la ligne, et entraîner l'apparition d'artefacts dans l'image prétraitée. Dans ce cas, le filtrage médian fait un compromis dépendant de la représentation, sur le support d'estimation, de chaque zone. Pour pouvoir affiner l'estimation, il faudrait chercher sur un voisinage plus étendu (en particulier, regarder dans les directions perpendiculaires à l'estimation) la valeur pertinente à choisir. Cependant, ce problème ne constitue pas une grande gêne puisque les variations petite échelle des composantes sont généralement très légères.

2.2.4 Conclusions et remarques sur le prétraitement

L'étape de prétraitement permet de s'affranchir de structures connues des images issues du détecteur. Ces structures nécessitent un traitement spécial pour savoir si elles sont acceptables ou non pour la qualité de l'image. En général, la décision est prise en fonction de spécifications propres au produit. Dans tous les cas, ces structures n'apparaissent pas dans l'image finale puisqu'elles sont corrigées par la chaîne de correction. Ces défauts particuliers doivent ainsi être traités à part. La méthode développée s'appuie sur des calculs de médian en lieu et place de la moyenne, ce qui lui donne une plus grande robustesse (voir notamment [10]) à la présence de défauts dans l'image $p(i, j)$. La procédure de prétraitement exposée est couramment utilisée sur les images Trixell. Même si elles conduisent parfois à l'introduction d'artefacts supplémentaires, les limitations exposées ne sont pas bloquantes.

2.3 Détection des zones défectueuses

2.3.1 Objectifs et enjeux

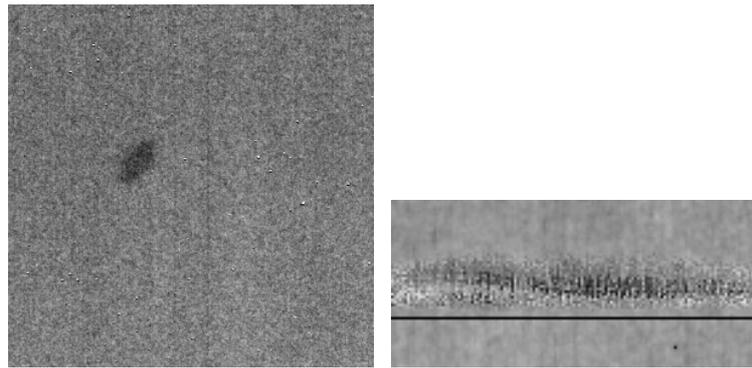
Une fois l'image d'origine débarrassée des structures récurrentes, il faut détecter les zones présentant des défauts inattendus. Ces défauts sont des pixels, ou des arrangements de pixels, qui, par leurs valeurs ou leur configuration spatiale, sont distinguables du bruit de fond présent sur toute l'image. La méthode de détection de zones défectueuses vise à identifier les zones contenant potentiellement des défauts. L'enjeu de cette méthode est de détecter toutes les zones où un défaut est présent, car toute zone non détectée ne sera plus analysée par la suite. En contrepartie, on peut accepter que quelques zones détectées par le bloc de détection ne contiennent en fait pas de défaut, mais ce taux de fausse alarme doit être faible car il va générer soit des calculs superflus, soit une observation superflue d'un observateur humain qui voudra vérifier visuellement toutes les zones défectueuses.

Nous livrons ici une méthode de détection pragmatique, s'appuyant sur la connaissance des défauts pouvant intervenir sur les images qui font l'objet de la thèse. Nous verrons dans les deux prochains chapitres des méthodes plus sensibles de détection, qui permettent de traiter le problème de manière plus générale, et qui s'attachent à la détection de défauts en général non détectés par la méthode que nous présentons dans ce chapitre.

2.3.2 Caractérisation des défauts à détecter et présentation des méthodes de détection

Les défauts susceptibles d'intervenir sur les images à analyser peuvent être classés en trois catégories :

1. les pixels défectueux isolés : c'est une déviation, à l'échelle du pixel, de la valeur attendue sur l'image $a(x, y)$ (voir figure 2.4(a)). La valeur attendue sur l'image est proportionnelle à la dose d'exposition. Idéalement, sur une image plate, on devrait obtenir une valeur constante. En l'absence de défaut, dans le cas réel, on obtient un niveau moyen $a(x, y)$ (celui correspondant à la dose choisie) qui a été bruité par des phénomènes physiques (source, effet de scintillement différent, etc) ou électroniques (bruit électronique de lecture, gains électroniques différents, etc). Un pixel isolé



(a) Défauts à détecter : tache sombre et pixels isolés (b) Défauts à détecter : ligne et groupement de pixels défectueux

FIGURE 2.4 – Exemples de défauts à détecter. A gauche : un groupement de pixels qui perturbe le niveau de fond (tache sombre), et des pixels isolés. A droite, une ligne défectueuse et un groupement de pixels défectueux pouvant être caractérisés en variance spatiale.

est une déviation ponctuelle de la valeur attendue. On peut modéliser cette déviation sous la forme d'un effet proportionnel à la dose appliquée. Cet effet correspondra simplement à une valeur de $a(x, y)$ différente de celle attendue.

2. groupement de pixels en valeur moyenne : ce sont des déviations grande échelle de la valeur moyenne attendue localement. Un défaut vient perturber le niveau moyen attendu (il s'agit d'une perte de sensibilité locale). Ces défauts se caractérisent par des taches claires ou sombres, comme celle de la figure 2.4(a)
3. groupement de pixels en variance : ce sont des déviations de la variance locale (voir figure 2.4(b)). Des défauts en général d'origine physique viennent se rajouter au bruit de fond naturel des images pour créer une zone de variance beaucoup plus forte. Ces défauts sont caractérisés par la présence, sur une échelle variable, de pixels de valeurs atypiques. Ces défauts peuvent s'accompagner d'une variation de la valeur moyenne locale, mais l'effet principal qui les rend visibles est une variation forte de la variance locale. Ces types de défauts sont en fait plusieurs pixels isolés défectueux qui sont regroupés dans une zone particulière, formant une zone étendue de pixels défectueux.

Nous proposons ici une stratégie en deux étapes, permettant d'abord de détecter les pixels isolés, puis les groupements de pixels, et de fusionner ensuite ces deux détections pour produire une seule cartographie des défauts. Cette cartographie est composée à la fois de pixels (pour la détection des pixels isolés), et de zones (pour la détection de groupements de pixels). Cette méthode est illustrée par la figure 2.5. Pour détecter les deux catégories de défauts, on s'appuie, en plus de la valeur des pixels utilisée pour la détection des pixels isolés, sur des estimations de différences de contraste (valeur moyenne différente entre certaines zones et le bruit de fond) et de variance. Ces attributs, comme détaillé dans les paragraphes suivants, sont assez classiquement utilisés pour la détection de défauts sur images issues de techniques de radiologie. Ils sont souvent combinés (logique floue dans [54]) afin de tendre vers les performances en détection du système visuel.

2.3.2.1 Détection des pixels isolés

L'objectif est de détecter les pixels isolés paraissant défectueux. Pour cette tâche, on s'appuie sur la valeur de chaque pixel, qu'on compare à une statistique locale. La méthode de détection de pixels s'appuie uniquement sur la valeur de chaque pixel composant l'image à analyser. Cette image est découpée en boîte d'analyse de taille 15×15 pixels. Cette taille est une échelle d'observation suffisamment grande pour que le pixel défectueux soit visible et pour que l'estimation des statistiques sur cette zone de l'image soit fiable. La valeur de chaque pixel va être comparée à un seuil calculé à partir de la moyenne et de

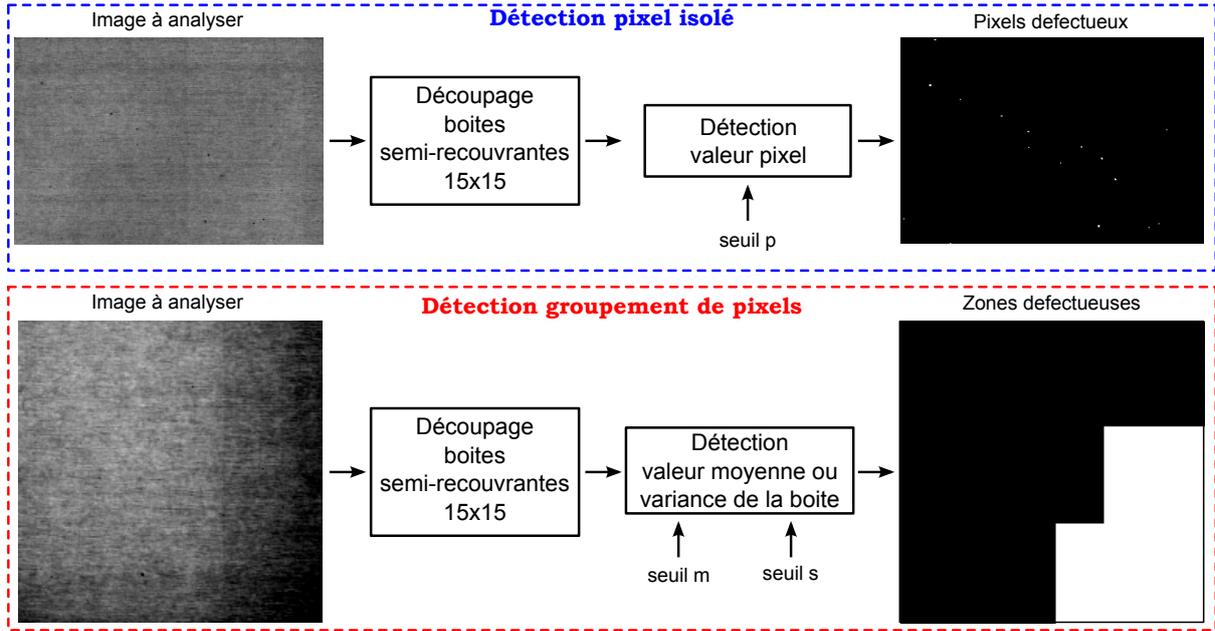


FIGURE 2.5 – Méthode de détection. Deux détections sont réalisées : une pour la détection des pixels et l'autre pour la détection de zones. Les deux cartes de défauts obtenues seront fusionnées pour obtenir la carte de détection finale (ici on a illustré chacune des détections sur deux zones différentes de l'image).

la variance de la boîte à laquelle il appartient. Ceci est une méthode de détection classique s'appuyant uniquement sur un critère statistique, le choix des seuils étant fixé soit afin d'obtenir un taux de fausse alarme donné, soit par une approche empirique. Dans notre cas, nous avons choisi une méthode empirique pour fixer les seuils, en nous appuyant sur l'observation visuelle. On considère l'image à analyser $\{p(i, j), (i, j) \in \llbracket 1; N_L \rrbracket \times \llbracket 1; N_C \rrbracket\}$ et $V_{i,j}$ un voisinage du pixel $p(i, j)$. La détection des pixels isolés se déroule en 2 étapes :

1. Calcul des statistiques (moyenne $\mu_{V_{i,j}}$ et variance $\sigma_{V_{i,j}}^2$) sur le voisinage de chaque pixel. On cherche à estimer les valeurs de la moyenne et de la variance du bruit de fond dans un voisinage du pixel à analyser. En pratique, afin de ne pas être perturbé par les pixels defectueux éventuellement présents dans le voisinage, on estime moyenne et variance par une méthode robuste [38]³ :

$$\mu_{V_{i,j}} = \text{médiane}(V_{i,j}) \quad (2.10)$$

$$\sigma_{V_{i,j}} = \frac{3}{2} \cdot \text{médiane}(|V_{i,j} - \mu_{V_{i,j}}|) \quad (2.11)$$

2. Décision sur chaque pixel : en fonction de la valeur de chaque pixel, on va décider si celui-ci est atypique ou non. Si le pixel est atypique, il aura la valeur $DMpix(i, j) = 1$ dans la carte des défauts pixels, sinon il aura la valeur 0.

$$DMpix_{i,j} = \begin{cases} 1 & \text{si } |p(i, j) - \mu_{V_{i,j}}| \geq t_p \sigma_{V_{i,j}} \\ 0 & \text{sinon} \end{cases} \quad (2.12)$$

On remarque la présence du seuil t_p qui permet de régler la sensibilité de la détection.

3. Le premier estimateur, la médiane, permet une estimation robuste et non biaisée de la tendance moyenne. Le second estimateur est la déviation absolue médiane (MAD), qui est un estimateur de la dispersion statistique. Pour qu'il soit un estimateur non-biaisé de l'écart-type dans le cas gaussien, il faut ajouter le facteur multiplicatif $K = \frac{3}{2}$

2.3.2.2 Détection de groupements de pixels

On souhaite ici détecter des groupements de pixels possédant des propriétés statistiques anormales : valeur moyenne locale ou variance locale anormale. Pour cela, on effectue la même opération de détection, mais en s'appuyant non pas sur la valeur de chaque pixel, mais sur la valeur des statistiques calculées sur des fenêtres d'analyse $F_{i,j}$ de taille $N_F \times N_F$. La stratégie se déroule toujours en deux étapes :

1. calcul des moyennes $m_{F_{i,j}}$ et des variances $s_{F_{i,j}}$ locales sur chacune des fenêtres d'analyse $F_{i,j}$ centrées sur (i, j) . Ici, on utilise les estimateurs de moyenne et de variance empiriques, et non les estimateurs robustes, car on ne cherche pas à estimer une statistique du fond en présence des perturbations, mais plutôt une caractéristique de la fenêtre analysée :

$$m_{F_{i,j}} = \frac{1}{N_F^2} \cdot \sum_{(m,n) \in F_{i,j}} p(m, n) \quad (2.13)$$

$$s_{F_{i,j}} = \frac{1}{N_F^2 - 1} \cdot \sum_{(m,n) \in F_{i,j}} (p(m, n) - m_{F_{i,j}})^2 \quad (2.14)$$

2. décision sur chaque fenêtre $F_{i,j}$: chaque pixel appartenant à la fenêtre $F_{i,j}$ sera marqué comme détecté dans la carte des détections $DMbox(i, j)$ si la fenêtre possède des valeurs de moyenne et de variance atypiques par rapport à celles du bruit de fond. Nous calculons donc d'abord les statistiques, de manière robuste, des valeurs de moyenne et de variance des fenêtres. Ces statistiques sont calculées localement, pour tenir compte de la non-stationnarité du bruit de fond. En effet, l'effet de dôme entraîne un bruit différent sur les bords de l'image. Il a été compensé en valeur moyenne par le prétraitement, mais en variance on observe encore une différence notable du bruit. Pour calculer les statistiques localement, on construit une image de moyennes et une image de variances respectivement à partir des valeurs de $m_{F_{i,j}}$ et $s_{F_{i,j}}$. Cette image est de dimension sept fois plus petite que l'image d'origine, du fait de l'utilisation de boîtes semi-recouvrantes. On note $M(i, j) = m_{F_{i,j}}$ et $S(i, j) = s_{F_{i,j}}$ ces deux images, respectivement. On passe ensuite un filtre médian glissant de taille 43×43 sur ces images. La taille est réglée sur celle du premier filtrage de l'étape de prétraitement, destiné à compenser l'effet de dôme. Les images M et S étant sous-échantillonnées d'un facteur 7 par rapport à l'image d'origine, le filtrage médian 43×43 est équivalent à un filtrage 301×301 sur l'image d'origine. Ce traitement permet de compenser deux effets néfastes pour l'étape de détection :

- (a) l'effet de non-stationnarité du bruit : la statistique du bruit de fond a une évolution sur une échelle du même ordre de grandeur que l'effet de dôme, puisqu'il en est la cause. D'autres phénomènes peuvent intervenir, rendant les statistiques du bruit dépendantes de la région où elles sont calculées.
- (b) artefacts de décomposition : nous avons vu que certains artefacts créés par l'étape de prétraitement peuvent être introduits. En prenant une taille de filtre identique pour calculer les références locales des statistiques du bruit, on pourra ainsi éviter la détection de ces artefacts lors de l'étape de détection.

On note :

- $\mu_m^F(i, j)$ et $\sigma_m^F(i, j)$ les images de la moyenne et l'écart type des valeurs de $m_{F_{i,j}}$ prises sur l'image traitée. Pour simplifier, on notera m_F l'image formée à partir des valeurs de $m_{F_{i,j}}$.
- $\mu_s^F(i, j)$ et $\sigma_s^F(i, j)$ les images de la moyenne et l'écart type des valeurs de $s_{F_{i,j}}$ prises sur l'image traitée. Pour simplifier on notera s_F l'image formée à partir des valeurs de $s_{F_{i,j}}$.

On a :

$$\mu_m^F(i, j) = \text{médiane}_{(m,n) \in V_{i,j}} (m_F(m, n)) \quad (2.15)$$

$$\sigma_m^F(i, j) = \frac{3}{2} \cdot \text{médiane}_{(m,n) \in V_{i,j}} (|m_F(m, n) - \mu_m^F(m, n)|) \quad (2.16)$$

$$\mu_s^F(i, j) = \text{médiane}_{(m,n) \in V_{i,j}} (s_F(m, n)) \quad (2.17)$$

$$\sigma_s^F(i, j) = \frac{3}{2} \cdot \text{médiane}_{(m,n) \in V_{i,j}} (|s_F(m, n) - \mu_s^F(m, n)|) \quad (2.18)$$

où $V_{i,j}$ est un voisinage de 43×43 pixels centré sur (i, j) .

On construit alors la carte des zones défectueuses de la manière suivante :

$$\begin{aligned} &\text{Soit } (i_0, j_0) \in F_{i,j}, \\ DMbox(i_0, j_0) &= \begin{cases} 1 & \text{si } |M(i, j) - \mu_m^F(i, j)| \geq t_m \sigma_m^F(i, j) \text{ ou } |S(i, j) - \mu_s^F(i, j)| \geq t_s \sigma_s^F(i, j) \\ 0 & \text{sinon} \end{cases} \end{aligned} \quad (2.19)$$

2.3.3 Choix des seuils de détection

Plusieurs méthodes sont envisageables pour fixer les seuils de détection des trois algorithmes présentés. La première, la plus naturelle, consiste à se fixer un taux de fausse alarme cible, et à paramétrer les trois méthodes afin qu'elle respecte ce taux de fausse alarme. La performance finale de l'algorithme sera alors qualifiée en termes de fausses alarmes, sans tenir compte de la sensibilité en termes de détection de chaque algorithme. Nous avons normalisé les méthodes de manière à ce que la statistique finale, à seuiller, suive dans les trois cas la même statistique (voir annexe A). Ainsi, pour un taux de fausse alarme donné, les trois seuils correspondants seront identiques. Par exemple, pour un taux de fausse alarme de l'ordre de 10^{-9} , il faudra prendre des seuils d'environ 6. Néanmoins, cette méthode souffre des défauts suivants :

- sensibilité différente de l'œil : chaque algorithme est spécialisé dans la détection d'un type de défaut. Il s'avère que la sensibilité de l'œil est différente selon le type de défaut. En particulier, l'œil détecte plus facilement une tache claire ou sombre, qu'un pixel un peu plus clair ou plus sombre que le bruit de fond. Ainsi, pour pouvoir se rapprocher des performances de l'œil pour la détection, l'algorithme de détection de groupements de pixels doit être réglé avec un seuil de détection plus faible, de manière à être plus sensible
- taux de fausse alarme purement théorique : l'hypothèse d'indépendance prise pour déterminer les seuils théoriques correspondant à des taux de fausse alarme fixés est généralement fausse. En effet, du fait de la chaîne de correction (voir section 1.4) et de l'aspect basse fréquence du bruit (figure 1.7), on ne peut pas dire que le bruit de fond de l'image soit réellement blanc : des structures spatialement corrélées apparaissent du fait du processus de correction des images, ainsi qu'un aspect basse fréquence du bruit, dû à la technologie même du détecteur.

La seconde méthode possible pour le réglage des seuils (voir annexe B) est de déterminer, pour chaque méthode, à partir de quel seuil le défaut ciblé par la méthode de détection est visible. Ceci peut être fait par une étude psychovisuelle mettant à contribution des observateurs humains. Ces études sont en général lourdes à mener. L'expérience menée pour déterminer le réglage des seuils a abouti sur des valeurs très surévaluées par rapport à celles que nous attendions sur le type d'image que nous avons traité. Ceci est principalement dû aux modèles choisis pour générer les images synthétiques à évaluer (bruit blanc, modèles simplistes de défauts). Il serait intéressant de pousser cette étude en s'appuyant sur l'évaluation d'artefacts réels.

Nous avons donc choisi de régler les seuils à partir de l'observation d'exemples : sur image réelle, ont été observées les valeurs des trois statistiques P , M et S , conjointement à l'image à analyser. A partir de cette observation, nous avons déterminé les seuils adéquats, permettant de détecter la plupart des défauts présents dans l'image. Les seuils suivants ont été adoptés : $t_p = 6$, $t_m = 4$ et $t_s = 5$. L'arbitraire de ces valeurs peut être évoqué, mais la tâche de détection fera l'objet d'une étude plus complète, avec une procédure d'évaluation adéquate, dans les chapitres suivants.

Néanmoins, en complément de ces seuils fixés de manière *ad hoc* on peut imaginer une stratégie expérimentale simple afin de qualifier rigoureusement les performances, et surtout les impacts en termes industriels, de cette méthode de détection. Il s'agit de mesurer le taux de fausse alarme expérimentalement, sur données réelles. Cette stratégie se décline en 2 étapes :

1. production de données : à partir d'images de spectre de puissance (NPS) d'images réelles (même type d'image que celles représentées sur la figure 1.7), on peut générer plusieurs réalisations de

Taille des boîtes d'analyse (en pixel)	t_p	t_m	t_s
15	6	4	5

TABLE 2.1 – Paramètres retenus pour la méthode de détection

bruit réel. C'est notamment la technique que nous utiliserons dans l'étape de création de contexte (section 2.7).

- test du taux de fausse alarme : sur ces réalisations de bruit, on peut alors appliquer l'algorithme de détection et compter le nombre de détections, ce qui nous donnera directement le nombre de fausses alarmes.

A l'aide de l'étude théorique menée dans l'annexe A, en prenant l'hypothèse d'un bruit blanc et gaussien, le nombre de fausses alarmes attendu, pour une image 3000×3000 , est de :

- 0.017 pixels pour l'algorithme de détection des pixels isolés, soit environ 1 fausse alarme toutes les 100 images traitées
- 11 ROIs de dimension 15×15 pour l'algorithme de détection de la valeur moyenne de groupements de pixels
- 0.1 ROI de dimension 15×15 (soit 1 ROI toute les 10 images traitées) pour l'algorithme de détection de la variance de groupements de pixels

Ce modèle théorique prenant des hypothèses qui ne se vérifient pas forcément sur les images analysées (notamment l'hypothèse de bruit blanc), il est attendu que les taux de fausse alarme théoriques soient différents des taux déterminés par le protocole expérimental décrit plus haut.

2.3.4 Limitations de la méthode de détection

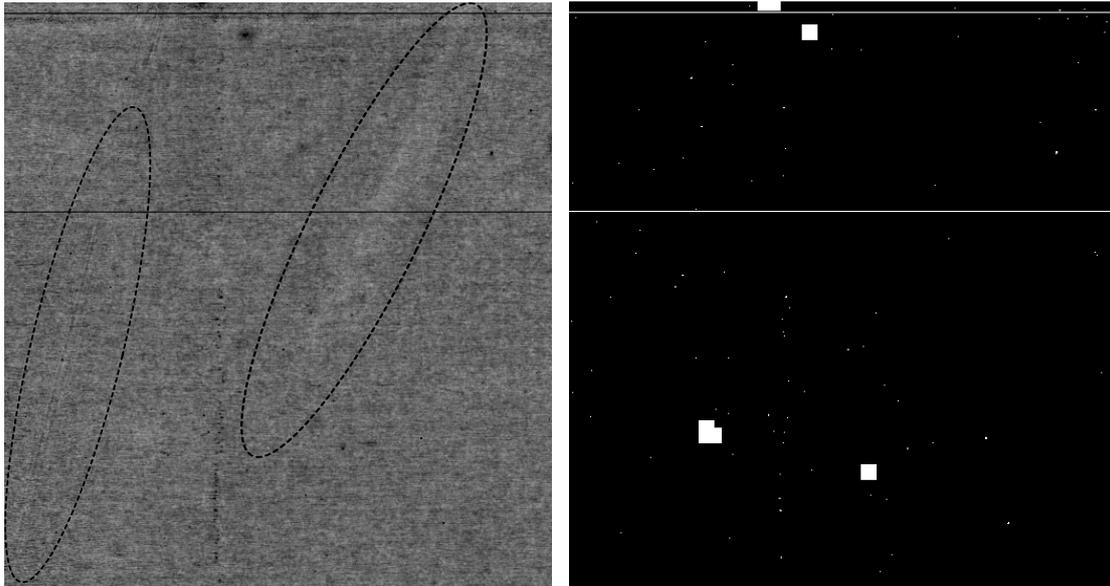
La méthode de détection développée ici possède un certain nombre de limitations, notamment :

- sensibilité en détection insuffisante : l'aspect purement statistique ne permet pas à cette méthode d'exploiter les corrélations spatiales locales, alors que l'œil y est particulièrement sensible. La seule corrélation spatiale utilisée est celle correspondant à la fenêtre d'analyse carrée 15×15 pixels. Même si la plupart des défauts sont détectés par la méthode proposée ici, celle-ci est insuffisante pour la détection de certains défauts présents sur image réelle (voir figure 2.6). La méthode permet de détecter les défauts vraiment inacceptables, et peut manquer certains défauts légers.
- Analyse bi-échelle : les deux échelles de défaut couvertes par la méthode de détection sont l'échelle du pixel (détection des pixels isolés) et l'échelle 15×15 pixels. L'utilisation de ces deux échelles est une amélioration significative par rapport à la méthode classique de détection pixel à pixel. Néanmoins certains défauts, de taille différente des deux échelles prises en considération, peuvent échapper à l'algorithme de détection.

2.3.5 Conclusions et remarques sur la détection

La méthode de détection définie dans cette section est une méthode naturelle qui permet de détecter la plupart des défauts présents sur les images. Néanmoins, certains défauts échappent à cette méthode de manière systématique, comme ceux présentant de grandes structures spatialement corrélées : ces défauts peuvent avoir un niveau très faible tout en étant visibles en raison de leur corrélation spatiale. Dans ce cas, les techniques s'appuyant uniquement sur des statistiques, même locales, sont inefficaces. L'objectif du chapitre 4 sera de mettre au point des méthodes de détection capables de détecter ce genre de défauts.

Néanmoins, en première approximation, la méthode de détection développée ici peut être utilisée. Nous l'avons implémentée au sein de la chaîne d'analyse de qualité, avec les paramètres de la table 2.1.



(a) Image sur laquelle est appliqué l'algorithme de dé- (b) Résultat de la détection : en blanc, les pixels détectés
tection

FIGURE 2.6 – La méthode de détection proposée ne permet pas de détecter les structures de rapport signal à bruit faible, mais visibles à l'œil du fait de leur corrélation spatiale. Ici, deux structures corrélées (lignes diagonales à gauche de l'image, et tache claire à droite de l'image) ne sont pas correctement détectées.

2.4 Extraction des zones défectueuses

2.4.1 Objectifs et enjeux

Une fois les zones défectueuses identifiées à l'aide de l'algorithme de détection, une étape d'extraction de ces zones permet de générer des captures de l'image d'origine centrées sur le défaut. Cette capture a pour but de minimiser l'espace de stockage nécessaire pour analyser des images de grande taille en mémorisant uniquement la zone suffisante à l'identification visuelle du défaut. Des images de petite taille permettront également de réduire le temps de calcul dans les étapes suivantes. Néanmoins, la taille de ces captures doit être déterminée de manière à ce que suffisamment de bruit de fond soit également présent dans la capture. Ceci pour les deux raisons suivantes :

- visuellement : l'œil a besoin de suffisamment de surface de contexte afin d'identifier la zone défectueuse
- numériquement : pour les étapes suivantes, notamment la localisation, nous allons avoir besoin d'estimer des statistiques du contexte. Les pixels de contexte doivent donc être suffisamment représentés afin d'estimer de manière fiable les statistiques.

2.4.2 Calcul adapté de la zone extraite

Afin de répondre à ces deux préoccupations, nous proposons de définir une règle de capture qui tient compte de la forme des objets détectés. En effet, on remarque que :

- quand l'objet détecté est de forme carrée, une extraction de forme carrée est appropriée (pas de préférence pour une direction), et une extraction de la taille double de l'objet détecté, dans les deux directions, est en général suffisante pour visualiser l'objet.
- Quand l'objet est plus étendu dans une direction donnée, on a en général besoin de visualiser une quantité plus importante de pixels de contexte dans la direction perpendiculaire à l'objet détecté de manière à bien identifier le défaut. Par exemple, lorsqu'on veut visualiser une ligne défectueuse, on

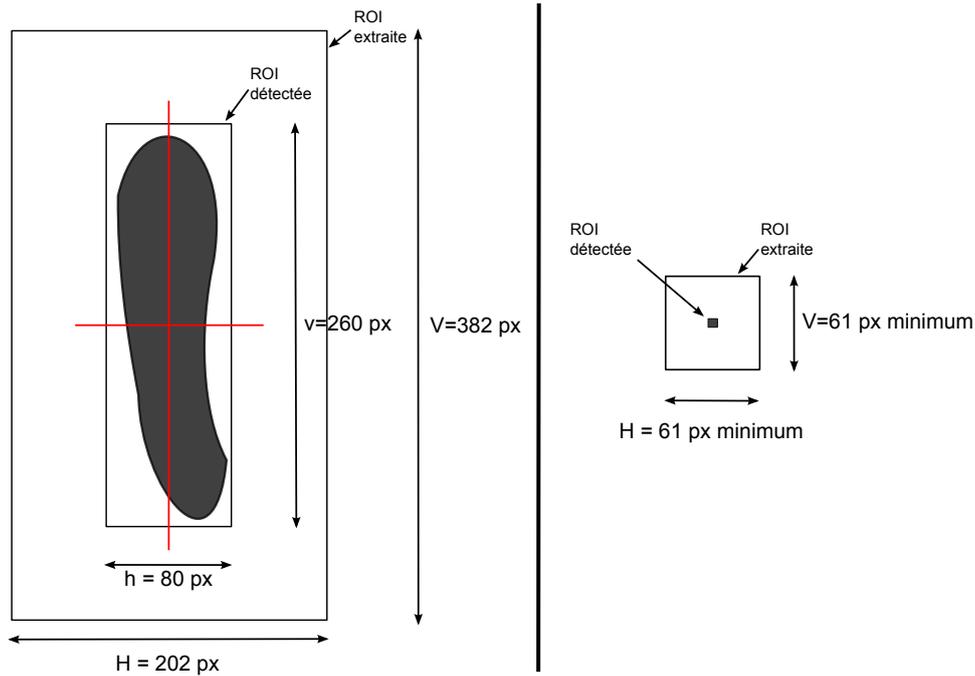


FIGURE 2.7 – Méthode d'extraction. Deux cas sont présentés : à gauche, extraction d'une forme à dominante verticale, et à droite extraction d'un pixel défectueux. La capture de la forme verticale va être étendue selon la composante horizontale pour une bonne visualisation. Le pixel défectueux sera toujours capturé dans une zone de 61×61 pixels, ce qui est suffisant à sa visualisation.

a besoin de voir suffisamment de lignes situées de part et d'autres de la ligne défectueuse. Dans le même temps, l'information des pixels situés sur des colonnes non impactées par la ligne défectueuse n'est en général pas nécessaire

La stratégie de capture adoptée se sert des dimensions horizontale, notée h , et verticale, notée v , de la zone détectée, et calcule les dimensions horizontale H et verticale V de la zone de capture, selon les formules suivantes :

$$H = h \left(1 + \frac{2}{\alpha + 1} \right) \quad (2.20)$$

$$V = v \left(1 + \frac{2}{\frac{1}{\alpha} + 1} \right) \quad (2.21)$$

$$\text{avec } \alpha = \frac{h}{v} \quad (2.22)$$

De plus, dans tous les cas nous avons fixé une taille de capture minimale, correspondant à 61 pixels *minimum* pour chacune des deux dimensions H et V . Une première étape vise à regrouper les zones détectées en composantes connexes, chacune de ces composantes étant ainsi extraite selon la règle énoncée ci-dessus.

La figure 2.7 donne un exemple d'extraction de deux défauts différents. Les zones capturées sont toujours centrées sur la ROI de détection. Pour une bonne visualisation, le format de la zone capturée est étendu selon la direction perpendiculaire à la direction principale de la ROI détectée. Sur la forme à gauche de la figure, on a donc capturé en ne suivant pas le format donné par la ROI détectée, mais en l'étendant selon la direction horizontale. En pratique, ceci permet une meilleure observation du défaut, car il sera visuellement comparé à des zones du contexte qui lui seront parallèles. En revanche, nous n'avons pas besoin, pour une bonne visualisation, de beaucoup de pixels situés dans l'axe du défaut. Pour l'extraction du pixel isolé, la ROI sera ajustée à sa taille minimale, à savoir 61×61 pixels.

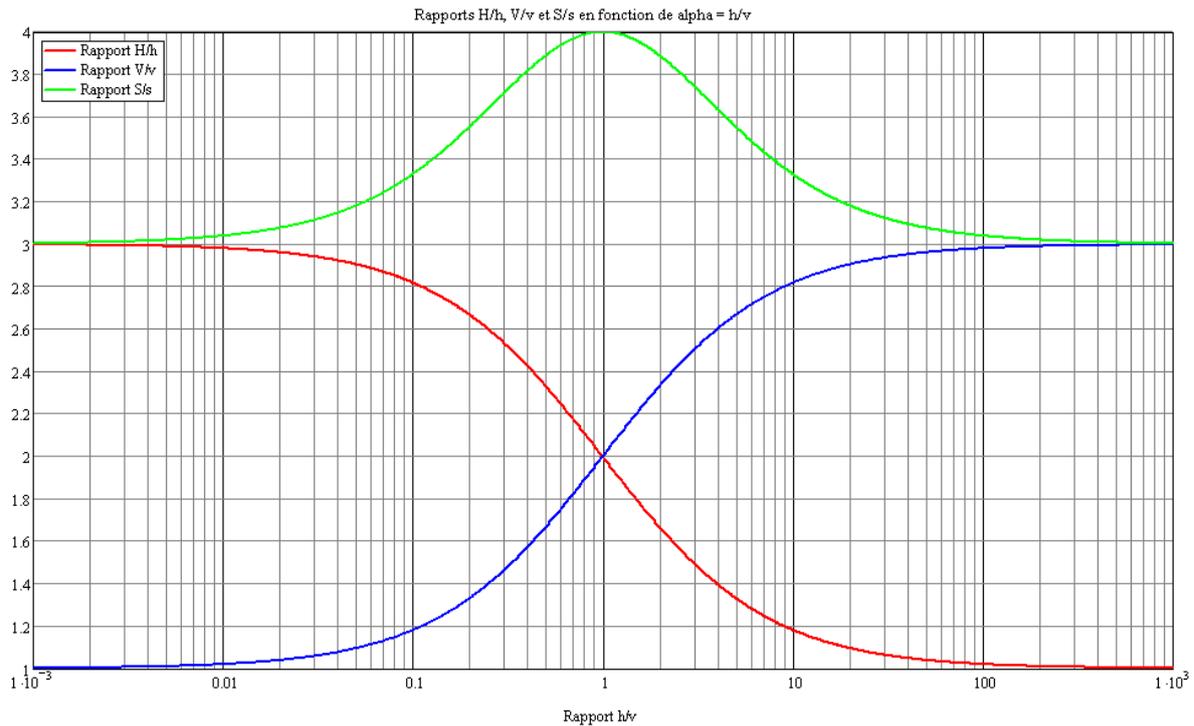


FIGURE 2.8 – Méthode d'extraction : courbes des rapports dimension extraite/dimension détectée. Selon la forme de l'objet détecté, on extrait une zone de taille comprise entre une fois et trois fois la taille détectée. La surface extraite varie en conséquence entre 3 fois et 4 fois la surface détectée.

Le graphe de la figure 2.8 donne les ratios entre longueur et largeur de la ROI extraite, et longueur et largeur de la ROI détectée. On observe que ces deux quantités varient entre 1 et 3 selon la forme de l'objet. La surface de l'objet capturée est elle entre 3 fois et 4 fois la surface de l'objet détecté, ce qui représente une taille raisonnable pour le stockage des données.

2.4.3 Conclusions et remarques sur l'extraction

La méthode présentée ici pour effectuer la capture des défauts est une méthode *ad hoc* qui, en pratique, donne des résultats satisfaisants puisque les défauts sont en général visibles sur toutes les captures effectuées.

Une autre méthode serait de capturer une zone de taille fixe, centrée sur chaque pixel détecté. Ceci revient à dilater les détectations par un élément structurant de la taille choisie, et de capturer ensuite chaque composante connexe par une ROI englobant cette composante.

L'avantage de la méthode présentée est de réduire la quantité de pixels enregistrés dans le cas de défauts de type ligne ou colonne défectueuse, puisque la différence selon l'axe horizontal et vertical entre la ROI de capture et la ROI détectée va dépendre de la forme de l'objet. Avec cette technique, on a également le contrôle de la quantité de pixels capturés par rapport à la quantité de pixels détectés. Enfin, l'aspect adaptatif à la taille de l'objet permet de réaliser des captures pertinentes du point de vue de l'observation : si on a détecté un défaut de grande taille, et qu'on applique une méthode de capture par dilatation par fenêtre fixe, la capture va être de taille quasiment identique à la détection. Dans ce type de configuration, il devient très difficile de bien visualiser le défaut, car il ne pourra pas être observé à la bonne échelle, il n'y a pas suffisamment de contexte pour permettre cette observation.

2.5 Localisation des pixels défectueux

2.5.1 Objectifs et enjeux

Dans cette étape, on applique aux ROIs extraites un algorithme de localisation susceptible de fournir une détection au pixel près des pixels défectueux. Cette étape a deux objectifs principaux :

- détection au pixel près : pour l'étape suivante de débruitage, on veut connaître la composition du défaut détecté. On veut donc obtenir une cartographie de la zone extraite, en deux classes : les pixels défectueux, et les pixels de fond.
- Suppression des fausses alarmes : certaines zones extraites ne présentent en définitive pas de défaut. Cette étape vise donc également à éliminer de l'analyse les fausses alarmes. Cette deuxième détection part de l'*a priori* défini par la première : on a une estimation de la zone de contexte, ce qui est utile pour affiner la méthode de détection.

La principale difficulté de cette étape repose dans la sensibilité de la détection : tout pixel non détecté à l'étape de localisation ne sera plus considéré comme pixel défectueux.

2.5.2 Méthode proposée

Pour traiter cette étape, nous proposons une méthode de détection basique, qui utilise l'*a priori* défini lors de l'étape de détection.

La méthode de localisation proposée s'appuie en effet sur une caractérisation du contexte (le fond bruité) par ses deux premiers moments statistiques, puis sur un seuillage de l'image à traiter à l'aide de cette caractérisation. L'étape précédente a permis de définir des captures $V_k(i, j)$ des défauts de l'image d'origine. Les pixels de chaque capture sont répartis en deux classes : les pixels détectés par la méthode de détection (classe D), et les pixels de fond (classe F). On note V_k^0 le sous-ensemble constitué par les pixels de fond dans V_k , et V_k^1 le sous ensemble constitué des pixels détectés. La procédure de localisation va affecter à chaque image V_k une image de localisation L_k avec $L_k(i, j) = 1$ si (i, j) est un pixel identifié comme défectueux, et $L_k(i, j) = 0$ si (i, j) est identifié comme un pixel de fond.

La méthode de localisation opère donc en deux étapes :

1. estimation de la statistique du fond : on calcule les deux quantités suivantes :

$$\mu_k^F = \text{médiane} \{V_k^0(i, j) | (i, j) \in F\} \quad (2.23)$$

$$\sigma_k^F = \frac{3}{2} \text{médiane} \{|V_k^0(i, j) - \mu_k^F| | (i, j) \in F\} \quad (2.24)$$

2. seuillage de l'image avec un seuil t_l : on détermine l'image de localisation L_k de la manière suivante :

$$L_k(i, j) = \begin{cases} 1 & \text{si } \left| \frac{V_k(i, j) - \mu_k^F}{\sigma_k^F} \right| \geq t_l \\ 0 & \text{sinon} \end{cases} \quad (2.25)$$

Pour des applications pratiques, nous avons choisi un seuil $t_l = 4.5$, ce qui permet, d'un point de vue théorique, de n'avoir qu'une seule fausse alarme en moyenne par capture $V(i, j)$. En effet, après passage par la méthode de prétraitement, on ne peut retrouver des défauts plus grands que 301×301 pixels. D'après la courbe de la figure 2.8, la taille maximale d'une capture est de $301 \cdot 301 \cdot 4 = 362404$ pixels. Le seuil théorique dans ce cas pour obtenir en moyenne une fausse alarme par image capturée est de l'ordre de $t_l = 4.5$ (voir figure A.1 en annexe). Nous avons donc choisi ce seuil pour appliquer la méthode de localisation.

2.5.3 Limitation de la méthode de localisation

L'étape de localisation souffre de problèmes similaires à l'étape de détection, à savoir un manque de sensibilité pour la détection de structures corrélées. En effet, la détection pixel à pixel, en utilisant

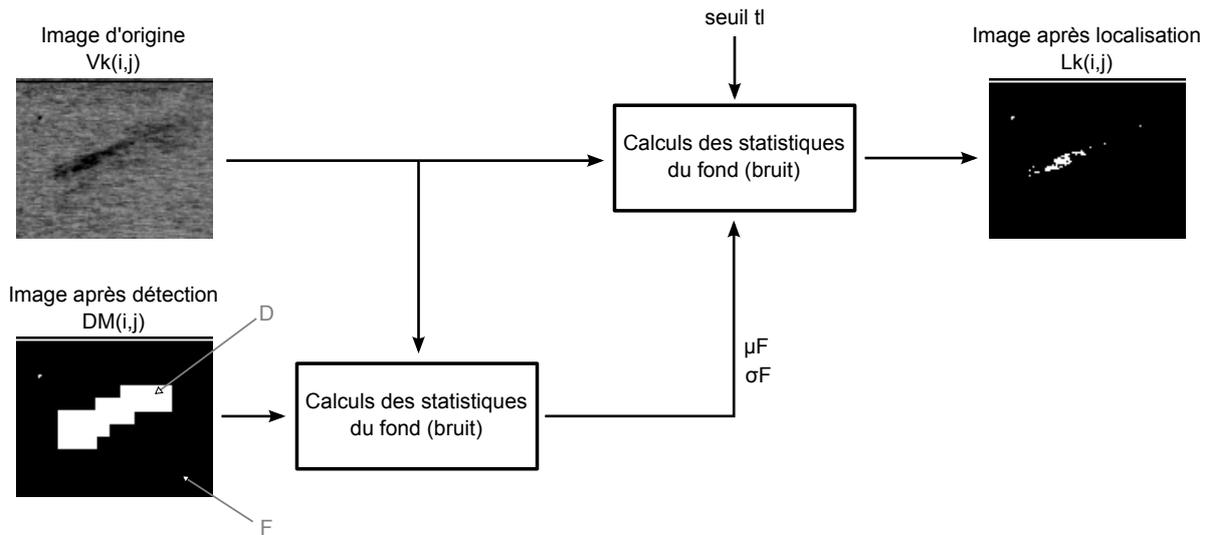


FIGURE 2.9 – Méthode de localisation : la valeur de moyenne et d'écart-type du bruit de fond est calculée à l'aide de l'image d'origine et de la carte de détection fournie par l'algorithme de détection. L'image d'origine est ensuite seuillée (voir équation 2.25) pour produire une image de localisation des pixels défauts.

uniquement un critère statistique, échoue généralement lorsqu'on veut détecter des structures plutôt visibles grâce à leur corrélation spatiale. La méthode de détection utilisée ici ne tient pas compte du voisinage de chaque pixel testé, et donc de l'information de corrélation spatiale. Dans ces circonstances, certains défauts sont mal détectés, comme illustré par la figure 2.10.

Le second problème découle plutôt de la procédure d'extraction choisie. En effet, notamment pour les pixels isolés, la détection n'est pas réalisée à la même échelle que la visualisation : on utilise une ROI de détection de taille 15×15 pixels, alors que la ROI de visualisation est beaucoup plus grande 61×61 . Ceci entraîne alors des résultats différents entre détection et localisation, particulièrement observables sur les défauts de type pixel isolé (dernière ligne de la figure 2.10). On observe que certains pixels isolés, détectés lors de l'étape de détection, ne le sont plus après l'étape de localisation. Il s'agit de pixels de niveau faible (ressortant du bruit de fond de moins de $4.5\sigma_{bruit}$), souvent à la limite du perceptible. Leur non détection peut être considérée comme acceptable.

2.5.4 Conclusion et remarques sur la localisation

L'algorithme de localisation présenté ici réalise une détection simple s'appuyant sur la statistique de chaque imagerie présentée. Elle utilise une estimation de segmentation en deux zones (défaut et fond) données par l'étape précédente de détection. Le seuil fixé pour réaliser cette étape permet sous l'hypothèse de bruit blanc, de n'obtenir pas plus d'une fausse alarme par imagerie. En pratique, le bruit étant plutôt filtré basse fréquence, on obtient un nombre de fausses alarmes encore plus réduit. Le seuil de la méthode de localisation pourrait donc être encore baissé, jusqu'à par exemple couvrir les problèmes de non détection des pixels isolés.

2.6 Débruitage de la zone défautueuse

2.6.1 Objectifs et enjeux

L'étape de débruitage doit être considérée comme une étape visant à faire une estimation du défaut, et en particulier de son niveau par rapport au bruit de fond. En effet, cette information est importante pour la suite du processus, dont le but sera d'introduire les défauts détectés dans un bruit de fond choisi

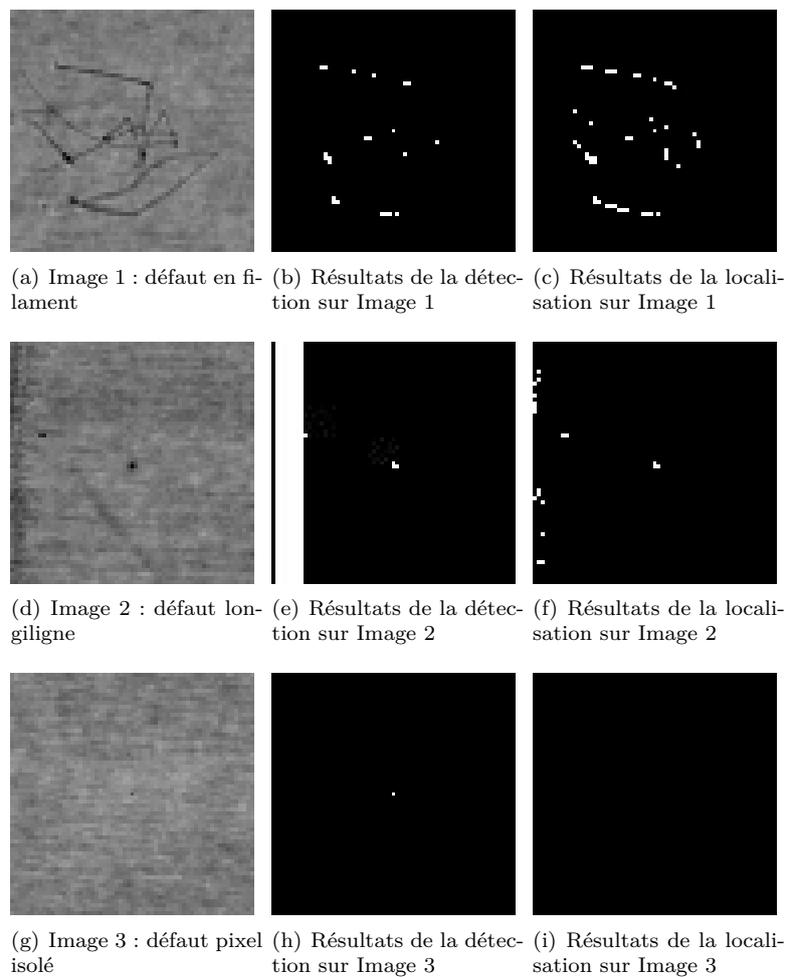


FIGURE 2.10 – La méthode de localisation proposée souffre principalement de deux défauts : le manque de sensibilité en détection sur défauts corrélés (deux premières lignes) et l'échelle de calcul pas forcément bien adaptée (dernière ligne : non détection d'un pixel isolé pourtant visible). A chaque fois on a représenté : l'image issue de l'algorithme de capture, le résultat associé de l'algorithme de détection et le résultat de l'algorithme de localisation appliqué sur la capture.

(un contexte) afin de visualiser les impacts visuels des défauts dans plusieurs configurations. L'estimation du niveau du défaut est également utile pour pouvoir le mesurer. L'objectif du débruitage de la zone défectueuse est donc de s'affranchir au *maximum* du contexte dans lequel le défaut a été détecté, pour pouvoir le replacer dans des contextes différents, et le mesurer. La méthode de débruitage s'appuie sur la connaissance *a priori* définie par la méthode précédente de localisation, à savoir qu'on connaît avec une bonne précision les pixels composant le défaut, et les pixels composant le bruit de fond.

Les principales difficultés de cette tâche sont de gérer les problèmes de défauts à localisation défaillante (en particulier, localisation partielle des pixels les composant, voir figure 2.10(c)), et de fournir à la méthode d'introduction dans le contexte une estimation pertinente du défaut, en éliminant l'effet du bruit de fond. Si cette difficulté n'est pas gérée, l'introduction dans un nouveau contexte pourra être faussée, et le défaut aura une visibilité mal estimée.

2.6.2 Quelques méthodes existantes

Les techniques de débruitage ont fait l'objet d'un nombre important de travaux. Pour traiter l'étape de débruitage dans notre cadre, nous avons testé deux techniques qui font partie des techniques les plus employées dans le domaine. La première est la technique simple de filtrage de Wiener [34, 76]. On considère l'image d'origine $i(x, y)$ à débruiter et sa transformée de Fourier qui sera notée $I(u, v)$. Cette image contient une image du défaut $d(x, y)$, de transformée de Fourier $D(u, v)$, dont on veut une estimation. Elle contient également un bruit $b(x, y)$, de transformée de Fourier $B(u, v)$ qu'on veut éliminer afin d'avoir une bonne estimation du défaut. On peut écrire l'image $I(x, y)$ sous la forme suivante :

$$i(x, y) = d(x, y) + b(x, y) \quad (2.26)$$

Une adaptation du filtre de Wiener au problème de débruitage mène à un filtrage par un filtre $f(x, y)$, dont la réponse fréquentielle $F(u, v)$ est donnée par :

$$F(u, v) = \frac{|I(u, v)|^2}{|I(u, v)|^2 + |B(u, v)|^2} \quad (2.27)$$

On obtient alors l'estimation du défaut $\hat{d}(x, y)$ par la formule de filtrage suivante :

$$\hat{D}(u, v) = F(u, v)I(u, v) \quad (2.28)$$

Pour appliquer le filtre de Wiener dans notre cas, nous avons estimé la densité spectrale du bruit $|B(u, v)|^2$ en découpant l'image à filtrer en blocs de taille 32×32 pixels, l'image d'origine $i(x, y)$ étant de taille 300×300 pixels. Cette densité est déterminée en moyennant fréquence à fréquence les transformées de Fourier de ces blocs. Les résultats de ce filtrage sont présentés dans la figure 2.11 On observe, sur les cas réels, que cette méthode conserve globalement le contenu fréquentiel des défauts, mais réalise un débruitage très faible.

Une alternative au filtrage de Wiener a été proposée par Lee[55, 57] dans les années 80. Son principal avantage est sa simplicité d'implémentation et sa rapidité, ce qui permet de l'adapter à des images de grande taille. Elle repose sur un calcul de statistique locale afin d'adapter le filtrage au contenu haute fréquence de l'image. Chaque pixel de l'image d'origine va être remplacé par la somme de deux termes : le premier est une estimation de la moyenne locale autour du pixel considéré, et l'autre est un terme de contraste auquel on applique un coefficient multiplicatif. On calcule donc $\hat{d}(x, y)$ tel que :

$$\hat{d}(x, y) = \mu + (x - \mu) \frac{\sigma^2}{\sigma^2 + \sigma_b^2} \quad (2.29)$$

où μ et σ^2 sont respectivement une moyenne locale et une variance locale calculées sur un voisinage de chaque pixel, et σ_b^2 est une estimation de la variance du bruit de fond, déterminée sur toute l'image. Le taille du voisinage pour calculer les statistiques locales est un paramètre de la méthode. Le coefficient

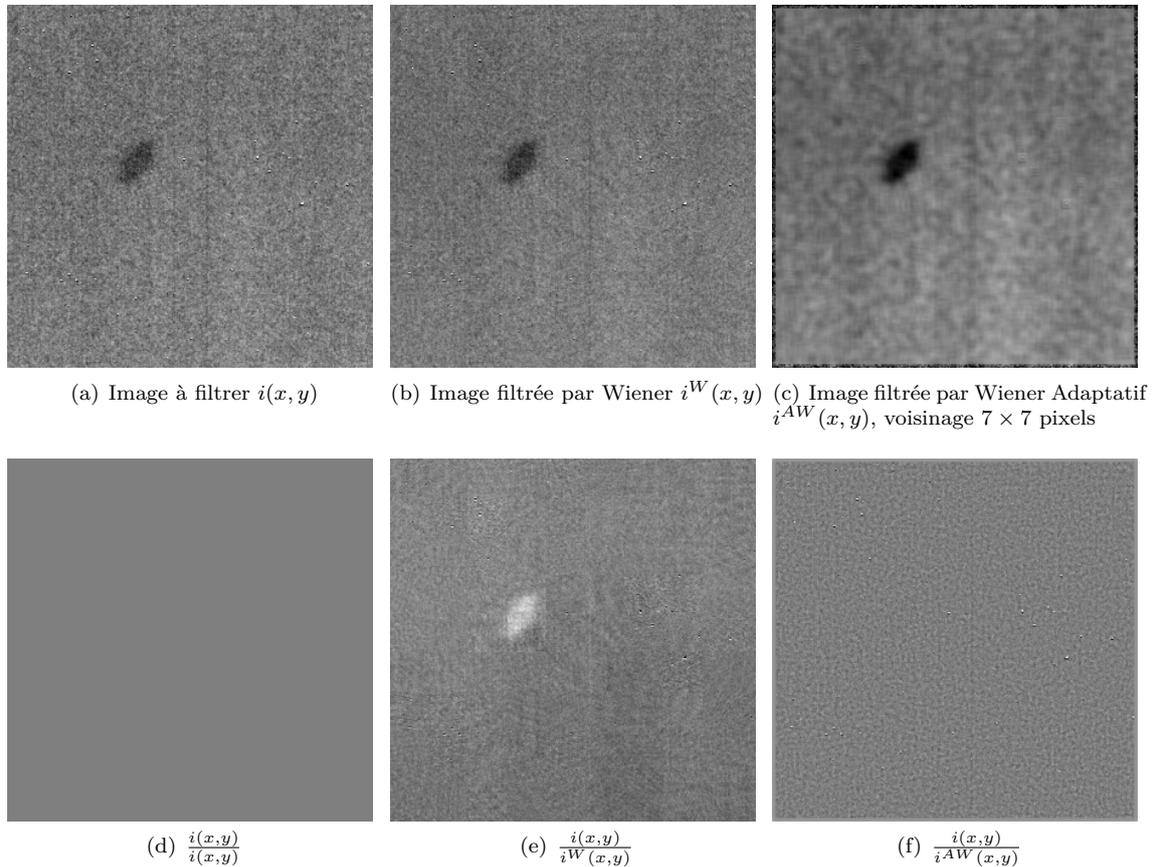


FIGURE 2.11 – Comparaison des méthodes de débruitage type Wiener et Wiener adaptatif sur défaut réel. Le filtrage de Wiener est inefficace : une partie de la tache noire a été filtrée en même temps que le bruit. Le rapport signal à bruit (RSB) pour ce défaut, à l'origine de 45dB (sur $i(x, y)$, figure 2.11(a)) est à 43dB après filtrage (figure 2.11(b)). Pour le même défaut, la méthode de Wiener adaptative donne un RSB de 55,5dB (figure 2.11(c)). Néanmoins, sur le défaut de type pixel isolé (partie droite de l'image d'origine), la méthode n'est pas adaptée puisqu'elle les filtre entièrement, en même temps que le bruit. La ligne du bas représente les ratios entre image d'origine et image débruitée. On s'attend idéalement à obtenir une image proche de celle de la figure 2.11(d) qui correspond à la division de l'image d'origine par un débruitage parfait. On constate que la méthode de Wiener estime mal une partie de la tache noire (encore visible sur l'image de ratio 2.11(e)). La méthode de Wiener adaptative, quant à elle, a bien estimé la tache noire, mais on constate que les pixels isolés ont été très mal estimés (il se retrouvent entièrement dans l'image de ratio 2.11(f)). Cette observation permet facilement de vérifier la qualité du débruitage des deux méthodes.

multiplicatif $k = \frac{\sigma^2}{\sigma^2 + \sigma_b^2}$ permet de régler le poids donné au terme de contraste par rapport au terme de moyenne (basse fréquence). Dans les zones où le contenu haute fréquence est prédominant par rapport à celui du bruit ($\sigma^2 \gg \sigma_b^2$), k va être proche de 1 et donc le poids du terme de contraste va être fort. À l'inverse, dans les zones de faible contraste ($\sigma^2 \ll \sigma_b^2$), le terme prédominant sera le terme basse fréquence.

Ce filtrage se montre très efficace sur les défauts basse fréquence (voir la tache sombre de la figure 2.11), mais inadapté aux défauts de type pixels isolés : l'impact d'un pixel isolé défectueux sur la variance, même en considérant un voisinage réduit, n'est en général pas suffisant pour augmenter significativement le terme k , conduisant à une élimination de ce type de défaut par le filtrage.

2.6.3 Méthode proposée

Les méthodes présentées ci-dessus sont des méthodes générales de débruitage, sans connaissance *a priori* sur l'image à débruiter. Dans notre cas, le but de l'étape de débruitage est uniquement de proposer une normalisation du défaut de manière à ce qu'il puisse être réintroduit dans un contexte différent (un bruit de valeur moyenne et de variance différente de celui dans lequel a été estimé le défaut) pour y être mesuré. Dans les cas réels, les défauts rencontrés sont principalement des perturbations sur le niveau moyen attendu dans l'image : l'image analysée est sous la forme $i(x, y) = d(x, y) \cdot b(x, y)$ où $d(x, y)$ est la valeur attendue sur chaque pixel (dans le cas d'une image plate, $d(x, y)$ est constante et proportionnelle à la dose appliquée) et $b(x, y)$ un bruit issu des différentes perturbations auxquelles est soumis le détecteur. Les défauts viennent perturber la valeur de $d(x, y)$. Le modèle multiplicatif adopté implique que le bruit de fond est multiplié à l'image attendue. Ceci correspond au type d'images traitées : ce sont des images sous éclairément X, corrigées des défauts électroniques (correction d'offset, voir section 1.4). Les défauts électroniques sont des défauts additifs, d'où la correction par soustraction d'une image d'offset, et les autres défauts, visibles sous éclairément X, sont généralement multiplicatifs.

L'objectif de la méthode de débruitage est donc de retrouver $d(x, y)$ à partir de $i(x, y)$ de manière à pouvoir introduire le défaut dans des nouveaux contextes. La méthode proposée opère en deux étapes :

1. extraction de la composante basse fréquence⁴ du défaut : en utilisant un filtre passe bas, on filtre le bruit de fond sur la surface de défaut définie par l'étape de localisation. On obtient l'image $d_{BF}(x, y)$. Dans notre application, le filtre choisi est une moyenne glissante 5×5 pixels, opérant uniquement sur les pixels définis par la localisation et stockés dans $L(x, y)$.
2. Estimation de la composante haute fréquence : après avoir corrigé l'image à traiter par la basse fréquence préalablement estimée, une détection semblable à celle utilisée pour la localisation est effectuée. Les valeurs des pixels détectés, stockés dans $d_{HF}(x, y)$ sont une estimation de la partie haute fréquence de $d(x, y)$.

La figure 2.12 illustre la méthode utilisée pour effectuer le débruitage. Cette méthode comporte trois traitements principaux : la transformation de la carte de détection originale $DM(x, y)$ en $DM_{dil}(x, y)$ (voir section suivante), l'estimation de la composante basse fréquence, et l'estimation de la composante haute fréquence.

2.6.3.1 Transformation de $DM(x, y)$

L'étape précédente de localisation fournit une carte des pixels défectueux $DM(x, y)$. Cette carte est déterminée pixel à pixel. Pour pouvoir estimer la basse fréquence sur des groupements de pixels défectueux, il est nécessaire de dilater cette carte de pixels de manière à prendre en compte les pixels appartenant à ce groupement, et non détectés comme défectueux. Pour cela, nous avons choisi d'effectuer une fermeture morphologique [34] (dilatation puis érosion) sur l'image $DM(x, y)$ afin de créer une carte des défauts élargie $DM_{dil}(x, y)$. Le choix de la fermeture, plutôt qu'une simple dilatation permet de

4. Dans toute la suite, on parlera de composante basse fréquence et haute fréquence, par abus de langage. En effet ces termes sont d'habitude réservés aux méthodes de filtrages linéaires, et donc incompatibles avec le modèle multiplicatif choisi ici

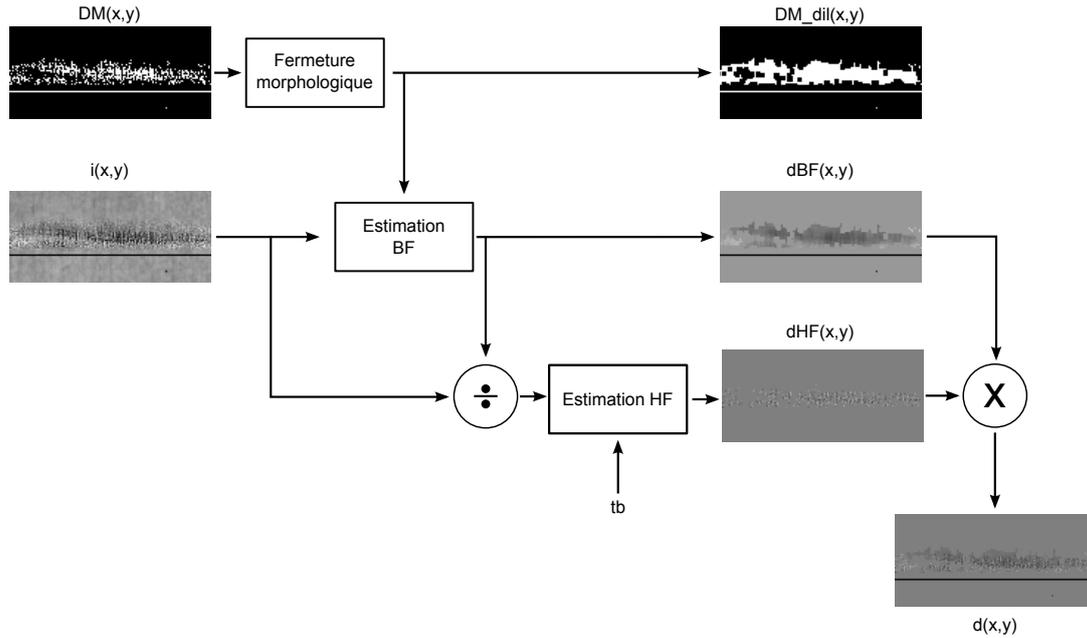


FIGURE 2.12 – Méthode de débruitage : l'estimation du défaut est effectuée par calcul d'une composante basse-fréquence puis de la composante haute fréquence associée. L'estimation basse fréquence est une moyenne glissante effectuée sur l'image d'origine sur les points définis par le masque de détection. L'estimation de la haute fréquence s'appuie sur une méthode de détection similaire à celle utilisée à l'étape de localisation. L'estimation finale est le produit des deux composantes basse fréquence et haute fréquence.

conserver les pixels isolés sans les étendre. Il est important que la méthode ne transforme pas ces pixels en des zones plus étendues, car le défaut serait alors mal estimé par la suite. La fermeture a été effectuée

avec un élément structurant 3×3 pixels $e = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \mathbf{1}_3$ ⁵

2.6.3.2 Estimation de la composante basse fréquence du défaut

L'estimation de la basse fréquence utilise les images $i(x, y)$ et $DM_{dil}(x, y)$. Le principe est d'effectuer une moyenne glissante des pixels de $i(x, y)$ en utilisant uniquement les pixels à 1 de $DM_{dil}(x, y)$. La moyenne glissante est dans notre cas effectuée sur un voisinage 5×5 pixels ($e = \mathbf{1}_5$). Pour calculer l'estimation basse fréquence $d_{BF}(x, y)$ de $d(x, y)$, on réalise donc les opérations suivantes :

1. masquage de l'image d'origine par la carte des défauts dilatée, afin de ne conserver que les zones défectueuses :

$$i_1(x, y) = i(x, y) \cdot DM_{dil}(x, y) \quad (2.30)$$

2. lissage de l'image pour déterminer la composante basse fréquence. A l'aide du noyau e on calcule en chaque point la somme locale des pixels voisins :

$$i_2(x, y) = (i_1 \otimes e)(x, y) \quad (2.31)$$

3. afin de normaliser la somme déterminée précédemment, il faut calculer le nombre de pixels sommés.

5. Afin de préserver les pixels détectés au bords de l'image $DM(x, y)$, il est préférable de l'inclure dans une image plus grande (n lignes en plus en haut et en bas de l'image, et n colonnes en plus à droite et à gauche, lorsqu'on utilise un élément structurant de taille $n \times n$), les bords étant remplis avec des 0.

Pour cela on utilise DM_{dil} en lui appliquant la même opération de filtrage :

$$nb(x, y) = (DM_{dil} \otimes e)(x, y) \quad (2.32)$$

4. on réapplique le masquage à i_3 de manière à ne considérer que les pixels détectés :

$$i_3(x, y) = i_2(x, y) \cdot DM_{dil}(x, y) \quad (2.33)$$

5. enfin, on normalise cette dernière image par le nombre de points sommés afin d'avoir, en chaque pixel, une valeur moyenne des pixels voisins :

$$d_{BF}(x, y) = \frac{i_3(x, y)}{nb(x, y)} \quad (2.34)$$

On a noté $A \otimes B(x, y)$ la corrélation de A par B et $A(x, y) \cdot B(x, y)$ est le produit pixel à pixel de A par B .

Dans l'optique de réintroduire le défaut dans un autre contexte (voir étape suivante de la chaîne d'analyse), il est judicieux de normaliser $d_{BF}(x, y)$ par la moyenne du bruit de fond de l'image traitée $m = \text{médiane} \{i(x, y) | DM(x, y) = 0\}$. On obtient alors $d_{BF}^{norm} = \frac{d_{BF}}{m}$.

2.6.3.3 Estimation de la composante haute fréquence du défaut

L'estimation de la composante basse fréquence est effectuée sur l'image d'origine corrigée des défauts haute fréquence, notée $i_{HF}(x, y)$:

$$i_{HF}(x, y) = \begin{cases} \frac{i(x, y)}{d_{BF}(x, y)} & \text{si } d_{BF}(x, y) \neq 0 \\ i(x, y) & \text{sinon} \end{cases} \quad (2.35)$$

Sur cette image des défauts haute fréquence, une détection est effectuée de manière à isoler les pixels défectueux. Cette détection est identique à celle effectuée lors de l'étape de localisation, sauf qu'on l'applique à $i_{HF}(x, y)$ et non pas à $i(x, y)$. On calcule préalablement, à l'aide de l'image $DM(x, y)$, la statistique du bruit :

$$m_{HF} = \text{médiane} \{i_{HF}(x, y) | DM(x, y) = 0\} \quad (2.36)$$

$$\sigma_{HF} = \frac{3}{2} \text{médiane} \{|i_{HF}(x, y) - m_{HF}| | DM(x, y) = 0\} \quad (2.37)$$

L'image $d_{HF}(x, y)$ est alors déterminée par :

$$d_{HF}(x, y) = \begin{cases} i_{HF}(x, y) & \text{si } \left| \frac{i_{HF}(x, y) - m_{HF}}{\sigma_{HF}} \right| \geq t_b \\ 1 & \text{sinon} \end{cases} \quad (2.38)$$

Une fois les deux composantes estimées, on reconstitue l'estimation du défaut par un simple produit :

$$\hat{d}(x, y) = d_{HF}(x, y) \cdot d_{BF}(x, y) \quad (2.39)$$

Les figures 2.13 et 2.14 illustrent la méthode de débruitage choisie sur défaut réel. On observe une estimation nettement meilleure comparée aux méthodes présentées auparavant. En particulier, la méthode choisie permet de capturer à la fois les défauts de type basse fréquence (la tache sombre dans l'image choisie pour illustration) et les pixels défectueux isolés. L'estimation du défaut peut facilement être normalisée par rapport au niveau du bruit de fond de l'image, facilitant ainsi l'introduction de ce défaut dans un nouveau contexte.

A partir de ce constat, nous avons imaginé une méthode d'estimation du défaut qui permet d'estimer

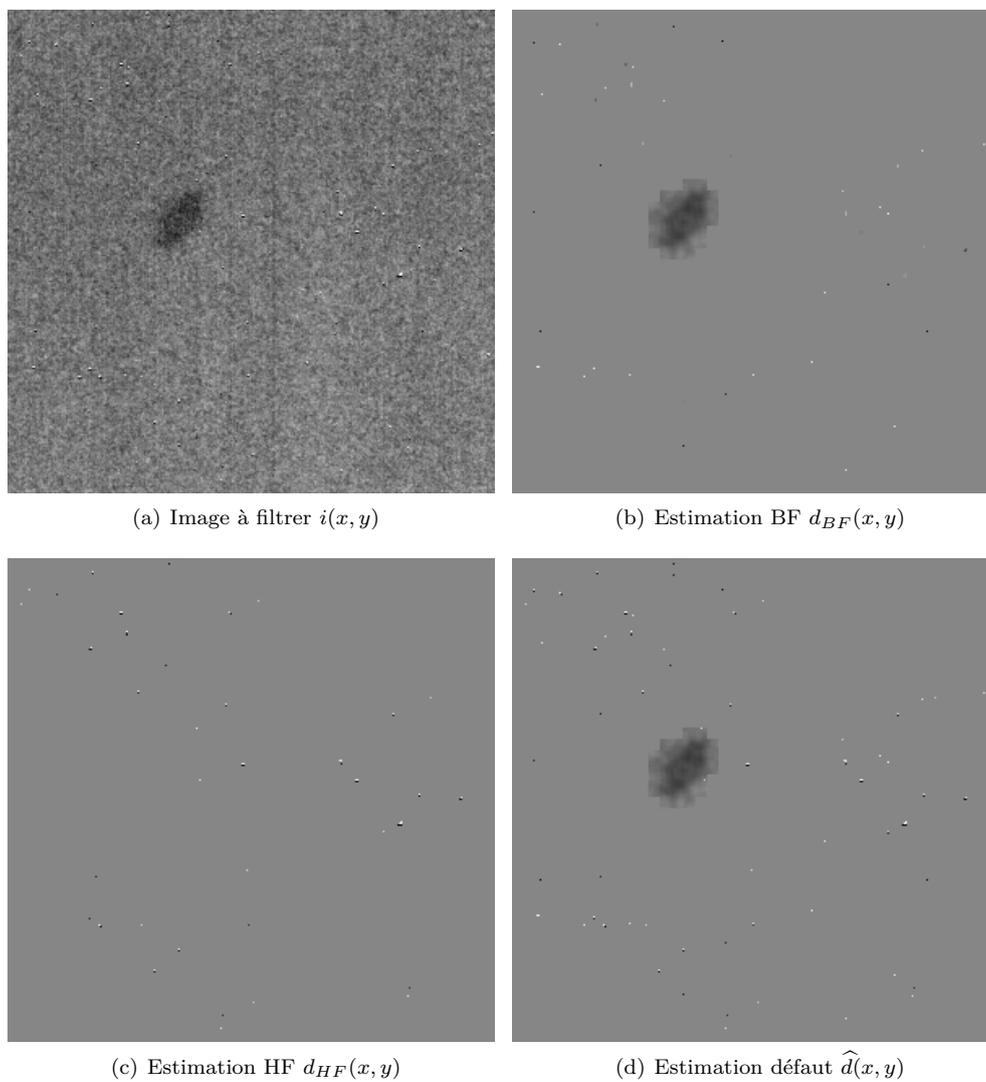


FIGURE 2.13 – Méthode de débruitage : illustration sur défaut réel. La tache sombre est estimée par la composante basse fréquence, comme les pixels isolés. La composante haute fréquence permet l'estimation des groupements de pixels défectueux.

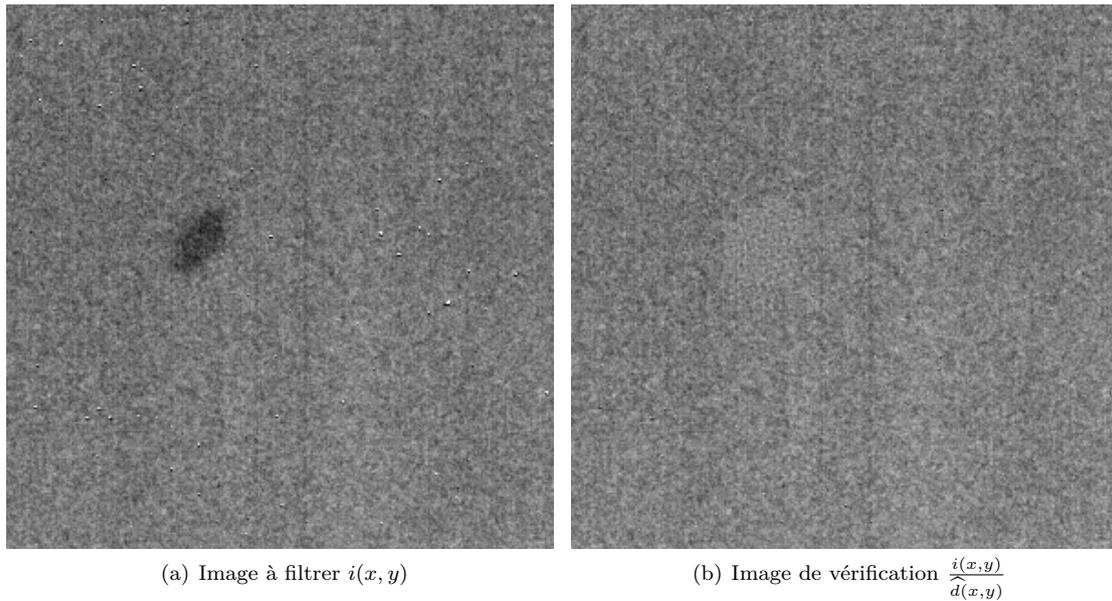


FIGURE 2.14 – Vérification de la méthode de débruitage : on divise l'image d'origine par l'image d'estimation du défaut. On observe uniquement du bruit de fond dans l'image résultante, témoignant d'une bonne estimation du défaut. Une tache de structure légèrement différente apparaît à la place de la tache sombre de l'image d'origine : à cet endroit, l'estimation basse fréquence a modifié la structure du bruit car une partie du bruit a été capturé dans la composante basse fréquence du défaut.

les trois effets à l'aide d'une mesure adaptée. Cette méthode s'appuie sur la connaissance *a priori* sur la composition du défaut, donnée par l'étape de localisation. La méthode utilise un filtrage passe bas afin d'isoler les déviations locales de valeur moyenne, ainsi que les pixels défectueux. Un filtrage passe haut vient ensuite extraire le contenu haute fréquence éventuellement présent du fait de la troisième catégorie de défaut : les déviations locales de variance.

2.6.4 Conclusions sur le débruitage

Trois méthodes de débruitage simples, et adaptées à notre problématique, ont été présentées. La dernière méthode, conçue de manière à être facilement intégrée à la chaîne d'analyse (utilise l'*a priori* fourni par l'étape de localisation, et prépare l'estimation du défaut pour l'étape d'introduction dans un nouveau contexte), donne des résultats satisfaisants sur nos images. De nombreuses méthodes de débruitage alternatives existent, notamment dans le cas d'un bruit multiplicatif. Les méthodes de débruitage par transformées en ondelettes [18, 61, 42] donnent de bons résultats sur images naturelles. Une autre catégorie de techniques vise à améliorer les propriétés de débruitage des ondelettes en ajoutant un filtrage multi-directionnel (en lieu et place des filtres horizontaux et verticaux des ondelettes) : il s'agit des algorithmes s'appuyant sur les contourlets [23, 60]. Cette catégorie de techniques est utilisée avec succès pour le débruitage d'images naturelles corrompues par du bruit. Dans notre cas, quelques essais utilisant les décompositions en ondelettes ont été menés (notamment pour la détection, voir le chapitre 4). Pour le débruitage extrême que nous voulons réaliser (on veut totalement éliminer le bruit de fond), les décompositions en ondelettes posent de nombreux problèmes, notamment des oscillations importantes : en effet, pour pouvoir débruiter suffisamment, il ne faut conserver que très peu de coefficients d'ondelettes non nuls, et la reconstruction du défaut présente donc de nombreuses oscillations : on observe le filtre d'ondelettes utilisé. La méthode *ad hoc* définie nous semble être une bonne base, pouvant éventuellement être améliorée, mais donnant d'ores et déjà des résultats satisfaisants sur toutes les images analysées.

2.7 Introduction du défaut dans un contexte d'observation

2.7.1 Objectifs et enjeux

L'étape de débruitage fournit une estimation normalisée des défauts. Dans cette étape d'introduction dans un nouveau contexte, on veut introduire chaque défaut détecté dans un bruit de fond ou dans une image clinique déterminée. Le but est de pouvoir, à partir d'une estimation d'un défaut réalisée sur une image technique, estimer l'impact visuel qu'aura ce défaut dans d'autres conditions, où dans d'autres contextes. Cet impact est en effet très variable selon la dose, et aussi selon la localisation du défaut dans l'image. Dans une image clinique notamment, cet impact diffère selon que le défaut est situé dans une zone exposée aux rayons X, où dans une zone non exposée (comme les zones sur les bords du détecteurs, se situant sous la collimation du tube à rayons X). Par une technique simple, on veut vouloir simuler ces situations d'observation différentes.

Dans notre cas, nous traiterons uniquement l'introduction dans un nouveau bruit de fond. L'introduction dans une image clinique suit le même principe général, mais des précautions supplémentaires doivent être prises car il faut appliquer les prétraitements (notamment : rehaussement de contours, « négatif » des images) qui sont appliqués par l'utilisateur final des détecteurs dans le cadre d'une utilisation clinique.

L'étape précédente donne donc une estimation du défaut. Pour pouvoir l'introduire dans un nouveau contexte, il faut :

1. calculer ce contexte à l'aide des informations disponibles,
2. introduire le défaut dans le contexte calculé.

Cette étape ne pose pas de difficulté majeure, car l'étape de débruitage fournit une estimation normalisée du défaut, prête à être introduite dans un nouveau contexte.

2.7.2 Calcul du contexte

La première étape pour introduire un défaut dans un nouveau contexte est de calculer une image de contexte de la taille de l'image d'estimation du défaut dont nous disposons. Dans notre cas, les contextes sont des bruits de fond, considérés à des doses différentes, dont on dispose de l'image de densité spectrale de puissance (NPS , voir images de la figure 1.7). On note $NPS(u, v)$ l'image de NPS du bruit que l'on veut synthétiser. On va dans un premier temps calculer la réponse impulsionnelle du bruit à synthétiser (notée $PSF(x, y)$, pour Point Spread Function), puis calculer le bruit dans le domaine réel⁶ $b(x, y)$ en convoluant la PSF avec une réalisation d'un bruit blanc $b_{blanc}(x, y)$.

$$PSF(x, y) = TF^{-1} \left\{ \sqrt{NPS(u, v)} \right\} \quad (2.40)$$

$$b(x, y) = (b_{blanc} * PSF)(x, y) \quad (2.41)$$

où $*$ est le produit de convolution⁷.

En prenant le bruit blanc $b_{blanc}(x, y)$ centré réduit, on peut régler la moyenne et la variance du bruit synthétisé respectivement à la valeur m et σ^2 pour obtenir $b_{m, \sigma}$ par :

$$b_{m, \sigma}(x, y) = \frac{\sigma}{\|PSF\|_2} b(x, y) + m \quad (2.42)$$

En pratique, on limite la PSF à une taille de 7×7 pixels, car les corrélations du bruit dans un rayon supérieur à 3 pixels sont négligeables.

6. L'intérêt ici d'effectuer le filtrage dans le domaine réel est que cette méthode s'adapte facilement à la création de contextes de taille quelconque. L'équivalent dans Fourier est de multiplier la transformée de Fourier de la réalisation de bruit blanc par la racine de notre image de NPS . Ceci nécessite de redimensionner l'image de NPS dont nous disposons (image 256×256 pixels) à la taille du contexte que l'on veut créer. Ce redimensionnement implique des interpolations et peut entraîner l'introduction de nouvelles fréquences spatiales, qui n'étaient pas présentes dans l'image de NPS . Il est donc préférable d'utiliser un filtrage dans le domaine réel pour éviter les artefacts de l'interpolation dans Fourier, le temps de calcul étant rendu acceptable par le support limité de la PSF .

7. $(A * B)(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A(m, n)B(x - m, y - n)$.

Les images de *NPS* sont calculées lors de certaines phases de test des détecteurs, et sont donc disponibles pour chaque produit. Les *NPS* sont calculés sur images corrigées, en découpant l'image entière en blocs de 256×256 pixels et en moyennant les modules des transformées de Fourier de ces blocs. Ceci permet de s'affranchir d'effets locaux, comme la présence d'éventuels défauts, dans la détermination de l'image de *NPS* : de cette manière, seul le spectre du bruit est calculé.

La figure 2.15 est un exemple de simulation de deux bruits à deux doses différentes : une dose nulle, et une dose de $1\mu Gy$. On constate que le bruit électronique (bruit à dose nulle) est négligeable à $1\mu Gy$. A cette dose, le bruit est plutôt basse fréquence, alors qu'il est proche d'être blanc pour une dose nulle. Dans ce dernier cas, on observe quand même des corrélations ligne et, dans une moindre mesure, colonne du bruit, dues à l'électronique de lecture.

2.7.3 Introduction du défaut dans le contexte

Le contexte ayant été défini, on veut maintenant introduire chaque défaut dans ce contexte. Les défauts débruités ont été centrés autour de la valeur 1, ces défauts étant multiplicatifs par rapport au bruit. Pour les introduire dans un nouveau contexte de bruit, il suffit donc de multiplier ces défauts par l'image de bruit. On appelle $b(x, y)$ le contexte défini dans l'étape précédente, et $\hat{d}(x, y)$ l'image de défaut débruitée définie lors de l'étape de débruitage. L'image du défaut dans son nouveau contexte, notée $i(x, y)$ est alors donnée par :

$$i(x, y) = \hat{d}(x, y) \cdot b(x, y) \quad (2.43)$$

La figure 2.16, présente un exemple d'introduction dans un nouveau contexte. L'étape nécessite, au niveau de la création du contexte, de définir la dose à partir de laquelle calculer le spectre du bruit (avec la technique de la section précédente), et les statistiques de ce bruit : moyenne et variance. Chaque produit ayant un couple dose/statistique du bruit différent, et variant en fonction des conditions d'utilisation du détecteur, la définition de ces grandeurs est un paramètre de la méthode d'introduction dans un nouveau contexte. D'une manière générale, du fait de l'effet multiplicatif des défauts, ceux-ci sont de plus en plus visibles au fur et à mesure que la dose augmente. Sur la figure 2.16, on a introduit l'artefact dans deux contextes de bruit différents. On constate évidemment que la visibilité du défaut dépend de l'énergie du bruit dans lequel il a été introduit, mais également du spectre de bruit, d'où l'intérêt de calculer chaque contexte à l'aide des densités spectrales de puissance du bruit.

2.8 Mesure de visibilité du défaut

2.8.1 Objectifs et enjeux

Les précédentes étapes ont permis de définir une estimation du niveau du défaut (étape de débruitage) et une image de contexte dans lequel on observe le défaut. La dernière étape de la chaîne d'analyse est d'affecter au couple (défaut, contexte d'observation) une valeur numérique représentative de la visibilité du défaut. Une étude spécifique sur la définition de la mesure de visibilité sera menée dans le chapitre 5. Nous donnons ici une première mesure, pouvant être considérée comme une bonne mesure de base. L'objectif de cette mesure est de quantifier la visibilité de chaque défaut, en étant la plus proche possible de l'évaluation qui serait faite par l'œil humain.

La bonne corrélation entre cette mesure objective et l'évaluation, subjective, par l'œil humain est la principale difficulté de la définition d'une mesure de visibilité, notamment dans les cas d'évaluation de défauts hétérogènes. En première approximation, nous donnons une mesure s'appuyant sur un calcul d'énergie, qui est la base de la mesure actuellement utilisée. Cette mesure s'appuie néanmoins sur la connaissance de l'estimation du défaut, ce qui est un plus par rapport aux techniques de base d'évaluation «en aveugle» de la visibilité.

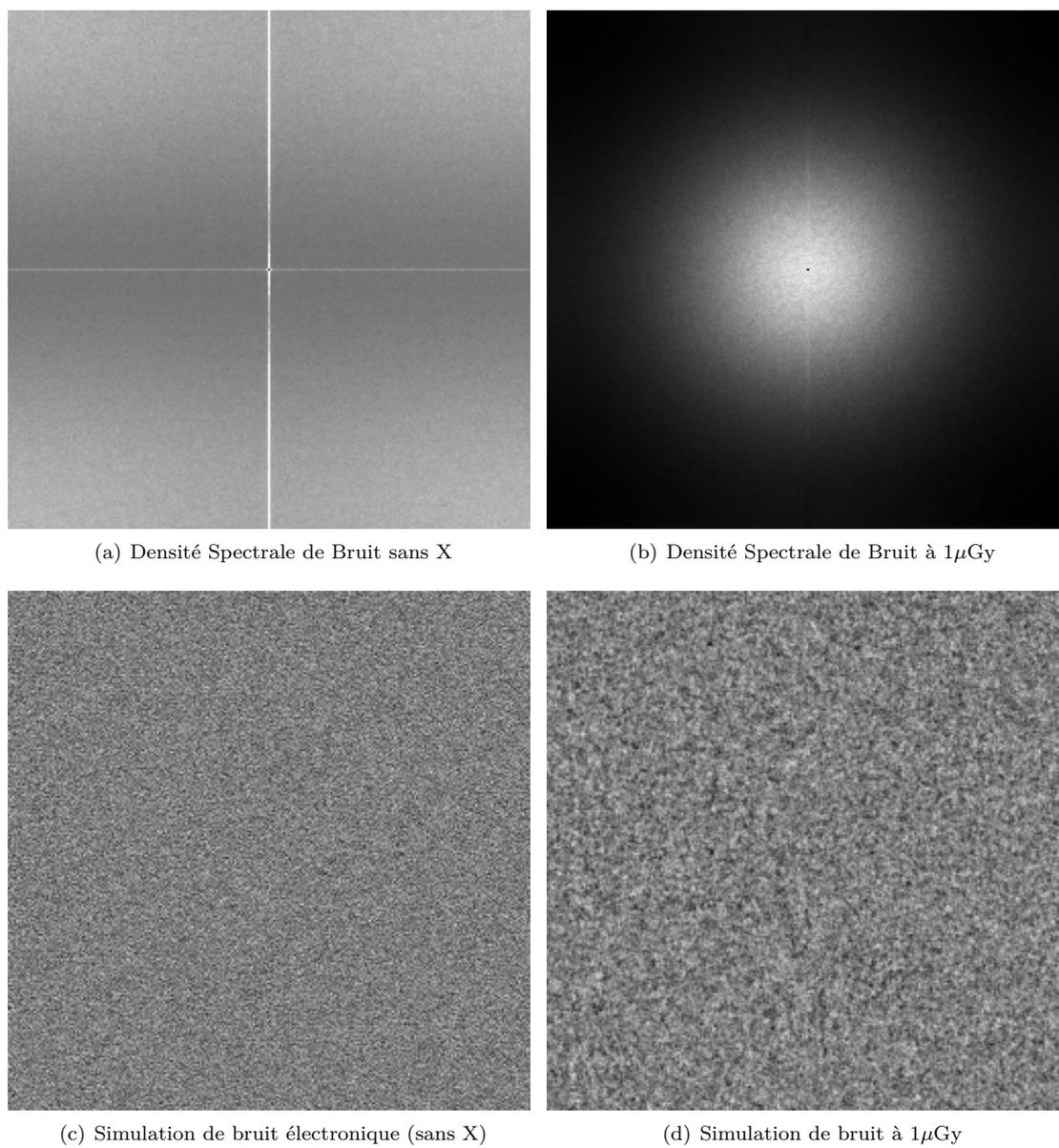


FIGURE 2.15 – Densités Spectrales de Bruit d'un détecteur numérique

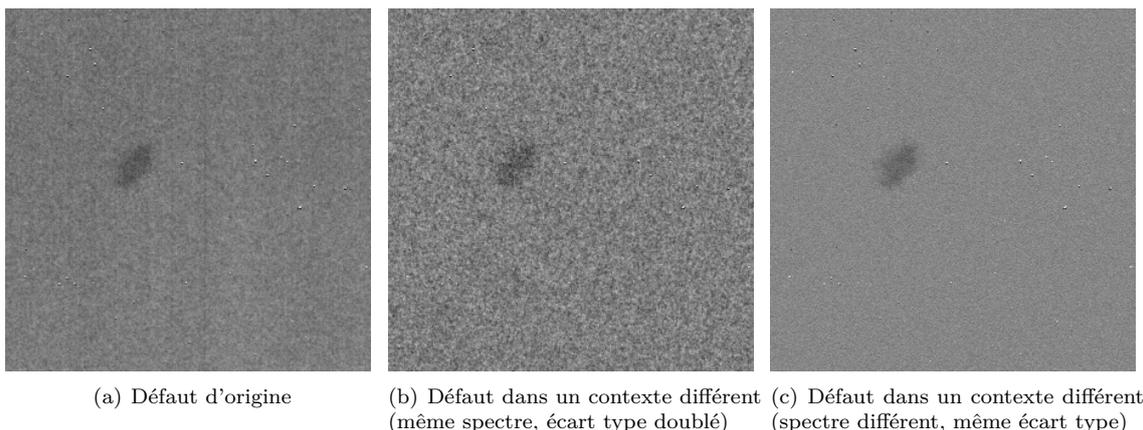


FIGURE 2.16 – Illustration d'un défaut introduit dans un nouveau contexte. Le défaut d'origine (à gauche) a été introduit dans deux contextes différents. Le premier (image 2.16(b)) est un contexte présentant un bruit d'énergie plus forte (spectre équivalent et un écart-type 2 fois plus fort que l'image d'origine). Le défaut est, dans ce nouveau contexte, beaucoup moins visible. Le second (image 2.16(c)) est un contexte présentant le même écart-type que le bruit de l'image d'origine, mais avec un spectre différent (on a pris un bruit correspondant à un bruit électronique). La visibilité du défaut augmente dans ce contexte.

2.8.2 Méthode proposée

On dispose de l'estimation du défaut $\hat{d}(x, y)$ et du contexte de bruit $b(x, y)$. La mesure de visibilité proposée consiste à :

1. normaliser $\hat{d}(x, y)$ en fonction du contexte défini,
2. calculer l'énergie du défaut normalisé $\hat{d}_{norm}(x, y)$,
3. calculer l'énergie moyenne du bruit,
4. calculer la mesure de visibilité, qui est le ratio entre énergie du défaut, et énergie moyenne du bruit.

On appelle $m(d)$ la mesure de visibilité du défaut $d(x, y)$. On a :

$$\hat{d}_{norm}(x, y) = \begin{cases} (\hat{d}(x, y) - 1) \cdot \text{moyenne}\{b(x, y)\} & \text{si } \hat{d}(x, y) \neq 1 \\ 0 & \text{sinon} \end{cases} \quad (2.44)$$

$$m(d) = \frac{\sum_x \sum_y \hat{d}_{norm}^2(x, y)}{\text{Var}\{b(x, y)\}} \quad (2.45)$$

Dans le cas où le contexte défini est un bruit de moyenne nulle, la mesure donnera un résultat très proche de zéro, voir nul, quel que soit le défaut considéré (équation 2.44). La modélisation multiplicative des défauts est dans ce cas inadaptée, il est donc souhaitable de décaler la moyenne du contexte de manière à la ramener à une valeur non nulle, ceci n'ayant aucun impact pour l'évaluation de la visibilité des défauts.

2.8.3 Résultats obtenus et commentaires

La mesure de visibilité décrite plus haut a été appliquée à plusieurs défauts issus de l'étape de ré-introduction dans un nouveau contexte. La figure 2.17 illustre les résultats obtenus sur un ensemble de défauts détectés sur une image acquise à une dose standard de $30\mu Gy$ et réintroduit dans un contexte de bruit électronique, donc à dose nulle. La mesure permet de faire ressortir principalement trois types de défauts :

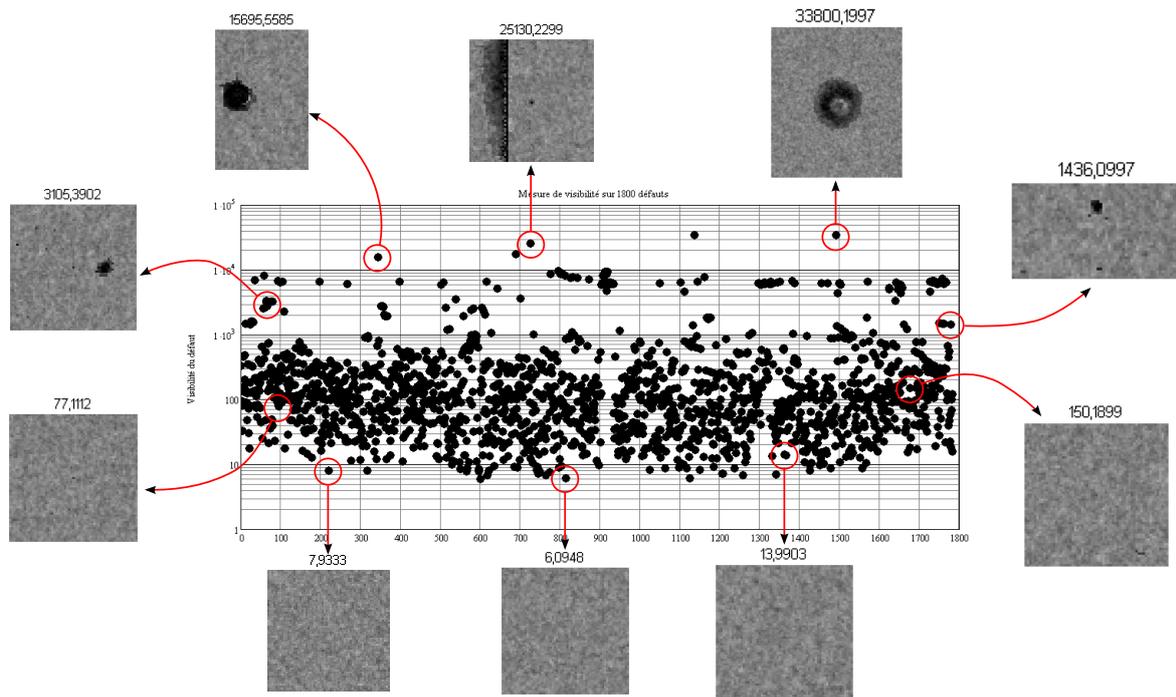


FIGURE 2.17 – Illustration de la mesure de visibilité sur 1800 défauts issus de l’analyse d’une image réelle. Les défauts de la partie supérieure (taches sombres, colonne défectueuse) sont nettement plus visibles que ceux de la partie inférieure, ce qui se traduit sur la mesure par une plus grande visibilité. Au dessus de chaque capture de défaut, on trouve la mesure de visibilité associée.

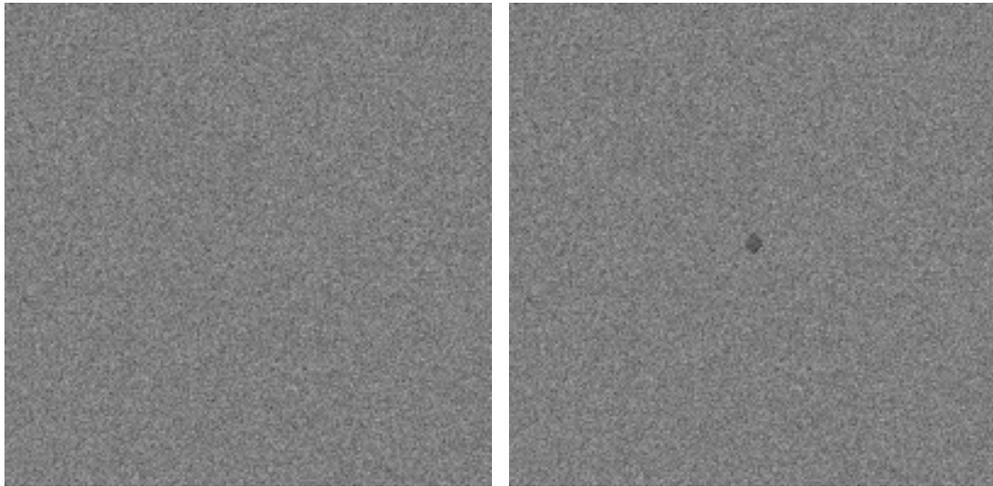
- lignes et colonnes défectueuses : ces défauts, en général très forts en niveau, sont les seuls ressortant clairement lorsqu’on applique la mesure au jeu de défauts considérés,
- taches : ce sont des amas circulaires de pixels défectueux, formant une tache pouvant atteindre une cinquantaine de pixels de diamètre,
- défauts de faible visibilité : ce soit des pixels défectueux soit des combinaisons de plusieurs défauts peu visibles. Il est très difficile de le discriminer du bruit car leur mesure est très faible. En pratique, ils sont en effet moins visibles que les défauts des deux premiers types.

Il faut noter que la mesure a été appliquée aux images telles qu’elles sont représentées à l’écran. Nous avons en effet appliqué le réglage de contraste choisi pour l’observation avant de calculer les mesures de visibilité. Cette étape est absolument indispensable pour obtenir une mesure représentative de la visibilité au sens de l’observation visuelle.

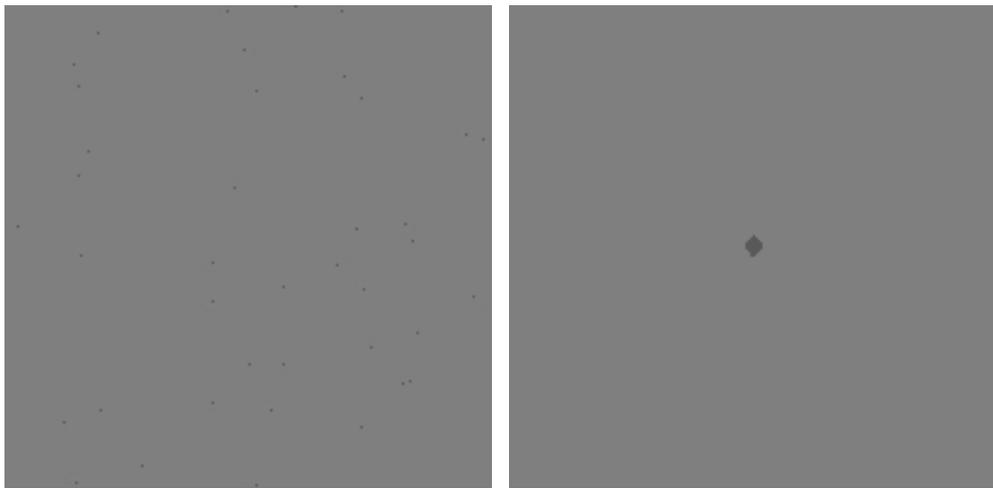
La principale limitation de cette mesure réside dans sa non prise en compte des aspects spatiaux des défauts. La mesure s’appuie uniquement sur un critère statistique, et est donc insensible notamment aux corrélations spatiales des défauts, dont on sait qu’elles jouent un rôle important dans leur visibilité. Cette limitation peut être mise en avant par l’exemple de la figure 2.18. Deux défauts de nature différente se voient affecter la même mesure de visibilité, alors que l’un est parfaitement invisible (les pixels isolés) et l’autre très visible (la tache sombre).

Pour nuancer ceci, dans le cas de défauts de nature différente, il est très difficile, même pour un observateur humain, de donner une quantification cohérente pour plusieurs types de défauts. Néanmoins, l’exemple présenté permet de rendre compte qu’une information essentielle à la mesure n’est ici pas utilisée : la corrélation spatiale.

La méthode présentée ici sera évaluée, et comparée à d’autres méthodes, utilisant, elles, cette information capitale, dans le chapitre 5.



(a) Défaut de type pixels isolés dans un contexte de bruit blanc. $m(d) = 360.5$ (b) Défaut de type tache dans un contexte de bruit blanc. $m(d) = 360.5$



(c) Défaut de type pixels isolés hors contexte (d) Défaut de type tache hors contexte

FIGURE 2.18 – Illustration de la mesure de visibilité sur deux exemples de défauts. Pour les deux types de défauts représentés (respectivement les pixels isolés, et la tache sombre), on obtient une mesure de visibilité strictement identique. Le défaut de type tache (figure 2.18(b)) est néanmoins beaucoup plus visible que les pixels isolés (figure 2.18(a)). Ces derniers sont en fait invisibles sur l'image, ils sont noyés dans le bruit de fond.

2.9 Conclusions sur la chaîne de traitement

Dans ce chapitre, nous avons décrit et illustré un ensemble de techniques permettant d'obtenir les résultats attendus de la chaîne automatique d'analyse et de contrôle des images, à savoir principalement :

- une liste des défauts présents grâce à l'étape de détection,
- une mesure de ces défauts grâce à l'ensemble de méthodes décrites entre la localisation et la mesure de visibilité.

Un autre résultat attendu de la chaîne d'analyse est la classification des défauts en plusieurs catégories. Cette étape peut être abordée en utilisant notamment les résultats des méthodes de localisation et de débruitage qui permettent une caractérisation facile de chaque défaut, et donc une extraction d'attributs aisée en vue d'une classification.

Nous avons vu que certaines méthodes décrites souffrent de limitations impactant la qualité des résultats de la chaîne d'analyse de manière variable. Parmi les limitations les plus critiques, on peut retenir le manque de sensibilité de l'étape de détection, et la difficulté, pour la méthode de mesure de visibilité, d'être pertinente pour des classes variées de défauts . Ces deux limitations nous paraissant importantes, nous avons choisi de chercher d'autres méthodes permettant d'améliorer les résultats de la méthode prise comme méthode de base, décrite dans ce chapitre. Les développements de ces nouvelles méthodes font l'objet des chapitres suivants.

Néanmoins, les méthodes présentées ici ont été appliquées à des cas réels, et ont permis d'obtenir une analyse satisfaisante des défauts présents. En particulier, la méthode d'extraction et de débruitage, qui sont deux contributions importantes de ce chapitre, donnent satisfaction. Nous ne chercherons donc pas à les améliorer par la suite, et nous préférons porter un éclairage particulier sur l'amélioration de la détection et de la mesure de visibilité dans les prochains chapitres.

Deuxième partie

Détection de pixels défectueux

Chapitre 3

Détection automatique de défauts : objectifs et évaluation des performances

Résumé

La détection de défauts est une tâche courante dans le domaine du contrôle qualité, et est un des points centraux de cette thèse. Compte tenu de la multitude de types de défauts à détecter, et des exigences en termes de performances de détection, nous sommes amenés à développer plusieurs algorithmes de détection. Ainsi, l'évaluation et la comparaison des performances de ces algorithmes est un point à traiter avec la plus grande attention. Ce chapitre décrit une méthodologie visant à mesurer les performances de ces algorithmes sur des images de grande taille, et présentant plusieurs défauts de différents types. En considérant les problèmes posés par l'utilisation des méthodes standard de comparaison, nous exposons une évolution de ces méthodes en s'appuyant sur une extension de la méthode du Receiver Operating Characteristic (ROC). L'approche pixel de cette méthode est modifiée au profit d'une approche basée objet, grâce à l'utilisation de dilations dans le calcul des taux de fausse alarme et de bonne détection. Le processus de dilatation proposé s'adapte automatiquement à la topologie des défauts à détecter et ne nécessite aucun réglage de paramètre. Au travers d'exemples de défauts à détecter, nous montrons que la méthodologie ainsi définie permet dans notre contexte une meilleure évaluation des performances des algorithmes de détection.

Sommaire

3.1	Objectifs de la détection	61
3.2	Méthode d'évaluation des performances de détection : introduction	61
3.3	La tâche d'inspection	62
3.4	L'analyse ROC	63
3.5	La comparaison des masques de détection	65
3.6	Exemples d'applications de la méthode souple	76
3.7	Commentaires et extensions possibles de la méthode	80
3.8	Conclusion	81

3.1 Objectifs de la détection

Dans le chapitre précédent, nous avons vu que l'étape de détection est une étape cruciale. Elle conditionne fortement la qualité de l'analyse effectuée sur l'image. En effet, un défaut détecté à cette étape fera généralement partie des défauts analysés, sauf si aucun pixel défectueux n'est trouvé lors de l'étape de localisation. Ainsi, toute détection entraînera une analyse de la zone détectée, et donc toute fausse alarme provoquera une perte de temps pour les utilisateurs et une confusion dans le rapport final d'analyse de qualité de l'image. Cette étape de détection doit donc comporter le moins possible de fausses alarmes. D'une manière générale, le taux de fausse alarme ciblé dans un contexte d'utilisation industrielle est de l'ordre de une fausse alarme tous les milliards de pixels analysés.

Dans le même temps, il est important de détecter tous les défauts car un défaut non détecté à cette étape ne sera plus considéré par la chaîne d'analyse.

Pour traiter l'étape de détection, il faut donc développer des algorithmes de détection sensibles à la présence de défauts et générant très peu de fausses alarmes. La détection au pixel près des défauts n'est pas requise à cette étape, car elle sera traitée par l'étape suivante de localisation. Néanmoins, les défauts doivent être détectés par l'intermédiaire de suffisamment de pixels défectueux afin que l'étape de capture soit efficace.

Ce contexte s'apparente à une tâche d'inspection semi-automatique des défauts : on veut identifier chaque défaut *via* plusieurs pixels qui seront détectés par la méthode de détection. Un observateur humain va alors venir inspecter chaque amas de pixels détectés en adaptant une fenêtre d'observation (équivalent de ce qui est fait lors de l'étape d'extraction du défaut) afin de vérifier la présence ou non d'un défaut.

Nous recherchons donc une méthode efficace dans ce contexte. Afin d'évaluer correctement les méthodes de détections qui seront exposées dans le chapitre suivant, nous avons besoin d'un outil adapté à notre contexte d'inspection semi-automatique de la qualité.

3.2 Méthode d'évaluation des performances de détection : introduction

Les tâches de contrôle de qualité sont des champs d'application particulièrement prolifiques pour le domaine du traitement des images, et en particulier celui des techniques de détection. Le nombre croissant de techniques conçues pour l'inspection de pièces manufacturées dans l'industrie est une preuve concrète de l'intérêt des industriels et de la communauté scientifique pour ce problème de contrôle qualité.

Les premières applications de ces techniques sont la détection de défauts sur des pièces textiles, en bois, ou d'autres matériaux industriels, par inspection automatique d'images numériques de leurs surfaces. L'acquisition de ces images s'effectue soit par simple imagerie optique (rayonnement visible), soit par rayonnement X comme dans notre cas, ou par d'autres méthodes de contrôle non destructif comme la réflexion d'ondes ultrasonores sur la surface à inspecter.

Différentes méthodes visant à mesurer les performances des algorithmes de détection existent et ont fait l'objet de communications scientifiques. Dans le contexte de détection et de reconnaissance de caractère, Wolf [93] mesure les performances des algorithmes par adéquation géométrique et graphes de performances. Liu [59] propose une méthode simple s'appuyant sur l'analyse du voisinage pour évaluer les performances des algorithmes de détection de bord. Nascimento [72] effectue lui un classement des erreurs de détection rencontrées afin de construire une métrique pour l'évaluation des algorithmes de détection d'objet embarqués dans des applications de vidéo-surveillance.

Des méthodes plus générales s'appuient sur la fusion de métrique de base soit par une logique classique [62], ou grâce à l'utilisation de la logique floue [47].

Même si ces techniques se révèlent très efficaces dans le contexte spécifique pour lequel elles ont été développées, leur fiabilité est remise en cause lorsqu'on considère le contexte de détection d'objets de forme et de nature variables, comme c'est le cas pour les défauts qui nous intéressent.

Dans la suite, nous allons voir comment évaluer les algorithmes de détection et comment utiliser la

méthode de l'analyse ROC, couramment utilisée pour ce type d'application. La méthode ROC classique sera d'abord présentée, et un dérivé de cette méthode, adapté à notre contexte sera ensuite présenté et illustré. Cette nouvelle méthode, appelée méthode ROC souple [69, 70] peut être appliquée à la fois à la comparaison d'algorithmes de détection, et au réglage des paramètres de ces algorithmes en fonction des performances voulues en termes de taux de bonne détection ou de fausse alarme.

3.3 La tâche d'inspection

Les détecteurs numériques pour radiographie produisent des images de grande taille, en général de l'ordre de 3000×3000 pixels. La tâche d'inspection consiste à trouver les défauts présents dans ces images. Les images analysées sont des images plates, sans sujet, et sont donc uniquement constituées de bruit de fond, et de potentiels défauts du détecteur. Dans notre cas, le but des algorithmes de détection est de détecter des zones défectueuses qui seront transmises à l'étape d'extraction ou analysées par un expert humain. Dans ce dernier cas, la tâche de contrôle de qualité se présente donc comme si on voulait attirer l'attention de l'expert sur des zones potentiellement défectueuses pour faciliter l'inspection des images. Les algorithmes de détection sont également destinés à faire partie de la chaîne d'analyse automatique de qualité décrite dans le chapitre précédent, et doivent donc fournir un résultat exploitable pour l'algorithme d'extraction, à savoir une estimation, même partielle, des pixels constituant les défauts. Le processus d'évaluation des algorithmes de détection doit ainsi prendre en compte ce contexte et fournir une mesure adaptée des performances pour des formes variées de défauts.

Les principaux enseignements donnés par ce contexte, et utiles pour construire une mesure de performance adaptée sont :

- la localisation parfaite de tous les pixels défectueux n'est pas requise : seuls quelques pixels détectés dans la zone défectueuse sont nécessaires pour l'identification du défaut. Par contre, au moins un pixel est nécessaire par défaut. On rappelle qu'un défaut est défini comme en ensemble connexe de pixels défectueux,
- les contours de chaque défaut ne peuvent pas être défini de manière univoque. La construction d'une vérité terrain reste subjective, et les pixels dans des zones aux alentours d'un défaut peuvent être vus soit comme des pixels défectueux, soit comme des pixels de fond. De plus, certains défauts ont naturellement des contours flous rendant caduque la construction d'une vérité terrain précise.

Ces observations impliquent une adaptation de la méthodologie ROC pour la rendre efficace. Nous décrirons l'adaptation développée dans la section 3.5.

Dans notre application, on peut différencier deux types de défauts :

- les défauts ponctuels : ce sont les pixels isolés ou des petits amas de pixels présentant des statistiques anormales (trop faible ou trop forte luminosité),
- défauts étendus : ce sont des pixels spatialement corrélés qui ne sont pas individuellement atypiques. En pratique, ces défauts prennent la forme de lignes, de colonnes ou de taches. Elles contiennent plusieurs pixels qui ne sont pas nécessairement des défauts ponctuels.

La figure 3.1 donne des exemples de défauts synthétiques avec leur vérité terrain («Defect Map») déterminée par un expert humain. Ces exemples ont été synthétisés afin d'illustrer des cas extrêmes de défauts pouvant apparaître dans n'importe quel système imageur.

Cette figure permet d'illustrer la difficulté de déterminer une vérité terrain selon le type de défaut considéré. Pour les défauts ponctuels, la vérité terrain peut être définie précisément, alors que pour les défauts étendus, comme la tache floue, la vérité terrain devient très subjective. Quand plusieurs défauts sont regroupés dans une zone limitée, comme sur la partie droite de la figure 3.1(a), la vérité terrain inclut les amas de pixels défectueux ainsi que quelques pixels non-défectueux situés aux alentours. Ainsi, les amas ne sont pas considérés comme plusieurs défauts, mais plutôt comme une seule zone défectueuse. Cette constatation doit être prise en compte lorsqu'on veut mesurer les performances de nos algorithmes de détection.

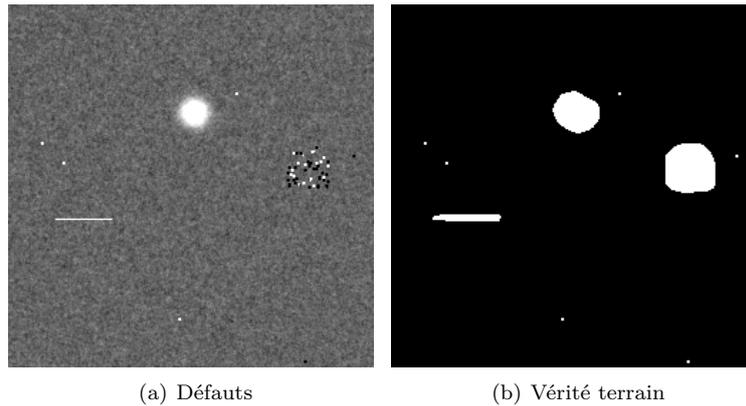


FIGURE 3.1 – Plusieurs catégories de défauts et leur vérité terrain déterminée à la main par un expert humain. On observe que la détermination de cette vérité terrain est une tâche subjective : les frontières des défauts sont définies grossièrement, les amas des pixels défectueux peuvent être regroupés au sein d'un seul et même défaut.

3.4 L'analyse ROC

3.4.1 Définitions et courbes ROC

Initialement développée pour l'évaluation des systèmes de surveillance radar dans les années 50, l'analyse ROC a été pour la première fois formalisée et intégrée à la théorie de la détection dans le milieu des années 1960 [36, 24]. Au fil des années, elle est devenue une technique standard pour l'évaluation des performances en détection [25, 39]. Son utilisation s'est d'abord développée dans le domaine de l'imagerie médicale, en particulier en radiologie [40, 68, 28], et s'est ensuite étendue à divers systèmes de détection.

Pour un problème à un seul défaut (ou une seule zone défectueuse dans notre cas), l'analyse ROC consiste à mesurer la réponse binaire d'un système de détection (défaut présent ou non) à un stimulus (*i.e.* l'image), en calculant le taux de bonne détection (noté tpr pour «true positive rate») et de fausse alarme (noté fpr pour «false positive rate») avec :

$$tpr = \frac{\text{vrais positifs}}{\text{total positifs}} \quad (3.1)$$

$$fpr = \frac{\text{faux positifs}}{\text{total négatifs}} \quad (3.2)$$

La figure 3.2 présente la représentation classique de la matrice de confusion.

Un couple $(fpr; tpr)$ correspond à un point dans le plan ROC. La courbe ROC d'un algorithme de détection est tracée en faisant varier les paramètres de cet algorithme, qui traduisent en général sa sensibilité en détection : à chaque valeur de ce paramètre correspond un point dans le plan ROC. L'analyse ROC est un outil approprié pour mesurer les performances en détection d'un système car elle ne tient pas compte de la répartition des deux classes traitées, ici défaut et fond, et fournit deux mesures intuitives et complémentaires qui peuvent être aisément utilisées pour calibrer le système de détection.

Dans un contexte de détection de défauts, il peut y avoir beaucoup plus d'un défaut dans une image. Les nombres de vrais positifs et de faux positifs peuvent être calculés dans cette image en utilisant la méthodologie dite du «Free-ROC» [45] (FROC). La méthodologie Free-ROC est une extension de la méthodologie ROC à la localisation des défauts, en général dans un contexte multi-cibles (Free-ROC est très utilisée dans le domaine militaire), alors que la méthodologie ROC concerne uniquement la détection de défaut. Dans notre application, la théorie du Free-ROC n'est pas nécessaire et nous considérerons ainsi uniquement les outils basiques fournis par la méthodologie ROC.

Si on considère une image avec plusieurs cibles -les défauts- et un algorithme de détection de défauts

		Système de détection		
		Défaut détecté	Pas de Défaut détecté	
Vérité terrain	Defect	Vrai positif TP	Faux négatif FN	$TP+FN$ = nombre total de positifs
	No Defect	Faux positif FP	Vrai négatif TN	$FP+TN$ = nombre total de négatifs

FIGURE 3.2 – La matrice de confusion définit les vrais positifs et les faux positifs pour une tâche de détection. Ces deux quantités sont celles qui seront représentées dans les courbes ROC

fournissant une classification pixel à pixel en deux classes (défaut ou fond), il y a quatre cas pour chaque pixel $p(i)$ de l'image :

- $p(i)$ est classé comme un défaut, et est réellement un défaut d'après la vérité terrain : c'est un vrai positif
- $p(i)$ est classé comme pixel de fond, et est réellement un pixel de fond d'après la vérité terrain : c'est un vrai négatif
- $p(i)$ est classé comme un défaut, et est en fait un pixel de fond d'après la vérité terrain : c'est un faux positif, aussi appelé fausse alarme
- $p(i)$ est classé comme pixel de fond et est en fait un pixel défectueux : c'est un faux négatif

Le principal avantage de l'analyse ROC est que les deux quantités calculées, tpr et fpr , sont normalisées respectivement par le nombre de pixels de défaut (nombre total de positifs) et le nombre de pixels de fond (nombre total de négatifs). Ainsi, contrairement aux autres mesures classiques comme la précision du système (le pourcentage de pixels correctement classés), le tpr et le fpr ne peuvent pas être biaisés par une faible représentation d'une classe par rapport à l'autre.

Pour mettre en évidence ceci, considérons une tâche de détection sur une image de 1000×1000 pixels contenant un seul pixel défectueux. Un algorithme ne détectant jamais de pixels sera en fait très précis : 99,9999% des pixels seront correctement classés. Dans le même temps, le tpr et le fpr seront tous deux nuls, traduisant de très mauvaises performances de détection. Ainsi, la seule donnée de la précision des algorithmes ne fournit pas assez d'informations pour évaluer les performances en détection, notamment car la précision est très sensible à la sur ou la sous-représentation d'une classe vis à vis de l'autre, ce qui est un cas très fréquent dans le contexte de la détection de défauts. Dans notre exemple, la sous-représentation de la classe défaut rend la mesure de précision peu pertinente pour évaluer les performances de détection.

L'analyse ROC représente en une courbe la sensibilité (quantité équivalente au taux de bonne détection tpr) des systèmes de détection en fonction du taux de fausse alarme fpr , qui sont deux informations pertinentes dans un contexte de contrôle qualité : considérées ensemble, elles permettent de dire combien de fausses alarmes vont être générées par l'algorithme pour une sensibilité donnée. De plus, la courbe ROC permet de rendre compte du comportement dynamique du système vis à vis d'un changement des paramètres de ce système (en général, le seuil de décision). Cette information peut être utilisée afin de comparer deux systèmes de détection : la courbe ROC du meilleur système sera celle qui se situe au dessus des autres courbes dans le plan ROC.

Les courbes ROC de quatre algorithmes de détection sont représentées sur la figure 3.3. L'algorithme de détection idéal est « algo1 » : sa courbe ROC est un échelon (tpr à 100% quel que soit le fpr). Dans le cas où les deux classes sont équilibrées (même nombre de pixels défectueux que de pixels de fond dans la vérité terrain), un algorithme correspondant à une décision aléatoire entre les deux classes donne la

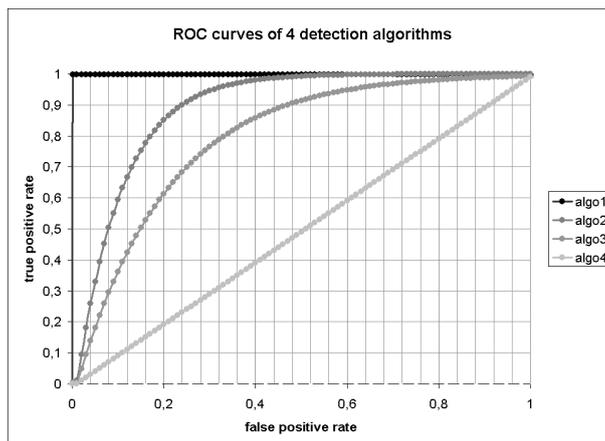


FIGURE 3.3 – Exemples de courbes ROC. Sur cet exemple, quatre courbes ont été tracées. La première, celle de l’algorithme 1, montre que cet algorithme de détection donne des résultats parfaits (100% de taux de bonne détection quel que soit le taux de fausse alarme, et donc quel que que soit les paramètres de l’algorithme). La courbe issue de l’algorithme 4 a été calculée à partir d’un algorithme qui choisit aléatoirement la classe du pixel traité parmi les deux classes possibles (défaut ou fond). Les deux autres courbes sont celles de deux algorithmes de détection quelconques. L’algorithme 2 a de meilleures performances que l’algorithme 3 car sa courbe ROC se situe au dessus de celle de l’algorithme 3

courbe ROC «algo4»(diagonale ascendante dans le plan ROC). Entre l’algorithme parfait et l’algorithme prenant une décision au hasard, l’algorithme 2 donne de meilleurs résultats que l’algorithme 3.

L’information donnée par les courbes ROC peut être synthétisé en une quantité, traduisant la performance de l’algorithme de détection associé : l’AUC pour «Area Under Curve »qui est simplement l’aire située sous la courbe ROC [40, 30, 16].

3.5 La comparaison des masques de détection

Dans la section 3.4, nous avons vu que l’analyse ROC peut être utilisée afin d’évaluer les performances d’un algorithme de détection. Cependant, un problème majeur persiste : comment estimer les deux grandeurs tpr et fpr ? En d’autres termes : quels pixels doivent être considérés comme vrais positifs et faux positifs ?

Pour chaque seuil de décision, tpr et fpr sont calculés en comparant un masque binaire de détection (avec les pixels détectés à 1 et les pixels non détectés à 0) à un masque constitué avec la vérité terrain. Dans toute la suite, le masque binaire de détection est appelé le masque test et noté $M_{i,j}$. Il est le résultat d’une détection pixel à pixel effectuée par l’algorithme de détection. Le masque de vérité terrain est noté $T_{i,j}$. En pratique, la manière la plus simple de comparer ces deux masques est d’effectuer une comparaison pixel à pixel, en utilisant la définition classique des vrais positifs et faux positifs donnée dans la section 3.4.

Comme exposé dans la section 3.3, l’expert humain n’a pas besoin que l’algorithme de détection lui fournisse une localisation au pixel près des défauts. En s’appuyant sur cette hypothèse, différentes façon de calculer tpr et fpr ont été proposées : Theiler [48] suggère de transformer le masque test afin de considérer chaque défaut du masque de vérité terrain comme 100% détecté si au moins un pixel est détecté dans ce défaut. D’une manière analogue, Harvey [41] propose de dilater le masque test à l’aide d’un élément structurant fixé *a priori*, afin d’inclure dans le calcul du tpr les pixels détectés se situant aux alentours de chaque défaut.

La définition de méthodes alternatives de calcul des taux de fausse alarme et de bonne détection prend sens dans notre contexte d’inspection semi-automatique auquel nous devons nous adapter afin de fournir une évaluation plus pertinente des performances de détection. Considérons d’abord les vrais positifs :

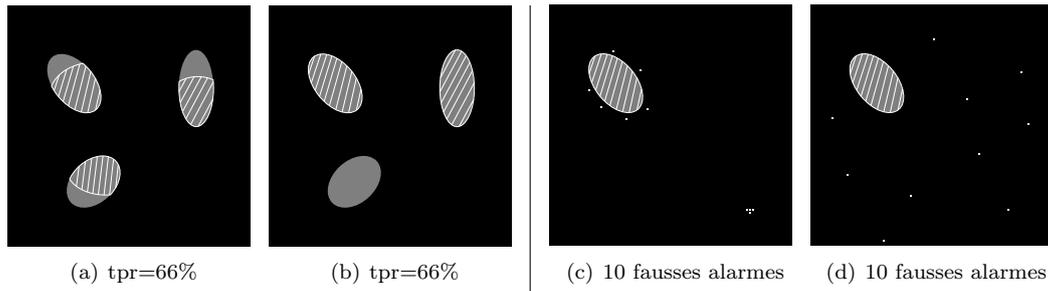


FIGURE 3.4 – sens du tpr et du fpr : sur chaque image, les cibles à détecter sont représentées en gris, et les pixels détectés par le système sont en blanc. Les deux premières images ont le même taux de bonne détection et les deux dernières le même nombre de fausses alarmes. Néanmoins, dans notre contexte, la configuration des figures 3.4(a) et 3.4(c) sont considérées comme meilleures en termes de résultat de détection car toutes les cibles ont été détectées et il n’y a qu’une seule réelle fausse alarme. En effet, sur la figure 3.4(c), les pixels détectés autour de la cible ne doivent pas être considérés comme des fausses alarmes car ils sont à proximité immédiate de la cible.

dans une situation à plusieurs défauts (*i.e.* il y a plus d’une cible dans le masque de vérité terrain), le tpr doit être de 100% si les pixels détectés conduisent à une localisation de toutes les cibles, même si tous les pixels défectueux ne sont pas détectés. Dans l’exemple de la figure 3.4(a), seulement 66% des pixels défectueux sont détectés (zone hachurée) alors que la détection permet la localisation de toutes les cibles. Le taux de bonne détection de la figure 3.4(b) est également de 66% alors que la détection effectuée est clairement moins bonne que la précédente car la cible située sur la partie inférieure de l’image n’a pas du tout été détectée.

Considérons maintenant le taux de fausse alarme : nous devons faire une distinction entre les fausses alarmes isolées, les fausses alarmes proches d’une cible, communément appelée «near-hit», et les amas de fausses alarmes. Pour l’expert humain, le nombre total de fausses alarmes est le nombre de fenêtres de visualisation, appelée AOI («Area Of Interest»), que le système affichera alors qu’aucun défaut n’apparaît sur ces fenêtres. Dans le cas de la chaîne d’analyse des images, une fausse alarme conduira à une extraction d’une AOI ne présentant aucun défaut visuellement apparent. Sur l’exemple de la figure 3.4, les deux images de la figure 3.4(c) et 3.4(d) contiennent dix fausses alarmes chacune. Cependant, dans le cas de la figure 3.4(c), l’expert humain considérera que l’algorithme de détection n’a produit qu’une seule fausse alarme : les quelques pixels détectés à proximité de la cible ne peuvent pas être considérés comme des fausses alarmes car l’AOI qui leur sera associée contiendra au moins une partie de la cible, et contiendra donc des pixels visuellement défectueux. Les autres pixels détectés forment un amas qui peut être visualisé à l’aide d’une seule AOI. Cette AOI détectée constitue alors la seule fausse alarme du système. Sur la figure 3.4(d), il n’y a que des fausses alarmes isolées : l’expert se verra présenté dix AOI ne contenant pas de pixels défectueux. Le système génère donc dix fausses alarmes.

En prenant en compte ces observations, nous proposons une méthode permettant de comparer simplement les masques binaires en considérant la spécificité des pixels situés à proximité des cibles. Ces pixels sont considérés comme des fausses alarmes par toutes les méthodes précédentes. Le but principal de notre méthode est d’étendre la définition au pixel près des deux quantités utiles dans l’analyse ROC à une définition basée objet. Cette méthode pour comparer les masques s’intègre parfaitement à la méthodologie ROC classique et ne requiert pas de paramétrage supplémentaire. De plus, notre méthode s’adapte automatiquement à tous les défauts, quelles que soient leur taille et leur forme.

Dans les paragraphes suivants, trois méthodes de comparaison de masques sont décrites et commentées : la comparaison simple basée pixel, la méthode définie par Theiler, et la méthode que nous avons développée et utilisée dans le cadre de cette thèse. La méthode proposée s’appuie sur les spécificités de la tâche d’inspection semi-automatique des images. La méthode de Harvey introduite plus haut requiert l’introduction d’une forte connaissance *a priori* sur la taille des cibles à détecter. Comme nous ne pouvons pas fournir dans notre cas de telles connaissances, cette méthode ne sera pas étudiée.

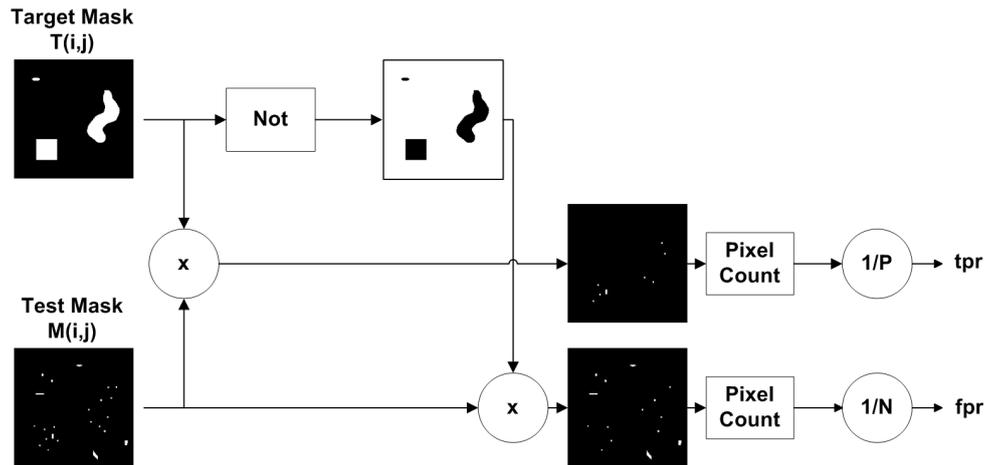


FIGURE 3.5 – Comparaison des masques pixel à pixel : le calcul de tpr et de fpr est effectué par comparaison directe entre masque de détection et masque de vérité terrain. La fonction «Not» est le complément binaire et la fonction «Pixel Count» retourne le nombre de pixels blancs (valeur à 1) dans l'image d'entrée.

3.5.1 Comparaison des masques pixel à pixel

La méthode la plus intuitive afin de comparer les masques de détection et de vérité terrain est d'effectuer un calcul pixel à pixel. Si on considère le masque de détection $T_{i,j}$, contenant P pixels défectueux (pixels ayant la valeur 1) et N pixels de fond (pixels ayant la valeur 0), la comparaison pixel à pixel peut être décrite comme sur la figure 3.5. La fonction «Pixel Count» retourne le nombre de pixels prenant la valeur 1 dans l'image passée en entrée, et la fonction «Not» et l'opérateur de complément : elle retourne 0 quand le pixel passé en entrée est à 1, et 1 quand le pixel passé en entrée est à 0. Le tpr et le fpr sont alors donnés par :

$$tpr = \frac{1}{P} \sum_{i,j} M_{i,j} \cdot T_{i,j} \quad (3.3)$$

$$fpr = \frac{1}{N} \sum_{i,j} M_{i,j} \cdot (1 - T_{i,j}) \quad (3.4)$$

La comparaison pixel à pixel des masques a été testée sur les trois défauts synthétiques des figures 3.6(a), 3.6(b) et 3.6(c). A chaque fois, la vérité terrain est tracée en gris sur la seconde ligne de l'illustration, avec l'ensemble des pixels détectés apparaissant en blanc, qu'ils soient des fausses alarmes ou des bonnes détections. Les trois défauts pris en exemple sont :

- des défauts ponctuels : six pixels isolés défectueux
- une tache floue : une tache brillante ayant des bords flous
- amas défectueux : un amas de pixels défectueux formant une zone identifiée visuellement comme un seul et même défaut

Par souci de clarté, sur la figure 3.6(d), les pixels défectueux définis par la vérité terrain sont pointés par des flèches.

La table 3.1 donne, pour les trois défauts, les taux de bonne détection et de fausses alarmes calculés avec la méthode de comparaison pixel à pixel.

Dans le cas des pixels isolés, la simple comparaison des masques pixel à pixel donne des résultats satisfaisants : pour ce type de défauts, la localisation exacte est requise et une fausse alarme est générée même lors de la détection d'un pixel proche du défaut. En pratique, le système doit être très précis pour ce type de défaut afin d'attirer l'attention de l'expert sur l'endroit précis du défaut car il s'agit d'un défaut de très petite taille, qui sera donc observé à petite échelle. Dans ces cas, la comparaison pixel à pixel est efficace : elle donne un tpr de $\frac{1}{2}$ et un fpr qui représente bien le nombre de fois où l'expert devra

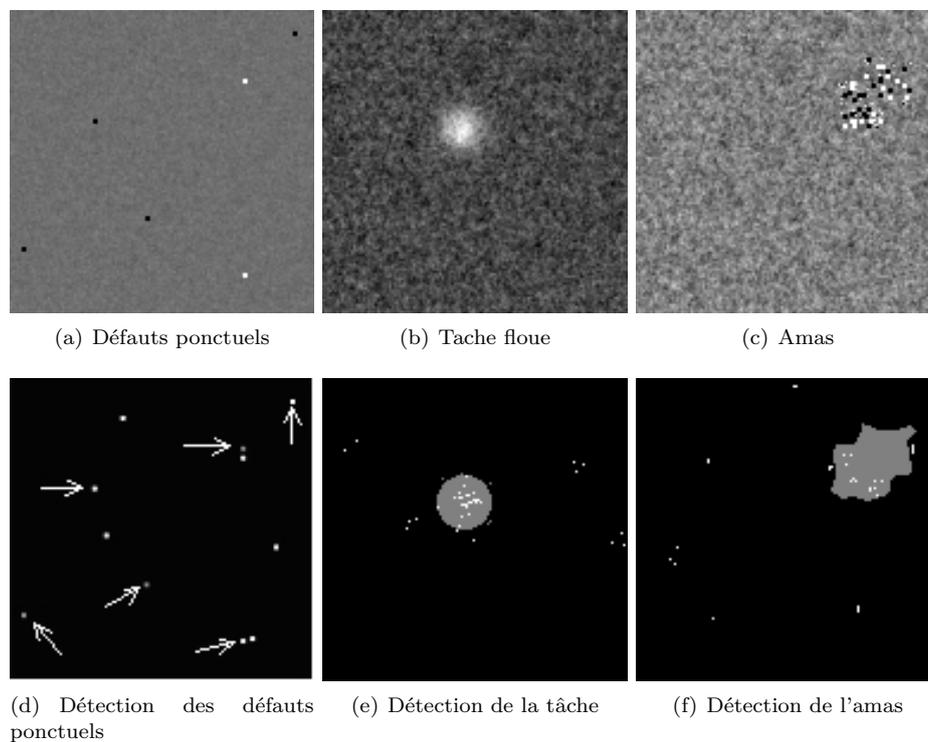


FIGURE 3.6 – Trois types de défauts et leurs masques de détection et de vérité terrain. Les pixels en gris sont les pixels du masque de vérité terrain, définie visuellement. Les pixels en blanc sont les pixels détectés comme défectueux par l’algorithme de détection.

Défaut	fpr (%)	tpr (%)
Ponctuel	0.12	50
Tache	0.03	3.8
Amas	0.12	1.5

TABLE 3.1 – Taux de bonne détection tpr et taux de fausse alarme fpr calculés avec la méthode de comparaison pixel à pixel des masques. Le seuil considéré pour calculer ces quantités est celui donnant les détections des figures 3.6(d), 3.6(e) et 3.6(f)

vérifier une zone détectée comme défectueuse alors qu'elle ne contient aucun défaut visible.

Le cas de la tache floue est plus ambigu. Dans l'exemple présenté, seulement quelques pixels de la tache sont détectés. Pour ce défaut, la vérité terrain est une zone circulaire qui inclut la tache claire. Dans l'exemple présenté, l'algorithme de détection est presque parfait : le nombre de pixels détectés et l'information sur leur position, au centre de la tache, suffisent pour que l'expert localise correctement le défaut. Néanmoins, en raison de la définition de la vérité terrain sous la forme d'un ensemble circulaire compact, les valeurs du taux de bonne détection et de fausse alarme sont loin de celles attendues pour une si bonne détection. De plus, on peut considérer le pixel détecté dans la partie inférieure gauche du défaut comme un pixel trop proche de la cible pour occasionner pour l'utilisateur une fausse alarme. Cet exemple montre que l'évaluation des performances de détection se heurte à la subjectivité de la définition de la vérité terrain. Cet élément n'est pas pris en compte dans la méthode de comparaison des masques pixel à pixel.

Nous pouvons faire les mêmes remarques pour les amas de défauts. Une fois encore, certains pixels (de part et d'autres de l'amas de pixels défectueux) sont comptabilisés comme étant des fausses alarmes alors qu'ils sont trop proche des frontières du défaut pour pouvoir être réellement considérés comme telles. Dans le même temps, un nombre important de pixels est détecté au sein de l'amas (principalement les défauts ponctuels constitutifs de l'amas défectueux), le taux de bonne détection atteignant à peine les 2%. Pour l'expert humain, dans ce cas, le défaut peut être considéré comme bien détecté, ce qui est en contradiction avec le faible taux de bonne détection donné par la méthode de comparaison pixel à pixel.

Ces exemples montrent que cette méthode n'est pas adaptée pour rendre compte des réelles performances des systèmes de détection dans notre cas. La raison principale est que la méthode de comparaison pixel à pixel n'intègre pas les spécificités détaillées dans la section 3.3.

3.5.2 Comparaison des masques par la méthode de Theiler

La méthode classique de comparaison pixel à pixel des masques conduisant à une mauvaise évaluation des performances, il faut développer une technique qui, en quelque sorte, mime le comportement de l'expert humain vis à vis de la tâche de détection. Dans cette idée, Theiler a proposé une métrique pour donner au masque de détection une interprétation plus proche de celle de l'expert humain. Cette technique utilise un procédé de remplissage : tous les pixels composant une cible sont considérés comme détectés si au moins un de ces pixels est détecté par l'algorithme de détection (voir la figure 3.7). Le taux de bonne détection et le taux de fausse alarme sont alors donnés par :

$$tpr = \frac{1}{P} \sum_{i,j} Fill(M_{i,j}) \cdot T_{i,j} \quad (3.5)$$

$$fpr = \frac{1}{N} \sum_{i,j} M_{i,j} \cdot (1 - T_{i,j}) \quad (3.6)$$

où *Fill* est l'opérateur de remplissage (marque tous les pixels de la cible à 1 si au moins un pixel appartenant à cette cible est à 1).

En suivant cette approche, les taux de bonne détection et de fausse alarme sont calculés sur l'exemple de la figure 3.6. Les résultats obtenus sont donnés dans la table 3.2.

Pour les défauts ponctuels, la méthode de Theiler fournit des résultats satisfaisants, similaires à ceux obtenus avec la comparaison des masques pixel à pixel.

Pour les deux autres exemples, le taux de fausse alarme d'après la méthode de Theiler est inchangé, mais le taux de bonne détection est maintenant, pour les deux types de défaut, de 100%. En effet, dans chaque cas, au moins un pixel a été détecté au sein de la cible. Cette stratégie du «tout ou rien» est bien adaptée à certains cas : pour la tache défectueuse, les pixels détectés sont situés au centre du défaut et sont distribués dans une zone qui est comparable à celle définie par la vérité terrain. Ainsi, du point de vue de l'expert, ces détections seront regroupées en une seule détection lui permettant de trouver le

Défaut	fpr (%)	tpr (%)
Ponctuel	0.12	50
Tâche	0.03	100
Amas	0.12	100

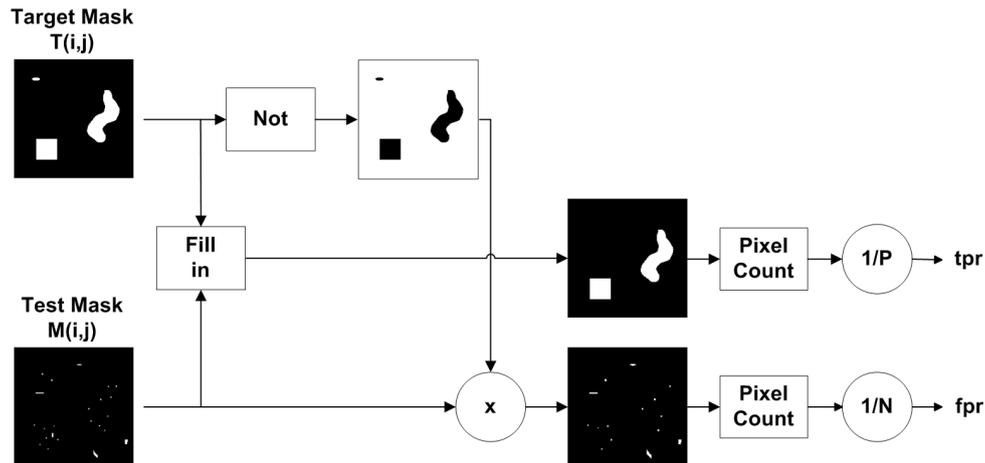
TABLE 3.2 – Taux de bonne détection tpr et taux de fausse alarme fpr calculés avec la méthode de Theiler

FIGURE 3.7 – Méthode de Theiler pour la comparaison des masques. La fonction de remplissage (nommée «Fill-in» dans le schéma) est ajoutée à la précédente méthode de comparaison pixel à pixel : si au moins un des pixels est détecté dans la cible, tous les pixels de cette cible seront décomptés comme vrais positifs.

défaut. Le taux de bonne détection de 100% rend bien compte de cette bonne performance de détection.

Dans le cas du défaut constitué d'amas de pixels défectueux, les limites de la stratégie de remplissage peuvent être mises en évidence. En effet, dans ce cas, il n'y a que la partie inférieure de l'amas qui est vraiment détectée. Dans son observation, en considérant la répartition des détections et la taille du défaut, il y a des risques que l'expert ne puisse pas identifier la partie supérieure du défaut. Dans la méthode de Theiler, de telles considérations n'ont aucun impact dans le calcul du taux de bonne détection, qui atteint donc les 100% pour le défaut traité. Cette défaillance se produit dans d'autres configurations, principalement celles mettant en jeu des défauts étendus.

Le calcul du taux de fausse alarme par la méthode de Theiler ne diffère pas de la méthode pixel à pixel, et comporte donc les mêmes inconvénients, comme décrit dans la section précédente.

En conclusion, la méthode de Theiler étend l'approche pixel à pixel à une approche plutôt orientée objet en enrichissant la méthode de base avec un procédé de remplissage opérant sur le masque de détection. Dans le cas d'une détection à une seule cible ou les cas de cibles de très petite taille par rapport à la dimension de l'image, cette méthode permet une évaluation acceptable des performances. Néanmoins, dans le cas plus général, comme la méthode de Theiler ne prend pas en compte la distribution spatiale des pixels détectés, qui est un élément important à considérer avant de déclarer une cible comme 100% détectée, on obtient des résultats non satisfaisants pour le calcul du taux de bonne détection. En particulier, ce taux est en général surestimé. De plus, le calcul du taux de fausse alarme reste inchangé et peu pertinent.

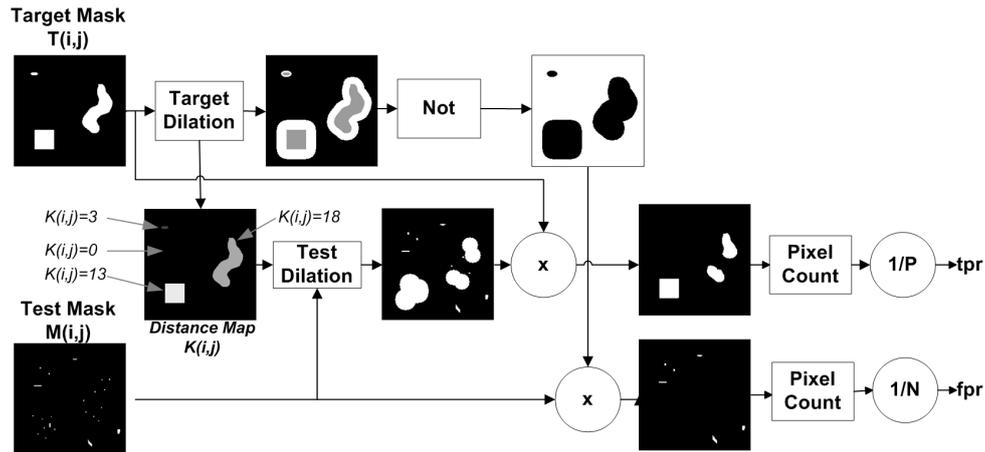


FIGURE 3.8 – Méthode de comparaison souple des masques. Les masques de détection et de vérité terrain sont dilatés d'un facteur calculé automatiquement en fonction de la forme et la taille de chaque cible de la vérité terrain. Ce procédé vise à mimer la perception du taux de bonne détection et de fausse alarme qu'aurait un expert humain dans un contexte d'inspection de défaut sur une image de grande taille.

3.5.3 Méthode proposée : comparaison souple

Dans cette section, nous présentons une nouvelle méthode qui permet d'apporter une réponse pragmatique au problème d'évaluation de performances de détection en dépassant les limitations des méthodes classiques décrites précédemment. Le principe de notre nouvelle méthode est illustré par la figure 3.8.

Cette méthode nécessite la définition de deux nouvelles fonctions : la dilatation des cibles («Target dilation») et la dilatation des détections («Test dilation»). La dilatation des cibles est appliquée sur le masque de vérité terrain, et vise à étendre les défauts de manière à recouvrir les pixels voisins des défauts pour tenir compte de la subjectivité dans la définition de la vérité terrain et des éventuels contours flous de certains défauts. La dilatation des détections est appliquée sur le masque de détection, qui contient les pixels détectés par l'algorithme de détection. L'opération ainsi réalisée mime la manière dont l'expert humain interprète et utilise les résultats de détection : à l'aide d'une fenêtre de visualisation, il va vérifier chacune des zones détectées, englobant également les pixels situés au voisinage de cette zone.

Ces dilatations nécessitent le calcul d'une carte de distance («Distance Map») comme détaillé plus bas. Cette carte de distance ne doit être calculée qu'une seule fois pour une vérité terrain donnée. Ainsi, le calcul de plusieurs points ROC (variation des paramètres de l'algorithme de détection, et donc du masque de détection) ne nécessite qu'une seule dilatation de la vérité terrain et un seul calcul de la carte de distance associée à la vérité terrain. La méthode proposée a un coût en calculs additionnels limité aux dilatations des masques de détection.

3.5.3.1 Calcul de la carte de distance

La carte de distance est calculée sur l'image de vérité terrain (target mask). L'opération fournit une image permettant d'assigner à chaque cible un facteur de dilatation. Le calcul de cette carte de distance est un procédé en deux étapes : la première consiste à appliquer une transformée en distance euclidienne [46] à chaque cible. On détermine ensuite dans une deuxième étape la distance de Hausdorff directe [78] sur chaque cible afin de trouver le facteur de dilatation. Ces deux étapes sont illustrées par la figure 3.9.

- *transformée en distance euclidienne* : la distance euclidienne entre les pixels (i, j) et (m, n) est définie par :

$$d_e((i, j), (m, n)) = \sqrt{(i - m)^2 + (j - n)^2}$$

La transformée en distance euclidienne sur le masque de vérité terrain calcule la distance euclidienne

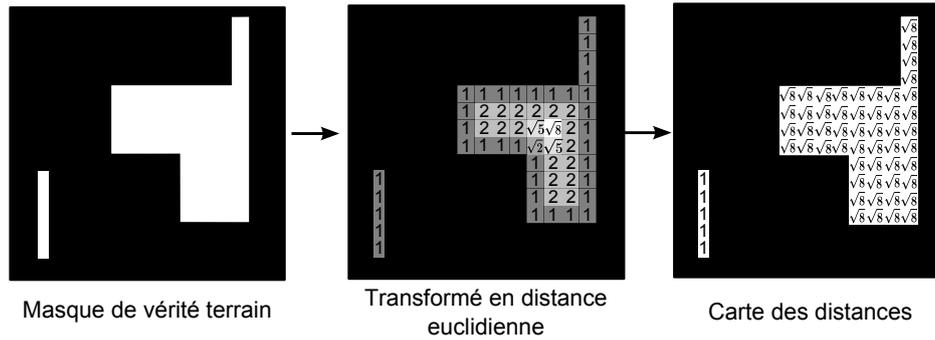


FIGURE 3.9 – Calcul de la carte de distance. La transformée en distance euclidienne du masque de détection est calculée. Elle représente la distance euclidienne de chaque pixel de la cible (en blanc) par rapport au fond (en noir). Pour chaque cible, le *maximum* de la transformée en distance euclidienne est calculé : il s’agit de la distance de Hausdorff directe de la cible, dont la valeur est ensuite stockée dans la carte de distance.

entre les pixels de la cible et les pixels de fond. L’image résultante $D(i, j)$ est déterminée par :

$$D(i, j) = \min_{m, n} \{d_e(T(i, j), T(m, n)) \mid T(m, n) = 0\}$$

Les valeurs de $D(i, j)$ représentent la distance minimale entre chaque pixel de la vérité terrain et le fond. Sur la figure 3.9, un exemple de transformée en distance euclidienne sur une image à deux cibles est présenté. La distance euclidienne d’une cible au fond de l’image dépend de la topologie de la cible.

- *distance de Hausdorff directe* : l’étape finale permet d’étendre les valeurs calculées pour chaque pixel à la cible entière. Pour chaque cible, le *maximum* de la transformée en distance des pixels la composant est calculé. Soit $l(i, j)$ une version labellisée du masque de vérité terrain ($L(i, j) = k$ si le pixel (i, j) appartient à la cible k , $k = 1 \dots n_{targets}$), et $L(i, j) = 0$ si (i, j) appartient au fond. La valeur stockée dans la carte de distance $K(i, j)$ est :

$$K(i, j) = \max_{m, n} \{D(m, n) \mid L(m, n) = L(i, j)\}$$

3.5.3.2 Dilatation de la vérité terrain

La dilatation de la vérité terrain consiste à dilater chaque cible du masque de vérité terrain $T(i, j)$, calculé par la procédure décrite ci-dessus (voir figure 3.10). Pour chaque cible, on utilise un élément structurant circulaire (approximation de la distance euclidienne) de rayon $K_{i, j}$. Cette dilatation de la vérité terrain permet de créer une zone autour de chaque cible dans laquelle les pixels détectés ne seront pas décomptés comme fausses alarmes. L’aspect adaptatif par rapport à chaque cible de cette méthode est sa principale originalité et son point fort : elle ne requiert aucun réglage de paramètre, processus qui est assez régulièrement arbitraire, contrairement à d’autres méthodes existantes (voir notamment [41]). De plus la dilatation utilisée par la méthode souple permet de l’utiliser dans le cas de cibles multiples puisque le facteur de dilatation est calculé automatiquement pour chaque cible de $T(i, j)$.

3.5.3.3 Dilatation des détections

La dilatation des détections est réalisée sur le masque binaire de détection $M(i, j)$. Chaque pixel de ce masque est dilaté à l’aide d’un élément structurant de rayon donné par la valeur $K(i, j)$ de la carte de distance, comme illustré par la figure 3.11. Cette dilatation des détections traduit le fait qu’une détection exhaustive de tous les pixels de la cible n’est pas requise dans notre situation, comme exposé dans la section 3.3. Ainsi, un seul pixel détecté dans la cible peut avoir un impact fort sur le taux de bonne

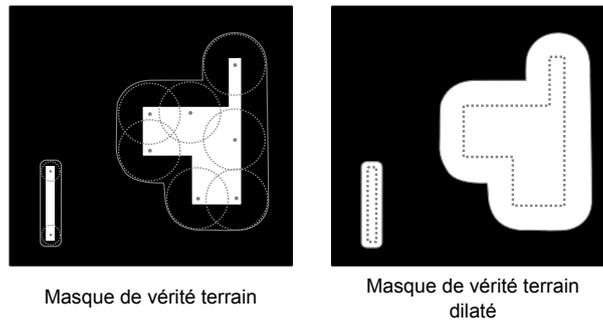


FIGURE 3.10 – Dilatation du masque de vérité terrain : les deux cibles (en blanc) sont dilatées par un facteur stocké dans la carte de distance (image de gauche). L'élément structurant utilisé pour ces dilatations est affiché en gris. Le résultat de la dilatation est représenté sur l'image de droite. La dilatation de la vérité terrain s'adapte automatiquement à la forme et à la taille de chaque cible, grâce à des facteurs de dilatation précédemment calculés et stockés dans la carte de distance.

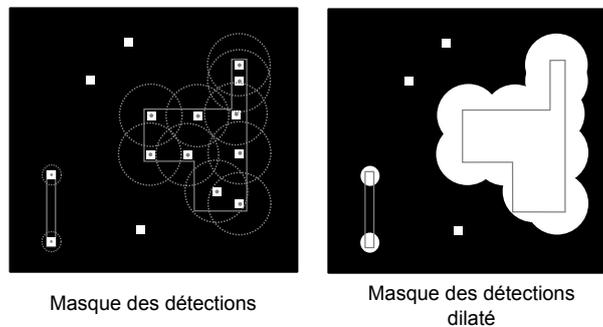


FIGURE 3.11 – Dilatation des détections : chaque pixel du masque de détection (pixels blancs sur l'image de gauche) est dilaté à l'aide d'un élément structurant circulaire (en pointillés gris) de taille correspondant au facteur de dilatation stocké dans la carte de distance. Le masque de vérité terrain est représenté en surimpression par des lignes grises. Les pixels détectés qui n'atteignent pas une cible ne sont pas dilatés.

détection pour notre méthode, comme pour celle de Theiler, contrairement à ce qu'on observe avec la méthode pixel à pixel. Cependant, contrairement à la méthode de Theiler, la dilatation des détections ne permet pas d'obtenir un taux de bonne détection de 100% dans toutes les configurations. La dilatation mime l'effet de la fenêtre d'observation utilisée par l'expert. Dans notre méthode, on considère que cette fenêtre a une taille similaire à la taille de l'objet qu'elle est censée observer. C'est pourquoi le facteur de dilatation nécessite un calcul morphologique sur chaque cible. Ce choix peut être raisonnablement accepté car pour l'inspection d'un pixel détecté, l'expert adaptera son échelle d'observation au contexte entourant cette détection dans l'image, *i.e.* à la cible.

Les taux de bonne détection et de fausse alarme tpr and fpr sont alors calculés comme suit :

$$tpr = \frac{1}{P} \sum_{i,j} TestDil(M_{i,j}, K_{i,j}) \cdot T_{i,j} \quad (3.7)$$

$$fpr = \frac{1}{N} \sum_{i,j} M_{i,j} \cdot (1 - TargDil(T_{i,j})) \quad (3.8)$$

où $TestDil$ est l'opération de dilatation des détections, et $TargDil$ l'opération de dilatation de la vérité terrain.

Les résultats de la méthode de comparaison souple des masques sur les trois défauts considérés sont

représentés sur la figure 3.12. La première ligne de cette figure représente la dilatation de la vérité terrain, nécessaire pour le calcul du taux de fausse alarme. La cible d'origine est en gris clair, et la cible après dilatation est en gris foncé. Les pixels détectés sont représentés en blanc. La dilatation de la vérité terrain est conçue pour étendre la taille des cibles tout en préservant leur forme globale. Ici, la dilatation de la vérité terrain conduit au doublement de la distance de Hausdorff de chaque cible. A la vue de notre méthode souple, deux interrogations peuvent persister :

1. pour quelle raison le facteur de dilatation doit-il dépendre de la taille de la cible ?
2. Comment justifier le choix de la procédure utilisée pour le calcul du facteur de dilatation ?

Les réponses pouvant être données à ces deux interrogations sont que :

1. un défaut de grande taille sera observé à grande échelle. Ainsi, la vérité terrain associée à de tels défauts est naturellement moins précise que celle associée à des défauts beaucoup plus petits (qui seront donc observés à une échelle zoomée). Ainsi, la marge d'erreur pour comptabiliser les fausses alarmes autour des défauts doit être proportionnelle à la taille de ces défauts.
2. Nous pouvons faire l'hypothèse qu'un défaut est observé avec une fenêtre d'observation de dimensions approximativement deux fois plus grandes que les dimensions du défaut. La dilatation de la vérité terrain réalisée mime ainsi cette fenêtre d'observation en autorisant la présence de pixels détectés dans une zone de cette taille autour du défaut. Cette dilatation est une procédure adaptée automatiquement aux défauts présents et ne nécessitant aucun réglage de la part de l'utilisateur.

Le résultat de la dilatation de la vérité terrain est que les pixels situés près des défauts amas et tache ne seront plus considérés comme étant de fausses alarmes car ils appartiennent maintenant, après dilatation, à la vérité terrain. La seconde ligne de la figure 3.12 montre le résultat de la dilatation des détectés (en blanc), la vérité terrain étant représentée en surimpression grise. Chaque pixel détecté appartenant à la cible initiale (en gris clair) est dilaté avant calcul du taux de bonne détection. Cette dilatation suit le raisonnement de l'expert humain : lors de l'observation d'une détection, la fenêtre d'observation englobe la détection et les pixels l'entourant. Ainsi, il est cohérent de considérer les zones voisines des zones de défaut lors du calcul du taux de bonne détection. Une des conséquences de cette dilatation des détectés est que seuls quelques points bien distribués sur le défaut sont suffisants pour qu'il soit détecté à 100%. En particulier, la détection des pixels composant le squelette [9, 34] de la cible sont suffisants pour obtenir 100% de détection pour cette cible.

3.5.3.4 Calcul de point dans le plan ROC par la méthode souple

Les valeurs des taux de bonne détection et de fausse alarme pour les exemples choisis sont indiqués dans la table 3.3.

Pour les défauts ponctuels, il n'y a aucun changement entre la méthode souple et les autres méthodes. Pour le défaut de type tache, le taux de bonne détection atteint 100%, ce qui est cohérent avec l'interprétation de la détection du point de vue de l'expert. En effet, la répartition des détectés dans ce défaut de taille modeste permet son identification aisée : l'expert humain ajustera sa fenêtre d'observation sur la zone des pixels détectés et visualisera ainsi le défaut dans sa totalité. Le taux de fausse alarme a légèrement baissé grâce aux pixels détectés en bas à gauche du défaut, qui ne sont plus considérés comme des fausses alarmes par la méthode souple.

Pour le défaut de type amas, le taux de fausse alarme a baissé pour les mêmes raisons, alors que le taux de bonne détection atteint maintenant 80% : en effet, la partie haute du défaut n'est pas considérée comme correctement détectée par notre méthode. Il s'agit d'une interprétation cohérente car l'amas est de grande taille et constitué de deux zones défectueuses : seulement celle du bas est correctement détectée par l'algorithme. Dans cette configuration, la méthode de Theiler donne un taux de bonne détection de 100%, ne prenant pas en considération que la partie haute du défaut a été manifestement mal détectée.

Dans les cas présentés, la méthode souple donne des *tpr* et *fpr* cohérents avec l'interprétation de l'expert en termes de performances de détection. Cette méthode effectue une dilatation morphologique des cibles et des détectés, et s'adapte automatiquement aux problèmes multi-cibles car les facteurs de

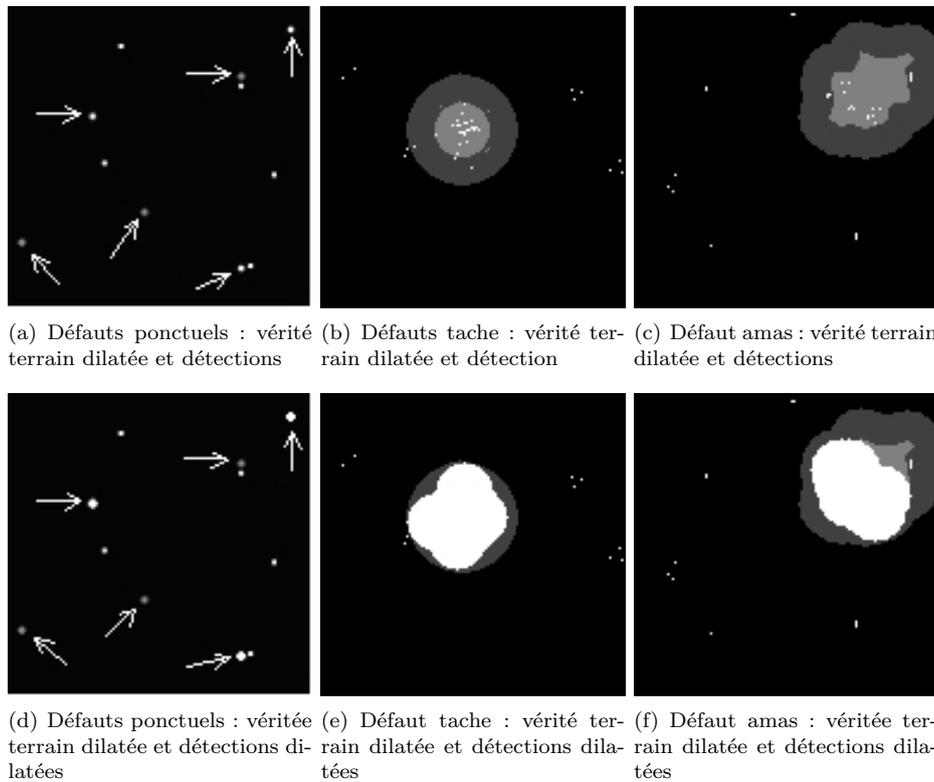


FIGURE 3.12 – Première ligne : masque de vérité terrain dilatée avec en surimpression le masque des détections (en gris clair le masque de vérité terrain original, en gris foncé le masque de vérité terrain dilaté, et en blanc les détections). Seconde ligne : masque de vérité terrain (en gris clair), masque de vérité terrain dilaté (en gris foncé) avec en surimpression le masque des détections dilaté (en blanc). La dilatation du masque de détection entraîne une augmentation du taux de bonne détection dans le cas des pixels détectés au sein d'une cible compacte. Pour les pixels isolés, une grande précision sur leur localisation étant requise, la dilatation du masque de détection n'a aucun effet sur le taux de bonne détection.

Défaut	fpr (%)	tpr (%)
Ponctuel	0.12	50
Tache	0.03	100
Amas	0.07	80

TABLE 3.3 – tpr et fpr pour la méthode de comparaison souple des masques

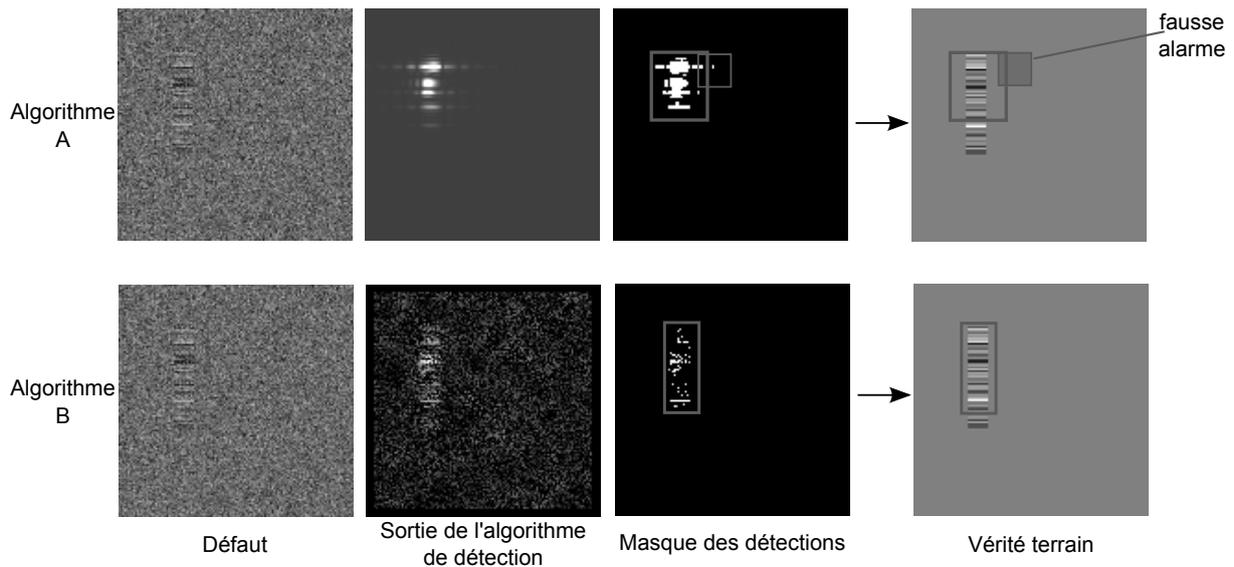


FIGURE 3.13 – Comparaison des performances de détection de deux algorithmes. La première colonne montre l’image d’origine contenant le défaut à détecter, la seconde colonne est l’image obtenue en sortie de l’algorithme de détection, et la troisième colonne est une version seuillée (choix d’un seuil de décision pour la détection) de l’image en sortie de l’algorithme. L’expert humain va inspecter à l’aide des AOIs représentées par des fenêtres grises. La dernière colonne est la superposition de la vérité terrain avec les AOI vérifiées par l’expert. L’algorithme A génère une fausse alarme.

dilatations sont calculés pour chaque cible individuellement. De plus, la méthode souple a le gros avantage de ne nécessiter aucun réglage de paramètres.

3.6 Exemples d’applications de la méthode souple

3.6.1 Mesure de performances d’algorithmes de détection par courbes ROC

Dans cette section, nous allons présenter en détails deux applications possibles de la méthode ROC souple dans le domaine de l’évaluation des performances de détection. Notre méthode est d’abord appliquée à l’évaluation de la performance de deux algorithmes sur une seule catégorie de défaut (section 3.6.1.1), puis pour comparer les résultats de détection d’un algorithme sur deux types de défauts (section 3.6.1.2). Notre méthode sera ensuite utilisée pour la calibration d’un algorithme de détection : nous verrons que notre méthode peut être efficace pour calibrer correctement un algorithme en accord avec le contexte d’inspection de la qualité (section 3.6.2).

3.6.1.1 Évaluation des performances de deux algorithmes pour un type de défaut

L’analyse ROC est un outil simple et efficace : pour comparer les performances de plusieurs algorithmes de détection, on peut comparer leur AUC respective. Dans cet exemple, on veut comparer deux algorithmes, appelés «Algorithme A» et «Algorithme B», à l’aide d’une analyse ROC. Les sorties de ces deux algorithmes sont représentées sur la seconde colonne de la figure 3.13. Les pixels détectés comme défectueux prennent de fortes valeurs (pixels clairs). Les courbes ROC correspondantes sont représentées sur la figure 3.14, les AUC calculées sur ces courbes étant reportées dans la table 3.4. Un seuillage des sorties des algorithmes est affiché dans la troisième colonne de la figure 3.13.

Les courbes ROC et les AUC calculées à l’aide de la méthode pixel à pixel montrent clairement que l’algorithme A réalise une meilleure détection que l’algorithme B. Notre méthode souple tire la conclusion inverse, la méthode de Theiler ne donnant dans la même temps aucune information discriminante en

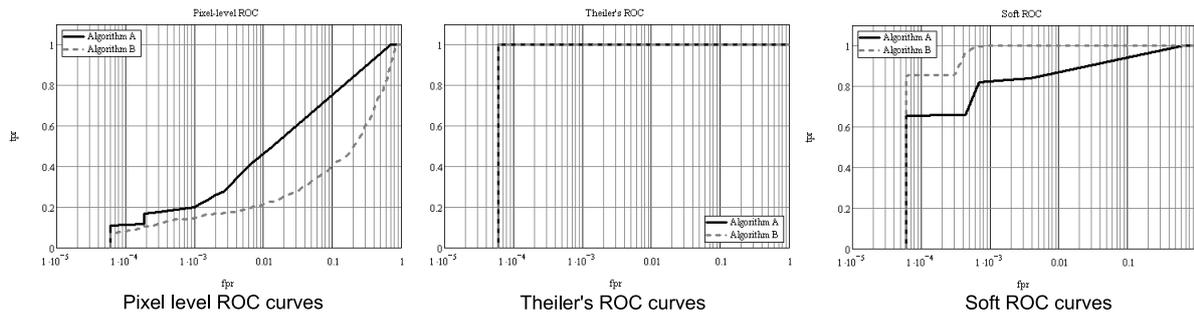


FIGURE 3.14 – Courbes ROC calculées par la méthode de comparaison pixel à pixel des masques, la méthode de Theiler, et la méthode souple. Chacune des courbes a été calculée pour l'algorithme A (courbe noire) et l'algorithme B (courbe en pointillés gris). Les courbes pour la méthode pixel à pixel suggèrent que l'algorithme A est meilleur, alors que notre méthode souple prêche de meilleures performances à l'algorithme B. La méthode de Theiler ne permet pas de discriminer les deux algorithmes : par cette méthode, ils sont tous deux considérés comme parfaits.

Défaut	AUC pixel à pixel	AUC Theiler	AUC souple
Algorithme A	0.81	1	0.945
Algorithme B	0.72	1	0.999947

TABLE 3.4 – Aire sous les courbes ROC pour les algorithmes A et B, calculées avec trois mesures ROC différentes : méthode pixel à pixel, méthode de Theiler et notre comparaison souple.

raison de l'effet néfaste sur le tpr dû à la stratégie du tout ou rien qu'elle utilise. Dans notre application, cette dernière méthode nous semble inappropriée. Afin de lever l'ambiguïté introduite par les résultats contradictoires de la méthode pixel à pixel et de la méthode souple, nous montrons sur la figure 3.13 que l'algorithme B donne de meilleurs résultats, dans la configuration étudiée, que l'algorithme A. La troisième colonne de cette figure représente le résultat d'un seuillage appliqué aux images de la deuxième colonne. Le seuil a été choisi de la manière suivante :

- pour l'algorithme A : le seuil de décision choisi est celui qui permet de maximiser le tpr en générant une seule fausse alarme
- pour l'algorithme B : le seuil de décision choisi est celui qui permet de maximiser le tpr sans générer de fausse alarme

En comparant les masques de détections obtenus à la vérité terrain (quatrième colonne de la figure 3.13), on peut noter que l'algorithme B permet une meilleure détection du défaut. Dans le même temps, l'algorithme A donne une fausse alarme qui est certes proche du défaut, mais suffisamment séparée pour être considérée comme une fausse alarme. L'algorithme B est lui sans fausse alarme. On peut donc déduire que pour notre application, l'algorithme B donne de meilleurs résultats que l'algorithme A. Dans ce contexte, notre méthode ROC souple permet une meilleure évaluation des performances de détection des algorithmes.

3.6.1.2 évaluation des performances d'un algorithme sur deux types de défaut

Comme détaillé dans la section 3.4, l'analyse ROC peut être menée dans des cas où une classe est sur-représentée, car le tpr et le fpr ne sont pas sensibles à une répartition des classes déséquilibrée, et c'est un des principaux avantages de la méthode ROC. Ainsi, on peut utiliser cette méthode afin de comparer les résultats d'un algorithme obtenus sur différents types de défauts, de taille et de forme différente. La figure 3.15 montre les résultats d'un algorithme de détection sur les défauts appelés respectivement «Lines» et «Spot». L'image du défaut, le résultat de la détection sur ce défaut (image en niveau de gris), le

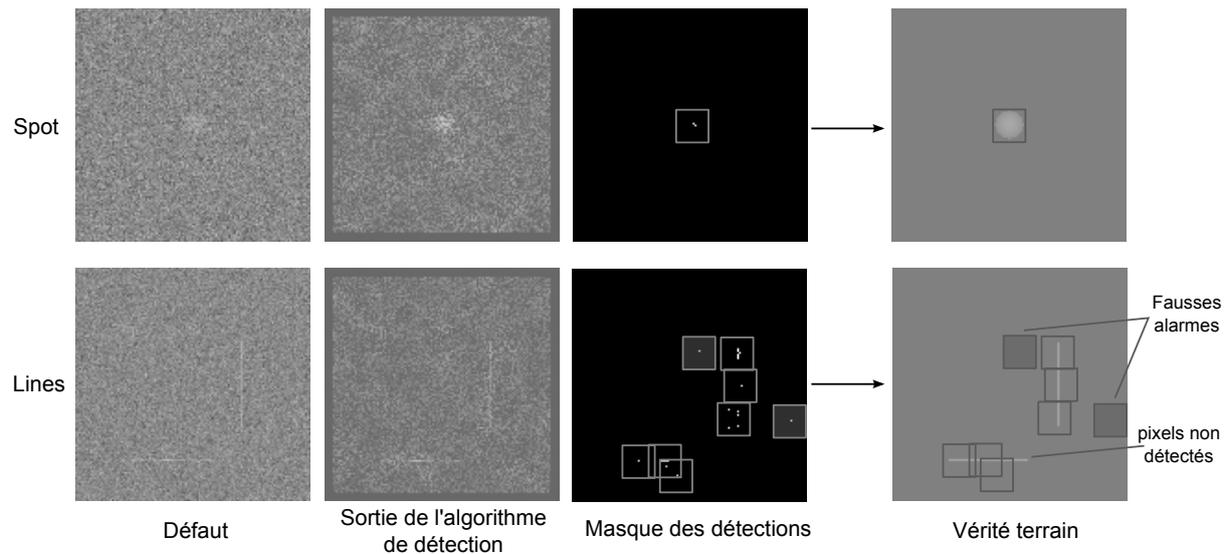


FIGURE 3.15 – Comparaison des performances de détection d’un algorithme sur les défauts Lines et Spot. La première colonne montre l’image d’origine du défaut, sur laquelle est réalisée la détection. La seconde colonne est la sortie de l’algorithme de détection (image en niveau de gris avec les parties détectées apparaissant en clair). La troisième colonne est le seuillage de l’image précédente pour un seuil de décision donné. En surimpression de ces images, on trouve les fenêtres d’observations (carrés gris) qui seront visualisées par l’expert pour vérification. La dernière colonne représente la vérité terrain avec ces mêmes fenêtres d’observations. Pour le défaut Lines, quelques unes de ces fenêtres sont des fausses alarmes.

Défaut	AUC pixel à pixel	AUC Theiler	AUC souple
Lines	0.885	1	0.986
Spot	0.763	1	1

TABLE 3.5 – Mesures des AUC sur les défauts Lines et Spot calculées par trois méthodes différentes : pixel à pixel, Theiler et souple

masque de détection (seuillage de l’image précédente) et la comparaison avec le masque de vérité terrain sont représentés.

Afin de déterminer sur quel défaut l’algorithme est efficace, ses courbes ROC sont tracées pour la méthode pixel à pixel, la méthode de Theiler et la méthode souple (voir figure 3.16). Les AUC correspondantes sont données dans la table 3.5. A partir des résultats donnés par la méthode ROC pixel à pixel, on pourrait dire que le défaut Lines est mieux détecté que le défaut Spot (son AUC associée est supérieure). Aucune conclusion ne peut être tirée de la méthode de Theiler car les AUC sont trop proches de 1.

Dans beaucoup de cas pratiques, on observe que la méthode de Theiler n’est pas assez discriminante : quasi systématiquement les performances de détection obtenues sont excellentes voire parfaites, nous empêchant de faire une comparaison efficace. Finalement, si on se fie aux résultats donnés par la méthode souple, on peut tirer une conclusion différente de celle donnée par la méthode pixel à pixel : le défaut Spot semble mieux détecté que le défaut Lines.

Comme montré dans la figure 3.15, le défaut Spot est en fait mieux détecté que le défaut Lines par l’algorithme. Sur cette figure, deux masques de détections ont été extraits (troisième colonne). Si on considère le défaut Spot, quelques pixels ont été détectés au centre de ce défaut, sans générer de fausse alarme. Pour l’expert humain qui vérifiera les pixels détectés à l’aide de fenêtres d’observation (boîtes grises sur la figure 3.15), le défaut Spot sera entièrement détecté. Dans le même temps, en ce qui concerne le défaut Lines, la figure 3.15 montre que ce défaut n’est détecté que partiellement. De plus, une fausse

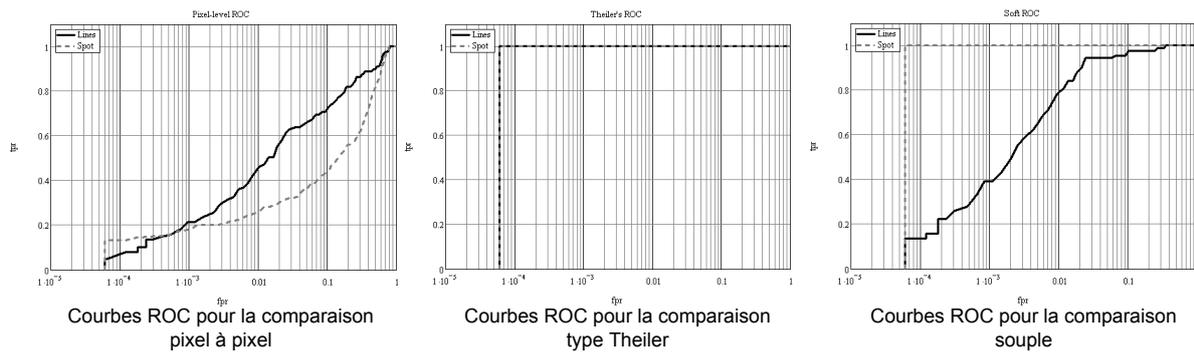


FIGURE 3.16 – Courbes ROC calculées avec les méthodes pixel à pixel, de Theiler, et souple pour les défauts Lines (courbe noire) et Spot (courbe pointillée grise). La méthode pixel à pixel indique que le défaut Lines est mieux détecté, alors que la méthode souple suggère une meilleure détection du défaut Spot. la méthode de Theiler ne permet pas de discriminer les deux détecteurs car elle les juge toutes deux parfaites.

alarme a été générée. Ce couple de masques de détection montre que dans notre contexte, l'algorithme de détection évalué donne de meilleurs résultats sur le défaut Spot que sur le défaut Lines, ce qui est la conclusion de notre méthode de comparaison ROC souple. L'utilisation de la méthode pixel à pixel aurait conduit à une erreur de jugement. On peut donc en déduire que notre méthode souple est ici mieux adaptée à notre besoin que la méthode classique pixel à pixel.

3.6.2 Calibration automatique des algorithmes de détection

Notre méthode de comparaison souple des masques introduite dans les paragraphes précédents peut également être utilisée afin de régler automatiquement le seuil de décision des algorithmes de détection. Ces algorithmes fournissent en sortie généralement des images en niveaux de gris où les pixels clairs sont ceux détectés, et les pixels sombres sont des pixels de fond. Afin d'obtenir un masque de détection assignant chaque pixel à la classe défaut ou fond, il faut régler un seuil de décision afin de binariser l'image de détection.

Quand le masque de vérité terrain est connu, nous pouvons utiliser la méthodologie ROC pour régler automatiquement ce seuil de décision à une valeur correspondant à un tpr et un fpr donné. Dans ce cas, nous allons voir que l'utilisation de la méthode souple permet d'obtenir des masques de détection plus cohérents avec les valeurs de tpr et de fpr choisies.

Si on considère les défauts introduits par la figure 3.6, nous voudrions l'image seuillée des images de détection fournies par l'algorithme. Cette image seuillée doit correspondre à un tpr donné, par exemple $tpr = 100\%$ (cible totalement détectée) avec un nombre *minimum* de fausses alarmes. C'est une image, correspondant à un point de fonctionnement, typiquement observée afin d'évaluer les performances de détection de l'algorithme. La figure 3.17 montre les images correspondant à un $tpr = 100\%$, obtenues pour les défauts amas et tache à profil gaussien, et calculées par deux méthodes de comparaison de masques : pixel à pixel (voir section 3.5.1) et méthode souple (voir section 3.5.3).

Pour le défaut de type amas, la figure 3.17(b) met en évidence la très grande sensibilité de la méthode pixel à pixel : pour avoir un $tpr = 100\%$ avec nos algorithmes de détection (images de la colonne de gauche dans la figure 3.17), strictement tous les pixels composant le défaut doivent être détectés. Ainsi, on doit se placer à un point de fonctionnement générant de nombreuses fausses alarmes, ce qui pourrait laisser penser que l'algorithme fonctionne mal sur le défaut traité. Ce n'est pourtant pas le cas, et la méthode de comparaison souple (figure 3.17(c)) permet d'éviter de telles erreurs d'interprétation : l'image seuillée montre une détection excellente dans notre contexte (les pixels détectés sont suffisants pour détecter le défaut dans sa totalité) avec seulement quelques fausses alarmes. La même conclusion peut être faite avec les images seuillées des figures 3.17(e) et 3.17(f). Dans ces deux cas, la comparaison pixel à pixel est trop contraignante sur la bonne localisation des détections, ce qui conduit à se placer à un point

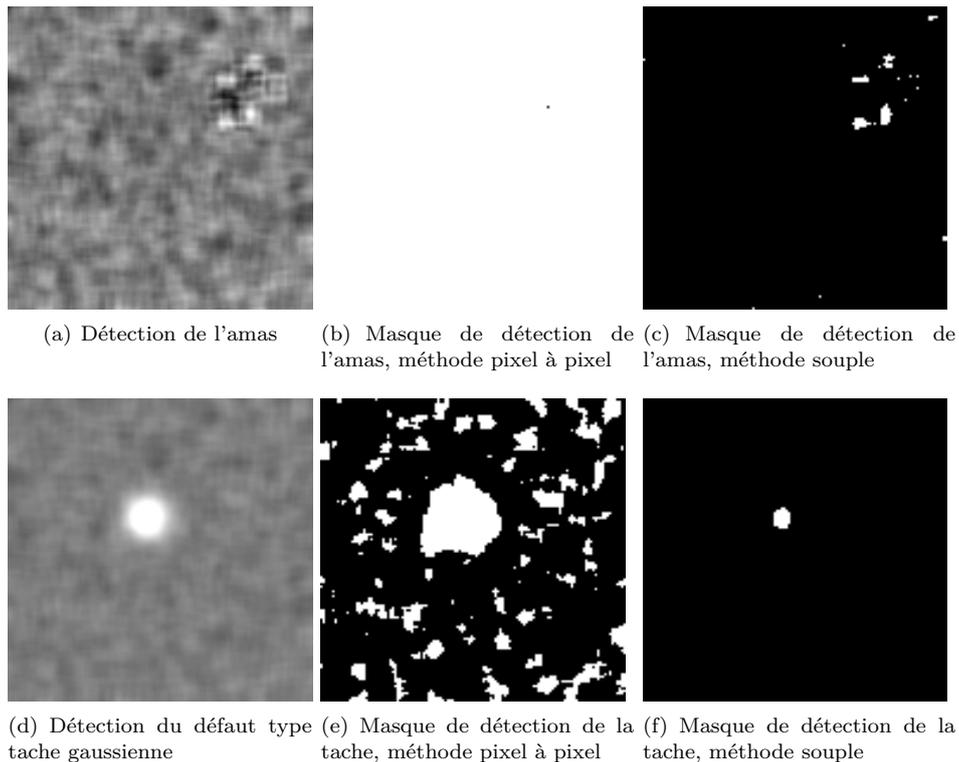


FIGURE 3.17 – Image seuillée pour un $tpr = 100\%$ selon deux méthodes de comparaison de masques : pixel à pixel et méthode souple. La méthode pixel à pixel donne une image avec trop de détections, conduisant à un fort taux de fausse alarme. À l'inverse, la méthode souple fournit une image où la cible est effectivement détectée (nous n'avons pas besoin de plus de pixels détectés pour identifier les deux cibles), avec beaucoup moins de fausses alarmes.

de fonctionnement où trop de pixels sont détectés (et donc avec trop de fausses alarmes) alors que la méthode souple donne des images seuillées satisfaisantes.

3.7 Commentaires et extensions possibles de la méthode

La méthode souple développée dans ce chapitre est un premier pas vers une comparaison haut niveau des masques, en opposition à l'aspect bas niveau de la méthode pixel à pixel, qui est bien adaptée à la tâche d'aide à l'inspection visuelle d'une image. Quelques améliorations immédiates peuvent être envisagées.

Tout d'abord, pour le calcul du taux de fausse alarme, on pourrait prendre en considération la taille des AOIs qui seront présentées à l'expert pour vérification visuelle. Notre méthode réalise un décompte pixel à pixel des fausses alarmes, ce qui correspondrait à une inspection pixel à pixel de ces fausses alarmes, *i.e.* l'expert utilise des AOIs de 1×1 pixels. Nous devrions plutôt considérer la vraie dimension des fenêtres d'observations du système d'inspection de manière à rassembler les fausses alarmes proches en une seule fausse alarme, puisque c'est tel qu'elles seront vues par l'expert. Ceci peut être fait en divisant l'image en fenêtres carrées de même taille que les AOIs réellement utilisées, et en dénombrant le nombre de fenêtres où apparaissent des fausses alarmes. Le nombre de fausses alarmes serait ainsi exactement le nombre de fois où une AOI sans défaut serait présentée à l'expert. En suivant la même idée, on peut également réaliser une dilatation des pixels de fausse alarme d'un facteur de la taille des AOIs afin de fusionner les fausses alarmes suffisamment proches en une seule fausse alarme. Dans ce cas, il faut cependant gérer le problème de normalisation afin d'obtenir le taux de fausse alarme.

Dans un second temps, dans notre approche le tpr a été normalisé par le nombre P de pixels défectueux

dans l'image (voir l'équation 3.7). Ce choix a été fait afin de garder une interprétation rapide de ce taux, mais ceci peut être trop restrictif dans certains cas. Si on considère une image contenant deux défauts : un de grande taille et un défaut ponctuel (un seul pixel défectueux). Un algorithme de détection qui détecterait correctement le premier défaut mais manquerait le second atteindrait néanmoins un *tpr* relativement élevé alors qu'il aura manqué un défaut. Afin d'éviter ce genre de situation contradictoire, il faudrait adopter une approche de normalisation encore plus orientée objet : le *tpr* serait calculé pour chaque cible (et donc à chaque fois normalisé par le nombre de pixel de la cible) et le *tpr* global sur l'image serait déterminé en moyennant les taux obtenus pour chacune des cibles de l'image. Ainsi, les cibles de tailles différentes avec la même part de pixels détectés auraient le même impact sur la valeur globale du taux de bonne détection. De plus, si la taille de l'AOI est connue, on pourrait également définir un facteur de dilatation constant pour effectuer la dilatation du masque de détection cohérent avec cette taille caractéristique d'observation.

Pour conclure, plusieurs améliorations de la méthode souple peuvent être réalisées en utilisant une connaissance *a priori* du système de détection. Cependant, l'évaluation globale des performances demeure inchangée (dilatation du masque de vérité terrain et du masque de détection).

3.8 Conclusion

La méthode exposée dans ce chapitre fournit un outil permettant de comparer les algorithmes de détections d'une manière facile et cohérente. En effet, plutôt que de réaliser une comparaison pixel à pixel des objets, nous utilisons une approche morphologique permettant de favoriser les systèmes réalisant une détection non exhaustive mais néanmoins efficace sur chacun des défauts à détecter. Dans la suite de l'étude des algorithmes de détection, nous utiliserons l'outil des courbes ROC, couplé à notre méthode de comparaison souple des masques, afin de comparer leurs performances et de se doter d'une mesure objective, via l'aire sous la courbe ROC, de ces performances. En pratique, cette procédure est un peu plus lente que la comparaison pixel à pixel, mais le coût est supportable car les calculs les plus lourds (calcul des facteurs de dilatation) ne sont effectués qu'une seule fois par tracé de courbe ROC. Enfin, la procédure définie s'adapte automatiquement à la taille et la forme des défauts, et ne nécessite aucun réglage de paramètre ce qui en fait un outil très efficace pour traiter le large spectre de défauts que nous avons à détecter.

Chapitre 4

Méthodes de détection de défauts

Abstract

L'étape de détection est un élément fondamental du contrôle qualité des images. Dans les cas réels, il faut appliquer cette détection avec peu d'a priori sur la nature des défauts à détecter, puisqu'une image peut contenir une variété importante de défauts. Dans ce chapitre, nous nous attachons à améliorer la brique de détection proposée dans la partie précédente. A partir d'un jeu de défauts synthétiques, représentatifs de la variété des défauts pouvant être observés dans le cas réel, plusieurs méthodes de détection sont proposées. A chaque fois, les performances des méthodes sont comparées aux meilleures performances constatées avec les méthodes décrites précédemment, à l'aide de la méthodologie ROC décrite dans le chapitre 3. La dernière méthode exposée, la méthode dite de la moyenne statistique pondérée, donne de bons résultats sur une large gamme de défauts. Néanmoins, des améliorations de cette méthode dans le cas des images réelles doivent être envisagées, notamment via une implémentation multi-échelle et une fusion des résultats issus de plusieurs algorithmes de détection.

Sommaire

4.1	Objectifs de la détection de défauts	85
4.2	Introduction aux méthodes de détection	87
4.3	La méthode de base : seuillage direct de l'image	92
4.4	Méthodes à base de statistiques locales : moyenne, variance, skewness et kurtosis	92
4.5	Détection par décomposition en ondelettes	100
4.6	Détection par filtrage directionnel	106
4.7	Détection par mesure d'anisotropie	112
4.8	Détection par moyenne statistique pondérée	116
4.9	Détection par snakes statistiques et rapports de vraisemblance	119
4.10	Synthèse	122
4.11	Conclusions et perspectives	124

4.1 Objectifs de la détection de défauts

Nous avons vu dans le chapitre 2 une méthode de détection relativement basique permettant de détecter les défauts présentant soit des pixels de niveaux très atypiques, soit des zones dont la moyenne locale ou la variance locale est atypique par rapport au reste de l'image. Nous avons également vu que dans certains cas, cette technique ne parvient pas à détecter correctement les défauts. C'est notamment le cas pour les défauts présentant des niveaux de gris proches de celui du bruit de fond, mais étant visibles par la corrélation spatiale qu'ils présentent. La figure 4.1 donne quelques exemples de défauts présents sur des images réelles.

Dans ce chapitre, nous allons détailler plusieurs techniques de détection permettant d'améliorer les performances de détection notamment sur des défauts présentant des structures corrélées spatialement. Afin de mesurer l'amélioration de performance réalisée, nous comparerons à chaque fois les résultats obtenus à ceux de la méthode présentée dans le chapitre 2.

Pour cela, on se munit d'échantillons de défauts synthétiques. Dans ces échantillons, le bruit de fond est blanc. On se place ainsi dans un cas particulier de contexte, correspondant à une image acquise sans dose de rayons X. Ce choix se justifie d'abord par la diversité de bruit qu'on peut rencontrer sur les images des détecteurs de la gamme Trixell : on observe une grande variété des fonctions de transfert des détecteurs en fonction du modèle (variation de la taille du pixel, variation des techniques d'assemblage) et en fonction du mode d'acquisition. Il faut donc faire un choix de manière à se trouver toujours en présence du même type de bruit. On a choisi le bruit blanc.

Afin de compléter l'étude des algorithmes de détection, il serait utile de voir comment varient leurs performances lorsqu'on fait varier le spectre du bruit. Cette étude n'a pas été conduite dans le cadre de cette thèse, où nous avons plutôt privilégié le développement de plusieurs algorithmes de détection.

On veut développer une nouvelle méthode de détection qui a les caractéristiques suivantes :

- elle ne nécessite pas d'information *a priori* sur la forme ou la taille du défaut,
- elle détecte tous les défauts visibles à l'œil nu,
- elle estime de manière assez fiable la forme des défauts détectés : une tolérance est acceptée puisqu'on se contenterait de détecter un ou plusieurs pixels au centre de chaque défaut, les étapes suivantes (notamment celle de localisation et de débruitage) se chargeant de réaliser une estimation du défaut. En pratique, on souhaite simplement que la détection puisse permettre à un observateur humain de retrouver rapidement dans l'image le défaut associé.
- elle génère très peu de fausses alarmes : le nombre de fausses alarmes ciblé est d'une fausse alarme pour cent images traitées. Ce taux correspond à une contrainte d'utilisation industrielle, où chaque fausse alarme génère une perte de temps.

La stratégie que nous avons adoptée pour traiter le problème est la suivante : étant donnée la généralité demandée, nous avons testé plusieurs méthodes dans l'optique que chaque type de défaut soit bien détecté par au moins une de ces méthodes. Par la suite, grâce à la palette d'algorithmes ainsi constituée, une

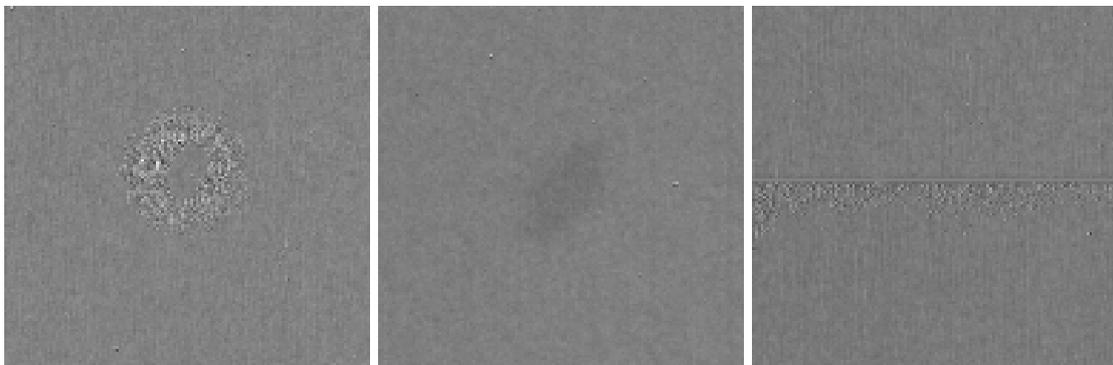


FIGURE 4.1 – Exemples de défauts sur images réelles

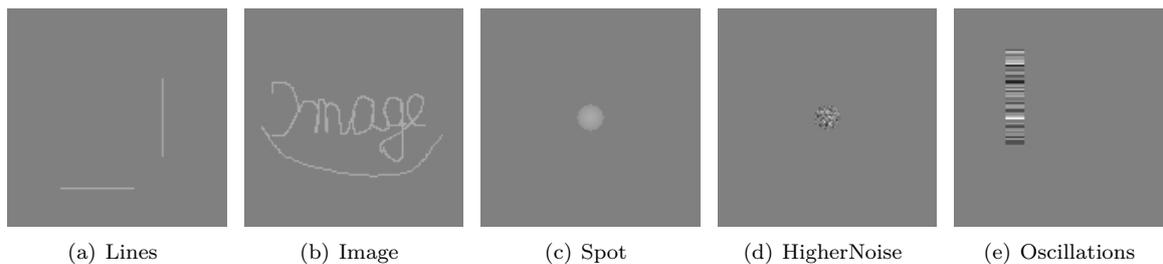


FIGURE 4.2 – Masque des défauts synthétiques

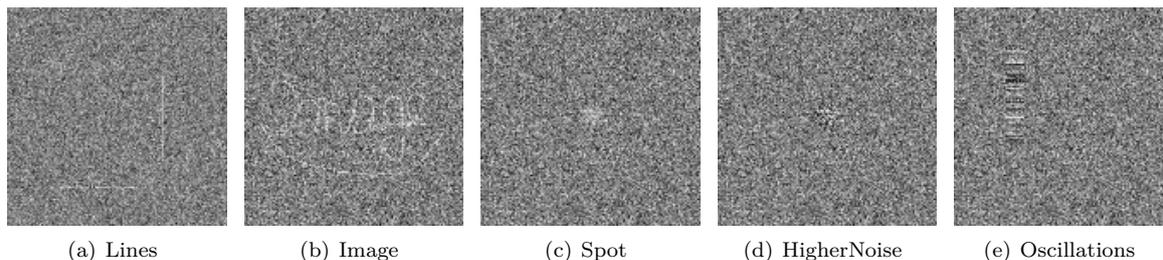


FIGURE 4.3 – Défauts synthétiques

stratégie de fusion peut permettre de former un détecteur efficace sur tout type de défaut.

4.1.1 Données de test

Afin de tester les méthodes de détection développées, nous avons construit une banque de cinq défauts synthétiques, de taille 128×128 pixels, représentatifs des défauts pouvant être rencontrés sur des images réelles. Les masques de ces défauts sont représentés sur la figure 4.2 : le fond de l'image est représenté en gris et la zone de défauts en plus clair ou plus sombre.

Ces masques ont ensuite été bruités par un bruit additif blanc gaussien (voir figure 4.3). Les masques des défauts ont été construits de différentes manières :

- les défauts Lines et Image sont constitués d'une zone claire uniforme de forme variable,
- le défaut spot est constitué d'une zone circulaire claire à contours flous (profil gaussien),
- le défaut Highernoise est constitué d'une zone circulaire contenant un bruit gaussien de variance plus élevée que le fond,
- le défaut Oscillations a été construit à partir d'une zone rectangulaire constitué de segments horizontaux de niveau constants, la valeur de ces segments variant selon l'axe vertical. Cette variance est supérieure à celle du bruit de fond (voir table 4.1).

En notant F la variable aléatoire représentant les pixels du fond et D la variable aléatoire représentant les pixels du défaut, on peut calculer les moyennes μ_F et μ_D , ainsi que les écarts-type σ_F et σ_D . À titre indicatif, on peut également calculer un rapport statistique, le ratio de Fisher $\mathcal{F}(D, F)$ défini dans [17] par :

$$\mathcal{F}(F, D) = \frac{(\mu_F - \mu_D)^2}{\sigma_F^2 + \sigma_D^2} \quad (4.1)$$

On peut également calculer un rapport signal à bruit classique :

$$RSB_{dB}(F, D) = 10 \log_{10} \frac{(\mu_D - \mu_F)^2}{\sigma_F^2} \quad (4.2)$$

Les valeurs de ces différentes mesures statistiques sont restituées dans le tableau 4.1. Le ratio de

Défauts	μ_F	σ_F	μ_D	σ_D	RSB (dB)	$\mathcal{F}(F, D)$
Lines	127,96	34,96	186,09	39,64	4,42	1,21
Image	127,38	42,43	188,77	39,64	3,21	1,10
Spot	127,35	42,43	156,38	43,33	-3,30	0,23
Oscillations	127,35	42,47	128,20	58,80	-33,94	0,00
HigherNoise	127,35	42,43	125,97	66,90	-29,76	0,00

TABLE 4.1 – Mesures statistiques sur les défauts synthétiques. Les ratio de Fisher des défauts Oscillations et HigherNoise sont quasi nuls (respectivement de l'ordre de 10^{-6} et 10^{-4}).

Fisher est une mesure permettant de traduire la «séparabilité statistique» de F et D ¹. Cette mesure est le critère à optimiser, utilisé en analyse discriminante, afin de séparer deux populations. Utilisé comme ici, il mesure la séparabilité des deux populations F et D par simple seuillage.

En observant attentivement ces défauts et les mesures associées, on constate que les pixels de défauts, pris individuellement, ne sont pas facilement différenciables du fond. En effet, leur valeur peut tout à fait être une valeur prise par la variable aléatoire F . Ce qui rend ces défauts visibles est leur agencement spatial, leur corrélation spatiale, qui est interprétée par l'œil. Les niveaux de RSB des défauts synthétiques confirment ce qu'on observe visuellement : les défauts sont au niveau du bruit de fond, d'où les RSB très faibles constatés. On peut considérer que les défauts sont repartis en deux classes distinctes :

- les défauts ponctuels : un pixel isolé qui possède une luminosité aberrante qui le rend visible par rapport à la texture de fond. Ces défauts sont facilement détectables par un simple seuillage s'appuyant sur une connaissance limitée de la statistique du fond de l'image.
- les défauts corrélés : un agencement de plusieurs pixels qui, interprété par l'œil, est identifié comme une structure différente du fond de l'image.

Les méthodes de détection développées dans ce chapitre auront pour objectif de détecter la deuxième classe de défauts.

4.2 Introduction aux méthodes de détection

4.2.1 Les méthodes classiques de détection ou de segmentation

La tâche de détection de défauts ou d'objets dans une image a été la cible de nombreuses études issues principalement de la communauté du contrôle qualité. On peut classer les travaux réalisés en six catégories :

- les méthodes à base de détection de contour : leur but est de détecter les objets en cherchant les contours séparant ces objets du fond de l'image,
- les méthodes d'extraction d'attribut (suivies ou non d'une classification) et les méthodes d'extraction d'objet dans des textures : elles s'appuient sur une connaissance de la texture du fond de l'image pour détecter ses irrégularités, caractéristiques de la présence d'un objet,
- les méthodes à base de changement d'espace de représentation : leur but est de transformer l'image d'origine par changement d'espace de représentation. Dans ce nouvel espace, la détection peut être plus aisée,
- les méthodes de segmentation : watershed, contours actifs, segmentation markovienne : elles ont

1. La mesure de visibilité à proprement parlée fera l'objet du chapitre suivant. Nous ferons juste remarquer que la comparaison des valeurs de ce ratio et des valeurs de visibilité que pourrait donner un observateur humain justifie le besoin d'une mesure de visibilité plus adaptée. En effet, les défauts aux ratio de Fisher les plus faibles ne sont pas forcément les moins visibles (c'est le cas du défaut Image). Néanmoins, ce chiffre permet de rendre compte, d'une manière statistique, de la difficulté de détection de certains défauts synthétiques.

pour but soit de séparer un ou des objets du fond de l'image, soit de réaliser une partition entière d'une scène.

Il existe bien sûr d'autres types de méthodes de détection ou de segmentation ne pouvant être classées dans ces catégories, mais ce sont généralement des méthodes très spécialisées (notamment type de défaut connu et modélisé).

Pour chaque catégorie, nous allons donner les principes des méthodes présentes dans la littérature en citant quelques références particulièrement axées sur notre problématique de détection de défauts.

4.2.1.1 Détection de contours

Les méthodes dérivatives Les méthodes à base de détection de contour sont les plus anciennes, et généralement simples à appliquer. Le principe est de détecter le contour de l'objet. Les premières méthodes de détection de contour développées s'appuient sur le calcul d'une approximation du gradient de l'image, afin de détecter les discontinuités d'intensité. Si on représente l'image sous la forme d'une fonction f définie dans le plan de l'image, ces méthodes cherchent typiquement à détecter les *maxima* locaux de la norme du vecteur gradient \vec{G} avec :

$$\vec{G} = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} (x, y) \quad (4.3)$$

Dans les zones de transition entre le fond et l'objet à détecter, la norme de \vec{G} est censée être élevée par rapport à sa valeur dans les zones de fond. Par cette méthode, on peut donc déterminer les pixels de contours (ce sont les maxima locaux de $\|\vec{G}\|$) ainsi que leur orientation (les contours suivent les lignes perpendiculaires au gradient).

Pour une image numérique, la difficulté repose sur l'estimation de \vec{G} , *i.e.* sur l'estimation des dérivées partielles de f . Pour cela, on utilise classiquement les masques définis par L. Roberts [77], Prewitt [75] et Sobel [81].

Selon le type de contours présents dans l'image, on utilisera les masques horizontaux et verticaux, ou diagonaux. Dans les deux cas, on obtient les deux composantes vectorielles du gradient, qui sont utilisées dans le calcul de sa norme. En pratique, on effectue sur l'image un prétraitement de type lissage avant d'appliquer les masques de Prewitt ou de Sobel, afin de limiter les fausses détections de contours liées au bruit.

Cette technique est également utilisé dans l'algorithme développé par J. Canny [14] Cet algorithme utilise les propriétés du gradient exposées ci-dessus, et implémente une phase de construction de contour plus élaborée, visant à produire des contours fins et mieux localisés. L'objectif de cet algorithme est de :

- minimiser le taux d'erreur (contours non détectés et détection de faux contours)
- bien localiser les contours (minimiser la distance entre contours détectés et contours réels)
- avoir unicité de la réponse du filtre à un contour (stabilité)

En pratique, l'algorithme de Canny peut être implémenté en quatre étapes. La première consiste à filtrer l'image dont on veut détecter les contours par une gaussienne². La seconde étape correspond au calcul de l'image de gradient. On calcule en chaque point la norme du vecteur gradient estimée à l'aide de masques dérivatifs. La troisième étape, appelée « suppression des non-maxima » vise à former des contours plus fins. Enfin, la dernière étape permet de fermer les contours, par un seuillage à hystérésis.

La fig 4.4 présente les résultats de l'algorithme de Canny sur plusieurs images de défauts synthétiques. Le jeu de paramètres choisi pour l'algorithme de Canny a été réglé empiriquement pour obtenir le meilleur résultat possible simultanément sur les cinq défauts testés (*i.e.* le jeu de paramètres est le même pour les cinq images résultats présentées).

2. La variance de cette gaussienne sera choisie selon le niveau de bruit dans l'image et le type de contours que l'on veut extraire : si l'image est très bruitée, ou si on veut détecter seulement les contours basse fréquence, on utilisera une gaussienne à faible variance

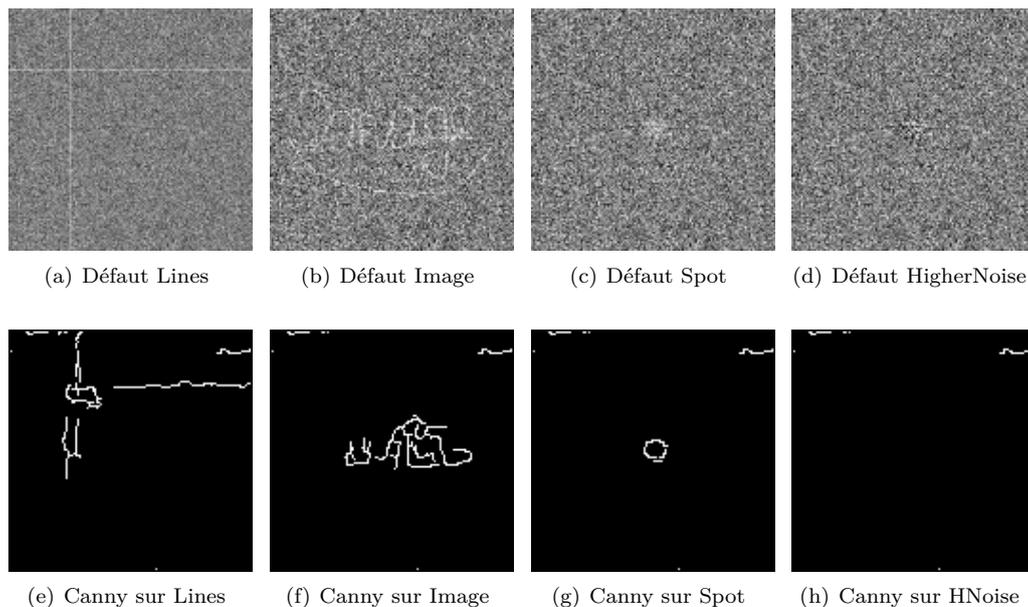


FIGURE 4.4 – Résultats de l’algorithme de Canny sur quelques défauts synthétiques

On constate naturellement que l’algorithme de Canny, bien que donnant un résultat honorable sur le défaut Spot (fig 4.4(g)), est loin d’être optimal pour notre problème, puisqu’il génère de nombreuses fausses alarmes qui se situent de surcroît relativement loin du défaut à détecter.

L’algorithme de Canny est le plus utilisé pour la détection de contours, même si il existe certaines alternatives. On citera notamment Lim [56] qui s’appuie sur une approche statistique afin de trouver le contour d’un objet en considérant une vignette de l’image que l’on cherche à séparer en deux classes (fond et objet) à partir de la simulation de contour séparant ces deux classes.

Toutes ces méthodes s’appuient sur le calcul du gradient \vec{G} de l’image, ce qui est responsable des mauvais résultats constatés sur la figure 4.4. En effet, dans notre cas le fond n’est pas uniforme mais fortement bruité, et le rapport signal à bruit n’est pas suffisant pour que l’information contenue par \vec{G} soit pertinente.

Les méthodes statistiques Une autre classe d’algorithmes se révèle particulièrement intéressante pour détecter, notamment, des contours, mais également toutes configuration particulière de pixels pouvant trahir la présence d’un défaut. Ces méthodes utilisent d’abord la théorie des gestalts [21] qui donne un ensemble de configurations géométriques particulières susceptibles d’être interprétées par l’œil. On trouve notamment dans ces configuration les alignements de pixels, et les taches. On utilise alors une statistique afin de tester la probabilité des gestalts, connaissant la valeur des pixels les composant et la loi qu’ils sont censés suivre. Cette technique est connue sous le nom de principe de Helmholtz [20, 22]. On obtient alors une méthode de détection par recherche d’une configuration connue de pixels, et comparaison statistique de l’occurrence de la configuration trouvée à un seuil dépendant du taux de fausse alarme. Ce taux de fausse alarme est un paramètre de la méthode. Dans le cas où la variable aléatoire régissant le niveau des pixels est indépendamment et identiquement distribuée selon une loi uniforme, ce taux de fausse alarme est calculable analytiquement et permet de régler le système de détection de manière très simple. Le cas d’autres lois de probabilité a été également étudié [37]. Néanmoins, la méthode de détection ainsi définie souffre d’inconvénients dont il faut apporter des solutions dans le cadre d’une utilisation industrielle. Le principal est qu’il ne semble pas y avoir d’implémentation rapide permettant d’éviter le test exhaustif de toutes les structures recherchées. Cette recherche exhaustive n’est pas réalisable pour le traitement d’images de grande taille.

Une méthode analogue a été testé sur une banque de huit défauts synthétiques englobant les défauts

synthétiques présentés en début de chapitre. Afin de s'affranchir du problème de recherche de gestalts, l'image à analyser était seuillée à deux fois l'écart-type du bruit puis dilatée afin de définir des ensembles connexes de pixels détectés. Sur chaque ensemble, un test du type principe de Helmholtz a été réalisé pour savoir si cet ensemble devait être détecté ou non. Les résultats ont permis de conclure que la méthode n'était pas assez sensible pour détecter des configurations de pixels ne s'adaptant pas au noyau de dilatation. En particulier, avec le noyau carré choisi, il s'est révélé impossible de détecter des structures linéaires.

4.2.1.2 Méthodes d'extraction d'attributs

Une autre classe de méthodes couramment utilisée est constituée d'algorithmes cherchant à différencier les pixels de défaut des pixels du fond à l'aide de critères qui peuvent être statistiques ou spectraux. Les critères statistiques simples pouvant être utilisés sont ceux de la moyenne et de la variance. Le principe est de comparer la valeur moyenne ou la variance du voisinage d'un pixel à celle de l'image entière. On peut par exemple former le ratio $\frac{\text{moyenne locale}}{\text{moyenne référence}}$ ou $\frac{\text{variance locale}}{\text{variance référence}}$. La référence est calculée soit sur toute l'image, soit sur un voisinage étendu autour du pixel traité dans les cas où l'image est de grande taille, afin de limiter les effets de non-stationnarité du fond. L'étape finale consiste à prendre une décision (par seuillage en général) sur la nature de chaque pixel selon la valeur du ratio de moyenne ou de variance associé. Cette méthode correspond à celle décrite dans le chapitre 2 pour la détection.

De manière moins triviale, on peut utiliser, selon le même principe, les statistiques d'ordre supérieur pour réaliser la détection. On calcule alors le skewness (moment d'ordre 3) et le kurtosis (moment d'ordre 4) locaux. Pour un bruit gaussien, ces quantités sont nulles, ce qui permet de ne pas avoir à calculer de valeur de référence si le défaut est plongé dans un bruit gaussien. On rappelle les définitions du skewness \mathcal{S}_X et du kurtosis³ \mathcal{K}_X pour une variable aléatoire X :

$$\mathcal{S}_X = \frac{\mu_3(X)}{\mu_2^{\frac{3}{2}}(X)} \quad (4.4)$$

$$\mathcal{K}_X = \frac{\mu_4(X)}{\mu_2^2(X)} - 3 \quad (4.5)$$

où $\mu_k(X) = E[(X - \bar{X})^k]$ est le moment centré d'ordre k de X , \bar{X} étant la moyenne de X .

Physiquement, le skewness est une mesure de l'asymétrie de la densité de probabilité P_X de X par rapport à sa moyenne et le kurtosis mesure l'aplatissement de cette densité par rapport à une densité gaussienne. Ainsi, pour une variable aléatoire suivant une loi gaussienne, $\mathcal{S}_X = 0$ et $\mathcal{K}_X = 0$. Pour une étude plus complète de ces statistiques d'ordre supérieur, se référer à [53].

L'utilisation de ces statistiques d'ordre supérieur (notées SOS) sur des simulations d'images radiographiques a été testée par F. Maussang [66]. Avec les mêmes méthodes, les SOS peuvent être avantageusement utilisées pour la détection d'objets enfouis sur des images sonar fortement bruitées [65].

Nous avons appliqué ces critères à notre problématique de détection des défauts synthétiques. Les résultats seront présentés en section 4.4.

D'autres méthodes de détection de défauts font l'objet de nombreuses études : il s'agit des méthodes de détection utilisées dans un contexte industriel de contrôle qualité qui utilise l'aspect texturé des surfaces à inspecter (structures tissées, structures de bois, ...). Dans [51], A. Kumar dresse une synthèse des différentes méthodes de détection de défauts sur fond texturé et applique une méthode à base de filtres linéaires à réponse impulsionnelle finie (RIF). Dans cette méthode l'information *a priori* sur la texture du fond, et notamment la fonction d'autocorrélation de cette texture, est largement utilisée. Dans notre cas, la texture de fond, à savoir un bruit, contient très peu d'information ce qui rendra inefficace toute méthode s'appuyant sur des informations de texture. Néanmoins, il faut avoir connaissance de ces méthodes dans le cas où on voudrait par exemple détecter les défauts dans des images intermédiaires où peuvent apparaître

3. Appellation prise par abus de langage. En fait il s'agit de l'excès de kurtosis

des structures connues à l'avance.

4.2.1.3 Méthodes de changement d'espace de représentation

Une méthode particulièrement efficace pour détecter tout type de défaut est de changer l'espace de représentation de l'image pour passer dans un espace adapté au défaut à détecter, dans lequel ce défaut sera porté par un faible nombre de coefficients, et serait donc facilement détectable. Un exemple bien connu de ce principe est réalisé par la méthode dite du filtrage adapté, où on se sert d'une connaissance *a priori* du défaut à détecter pour fabriquer un filtre RIF qui est utilisé pour filtrer l'image. La présence d'un défaut du type recherché est alors détectable sur un nombre très réduit de coefficients, dans le cas idéal sur un seul point de l'image résultat. Cette méthode de détection optimale n'est pas utilisable dans notre cas puisqu'on ne dispose pas de modèle du défaut à détecter. De plus il peut y avoir plusieurs défauts hétéroclites dans une seule image. Dans ce cas général, on peut chercher à décomposer le signal à détecter dans la base la mieux adaptée, *i.e.* la base qui offrira une représentation la plus «parcimonieuse» possible du défaut. Par abus de langage, on entend par représentation parcimonieuse que la nouvelle base permette la décomposition du défaut sur un nombre significativement plus faible de vecteur de base (donc avec plus de coefficients nuls) que la base de représentation d'origine.

Outre la base de représentation naturelle des images (base de diracs), on peut citer la classique base de Fourier (base d'exponentielles complexes) qui est adaptées à la représentation des signaux périodiques. Dans notre cadre de détection, ces deux bases sont peu intéressantes : seuls les défauts de type pixels isolés sont vraiment détectables directement dans la base naturelle, alors que la base de Fourier est difficilement utilisable pour la détection car il y est difficile de séparer le fond bruité de l'image (non texturé il est réparti sur l'ensemble des vecteurs de la base de Fourier) du défaut. Le défaut a en général une représentation peu parcimonieuse dans Fourier.

Une autre base de représentation particulièrement intéressante dans le cadre de la détection est la base d'ondelettes. Elle permet la représentation parcimonieuse de signaux compacts en localisation et en fréquence, ce qui est le cas des défauts qu'on veut détecter. De plus, la multitude de vecteurs de base possibles permet d'adapter la décomposition aux types de signaux à détecter, sans trop perdre en généralité. La transformation en ondelettes [61], théorie développée au début des années 1980 par Morlet puis largement popularisée dix ans plus tard notamment par Daubechies [18], a connu depuis un fort développement grâce à la mise au point d'algorithmes de calcul rapides et efficaces par banc de filtres. Pour notre application, nous avons élaboré une méthode de détection dans une base d'ondelettes : le principe et les résultats sont exposés en section 4.5.

D'autres transformations sont appliquées de manière classique en détection. On peut citer la transformée de Hough [43], qui permet de détecter des lignes par projection dans un espace de paramètres (typiquement : angle de la droite dans le plan de l'image et un paramètre d'offset). Cette transformée se montre très efficace pour la détection de lignes de grande taille puisqu'elle transforme ces lignes en un point dans l'espace des paramètres. Cependant, cette transformée n'est pas adaptée à notre problématique car :

- elle n'est vraiment adaptée qu'à la détection de lignes ou de cercles,
- elle nécessite que les lignes à détecter soient de grande dimension par rapport à celle de l'image, de manière à être suffisamment robuste au bruit,
- la transformée de Hough est non-réversible dans un cadre général (image en niveaux de gris) : en particulier, les segments détectés dans la base de Hough ne peuvent pas être différenciés des droites parcourant toute l'image.

4.2.1.4 Les méthodes de segmentation

Enfin, une dernière classe de méthodes peut être identifiée : les méthodes de segmentation d'objets.

Une segmentation de \mathbb{R}^2 est une partition P de \mathbb{R}^2 avec :

$$\begin{cases} \mathcal{P} = \{P_i\}_{i=1\dots n} \\ \forall (i, j) \in \mathbb{N}^2, i \neq j, P_i \cap P_j = \emptyset \end{cases}$$

De nombreuses méthodes de segmentation existent, la plupart ayant pour but de décomposer la scène représentée en objets. Ici, notre problématique est différente car on veut extraire les défauts d'une image en présence d'un fond bruité. Dans ce cas, les méthodes de segmentation les plus adaptées sont celles orientées segmentation d'objet et non de scènes complètes. Dans le cadre de ces travaux, deux méthodes de segmentation ont été testées *via* des partenariats avec des équipes expertes dans leurs domaines respectifs : la segmentation par snakes statistiques [32, 73] avec la société KaoLab, et la segmentation par champs de Markov [33, 74] avec l'INRIA Rhône-Alpes. Les résultats obtenus par ces deux méthodes seront présentés en section 4.9.

4.3 La méthode de base : seuillage direct de l'image

La méthode classique de détection, qu'on va prendre comme méthode de base, consiste en un simple seuillage de l'image. On va choisir un seuil s , et on va décider pour chaque pixel (i, j) de l'image I :

$$\begin{cases} (i, j) \in D \text{ si } |I(i, j) - \hat{\mu}| \geq s\hat{\sigma} \\ (i, j) \in F \text{ sinon} \end{cases} \quad (4.6)$$

où $\hat{\mu}$ et $\hat{\sigma}$ sont respectivement la moyenne et la variance du bruit de fond estimées. Cette estimation peut soit s'appuyer sur une connaissance *a priori*, soit être le résultat de l'estimation sur l'image traitée, comme c'est généralement le cas dans nos applications. Afin d'être plus robuste à la présence de défauts, on utilise, comme dans le chapitre 2, les estimateurs robustes (voir l'équation 2.11).

Les courbes ROC de la détection des cinq défauts synthétiques, par la méthode de seuillage simple, sont représentées sur la figure 4.6. On représente également les images issues de cette détection sur la figure 4.5. On trouve sur cette figure les images en sortie de l'algorithme de détection (sur la deuxième ligne). A partir de ces images, on détermine le seuil de décision conduisant à un taux de fausse alarme nul, puis à un taux de bonne détection de 100% (respectivement quatrième et cinquième ligne). On constate que les artefacts sont globalement mal détectés, ce qui prouve que la tâche de détection est ardue. En particulier, les défauts Image et Lines sont très mal détectés. Les défauts Spot et Oscillations sont assez bien détectés tant qu'on se contente d'un taux de bonne détection du défaut réduit. Par contre, pour une identification plus complète du défaut, on constate qu'il faut tolérer un nombre très important de fausses alarmes (voir cinquième ligne de la figure 4.5).

Cet algorithme trivial va servir de base de comparaison pour les autres algorithmes de détection que nous allons détailler.

4.4 Méthodes à base de statistiques locales : moyenne, variance, skewness et kurtosis

4.4.0.5 Principe

Le but des méthodes à base de statistiques locales est de détecter les défauts à l'aide d'un changement localisé d'une mesure statistique induit par la présence d'un défaut. Le principe est le suivant : on va parcourir pixel par pixel l'image à analyser et calculer pour chaque pixel une statistique locale sur un voisinage à définir, qu'on va comparer à la même statistique calculée sur toute l'image ou sur un voisinage étendu, qui sera appelée statistique de référence. Dans toute la suite, on prendra les notations suivantes :

- $I(i, j)$ l'image à analyser,

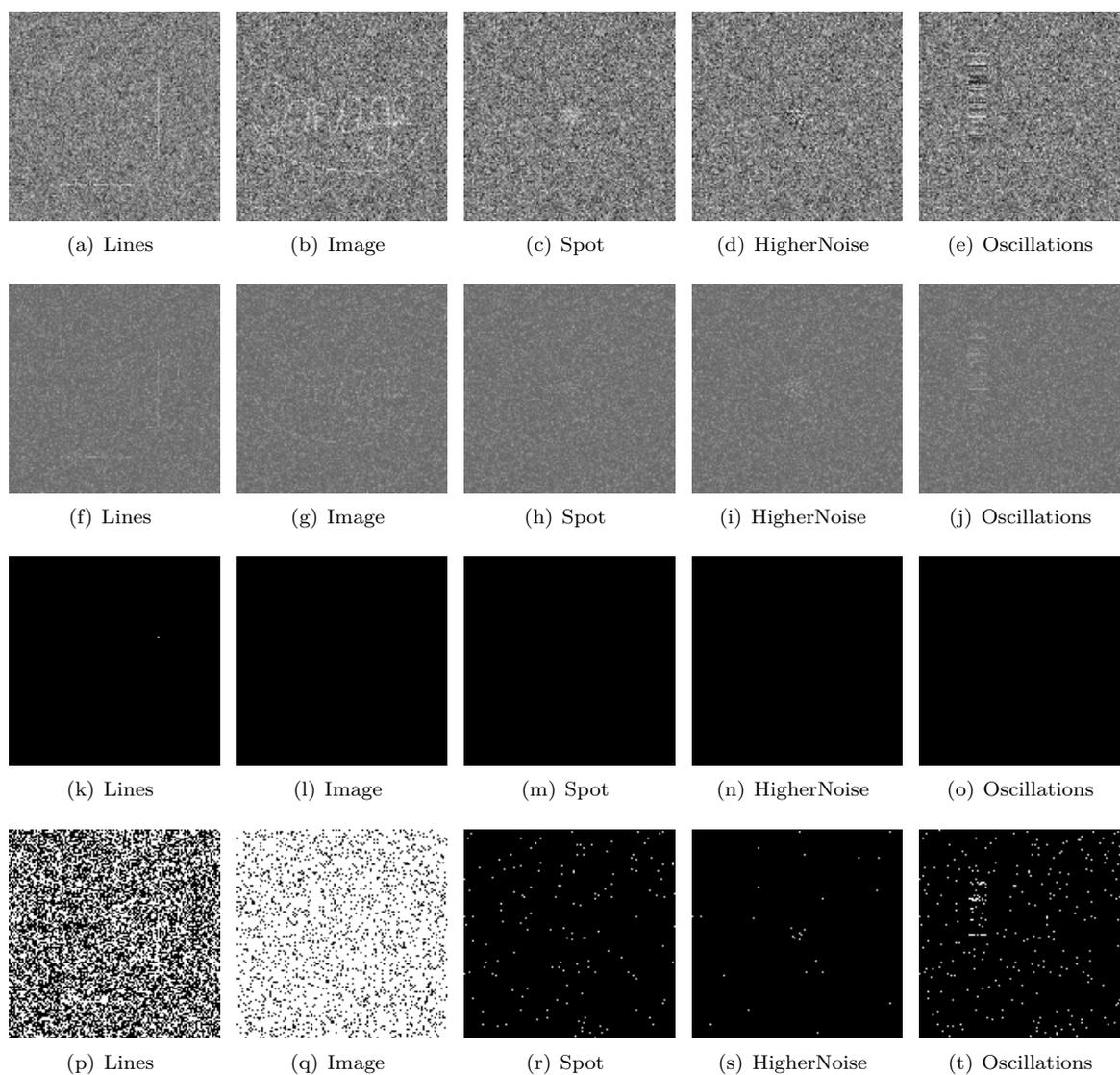


FIGURE 4.5 – Résultats de la méthode de seuillage simple. Première ligne : les images des défauts synthétiques à détecter. Deuxième ligne : les images, en niveaux de gris, issues de l'algorithme de détection. Les pixels brillants sont ceux considérés comme des défauts. Troisième ligne : le masque binaire issu de la détection, pour un taux de fausse alarme nul. Quatrième ligne : le masque binaire issu de la détection, pour un taux de bonne détection de 100%.

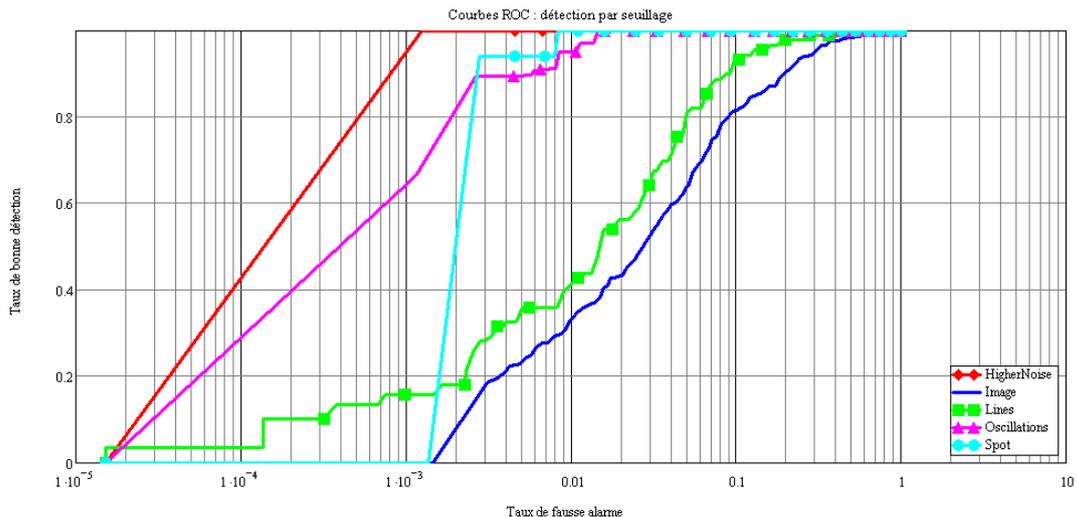


FIGURE 4.6 – Courbes ROC pour la méthode de détection par seuillage simple

- $V^I(i, j)$ un voisinage du pixel $I(i, j)$,
- $Moy(V^I(i, j))$ est la moyenne sur le voisinage $V^I(i, j)$,
- $MoyRef(i, j)$ la moyenne de référence au point (i, j) ,
- $Var(V^I(i, j))$ est la variance sur le voisinage $V^I(i, j)$,
- $VarRef(i, j)$ la variance de référence au point (i, j) ,
- $Skew(V^I(i, j))$ est le skewness (voir eq 4.4) sur le voisinage $V^I(i, j)$,
- $Kurt(V^I(i, j))$ est le kurtosis (voir eq 4.5) sur le voisinage $V^I(i, j)$.

4.4.1 Description de l'algorithme

Les algorithmes de détection à base de statistiques locales ont une structure commune représentée par la fig 4.7. On trouve en entrée :

- Image d'origine : image brute de taille $N \times M$ pixels contenant le ou les défauts à détecter
- d : paramètre contrôlant la taille du voisinage à considérer pour le calcul des statistiques locales. Le voisinage retenu sera de taille $(2d + 1) \times (2d + 1)$
- Image ref : image de même taille que l'image d'origine contenant une mesure statistique de référence au pixel (i, j) . Cette mesure de référence contient en général la statistique sur un voisinage étendu (de taille $(2D + 1) \times (2D + 1)$ avec $D \gg d$)

Les blocs de calculs intervenant dans ces méthodes peuvent être décrits comme suit :

- Extraction voisinage : extrait un voisinage centré autour du pixel d'indice (i, j)
- Calcul de la statistique : sur le voisinage extrait, on calcule une statistique (moyenne, variance, skewness ou kurtosis)
- Division par la statistique de référence : la valeur de la statistique précédemment calculée est divisée par la valeur de référence
- Écriture dans l'image résultat : le résultat du rapport précédent est affecté au pixel d'indice (i, j) dans l'image résultat

En sortie, les méthodes de détection fournissent une image dans laquelle les zones anormales apparaîtront en clair (valeurs élevées) et les zones assimilables au fond en sombre (valeurs aux alentours de 1).

	Seuillage	Moyenne Locale	Variance Locale	Skewness Local	Kurtosis Local
HigherNoise	0,999	0,984	1,000	0,842	0,822
Image	0,931	0,856	0,810	0,663	0,640
Lines	0,964	0,795	0,841	0,661	0,445
Oscillations	0,998	0,995	0,998	0,889	0,924
Spot	0,998	1,000	0,995	0,917	0,879

TABLE 4.2 – Aires sous courbe ROC des méthodes de détection par statistique locale

La détection par statistique locale est très proche de la méthode de détection définie dans le chapitre 2 : cette dernière méthode coïncide avec la détection par moyenne locale ou variance locale⁴.

Le calcul de la statistique de référence nécessite l'utilisation d'estimateurs robustes de la moyenne, de la variance, du skewness et du kurtosis. Les estimateurs robustes de la moyenne et de la variance ont été donnés dans les chapitres précédents (voir équation 2.11). Pour les estimations robustes du skewness et du kurtosis, nous avons utilisé respectivement les estimateurs de Bowley [11] et de Moors [71]. Ces estimateurs s'avèrent particulièrement robustes en présence d'échantillons atypiques [50].

$$\widehat{S}_X = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} \quad (4.7)$$

$$\widehat{K}_X = \frac{(O_7 - O_5) + (O_3 - O_1)}{O_6 - O_2} - 1.23 \quad (4.8)$$

où Q_i est le $i^{\text{ème}}$ quartile, et O_j est le $j^{\text{ème}}$ octile de l'échantillon X . Soit F la densité de probabilité de la loi de X , et F^{-1} sa fonction réciproque (fonction de répartition), on a :

$$E_i = F^{-1}\left(\frac{i}{4}\right)$$

$$O_j = F^{-1}\left(\frac{j}{8}\right)$$

4.4.2 Résultats sur défauts synthétiques

Les figures 4.8 et 4.9 montrent les résultats de la détection par statistique locale pour les cinq défauts synthétiques et pour quatre mesures statistiques : moyenne, variance, skewness et kurtosis. Les calculs ont été réalisés pour une taille de fenêtre d'analyse de 15×15 pixels, comme pour la méthode de détection décrite dans le chapitre 2.

Dans chaque cas, la statistique de référence a été calculée à l'aide des estimateurs robustes sur toute l'image, tandis que la statistique locale est calculée sur un voisinage de 15×15 pixels autour du pixel traité.

Pour chacune des quatre mesures, on représente d'abord l'image en niveaux de gris (figure 4.8), en sortie des algorithmes de détection, où les zones de défauts sont en clair. On a également représenté les images seuillées correspondant à un taux de fausse alarme nul (figure 4.9).

Sur les images synthétiques, les résultats obtenus sont globalement mauvais. De manière ponctuelle,

4. Dans ce chapitre, comme les images traitées sont de petite taille, nous avons choisi de normaliser chaque statistique par une référence calculée sur toute l'image. La statistique est alors normalisée par division par cette référence. Dans le chapitre 2, la statistique de référence a été estimée sur l'ensemble des valeurs prises par cette statistique, puis nous avons réalisé une normalisation par centrage-réduction. Cette procédure de normalisation différente ne change pas les résultats des courbes ROC (conservation de la relation d'ordre).

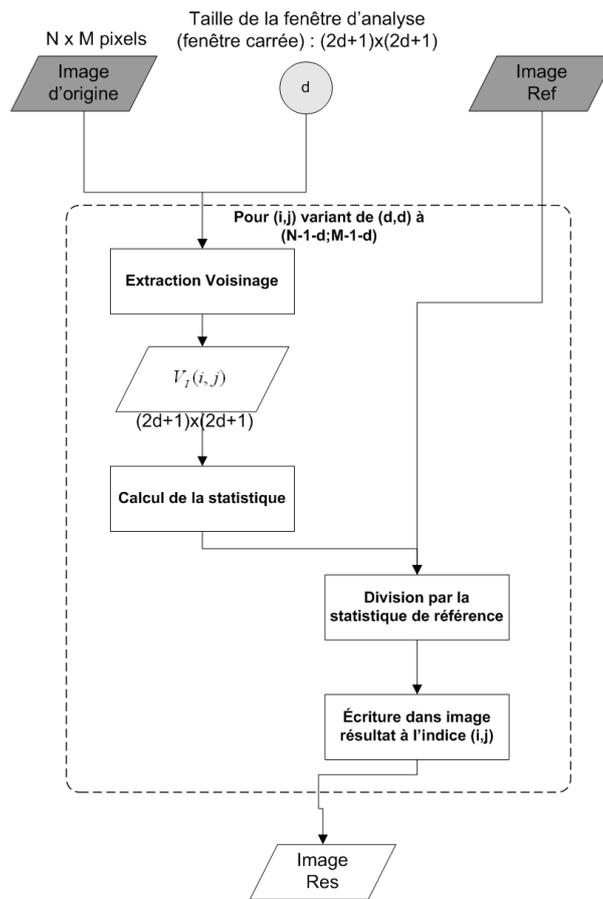
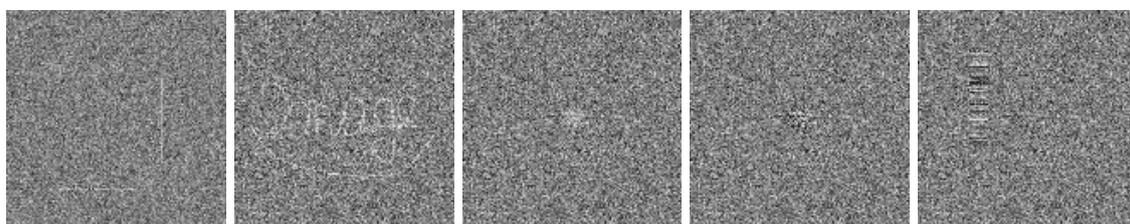
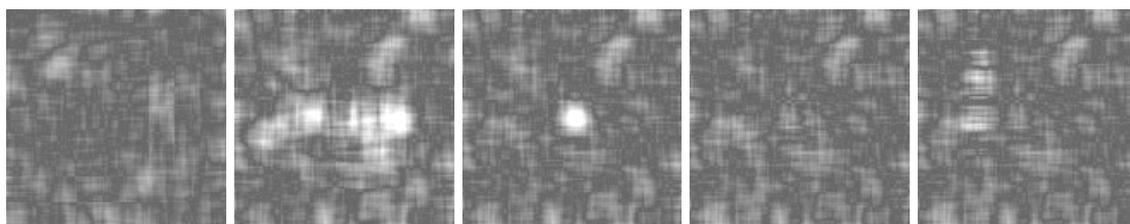


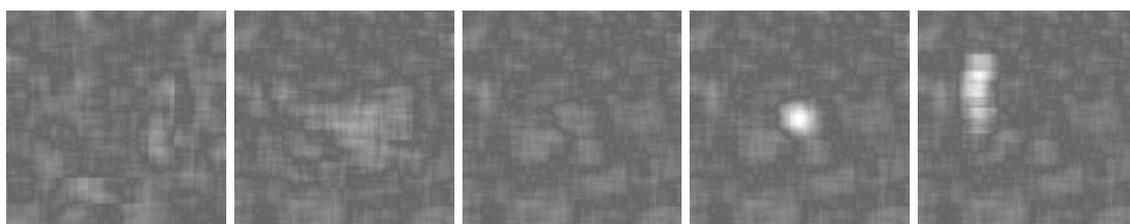
FIGURE 4.7 – Schéma descriptif des méthodes de détection par statistique locale



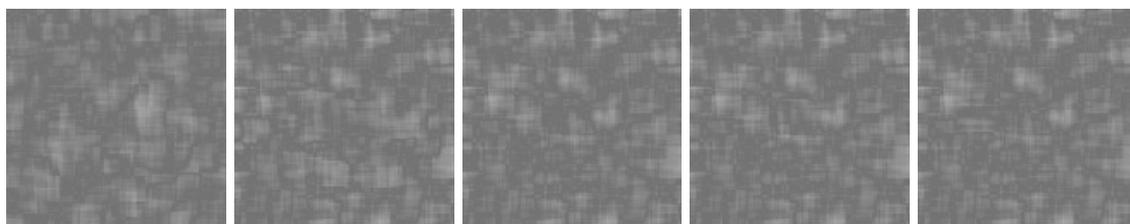
(a) Images d'origine



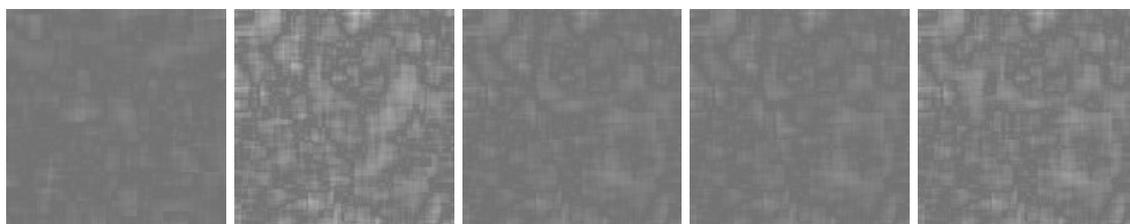
(b) Images en sortie du détecteur par moyenne locale



(c) Images en sortie du détecteur par variance locale



(d) Images en sortie du détecteur par skewness local



(e) Images en sortie du détecteur par kurtosis local

FIGURE 4.8 – Images résultats en sortie des algorithmes de détection par statistique locale. Les défauts Spot, HigherNoise et Oscillations sont bien détectés par au moins une méthode.

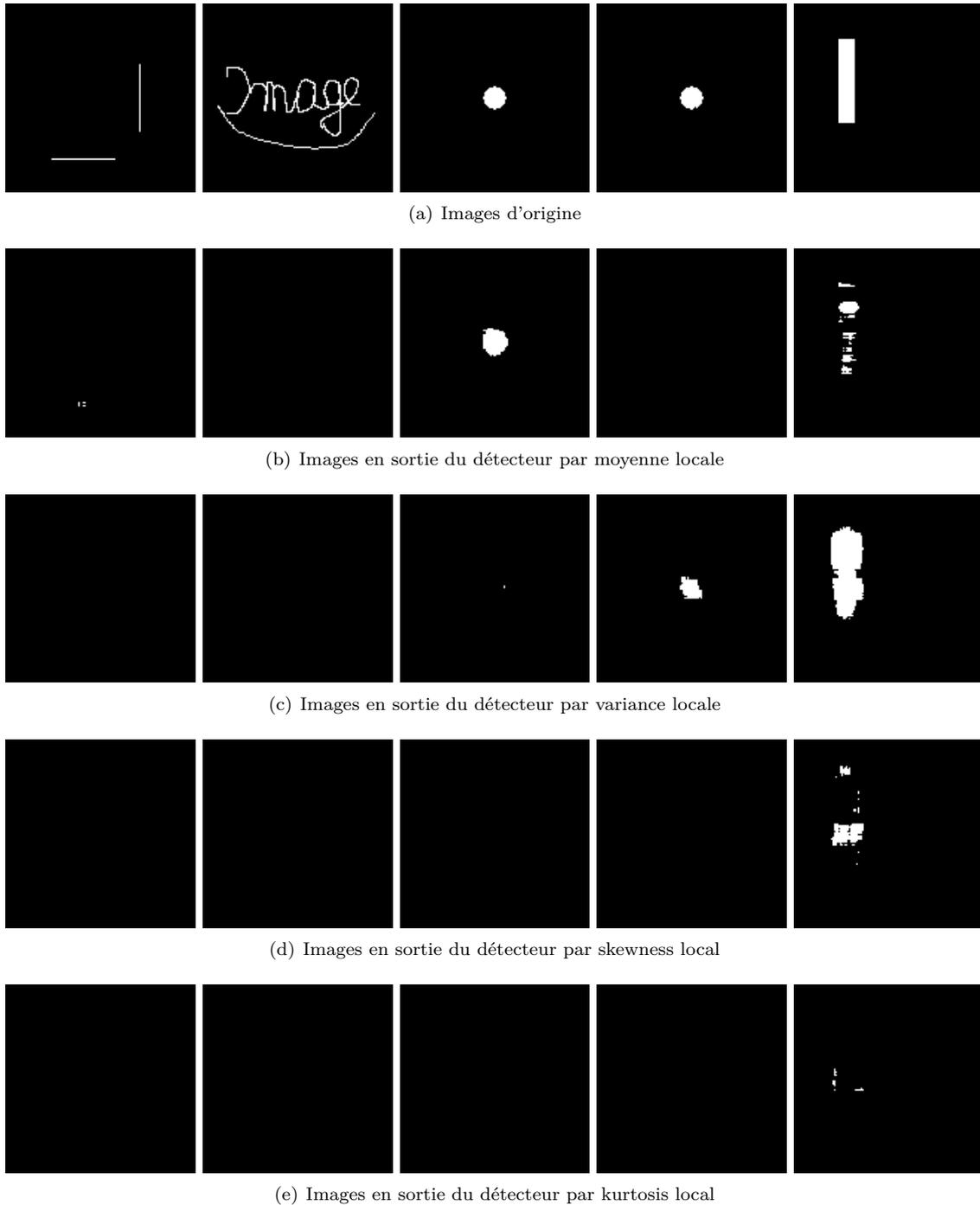
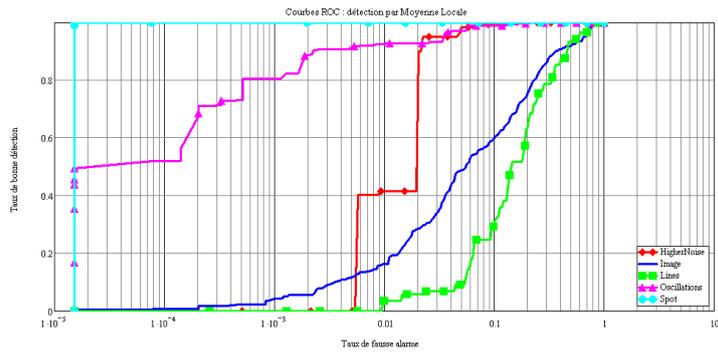
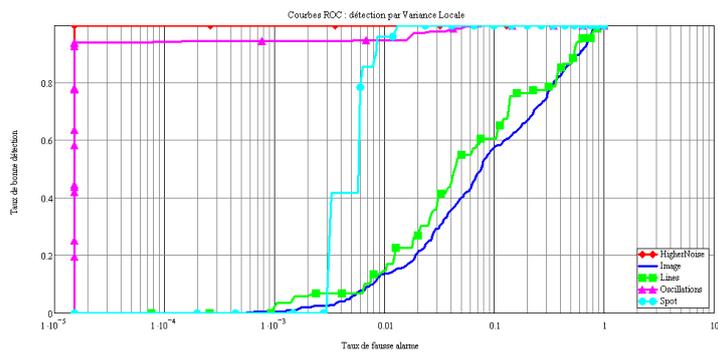


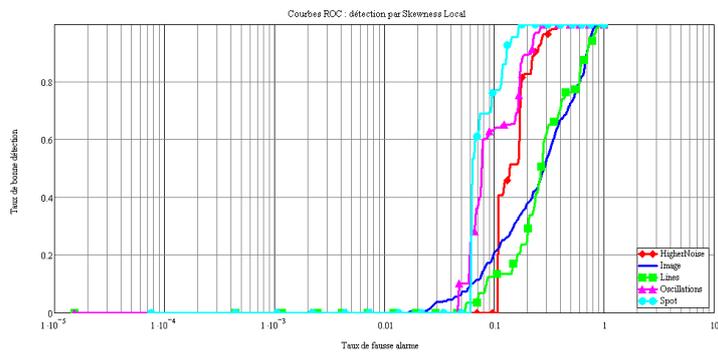
FIGURE 4.9 – Images des détecteurs par statistique locale, seuillées pour un taux de fausse alarme nul



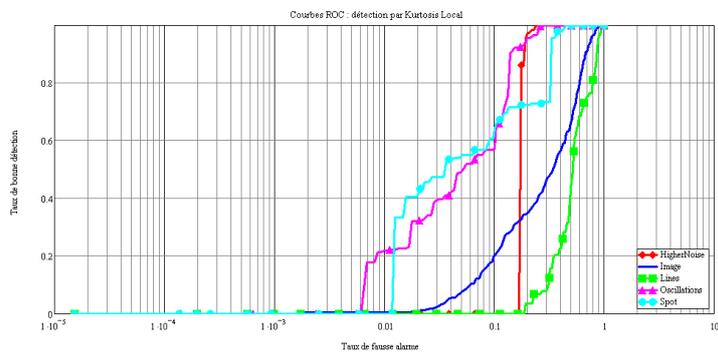
(a) Courbes ROC de la détection par moyenne locale



(b) Courbes ROC de la détection par variance locale



(c) Courbes ROC de la détection par skewness local



(d) Courbes ROC de la détection par kurtosis local

FIGURE 4.10 – Courbes ROC des méthodes de détection par statistique locale. La méthode de moyenne locale donne des résultats parfaits sur les défauts Spot et Oscillations. Le défaut HigherNoise est lui parfaitement détecté par la méthode de la variance locale. Aucune des méthodes proposées ici ne permet de détecter efficacement les défauts Lines et Image.

certaines méthodes simples, comme la variance sur les défauts HigherNoise ou Oscillations ou la moyenne sur le défaut Spot, peuvent se révéler efficaces. Les méthodes à base de statistiques d'ordre supérieur, à savoir le skewness et le kurtosis, sont, par contre, inadaptées. Ce résultat, qui peut paraître surprenant étant donné les échos favorables dont font l'objet les SOS pour la détection sur des images sonar notamment, s'explique par les deux constatations suivantes :

- le contexte différent : les SOS ont été appliquées de manière régulière pour détecter des défauts noyés dans du bruit blanc gaussien. Dans notre cas, les défauts synthétiques sont bien noyés dans un tel bruit, mais les pixels de défauts sont eux mêmes répartis sur une gaussienne (ce sont des défauts additifs rajoutés au bruit de fond). Le skewness et le kurtosis, efficaces pour détecter les non-gaussianités, deviennent donc inefficaces dans notre cas.
- La variance des estimateurs : en théorie, on pourrait néanmoins détecter des variations des SOS au niveau des contours des défauts. Cependant, en pratique, du fait de la variance élevée des estimateurs du skewness et du kurtosis, cette variation est indécélable. Il faut en effet faire un compromis sur la taille de la fenêtre de calcul des statistiques : une plus grande taille donnera une variance plus faible des estimateurs, mais rendra l'impact des défauts présents dans la fenêtre négligeable. On peut montrer que, dans notre cas, ce compromis conduit inévitablement à une mauvaise détection des défauts.

4.4.3 Conclusions sur les méthodes à base de statistiques locales

Les méthodes à base de statistiques locales, notamment celles s'appuyant sur un calcul de moyenne ou de variance locale, peuvent ponctuellement se révéler très efficaces. En ce sens, elles constituent des méthodes de base pour la détection de défaut dans notre contexte. En particulier, ces deux méthodes permettent d'exploiter deux caractéristiques fréquentes qu'on retrouve dans la majorité des défauts contenus dans les images de radiologie numérique :

- la perte de sensibilité locale : cette caractéristique se traduit par une variation localisée de la valeur moyenne de l'image. Elle est naturellement détectée par une méthode discriminante sur la valeur de cette statistique locale. La difficulté est en général le calcul de la statistique de référence.
- la pixellisation locale : cette caractéristique entraîne une augmentation localisée de la variance du bruit. Ces défauts sont causés par une contrainte de fabrication des détecteurs. Comme pour la perte de sensibilité, on peut détecter assez aisément ce type de défauts par un calcul de la variance locale, avec la même difficulté à gérer pour le calcul de la valeur de référence de cette variance.

Dans la suite, afin de mesurer l'effet positif des méthodes proposées sur les performances de détection, nous tracerons pour chaque défaut synthétique la meilleure courbe obtenue à l'aide des trois méthodes de détection de base que sont le seuillage direct et les méthodes de moyenne et de variance locale. Ainsi, pour les cinq défauts étudiés, on aura :

- défaut HigherNoise : méthode de variance locale ($AUC \approx 1$),
- défaut Image : méthode de seuillage direct ($AUC = 0,93$),
- défaut Lines : méthode de seuillage direct ($AUC = 0,964$),
- défaut Oscillations : méthode de variance locale ($AUC \approx 1$),
- défaut Spot : méthode de moyenne locale ($AUC = 1$).

4.5 Détection par décomposition en ondelettes

4.5.1 Principe

La méthode de détection par ondelette a pour but de détecter des structures de type lignes (hautes fréquences horizontales, verticales ou diagonales) ou des taches (basses fréquences) en utilisant une analyse multi-résolution s'appuyant sur la transformée en ondelettes discrètes (TOD). Le but de cette méthode est de se rapprocher d'une base de représentation optimale pour les défauts courants, les rendant alors plus facilement détectables, comme expliqué en section 4.2.1.3. De plus, la TOD fournit un schéma de

traitement des images permettant une analyse de celles-ci à plusieurs échelles, ce qui est particulièrement intéressant pour notre problématique. En effet, certains défauts comme le défaut Image sont visibles uniquement lorsqu'ils sont observés à grande échelle. Cette méthode possède en outre des avantages pratiques évidents : simplicité d'implémentation de la TOD par banc de filtres à RIF, rapidité d'exécution et parcimonie de la représentation à une échelle donnée (le nombre de pixels à traiter reste constant tout au long du calcul). De plus, la TOD étant une opération inversible, il est très aisé de travailler dans l'espace transformé, puis de revenir dans l'espace réel pour visualiser les résultats.

4.5.2 Description de l'algorithme

Le principe général de l'algorithme de détection par ondelettes est décrit par la fig 4.11. En entrée de l'algorithme, on trouve :

- l'image d'origine Im ,
- l'ondelette choisie w : à chaque type d'ondelette va correspondre une paire de filtres de décomposition (ou filtre d'analyse) basse fréquence et haute fréquence, ainsi qu'une paire de filtres de reconstruction (ou filtre de synthèse). Ces deux paires sont identiques si on choisit une famille d'ondelettes orthogonales. Dans notre étude, nous avons choisi, sauf indication contraire, une ondelette de la famille Battle-Lemarié d'ordre 2. Ce choix est empirique. Le désavantage de cette famille est que les filtres de décomposition et de reconstruction correspondants ne sont pas à support finis. Néanmoins, le temps de calcul n'étant pas une contrainte dans un premier temps, nous avons choisi cette famille car nous avons constatés dans notre cas qu'elle permet d'obtenir des résultats de reconstruction plus fidèles aux défauts d'origine.
- Un paramètre de détection nd : ce paramètre détermine un seuil de détection en nombre d'écart-types des pixels dans chaque sous-bande (voir fonction Peeling). Pour les résultats exposés, on a pris systématiquement $nd = 5$.

L'image d'origine est décomposée jusqu'à une échelle qui sera déterminée par sa taille. Si l'image est de taille $2^n \times 2^n$:

$$EchelleMax = N = \log_2(2^n) = n$$

En pratique, on se restreint à la décomposition à $EchelleMax - 2$ pour garder suffisamment de pixels afin de pouvoir extraire dans chaque bande des statistiques fiables.

Les fonctions intervenant dans ce traitement peuvent être résumées comme suit :

- $VHDecomp(Im, w, K)$ réalise une décomposition en ondelettes à l'échelle $K + 1$ en partant de la décomposition à l'échelle K ,
- $Peeling$: réalise une détection par sous-bandes dans la base d'ondelettes et extrait «l'excédent» pour chaque coefficient d'ondelette détecté,
- $DecompStorage$: stocke les résultats de la fonction $Peeling$ à chaque échelle en vue d'une reconstruction.

On obtient alors en sortie $EchelleMax - 2$ images de taille $2^n \times 2^n$ contenant les images de détection à chaque échelle : tous les pixels de ces images sont à zéro, sauf ceux qui ont été détectés comme anormaux par la fonction $Peeling$.

D'une manière générale, à chaque échelle k , on a :

$$D_k = VHDecomp(D_{k-1}, w, k) \quad (4.9)$$

$$P_k = Peeling(D_k, k) \quad (4.10)$$

$$R_k = D_k - P_k \quad (4.11)$$

4.5.2.1 Fonction VHDecomp

La fonction $VHDecomp$ réalise une décomposition en ondelettes à l'échelle supérieure (échelle plus grossière) à partir d'une décomposition donnée. Le découpage réalisé par $VHDecomp$ est représenté sur la fig 4.12. L'intérêt d'une telle décomposition, par rapport à la décomposition classique DWT (Discrete

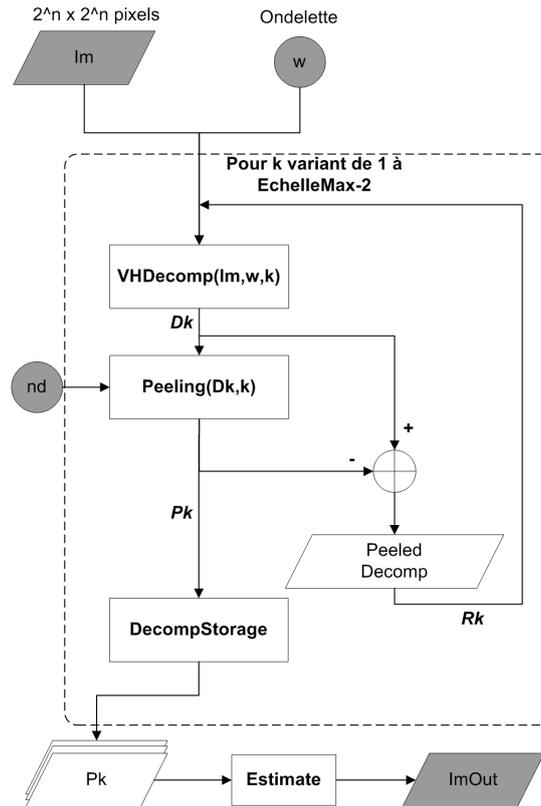


FIGURE 4.11 – Schéma de principe : détection par ondelettes

Wavelet Transform) repose dans l'analyse plus fine des cadrans hautes-fréquences (cadrant supérieur droit : hautes fréquences verticales et cadrant inférieur gauche : hautes fréquence horizontales) alors que la décomposition DWT classique se contente de décomposer le cadrant basse-fréquence. De plus, les cadrans des hautes fréquences verticales et horizontales ne sont filtrés qu'avec un filtre de synthèse à une dimension (passe haut ligne ou passe haut colonne), contrairement à la transformée en ondelette classique qui utilise à chaque décomposition un filtrage séparable suivant les deux dimensions de l'image.

Une décomposition VH à l'échelle n génère ainsi k sous bandes avec :

$$k = \begin{cases} 4 & \text{si } n = 1 \\ 2 + 2^n - 1 & \text{si } n > 1 \end{cases} \quad (4.12)$$

La fig 4.12 met bien en évidence l'intérêt de la décomposition VH : on voit que lorsqu'on décompose à des échelles de plus en plus élevées, le signal (ici, il s'agit d'un trait vertical et d'un trait horizontal) se concentre sur un faible nombre de coefficients, tout en conservant globalement son énergie. Dans l'exemple présenté, les deux traits sont concentrés en quelques points dans une seule bande, les autres bandes ne comportant que du bruit, à l'exception de la bande BF générée à l'échelle 1. Nous avons donc :

- pour le bruit : dans chaque bande, la puissance de bruit est inchangée puisque le bruit est blanc
- pour le signal : dans les bandes HF, la puissance du signal se concentre sur très peu de coefficients

On a ainsi considérablement augmenté la détectabilité du défaut.

4.5.2.2 Fonction Peeling

La fonction de Peeling est chargée de la détection des pixels défectueux dans toutes les sous-bandes d'une décomposition, ainsi que de l'extraction du signal de défaut. Cette fonction réalise un seuillage dans

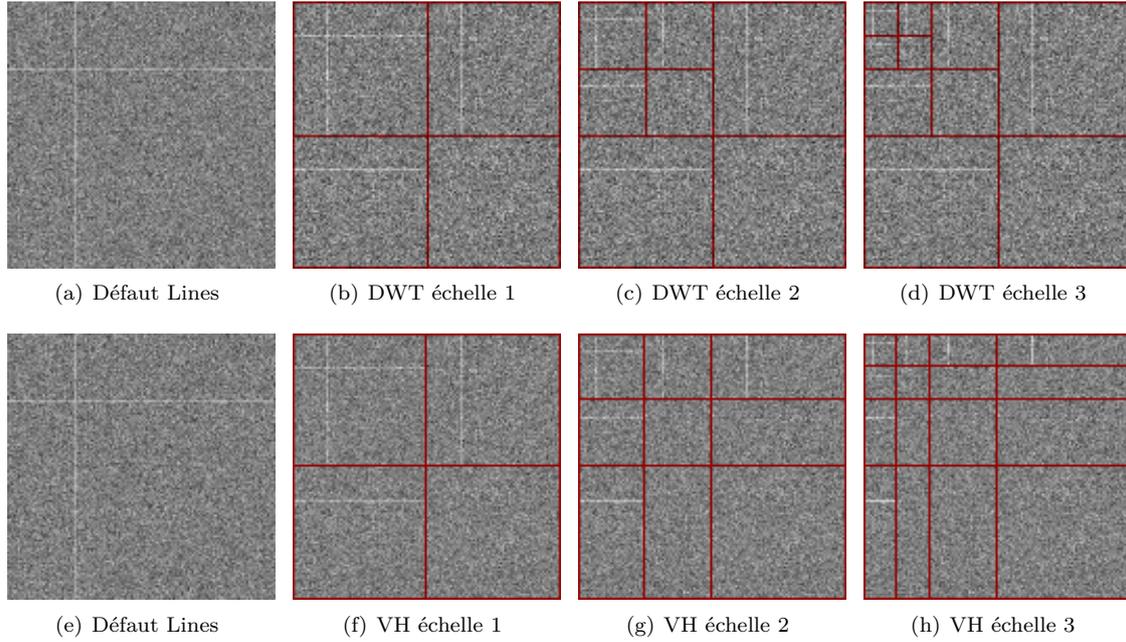


FIGURE 4.12 – Comparaison de deux décompositions en ondelettes : décomposition classique (DWT) et décomposition VH

chaque sous-bande de la décomposition. Soit $I(i, j)$ l'image en entrée de la fonction de Peeling, et $O(i, j)$ l'image en sortie. On appelle $\mathcal{S} = \{S_k\}$ la partition de $I(i, j)$ définissant la décomposition en ondelette de $I(i, j)$.

On a pour un pixel de coordonnées (i, j) appartenant à S_k :

$$O(i, j) = \begin{cases} I(i, j) - (m_k + nd\sigma_k) & \text{si } |I(i, j) - m_k| \geq nd \cdot \sigma_k \\ 0 & \text{sinon} \end{cases} \quad (4.13)$$

m_k étant la moyenne de la sous-bande S_k , et σ_k son écart-type.

On récupère ainsi dans $O(i, j)$ une image résiduelle correspondant à l'excès présent sur les coefficients d'ondelette anormalement élevés.

4.5.2.3 Fonction Estimate

La fonction Estimate permet de synthétiser une image dans la base réelle (par opposition à la base d'ondelettes) à partir des décomposition successives pelées produites par la fonction Peeling aux différents niveaux de décomposition analysés. Pour cela, la fonction utilise une transformation VH inverse, $VHRecomp$, qui réalise l'opération inverse de $VHDecomp$. A partir d'une décomposition à l'échelle n , elle calcule la décomposition à l'échelle $n - 1$. Cette opération est effectuée de manière similaire à l'opération de transformation en ondelette inverse, à l'aide des filtres de synthèse. L'image finale $ImOut$ est alors calculée par la formule de récurrence suivante :

$$\begin{cases} x_N = P_N \\ x_k = VHRecomp(x_k) + P_k, k \in [0; N - 1] \end{cases} \quad (4.14)$$

on obtient $ImOut = x_0$ en prenant $\forall(i, j) P_0(i, j) = 0$.

Cette opération consiste à rassembler les détections effectuées aux différentes échelles. Notons que dans cette optique, il était important d'effectuer une opération de pelage, afin que la partie du signal

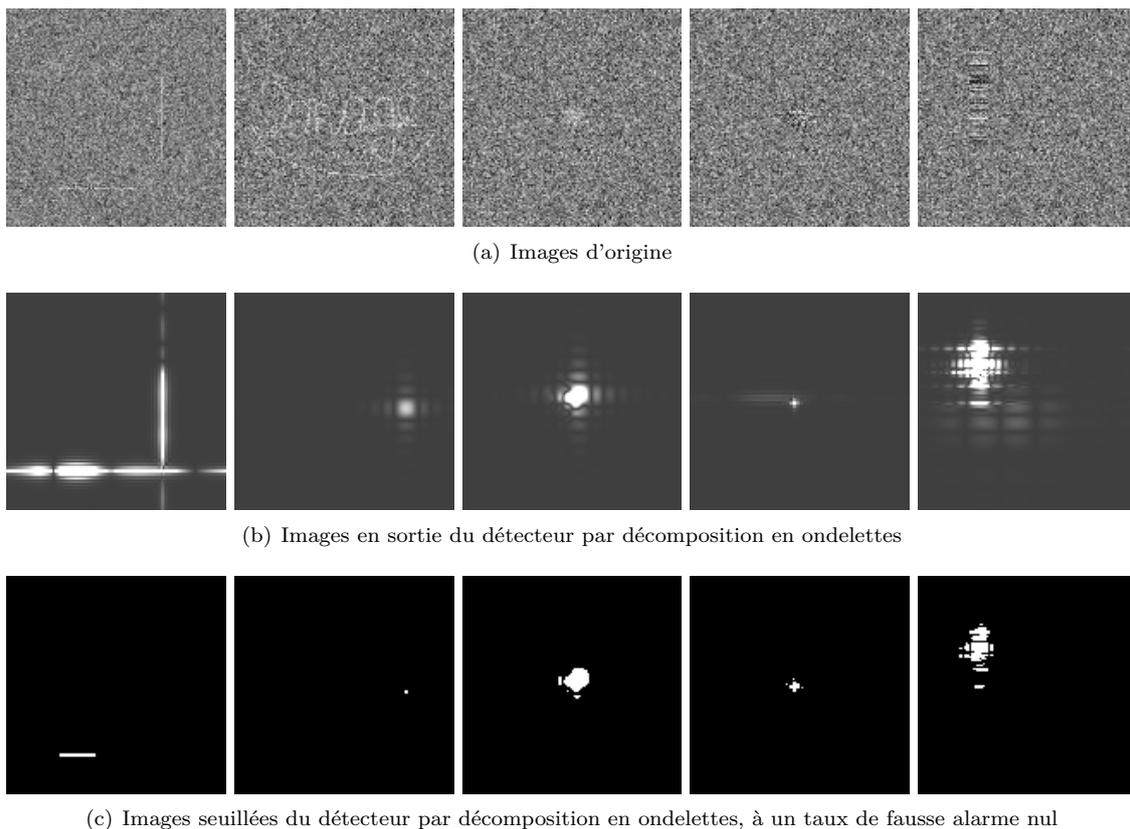


FIGURE 4.13 – Résultats et images seuillées des détecteurs par décomposition en ondelettes

détecté à une échelle ne se retrouve pas détecté à l'échelle suivante. De cette manière, le signal reconstitué dans cette étape d'estimation ne prend pas en compte d'information redondante.

4.5.3 Résultats sur défauts synthétiques

La fig 4.13 présente les résultats de la méthode de détection par décomposition en ondelettes sur les défauts synthétiques. On trace également les courbes ROC correspondantes (figure 4.14). Sur ces images 128×128 pixels, on a effectué une décomposition VH allant jusqu'à l'échelle 5. Dans chaque sous-bande de la décomposition en ondelettes, on a effectué une détection avec le paramètre $nd = 4, 5$. La valeur de ce paramètre est la seule variable à ajuster de cette méthode. Elle permet de contrôler le nombre de fausses alarmes générées. Typiquement, on peut faire varier cette valeur de 4 à des valeurs élevées. On utilisera des valeurs élevées de ce paramètre pour isoler uniquement les défauts de grande amplitude. Dans notre application, nous voulons détecter toute structure ressortissant du bruit de fond, d'où le choix d'un paramètre relativement bas.

On observe de bons résultats sur le défaut Lines par rapport aux méthodes précédentes. La courbe ROC de la meilleure méthode de détection précédente pour ce défaut et assez nettement en dessous de celle de la méthode par décomposition en ondelettes. Pour les défauts HigherNoise et Spot, la méthode en ondelettes est équivalente aux meilleures méthodes statistiques : ces deux défauts sont très bien détectés.

Pour le défaut image, la méthode en ondelettes est légèrement en dessous de la méthode de base par seuillage direct, qui est jusqu'à maintenant la meilleure méthode présentée pour ce défaut. Ce défaut présente des structures hautes fréquences mais leur nature curviligne fait que celles-ci ne sont pas adaptées aux bases d'ondelettes. En effet, les ondelettes sont très efficaces sur les structures horizontales, diagonales ou verticales, mais perdent rapidement de leur pouvoir de détection dans le cas d'orientations plus variées. Une voie possible à explorer est l'utilisation par exemple d'une transformée en curvelets qui s'adaptent

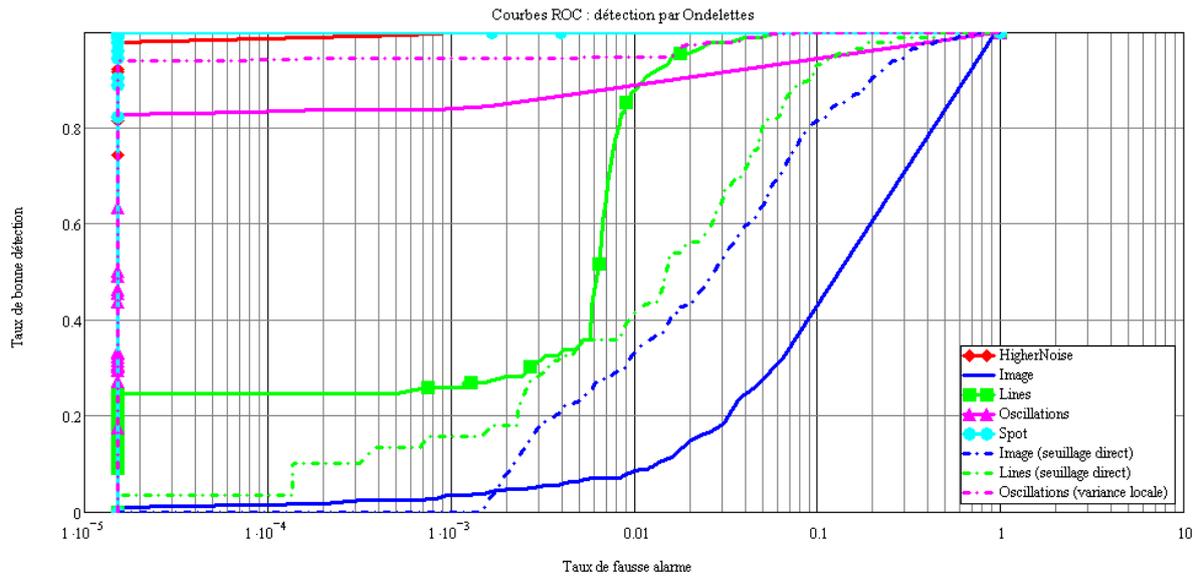


FIGURE 4.14 – Courbe ROC de la méthode de détection par ondelettes. Pour les défauts Oscillations, Lines et Image, les courbes de la meilleure méthode statistique ont été également représentées.

	HigherNoise	Image	Lines	Oscillations	Spot
AUC max	1,000	0,931	0,964	0,998	1,000
des méthodes statistiques	Variance Locale	Seuillage	Seuillage	Variance Locale	Moyenne Locale
AUC méthode ondelettes	0,999	0,649	0,993	0,940	1,000

TABLE 4.3 – Aires sous courbe ROC de la méthode de détection par ondelettes

aux géométries locales des objets.

Le tableau 4.3 donne les AUC de la méthode de détection par ondelettes comparées à celles obtenues par la meilleure méthode statistique. On constate l'amélioration sur le défaut Lines.

4.5.4 Conclusion sur la méthode de détection par décomposition en ondelettes

La méthode proposée ici détecte relativement bien la variété de défauts synthétiques à laquelle elle a été confrontée, à l'exception notable du défaut Image. Cette méthode s'appuie sur une transformée en ondelettes dont la base de décomposition a été spécialement adaptée pour détecter les lignes verticales ou horizontales, de largeur variable. L'avantage de cette méthode est sa polyvalence et son aspect multi-échelle :

- polyvalence : un large spectre de défauts peuvent être détectés par cette méthode : les défauts présentant de hautes fréquences horizontales et verticales comme les défauts présentant des structures plus basse fréquence.
- multi-échelle : l'utilisation de la transformée en ondelette discrète permet un parcours de toutes les échelles de détection possible, ce qui est difficile à faire avec les méthodes statistiques⁵. L'analyse multi-échelle s'effectue pour un coût faible en temps de calcul.

5. Ceci est fait en particulier en adaptant la taille de la fenêtre d'analyse des méthodes statistiques. Néanmoins, ce procédé a un coût important en terme de complexité et de temps de calcul

Les principaux défauts de la méthode sont ses faibles performances pour la détection d'orientations quelconques (comme on peut le noter sur les résultats de détection sur le défaut Image) et les problèmes de détection dus à la décimation : en effet, la transformée en ondelette n'est pas invariante par translation de l'image de départ, et en particulier, certains défauts qui, dans une configuration donnée, sont très bien détectables, peuvent devenir quasi-indétectables dans la base d'ondelettes après translation de ces défauts. Cet effet est lié à la décimation qui est effectuée dans la transformation en ondelette discrète, et est important dans le cas où on utilise une ondelette à support très compact de type Haar. Dans notre cas, on a limité l'effet en utilisant l'ondelette de Battle-Lemarié, qui est à large support.

Globalement, la méthode par transformation en ondelette est efficace sur une large gamme de défauts, et améliore globalement les résultats des méthodes à base de statistique locale. Ces dernières restent néanmoins les plus efficaces dans le cas de défaut comme HigherNoise qui est parfaitement adapté, par sa taille et sa nature, à ces dernières méthodes.

4.6 Détection par filtrage directionnel

4.6.1 Principe

La méthode précédente a montré qu'on peut aisément détecter des défauts de type lignes horizontales, verticales ou diagonales. De même, les taches de type Spot peuvent être détectées en basse fréquence par ces mêmes méthodes basées ondelettes. La nouvelle méthode présentée dans cette section vise à détecter les structures filiformes et curvilignes du type de celles présentes dans le défaut Image. Pour cela on va chercher à décomposer chaque pixel de l'image à analyser sur un nombre de fréquences spatiales limité. Chaque fréquence correspond à une structure ligne d'orientation donnée. En pratique, on va filtrer l'image par un banc de filtres directionnels.

La figure 4.15 présente le principe de ce filtrage. La première ligne contient les fonctions de transfert (domaine fréquentiel, fréquence (0;0) au centre de l'image) des filtres utilisés. Ici on a choisi d'analyser huit directions différentes. La troisième ligne contient les résultats des filtrages du défaut par chacun de ces huit filtres. On constate que les lignes horizontale et verticale du défaut apparaissent chacune nettement sur une seule image filtrée. A l'aide des ces images filtrées, on va déterminer l'image de détection en combinant pixel à pixel une ou plusieurs images filtrées.

Cette méthode revient à faire des moyennes locales selon plusieurs directions dans l'image d'origine, et à chercher si une ou plusieurs directions analysées présente une corrélation suffisamment forte. Cette méthode requiert donc que le bruit de fond de l'image analysée soit blanc, ou du moins qu'il ne présente pas de corrélations spatiales linéaires fortes.

4.6.2 Description de l'algorithme

Le principe général de l'algorithme est décrit par le schéma de la figure 4.16. On distingue principalement trois blocs de traitement, qui seront décrits ci-après : un module de calcul des filtres, un module de filtrage simple, et un module de sélection pixel à pixel des filtres retenus et à partir desquels l'image résultat sera calculée.

4.6.2.1 Calcul des filtres

Si on choisit de réaliser l'analyse selon n orientations, on calcule un banc de filtres F_k , $k = 1 \dots n$. Dans la suite, les filtres sont exprimés à l'aide des coordonnées polaires (r, θ) dans Fourier. On note \widehat{F}_k la transformée de Fourier de F_k . L'origine du repère pour les coordonnées polaires est la fréquence nulle dans la représentation de Fourier. On a donc $\widehat{F}_k(u, v) = \widehat{F}_k(r \cos \theta, r \sin \theta)$. Dans notre application on a pris les F_k tels que :

$$\widehat{F}_k(r, \theta) = \begin{cases} \cos^2\left(\frac{\pi}{2}(\theta - \theta_k)\right) & \text{si } \min(|\theta - \theta_k - \text{sgn}(\theta)\pi|, |\theta - \theta_k|) \leq \frac{\pi}{n} \\ 0 & \text{sinon} \end{cases} \quad (4.15)$$

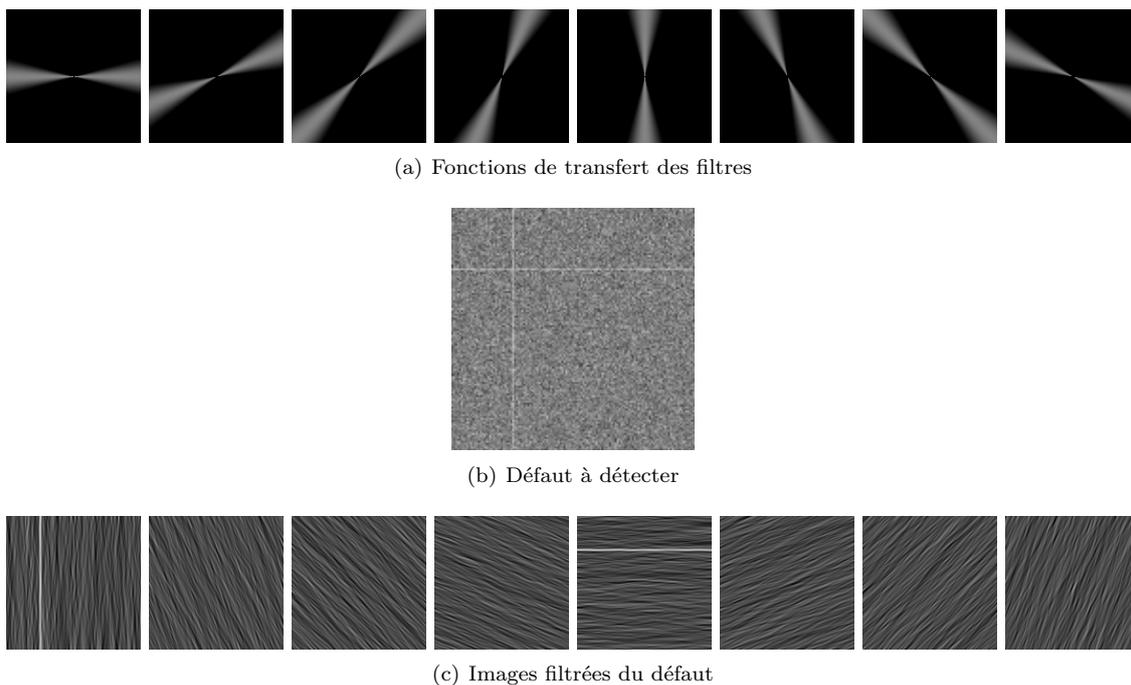


FIGURE 4.15 – Filtrage d'un défaut par banc de filtres directionnels

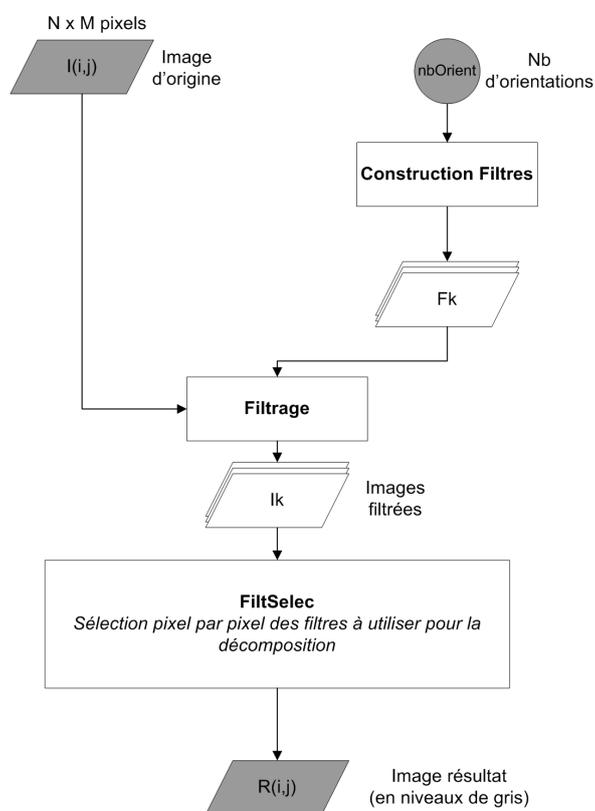


FIGURE 4.16 – Schéma de principe : détection par filtrage fréquentiel

avec $\theta_k = \frac{\pi}{n}(k-1)$. Pour des raisons pratiques (éliminer la composante constante des images), on prend de plus $\widehat{F}_k(0,0) = 0$.

Un exemple de tels filtres est présenté dans la figure 4.15(a). On remarque de plus que, pour les filtres définis, on a :

$$\forall (i, j) \neq (0, 0), \sum_{k=1}^n \widehat{F}_k(i, j) = 1 \quad (4.16)$$

Ces filtres ont un profil de type sinusoïdal centré sur chacune des orientations analysées et sont à supports fréquentiels recouvrants.

4.6.2.2 Filtrage

Une fois les filtres calculés, la méthode applique chacun de ces filtres à l'image d'origine $I(i, j)$ pour générer des images filtrées $I_k(i, j)$. Le bloc de filtrage réalise cette opération simple, générant à partir de $I(i, j)$ et des filtres $F_k(i, j)$ les images filtrées $V_k(i, j)$ avec :

$$\forall k \in [1; n], V_k(i, j) = I(i, j) * F_k(i, j) \quad (4.17)$$

$$= TF^{-1} \left\{ \widehat{I}(u, v) \cdot \widehat{F}_k(u, v) \right\} \quad (4.18)$$

En pratique, l'opération de filtrage est réalisée à l'aide de l'équation 4.18.

4.6.2.3 FiltSelec : sélection des filtres à utiliser

Enfin, une dernière étape consiste à déterminer quels filtres seront conservés pour générer l'image résultat. Dans cette étape, on opère pixel à pixel.

Pour chaque pixel (i, j) , on a donc un vecteur

$$V(i, j) = \begin{pmatrix} V_1(i, j) \\ V_2(i, j) \\ \vdots \\ V_n(i, j) \end{pmatrix}$$

Ce vecteur contient deux composantes, $V(i, j)^+$ et $V(i, j)^-$ contenant respectivement uniquement les valeurs positives de V_i, j et uniquement ses valeurs négatives. Ces deux composantes correspondent à la présence d'orientations présentant des structures corrélées respectivement claires ou sombres par rapport à la moyenne de l'image. Afin de former l'image de détection finale, il y a deux manières possibles de traiter cette information :

- soit on se place dans le cas général où on dispose de peu d'*a priori* sur les défauts à détecter, et on cherche donc à détecter dans le même temps les structures claires et sombres. Cette approche est la plus générale : elle permet de ne pas distinguer deux classes de défauts.
- Soit on cherche uniquement à détecter les structures claires ou sombres, et on ne traite qu'une composante. Cette méthode permet de réduire l'influence du bruit dans le processus de détection, au prix d'une perte de généralité de la méthode.

Dans notre cas, nous avons choisi la première méthode, afin de nous placer dans le cas le plus général possible. Néanmoins la deuxième méthode a été testée, et donne effectivement de meilleurs résultats sur les défauts ne présentant que des structures plus claires que le bruit de fond, comme c'est le cas pour tous les défauts synthétiques étudiés, sauf le défaut Oscillations.

On rappelle qu'on a pris $\widehat{F}_k(0,0) = 0$ donc les images filtrées $\{I_k\}$ sont de moyenne nulle. Les éléments dans chaque vecteur $V(i, j)^+$ et $V(i, j)^-$ sont alors regroupés et triés par ordre décroissant de leur valeur absolue.

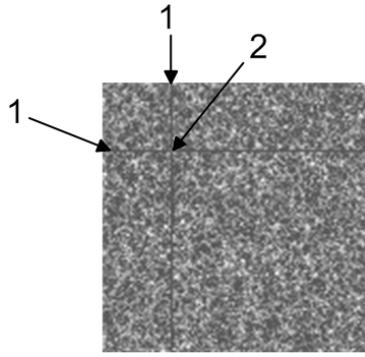


FIGURE 4.17 – Nombre de filtres sélectionnées sur défaut de type Lines

On obtient le vecteur $V^{tri}(i, j)$:

$$V^{tri}(i, j) = \begin{pmatrix} V_{k_1}(i, j) \\ V_{k_2}(i, j) \\ \vdots \\ V_{k_n}(i, j) \end{pmatrix} \text{ avec } V_{k_1}(i, j) \geq V_{k_2}(i, j) \geq \dots \geq V_{k_n}(i, j) \quad (4.19)$$

Après cette opération de classement, par souci de simplicité, on conserve la numérotation des filtres de 1 à n , $F_1(i, j)$ étant alors le filtre à la réponse maximale au pixel (i, j) considéré.

La dernière étape consiste à déterminer quelles orientations garder dans l'image finale. On détermine ainsi α^* tel que⁶ :

$$\alpha^* = \arg \max_{\alpha} \left\{ \frac{\sum_{k=1}^{\alpha} |V_k^{tri}(i, j)|}{\sqrt{\sum_{k=1}^{\alpha} \|F_k\|^2}} \right\} \quad (4.21)$$

Cette étape revient à choisir si la puissance récupérée par filtrage au point (i, j) est plus importante lorsqu'on utilise un filtre, ou la somme de plusieurs filtres. La normalisation par l'énergie du filtre ou de la combinaison de filtres utilisés est impérative pour pouvoir comparer de manière cohérente les réponses du ou des filtres en (i, j) . Cette étape de sélection de filtres revient à chercher s'il y a une direction locale privilégiée autour du pixel (i, j) . Si en ce point, il faut deux ou plusieurs filtres de base pour approcher la direction présente, ou si il y a effectivement deux ou plusieurs directions présentes, il faudra utiliser plusieurs filtres pour les récupérer. Ceci est illustré par la figure 4.17. Sur cette image, on a représenté en chaque point le nombre de filtres sélectionnés par la procédure décrite auparavant et appliquée au défaut constitué d'une ligne horizontale et d'une ligne verticale. Pour les pixels se situant sur l'une ou l'autre des deux lignes, un seul filtre a été à chaque fois sélectionné : le filtre correspondant à l'orientation verticale ou horizontale. L'exception est le point de croisement des deux lignes, où deux filtres ont été sectionnés : le filtre horizontal et vertical, ce qui traduit bien la présence de deux structures corrélées horizontalement et verticalement en ce point. Dans les zones de bruit, le nombre de filtres choisis est aléatoire.

Notons que lorsqu'on parle de direction privilégiée, il s'agit de direction sur laquelle les pixels ont une valeur relativement constante et plus ou moins élevée par rapport à la moyenne de l'image. En effet, les filtres utilisés réalisent une opération de moyennage selon plusieurs directions, et ne peuvent donc pas

6. Cette formule s'applique directement dans le cas d'un bruit blanc. Dans le cas d'un bruit non blanc, il convient d'adapter la normalisation à la distribution particulière en énergie du bruit selon les canaux fréquentiels. Pour un bruit B , on aurait alors

$$\alpha^* = \arg \max_{\alpha} \left\{ \frac{\sum_{k=1}^{\alpha} |V_k^{tri}(i, j)|}{\sqrt{\sum_{k=1}^{\alpha} \|F_k * B\|^2}} \right\} \quad (4.20)$$

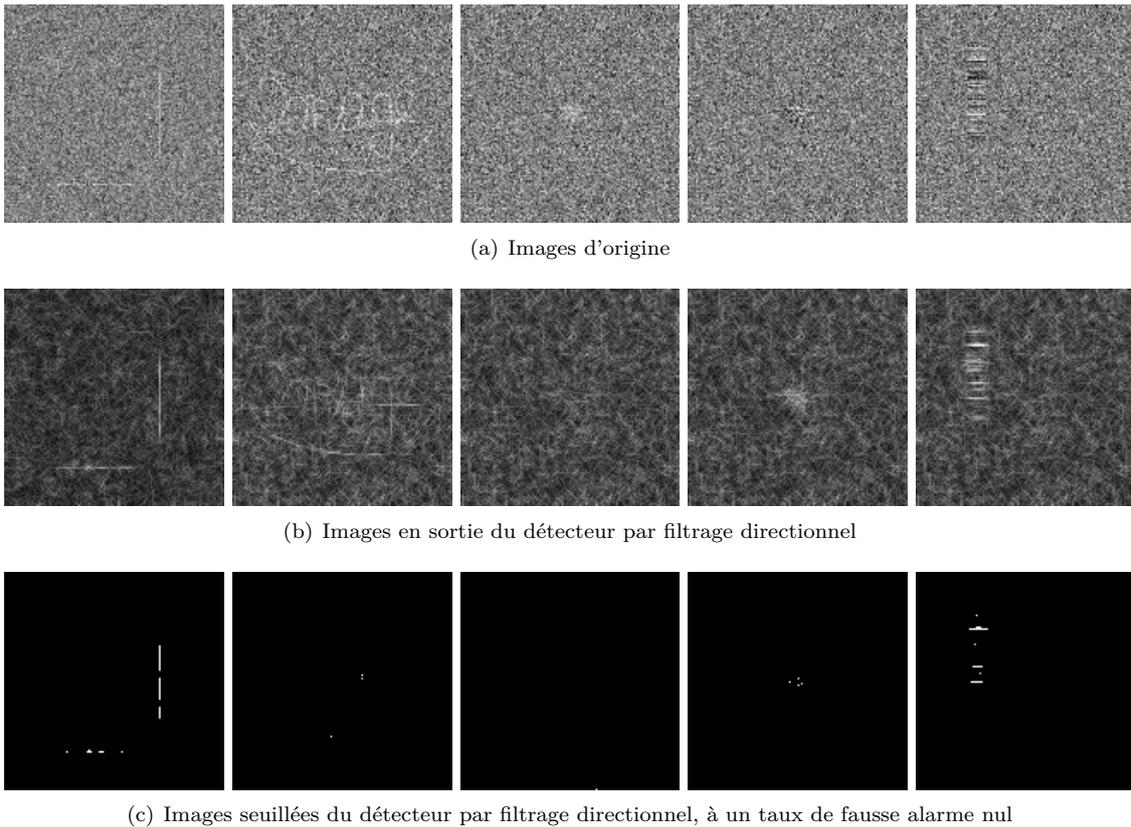


FIGURE 4.18 – Résultats et images seuillées du détecteur par filtrage directionnel

détecter de fluctuations haute fréquence selon une direction.

4.6.3 Résultats sur défauts synthétiques

Les résultats présentés ont été obtenus à l'aide d'une analyse pour $n = 8$ par les filtres décrits ci-dessus. Le filtrage a été effectué sur la totalité des images (pas d'analyse locale) de taille 128×128 pixels. Pour chaque défaut, on fait apparaître, sur la figure 4.18, l'image résultat en niveaux de gris, puis le seuillage de ces images conduisant à un taux de fausses alarmes nul.

Sur la figure 4.19, on a tracé les courbes ROC de l'algorithme pour les cinq défauts synthétiques. Pour les défauts Oscillations, Image et Lines, on a également tracé la meilleure courbe obtenue avec les algorithmes décrits précédemment. On rappelle de plus que pour les défaut Spot et Highernoise, on a obtenu, avec les méthodes précédentes, des résultats parfaits.

On constate une amélioration notable de la détection du défaut Lines. Ce type de défaut est naturellement très bien détecté par le filtrage directionnel, tout comme le défaut Oscillations. L'amélioration est également perceptible sur le défaut Image : du fait de l'aspect multi-directionnel de la méthode, on arrive à améliorer légèrement la détection de ce défaut, qui reste le plus difficile à détecter.

En contrepartie, l'artefact Spot n'est pas du tout détecté. En effet, ce défaut ne présentant pas de direction privilégiée, il est en général décomposé selon une direction assez aléatoire, comme le bruit, et ne ressort donc pas nettement dans le résultat. Le défaut Highernoise possède la même géométrie, mais est lui bien détecté. Ceci résulte de sa nature particulière : les pixels de ce défaut seront eux aussi décomposés selon une direction assez aléatoire, mais comme leur valeur individuelle est forte, leur amplitude dans l'image résultat sera plus importante que celle des pixels du défaut Spot.

La table 4.4 témoigne de l'amélioration sensible des performances sur les défauts Lines et Image. Sur ces défauts, la méthode proposée améliore assez nettement les résultats obtenus précédemment. En

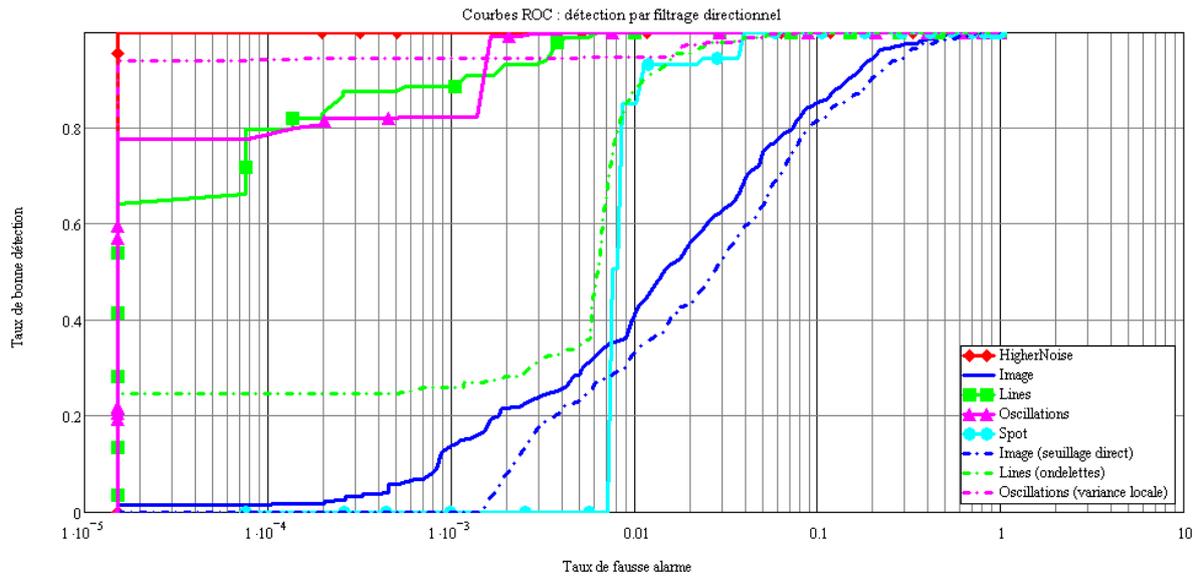


FIGURE 4.19 – Courbe ROC de la méthode de détection par filtrage directionnel. Pour les défauts Oscillations, Lines et Image, les courbes de la meilleure méthode précédemment décrite ont été représentées.

	HigherNoise	Image	Lines	Oscillations	Spot
AUC max	1,000	0,931	0,993	0,998	1,000
des méthodes précédentes	Variance Locale	Seuillage	Ondelettes	Variance Locale	Moyenne Locale
AUC filtrage directionnel	0,999	0,979	1,000	1,000	0,997

TABLE 4.4 – Aires sous courbe ROC de la méthode de détection par filtrage directionnel, comparées aux méthodes précédentes.

particulier, cette méthode est la première à réellement détecter le défaut Image.

4.6.4 Conclusions sur la détection par filtrage directionnel

La méthode de détection par filtrage directionnel présente l'avantage de pouvoir détecter des structures linéaires d'orientations variées. Le nombre d'orientations analysées est un des paramètres de la méthode. Le temps de calcul court de cette méthode permet d'envisager un traitement par bloc sur des images de grande taille, ce qui permet de gérer un de ses inconvénients, qui est son manque de prise en compte de diverses échelles. En effet, le filtrage réalisé s'apparente à un calcul de moyennes sur toute l'image selon plusieurs directions. On ne peut donc pas adapter directement cette méthode sur des images de grande taille, car on serait alors incapable de détecter des structures linéaires de dimension modeste.

Outre le manque d'aspect multi-échelle, le principal inconvénient de la méthode présentée ici est son inefficacité sur les structures circulaire, comme le défaut Spot.

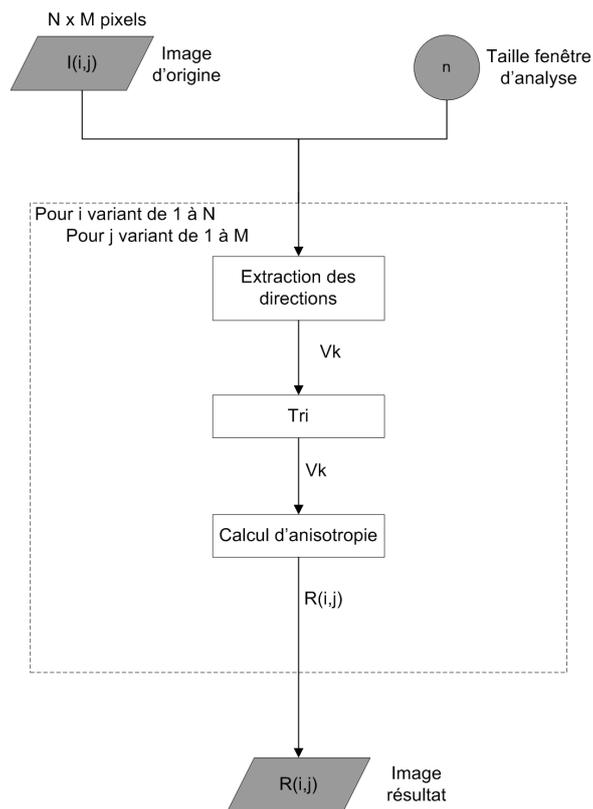


FIGURE 4.20 – Schéma de principe : détection par mesure d’anisotropie

4.7 Détection par mesure d’anisotropie

4.7.1 Principe

La méthode de filtrage directionnel décrite ci-dessus présente un moyen de retrouver les structures basse fréquence présentes autour de chaque pixel de l’image à analyser. A partir de cette méthode, nous avons voulu analyser de manière plus fine les structures filiformes locales en calculant une quantité pouvant s’apparenter à une mesure d’anisotropie locale. L’anisotropie est ici vue comme la modification brutale d’une propriété locale (autour du pixel analysé) lorsqu’elle est évaluée selon plusieurs directions.

La méthode proposée repose donc sur l’évaluation, autour d’un pixel, d’une même quantité selon plusieurs directions : la présence d’un défaut se traduira alors par un changement brutal de cette quantité, témoin d’une anisotropie locale. En effet, on a fait l’hypothèse d’un bruit non spatialement corrélé : dans le cas de nos images de test, il est blanc. Les paramètres de la méthode sont donc la mesure choisie pour calculer cette anisotropie, ainsi que la taille de la fenètre d’analyse. Par opposition à la méthode précédente, nous avons choisi d’analyser toutes les directions possibles pour un taille de fenètre d’analyse fixée. Néanmoins la dimension de cette fenètre sera forcément réduite (classiquement, une quinzaine de pixels) pour des raisons pratiques.

4.7.2 Description de l’algorithme

Le schéma de principe général de l’algorithme de détection par mesure d’anisotropie est représenté sur la fig 4.20. A partir de l’image initiale $I(i, j)$ et de n , la taille de la fenètre d’analyse (n est un nombre impair), on va calculer pour chaque pixel de $I(i, j)$ une mesure locale d’anisotropie dans l’image résultat $R(i, j)$.

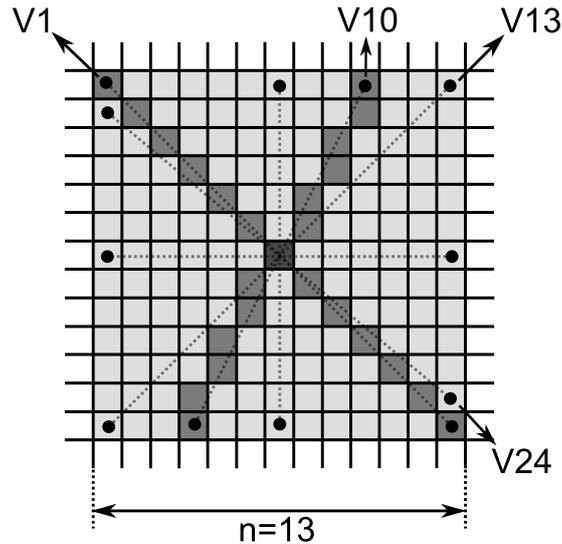


FIGURE 4.21 – Exemple d'extraction des directions

4.7.2.1 Extraction des directions

La première étape consiste à extraire les directions autour du pixel (i, j) traité. avec une fenêtre d'analyse de taille $n \times n$ centrée sur le pixel traité, il y a exactement $2(n - 1)$ directions à extraire. Ces directions correspondent aux droites discrètes contenant n pixels et passant par le pixel traité. La fig 4.21 montre une extraction des directions pour une fenêtre d'analyse de taille $n = 13$. Les pixels des $2(n - 1) = 24$ droites discrètes sont stockés dans les vecteurs $V_k(i)$, avec $1 \leq k \leq 2(n - 1)$ et $1 \leq i \leq n$. Sur la fig 4.21, on a tracé les droites discrètes pour $k = 1$ et $k = 10$. Ces droites discrètes sont déterminées à l'aide d'algorithmes type Bresenham [12].

4.7.2.2 Tri

Chaque vecteur V_k est ensuite trié par valeurs croissante ou décroissante.

4.7.2.3 Calcul d'anisotropie

A partir de la famille de vecteurs triés $\{V_k(i)\}_{1 \leq k \leq 2(n-1)}$, on va calculer une mesure locale d'anisotropie. Pour notre application, la mesure \mathcal{A} retenue est :

$$\mathcal{A} = \max_k \sum_{i=1}^n |V_k(i) - \bar{V}(i)| \quad (4.22)$$

avec $\bar{V}(i) = \text{médiane} \{V_k(i)\}_{1 \leq k \leq 2(n-1)}$.

Cette mesure revient à comparer la dynamique des pixels de chaque direction à celle d'un vecteur médian \bar{V} . Cette opération peut s'apparenter au calcul d'une distance entre histogrammes à très petite échelle.

4.7.3 Résultats sur défauts synthétiques

La figure 4.22 montre les résultats de cette méthode de détection sur les défauts synthétiques, ainsi que le seuillage obtenu selon les mêmes critères que précédemment. Le paramètre de taille de la fenêtre d'analyse a été réglé à $n = 13$.

Les résultats sont très bons sur le défaut Lines. Les défauts Spot et HigherNoise sont en revanche très mal détectés, la méthode étant inadaptée dans le cas de défauts de ce type. Le défaut Image est

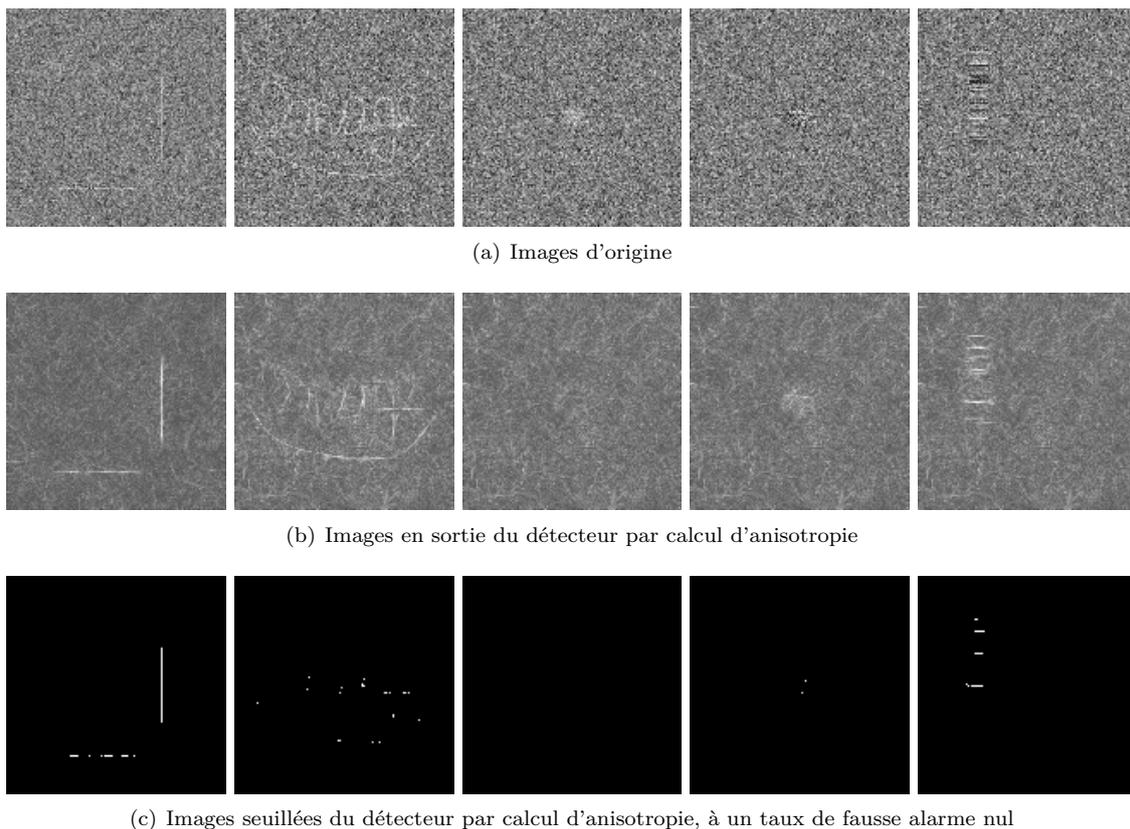


FIGURE 4.22 – Résultats et images seuillées du détecteur par mesure d'anisotropie

bien détecté : on obtient un résultat comparable à celui obtenu avec la méthode du filtrage directionnel. Enfin, le défaut Oscillations est moyennement détecté : pour ce type de défaut la mesure d'anisotropie est visiblement moins efficace que le filtrage fréquentiel ou la variance locale. L'avantage de cette méthode repose dans la détection relativement bonne du défaut Image, qui est un des plus difficiles à traiter, ainsi que du défaut Lines. Pour ce type de défaut, l'analyse exhaustive des directions locales ainsi que l'échelle très petite (fenêtre de 13×13 pixels) choisie pour le calcul de la mesure d'anisotropie permettent de détecter efficacement les orientations de petite taille.

Le principal désavantage de cette méthode est le temps de calcul important, qui, même optimisé, requiert 50 secondes sous Mathcad tournant sur un P4@2.8Ghz, 2Go Ram et pour une image 128×128 pixels, tandis que les méthodes précédentes ne requièrent qu'une dizaine de secondes pour traiter une image.

4.7.4 Conclusion sur la méthode de détection par calcul d'anisotropie

La méthode de détection par calcul d'anisotropie permet de détecter très efficacement les alignements de taille réduite de pixels défectueux. Elle utilise une analyse exhaustive des orientations possibles de ces alignements puis en dérive, à l'aide d'un critère d'anisotropie, une mesure qui est particulièrement efficace pour détecter ces alignements. Encore une fois, cette méthode n'est pas intrinsèquement multi-échelle, mais peut être adaptée en faisant varier la taille de la fenêtre d'analyse. Néanmoins, contrairement à la méthode précédente, le temps de calcul pose ici problème car les temps constatés pour traiter une image de petite taille, à une échelle donnée, semble être un frein pour l'adaptation multi-échelle de cette méthode sur des images de grande taille.

De plus, cette méthode est très spécialisée dans la détection des alignements de pixels défectueux, et manque ainsi de polyvalence, notamment pour les défaut de type tache claire ou sombre.

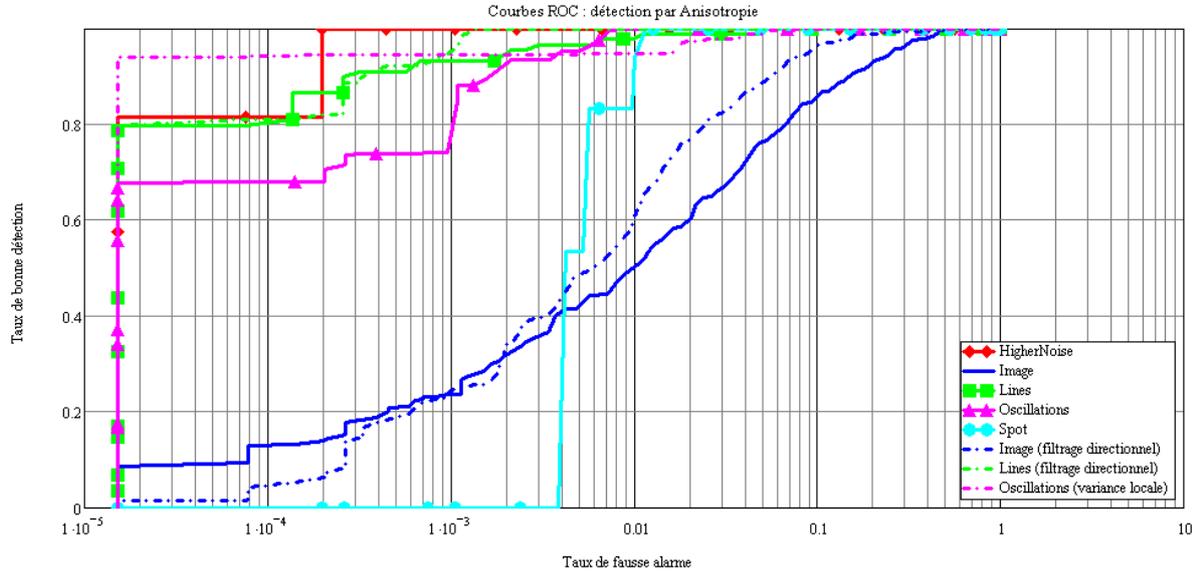


FIGURE 4.23 – Courbe ROC de la méthode de détection par calcul d'anisotropie. Pour les défauts Oscillations, Lines et Image, les courbes de la meilleure méthode précédemment décrite ont été représentées.

	HigherNoise	Image	Lines	Oscillations	Spot
AUC max	1,000	0,979	1,000	0,998	1,000
des méthodes	Variance	Filtrage	Filtrage	Variance	Moyenne
précédentes	Locale	directionnel	directionnel	Locale	Locale
AUC anisotropie	0,999	0,951	0,999	0,999	0,994

TABLE 4.5 – Aires sous courbe ROC de la méthode de détection par calcul d'anisotropie, comparée aux méthodes précédentes.

4.8 Détection par moyenne statistique pondérée

4.8.0.1 Principe

Les méthodes précédentes étaient principalement dédiées à la détection de structures filiformes type ligne. En effet, ce type de défaut corrélé présente un ratio visibilité/déteçtabilité particulièrement élevé : les pixels constituant une ligne visible dans l'image ne sont pas individuellement atypiques alors qu'ils sont en général très visibles.

La méthode présentée dans cette section a été conçue de manière à traiter une plus large catégorie de défauts. Elle repose sur la constatation suivante : lorsque plusieurs pixels défectueux sont proches, ils sont identifiés plus facilement par l'œil comme défectueux. L'idée de cette méthode est alors de calculer une nouvelle valeur pour chaque pixel de l'image à traiter en effectuant une moyenne pondérée des pixels voisins. Les coefficients de pondération mesurent une similarité statistique entre pixels. Cette méthode est principalement paramétrée par :

- n : la taille du voisinage considéré
- $\mathcal{S}(p_i, p_j)$: la fonction permettant de mesurer la similarité statistique entre 2 pixels p_i et p_j de l'image d'origine.

4.8.1 Description de l'algorithme

Le schéma de principe de la méthode est représenté sur la fig 4.24. A partir d'une image d'origine I et du paramètre n contrôlant la taille de la fenêtre définissant le voisinage de chaque pixel, on va calculer une nouvelle image R où les pixels de défauts seront détectés par simple seuillage. Pour cela, on travaille sur un voisinage de chaque pixel $I(i, j)$ traité, voisinage qu'on notera $V_{i,j}$. Ce voisinage est de taille $n \times n$ pixels, et est centré sur le pixel (i, j) . On parcourt les valeurs de ce voisinage en prenant $V_{i,j}(r, c)$ avec $(r, c) \in \llbracket 1; n \rrbracket^2$.

La formule générale donnant $R(i, j)$ peut alors s'écrire :

$$R(i, j) = \sum_{r=1}^n \sum_{c=1}^n K_{i,j}(r, c) V_{i,j}(r, c) \quad (4.23)$$

4.8.1.1 Extraction du voisinage

La première étape est une simple extraction des pixels autour du pixel (i, j) traité. On récupère ainsi une ROI notée $V_{i,j}$ de taille $n \times n$ centrée sur le pixel traité.

4.8.1.2 Calcul des coefficients du filtre

A partir du voisinage récupéré, on va calculer les coefficients $K_{i,j}(r, c)$, $(r, c) \in \llbracket 1; n \rrbracket^2$ par la formule suivante :

$$K_{i,j}(r, c) = \frac{1}{C} \cdot \frac{1}{1 + \left(\frac{V_{i,j}(r, c) - x}{\sigma} \right)^2} \quad (4.24)$$

σ étant l'écart-type de l'image analysée (ou un écart-type locale dans le cas d'image de grande taille), et le facteur $\frac{1}{C}$ permettant la normalisation du filtre $K_{i,j}$. On a donc $\sum_{r,c} K_{i,j}(r, c) = 1$.

D'un point de vue général, le coefficient $K_{i,j}(r, c)$ varie entre 0 et 1, et est d'autant plus grand que la valeur de $V_{i,j}(r, c)$ est proche de la valeur x du pixel traité.

4.8.1.3 Sommation

La dernière étape consiste à calculer la somme des pixels de $V_{i,j}$ pondérée par les coefficients $K_{i,j}$ par la formule 4.23. Ainsi, on donnera plus de poids aux pixels ayant des valeurs proches de celle du pixel traité.

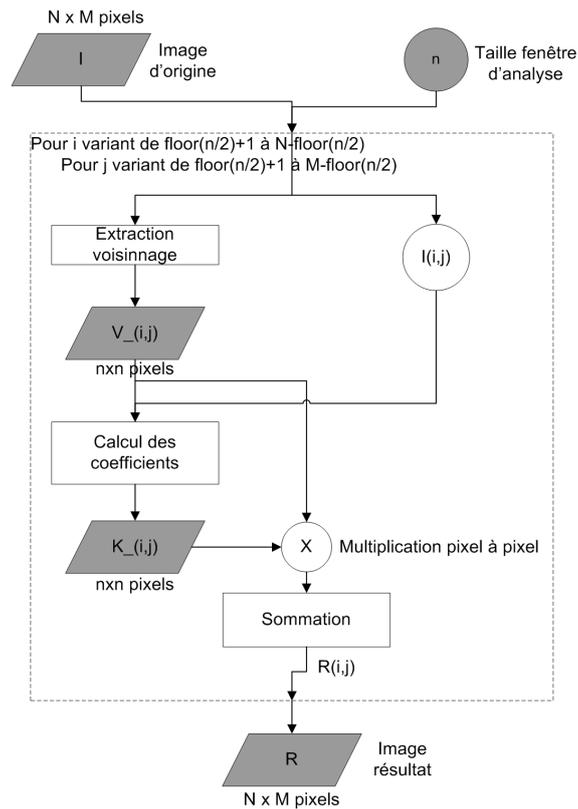


FIGURE 4.24 – Schéma de principe : détection par moyenne pondérée

4.8.2 Résultats sur défauts synthétiques

Les résultats obtenus avec la méthode de moyenne statistique pondérée sont présentés par la fig 4.25.

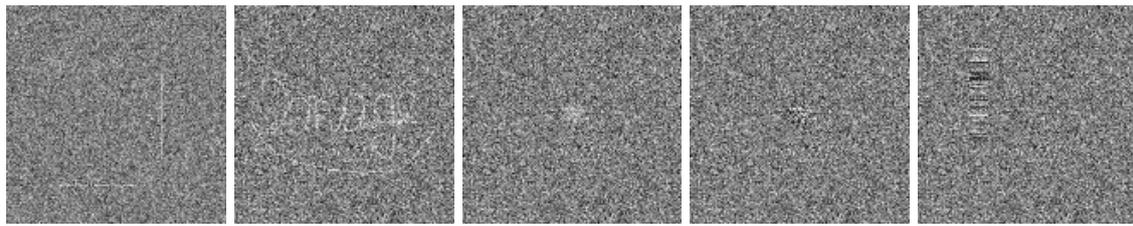
Sur les défauts de type Lines, les résultats de détection sont moins bons que les méthodes de filtrage fréquentiel, d'anisotropie ou d'ondelettes. Il faut noter que pour ces types d'artefacts, la méthode de moyenne pondérée n'est pas censée donner de bons résultats car ces défauts ne sont pas compacts, et donc peu de pixels de défauts sont présents dans la fenêtre de voisinage carrée analysée. Les bons résultats constatés ici sont donc plutôt une conséquence du débruitage effectué, plutôt que d'un renforcement de la zone de défaut.

Sur les défauts Spot et HigherNoise, les résultats sont très satisfaisants : les images seuillées ne présentent aucune fausse alarme et la détection réalisée est très satisfaisante pour notre application. Le défaut Oscillations est assez bien détecté, mais les performances restent néanmoins en dessous de l'algorithme de variance locale, qui est très efficace sur ce genre de défaut.

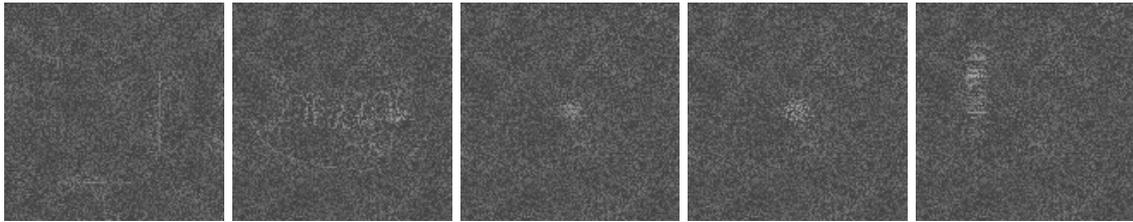
Enfin, le défaut Image est détecté de manière équivalente à la meilleure méthode observée jusqu'à présent sur ce défaut (à savoir la méthode du filtrage directionnel). En particulier, on observe de bonnes performances pour des taux de fausse alarme très faible.

4.8.3 Conclusions sur l'algorithme de détection par moyenne pondérée

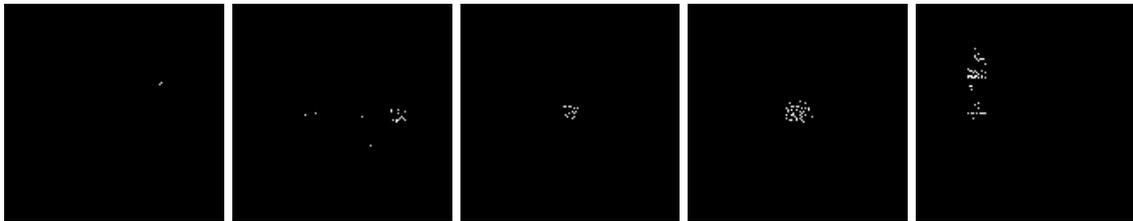
L'algorithme décrit dans cette section est l'un des plus polyvalents élaborés lors de cette thèse. Son efficacité sur les défauts synthétiques est satisfaisante, à l'exception notable du défaut Lines. En effet, pour ce type de défauts, il faudrait adapter la fenêtre d'analyse à la géométrie du défaut, en considérant par exemple des fenêtre rectangulaire suivant l'extension spatiale des alignements à détecter. Cette adaptation, forcément coûteuse en temps de calcul, n'est pas réellement satisfaisante dans le cadre d'un contrôle qualité d'images de grande taille. De plus, même si les performances de l'algorithme de détection



(a) Images d'origine



(b) Images en sortie du détecteur par moyenne statistique pondérée



(c) Images seuillées du détecteur par moyenne statistique pondérée, à un taux de fausse alarme nul

FIGURE 4.25 – Résultats et images seuillées du détecteur par moyenne statistique pondérée

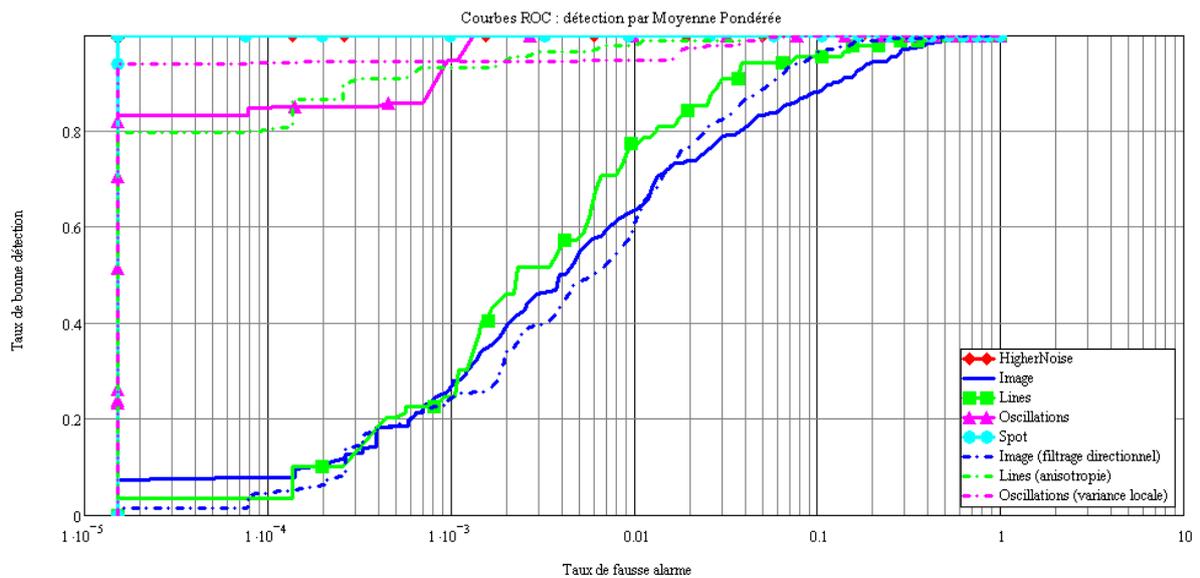


FIGURE 4.26 – Courbes ROC de la méthode de détection par moyenne statistique pondérée. Pour les défauts Oscillations, Lines et Image, les courbes de la meilleure méthode précédemment décrite ont été représentées.

	HigherNoise	Image	Lines	Oscillations	Spot
AUC max des méthodes précédentes	1,000 Variance Locale	0,979 Filtrage directionnel	1.000 Filtrage directionnel	0,998 Variance Locale	1,000 Moyenne Locale
AUC moyenne statistique pondérée	1,000	0,960	0.983	1,000	1,000

TABLE 4.6 – Aires sous courbe ROC de la méthode de détection par moyenne statistique pondérée, comparée aux méthodes précédentes.

par moyenne statistique pondérée sont moyennes sur ces types de défauts, on observe qu'on arrive à détecter partiellement le défaut Lines, sans générer de fausse alarme. Ceci est généralement suffisant pour une méthode implémentée dans la première étape de détection de la chaîne d'analyse.

4.9 Détection par snakes statistiques et rapports de vraisemblance

4.9.0.1 Principe

En plus des méthodes développées et testées dans le cadre des travaux de thèse, nous avons exploré d'autres types de méthodes de détection avec l'aide de partenaires externes⁷. Ces deux types de méthodes sont :

- les snakes statistiques [73, 63] et les tests de maximum du rapport de vraisemblance [83] par l'intermédiaire de la société KaoLab
- la segmentation markovienne [33, 74, 8]

Les résultats de la segmentation markovienne ont montré que ces méthodes ne sont pas directement applicables à notre problème, principalement du fait de la topologie des défauts à détecter. En effet, la segmentation markovienne testée dans le cadre du partenariat nécessite des défauts compacts et homogènes pour être efficace. Ainsi les résultats obtenus étaient bons uniquement sur le défaut Spot.

A l'opposé, le couplage des méthodes de segmentation par snakes statistiques et par maximum de vraisemblance a permis d'obtenir de très bons résultats sur tous les défauts synthétiques traités.

4.9.1 Snakes statistiques et rapport de vraisemblance : principes

4.9.1.1 Le test du maximum du rapport de vraisemblance

Les résultats obtenus par KaoLab sur nos images de défauts synthétiques reposent principalement sur deux méthodes. La première utilise un principe simple de calcul d'un rapport de vraisemblance, les défauts à détecter maximisant alors ce rapport. Cette méthode nécessite de déterminer une estimation de la forme du défaut à rechercher, à partir de laquelle on va construire un masque de ce défaut noté w .

L'image où on recherche ce défaut sera notée, pour simplifier, de manière unidimensionnelle, $S = \{s_i, i \in [0, N - 1]\}$. Elle comporte deux classes de pixels : les pixels de défaut et les pixels de fond. Les pixels de défaut ont une densité de probabilité notée $P(\theta_t, s_i)$ et ceux du fond une densité notée $P(\theta_b, s_i)$. On doit également déterminer une fenêtre d'analyse notée F , dont la taille représente quelque fois la taille de w (typiquement 2 à 3 fois).

Le test du rapport du maximum du rapport de vraisemblance (MLRT) consiste à chercher le site optimal j^* maximisant la log-vraisemblance :

7. Partenaires : société KaoLab et INRIA Rhône-Alpes (Florence Forbes et Juliette Blanchet)

$$j^* = \arg \max_j \left\{ \sum_{i \in w^j} \log(P(\theta_t, s_i)) + \sum_{i \in \overline{w^j}} \log(P(\theta_b, s_i)) - \sum_{i \in F^j} \log(P(\theta_b, s_i)) \right\} \quad (4.25)$$

où w^j est une version du masque w centrée sur le pixel j , F^j est une version de la fenêtre d'analyse centrée sur le pixel j , et $\overline{w^j}$ est le complémentaire de w^j dans F^j

La quantité à maximiser correspond au test d'hypothèse H_0^j vs H_1^j avec :

- H_0^j : il y a uniquement des pixels de fond dans F^j ,
- H_1^j : il y a dans F^j un défaut de forme w centré sur le pixel j .

Ce test se révèle être quasi-équivalent à l'algorithme optimal du maximum de vraisemblance (ML). La différence se situant sur les fenêtres d'évaluation des paramètres de loi θ_t et θ_b : l'algorithme ML nécessite une estimation de θ_b sur l'image entière S , alors que le MLRT se contente de l'estimation sur $\overline{w^j}$. On montre que le MLRT garde de bonnes performances dans le cas, idéal pour l'algorithme ML, où le bruit de fond est homogène sur S . De plus, il est particulièrement efficace dans un cas de bruit non homogène.

Bien que cet algorithme nécessite la connaissance de w , on peut raisonnablement penser qu'une estimation grossière de ce masque ne détériore pas de manière significative les performances de détection. Un inconvénient plus contraignant est la nécessité de connaître *a priori* les lois suivies par les classes de défaut et de bruit de fond.

L'algorithme du MLRT a été utilisé pour la détection des défauts Spot et HigherNoise (voir résultats sur la figure 4.27).

4.9.1.2 Les snakes statistiques

La seconde méthode utilisée par les algorithmes de détection de KaoLab repose sur un principe de contours actifs (snakes) dont le critère d'évolution est de type statistique. Les méthodes de contours actifs réalisent une segmentation de l'image en deux régions (l'intérieur du contour et l'extérieur du contour) par la minimisation d'un critère qui sert de critère d'évolution des nœuds du contour. Les premières méthodes s'appuyaient principalement sur un critère à base de détection des contours des régions, et se révélaient donc peu efficaces dans le cas d'images fortement bruitées. Dans [73], il est proposé un critère original reposant sur l'évaluation de grandeurs statistiques et ne nécessitant pas d'évaluation paramétrique des lois des régions à segmenter. Ce critère est appelé complexité stochastique, Δ avec :

$$\Delta = \Delta_S + \Delta_P + \Delta_C \quad (4.26)$$

avec :

- Δ_S : le nombre de bits nécessaire pour coder les niveaux de gris de l'image connaissant le contour actif et les densités de probabilité de chaque région de l'image définie par ce contour,
- Δ_P : le nombre de bits nécessaire pour coder les densités de probabilités des régions,
- Δ_C : le nombre de bits nécessaire pour coder le contour actif.

La segmentation finale est alors déterminée par convergence du contour actif, une méthode d'optimisation de cette convergence étant développée dans la référence citée précédemment.

La puissance de cette méthode repose dans la modélisation des densités de probabilités des régions par des fonctions constantes par morceaux qui permettent d'évaluer facilement Δ_P sans perdre en généralité. Les paramètres de ces fonctions sont dans le même temps évalués par la méthode de segmentation. Ainsi, on réalise une segmentation non-paramétrique de l'image particulièrement robuste vis à vis des lois de probabilité suivies par les différentes régions de l'image. Néanmoins, dans les cas de très faible contraste entre ces régions, les approches paramétriques restent naturellement plus efficaces.

Cette méthode a été utilisée pour la détection des défauts Square et Oscillations (voir figure 4.27).

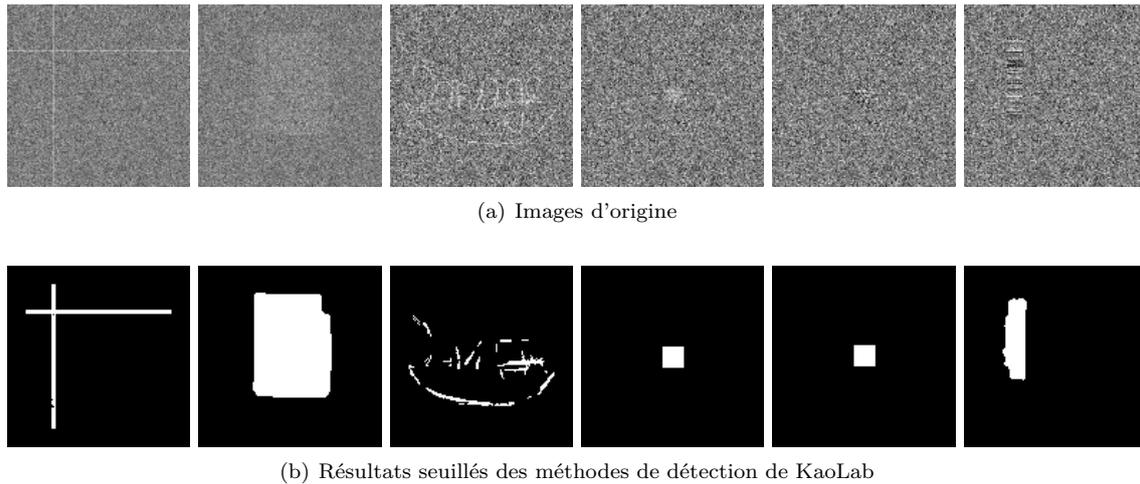


FIGURE 4.27 – Résultats des détecteurs KaoLab. Pour chaque défaut, on a représenté le meilleur résultat donné entre la méthode des snakes statistiques et du rapport de vraisemblance.

4.9.2 Résultats sur défauts synthétiques

Les résultats des méthodes élaborées par Kaolab sont présentés sur la fig 4.27. Seules les images résultats seuillées sont représentées. Les quatre dernières images représentées sont quatre des cinq défauts synthétiques servant de base de référence pour les méthodes de détection présentées précédemment. Le défaut Lines est absent pour des raisons techniques, mais est remplacé par un défaut du même type (première image).

Il faut noter que les résultats obtenus ne l'ont pas été avec un seul et même algorithme de détection : pour chaque défaut, seule la méthode de détection (parmi les trois disponibles) donnant le meilleur résultat a été sélectionnée. Les défauts Spot et HigherNoise ont ainsi été détectés vraisemblablement à l'aide d'une méthode à base de MLRT, les défauts Oscillations et Square ont été eux traités avec une méthode à base de snakes statistiques, et les défauts type Lines et Image ont été traités à l'aide d'une méthode dédiée à la détection de lignes.

Les résultats sont bons, à l'exception notable du défaut Image, qui semble correctement détecté mais qui contient en fait un nombre important de fausses alarmes. Sur la figure 4.27, on a confronté les résultats obtenus avec la méthode la plus efficace testée jusqu'ici. On observe que les défauts Spot, Oscillations et HigherNoise sont parfaitement détectés. L'apport de la méthode de détection par snakes statistiques est surtout important pour le défaut Oscillations. Le défaut Image reste assez mal détecté : pour ce défaut, on ne fait pas mieux que la méthode de la moyenne statistique pondérée décrite dans la section précédente.

Globalement, les résultats obtenus à l'aide de cette nouvelle méthode sont très bons. Il faut néanmoins les nuancer étant donné qu'ils sont le fruit de plusieurs algorithmes dédiés, et que des tests menés sur des images présentant plusieurs défauts hétérogènes ont montré la grande difficulté à combiner efficacement les trois algorithmes en une seule méthode de détection cohérente.

4.9.3 Conclusions sur la méthode de détection par snakes statistiques et MLRT

La méthode testée ici donne des résultats très convaincants sur les défauts synthétiques. Néanmoins, si on considère que chacun des résultats est issu de la sélection d'un algorithme adapté à la nature de chaque défaut, parmi trois algorithmes différents, on peut dire que la méthode de détection par snakes ou par MLRT est simplement au niveau des méthodes élaborées précédemment. Les essais menés ici ont permis de confirmer la difficulté de la tâche de détection, lorsqu'on ne se donne aucun *a priori* sur les défauts à détecter. Pour obtenir de bons résultats, on a dû choisir parmi trois algorithmes différents, tous

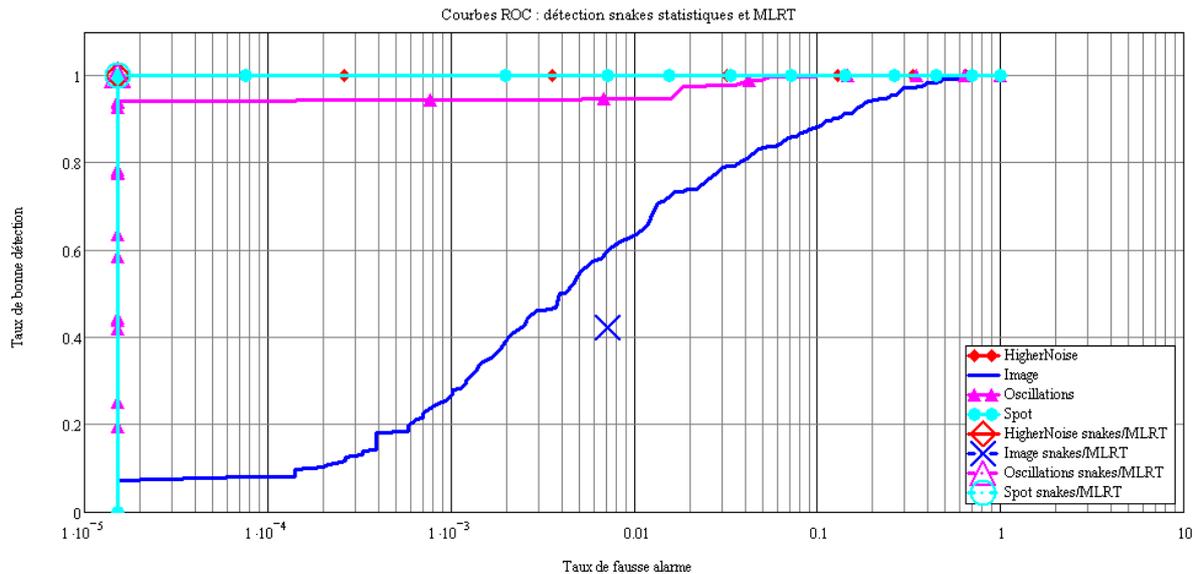


FIGURE 4.28 – Courbe ROC de la méthode de détection par snakes statistiques ou Maximum Likelihood Ratio Test. Les courbes des meilleurs algorithmes pour les quatre défauts synthétiques (HigherNoise, Image, Oscillations et Spot) sont représentées ainsi que les points correspondant aux détections par snakes ou MLRT.

trois étant au niveau de l'état de l'art.

4.10 Synthèse

Les algorithmes exposés dans ce chapitre permettent d'améliorer de manière importante les performances obtenues avec l'algorithme classique de détection à base de statistiques de moyenne ou de la variance locales. En particulier, l'algorithme de moyenne statistique pondérée offre une grande polyvalence sur la classe de défauts traitée : il permet de détecter à chaque fois la présence de défaut, avant toute fausse alarme. De plus, au prix d'un taux de fausse alarme un peu plus élevé, il permet une bonne détection des pixels composant le défaut. La figure 4.29 représente les meilleures courbes ROC obtenues avec les méthodes décrites dans ce chapitre. On voit en effet que la méthode par moyenne statistique pondérée est la meilleure, ou comparable à la meilleure méthode, dans trois cas sur cinq.

La table 4.7 donne une synthèse quantitative des résultats obtenus avec les algorithmes décrits. Pour chaque défaut, et pour chaque algorithme de détection, on donne la mesure de l'aire sous la courbe ROC. Pour réaliser une synthèse mieux adaptée à notre contexte, on a calculé dans cette table l'aire de la courbe restreinte à un taux de fausse alarme se trouvant entre 0 et $\tau fa_{max} = 10^{-4}$. En effet, des taux de fausse alarme supérieurs ne sont en pratique pas acceptables. Les aires données ont été normalisées⁸ de manière à ce qu'elle varient entre 0 et 1.

On peut tirer plusieurs enseignements de ces résultats. Le premier est qu'aucun algorithme n'est meilleur dans tous les cas, ce qui était attendu étant donnée la diversité des défauts considérés. En second lieu, on observe bien que l'algorithme de détection par moyenne statistique pondérée est un bon compromis dans le cas où on veut traiter tous ces défauts à l'aide d'un seul algorithme. Encore une fois, on peut dire que cet algorithme est le plus polyvalent, et donne de plus des résultats excellents dans certains cas. Enfin, on remarque la difficulté rencontrée pour la détection du défaut Image via les faibles AUC obtenues sur ce défaut. Ce défaut est très difficile à détecter de manière exhaustive. Dans les tests réalisés, on constate que seulement une petite partie de ce défaut est détectée, à la jonction du «g» et du «e». En général, le reste du défaut est détectable seulement à des taux de fausse alarme très hauts.

8. Sans normalisation, en gardant la définition naturelle de l'aire sous une courbe, les aires calculées varieraient entre 0

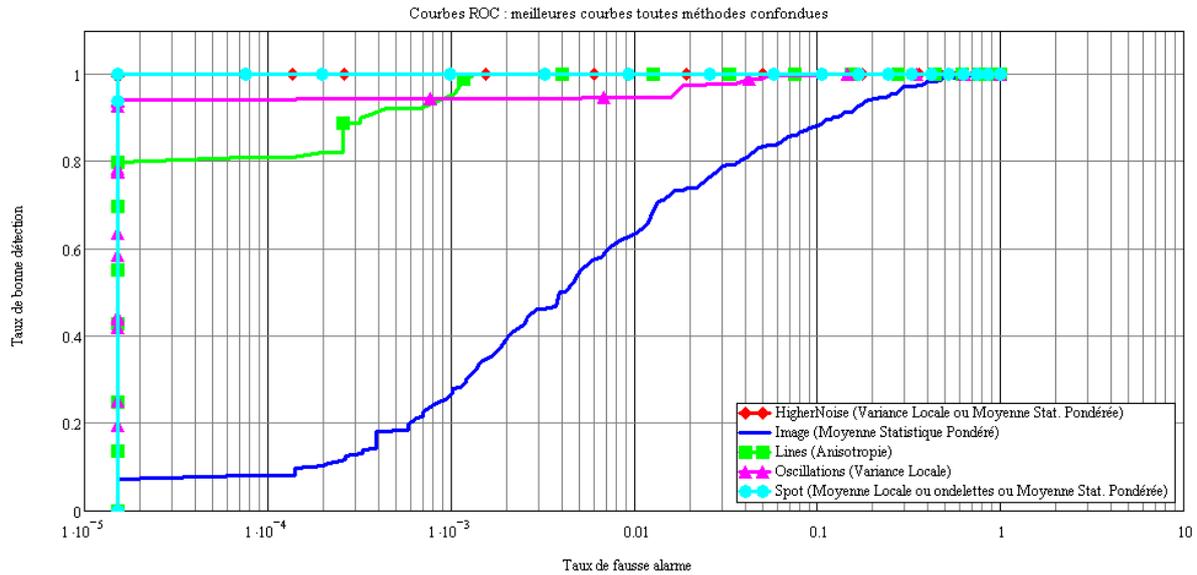


FIGURE 4.29 – Meilleures courbes ROC obtenues avec les méthodes décrites dans ce chapitre. Pour chaque courbe, la méthode correspondante est renseignée en légende.

	HigherNoise	Image	Lines	Oscillations	Spot
Seuillage	0,000	0,000	0,034	0,000	0,000
Moyenne Locale	0,000	0,006	0,000	0,512	1,000
Variance Locale	1,000	0,000	0,000	0,942	0,000
Skewness Local	0,000	0,000	0,000	0,000	0,000
Kurtosis Local	0,000	0,000	0,000	0,000	0,000
Ondelettes	0,978	0,012	0,247	0,831	1,000
Filtrage directionnel	0,349	0,025	0,806	0,643	0,000
Anisotropie	0,816	0,103	0,798	0,678	0,000
Moyenne Statistique Pondérée	1,000	0,077	0,034	0,838	1,000

TABLE 4.7 – Synthèse des aires sous courbes ROC des méthodes présentées dans ce chapitre. Les aires calculées correspondent aux aires sous les courbes ROC restreintes à la plage de fausse alarme $[0; \tau fa_{max}]$ avec $\tau fa_{max} = 10^{-4}$. Pour chaque défaut, la meilleure méthode est signalée par une cellule grisée.

4.11 Conclusions et perspectives

Dans cette partie, nous avons donc vu comment améliorer, pour les cas de détection difficiles, les performances de détection au travers d'approche différentes. Ces approches mêlent des aspects statistiques aux aspects spatiaux : en effet, il nous paraît obligatoire de prendre en compte de manière importante les corrélations spatiales lors de la détection, la difficulté étant de coupler des critères spatiaux aux critères statistiques. Ceci doit être fait sans perdre en généralité puisque la nature des défauts à détecter est très variable. Ainsi, nous nous sommes attachés, dans les méthodes exposées, à ne pas tendre vers l'élaboration de méthodes trop spécialisées dans la détection d'un type précis de défaut. C'est ainsi que la méthode de la moyenne statistique pondérée permet de traiter une large gamme de défauts.

Afin d'appliquer cette méthode à la chaîne d'analyse, il reste certaines questions que nous n'avons malheureusement pas pu aborder.

Dans un premier temps, la plupart des méthodes décrites dans ce chapitre ne considèrent qu'une seule échelle de détection. Sur les défauts synthétiques dont nous disposons, ceci ne posait pas de problème étant donnée la taille réduite des images traitées. Dans les cas réels, on s'aperçoit que certains défauts ne sont visibles qu'à une échelle donnée. Afin de prendre en compte cette constatation, il faudrait élaborer un schéma d'implémentation multi-échelle des méthodes de détection. Ceci peut passer par l'utilisation de la décomposition en ondelettes : l'algorithme étant alors appliqué aux versions sous-échantillonnées de l'image, en haute et en basse fréquence. Afin de réduire le nombre d'échelles traitées, on peut aussi uniquement travailler sur trois zooms différents de l'image d'origine. On choisirait alors des facteurs de zoom pertinents par rapport aux échelles d'observation des images de taille 3000×3000 pixels.

Un autre aspect à prendre en considération lorsqu'on traite des images réelles est la normalisation locale par rapport au bruit de fond. En effet, dans la pratique, le bruit de fond n'est souvent pas homogène : on observe, même sur les images débarrassées de leurs structures (voir section 2.2), des variations locales de la variance et du spectre du bruit. Lorsqu'on veut appliquer les algorithmes de détection à ces images, il faut prendre en compte ces variations afin d'effectuer une normalisation non plus globale, mais locale de l'image de détection. Ceci a d'ailleurs été pris en considération dans la méthode de détection de base exposée plus tôt (section 2.3).

D'un point de vue pratique, on ne peut faire abstraction de la contrainte de temps de calcul, même si celle-ci peut paraître accessoire dans un premier temps. En effet, certaines méthodes de détection s'appuient sur une analyse exhaustive de configurations particulières de pixels, ce qui peut être, en termes de performances de détection, très efficace. Néanmoins, sans schéma d'implémentation rapide, on ne peut utiliser ces méthodes pour détecter des défauts de forme et de nature inconnues sur des images de grande taille.

Enfin, même si à chaque fois nous nous sommes attachés à donner aux méthodes de détection une grande généralité, il est clair que certaines sont efficaces sur une plage donnée de défauts. La construction d'un algorithme efficace reposant sur un seul principe de détection nous paraît improbable. Afin de tendre vers un schéma de détection très efficace, il faudrait donc étudier les possibilités de fusion des résultats donnés par les algorithmes de détection. Ceci représente, à notre avis, l'axe de travail le plus important. C'est aussi le plus prometteur dans le contexte industriel puisqu'il permettrait d'ajouter une grande robustesse (suppression des fausses alarmes générées par certains algorithmes, grâce à l'information donnée par d'autres algorithmes) et des possibilités de classification (classification des défauts selon les réponses des méthodes de détection, gestion de l'apparition d'une nouvelle classe de défauts, *etc*).

Troisième partie

Mesure de visibilité des défauts dans une image

Chapitre 5

Mesure objective de visibilité de défauts

Résumé

Ce chapitre est introductif à l'élaboration d'une mesure de visibilité, élément final de la chaîne d'analyse de la qualité des images. Dans un premier temps, un aperçu bibliographique de la problématique de la mesure objective de la qualité est donné. A travers trois catégories de mesures objectives, faisant également partie de trois champs d'application différents, nous donnons les architectures les plus courantes servant de base de développement des mesures objectives. Les résultats de tests menés sur nos défauts synthétiques sont ensuite exposés, pour deux mesures choisies. Ces tests montrent que l'ajout de mécanismes psycho-visuels du système visuel humain semble améliorer sensiblement les résultats de la mesure. Une base prometteuse de futurs développements, s'appuyant sur le Channelized Hotelling Observer, est également brièvement décrite : cette base permet d'incorporer de manière relativement aisée de tels aspects psycho-visuels.

Sommaire

5.1	But d'une mesure objective de visibilité	129
5.2	Mesures objectives : éléments bibliographiques	130
5.2.1	Mesures classiques d'artefacts	130
5.2.2	Les mesures statistiques	133
5.2.3	Les mesures de type «Model Observer»	135
5.3	Application de mesures de visibilité sur les défauts synthétiques	138
5.3.1	Mesure de référence : énergie avec normalisation locale	139
5.3.2	Mesure de l'Hotelling Observer (HO)	140
5.3.3	L'Hotelling Observer non blanchisseur avec filtre de l'œil (NPWE)	140
5.4	Conclusions	143

La construction d'une mesure objective de qualité d'une image est au cœur de très nombreuses études en traitement d'images. Malgré cette quantité importante de travaux, il n'y a aujourd'hui aucune méthode universelle permettant de mesurer la qualité d'une image de manière fiable, *i.e.* de manière suffisamment proche de l'évaluation qui serait faite par l'œil humain. La principale raison permettant d'expliquer cette situation repose certainement sur la diversité des situations, des natures des objets à évaluer, et des conditions d'observations, qui empêche l'unification de toutes les mesures objectives de qualité existantes en une seule mesure cohérente. En effet, l'efficacité de la mesure objective passe bien souvent par la prise en compte des spécificités de la tâche d'évaluation, ce qui rend difficile toute unification. D'ailleurs, selon la tâche et les conditions d'observation, les mécanismes mis en jeu dans la mesure de la qualité par l'œil humain sont très variables.

Dans ce chapitre, nous allons d'abord définir le contexte et les objectifs de la mesure objective de qualité que nous souhaitons mettre en place. Nous évoquerons également, dans les grandes lignes, les travaux déjà menés dans le domaine, notamment les principales classes de mesures de qualité existantes, ainsi que les moyens à mettre en œuvre pour valider correctement une mesure objective. Enfin, nous terminerons par l'illustration de la méthode de validation que nous avons mis en place, sur une mesure objective basique de la qualité, ainsi que sur deux mesures de type Hotelling Observer, dérivées de la littérature. Deux moyens de valider ces mesures seront présentés. Ce chapitre ouvre la voie à la perspectives de développements ultérieurs sur l'amélioration de la mesure de qualité, dans le contexte défini par ce chapitre.

5.1 But d'une mesure objective de visibilité

L'objectif premier de la mesure objective de visibilité est de permettre une évaluation automatique des défauts présents dans les images. Dans le chapitre 2, nous avons vu à quel niveau doit intervenir la mesure de qualité, que nous avons appelée mesure de visibilité, car elle traduit la présence ou non d'un défaut au sens de l'observation visuelle. Cette mesure automatique permet donc de quantifier la visibilité de chaque défaut, sans que ceci nécessite l'intervention d'un observateur humain, coûteuse en temps, et souvent rébarbative pour l'observateur qui doit analyser des milliers d'images de grande taille pouvant contenir des défauts de quelques pixels.

Afin de mieux définir ce que l'on attend de cette mesure, nous avons besoin de préciser les éléments suivants :

1. la tâche visuelle que l'on veut quantifier par l'élaboration de la mesure
2. le bruit de fond et le type de signal à considérer, ainsi que la méthode permettant d'introduire le signal dans le bruit de fond
3. la mesure à appliquer aux images
4. la figure de mérite à dériver à partir de cette mesure
5. le moyen de comparer cette figure de mérite aux performances humaines

Dans [7] (pages 593-629), on trouvera les détails de chacun de ces éléments à définir afin de construire une mesure de visibilité adaptée. Dans notre cas, on peut construire notre méthode de mesure en considérant que :

1. on désire traiter le cas de la tâche de détection d'un signal dans du bruit. Le signal peut soit être présent, soit absent, et peut apparaître à des positions aléatoires dans l'image. Néanmoins, dans le cadre de cette étude, nous avons considéré des vignettes de petite taille, approximativement centrées sur le défaut éventuel par la procédure d'extraction. On peut donc dire que notre tâche est de détecter un signal au centre des vignettes qui seront présentées. Ceci mime l'inspection visuelle d'une image, l'expert étant visuellement attiré successivement sur des zones différentes de l'image. Par cette approximation, on ne tient pas compte des phénomènes de masquages visuels pouvant modifier la perception globale d'une image. En résumé, la tâche consiste à assigner à chaque défaut (ou à chaque vignette présentée) un scalaire proportionnel au degré de visibilité du défaut.

2. Le bruit de fond et le signal à détecter sont connus, et déterminés par les étapes précédentes de la chaîne d'analyse de la qualité (voir chapitre 2 : méthode de localisation ou méthode de création du contexte). Les signaux à détecter sont très variables, car les défauts sont en pratique de nature très différente. Le fond peut être lui facilement synthétisé à partir de nos connaissances *a priori* des détecteurs.
3. La mesure objective de visibilité à appliquer est à déterminer.
4. La figure de mérite issue de cette mesure sera directement la sortie de la mesure de visibilité. En théorie, cette figure de mérite doit prendre en compte les propriétés statistiques des différentes réponses possibles de la mesure (les valeurs typiques en présence d'un artefact, en absence d'un artefact, et la variance de la mesure, voir [7], p639-643). Pour simplifier, nous laisserons en perspective la formation de cette figure de mérite, qui est une étape nécessaire afin de comparer les résultats de la mesure à des données subjectives.
5. Les moyens de comparer les performances de la mesure à des données subjectives est la dernière étape de validation d'une mesure objective. C'est également l'étape la plus longue et la plus coûteuse. De premières données d'évaluation subjective ont été recueillies. Néanmoins, la maturité insuffisante de cette partie fait que la validation de la mesure sera également citée en perspective de futurs développements.

La principale difficulté repose en la variabilité du signal à détecter. En effet, l'utilisation pratique d'une mesure de visibilité nécessite qu'elle soit suffisamment générique pour être appliquée avec succès à la variété de défauts pouvant intervenir dans notre contexte industriel. Parallèlement, cette universalité de la mesure permet aux utilisateurs de former mentalement une échelle de valeur cohérente de la mesure, constituant ainsi un vecteur puissant de communication technique sur la qualité des images. Actuellement, la mesure classique implémentée, qui va servir de premier exemple dans ce chapitre, souffre d'une faible généralité.

5.2 Mesures objectives : éléments bibliographiques

L'étude, et parfois la quête, d'une mesure objective de qualité des images a suscité une quantité phénoménale de travaux, et on peut donc trouver dans la littérature de nombreuses références, traitant de l'élaboration d'une telle mesure, dans des contextes variés. Les éléments bibliographiques donnés ici sont donc loin d'être exhaustifs. Afin de disposer des bases nécessaires à la construction d'une mesure de visibilité, nous allons donner quelques éléments qui nous ont semblé les plus pertinents dans le cadre de la mesure de visibilité.

On peut séparer les mesures objectives de qualité en trois grandes catégories.

La première regroupe les méthodes permettant de mesurer une dégradation de qualité sur une image dont la version non dégradée est connue, ou une dégradation de la qualité due à la présence d'artefacts spécifiques. Il s'agit là de mesure de qualité au sens général. La seconde catégorie est constituée des mesures s'appuyant sur les attributs statistiques des artefacts présents afin de mesurer leur visibilité. Il s'agit donc là de méthodes plus susceptibles de répondre à notre besoin puisqu'elles mesurent non pas une dégradation de la qualité, mais un degré de visibilité d'objets particuliers. Ces mesures ne nécessitent en général pas de connaissance *a priori* forte, contrairement aux méthodes de la première catégorie. Enfin, la troisième catégorie pouvant être formée regroupe les méthodes s'appuyant sur des attributs fréquentiels des objets à mesurer.

5.2.1 Mesures classiques d'artefacts

L'élaboration d'une mesure objective de la qualité a été, au moins à ses débuts, principalement guidée par les développements des communications numériques, et en particulier l'avènement des algorithmes de compression. En effet, afin de qualifier la qualité d'un canal de transmission, ou la qualité d'un algorithme de compression, il est souvent nécessaire de disposer d'une mesure objective qu'on puisse facilement appliquer au signal ou aux images transmises. De nombreuses méthodes ont donc été proposées.

La majorité des méthodes de ce champ d'application appartient à la classe de méthodes dites «Full Reference»(FR) ou «Reduced Reference »(RR). Ce sont en effet des méthodes qui nécessitent :

- soit une référence complète, qui est l'image non dégradée par la transmission ou la compression, pour pouvoir mesurer la qualité de l'image dégradée,
- soit une référence réduite, représentée par exemple par la connaissance de la variété d'artefacts présents dans l'image à mesurer, adjointe de la connaissance de l'impact de chaque artefact sur la qualité globale de l'image.

Les mesures les plus répandues sont l'erreur quadratique moyenne, ou MSE (Mean Square Error), et sa dérivée, le PSNR (Peak Signal to Noise Ratio). Ces deux mesures nécessitent la connaissance d'une référence complète pour calculer la dégradation. Soit $R(x, y)$ l'image de référence non dégradée, de taille $N_x \times N_y$, et $D(x, y)$ l'image dégradée :

$$MSE = \frac{1}{N_x \cdot N_y} \sum_x \sum_y |R(x, y) - D(x, y)|^2 \quad (5.1)$$

$$PSNR = 10 \log_{10} \frac{a^2}{MSE} \quad (5.2)$$

où a est la valeur maximale des pixels (si les images sont codées sur 8 bits, on aura $a = 255$). Ces deux mesures sont très largement utilisées, en dépit de leurs limitations importantes, grâce à leur simplicité de calcul et à leur simplicité d'interprétation physique. Néanmoins, on peut facilement générer des images dont les dégradations visuelles sont très différentes, et à même erreur quadratique moyenne (voir par exemple [3, 85]).

Ces deux premières mesures, assez générales, ont été complétées par un grand nombre d'autres particulièrement adaptées pour mesurer des artefacts classiques de compression. Ainsi, dans [86, 64] on propose des méthodes de mesure des deux artefacts de compression les plus courants : le flou et l'artefact en anneaux («ringing»). Le flou est ainsi mesuré sans référence, alors que la mesure de l'artefact en anneaux nécessite la connaissance de l'image d'origine, avant compression. Les méthodes de mesure sont confrontées avec succès à des résultats d'évaluation subjective.

Afin de prendre en compte une variété d'artefacts plus importante, des mesures plus générales ont été développées. On en citera trois qu'on retrouve assez fréquemment dans la littérature : la métrique hybride normalisée (NHIQM : Normalized Hybrid Image Quality Metric), la mesure de qualité à référence réduite (RRIQA : Reduced-Reference Image Quality Assessment) et l'index de similarité structurel (SSIM : Structural Similarity Index Metric).

5.2.1.1 NHIQM [52, 26]

Cette mesure utilise la quantification d'attributs spatiaux classiques f_n , parmi lesquels l'effet de bloc, l'effet de flou, le masquage, l'activité basée sur le gradient, *etc.* Ces attributs sont pondérés puis sommés :

$$NHIQM = \sum_{n=1}^5 w_n \cdot f_n \quad (5.3)$$

Les poids w_n sont déterminés afin de régler l'impact de chaque attribut sur la mesure globale. En général, ces poids sont réglés pour chaque application, à l'aide d'étude subjective de la qualité. Ceci représente d'ailleurs le principal défaut de cette méthode, qui repose entièrement sur la détermination des poids : ceci nécessite la mise en place d'un protocole long et complexe d'évaluation subjective. De plus, les attributs sélectionnés ne sont pas forcément pertinents dans toutes les applications. Dans notre cas, il est inutile d'essayer de dériver une bonne mesure à partir d'attributs comme l'effet de bloc ou l'activité du gradient. D'ailleurs, ces attributs nécessitent la connaissance de la référence (attributs de type FR).

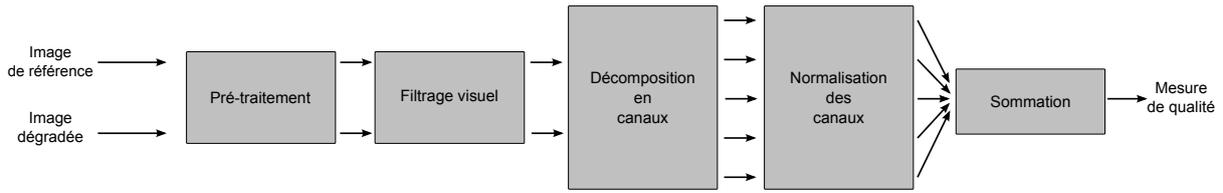


FIGURE 5.1 – Schéma de principe général d’une mesure de qualité par décomposition en canaux : après avoir pré-traité et appliqué des filtrages mimant le système visuel humain, les images sont décomposées en canaux. L’information de chaque canal est sommée, après normalisation, afin d’obtenir la mesure finale.

5.2.1.2 RRIQA [87]

La mesure RRIQA est une mesure particulière dans le sens où elle ne nécessite pas la connaissance d’une référence complète. Par contre, il faut connaître certaines propriétés statistiques entre les deux images afin de déterminer leur divergence de Kullback-Leibler (voir section 5.2.2). La mesure utilise une décomposition en ondelettes des images, et les calculs sont effectués dans chaque sous-bande de cette décomposition :

$$RRIQA = \log_2 \left(1 + \frac{1}{C_0} \sum_{k=1}^K |d^k(R^k, D^k)| \right) \quad (5.4)$$

où C_0 est une constante et $d^k(R^k, D^k)$ est une estimation de la distance de Kullback-Leibler sur la sous-bande k de l’image de référence R et de l’image dégradée D . On voit que la connaissance de cette distance dans chaque sous-bande n’est pas aisée. Néanmoins, dans certains cas, on peut la déterminer à partir des moments statistiques des deux images, et il peut être alors assez facile de la calculer.

5.2.1.3 SSIM [85]

La mesure de similarité structurelle est une mesure de type statistique simple s’appuyant sur les moments statistiques d’ordre 1 et 2 des deux images. Cette mesure tente de capter les changements de structures spatiales [84] dans les images à travers une mesure de luminosité, de contraste et de corrélation. Une expression simple de la SSIM est alors donnée par :

$$SSIM = \frac{(2\mu_R\mu_D + C_1)(2\sigma_{RD} + C_2)}{(\mu_R^2 + \mu_D^2 + C_1)(\sigma_R^2 + \sigma_D^2 + C_2)} \quad (5.5)$$

où μ et σ sont les moyennes et écarts-types des images. On a donc noté μ_R et σ_R la moyenne et l’écart-type de l’image de référence et μ_D et σ_D la moyenne et l’écart-type de l’image dégradée. σ_{RD} est la covariance¹ entre l’image de référence et l’image dégradée.

Les constantes C_1 et C_2 sont introduites pour des raisons de stabilité numérique de la mesure lorsque μ ou σ sont proches de zéro. Le SSIM a été confronté au MSE et à d’autres mesures afin d’approcher les résultats d’une étude subjective sur la qualité d’images compressées [87]. Il a été montré que le SSIM donne les meilleurs résultats. Le calcul de cette mesure nécessite néanmoins le calcul du coefficient de corrélation σ_{RD} , ce qui peut représenter un frein à son application.

En plus de ces mesures classiques, un certain nombre d’études ont été menées afin de sommer plusieurs de ces attributs pour en déduire une seule mesure. Dans [27], cette sommation est effectuée avec en plus une distinction de zones d’intérêt : les images à évaluer sont divisées en une zone définie comme celle attirant l’attention de l’observateur, et une zone de fond, dans laquelle les dégradations sont moins perceptibles.

D’une manière plus générale, le principe commun à beaucoup de mesures de visibilité est celui illustré par la figure 5.1.

1. $\sigma_{RD} = \mathbf{E}\{(R - \mu_R) \cdot (D - \mu_D)\}$

De nombreuses mesures de qualité suivent ce schéma général, où des aspects liés au fonctionnement du système visuel humain interviennent. La mesure est formée en cinq étapes :

1. un prétraitement des images en entrée est effectué. Dans cet exemple, on s'est placé dans le cas où on veut quantifier une détérioration apparue sur une image dont on connaît la version de référence, non dégradée. Le prétraitement contient en général des opérations purement numériques (conversion de types de données, réglage de la dynamique en cohérence avec le contraste affiché) mais également des opérations liées à la chaîne de vision. Par exemple, on peut appliquer un filtrage passe-bas sensé mimer le passage par le système optique de l'œil.
2. L'opération de filtrage visuel consiste en général à appliquer la fonction de sensibilité en contraste de l'œil. En effet, l'œil est différemment sensible aux contrastes des objets présents, en fonction de leurs fréquences spatiales. Cette fonction est connue et modélisée (voir par exemple [6]). Cette opération peut néanmoins intervenir à la quatrième étape, en tant que normalisation, après la décomposition en canaux.
3. La décomposition en canaux permet de séparer l'information en canaux d'orientation et de fréquence différentes. Ces canaux sont reconnus comme étant liés aux mécanismes des réponses neuronales de la cortex primaire du système visuel. La base de fonctions choisie pour cette décomposition peut être constituée de gaborettes [88], de différences de gaussiennes [92], etc.
4. Normalisation des canaux : à cette étape, on peut appliquer la fonction de sensibilité en contraste de l'œil, comme nous l'avons vu plus tôt, mais également des mécanismes de masquage entre canaux. En effet, une forte réponse dans certains canaux peut occasionner le masquage des fréquences spatiales proches. C'est à cette étape qu'on peut appliquer ces fonctions de masquage, en tirant partie de la décomposition en canaux. D'autres opérateurs non-linéaires de type saturation peuvent également avoir leur place ici.
5. Somme : la somme est réalisée en général à l'aide d'un opérateur de fusion de type norme de Minkowski²

La mesure implémentée dans l'outil DCTune [89, 90] suit ce principe, et incorpore les mécanismes de masquage en luminosité et en contraste. Cette mesure est fréquemment utilisée afin d'optimiser les paramètres de compression JPEG.

Enfin, un certain nombre d'essais ont été menés afin de pouvoir mesurer la qualité d'une image sans connaître l'image de référence. Même si dans le domaine de la compression ces techniques sont très minoritaires, elles peuvent représenter un intérêt dans notre situation. En effet, si l'étape de détection n'a pas fonctionné, on peut vouloir quand même tenter de mesurer la visibilité d'un artefact, sans connaître sa position et les pixels qui le composent. Parmi ces techniques, on citera notamment Sheikh et Bovik [80], qui modélisent les statistiques des scènes naturelles dans un espace d'ondelettes afin de mesurer les dégradations apportées par des algorithmes de compression opérant dans cet espace, comme la compression JPEG2000.

5.2.2 Les mesures statistiques

Une autre catégorie de mesures de qualité peut être formée à partir de toutes les mesures dites statistiques. Ces mesures s'appuient uniquement sur les histogrammes ou les moments statistiques de zones de l'image afin de mesurer une distance entre ces zones et le fond. Ces mesures sont également utilisées afin de suivre des changements temporels, notamment en imagerie spatiale ou SAR, de zones bien définies.

L'utilisation du ratio de Fisher et le rapport signal à bruit définis au début du chapitre 4 (section 4.1.1) nécessite la connaissance des deux ensembles à comparer. Dans notre cas, on veut comparer le fond de l'image, noté F , aux pixels définissant le défaut à mesurer, noté D .

Deux mesures statistiques communément utilisées sont la divergence de Kullback-Leibler et la distance de Bhattacharyya.

2. $m = \left(\sum_l \sum_k c_{l,k}^\beta \right)^{\frac{1}{\beta}}$ avec $c_{l,k}$ le k^{eme} coefficient du canal l et β un paramètre de la norme. Ce paramètre peut être réglé par accord de la mesure avec des données d'évaluation subjectives.

5.2.2.1 Divergence de Kullback-Leibler [49]

On donne ici la version symétrisée de la divergence de Kullback-Leibler (DKL), notée \mathcal{K} :

$$\mathcal{K} = \mathcal{D}(P_D \| P_F) + \mathcal{D}(P_F \| P_D) \quad (5.6)$$

où $\mathcal{D}(P_D \| P_F)$ est la *DKL* non symétrique :

$$\mathcal{D}(P_D \| P_F) = \sum_x P_D(x) \frac{P_D(x)}{P_F(x)} \quad (5.7)$$

P_F est la densité de probabilité des pixels du fond, et P_D est la densité de probabilité des pixels de défaut. La DKL est une mesure couramment utilisée en imagerie radar, donc en présence d'un fort bruit (speckle) pour mesurer des évolutions temporelles de scènes [44]. Même si sa corrélation avec les performances des modèles d'observateur est moins forte, d'après la littérature [35], que celle obtenue par la distance de Bhattacharyya, la divergence de Kullback-Leibler présente l'avantage de pouvoir être calculée de manière non-paramétrique, simplement à l'aide des moments d'ordre 1 à 4 des variables aléatoires F et D [58]. La DKL, dans son expression non-symétrique originelle (voir équation 5.7, est une mesure d'entropie mutuelle : la DKL entre D et F représente le nombre de bits nécessaires pour coder, par un codage entropique calculé sur F , les réalisations de D . Le principal défaut de cette mesure est son asymétrie. Ce problème peut être résolu très simplement en rajoutant la divergence réciproque, comme noté dans la formule de l'équation 5.6.

5.2.2.2 Distance de Bhattacharyya

La distance de Bhattacharyya \mathcal{B} présente, quant à elle, l'avantage d'être invariante par transformation bijective sur F et D , ce qui n'est pas le cas par exemple du ratio de Fisher. On a :

$$\mathcal{B} = -\ln \sqrt{\sum_x P_D(x) \cdot P_F(x)} \quad (5.8)$$

Dans [4, 35], il est notamment montré que la distance de Bhattacharyya est une bonne mesure de contraste entre D et F . En effet, on constate que la relation entre les performances d'un modèle d'observateur, mesurée par une aire sous courbe ROC ou un taux de mauvaise classification, et la distance de Bhattacharyya, est bijective et identique quelles que soient les lois de D et F . En particulier, dans le cas d'un algorithme de détection à base de calcul de *maximum* de vraisemblance, il a été montré que la distance de Bhattacharyya est directement liée à l'aire sous courbe ROC par une relation analytique. Elle est donc un indicateur privilégié du contraste dans un contexte de détection (mesure de visibilité), comme le notre.

Des tests rapides ont été menés afin d'utiliser ces mesures statistiques sur les défauts synthétiques, dans un premier temps. Les résultats montrent une certaine faiblesse de ces mesures, qui repose sans doute sur l'absence de prise en compte de la configuration spatiale du défaut. En effet, si on suppose connus les pixels composant le défaut, on peut former différentes configurations spatiales menant à la même mesure de Bhattacharyya ou de DKL, mais possédant des visibilités bien différentes. Dans des configurations de type lignes défectueuses en particulier, les mesures n'ont pas donné satisfaction. Néanmoins, il faudrait étendre cette étude, en introduisant un paramètre spatial. On pourrait par exemple s'inspirer de ce qui a été fait avec la théorie de *gestalts* (voir page 89) afin d'isoler certaines configurations spatiales, et de mesurer, dans cette configuration, une distance entre les pixels de défaut et le fond à l'aide des mesures statistiques définies plus haut.

5.2.3 Les mesures de type «Model Observer»

Outre la compression d'image, d'autres champs d'application des mesures de qualité ont permis de nombreux développements : l'imagerie militaire et l'imagerie médicale. Dans l'imagerie médicale en particulier, avec l'augmentation du nombre de modalités d'acquisition d'images (scanner, tomographie, mammographie), il a été nécessaire de mettre en place une méthodologie complète d'évaluation de la qualité. Cette méthodologie est destinée à remplacer les campagnes d'évaluation subjective, nécessitant la mobilisation de plusieurs experts humains, et donc très coûteuses en temps. De plus, pour certaines applications où on cherche à optimiser des paramètres d'acquisition, ou de post-traitement, il devient impossible de recourir à de telles campagnes d'évaluation.

Un grand nombre de «Model Observer»(MO) ont ainsi été développés, remplaçant le plus fidèlement possible les campagnes d'évaluation subjective. Dans le cadre de ce chapitre, nous allons nous attarder principalement sur trois mesures de type MO, dont deux seront appliquées aux défauts synthétiques. Pour un panorama complet des autres MO, nous renvoyons aux chapitres 10 et 11 du «Handbook of Medical Imaging» [7]. Il faut noter que pour ces trois mesures, nous nous sommes placés dans l'hypothèse d'un signal connu, de transformée de Fourier $S(u, v)$. En effet, la mesure intervient après l'étape de détection : on dispose donc d'une estimation du signal à mesurer.

5.2.3.1 L'Hotelling Observer (HO)

La mesure du rapport signal à bruit issue de l'Hotelling Observer est celle correspondant à un filtrage adapté, normalisée par la densité spectrale du bruit [5, 13] :

$$HO_{SNR}^2 = \sum_{u,v} \frac{|S(u,v)|^2}{N(u,v)} \quad (5.9)$$

où $S(u, v)$ est la transformée de Fourier du signal à mesurer, et $N(u, v)$ est la densité spectrale de puissance du bruit de fond de l'image, dans lequel le signal est noyé.

Cette mesure requiert la connaissance complète du signal à mesurer (pour réaliser le filtrage adapté) ainsi que du bruit (pour effectuer la normalisation par la densité spectrale du bruit). Cette mesure est l'une des plus pertinentes en présence de bruit corrélé, car elle permet de tenir compte de ces corrélations pour extraire la mesure de visibilité du signal considéré. Néanmoins, son application est très délicate dans ce cas car il faut connaître la densité spectrale de l'image de fond. Si cette image est uniquement constituée de bruit, comme nous le considérerons dans l'application de cette mesure à la section suivante, il est possible de déterminer cette densité spectrale. Par contre, lorsque l'on ne dispose pas d'une modélisation complète de ce fond, et en particulier en présence d'un contenu clinique, la détermination de cette densité spectrale nécessite la création de nombreuses réalisations de l'image de fond afin d'en obtenir une estimation fiable.

5.2.3.2 L'Hotelling Observer non blanchisseur avec filtre de l'œil (NPWE)

Le NPWE est une modification de l'Hotelling Observer afin de prendre en compte la fonction de sensibilité en contraste de l'œil (CSF). En effet, des expériences physiologiques ont permis de déterminer cette fonction, mettant en évidence un pic de sensibilité de l'œil à la fréquence de $4 \text{ cycles.deg}^{-1}$. On appelle cette fonction E et f la fréquence radiale. $E(f)$ peut être approchée par [15] :

$$E(f) = f^{1,5} \exp -cf^{0,68} \quad (5.10)$$

Le facteur c est un scalaire à ajuster en fonction de la distance de visualisation, de la taille et de la résolution de l'écran de visualisation, afin d'obtenir une réponse maximale à $4 \text{ cycles.deg}^{-1}$.

On peut déterminer cette constante c en fonction de la position du pic de sensibilité f_0 . A une distance

d'observation d_0 , la constante c est alors déterminée par :

$$c = \frac{1,5}{0,68 \cdot \frac{f_0}{d_0 \cdot 180}^{0,68}} \quad (5.11)$$

f_0 étant exprimée en cycles.deg^{-1} , d_0 en mm et c en mm^2 .

Si il est reconnu que la fréquence radiale f_0 du pic de sensibilité de la fonction doit être pris à $f_0 = 4 \text{ cycles.deg}^{-1}$, la forme de la fonction $E(f)$ diffère selon les sources. Le modèle de fonction $E(f) = f^\alpha \exp -cf^\beta$ revient toujours, mais avec des valeurs des paramètres α et β variables [94, 13, 15]. Nous nous sommes fiés à la référence la plus récente en prenant $\alpha = 1,5$ et $\beta = 0,68$. La figure 5.2 montre les courbes de la CSF à différentes distances de visualisation (20 centimètres, 50 centimètres et un mètre). On a représenté, en trait vertical, la fréquence spatiale maximale affichable par l'écran. Cette fréquence dépend de la taille des pixels de l'écran (pitch). Ici, il s'agit d'un écran avec une taille de pixel de $0,29\text{mm}$. On observe que lorsque l'on s'éloigne de l'écran, le pic de sensibilité se décale vers les basses fréquences spatiales, ce qui est logique puisque ce pic est à $4 \text{ cycles;deg}^{-1}$ quelle que soit la distance de visualisation. Ce pic est également moins large à des distances de visualisation élevées.

Ce filtrage de l'œil est alors intégré au NPWE et conduit à la mesure suivante :

$$NPWE_{SNR}^2 = \frac{\left(\sum_{u,v} |S(u,v)|^2 |E(u,v)|^2 \right)^2}{\sum_{u,v} |S(u,v)|^2 |E(u,v)|^4 N(u,v)} \quad (5.12)$$

où $E(u,v)$ est le filtre de l'œil en coordonnées cartésiennes.³

Le signal à mesurer $S(u,v)$ est corrélé avec le filtre de l'œil, ainsi qu'avec le bruit filtré. On mesure le rapport signal à bruit de ce procédé par division de ces deux quantités. Comme son nom l'indique, la mesure ne compense pas les éventuelles corrélation du bruit et n'est donc pas optimale en présence d'un bruit non blanc.

5.2.3.3 Le Channelized Hotelling Observer (CHO)

La dernière mesure que nous détaillerons ici est celle dite du Channelized Hotelling Observer. Le processus intègre des aspects du système visuel humain en décomposant les signaux sur un nombre réduit de sous-bandes fréquentielles, les canaux. Les réponses dans chacun des canaux sont ensuite sommées.

Plusieurs fonctions sont possibles pour former les canaux du CHO. Nous nous attacherons ici à illustrer les canaux de type filtres de Gabor, couplés à la fonction de sensibilité de l'œil décrite auparavant. Les fonctions de Gabor sont couramment utilisées pour cette tâche [95, 15], grâce aux résultats d'études psycho-visuelles qui ont montré qu'une bande passante d'un octave en fréquence⁴ estime relativement bien le fonctionnement du système visuel humain : or, les filtres de Gabor permettent de contrôler assez facilement cette bande-passante. Couplées à une modulation par la CSF, ce mécanisme en canaux permet de construire un modèle simple du système visuel. La formule, en fréquence, d'un canal est donc donnée par :

$$cho(f) = \exp \left(-\frac{1}{2} \frac{(f - f_c)^2}{\sigma_{f_c}^2} \right) \cdot E(f) \quad (5.14)$$

avec f_c la fréquence centrale en cycles.deg^{-1} , $E(f)$ la CSF comme définie par l'équation 5.10, et $\sigma_{f_c} = \frac{f_c}{3\sqrt{2 \ln 2}}$. Nous avons ici calculé la valeur de σ_{f_c} de manière à respecter la bande passante d'un octave.

En formant plusieurs de ces canaux, à des fréquences centrales et des orientations variables (la littérature donne en général cinq ou six fréquences centrales, entre 1 et 16 cycles par degré, en puissance de

3. A l'aide de l'équation 5.10, on peut déduire $E(u,v)$:

$$E(u,v) = E \left(\sqrt{u^2 + v^2} \right) = \sqrt{(u^2 + v^2)}^{1,5} \exp \left\{ -c \left(\sqrt{u^2 + v^2} \right)^{0,68} \right\} \quad (5.13)$$

4. La bande passante, en octave, est le \log_2 du rapport entre les deux fréquence de coupure à mi hauteur

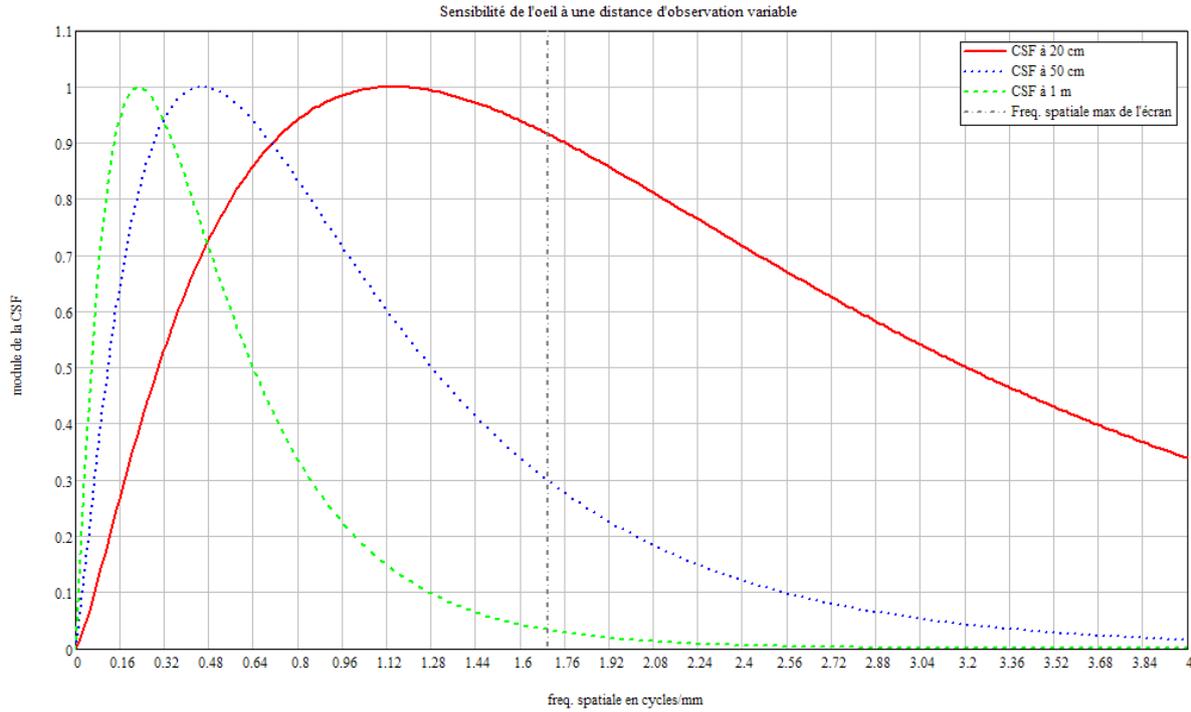


FIGURE 5.2 – Sensibilité de l’œil en fonction de la fréquence spatiale des objets et à différentes distances d’observation. Les 3 courbes ont été normalisées en amplitude afin de pouvoir être affichées sur le même graphique.

2, et huit orientations uniformément échantillonnées entre 0 et π , voir par exemple [29]), on crée le banc de filtres constituant les différents canaux, qu’on note C_k . La figure 5.3 montre la fonction de transfert de cinq canaux différents, pour des fréquences centrales variables. On observe que la bande passante est variable en fonction de cette fréquence centrale, ce qui est censé représenter une caractéristique du système visuel humain. La CSF est également représentée. Afin d’en tenir compte lors de la décomposition des canaux, on peut moduler l’amplitude de chaque canal par cette CSF. Il y a deux possibilité pour cela :

- modulation en amplitude : l’expression du canal est simplement modulée en amplitude par la CSF :

$$C_k^{mod}(f) = C_k(f) \cdot E(f) \quad (5.15)$$

- pondération du canal entier par l’amplitude de la CSF à la fréquence correspondant à la fréquence centrale du canal :

$$C_k^{mod}(f) = C_k(f) \cdot E(fc_k) \quad (5.16)$$

où fc_k est la fréquence centrale du canal k .

Nous avons choisi cette deuxième solution, qui se retrouve dans la littérature [15]. Une explication possible est que la CSF, qui traduit plutôt un phénomène optique (filtrage par l’optique de l’œil, conversion des photons en signal électrique par les photorécepteurs de la rétine, etc) intervient avant le mécanisme des canaux, qui modélise les mécanismes du cortex visuel primaire.

Le signal est alors filtré par les différents canaux :

$$V_k(u, v) = S(u, v) \cdot C_k(u, v) \quad (5.17)$$

Le rapport signal à bruit final est alors formé de la manière suivante :

$$CHO_{SNR}^2 = \sum_k \left\{ \frac{\sum_{u,v} |S(u, v) C_k(u, v)|^2}{\sum_{u,v} N(u, v) |C_k(u, v)|^2} \right\} \quad (5.18)$$

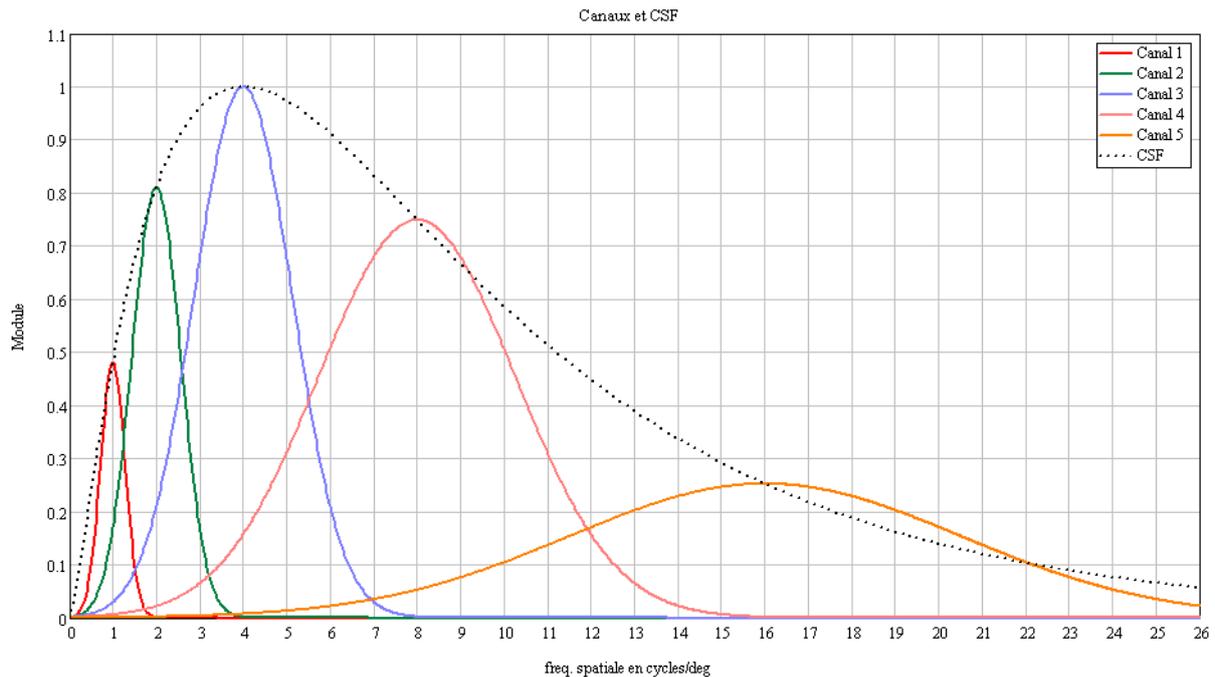


FIGURE 5.3 – Représentation des différents canaux pour cinq valeurs de fréquence centrale différentes : $f_c = 1; 2; 4; 8; 16 \text{ cycles.deg}^{-1}$. La CSF est également représentée.

L'intérêt que l'on peut voir dans cette méthode est qu'elle permet notamment une meilleure évaluation des défauts de type structure linéaire. En effet, ces structures seront bien capturées par un seul canal. La formule du CHO_{SNR} (equation 5.18) montre qu'il y a une intégration de l'énergie sur chacun des canaux, avant sommation finale. La structure linéaire va donc être intégrée sur le canal qui la capture, et ainsi son poids sera bien supérieur dans la sommation finale, comparée à la méthode de l'Hotelling Observer qui réalise une sommation sur toutes les fréquences spatiales, fréquence à fréquence.

Un autre intérêt de la méthode repose dans la possibilité d'évolutions facilitée par l'utilisation de la décomposition en canaux. En particulier, cette décomposition est particulièrement utilisée pour introduire des mécanismes non-linéaires du système visuel humain [95]. On peut également imaginer des méthodes utilisant une pondération différentes des canaux, en fonction des propriétés de visibilité de chacune des plages fréquentielles constituant ces canaux. En effet, à chaque canal correspond un filtre spatial dont l'impact visuel individuel peut être facilement évalué. On peut alors effectuer une pondération de chaque canal en fonction d'évaluations subjectives de la visibilité. Certaines variantes remplacent même la sommation finale sur les canaux, par un opérateur de *maximum* : seul le rapport signal à bruit maximal sur les canaux est utilisé pour former la mesure.

Faute de temps, nous n'avons pas pu appliquer le potentiel du CHO à la mesure de nos artefacts, mais il s'agit là d'une perspective envisageable à court terme.

5.3 Application de mesures de visibilité sur les défauts synthétiques

Afin de vérifier les performances de trois différentes mesures définies auparavant, nous allons appliquer ces mesures aux défauts synthétiques présentés dans le chapitre 4. Nous rappelons qu'il s'agit ici de défauts synthétiques dont on connaît parfaitement la localisation (ils ont été construits à partir de masques connus), et noyés dans un bruit synthétique blanc. Nous allons appliquer à ces défauts les mesures suivantes :

1. la mesure de référence, qui est une mesure d'énergie avec estimation du bruit de fond local (voir section suivante),
2. la mesure de l'Hotelling Observer (voir équation 5.9),
3. la mesure de l'Hotelling Observer non blanchisseur avec filtre de l'œil (voir équation 5.12).

Afin de comparer les résultats, nous allons former pour chaque défaut un échantillonnage de la mesure entre deux valeurs choisies à la main : une valeur faible réglée de manière à être en présence de défauts non visibles, et une valeur élevée correspondant à des défauts bien visibles à l'écran. Les différentes valeurs des mesures sont obtenues en réglant la puissance du bruit de fond.

5.3.1 Mesure de référence : énergie avec normalisation locale

En guise de mesure de référence, nous avons considéré la mesure existante à Trixell, notée M_{Trix} . Cette mesure est assez différente des deux mesures de type Hotelling Observer que nous avons également testées. En effet, la mesure de l'énergie ne nécessite aucune connaissance *a priori* sur le bruit ou sur le défaut à mesurer. La mesure s'effectue en trois étapes :

1. l'image à mesurer, dont on a préalablement retranché la valeur moyenne (cette opération peut être effectuée par estimation locale ou globale en fonction de la taille de l'image), est découpée en blocs B_k de 60×60 pixels semi-recouvrants. Sur chacun de ces blocs, on calcule l'énergie des pixels (somme des valeurs au carré des pixels),
2. normalisation : les valeurs des énergies sont normalisées par la valeur médiane des énergies calculées sur tous les blocs. Dans notre cas, comme on a appliqué cette mesure sur des images de petite taille, nous avons normalisé par une référence déterminée à partir de tirage aléatoire de bruit de fond ne contenant pas de défaut. Ceci permet de se placer dans un cas comparable à l'utilisation pratique, où la mesure est appliquée sur image de grande taille.
3. La mesure retournée est la valeur maximale des énergies normalisées.

On a finalement :

$$M_{Trix}^2 = \max_k \frac{\sum_{i,j} B_k(i,j)^2}{\text{median}_r (N_r(i,j)^2)} \quad (5.19)$$

où N_r sont les tirages aléatoires de bruit de fond et B_k les matrices issus du découpage en blocs de l'image à mesurer. Pour notre application, on a tiré 112 blocs de 60×60 pixels de bruit pour déterminer la valeur de normalisation.

Sur la figure 5.4 sont représentés les résultats de cette mesure appliquée aux défauts synthétiques. Cette présentation permet une vérification visuelle simple et rapide de l'efficacité de la mesure, bien qu'elle ne peut se substituer à des confrontations aux évaluations subjectives de la qualité (voir notamment [1]).

A partir de cette figure, on peut identifier deux problèmes principaux :

1. problème d'échelle de la mesure : comme nous venons de le voir, l'échelle de la mesure n'est pas adaptée aux défauts, et ce comportement est dépendant du type de défaut mesuré. En effet, ce manque d'adaptation se fait en particulier ressentir sur des défauts filiformes (défauts Image ou Lines par exemple) pour lesquels le défaut est présent dans tous les blocs. Dans cette configuration, l'impact du défaut sur la valeur du bloc est très faible, ce qui rend la mesure peu sensible.
2. Comportement différent de la mesure selon la nature du défaut : par exemple, à $M_{Trix} = 1.015$, on voit apparaître, certes assez légèrement, les défauts Square et Spot, mais pas du tout les défauts Highernoise et Oscillations. Ceci s'explique par le fait que la mesure capture assez bien la caractéristique principale de ces deux derniers défauts, qui sont des défauts possédant un niveau de variance locale plus élevée que le fond. Par contre, sur des défauts comme Lines et Spot, la mesure est inadéquate car elle ne peut mesurer la caractéristique de ces défauts, qui est une faible déviation de la valeur moyenne. Ainsi, pour des valeurs de la mesure similaires, la puissance de bruit est supérieure pour ces défauts, et donc ils sont très peu visibles.

Ces problèmes étaient attendus, étant donné que la mesure est surtout adaptée à des défauts à forte variance locale. Néanmoins, dans le cas pratique, cette mesure est utilisée fréquemment pour quantifier

l'impact visuel d'un défaut à Trixell, car les défauts étendus ou filiformes sont moins fréquents que les défauts présentant des structures à forte variance (type pixellisation). Le point fort de cette mesure est qu'elle ne nécessite d'aucune connaissance *a priori*, et que le bruit de fond servant de référence est estimé localement.

5.3.2 Mesure de l'Hotelling Observer (HO)

La mesure de l'Hotelling Observer, comme définie par l'équation 5.9 est appliquée aux défauts synthétiques. Le bruit de fond étant blanc, on a $N(u, v) = \frac{\sigma^2}{n^2}$, n^2 étant le nombre de pixels des images, et σ^2 la variance du bruit de fond. Dans ces conditions, la formulation de la mesure revient à un rapport signal à bruit classique, tel que celui défini au début du chapitre 4 pour caractériser les défauts synthétiques :

$$RSB(F, D) \propto \frac{(\mu_D - \mu_F)^2}{\sigma_F^2} \quad (5.20)$$

Les résultats obtenus pour cette mesure sont synthétisés sur la figure 5.5.

Cette vérification permet d'étudier deux comportements de la mesure :

- sa capacité à discriminer les défauts du bruit de fond
- sa faculté de généralisation à différents types de défauts

Sur les résultats de la figure 5.5, on peut noter que la mesure du rapport signal à bruit mesure assez mal les défauts Oscillations, Highernoise et Image. En particulier, si on observe la deuxième colonne de la figure, correspondant à une mesure de $HO_{SNR} = 6.6$, on voit que la mesure affecte la même valeur à des défauts visibles (Lines, Spot et Square) qu'à des images ne contenant visuellement que du bruit de fond. Ceci permet de dire que la discrimination des défauts par rapport au bruit de fond est mauvaise pour cette mesure. D'un point de vue général, on observe que le défaut Image est surévalué : il est mesuré avec des valeurs élevées, alors qu'il n'apparaît visuellement pas sur les images représentées. Il en est de même, dans une moindre mesure, pour les défauts Oscillations et Highernoise. Dans le même temps, on visualise clairement la visibilité croissante des défauts Lines, Square et Spot. Ce comportement hétérogène en fonction de la nature du défaut mesuré traduit une mauvaise généralisation de la mesure à des défauts de nature variable.

Ces résultats étaient attendus car la mesure prend en compte l'information spatiale du défaut uniquement dans son aspect global : l'information est sommée sur toute l'image. Ainsi, les défauts compacts (Spot et Square) ont une mesure sous estimée : même si statistiquement ils n'apparaissent pas très forts, leur nature spatiale (concentration sur une petite distance) les rend très visibles. *A contrario*, le défaut Image est, par les valeurs des pixels qui le composent, assez fort, mais ces pixels sont dispersés sur une grande surface. La mesure surestime sa visibilité. Néanmoins, les résultats sont globalement meilleurs comparés à la méthode précédente de calcul d'énergie avec normalisation locale.

5.3.3 L'Hotelling Observer non blanchisseur avec filtre de l'œil (NPWE)

On applique maintenant la mesure de l'Hotelling Observer non blanchisseur, mais avec filtrage de l'œil. Dans notre cas, le bruit étant blanc, l'absence d'une processus de blanchiment dans le calcul de la mesure ne pose pas problème. On a appliqué la formule de l'équation 5.12, avec une distance de visualisation $d_0 = 50cm$ pour appliquer le filtre de l'œil (équations 5.10 et 5.11). Il en résulte le filtre représenté sur la figure 5.6.

Les résultats de la mesure du NPWE sur les six défauts synthétiques sont représentés sur la figure 5.7.

On observe clairement les effets bénéfiques du filtre de l'œil, notamment sur le défaut Oscillations. Le filtrage des hautes fréquences a permis de réduire la surestimation de ce défaut, qui se rapproche maintenant, en termes de mesure, des défauts Spot et Square. L'effet est le même, certes dans une moindre mesure, sur les défauts Image et Highernoise. Néanmoins, ces défauts restent sous-estimés en visibilité par rapport aux défauts Spot et Square, qui ont eux aussi été filtrés par le filtre, mais dans sa plage basse-fréquence. Ils sont donc devenus légèrement plus visibles, à même niveau de la mesure.

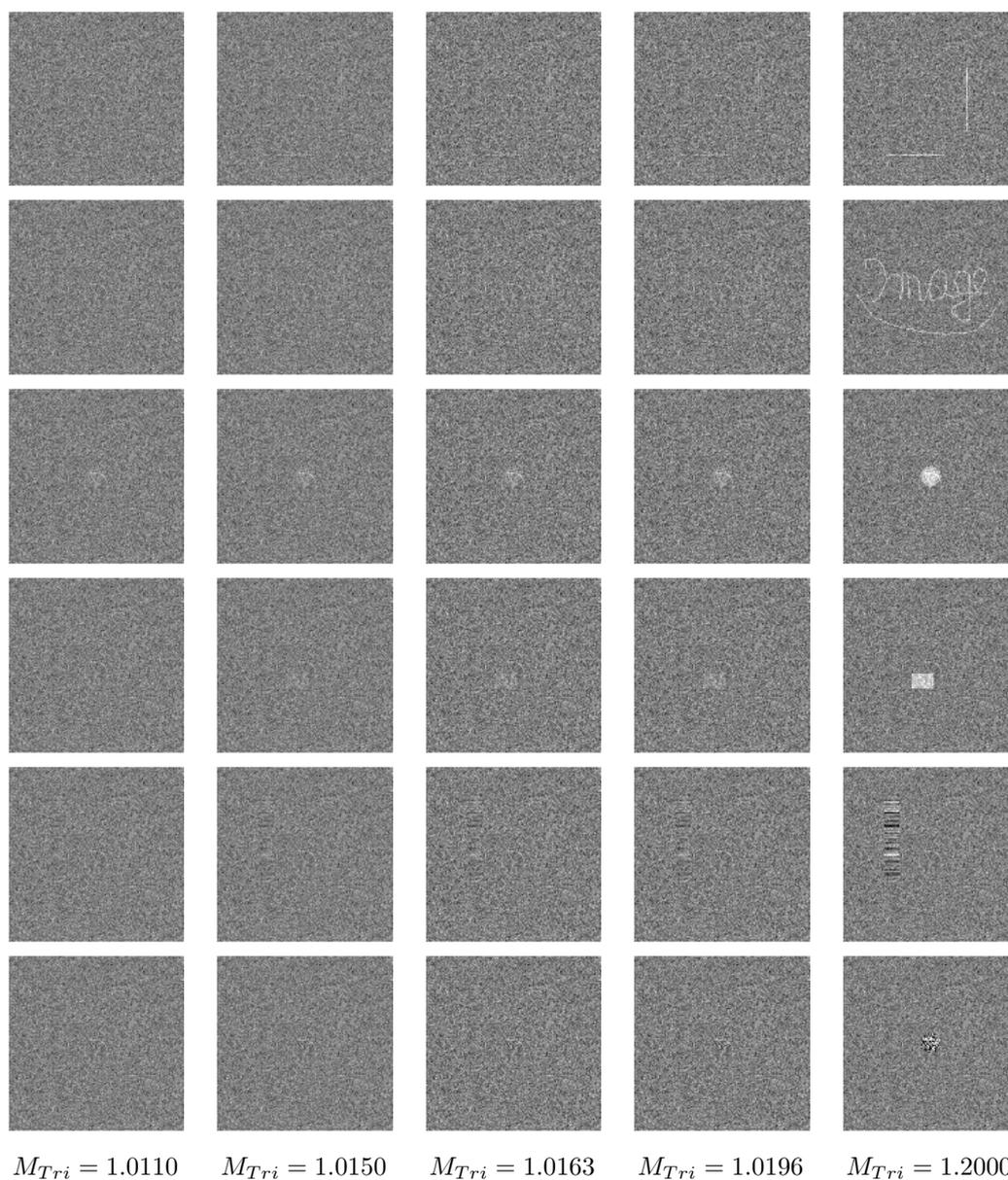


FIGURE 5.4 – Défauts synthétiques (de bas en haut : Lines, Image, Spot, Square, Oscillations et Highernoise) à mêmes valeurs de mesure de M_{Tri} . La puissance du bruit de fond est ajustée de manière à obtenir les valeurs de M_{Tri} fixées.

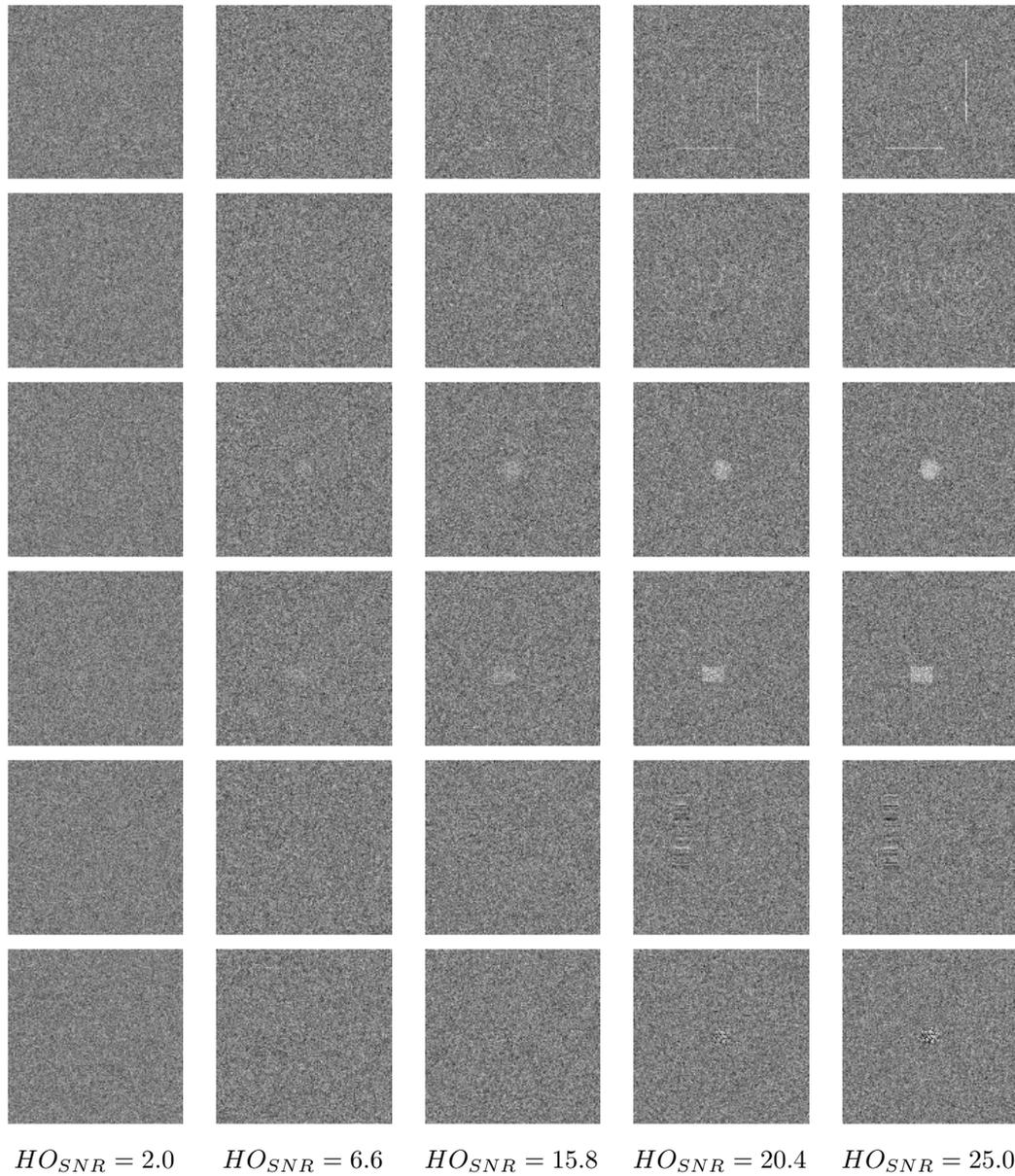
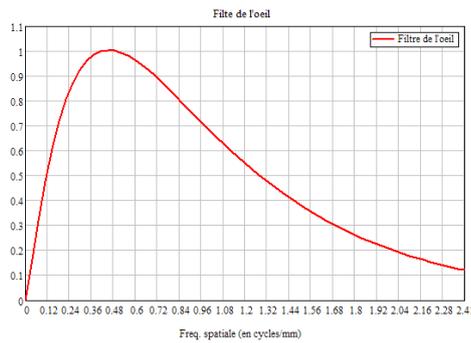
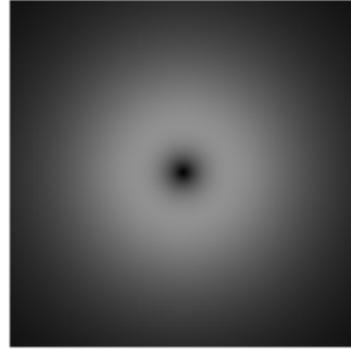


FIGURE 5.5 – Défauts synthétiques (de bas en haut : Lines, Image, Spot, Square, Oscillations et Highernoise) à mêmes valeurs de mesure de HO_{SNR} . La puissance du bruit de fond est ajustée de manière à obtenir les valeurs de HO_{SNR} fixées.



(a) Filtre de l'oeil (coupe radiale)



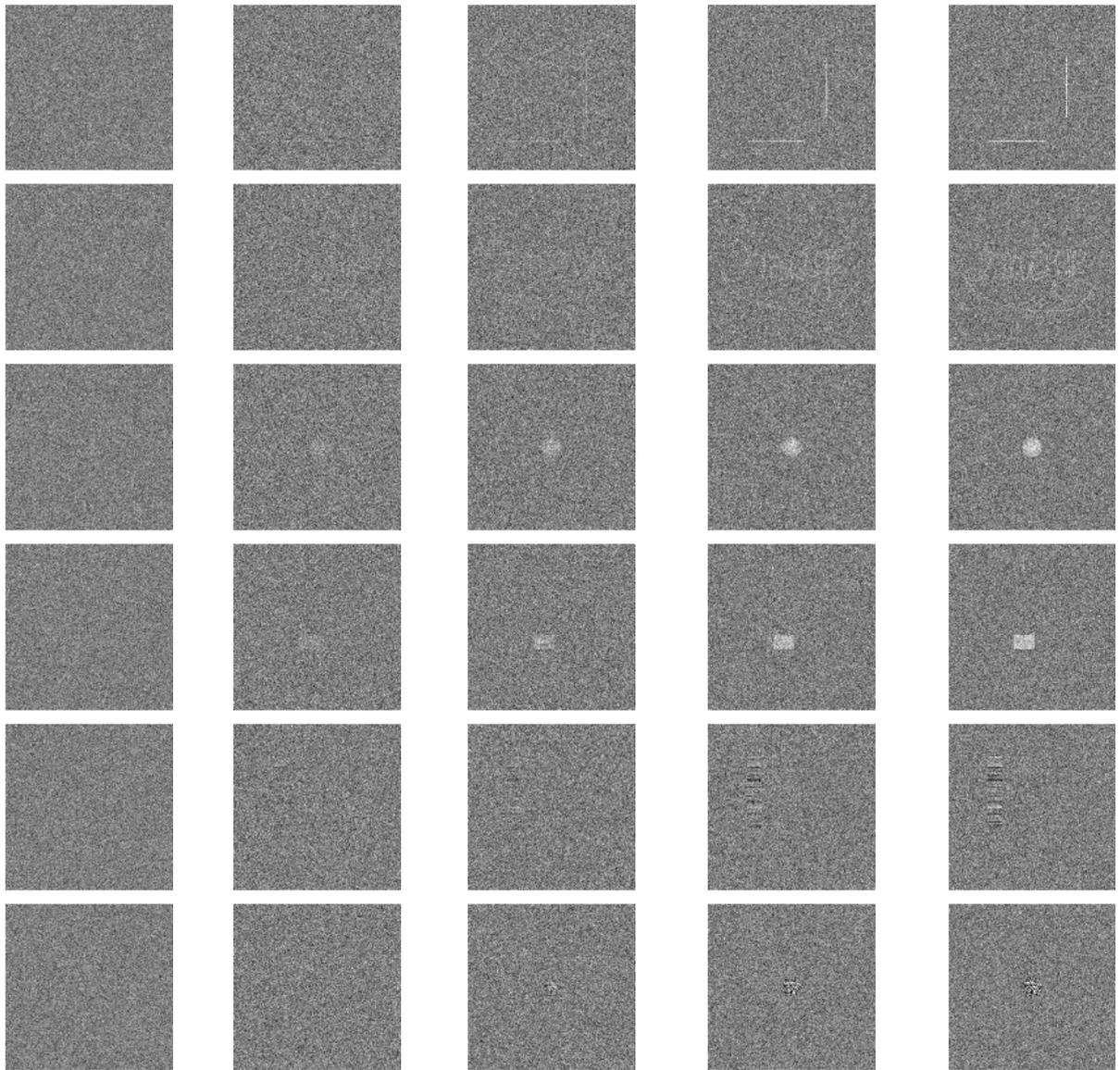
(b) Filtre de l'oeil (représentation fréquentielle)

FIGURE 5.6 – Filtre de l'oeil calculé à $d_0 = 50\text{cm}$: coupe radiale et transformée de Fourier du filtre (fréquence nulle au centre de l'image).

5.4 Conclusions

Dans cette partie, nous avons fait une revue de trois grandes classes de mesure de qualité, qui peuvent être utilisées comme mesure de visibilité, apparaissant dans la littérature. La quantité de travaux menés sur le sujet, et la grande diversité des méthodes proposées, reflète la complexité de la tâche, et surtout la grande difficulté à trouver une méthode de mesure pouvant s'appliquer à une large variété de défauts. Pour illustrer ceci, nous avons mené des tests de vérification de deux mesures issues des mesures de type Model Observer : l'Hotelling Observer et l'Hotelling Observer non blanchisseur avec filtre de l'oeil. La variété des défauts synthétiques suffit à mettre en échec ces deux mesures, dans un contexte pourtant idéal de bruit de fond blanc. Néanmoins, l'amélioration apportée par l'ajout du filtre de l'oeil est encourageante, et nous pousse à tenter d'incorporer des aspects psychovisuels dans l'élaboration d'une mesure. En ce sens, la mesure du Chanelized Hotelling Observer offre un cadre simple pratique, et est d'ailleurs la mesure de base de nombreuses études visant à enrichir les mesures de qualité avec des éléments mimant le système visuel humain.

L'introduction à la mesure de visibilité que nous avons donnée dans ce chapitre appelle naturellement à la conduite d'autres travaux. Dans un premier temps, on pourrait se placer dans un cadre restreint, en choisissant un type de défaut bien précis à mesurer, quitte à faire varier éventuellement le bruit de fond et le contexte. En effet, il est très difficile de mesurer avec une même méthode des défauts de natures différentes, d'où l'idée de mesurer d'abord correctement une certaine classe de défauts, pour ensuite essayer de généraliser cette mesure en optant par exemple pour l'utilisation de plusieurs mesures et leur fusion en un seul critère. Enfin, l'élaboration d'une telle mesure nécessite qu'elle soit confrontée à des données d'évaluation subjective de la visibilité. Des premières données ont été recueillies dans ce but. Leur exploitation, et éventuellement la définition d'une base de données de validation subjective plus large, constitue un élément important afin de confirmer l'efficacité des mesures développées.



$NPWE_{SNR} = 2.0$ $NPWE_{SNR} = 6.6$ $NPWE_{SNR} = 15.8$ $NPWE_{SNR} = 20.4$ $NPWE_{SNR} = 25.0$

FIGURE 5.7 – Défauts synthétiques (de bas en haut : Lines, Image, Spot, Square, Oscillations et Highernoise) à mêmes valeurs de mesure de $NPWE_{SNR}$. La puissance du bruit de fond est ajustée de manière à obtenir les valeurs de $NPWE_{SNR}$ fixées.

Conclusions et perspectives

Conclusions

Cette thèse constituait un projet ambitieux visant à obtenir des méthodes et des outils pour traiter la thématique globale de l'analyse de qualité d'images. En effet, Trixell produit des milliers de détecteurs numériques chaque année, et l'enjeu industriel est de contrôler les images issues de tous ces détecteurs afin de fournir une analyse fine des défauts pouvant apparaître aléatoirement sur ces images. La principale problématique de cette thèse était de mettre en place une chaîne d'analyse de qualité d'images, en s'appuyant sur la spécificité des images Trixell. Ces images sont généralement des images techniques, sans contenu clinique, où les seules structures apparaissant, en plus des éventuels défauts de fabrication, étant du bruit issu soit de la technologie des panneaux plats (bruit structuré), soit de la source de rayons X, soit de l'électronique de lecture.

Nous nous sommes ainsi placés dans ce cadre, et nous avons défini, dans le chapitre 2, une chaîne complète d'analyse de la qualité. Cette chaîne permet, à partir des images techniques issues des détecteurs plats, de :

1. retirer les structures de bruit issues de la technologie du détecteur : ces structures apparaissant de manière systématique, elles n'étaient pas l'objet premier de l'analyse de la qualité, et donc de cette thèse
2. détecter les défauts apparaissant aléatoirement dans ces images
3. donner une estimation de ces défauts à l'aide d'un débruitage de la zone défectueuse
4. mesurer l'impact visuel de chaque défaut dans un contexte de fond défini

Cette chaîne a été décrite dans le chapitre 2, et illustrée à l'aide d'exemples réels de défauts. Le principe de cette chaîne est de détecter et d'extraire chaque défaut, qui est ensuite traité indépendamment de l'image de laquelle il est issu. Cette extraction est une étape assez délicate car il faut capturer correctement le défaut, à l'échelle adéquate, afin de minimiser le volume de données capturées, mais de conserver un défaut visible à l'œil. Une méthode simple de capture a donc été proposée (page 36). Chaque défaut subit alors une étape de débruitage, destinée à estimer le plus finement possible les défauts, indépendamment du bruit de fond. Une méthode de débruitage, s'appuyant sur la nature spécifique des défauts rencontrés, a donc été développée, et testée avec succès sur de nombreuses images (page 40).

La chaîne ayant été appliquée à plusieurs images réelles nous avons pu, grâce à cette expérimentation, identifier un certain nombre de limitations. La principale porte sur le manque de sensibilité de l'étape de détection, qui a des difficultés à détecter les défauts de faible intensité, mais présentant des corrélations spatiales importantes.

Ainsi, dans le chapitre 4, nous nous sommes attachés à définir des méthodes de détection plus efficaces, utilisant notamment les aspects spatiaux des défauts en plus de leurs aspects purement statistiques. Après avoir décrit, dans le chapitre 3, la méthode de comparaison d'algorithmes de détection s'appuyant sur l'analyse ROC, développée au cours de cette thèse, nous avons donc pu tester un certain nombre de nouvelles méthodes de détection. A la suite de ces tests, il apparaît que certaines méthodes, en particulier la méthode de moyenne statistique pondérée (page 116), ou la méthode de filtrage directionnel (page 106), donnent des résultats encourageants sur des défauts synthétiques représentatifs de la difficulté de la tâche de détection.

Enfin, dans un dernier chapitre, nous avons décrit les possibilités offertes dans la littérature afin d'améliorer l'étape de mesure de visibilité des défauts. En effet, la mesure classique, s'appuyant sur un calcul d'énergie, se révèle insuffisante dans le cas de défauts possédant des corrélations spatiales fortes. Nous avons donc détaillé les différentes classes d'architecture de mesure objective de la qualité se retrouvant dans la littérature. Un éclairage particulier a été donné à une certaine classe de mesures, celles s'appuyant sur des modèles d'observateurs. Ces mesures sont très fréquemment utilisées en imagerie médicale afin d'évaluer les performances, en termes de détectabilité d'objets du point de vue du praticien, de différentes modalités d'acquisition d'images ou différentes conditions d'acquisition. Après avoir décrit trois mesures de bases s'appuyant sur de tels modèles, nous en avons testé deux sur défauts synthétiques. Les tests de la deuxième mesure sont encourageants. Les capacités de cette mesure à dissocier les défauts du bruit de fond, et à mesurer une large gamme de défauts, semblent être supérieures à celles de la

mesure de base, semblable à la première mesure testée. Des perspectives d'amélioration ont également été données, à travers l'utilisation d'un modèle décomposant les images sur différents canaux fréquentiels, censés être représentatifs du processus effectué par le système visuel humain.

Les principaux résultats de cette thèse reposent dans la définition de l'architecture globale d'une chaîne d'analyse de la qualité et la définition d'algorithmes de détection susceptibles d'améliorer significativement les performances en détection de la chaîne. D'un point de vue scientifique, ces algorithmes de détection, ainsi que la méthode développée afin de les comparer, constituent l'apport principal de cette thèse. Du côté industriel, le développement d'une chaîne complète d'analyse de la qualité, prenant en compte les spécificités des images générées dans l'entreprise, et d'une grande modularité, constitue un outil essentiel au développement de l'analyse qualité des images issues des détecteurs en production. Les développements annexes, comme la définition des formats de fichiers nécessaires pour l'archivage des défauts extraits par la chaîne, sont également des apports notables dans ce but.

Perspectives

Les travaux exposés dans cette thèse sont donc une première voie en vue du développement industriel de la chaîne complète d'analyse de qualité des images. Les aspects très transverses ouvrent à de nombreuses perspectives. Elles sont d'ordre industriel et d'ordre scientifique.

Du point de vue industriel, la suite logique des développements décrits dans ce manuscrit est naturellement la confrontation d'un certain nombre de méthodes à la réalité. En particulier, les méthodes de détection décrites au chapitre 4 doivent être testées en grande échelle et sur des cas réels. L'application de ces méthodes à des images de grande taille posera certainement le problème de l'échelle de détection. En effet, nous avons appliqué ces méthodes sur des défauts synthétiques, à petite échelle, mais dans le cas réel les défauts peuvent être de grande échelle, jusqu'à une extension de mille pixels. On pourrait alors réfléchir à l'intégration d'une stratégie rapide, comme celle développée avec la méthode basée ondelettes, qu'on couplerait alors avec la méthode de détection dans chacune des sous-bandes. Cette approche permet de garantir un temps de calcul et un volume de données à traiter raisonnable, mais souffre également du problème de la transformée en ondelettes classique : la variance par translation. Pour palier ce problème, il faudrait également étudier d'autres transformations, comme l'ondelette à trous. Dans la même idée, des transformées de type curvelets pourraient être utilisées, toujours dans l'idée de se doter d'un outil de détection multi-échelle tout en limitant la puissance de calcul nécessaire et la quantité de données à traiter, qui sont des contraintes industrielles importantes dans notre cas.

Toujours du point de vue industriel, il pourrait être utile de définir une méthode de classification des défauts. Plusieurs solutions peuvent être envisagées dans ce cas, mais il semble qu'une méthode semi-supervisée soit la plus appropriée. Il faudrait en effet être capable de définir des classes de défauts revenant de manière très fréquente, à l'aide d'attributs géométriques par exemple. Chaque nouveau défaut serait alors confronté à cette base de défauts connus, et associé à une classe. La difficulté résiderait ici dans la gestion de l'apparition d'une nouvelle classe de défaut. Une telle méthode de classification automatisée des défauts serait d'une grande valeur ajoutée du point de vue industriel.

Concernant la mesure de visibilité, les perspectives industrielles et scientifiques sont multiples. En effet, dans un premier temps il faudrait définir une méthode de validation de mesures par des données d'évaluation subjective. Ensuite, se pose le problème de la mesure, et en particulier de comment mesurer avec une seule méthode des défauts de nature très variable. Sur ce point, on peut légitimement imaginer l'utilisation de fusions de plusieurs critères de qualité, comme c'est d'ailleurs fait dans de nombreuses études publiées. Dans ce cadre, les données d'évaluations subjectives prennent toute leur importance car les paramètres de fusion devront certainement être optimisés de manière à ce que la mesure soit une évaluation fidèle à celle donnée par les observateurs humains. Pour le développement de la mesure elle-même, la voie naturelle est d'utiliser une architecture de type décomposition en canaux, qui permet de

contrôler assez facilement une diversité des structures à mesurer, et de leur affecter des poids différents sur la mesure de visibilité globale.

En conclusion, les perspectives d'amélioration de la chaîne d'analyse de la qualité sont très nombreuses. La multiplication des sources d'images numériques est un élément moteur pour le développement de techniques de traitement d'images, et notamment celles relatives à l'analyse de la qualité de ces images. Il est évident que bon nombre de ces techniques peuvent être utilisées à bon escient dans le monde industriel. La grande difficulté, comme souvent dans le domaine du traitement d'images, étant d'arriver à implémenter ces techniques et à les adapter aux spécificités et aux contraintes industrielles. Dans ce contexte, le pragmatisme est un adversaire très difficile à surpasser !

Quatrième partie

Annexes

Annexe A

Calculs des seuils théoriques pour un taux de fausse alarme fixé

Sommaire

A.1	Détermination de t_p	155
A.2	Détermination de t_m	156
A.3	Détermination de t_s	158

Ce chapitre détaille les calculs des seuils théoriques pour les méthodes de détection d'un pixel isolé, et d'un groupement de pixels par calcul de moyenne et de variance. On notera, par analogie avec le chapitre 2.3, les seuils des trois méthodes respectivement t_p , t_m et t_s . Les méthodes s'appuient sur le calcul d'une statistique qui sera seuillée pour réaliser la détection. Les statistiques des méthodes seront notées comme suit :

- pour la méthode de détection d'un pixel isolé (voir section 2.3.2.1), on appellera P la statistique calculée. On détectera ainsi un défaut dans l'image si $|P| \geq t_p$.
- pour la méthode de détection d'un groupement de pixels (voir section 2.3.2.2), deux statistiques sont calculées : M et S correspondant respectivement à un calcul de moyenne et de variance sur la ROI à tester. La ROI sera détectée si $|M| \geq t_m$ ou $|S| \geq t_s$.

Dans toute la suite de ce chapitre, on souhaite que chaque méthode possède le même taux de fausse alarme noté τfa . On va déterminer les seuils des trois méthodes pour ce taux de fausse alarme fixé. Ces calculs ont été faits dans notre cadre particulier, mais ils sont néanmoins assez classiques. Pour une approche plus générale de ces problèmes statistiques, on pourra se référer à [79].

A.1 Détermination de t_p

Comme dans le chapitre 2.3, on considère que les pixels $p_{i,j}$ de l'image sur laquelle est appliquée la détection sont spatialement indépendants, et suivent une loi gaussienne de moyenne μ et de variance σ^2 . La statistique P , pour la tâche de détection d'un pixel isolé, est simplement la valeur du pixel centrée et réduite, les paramètres de la loi que suivent les pixels étant calculés à chaque fois sur un voisinage du pixel traité. On détermine ainsi les estimations des deux paramètres de la loi $\hat{\mu}$ et $\hat{\sigma}$. Dans le cas général ces deux estimations dépendent du pixel (i, j) autour duquel elles sont calculées. Néanmoins, pour simplifier les calculs, dans le cadre de cette étude théorique, on ne tiendra pas compte de cette dépendance. On a donc $P = \left| \frac{p_{i,j} - \hat{\mu}}{\hat{\sigma}} \right|$ et $P \sim \mathcal{N}(0, 1)$. La densité de probabilité f_g de la loi gaussienne centrée réduite étant classiquement donnée par :

$$\begin{aligned} f_g : \mathbb{R} &\longrightarrow [0; 1] \\ x &\longmapsto \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} x^2 \right\} \end{aligned}$$

Si on appelle F_g la fonction de répartition de la loi gaussienne, et F_g^{-1} sa fonction réciproque :

$$\begin{aligned} F_g : x &\longmapsto \Pr(P \leq x) \\ F_g^{-1} : y &\longmapsto \arg \{ \Pr(P \leq x) = y \} \end{aligned}$$

On cherche alors le seuil t_p à appliquer à P tel que $\Pr(|P| \geq t_p) = \tau fa$. En tenant compte de la symétrie de la loi gaussienne, on sait que, pour tout $x \in \mathbb{R}$, $\Pr(P \leq -x) = \Pr(P \geq x)$. Ceci nous permet de choisir un unique seuil t_p afin de détecter des valeurs anormalement basses ou anormalement hautes de la statistique P . Ce seuil peut être calculé à l'aide de l'inverse de la fonction de répartition :

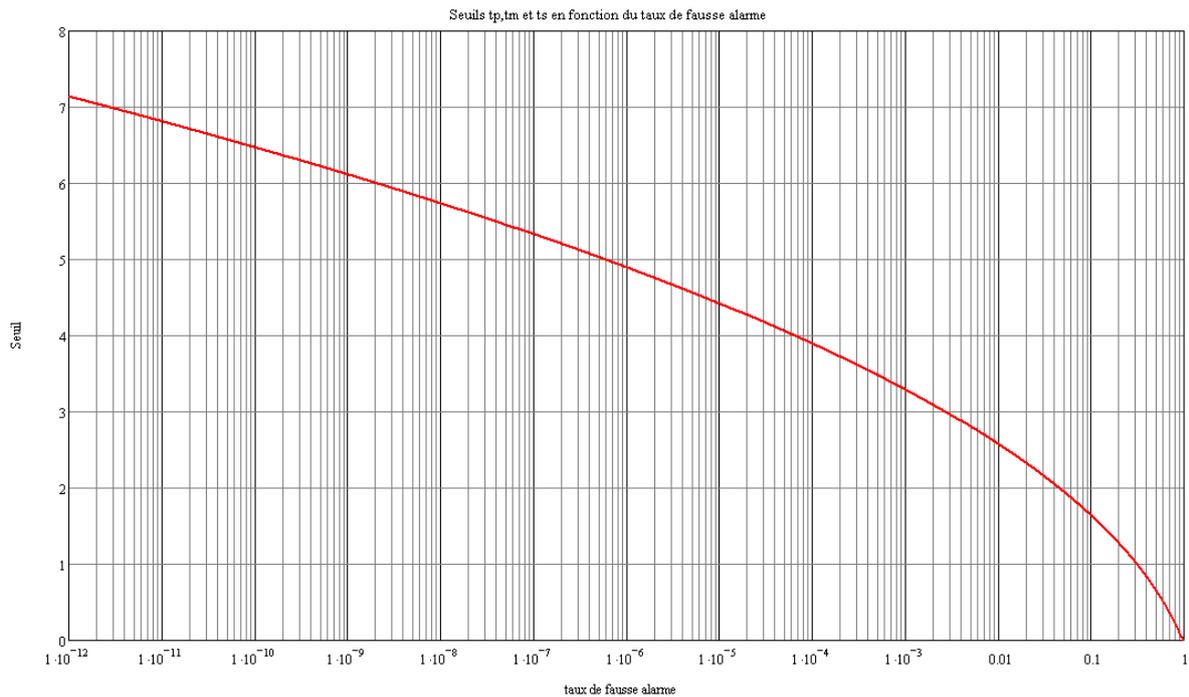
$$\Pr(P \geq t_p) + \Pr(P \leq -t_p) = \tau fa \quad (\text{A.1})$$

$$\Pr(P \leq -t_p) = \Pr(P \geq t_p) = \frac{\tau fa}{2} \quad (\text{A.2})$$

et donc

$$t_p = -F_g^{-1} \left(\frac{\tau fa}{2} \right) = F_g^{-1} \left(1 - \frac{\tau fa}{2} \right) \quad (\text{A.3})$$

En pratique, les valeurs de F^{-1} sont calculées par approximation [91, 2]. Le tableau A.1 et la figure A.1 donne les valeurs de t_p à appliquer en fonction du taux de fausse alarme. Pour la détection d'un pixel isolé, on utilise un seuil de $t_p = 6$ ce qui correspond à un taux de fausse alarme théorique de l'ordre de

FIGURE A.1 – Courbe du seuil à appliquer à la statistique P , M ou S en fonction du taux de fausse alarme τfa .

τfa	10^{-5}	10^{-6}	10^{-7}	10^{-8}	$1.973 \cdot 10^{-9}$	10^{-9}	10^{-10}	10^{-11}
t_p	4.417	4.892	5.327	5.731	6.000	6.109	6.467	6.807

TABLE A.1 – Seuils t_p à appliquer en fonction du taux de fausse alarme. En gras, le seuil utilisé en pratique.

10^{-9} : pour un milliard de pixels testés, nous obtiendrons en moyenne un pixel faussement détecté comme défectueux.

A.2 Détermination de t_m

En prenant toujours l'hypothèse de gaussianité et d'indépendance des pixels $p(i, j)$, $p(i, j) \sim \mathcal{N}(\mu, \sigma^2)$ on peut également calculer théoriquement le seuil t_m à appliquer à la statistique M . On rappelle que pour un pixel (i, j) et son voisinage $F_{i,j}$ de taille $N_F \times N_F$ pixels, on calcule pour chaque (i, j) la quantité $m_{F_{i,j}} = \frac{1}{N_F^2} \cdot \sum_{(m,n) \in F_{i,j}} p(m, n)$. On peut alors montrer que la statistique $M = \{m_{i,j}\}$ suit une loi gaussienne : en effet, les $m_{i,j}$ sont des combinaisons linéaires de variables aléatoires gaussiennes

indépendantes. De plus on a :

$$\begin{aligned}
 \mathbf{E}\{M\} &= \mathbf{E}\left\{\frac{1}{N_F^2} \cdot \sum_{(m,n) \in F_{i,j}} p(m,n)\right\} \\
 &= \frac{1}{N_F^2} \cdot \sum_{(m,n) \in F_{i,j}} \mathbf{E}\{p(m,n)\} \\
 &= \frac{1}{N_F^2} \cdot \sum_{(m,n) \in F_{i,j}} \mu \\
 &= \mu
 \end{aligned}$$

La variance de M s'écrit également en fonction des paramètre de la loi suivie par les pixels :

$$\begin{aligned}
 \text{Var}\{M\} &= \mathbf{E}\{(M - \mu)^2\} \approx \mathbf{E}\left\{\frac{1}{N_F^2} \cdot \sum_{(m,n) \in F_{i,j}} (p(m,n) - \mu)^2\right\} \\
 &= \frac{1}{N_F^4} \mathbf{E}\left\{\sum_{(m,n) \in F_{i,j}} (p(m,n) - \mu)^2\right\} \\
 &= \frac{1}{N_F^4} \cdot N_F^2 \sigma^2 \\
 &= \frac{1}{N_F^2} \cdot \sigma^2
 \end{aligned}$$

Finalement, on déduit la loi de M : $M \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{N_F^2}\right)$. Les paramètres de cette loi dépendent de la taille N_F de la fenêtre d'analyse $F_{i,j}$. Cette taille est paramétrée par l'utilisateur en fonction de l'échelle de détection qu'il souhaite traiter, en sachant que les défauts de taille similaire à celle choisie seront détectés de manière optimale. Dans la méthode de détection présentée dans la section 2.3.2.2, la moyenne et la variance de P est estimée par respectivement μ_m et σ_m .

Le seuil de détection est alors appliqué sur M de telle manière qu'un groupement de pixel centré en (i, j) , est détecté comme défectueux, si $\left|\frac{m_{i,j} - \mu_m}{\sigma_m}\right| \geq t_m$. Pour avoir une tache de détection possédant un taux de fausse alarme de τfa , il faut calculer le seuil t_m tel que :

$$\Pr\left(\left|\frac{M - \mu_m}{\sigma_m}\right| \geq t_m\right) = \tau fa$$

Comme M suit une loi gaussienne, de la même manière que la statistique P précédemment étudiée, on peut se placer dans le même cas. Les seuils de détection et les taux de fausse alarme (figure A.1 et table A.2) sont identiques et ne dépendent pas de la taille de la fenêtre d'analyse. Le seuil de détection utilisé pour paramétrer cet algorithme de détection a été déterminé empiriquement à $t_m = 4$, ce qui correspond à un taux de fausse alarme de l'ordre de 10^{-5} qui est significativement plus élevé que celui obtenu à l'aide du seuillage de la statistique P . Ceci traduit le fait que l'oeil est plus sensible à des groupements de pixels défectueux qu'à des pixels isolés, ce qui oblige à augmenter la sensibilité de la détection. En contrepartie, on obtient un taux de fausse alarme significativement supérieur. Néanmoins, la statistique utilisée est calculée sur des groupement de pixels : on obtiendra ainsi avec le seuil de $t_m = 4$ une fausse alarme toutes les 100000 fenêtres d'analyse traitées. Pour la méthode retenue, dans une image de taille 3002×3002 , il y a 183184 ROIs à analyser (découpage par blocs de taille 15×15 semi-recouvrants). On obtient alors en moyenne 11 fausses alarmes (qui sont des ROI de 15×15 pixels) par image analysée.

τfa	$6.334 \cdot 10^{-5}$	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}
t_p	4.000	4.417	4.892	5.327	5.731	6.109	6.467	6.807

TABLE A.2 – Seuils t_m à appliquer en fonction du taux de fausse alarme. En gras, le seuil utilisé en pratique.

A.3 Détermination de t_s

Le seuil t_s est appliqué sur la statistique $S = \{s_{F_{i,j}}\}$. On note $\mu_s = \mathbf{E}\{S\}$ et $\sigma_s^2 = \mathbf{E}\{(S - \mu_s)^2\}$. On suppose ainsi que la statistique S est stationnaire sur l'image analysée, et donc que ses paramètres ne dépendent pas de la fenêtre $F_{i,j}$ sur laquelle elle est calculée. Un groupement de pixels centré en (i, j) sera alors détecté si $\frac{|s_{F_{i,j}} - \mu_s|}{\sigma_s} \geq t_s$.

On a (estimateur sans biais de la variance) :

$$s_{F_{i,j}} = \frac{1}{N_F^2 - 1} \cdot \sum_{(m,n) \in F_{i,j}} (p(m,n) - m_F)^2 \approx \frac{1}{N_F^2} \cdot \sum_{(m,n) \in F_{i,j}} (p(m,n) - m_F)^2 \quad (\text{A.4})$$

Par hypothèse, on sait que les $p(m, n)$ sont indépendants et suivent une loi gaussienne de moyenne μ et de variance σ^2 .

Théorème A.1 (Fisher-Cochrane) Soit un vecteur aléatoire gaussien centré réduit $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, I_k)$. Soient q matrices symétriques définies non négatives, $\{A_i\}_{1 \leq i \leq q}$, de rangs r_1, \dots, r_q et telles que : $A_1 + \dots + A_q = I_k$ et $r_1 + \dots + r_q = k$ (dimension de l'espace vectoriel)

Alors :

les formes quadratiques $\mathbf{Z}_j = \mathbf{X}^\top A_j \mathbf{X}$ sont des variables aléatoires indépendantes qui suivent des lois de Pearson (aussi appelées lois du chi-deux) à r_j degrés de liberté

En appliquant le théorème de Fisher-Cochrane avec :

- $\mathbf{X} = (X_1 \dots X_k \dots X_{N_F^2})^\top$ avec $X_{mN_F+n} = \frac{p(m,n) - m_{F_{i,j}}}{\sigma}$, où $\mathbf{e} = (1 \ 1 \ \dots \ 1)^\top$ et $(m, n) \in F_{i,j}$.
- On a bien $\mathbf{X} \sim \mathcal{N}(0, 1)$ en prenant l'approximation $\mu \approx m_{F_{i,j}}$
- $A_1 = \frac{1}{N_F^2} \mathbf{e} \mathbf{e}^\top$ et $A_2 = I_{N_F^2} - \frac{1}{N_F} \mathbf{e} \mathbf{e}^\top$

on a bien $r_1 + r_2 = N_F^2$, et A_1 et A_2 des matrices définies non négatives (ce sont des matrices de projection de \mathbf{X} , respectivement sur le vecteur \mathbf{e} et sur le sous espace orthogonal à \mathbf{e}). On applique donc le théorème

τfa	10^{-5}	10^{-6}	5.73×10^{-7}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}
t_p	4.417	4.892	5.000	5.327	5.731	6.109	6.467	6.807

TABLE A.3 – Seuils t_s à appliquer en fonction du taux de fausse alarme. En gras, le seuil utilisé en pratique.

de Fisher-Cochrane à \mathbf{X} , on obtient alors notamment :

$$Z_2 = \mathbf{X}^\top A_2 \mathbf{X} = \mathbf{X}^\top \left(I_{N_F^2} - \frac{1}{N_F^2} \mathbf{e} \mathbf{e}^\top \right) \mathbf{X} \sim \chi^2(N_F^2 - 1) \quad (\text{A.5})$$

$$= \mathbf{X}^\top \mathbf{X} - \frac{1}{N_F^2} \mathbf{X}^\top \mathbf{e} \mathbf{e}^\top \mathbf{X} \quad (\text{A.6})$$

$$= \|\mathbf{X}\|^2 - \left(\frac{1}{N_F} \sum_{k=1}^{N_F^2} X_k \right)^2 \quad (\text{A.7})$$

$$= \sum_{k=1}^{N_F^2} \left(\frac{p(\alpha_k, \beta_k) - m_F}{\sigma} \right)^2 \text{ avec } k = \alpha_k N_F + \beta_k, (\alpha_k, \beta_k) \in \mathbb{N}^2, \beta_k < N_F \quad (\text{A.8})$$

$$= \frac{1}{\sigma^2} \sum_{k=1}^{N_F^2} (p(\alpha_k, \beta_k) - m_F)^2 \quad (\text{A.9})$$

$$= \frac{1}{\sigma^2} \sum_{(m,n) \in F_{i,j}} (p(m, n) - m_F)^2 \quad (\text{A.10})$$

$$= \frac{N_F^2}{\sigma^2} s_{F_{i,j}} \quad (\text{A.11})$$

On déduit alors que $\frac{N_F^2}{\sigma^2} S$ suit une loi du χ^2 à $N_F^2 - 1$ degrés de liberté. On a seulement $N_F^2 - 1$ degrés de liberté car la moyenne m_F doit être estimée à partir des données, ce qui enlève un degré de liberté (ceci se traduit par la projection sur l'espace orthogonal à \mathbf{e} à l'aide de la matrice de projection A_2).

Finalement, on peut dire que $S \sim \frac{\sigma^2}{N_F^2} \chi^2(N_F^2 - 1)$. En pratique, on a $N_F \gg 1$ (typiquement, on prend des tailles de ROIs variant entre $N_F = 15$ et $N_F = 60$). On peut donc approcher la loi du chi-deux obtenue pour S par une loi normale : $S \sim \frac{\sigma^2}{N_F^2 - 1} \mathcal{N}(N_F^2 - 1, 2(N_F^2 - 1)) \sim \frac{\sigma^2}{N_F^2} \mathcal{N}(N_F^2, 2N_F^2)$. On déduit donc de cette approximation que S est gaussienne, puisque elle peut s'écrire comme combinaison linéaire de variables gaussiennes.

Pour déterminer le seuil t_s , on est donc ramené aux cas précédents car $\frac{S - \mu_S}{\sigma_S} \sim \mathcal{N}(0, 1)$.

La courbe du choix du seuil t_s en fonction du taux de fausse alarme (voir figure A.1) est identique à celles utilisées pour fixer t_p et t_m .

Pour une utilisation pratique, nous avons choisi $t_s = 5$ ce qui correspond à un taux de fausse alarme théorique de $\tau fa = 5.73 \times 10^{-7}$. Avec un raisonnement identique à celui mené pour le réglage de t_m , sur une image de 3002×3002 pixels et un découpage en ROIs de 15×15 pixels, on obtient en moyenne 1 fausse alarme (qui est une ROI de 15×15 pixels par image traitée).

Annexe B

Etude psychovisuelle pour la fixation des seuils t_p , t_m et t_s

Sommaire

B.1	Choix du seuil t_p	163
B.2	Choix des seuils t_m et t_s	163

On remarque la présence des deux seuils t_m et t_s dans l'équation 2.19 qui permettent de paramétrer la sensibilité de détection. Comme pour le seuil t_p de l'équation 2.12, ces seuils peuvent être réglés de manière à assurer, sous l'hypothèse que $p_{i,j}$ suit un modèle gaussien, un taux de fausse alarme constant sur les trois détections (détection d'un pixel, détection d'une zone à valeur moyenne atypique, détection d'une zone à valeur de variance atypique). Le calcul théorique des trois seuils par cette méthode peut être trouvé en annexe A.

Néanmoins, dans cette section, nous avons préféré privilégier une approche empirique pour le réglage de ces seuils. En effet, la méthode de détection exposée vient en début de chaîne, et nous avons vu qu'elle doit avoir une bonne sensibilité car les défauts non détectés ici ne pourront plus l'être dans les étapes suivantes, et seront donc définitivement exclus de l'analyse de qualité de l'image. Dans ce contexte, il est alors préférable de régler les seuils de détection afin d'avoir la performance en terme de détection voulue, plutôt que de se fixer un taux de fausse alarme cible.

Afin de régler les trois seuils t_p , t_m et t_s , nous procédons donc par observation visuelle : le but est de déterminer le seuil se trouvant à la limite de visibilité des défauts détectés par chacune des méthodes. Dans cette section, nous allons, à partir d'observations faites sur des images simulées, décider empiriquement de la valeur des seuils à adopter.

B.1 Choix du seuil t_p

Le choix du seuil t_p a été obtenu de la manière suivante : on a simulé des images de bruit blanc gaussien, de variance σ^2 de taille 60×60 pixels, dans lesquels on a introduit un pixel atypique à une position aléatoire sur chaque image. Nous avons fait croître la valeur de ce pixel de moyenne de l'image $+ \sigma$ à moyenne de l'image $+ 8\sigma$. Nous avons ensuite calculé dix réalisations différentes du bruit (échantillons 1 à 10), créant ainsi $8 \times 10 = 80$ imajettes à évaluer. Une série d'imajettes pour une seule réalisation de bruit est représentée sur la figure B.1.

Chacune de ces imajettes a été évaluée par quatre experts humains. Il leur a été demandé d'identifier les imajettes où ils voyaient un pixel défectueux. Ces experts industriels sont initiés au problème de détection de défauts dans notre contexte, et savent donc à partir de quel seuil un pixel peut être considéré comme défectueux.

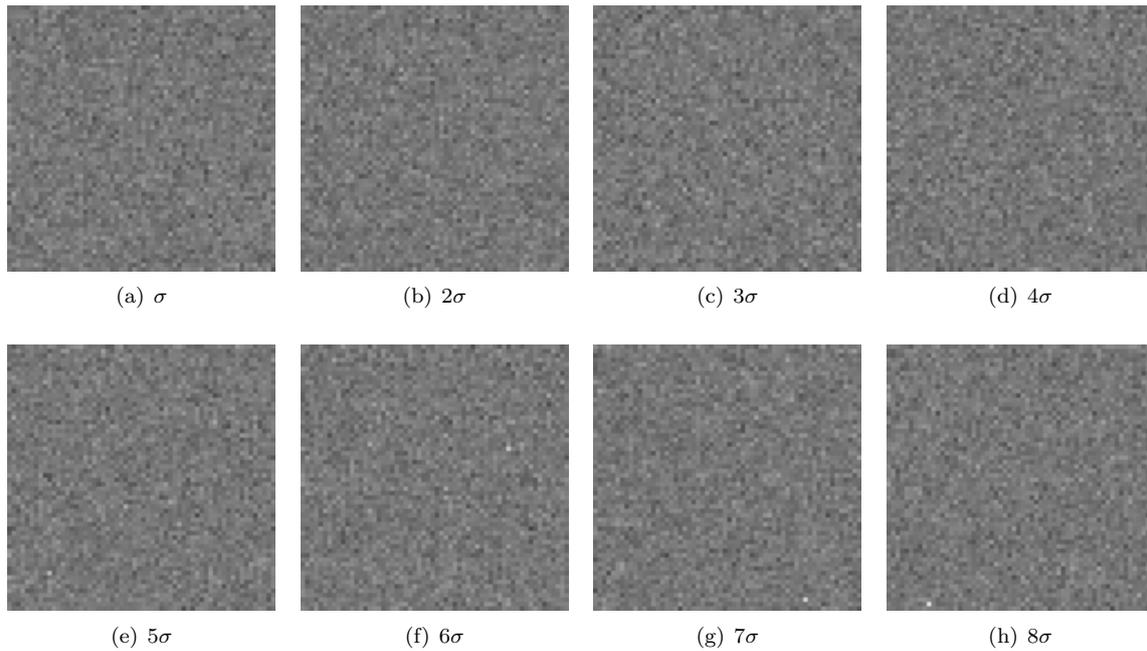
Le contraste des images visualisées a été réglé de manière à ce que toute la dynamique de l'image soit visualisable sur les 256 niveaux de gris affichés par l'écran.

Les tables B.1 et B.2 représentent les résultats de ce test effectué par quatre observateurs humains, sur dix réalisations de bruit différentes (la position du pixel défectueux a été changée à chaque réalisation). Pour chaque réalisation, lorsque la majorité des observateurs, c'est à dire au moins trois observateurs sur quatre, a indiqué la présence d'un pixel défectueux, on a considéré que le pixel inséré est un pixel à détecter par la méthode de détection. Ces pixels sont recensés dans la table B.1. On constate que tous les pixels ayant un niveau de 7σ au dessus du niveau moyen sont détectés, comme la grande majorité des pixels ayant un niveau de 6σ au dessus du niveau moyen. Le seul cas où ces derniers n'ont pas été détectés correspond à une configuration où le pixel défectueux a été placé au bord de l'image, ce qui l'a rendu difficilement détectable.

La table B.2 synthétise les résultats. Elle permet de voir, si on exclut le cas du pixel défectueux présent dans un bord de l'image, que dans tous les cas le pixel a été détecté à partir d'un seuil correspondant à $t_p = 6$. Dans un cas, un pixel à $t_p = 5$ a été détecté par les observateurs. A partir de cette étude, nous pouvons raisonnablement choisir $t_p = 6$, ce qui correspond à la grande majorité des défauts de type pixel isolé à détecter.

B.2 Choix des seuils t_m et t_s

Les seuils t_m et t_s sur la valeur moyenne et la valeur de variance ont été déterminés par la même méthode. Les résultats sont exposés dans les tableaux B.3 à B.6. Des illustrations des images à évaluer

FIGURE B.1 – Détermination empirique du seuil t_p : à partir de l'observation visuelle, on détermine $t_p = 6$.

Valeur pixel (en multiples de σ)	1	2	3	4	5	6	7	8
Échantillon 1						x	x	x
Échantillon 2						x	x	x
Échantillon 3						x	x	x
Échantillon 4					x	x	x	x
Échantillon 5						x	x	x
Échantillon 6							x	x
Échantillon 7						x	x	x
Échantillon 8						x	x	x
Échantillon 9						x	x	x
Échantillon 10						x	x	x

TABLE B.1 – Évaluation des pixels défectueux visibles sur 10 échantillons, en fonction de la valeur appliquée au pixel. Le symbole «x» représente les échantillons où la majorité des experts ont identifié la présence du pixel défectueux.

Valeur pixel (en multiples de $\sigma_{V_{i,j}}$)	1	2	3	4	5	6	7	8
Taux de détection (tous experts confondus) en %	0	0	0	5	40	87.5	95	100
Cas où les experts ont jugé le défaut visible (%)	0	0	0	0	10	90	100	100

TABLE B.2 – Résultats de l'expérience de détermination du seuil t_p . Sont représentées les valeurs appliquées au pixel défectueux, le taux de détection pour chacune de ces valeurs, et le nombre de cas où la réponse des experts est qu'un défaut est présent. La valeur du seuil choisie est $t_p = 6$.

Échelle k des valeurs de $m_{F_{i,j}}$ ($m_{F_{i,j}} = \mu_m^F + k\sigma_m^F$)	8	10	12	14	16	18	20	22
Échantillon 1		x		x	x	x	x	x
Échantillon 2					x	x	x	x
Échantillon 3						x	x	x
Échantillon 4				x		x	x	x
Échantillon 5					x	x	x	x
Échantillon 6					x	x	x	x
Échantillon 7				x	x	x	x	x
Échantillon 8						x	x	x
Échantillon 9		x			x	x	x	x
Échantillon 10					x	x	x	x

TABLE B.3 – Évaluation des groupements de pixels visibles (en valeur moyenne) sur 10 échantillons, en fonction de la déviation appliquée. Le symbole «x» représente les échantillons où la majorité des experts ont identifié la présence d'une zone défectueuse.

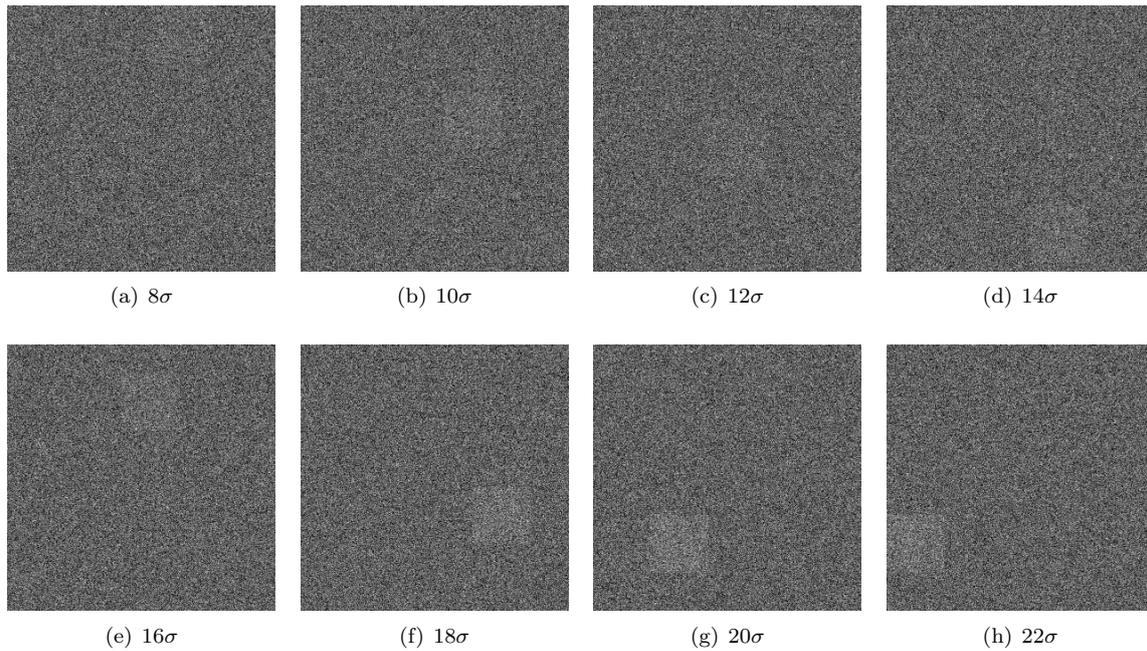
Échelle k des valeurs de $m_{F_{i,j}}$ ($m_{F_{i,j}} = \mu_m^F + k\sigma_m^F$)	8	10	12	14	16	18	20	22
Taux de détection (tous experts confondus) en %	5	25	5	40	75	97.5	100	97.5
Cas où les experts ont jugé le défaut visible (%)	0	20	0	30	70	100	100	100

TABLE B.4 – Résultats de l'expérience de détermination du seuil t_m . Sont représentées les valeurs appliquées à la zone défectueuse, le taux de détection pour chacune de ces valeurs, et le nombre de cas où la réponse des experts est qu'un défaut est présent. La valeur du seuil choisie est $t_m = 16$.

pour chacune des détections (détection sur la moyenne à partir de la statistique $m_{F_{i,j}}$ et détection sur la variance à partir de la statistique $s_{F_{i,j}}$) sont représentées sur les figures B.2 et B.3. Les résultats de cette évaluation nous pousse à choisir comme seuils $t_m = 16$. En effet on constate que dans la majorité des cas, les experts ont identifiés un défaut à partir de ce seuil. Certains échantillons ont été identifiés comme défectueux pour des seuils plus faibles, mais leur nombre très réduit nous pousse à dire qu'il s'agit plutôt d'une position du défaut présenté (notamment une position très centrale dans l'image) et une configuration de bruit particulière qui ont conduit à la détection. Dans le cas général, des défauts de niveaux plus bas ne seront pas détectés. Le seuil de 16 nous semble être un bon compromis.

Le même raisonnement nous conduit à choisir $t_s = 24$. Dans cette situation, le choix est encore plus clair puisqu'une seule image présentée, à un seuil inférieur, a conduit à l'identification d'un défaut par les experts. Ceci est clairement un cas atypique pour lequel la configuration du bruit et la position particulière du défaut ont été favorables à la détection à un seuil aussi bas. Dans le cas général, on observe que le seuil pertinent est plutôt $t_s = 24$.

Les seuils ainsi déterminés correspondent à une configuration typique de l'algorithme de détection, s'appuyant sur une courte étude du niveau pertinent de détection vis à vis de la tâche d'analyse de qualité. L'évaluation, faite sur une population d'observateurs connaissant le contexte de cette tâche permet d'accroître la pertinence des seuils choisis. Bien entendu, ce choix nécessiterait une étude complète, et un protocole de validation adapté. Néanmoins, dans ce chapitre où nous nous attachons à décrire des moyens simples d'implémenter la chaîne complète de détection, une étude sur les seuils de détection à adopter ne nous a pas semblé indispensable. Pour pousser plus loin cette évaluation, il faudrait approfondir les points suivants :

FIGURE B.2 – Détermination empirique du seuil t_m : à partir de l’observation visuelle, on a déterminé $t_m = 16$.

Échelle k des valeurs de $s_{F_{i,j}}$ ($s_{F_{i,j}} = \mu_s^F + k\sigma_s^F$)	16	18	20	22	24	26	28	30
Échantillon 1					x	x	x	x
Échantillon 2	x				x	x	x	x
Échantillon 3					x	x	x	x
Échantillon 4					x	x	x	x
Échantillon 5					x	x	x	x
Échantillon 6					x	x	x	x
Échantillon 7					x	x	x	x
Échantillon 8					x	x	x	x
Échantillon 9					x	x	x	x
Échantillon 10					x	x	x	x

TABLE B.5 – Évaluation des groupements de pixels visibles (en valeur de variance) sur 10 échantillons, en fonction de la déviation appliquée. Le symbole «x» représente les échantillons où la majorité des experts ont identifié la présence d’une zone défectueuse.

Échelle k des valeurs de $s_{F_{i,j}}$ ($s_{F_{i,j}} = \mu_s^F + k\sigma_s^F$)	16	18	20	22	24	26	28	30
Taux de détection (tous experts confondus) en %	20	12.5	25	7.5	97.5	100	97.5	100
Cas où les experts ont jugé le défaut visible (%)	10	0	0	0	100	100	100	100

TABLE B.6 – Résultats de l’expérience de détermination du seuil t_m . Sont représentées les valeurs appliquées à la zone défectueuse, le taux de détection pour chacune de ces valeurs, et le nombre de cas où la réponse des experts est qu’un défaut est présent. La valeur du seuil choisie est $t_s = 24$.

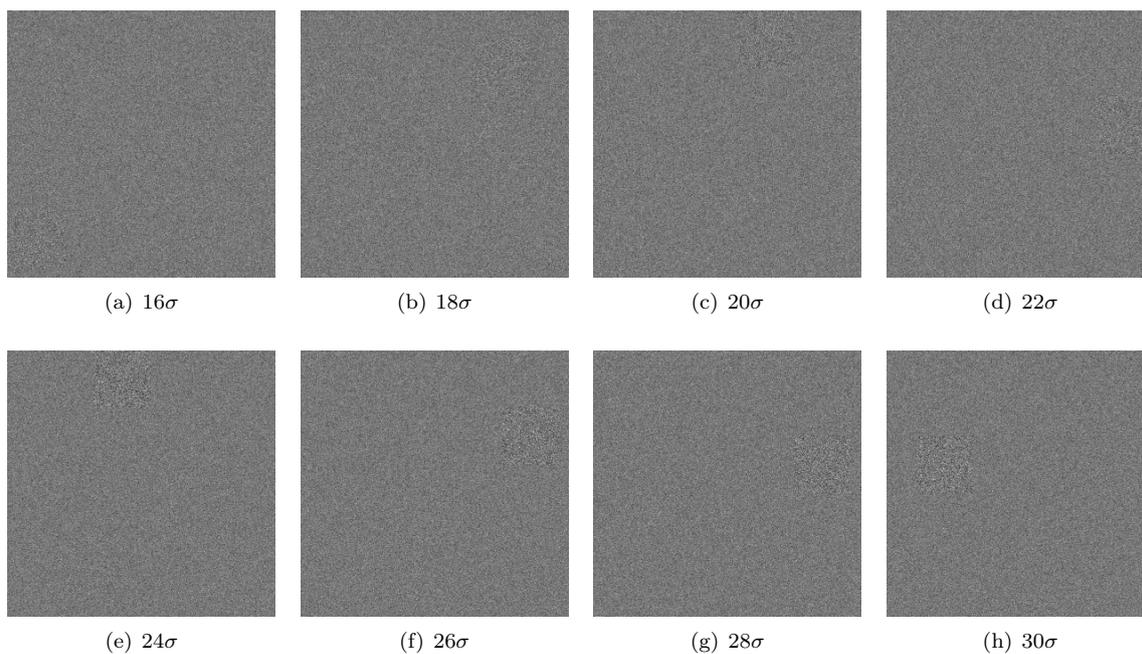


FIGURE B.3 – Détermination empirique du seuil t_s : à partir de l'observation visuelle, on a déterminé $t_s = 24$.

- contrôle des moyens d'évaluation : distance à l'écran, luminosité de la pièce, calibration de l'écran, notamment pour être conforme aux spécifications de la norme ITU-BT500 [1],
- étude de l'effet du gamma de l'écran : seules des déviations positives (pixels clairs, taches apparaissant en plus clair, tache de bruit avec une variance plus élevée) ont été évaluées. Il faudrait pousser l'étude sur des déviations négatives. On observerait notamment que le seuil à choisir n'est pas forcément symétrique par rapport à la moyenne de la statistique évaluée.
- Contrôle de l'effet d'apprentissage à l'artefact à évaluer : nous avons constaté au cours de l'évaluation que les experts s'habituent au type d'artefact à détecter, et améliorent ainsi leur sensibilité de détection
- Étude dans le cas de bruits différents : l'évaluation a été menée sur des images de bruit blanc. Il faudrait étudier l'impact du contenu fréquentiel du bruit de fond sur les seuils de détection à adopter.

Bibliographie

- [1] Methodology for the subjective assessment of the quality of television pictures, Recommendation ITU-R BT.500-11, 2002.
- [2] Peter J. Acklam. Inverse normal cumulative distribution function. <http://home.online.no/~pjacklam/notes/invnorm/index.html>.
- [3] Harrison H. Barrett. An introduction to Objective Assessment of Image Quality.
- [4] Harrison H. Barrett, Craig K. Abbey, and Eric Clarkson. Objective assessment of image quality. iii. roc metrics, ideal observers, and likelihood-generating functions. *Journal of the Optical Society of America A*, 15 :1520–1535, June 1998.
- [5] Harrison H. Barrett, Jie Yao, Jannick P. Rolland, and Kyle J. Myers. Model observers for assessment of image quality. *Proceedings of the National Academy of Science*, 90 :9758–9765, November 1993.
- [6] Peter G. J. Barten. *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. SPIE - The International Society for Optical Engineering, Bellingham, Washington, 1999.
- [7] Jacob Beutel, Harold L. Kundel, and Richard L. Van Metter, editors. *Handbook of Medical Imaging*, volume 1. SPIE - The International Society for Optical Imaging, 2000.
- [8] Juliette Blanchet. *Modèles markoviens et extensions pour la classification de données complexes*. PhD thesis, Université Joseph Fourier, Grenoble 1, Octobre 2007.
- [9] Harry Blum. A Transformation for Extracting New Descriptors of Shape. In Weiant Wathen-Dunn, editor, *Models for the Perception of Speech and Visual Form*, pages 362–380. MIT Press, Cambridge, 1967.
- [10] Philippe Bolon. Filtrage d’ordre vraisemblance et optimalite des prétraitements d’image. *Traitement du Signal*, 9(3) :225–250, 1992.
- [11] Arthur L. Bowley. *Element of Statistics*. P.S. King, London, 1920.
- [12] J.E. Bresenham. Algorithm for computer control of a digital plotter. *IBM System Journal*, 4(1) :25–30, 1965.
- [13] Arthur E. Burgess. Comparison of non-prewhitening and hotelling observer models. In H. L. Kundel, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 2436 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 2–9, April 1995.
- [14] J.F. Canny. A computational approach of edge detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8, November 1986.
- [15] Amarpreet S. Chawla, David Delong, Robert Saunders, Craig Abbey, and Ehsan Samei. Analyzing the effect of dose reduction on the detection of mammographic lesions using mathematical observer models. *Progress in biomedical optics and imaging*, 7(32), 2006.
- [16] Corinna Cortes and Mehryar Mohri. AUC optimization vs. error rate minimization. In *Advances in Neural Information Processing Systems (NIPS 2003)*, volume 16. MIT Press, 2003.
- [17] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley, New York, 1991.
- [18] Ingrid Daubechies. *Ten Lectures on Wavelets*. Soc for Industrial & Applied Math, December 1992.

- [19] S. E. Derenzo, M. J. Weber, E. Bourret-Courchesne, and M. K. Klintonberg. The quest for the ideal inorganic scintillator. *Nuclear Instruments and Methods in Physics Research A*, 505 :111–117, June 2003.
- [20] Agnes Desolneux, Lionel Moisan, and Jean-Michel Morel. Edge detection by helmholtz principle. *Journal of Mathematical Imaging and Vision*, 14 :271–284, 2001.
- [21] Agnes Desolneux, Lionel Moisan, and Jean-Michel Morel. Gestalt theory and computer vision. In A. Carsetti, editor, *Seeing, Thinking and Knowing*, pages 71–101. Kluwer Academic Publishers, 2004.
- [22] Agnès Desolneux, Lionel Moisan, and Jean-Michel Morel. Maximum meaningful events and applications to image analysis. *The Annals of Statistics*, 31(6) :1822–1851, Dec. 2003.
- [23] M. N. Do and M. Vetterli. Contourlets : a directional multiresolution image representation. In *Proceedings of 2002 IEEE International Conference on Image Processing*, pages 357–360, Rochester, New York, USA, Sept. 2002.
- [24] D.D. Dorfman and E.J. Jr Alf. Maximum likelihood estimation of parameters of signal detection theory - a direct solution. *Psychometrika*, 33 :117–124, 1968.
- [25] J.P. Egan. *Signal detection theory and ROC analysis*. Academic Press, New York, 1975.
- [26] U. Engelke and H. J. Zepernick. Quality evaluation in wireless imaging using feature-based objective metrics. In *Proc. of IEEE ISWPC*, Feb. 2007.
- [27] Ulrich Engelke, Vuong Xuan Nguyen, and Hans-Jurgen Zepernick. Regional attention to structural degradations for perceptual image quality metric design. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 869–872, 2008.
- [28] Tom Fawcett. ROC graphs : Notes and practical considerations for researchers. Technical report, HP Labs, 2004. http://home.comcast.net/~tom.fawcett/public_html/papers/ROC101.pdf.
- [29] Sylvain Fischer, Filip Šroubek, Laurent Perrinet, Rafael Redondo, and Gabriel Cristóbal. Self-invertible 2d log-gabor wavelets. *Int. J. Comput. Vision*, 75(2) :231–246, 2007.
- [30] Peter A. Flach. Tutorial on the many faces of roc analysis in machine learning. In *International Conference of Machine Learning*, <http://www.cs.bris.ac.uk/flach/ICML04tutorial/>, 4 Juillet 2004.
- [31] I. Fujieda, G. Cho, J. Drewery, T. Gee, and al. X-ray and charged particle detection with CsI(Tl) layer coupled to aSi :H photodiode layers. *IEEE Transactions on Nuclear Science*, 38(2) :255–262, April 1991.
- [32] Ph. Réfrégier G. Delyon, P. Martin and F. Guérault. Segmentation d’images par minimisation de la complexité stochastique avec un modèle de bruit non paramétrique. In *Congrès Reconnaissance de Formes et Intelligence Artificielle (RFIA2006)*, Janvier 2006.
- [33] Florence Forbes Gilles Celeux and Nathalie Peyrard. Em procedures using mean field-like approximations for markov model-based image segmentation. Technical report, INRIA Rhone-Alpes, December 2001.
- [34] Rafael C. Gonzales and Richard E. Woods. *Digital Image Processing Second Edition*. Prentice Hall, 2002. ISBN 0201180758.
- [35] François Goudail, Philippe Réfrégier, and Guillaume Delyon. Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images. *J. Opt. Soc. Am. A*, 21(7) :1231–1240, 2004.
- [36] D.B. Green and J.A. Swets. *Signal detection theory and psychophysics*. Wiley, New York, 1966.
- [37] Bénédicte Grosjean and Lionel Moisan. *A-contrario* detectability of spots in textured backgrounds. *Journal of Mathematical Imaging and Vision*, 33(3) :313–337, 2009.
- [38] Frank R. Hampel. *Contributions to the theory of robust estimation*. PhD thesis, University of California, 1968.
- [39] J. Hanley. Receiver operating characteristic (roc) methodology : The state of the art. *Critical Reviews in Diagnostic Imaging*, 29(3) :307–335, 1989.

- [40] J.A. Hanley and B.J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(29) :29–36, 1982.
- [41] N. R. Harvey and J. Theiler. Focus-of-attention strategies for finding discrete objects in multispectral imagery. In S. S. Shen and P. E. Lewis, editors, *Proceedings of the SPIE, Volume 5546*, volume 5546, pages 179–189, oct 2004.
- [42] Y. Hawwar and A. Reza. Spatially adaptative multiplicative noise image denoising technique. *IEEE Transactions on Image Processing*, 11 :1397–1404, Dec. 2002.
- [43] P.V.C. Hough. A method and means for recognizing complex patters. *U.S. Patent 3,069,654*, 1962.
- [44] Jordi Inglada. Change detection on SAR images by using a parametric estimation of the Kullback-Leibler divergence. In *Proc. of Geoscience and Remote Sensing Symposium (IGARSS)*, volume 6, pages 4104–4106, 2003.
- [45] John M. Irvine. Assessing target search performance : the free-response operator characteristic model. *SPIE Optical Engineering*, 43(12) :2926–2934, December 2004.
- [46] A.K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, 1995.
- [47] Jerry Moore James Keller and Paul Gader. A fuzzy logic approach to detector scoring. In *Fuzzy Information Processing Society - NAFIPS, 1998 Conference of the North American*, volume 20, pages 339–344, 20–21 August 1998.
- [48] Nancy A. David James Theiler, Neal Harvey and John M. Irvine. Approach to target detection based on relevant metric for scoring performance. *33rd Applied Imagery Pattern Recognition Workshop (AIPR'04)*, 00 :184–189, 2004.
- [49] Don H. Johnson and Sinan Sinanovic'. Symmetrizing the kullback-leibler distance. Technical report, IEEE Transactions on Information Theory, 2001.
- [50] Tae-Hwan Kim and Halbert White. On more robust estimation of skewness and kurtosis : Simulation and application to the S&P500 index. University of California at San Diego, Economics Working Paper Series 2003-12, Department of Economics, UC San Diego, Sep 2004.
- [51] Ajay Kumar and Grantham K.H. Pang. Defect detection in textured materials using optimized filters. *IEEE Transaction on Systems, Man, and Cybernetics*, 32(5), October 2002.
- [52] T.M. Kusuma and H.J. Kusuma. A reduced-reference perceptual quality metric for in-service image quality assessment. In *Joint First Workshop on Mobile Future and Symposium on Trends in Communications*, pages 71–74, 2003.
- [53] J.L. Lacoume and P.O. Amblard. *Statistiques d'ordre supérieur pour le traitement du signal*. Masson, Paris, 1997.
- [54] V. Lashkia. Defect detection in X-ray images using fuzzy reasoning. *Image and Vision Computing*, 19(5) :261–269, Avril 2001.
- [55] J.-S. Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2 :165–168, March 1980.
- [56] D.H. Lim. Robust edge detection in noisy images. *Computational Statistics and Data Analysis*, 50, 2006.
- [57] Jae S. Lim. *Two-dimensional signal and image processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1990.
- [58] Jen-Jen Lin, Naoki Saito, and Richard A. Levine. Edgeworth approximations of the kullback-leibler distance towards problems in image analysis. Technical report, University of California, Davis, CA, 1999.
- [59] G. Liu and R.M. Haralick. Assignment problem in edge detection performance evaluation. In *Proceedings of Conference on Computer Vision and Pattern Recognition CVPR 2000*, pages 1026–1031, Hilton Head, SC, USA, June 2000. IEEE Computer Society.

- [60] Zhiling Long and Nicolas H. Younan. Denoising of images with multiplicative noise corruption. In *Proceedings of EUSIPCO 2005*, Antalya, Sept. 2005.
- [61] Stéphane Mallat. *Une exploration des signaux en ondelettes*. Editions de l'école Polytechnique, 2000.
- [62] Vladimir Y. Mariano, Junghye Min, Jin-Hyeong Park, Rangachar Kasturi, David Mihalcik, Huiping Li, David Doermann, and Thomas Drayer. Performance evaluation of object detection algorithms. *International Conference on Pattern Recognition*, 03 :30965, 2002.
- [63] Pascal Martin. *Application du principe de minimisation de la complexité stochastique à la segmentation d'images bruitées par contour actif*. PhD thesis, Université Paul Cézanne, Aix-Marseille, Février 2006.
- [64] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. Perceptual Blur and Ringing Metrics : Application to JPEG2000. *Signal Processing : Image Communication*, 19(2) :163–172, 2004.
- [65] F. Maussang. *Traitement d'images et fusion de données pour la détection d'objets enfouis en acoustique sous-marine*. PhD thesis, Université Joseph Fourier, Grenoble, Novembre 2005.
- [66] F. Maussang. Utilisation des statistiques d'ordre supérieur en contrôle qualité de détecteurs de rayons x. In *Actes du 20e colloque GRETSI*, Septembre 2005.
- [67] Richard L. Van Metter and Jacob Beutel, editors. *Amorphous silicon x-ray image sensor*, volume 2708. SPIE, 1996.
- [68] Charles E. Metz. ROC methodology in radiologic imaging. *Investigative Radiology*, 21 :720–733, 1986.
- [69] Yann Le Meur, Jean-Michel Vignolle, and Jocelyn Chanussot. A practical use of roc analysis to assess the performances of defects detection algorithms. In SPIE, editor, *Proceedings of Eighth International Conference on Quality Control by Artificial Vision*, volume 6356, may 2007.
- [70] Yann Le Meur, Jean-Michel Vignolle, and Jocelyn Chanussot. A practical use of receiver operating characteristic analysis to assess the performances of defects detection algorithms. *Journal Of Electronic Imaging*, 17, Jul. 2008.
- [71] J.J.A. Moors. A Quantile Alternative for Kurtosis. *The Statistician*, 37(1) :25–32, 1988.
- [72] J. C. Nascimento and J. S. Marques. Novel metrics for performance evaluation of object detection algorithms. In *1st ISR Workshop on Systems, Decision and Control Robotic Monitoring and Surveillance, Lisbon*, June 2005.
- [73] F. Galland P. Martin, Ph. Réfrégier and F. Guéroult. Nonparametric statistical snake based on the minimum stochastic complexity. *IEEE Transactions on Image Processing*, 15(9) :2762–2770, September 2006.
- [74] Nathalie Peyrard. *Approximations de Type Champ Moyen des Modèles de Champ de Markov pour la Segmentation de Données Spatiales*. PhD thesis, Université Joseph Fourier, Grenoble, France, October 2001.
- [75] J.M.S. Prewitt. Object enhancement and extraction. In B.S. Lipkin and A. Rosenfeld, editors, *Picture Processing and Psychopictorics*. Academic Press, New York, 1970.
- [76] Philippe Réfrégier. *Théorie du signal*. Masson, 1993.
- [77] L.G. Roberts. Machine perception of three-dimensional solids. In *Optical and Electro-Optical Information Processing*. MIT Press, Massachusetts, 1965.
- [78] William Rucklidge. *Efficient Visual Recognition Using the Hausdorff Distance*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1996.
- [79] Gilbert Saporta. *Probabilités, analyses de données et statistiques, 2eme Édition*. Editions Technip, 2006. ISBN 2710808145.
- [80] Hamid Rahim Sheikh, Alan Conrad Bovik, and Lawrence Cormack. No-reference quality assessment using natural scene statistics : Jpeg2000. *IEEE Trans. on Image Processing*, 14(11), November 2005.

- [81] I.E. Sobel. *Camera Models and Machine Perception*. PhD thesis, Stanford University, California, 1970.
- [82] Arthur Stanton. Wilhem Conrad Röntgen, On a New Kind of Rays. *Nature*, 53 :274–276, 1896. traduction d’un article à la Société de Physique et de Medecine de Würzburg.
- [83] F. Goudail V. Pagé and Ph. Réfrégier. Improved robustness of target location in nonhomogeneous backgrounds by use of the maximum-likelihood ratio test location algorithm. *Optics Letters*, 24(20) :1385, October 1999.
- [84] Zhou Wang and Alan C. Bovik. Why is image quality assessment so difficult. In *in Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, pages 3313–3316, 2002.
- [85] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, Student Member, and Eero P. Simoncelli. Image quality assessment : from error measurement to structural similarity. *IEEE Transactions on Image Processing*, 13 :600–612, 2004.
- [86] Zhou Wang, Hamid R. Sheikh, and Alan C. Bovik. No-reference perceptual quality assessment of jpeg compressed images. In *Proceedings of IEEE 2002 International Conferencing on Image Processing*, pages 477–480, 2002.
- [87] Zhou Wang and Eero P. Simoncelli. Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. In *Human Vision and Electronic Imaging X Conference*, pages 149–159, 2005.
- [88] A.B. Watson. Detection and recognition of simple spatial forms. Technical report, NASA, Ames Res. Center, Moffett Field, CA, 1983.
- [89] Andrew B. Watson. Dctune : A technique for visual optimization of dct quantization matrices for individual images. *Society for Information Display Digest of Technical Papers*, XXIV :946–949, 1993.
- [90] Andrew B. Watson. Image data compression having minimum perceptual error, May 1997. US Patent n°5629780.
- [91] Michael J. Wichura. Algorithm as 241 : The percentage points of the normal distribution. *Applied Statistics*, 37(3) :477–484, 1988.
- [92] H. Wilson and J. Bergen. A four mechanism model for threshold spatial vision. *Vision Research*, (19) :19–32, 1979.
- [93] Christian Wolf and Jean-Michel Jolion. Object count/area graphs for the evaluation of object detection and segmentation algorithms. *Int. J. Doc. Anal. Recognit.*, 8(4) :280–296, 2006.
- [94] Yani Zhang, Binh T. Pham, and Miguel P. Eckenstein. Automated optimization of jpeg 2000 encoder based on model observer performance for detecting variable signals in x-ray coronary angiograms. *IEEE Transactions on Medical Imaging*, 23(4), Apr. 2004.
- [95] Yani Zhang, Binh T. Pham, and Miguel P. Eckstein. The effect of nonlinear human visual system components on performance of a channelized hotelling observer in structured backgrounds. *IEEE Transactions on Medical Imaging*, 25(11) :1348–1362, Oct. 2006.

Résumé

Durant les dix dernières années, l'imagerie médicale a profité pleinement de l'essor technologique des techniques d'imagerie numérique. Dans le domaine de la radiographie X, les nouveaux détecteurs, possédant un encombrement réduit et une meilleure sensibilité, détrônent progressivement les anciennes techniques à film ou à tube. Cette thèse, préparée en partenariat avec l'entreprise Trixell, leader mondial des détecteurs plats pour radiographie médicale, traite du contrôle qualité des images issues de ces détecteurs. L'exigence du domaine médical impose un contrôle strict des défauts pouvant apparaître sur les images. Cette thèse est donc axée sur la conception d'une chaîne complète d'analyse de la qualité des images issues des détecteurs numériques. Un éclairage particulier est donné à la détection des défauts, parfois peu visibles, à l'aide de techniques de détection adaptées à notre contexte de défauts spatialement corrélés dans un fort bruit de fond.

Mots clés : détection de défauts, contrôle qualité, radiographie numérique, mesure de visibilité

Abstract

Since last decade, medical imaging has grown up with the development of new digital imaging techniques. In the field of X-ray radiography, new detectors replace progressively older techniques, based on film or x-ray intensifiers. These digital detectors offer a higher sensibility and reduced overall dimensions. This PhD thesis has been prepared with Trixell, the world leading company in flat detectors for medical radiography. It deals with quality control on digital images stemming from these detectors. High quality standards of medical imaging impose a close analysis of the defects that can appear on the images. This PhD thesis describes a complete process for quality analysis of such images. A particular focus is given on the detection task of the defects, thanks to methods well adapted to our context of spatially correlated defects in noise background.

Keywords : defects detection, quality control, digital radiography, visibility measurement