



HAL
open science

La qualité d'image dans le contexte de la numérisation de livres anciens

Emile Vinsonneau

► **To cite this version:**

Emile Vinsonneau. La qualité d'image dans le contexte de la numérisation de livres anciens. Traitement des images [eess.IV]. Université de Bordeaux, 2015. Français. NNT : 2015BORD0057. tel-01233181

HAL Id: tel-01233181

<https://theses.hal.science/tel-01233181>

Submitted on 24 Nov 2015

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.

THÈSE CIFRE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUE ET INFORMATIQUE
SPÉCIALITÉ INFORMATIQUE

Par Emile Vinsonneau

La qualité d'image dans le contexte de la numérisation de livre ancien

Sous la direction de : Jean-Philippe Domenger
(Co-encadrante : Anne Cherif)

Soutenue le 30 mars 2015

Membres du Jury :

Mme Delest Maylis	Professeur	LaBRI, université de Bordeaux	Présidente
Mr Ogier Jean-Marc	Professeur	L3i, université de La Rochelle	Rapporteur
Mr Ramel Jean-Yves	Professeur	LI, université de Tour	Rapporteur
Mme Eglin Véronique	Maître de conf.	LIRIS, université de Lyon	Examinatrice

préparée au sein

du **Laboratoire Bordelais de Recherche en Informatique**
et de la société **Innovative, Imaging, Solutions (i2S)**

Résumé : L'objectif de cette thèse est de proposer des méthodes ou des outils de calibration permettant d'améliorer la qualité d'image dans le cadre de la numérisation de livres anciens.

Le premier chapitre concerne la mise en œuvre d'un "scanner sans éclairage". Ce scanner est matriciel et il ne contrôle pas la lumière via un éclairage artificiel. L'objectif du projet est de pouvoir corriger, par calibration ou par la mise en place d'un traitement d'image les défauts d'éclairage apparaissant sur le document. Nous verrons dans un premier chapitre les solutions possibles pour y répondre. Nous y proposerons également un moyen de segmenter le fond du document en se basant sur des pixels caractérisants le fond du document. Ce résultat permettra de reconstruire le fond puis de corriger les non-uniformités.

Le deuxième chapitre concerne la mise en place d'un contrôle qualité de numérisation de livres anciens. Effectivement, la manipulation du matériel dans le contexte de la numérisation industrielle engendre des erreurs possibles de réglage. De ce fait, nous obtenons une numérisation de moindres qualités. Le contrôle se focalisera sur le flou de focus qui est le défaut le plus présent. Nous proposons dans ce chapitre de combiner différentes approches du problème permettant de le quantifier. Nous verrons que la combinaison de ces informations permet d'estimer avec précision la qualité de netteté de l'image.

Le troisième chapitre évoque les problématiques de la gestion de la couleur. Pour avoir une homogénéité colorimétrique entre tous les appareils, il est indispensable que la calibration soit inférieure à un seuil visuel. Les contraintes industrielles ont de nombreux critères et il est difficile de tous les respecter. L'objectif de ce chapitre est de récapituler comment mettre en place un système permettant de calibrer la couleur avec toutes ces contraintes. Puis, il s'agit de comprendre comment créer une transformation entre l'espace de couleur de l'appareil et l'espace de couleur de connexion (l'espace $L^*a^*b^*$). Nous verrons que la solution de ce problème se résout par une régression polynomiale dont le degré du polynôme varie en fonction du nombre de mesures faite sur la mire colorimétrique.

mots-clés : numérisation, qualité d'image, images de document anciens, netteté, descripteur de qualité, correction d'éclairage, fond d'un document, calibration couleur

Abstract :

Image quality in document scanner

The goal of this thesis is to add some tools in order to upgrade image quality when scanning with book digitization.

First Chapter talks about image scanner without lighting control. This problem focuses on document camera. The goal is to correct lighting. We will see some corrections and we will suggest our method. For this part, we detect pixel's background document and we will rebuild the background of the image by them. With this information, we can correct lighting.

Second chapter presents some way to do quality control after digitization, specially out of focus problem. We will enumerate different points of view to analyse and to estimate this information. To validate descriptors, we suggest to blur any picture and to compute blur estimation in order to evaluate precision. After that, we propose to combine descriptors by machine learning.

Third chapter mentions color management problem. Every image device needs to be calibrated. This chapter will expose how to calibrate scanner and explain it. We will see that $L^*a^*b^*$ color space is the connection profile space. To calibrate color, we must transform scanner color space to $L^*a^*b^*$. We will see, in order to convert information, solution depends on color chart used but we show a link between the function and the number of patches.

Keywords : digitization, image quality, old document, sharpness, quality feature, lighting correction, background of document, color calibration

Remerciements :

Je tiens particulièrement à remercier mes directeurs de thèse, Anne Chérif et Jean-Philippe Domenger, pour leur expertise scientifique mais aussi pour leurs qualités humaines. Je leur remercie de m'avoir aidé à réaliser cette thèse sans qui je n'aurais probablement pu finir cette expérience difficile.

Je remercie également les personnes qui m'ont formé durant mes années universitaires et me donnant l'envie de travailler dans ce domaine.

Lors de mes années à i2S, je tiens à remercier les personnes qui m'ont aidé de près ou de loin pour la réalisation de ce projet. Je ne citerai probablement pas tout le monde, mais je remercie Jean-Luc, Michael, Romain, Nicolas, Dominique et Vincent pour m'avoir aidé et conseillé pour programmer. Je remercie également Jean-Luc, Pierre, Philippe et Bruno pour m'avoir apporté divers conseils et aidé à comprendre le fonctionnement des appareils de numérisation ainsi qu'Eric parce qu'il faut savoir dire "xxx" quand il faut. Et je finirai par remercier Anne pour son soutien ainsi que Rémi, Anthony et Sébastien.

Je tiens aussi à remercier les personnes qui m'ont aidé dans la réflexion de ma thèse : Rémi et Bastien, et les doctorants, Marc et Aurélien, à qui je souhaite tout plein de réussite.

Je pense aussi à mes amis pour leur attention et leur encouragement qui m'ont porté durant ces années. Je ne les citerai probablement pas tous mais je pense particulièrement à Hervé, Keup', Romu, Fred, Bulo, Michou, Crabbi, Franky, Kenny, Flo, BeBer ... Ainsi que Béatrice et Guillaume qui ont regardé si j'accordais bien mes participes passés. Merci.

Je termine en remerciant ma famille (pour avoir créé un être tel que moi même s'il y a quelques défauts de fabrication ...), mais surtout parce qu'ils ont pris le temps d'écouter mes réponses à la fameuse question "T'en es où dans ta thèse ?" alors que probablement même mes responsables ne comprendraient pas la réponse.

Table des matières

Résumé	i
Abstract	ii
Remerciements	iii
Table des matières	iv
Table des figures	vii
Liste des tableaux	xi
1 Introduction	1
2 La correction d'illumination	6
2.1 Introduction	6
2.2 Correction d'illumination	8
2.2.1 Les images intrinsèques	8
2.2.2 La Carte de correction d'éclairage	15
2.2.2.1 La transformation	16
2.2.2.2 La binarisation dans le contexte des défauts d'illumination	21
2.3 Correction d'éclairage par une carte du fond	25
2.3.1 Les pixels caractéristiques du fond d'un document :	27
2.3.2 Segmentation	31
2.3.2.1 Split	31
2.3.2.2 Merge	32
2.3.3 Analyse du fond	33
2.3.4 Reconstruction du fond	39
2.3.4.1 Reliure du livre	39
2.3.4.2 Reconstruction du fond du document	42
2.3.5 Application	43
2.3.6 Conclusion	46
2.4 Conclusion	48
3 Flou de mise au point	49
3.1 Introduction	49
3.2 La mise au point	50
3.2.1 Fonction d'étalement du point	55
3.3 Numérisation industrielle	57

3.3.1	Sampling Rate	59
3.3.2	Resolution	59
3.3.3	Sharpening	60
3.3.4	Acutance	60
3.3.5	Conclusion	62
3.4	Mesure de netteté sur une image unique sans référence	63
3.4.1	Étalement de contour	63
3.4.2	Descripteurs	68
3.4.3	Résultats et analyses	73
3.4.4	Apprentissage supervisé	76
3.4.5	Résultats expérimentaux	79
3.5	Conclusion	85
4	La calibration couleur	87
4.1	Introduction	87
4.2	La gestion de la couleur	89
4.2.1	La gestion de la couleur des appareils d'acquisition - le scanner	91
4.2.2	Les différents espaces de couleur utilisés	92
4.2.2.1	CIE RVB	92
4.2.2.2	CIE XYZ 1931	94
4.2.2.3	CIE $L^*a^*b^*$	97
4.2.3	La différence de deux couleurs : ΔE	99
4.2.4	Les contraintes industrielles	102
4.2.4.1	Metamorfoze	102
4.2.4.2	FADGI	103
4.2.5	Processus de calibration d'un matériel de numérisation	106
4.3	La calibration colorimétrique d'un scanner	110
4.3.1	Les caractéristiques de la mire de calibration	111
4.3.2	Le gamma du scanner	113
4.3.3	Estimation par un modèle linéaire et non-linéaire pour la transformation de RVB à $L^*a^*b^*$	116
4.3.4	L'évaluation de la calibration	121
4.3.5	Analyse des divergences de résultat avec le logiciel Imatest (référence de l'évaluation colorimétrique)	122
4.3.6	Résultats et analyses sur le ΔE_{00}	124
4.3.6.1	Analyse du degré du polynôme pour la résolution du problème linéaire par la méthode des moindres carrés	125
4.3.6.2	Analyse sur les coefficients de pondération de la méthode des moindres carrés pondérées	128
4.4	Conclusion	130
5	Conclusion	132
6	Annexes	135
6.1	Exemples de la correction d'éclairage	135
6.1.1	Exemples utilisés dans le chapitre	135

6.1.2	Exemples	137
6.1.3	Exemple d'un livre ancien	138
6.1.4	Exemple de livres récents	139
6.1.5	Exemple détaillé d'une erreur de correction d'éclairage	141
6.1.6	Exemple détaillé d'une suppression d'ombre	142
6.1.7	Exemple détaillé de correction d'éclairage de livre récent	143
6.2	Tableau des résultats concernant la calibration couleur	145
6.2.1	Comparaison des conversions <i>sRGB</i> [4.3.5]	145
6.2.2	L'icc <i>sRGB</i> officiel [4.3.5]	147
6.2.3	Comparaison en delta E des mires old et New d'imatest avec la référence de la mire utilisée [4.3.5]	148
6.3	Détails des calculs du deltaE	148
6.3.1	ΔE CMC [4.2.3]	148
6.3.2	ΔE 94 [4.2.3]	150
6.3.3	ΔE 00 [4.2.3]	151
6.4	Les logiciels	152
6.4.1	La correction d'illumination	152
6.4.2	Le flou de mise au point	154
6.4.3	La calibration couleur	156

Table des figures

1.1	POLINUM - Chaîne de la numérisation à la valorisation	1
1.2	Scanner linéaire	2
1.3	Scanner matriciel	3
1.4	Scanner linéaire - SupraScan*	3
1.5	Scanner matriciel - CopiBook*	4
1.6	Nous présentons, ici, une zone d'une image ayant un problème d'éclairage. Nous observons que le côté gauche de l'image est plus foncée que le côté droit alors que le fond du document est uniforme.	4
2.1	Exemple d'un défaut d'éclairage ambiant	7
2.2	Exemple d'un défaut d'éclairage créé par l'utilisateur	7
2.3	Décomposition de l'image en images intrinsèques [Weiss, 2001]	9
2.4	Filtre passe haut avec les paramètres $r_H = 2.5$, $r_L = 0.5$ et $\delta_w = 0.2$	10
2.5	Schéma des différentes étapes du filtre homomorphique	11
2.6	Application du filtrage homomorphique dans le but de corriger l'éclairage	11
2.7	Correction d'éclairage à partir d'un filtre retinex	14
2.8	Mise en place d'une procédure de suppression de fond à l'aide d'une référence	16
2.9	Résultats avec les opérateurs euclidiens	17
2.10	Mise en place d'une procédure de suppression de fond à partir de l'opérateur LIP	20
2.11	Comparaison des opérateurs \times , γ et \odot	21
2.12	Premier exemple de problème d'éclairage associés à son histogramme	22
2.13	Deuxième exemple de problème d'éclairage associés à son histogramme	22
2.14	Histogrammes locaux	23
2.15	Étapes de la correction d'éclairage par une carte de fond dans le cas	26
2.16	Schéma d'une coupe 1D d'un fond	28
2.17	Exemple de pixels plats	30
2.18	Résultat du split	34
2.19	Les différentes régions continues	35
2.20	Cas où la région centrale est supérieure en luminance car elle est sur-exposée par la lumière	37
2.21	Cas où la région supérieure est bleu avec une luminance de 230 et la région inférieure est grise avec une luminance de 180	38
2.22	Exemple de problème de sélection dans un cas de saturation de luminosité	39
2.23	Reliures de livre ancien	40
2.24	Reliures de livre nouveau	40
2.25	Courbe de reliure de livre ancien	41
2.26	Courbe de reliure de livre nouveau	42
2.27	Reconstruction partiel du fond	43

2.28	Reconstruction partielle du fond	44
2.29	Utilisation de la correction d'éclairage pour enlever une ombre humaine	45
2.30	Utilisation de la correction d'éclairage	46
2.31	Nous observons dans la correction une aberration de la correction d'éclairage au niveau de l'image et son dégradé artificiel (voir annexe 6.1.5)	47
3.1	Rayons particuliers dans le cas d'une lentille convergente	50
3.2	Mise au point	51
3.3	cercle de confusion	51
3.4	Focale - Ici, nous pouvons voir que la modification de la focale impacte directement l'information sur le capteur par un étalement du point.	52
3.5	La convergence des rayons de la lentille. Nous pouvons voir que la convergence des rayons lumineux modifie la position du foyer principal image, par conséquent, l'information du point sur le capteur s'étale.	52
3.6	La profondeur de champ	53
3.7	Zone nette d'un scanner	54
3.8	Tableau récapitulatif des normes lié à la représentation du signal	58
3.9	Mire de résolution - Sampling Rate de SilverFast	59
3.10	Fonction de transfert de modulation. De haut en bas, nous avons représenté la mire sinusoïdale avec le profil associé. Ensuite, nous avons l'observation théorique d'une caméra et le profil associé. Nous constatons que le premier cycle est bien représenté puis le signal converge à cause de la diffraction optique. La dernière courbe est la FTM.	60
3.11	Exemples de deux profils de contours. Nous observons que le second profil présente un contraste plus important au niveau des contours.	61
3.12	Exemple d'images mettant en évidence l'acutance. La première image est l'image originale sans traitement. La seconde et la troisième image sont les résultats de l'image avec un filtre passe-haut faible et fort. On a une sensation de netteté qui est de plus en plus marquée.	61
3.13	Exemple de FTM avec des courbes ayant des acutances différentes. Nous observons que la courbe A et B ont la même capacité à différencier les détails, mais A conserve un contraste plus élevé.	62
3.14	L'approche du defocus blur estimation [Zhuo and Sim, 2011] : ici, \otimes est le produit de convolution, Δ est l'opérateur du gradient et la ligne en pointillé est la position du contour.	65
3.15	Nous avons l'image (a) nette et l'image (c) floue. Les images (b) et (d) représentent les transformées de Fourier et leur puissance spectrale de puissance associées aux image (a) et (c). [Russ, 2010]	66
3.16	Schématisation de la binarisation faite dans l'article [Lelegard et al., 2010]	67
3.17	Profils de contour 1D	68
3.18	Histogrammes des épaisseurs de contours	69
3.19	Pyramide d'ondelette sur trois niveaux	70
3.20	Les différents types de structure de contour	70
3.21	Exemple d'évolution du flou sur une même image et les binarisations dans l'espace de Fourier. Nous constatons que les hautes fréquences, au centre de l'image, diminuent lorsque le flou augmente.	72
3.22	Binarisation dans l'espace de Fourier après filtrage	72
3.23	Représentation des valeurs normalisées des descripteurs en fonction de σ_{th}	73

3.24	Représentation de 5 mêmes régions d'une image lissée par un filtre gaussienne de noyau 7x7 et de σ variant d'un pas de 0.4. Le morceau d'image du haut est supposée nette, la troisième image correspondant à un filtrage de σ de 0.8 et elle est considérée comme floue par une organisation telle que la Bibliothèque Nationale de France, tandis qu'une personne non spécialisée dans le domaine va davantage considérer qu'une image est floue pour un σ allant de 0.8 et 1.2 soit entre la troisième et quatrième image.	74
3.25	Représentation des erreurs d'estimation de σ en fonction du σ théorique	75
3.26	Représentation de trois descripteurs et la formation de cluster pouvant permettre l'estimation de σ	77
3.27	Représentation des erreurs d'estimation de σ obtenu par apprentissage supervisé en fonction du σ théorique	78
3.28	Exemple d'image d'acquisition. Dans la partie droite de l'image, nous avons mis une mire FTM, une mire couleur et une mire de résolution.	79
3.29	Évaluation de la FTM selon la FADGI	81
3.30	Évolution des descripteurs en fonction de la modification de la distance de mise au point	81
3.31	Évolution des descripteurs en fonction de la modification de la focale	82
3.32	Évolution de la FTM et de notre estimateur en fonction de la modification de la distance de mise au point	83
3.33	Évolution de la FTM et de notre estimateur en fonction de la modification de la focale	84
4.1	L'une des problématiques de la gestion de la couleur est de pouvoir communiquer la couleur entre les différents outils graphiques.	89
4.2	L'utilisation d'un fichier icc permet la conversion et il est invisible pour l'utilisateur. .	90
4.3	Un fichier icc se découpe en deux parties. La partie supérieure donne les informations d'utilisation. Principalement, le type de profile indique le sens de conversion entre le "Data Color Space" et "PCS Color Space". La deuxième partie donne les informations de conversion. Dans le fichier icc présenté, nous avons les luts de conversion explicite AtoB0 et les informations de conversion avec AtoB1 et les tags mediaWhitePointTag et chromaticAdaptationTag.	91
4.4	Triangle des couleurs dans le repère $(0, \vec{R}, \vec{V}, \vec{B})$ [pierre.nerzic, 2008].	94
4.5	Diagramme Chromatique de Wright - Projection du Triangle des couleurs dans le repère (OR, OG) [pierre.nerzic, 2008].	94
4.6	Représentation des Fonctions colorimétriques RVB [Wik, 2014].	95
4.7	Représentation des fonctions colorimétriques XYZ [Wik, 2014].	95
4.8	Diagramme des couleurs RVB et XYZ [Wik, 2014].	96
4.9	Représentation graphique de l'espace $L^*a^*b^*$ [Wik, 2014]	98
4.10	Conversion d'une image en RVB en $L^*a^*b^*$	99
4.11	Tableau des normes de la FADGI concernant la reproduction de la couleur [[Initiative, 2010]]	104
4.12	Mesure de la courbe des niveaux des gris [Initiative, 2010]. Nous avons en entrée (abscisse) la densité physique des patches de niveaux de gris et en sortie la valeur RVB au format standard. La valeur mesurée doit se situer entre les deux courbes en pointillés. .	105
4.13	Nous avons représenté la répartition colorimétrique de la <i>colorchecker classic</i> dans l'espace RVB de l'appareil et nous le comparons à celle d'un format standard dans la figure 4.14	107
4.14	Représentation de la répartition colorimétrique de la <i>colorchecker classic</i> dans l'espace RVB standard	108
4.15	Schéma qui résume la problématique de la calibration couleur	108
4.16	Charte colorimétrique de la <i>colorchecker classic</i>	109

4.17	Mire de Kodak	111
4.18	Deux autres exemples de mires couleur. La mire Agfa <i>it8.7/2</i> en haut et la mire colorchecker SG en bas.	112
4.19	Représentation de la fonction gamma avec deux exemples caractéristiques de la représentation couleur	114
4.20	Mire de référence Blanc/gris/Noir fabriqué par X-rite	115
4.21	Deux exemples de mires. La mire de gauche est la Colorchecker Classic de X-rite et la mire de droite est la <i>Q13</i> de Kodak.	116
4.22	Transformation de l'espace R'V'B' à l'espace de référence $L^*a^*b^*$	117
4.23	Application de deux fichiers .icc pour obtenir un visuel dans un espace RVB normé	121
4.24	Les logiciels d'analyses d'images utilisent nécessairement trois profils ICC alors qu'un seul est nécessaire pour faire le calcul du ΔE qui se fait dans l'espace $L^*a^*b^*$	122
4.25	Représentation en deltaE des écarts de conversions à partir des patches de la <i>colorchecker classic</i>	123
6.1	Correction du manuel d'utilisation utilisé dans les illustrations pour les méthodes de l'état de l'art	136
6.2	Correction de la page à forte variation d'illumination	136
6.3	Correction d'une page à forte variation d'illumination	137
6.4	Correction d'un livre ancien ayant des taches sur le fond du livre	138
6.5	Correction d'un livre récent	139
6.6	Deuxième correction d'un livre récent	140
6.7	Exemple détaillé d'une erreur de segmentation du fond [2.3.6]	141
6.8	Exemple détaillé d'une suppression de fond [2.3.5]	142
6.9	Exemple détaillé d'une suppression de fond	143
6.10	Exemple détaillé d'une suppression de fond [2.4]	144
6.11	Interface de la correction d'éclairage	152
6.12	Interface de la mesure du flou de mise au point	154
6.13	Interface de départ du logiciel de calibration couleur	156
6.14	Interface de gestion du Gamma	157
6.15	Viewer	157
6.16	Courbes des niveaux de gris	158
6.17	Valeurs de calibration	158
6.18	Résultats des ICCs selon plusieurs modèles	159
6.19	Modification locale d'une LUT	159
6.20	Calcul du Delta E	160

Liste des tableaux

3.1	Tableau récapitulatif des calculs des FTM en fonction de la modification des distance de mise au point	80
3.2	Tableau récapitulatif des calculs des FTM en fonction de la modification de la focale	80
4.1	Couleurs primaires et blanc de référence du CIE RVB 1931	93
4.2	L'erreur de la balance des blancs	105
4.3	L'erreur des niveaux de gris	105
4.4	Tableau représentant les évaluations de la couleur en fonction du ΔE_{00} [Initiative, 2010]	106
4.5	Tableau représentant un triplet Blanc/Gris/Noir basé sur les densités données par une mire ColorChecker que l'on peut voir dans la figure 4.20	115
4.6	Tableau des moyennes des ΔE_{00} entre les différentes conversions	123
4.7	Tableau des moyennes des ΔE_{00} entre les différentes mires	124
4.8	Calcul du ΔE_{00} à partir de la mire <i>colorchecker classic</i>	126
4.9	Résultats du ΔE_{00} à partir de la mire <i>colorchecker SG 140</i>	127
4.10	Utilisation de la pondération dans la méthode des moindres carrés permettant d'obtenir une minimisation de l'erreur en ΔE_{00} sur les patches de niveaux de gris	129
6.1	Conversion de RVB en $L^*a^*b^*$ selon le modèle standard	145
6.2	Tableau des calculs du ΔE_{00} entre les différentes conversions	146
6.3	Tableau des calculs du ΔE_{00} entre les différentes conversions de sRGB.icc	147
6.4	Calcul du ΔE_{00} en fonction de la mire de référence d' <i>imatest (new)</i> et de la référence actuelle	148
6.5	Calcul du ΔE_{00} en fonction de la mire de référence d' <i>imatest (old)</i> et de la référence actuelle	149

Chapitre 1

Introduction

Cette thèse s'est déroulée dans le cadre d'une convention CIFRE (Convention Industrielle de Formation par la REcherche) entre la société *i2S*, fabricant de scanner de livres anciens et le LaBRI (Laboratoire Bordelais de Recherche en Informatique). Elle fait partie du projet de recherche POLINUM. POLINUM regroupe un consortium de 8 partenaires français, privés et publics. Cette collaboration a pour but de valoriser les fonds patrimoniaux. Pour atteindre son objectif, POLINUM crée une chaîne de numérisation modulaire (cf. figure 1.1) pour permettre la mise en valeur de ces documents.



FIGURE 1.1: POLINUM - Chaîne de la numérisation à la valorisation

Cette chaîne est divisée en plusieurs étapes :

- L'étape de numérisation regroupe l'ensemble des processus de numérisation permettant de passer du livre physique à une image numérique.
- L'extraction d'information s'appuie principalement sur la reconnaissance des caractères et des images qui constitue l'information importante d'un document.

- Les métadonnées sont des informations descriptives liées à l'image. La dynamique de cette étape a pour objectif de donner les métadonnées de l'image en se basant sur les besoins et les informations récupérées dans l'étape précédente.
- La quantité importante d'informations impose de réduire la taille des fichiers numériques par des méthodes de compression d'images. Cette étape convertit l'image dans tous les formats possibles.
- La profusion d'informations pose un problème de lisibilité qui peut être corrigé par la capacité à ordonner les informations. Elle peut être indexée en fonction des images, du texte ou des métadonnées.
- L'étape de diffusion regroupe tous les supports numériques permettant l'accès aux informations.
- L'ensemble de ces processus va permettre d'exploiter au mieux le contenu en fonction des motivations de chacun.

Nos travaux se focaliseront sur l'étape de numérisation. L'objectif est de proposer des solutions permettant de faciliter la suite de la chaîne. Pour comprendre les défaillances possibles de la numérisation, il faut comprendre le fonctionnement d'un scanner de document.

Dans cette introduction, nous présentons la conception d'un scanner qui est dédié à numériser des livres anciens. Nous introduirons ensuite les différents axes de recherche que nous traiterons.

Un scanner est un appareil qui enregistre un document en une image numérique. Il est bien évident qu'il n'est pas question de scanner de bureau dans ce sujet car il est trop limité en qualité d'image et productivité. Pour expliquer le fondement des appareils de numérisation industriels, nous vous présentons les deux types de numérisation de documents :

- **le scanner linéaire** - Ce type d'acquisition est faite ligne par ligne. Le capteur ou l'objet se déplace permettant de numériser l'ensemble du livre. Nous présentons le processus dans la figure 1.2 où la zone d'acquisition (la zone rouge) est une ligne sur la scène. Cette ligne se déplace pour numériser l'ensemble de l'objet.

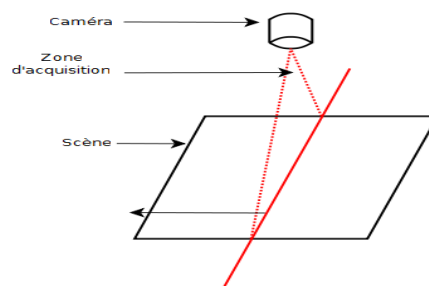


FIGURE 1.2: Scanner linéaire

- **le scanner matriciel** - Ce type d'acquisition est faite en une fois. Celle-ci consiste à positionner le capteur et de faire l'acquisition du document. Dans la figure 1.3, nous observons bien que la scène est entièrement incluse dans la zone à capter.

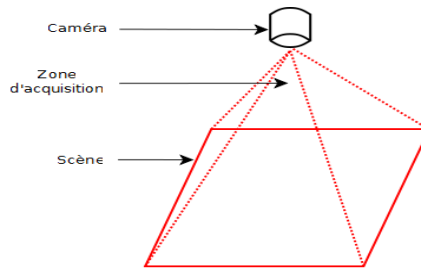


FIGURE 1.3: Scanner matriciel

Les différents types d'acquisition ont des avantages et des inconvénients. Par exemple, le scanner matriciel est très rapide mais il ne peut être aussi précis qu'un scanner linéaire qui fait l'acquisition ligne par ligne. La gestion de l'éclairage est différent pour les deux types de scanner. Effectivement, utiliser un éclairage artificiel sur une longueur, qui suit le mouvement et proche du livre à numériser est accessible (cf. figure 1.4) ; par contre, le contrôle de l'éclairage sur un scanner matriciel est moins évident car il est possible d'avoir des interférences avec l'extérieure et la forme du livre (cf. figure 1.5).

La numérisation de livres anciens consiste à poser le livre ouvert sur un support afin de protéger la reliure du livre et aussi d'éviter les contacts possibles avec les pages.

Nous notons que les formats de numérisation utilisés par la société *i2S* varient et les formats standard sont du format *A3* ($297mm \times 420mm$) pour les matriciels, au format double *A0* ($841mm \times 1189mm$, soit 2 fois le format *A0*) pour les linéaires et des caméras sont utilisées pour les deux types de numérisation.



FIGURE 1.4: Scanner linéaire - SupraScan*

Dans cette thèse, nous traiterons trois problématiques correspondant aux trois chapitres : la correction d'éclairage pour un scanner sans contrôle de l'éclairage, un contrôle de netteté après acquisition et enfin un moyen d'améliorer la calibration colorimétrique pour un scanner contrôlant l'éclairage.



FIGURE 1.5: Scanner matriciel - CopiBook*

Le première chapitre portera sur les scanners matriciels ne contrôlant pas l'éclairage. La maîtrise de l'éclairage pour un tel appareil est possible mais le coût de l'appareil augmente. Ce dernier dispose bien souvent que de l'éclairage ambiant, c'est-à-dire de l'éclairage du lieu où se situe le scanner. Il suffit d'une simple numérisation pour se rendre compte de l'impact sur la qualité de l'image dans ces conditions comme nous le voyons sur l'image 1.6.

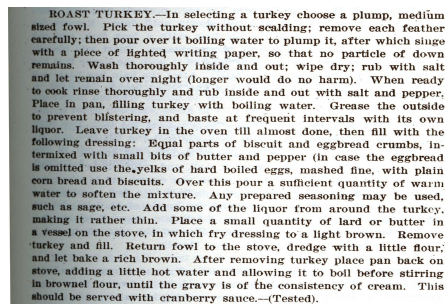


FIGURE 1.6: Nous présentons, ici, une zone d'une image ayant un problème d'éclairage. Nous observons que le côté gauche de l'image est plus foncée que le côté droit alors que le fond du document est uniforme.

Les conséquences du problème d'illumination sont doubles. En effet, d'une part, l'image peut être désagréable à voir puisque l'image sur notre écran est très différente de l'image que nous voyons, d'autre part, ces défauts modifiant le contenu, plus particulièrement les effets d'ombre sur les images. Cela a pour conséquence d'obtenir une image surexposée (ou sous-exposée). Une image sur-exposée ou un texte mal-segmenté sont des problématiques pouvant se répercuter sur la valorisation du livre.

L'entreprise avait pour objectif de corriger les non-uniformités engendrées par un problème d'éclairage.

Dans ce chapitre, nous verrons les moyens qui existent pour répondre à cette problématique, ainsi que la méthode que nous proposerons.

Le deuxième chapitre se situera dans l'étape de numérisation, nous proposerons de contrôler la qualité de l'image avant de lancer la suite de la chaîne de numérisation de POLI-NUM. Le système de productivité peut engendrer des erreurs humaines où la calibration

de l'appareil n'a pas été faite correctement. L'image deviendra souvent flou. Une telle image aboutira à une analyse de l'information moins précise, voir même impossible. Ce contrôle est aussi indispensable car il n'est pas possible d'attendre la fin de la chaîne pour se rendre compte d'une erreur. De plus, plus la numérisation est nette, plus la distinction entre les éléments sera bonne et alors l'analyse sera pertinente.

Nous allons donc nous intéresser à quantifier le flou d'une numérisation de livres anciens. Cette analyse se fera donc après acquisition. Nous verrons dans ce second chapitre plusieurs approches de la problématique et nous proposerons de les combiner afin d'affiner la mesure possible.

Le dernier chapitre concerne la calibration couleur d'un scanner. C'est une étape qui se fait lors du réglage de l'appareil à l'aide d'une mire colorimétrique dont les références sont connues. La motivation de cette partie est de représenter avec une précision suffisante, c'est-à-dire non perceptible par notre œil, les images incluses dans les livres. La représentation colorimétrique d'une image est importante car il faut avoir la sensation que l'image correspond aux couleurs utilisées dans le livre. De plus, l'impression peut être différente d'une époque à une autre, et il faut que la couleur puisse inclure cette information.

Afin de représenter les mêmes couleurs d'un appareil à un autre, il faut pouvoir se soumettre aux normes de la gestion de la couleur. Le but de POLINUM est de diffuser les fonds patrimoniaux sur n'importe quel support. Il faut alors que les couleurs retranscrits par l'appareil de numérisation soient les plus fidèles possibles pour accomplir cette tâche. Pour éviter ce type de problème, nous allons reprendre dans ce chapitre les notions de la gestion de la couleur, expliquer comment calibrer et comment appliquer une transformation qui doit respecter les contraintes industrielles que nous énumérerons.

L'ensemble de ces problématiques sont dans le thème de la qualité d'image ; chacune d'entre elles se situent hors de l'étape de d'acquisition. La problématique de la qualité d'image est vaste et nous traiterons ici quelques points à l'image du projet POLINUM

Chapitre 2

La correction d'illumination

2.1 Introduction

La qualité de numérisation est dépendante des conditions d'éclairage du document à numériser. Si l'éclairage est homogène sur l'ensemble du document alors les défauts d'illumination ne seront pas visible.

Les scanners matriciel et linéaire doivent impérativement contrôler l'éclairage pour obtenir uniquement l'information du livre. L'éclairage de la scène est fait à l'aide d'un éclairage artificiel. Une mire blanche de référence peut aider à calibrer l'éclairage pour corriger les défauts d'illumination. Toutefois, il existe une différence entre le scanner matriciel et le scanner linéaire. Le scanner linéaire utilise un éclairage artificiel qui suit le capteur afin de produire un éclairage constant. L'éclairage est proche du document, il est alors précis et peut être étalonné. Pour le cas du scanner matriciel, l'éclairage est plus complexe a maîtriser. Effectivement, l'éclairage doit se situer à l'extérieure du champ de vision de la caméra ce qui implique que l'éclairage artificiel est relativement loin de l'objet, par conséquent d'autres sources de luminosité (éclairage de la pièce, soleil, utilisateur) interfèrent sur l'éclairage de la scène. Le réglage de l'éclairage peut être étalonné à nouveau par une mire blanche mais il devient nécessaire de le faire fréquemment à cause de la variation de l'éclairage ambiant. Le deuxième problème provient de la forme du livre, celle-ci a un impact important parce qu'il n'est jamais plat et, par conséquent, l'illumination varie. L'étalonnage fait précédemment devient incorrecte dans ce cas. Nous pouvons observer deux images obtenues sans réglage de l'illumination dans la figure 2.1 et dans la figure 2.2 pour illustrer les variations d'éclairages possibles.

Même si la maîtrise de l'éclairage est un élément important, dans le cadre des acquisitions faites par un scanner matriciel, il sera nécessaire de corriger l'éclairage en utilisant



FIGURE 2.1: Exemple d'un défaut d'éclairage ambiant



FIGURE 2.2: Exemple d'un défaut d'éclairage créé par l'utilisateur

des méthodes de traitements d'images.

I2S veut proposer un algorithme de correction d'éclairage pour leur gamme de scanner matriciel. Le cahier des charges imposait que la correction d'éclairage puisse transformer le fond du livre en une région avec un éclairage homogène. Elle souhaite que la transformation puisse s'appliquer sur des livres anciens et récents. Le temps de calcul doit être faible, l'objectif est de traiter en moins de 2 secondes pour une page de format *A3* en *300DPI* (soit environ *9M* de pixels).

Nous considérons que le fond du document d'un livre ancien est découpé en deux à cause de la reliure qui peut être usée ou épaisse, ce qui sépare la page de gauche et de droite. Dans le cas du livre récent, il peut exister plusieurs fonds, il n'y a pas de séparation nette entre les images et le fond du livre, et la qualité du papier génère des zones de saturation ne permettant pas de différencier certaines informations.

Nous allons étudier à présent deux approches de l'état de l'art, puis nous proposerons une nouvelle approche.

2.2 Correction d'illumination

Une image peut être obtenue à partir de deux informations : la luminance et la réflectance (cf. [Jain, 1989]). La luminance est créée à partir des interactions de la surface de la scène et de l'illumination. La réflectance décrit comment chaque point reflète la lumière de l'objet.

Le modèle considère que l'image $I(x, y)$ est déterminée par l'image de la réflectance $R(x, y)$ et l'image de luminance $L(x, y)$ selon la relation 2.1.

$$I(x, y) = R(x, y)L(x, y) \quad (2.1)$$

Ces deux images sont appelées les images intrinsèques.

La première approche corrigera les défauts d'éclairage en cherchant à différencier la composante réflectance R de la composante luminance L . Nous verrons, dans une première partie, les démarches et hypothèses permettant d'effectuer cette opération sur tout type d'image.

La deuxième approche consiste à déterminer uniquement la composante L dans le cas de la numérisation de livre ancien, puis à corriger l'image à l'aide de cette information. La composante L peut être déterminée par une simple mire blanche permettant d'obtenir une telle information dans notre contexte de numérisation. A partir de cette information, nous pouvons corriger l'image en se basant sur la relation précédente pour obtenir :

$$I(x, y) = R(x, y)L(x, y) \quad (2.2)$$

Nous citerons d'autres modèles de transformation proposés dans l'état de l'art qui remplacent la transformation qui est ici la division par L . Le résultat que nous souhaitons obtenir est une carte de correction d'éclairage, c'est-à-dire l'ensemble des coefficients de correction d'éclairage pour tous les pixels obtenus par la transformation choisie.

Ainsi, nous terminerons sur les méthodes qui permettent d'extraire la composante de luminance dans le contexte de la numérisation, c'est-à-dire la segmentation du fond du document et sa reconstruction.

2.2.1 Les images intrinsèques

Les images intrinsèques sont la décomposition de l'image en deux caractéristiques luminance-réflectance, cette approche est valable pour n'importe quel type d'image. Cette hypothèse, que nous pouvons retrouver dans [Jain, 1989], aboutit sur des méthodes ayant pour

but d'avoir uniquement la composante de la réflectance. Nous présentons un exemple caractéristique de la méthode dans la figure 2.3.

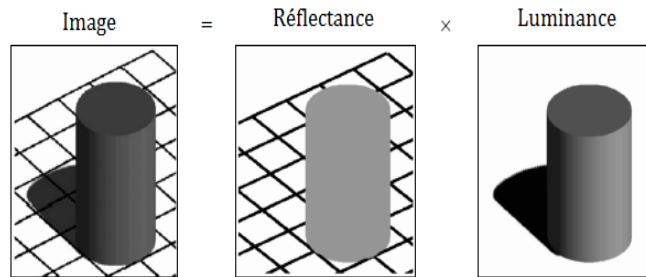


FIGURE 2.3: Décomposition de l'image en images intrinsèques [Weiss, 2001]

L'algorithme des images intrinsèques a pour objectif de labéliser les pixels à partir des dérivées de l'image. Cette information est la plus pertinente pour déterminer les changements de luminance.

L'article de [Tappen et al., 2005] suggère d'utiliser un classifieur supervisé permettant de distinguer les informations de luminance et de réflectance avec les deux descripteurs suivants :

Le premier descripteur est basé sur l'information colorimétrique. Si la luminance varie, alors, il existe une relation de proportionnalité entre les deux points adjacents, dans le cas inverse, il s'agit de réflectance. Si on pose deux triplets RVB c_1 et c_2 où seul la luminance varie, alors il existe un réel α tel que $c_2 = \alpha c_1$.

Le deuxième descripteur est basé sur les informations de niveaux de gris. A partir du filtrage de [Tieu and Viola, 2004] qui permet de classifier les textures d'une image, les auteurs déterminent si le résultat correspond à une information de luminance ou de réflectance.

Avec l'ensemble des pixels de luminance, il est possible de recomposer une image des variations d'illumination, et donc de retrouver l'image réflectance en faisant la différence. Le coût avantage/temps de calcul nous paraît trop faible pour mettre en place un tel système.

Toutefois, ils existent des alternatives à cette théorie comme le filtrage homomorphique et le filtrage Retinex. Ces alternatives se basent sur des hypothèses concernant la composante L afin de la lisser.

Filtrage homomorphique

Ce filtrage fréquentiel permet de supprimer les défauts d'illumination non uniforme et de rehausser les contrastes. Nous retrouvons les explications de cette méthode dans le livre de A. K. Jain [Jain, 1989].

En se basant sur la décomposition luminance-réflectance de l'image, l'information se décompose en deux parties : la réflectance influe sur les moyennes et hautes fréquences, et la luminance sur les basses fréquences. La composante d'illumination subit de faibles variations locales et l'objectif est de supprimer ces fréquences pour atténuer les irrégularités d'illumination, et par la même occasion de réajuster le contraste.

L'algorithme utilise, tout d'abord, le *log* pour séparer les composantes luminance et réflectance.

Soit la décomposition de l'image I :

$$\log(I(x, y)) = \log(R(x, y)) + \log(L(x, y)) \quad (2.3)$$

Un filtre passe-haut F est appliqué sur le signal $\log(I(x, y))$ (voir figure 2.4). Soit F , tel que :

$$F(w_x, w_y) = (r_H - r_L) \left(1 - \exp^{-\left(\frac{w_x^2 + w_y^2}{2\delta_w^2}\right)} \right) + r_L \quad (2.4)$$

avec (w_x, w_y) les coordonnées fréquentielles, r_H et r_L , les coefficients fréquentiels où r_L détermine le gain des basses fréquences (< 1) et r_H celui des moyennes et hautes fréquences (> 1) et δ_w représente la fréquence de coupure.

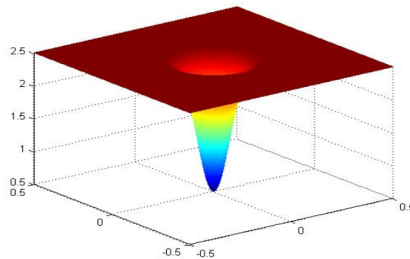


FIGURE 2.4: Filtre passe haut avec les paramètres $r_H = 2.5$, $r_L = 0.5$ et $\delta_w = 0.2$

Après filtrage, l'information de luminance a été atténuée car elle est essentiellement composée de basse fréquence ; alors que la composante R a été renforcée par le coefficient r_H . Par conséquent, on considère, selon cette hypothèse, que la composante de la luminance s'annule et nous obtenons :

$$F[\log(I(x, y))] \approx \log(R(x, y)) \quad (2.5)$$

Pour revenir dans le domaine spatial, il suffit d'appliquer, dans une dernière étape, la fonction exponentielle. Cette méthode, appelée filtrage homomorphique, est résumée dans la figure 2.5.

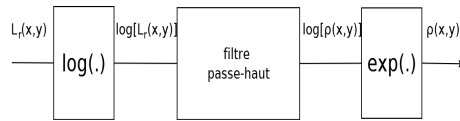
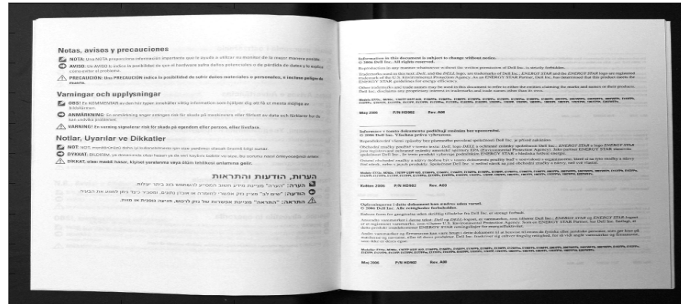
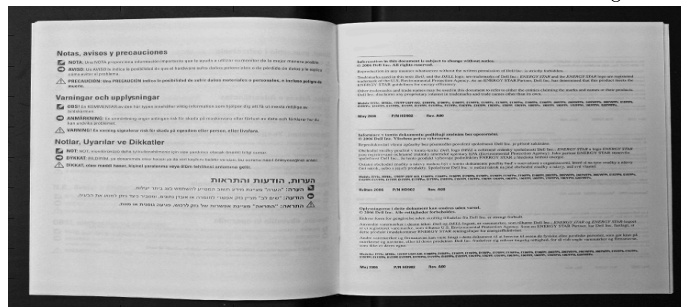


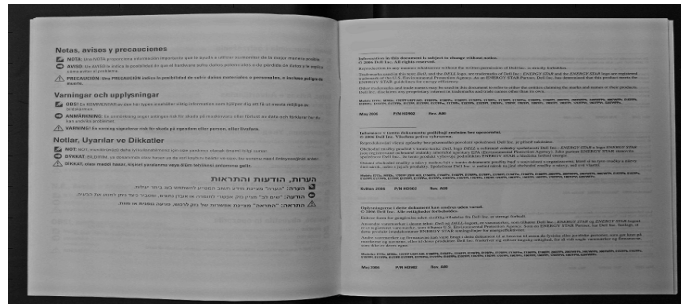
FIGURE 2.5: Schéma des différentes étapes du filtre homomorphe



a) Manuel d'utilisation scanné dans des conditions d'éclairage ambiant



b) Image filtrée ($\delta_1 = 0,3$)



c) Image filtrée avec différents paramètres ($\delta_2 = 0,7$)

FIGURE 2.6: Application du filtrage homomorphe dans le but de corriger l'éclairage

Nous avons appliqué deux traitements que nous pouvons voir dans la figure 2.6 avec les paramètres communs $r_H = 1,1$ et $r_L = 0,75$ pour les deux images et $\delta_1 = 0,3$ et $\delta_2 = 0,7$. Nous voyons que le résultat de la première image 2.6.b tend à uniformiser le fond du document (mais pas totalement) et le document devient plus foncé à cause des coefficients r_H et r_L . Le deuxième filtrage (2.6.c) a un fond plus homogène, mais l'apparition d'artefacts aux extrémités du livre et un effet ondulateur sur l'ensemble de l'image font que la transformation ne peut être utilisée dans le cadre du document. De plus, les coefficients r_H et r_L devraient accentuer cet effet.

Les hypothèses du filtrage ne sont pas adéquates à notre problématique. Effectivement, l'hypothèse considère que l'illumination varie lentement ce qui permet de corriger les basses fréquences d'illumination d'où l'utilisation d'un filtre passe-haut, or le défaut d'éclairage combiné à la forme du livre donne une variation beaucoup plus importante de l'éclairage. L'information de luminance, dans notre contexte d'éclairage, influe sur les basses et moyennes fréquences. Or l'information de réflectance se situe aussi au niveau des moyennes fréquences, par conséquent, un lissage ne peut pas résoudre ce problème sous peine de modifier les deux types d'informations.

Le cahier des charges validait la transformation lorsque le fond du document tendait vers une même valeur, ce filtrage répond à ce critère mais l'apparence d'artefact ne permet pas de le valider.

Retinex

La théorie du Retinex est basée sur le même concept de la décomposition de l'image en luminance-réflectance, mais elle se différencie par rapport à l'hypothèse faite par le filtre homomorphique. En effet, elle ne considère pas qu'il existe une corrélation entre les bandes de fréquences et l'illumination car elle ne prend pas en compte la vision humaine. Ces deux informations ne sont pas mélangées mais elle sont comparées pour générer une sensation visuelle. Les fondements du modèle appelé Retinex a été proposé par Edwin Land que nous retrouvons dans l'article [Land, 1964], il redéfinit un modèle de la lumière et de la perception des couleurs. Par la suite, plusieurs propriétés sont apparues se basant sur ce modèle permettant d'extraire la réflectance.

L'article de [Jobson et al., 1997] est une des variantes la plus connue. Cette variante repose sur deux propriétés. La première propriété propose un traitement en prenant en compte l'ensemble des informations colorimétriques c'est-à-dire que les traitements vont s'appliquer sur les trois canaux R , V et B ; la séparation luminance-réflectance est souvent basée sur des hypothèses qui considère qu'un niveau de gris correspond à une information de luminosité mais les hypothèses deviennent incorrecte lorsqu'il y a un éclairage artificiel où une couleur est plus importante que les autres. Par conséquent, le traitement de chaque couleur est séparé et il n'y a pas de conversion RVB en niveaux de gris possible.

La deuxième propriété est la théorie de Hurlbert (cf. [Hurlbert, 1989]) qui montre que le problème d'éclairage a une solution qui a une forme centrale/encerclement (center/-surround dans les articles, que nous pouvons associer à la propagation de la lumière sur

la scène) qui permet d'estimer une réflectance relative en fonction des conditions d'éclairages. Cette propriété met en avant que pour déterminer l'information de luminance en fonction de la scène, il est pertinent d'utiliser plusieurs échelles pour pouvoir l'identifier.

Dans la première partie de l'article [Jobson et al., 1997], la forme du Retinex est présentée, celle-ci est dite multiscale center/surround (MSR) et elle est décomposée en deux : la fonction center/surround et la notion multi-échelle. Pour une échelle, qui est appelée single scale retinex (SSR), nous avons l'application de la fonction de center/surround qui est la suivante :

$$R_i(x, y) = \log(I_i(x, y)) - \log[F(x, y) * I_i(x, y)] \quad (2.6)$$

où $R_i(x, y)$ est la sortie Retinex, $I_i(x, y)$ est l'image pour sa i-ème bande colorimétrique. "*" est l'opérateur de convolution, et F est la fonction de surround.

$$F(x', y') = K \exp \left[-\frac{x'^2 + y'^2}{\sigma^2} \right] \quad (2.7)$$

où σ détermine l'échelle (et la force de la gaussienne), et ce filtre doit être normalisé par K qui vérifie :

$$\sum_{x'} \sum_{y'} F(x', y') = 1 \quad (2.8)$$

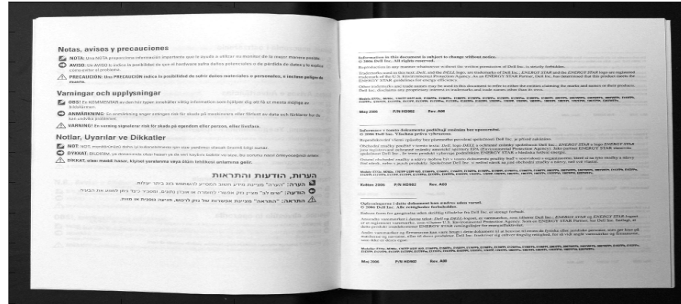
Le résultat du multiscale retinex est une pondération entre les différentes single scale retinex. Nous obtenons alors :

$$R_{M_i}(x, y) = \sum_{n=1}^N w_n R_{n_i}(x, y) \quad (2.9)$$

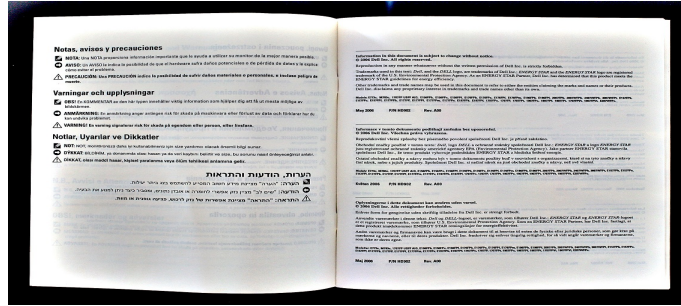
où N est le nombre d'échelle. $R_{n_i}(x, y)$ est la i-ème composante de la n-ème échelle. $R_{M_i}(x, y)$ est la i-ème composante de la sortie du multiscale retinex, et w_n est le poids associé à la n-ème échelle.

La transformation du Retinex 2.9 est déterminée par les couples : valeur d'échelle et coefficient de pondération.

La deuxième partie de l'article utilise des coefficients de pondération identique pour chacune des échelles et deux paramètres d'échelles : une petite et une grande échelle. L'auteur montre que l'utilisation d'une grande échelle lisse les zones uniformes et crée des formes d'halo au niveau des forts contours lorsque l'on associe les différentes échelles (voir figure 2.7). Les deux paramètres sont trop différents, il paraît évident qu'il faut utiliser plus que deux échelles.



a) Manuel d'utilisation scanné dans des conditions d'éclairage ambiant



b) Image filtrée par Retinex

FIGURE 2.7: Correction d'éclairage à partir d'un filtre retinex

La suite de l'article met en évidence la manipulation des coefficients de pondération étant donnée que c'est à l'aide de ces variables que l'on peut changer la dynamique de l'image. L'auteur montre des images pour lesquelles le résultat n'est pas convaincant car l'image devient trop grise sur des images où il y a une grande différence de luminosité. Il suggère alors une méthode permettant d'appliquer des coefficients supplémentaire permettant de corriger les couleurs (CR, color restoration), cette dernière permet essentiellement de normaliser l'espace colorimétrique et de contrôler la saturation du rendu final. L'auteur propose de calculer de la manière suivante : l'espace de couleur est transformé dans un premier temps :

$$I'_i(x, y) = \frac{I_i(x, y)}{\sum_{i=1}^N I_i(x, y)}, i = 1, \dots, N \quad (2.10)$$

Pour ensuite calculer la fonction de la correction de couleur est définie par $C(x, y)$

$$C(x, y) = f[I'_i(x, y)] \quad (2.11)$$

où $f[]$ est la fonction qui permet de normaliser l'espace colorimétrique. L'article ne donne pas de définition mais l'ensemble des articles fait par Rahman, Jobson et Woodell [Rahman et al., 1996], [Rahman et al., 1997] et [Rahman et al., 2004] discute en partie

sur cette fonction.

Ce dernier propose :

$$C_i(x, y) = \log \left(\frac{NI_i(x, y)}{\sum_{i=1}^N I_i(x, y)} \right) \quad (2.12)$$

Le multiscale center/surround color restoration (MSRCR) s'écrit alors :

$$R_i(x, y) = C_i(x, y) \sum_{k=1}^k W_k \log(I_i(x, y)) - \log(F_k(x, y) * I_i(x, y)) \quad (2.13)$$

L'usage d'une telle fonction va nécessiter d'améliorer plusieurs points :

- le système de pondération qui permettra d'obtenir un fond clair et homogène.
- déterminer une échelle pour ne pas avoir de présence de halo aux niveaux des contours, ainsi que le fond noir du support.
- de modifier la fonction de normalisation de l'espace colorimétrique.

La décomposition de l'image en luminance-réfectance propose deux alternatives : la première consiste à faire une analyse des dérivées de l'image ce qui permet de labéliser les pixels de l'image en luminance ou en réfectance, à partir de l'ensemble des pixels de luminance, il est alors possible de construire un profil de luminance et de le soustraire à l'image originale pour récupérer la réfectance ; la deuxième suggère de corriger directement ces défauts à partir du logarithme de l'image selon certaines propriétés faites de la lumière sur la scène.

L'autre approche est basée sur le contexte de la numérisation de livre. L'objectif est d'identifier les variations d'illumination puis de les corriger. Nous verrons dans la sous-section suivante la mise en place d'un tel système de correction.

2.2.2 La Carte de correction d'éclairage

La carte de correction d'éclairage est une image contenant pour tous les pixels les gains à appliquer pour corriger l'illumination. Ces gains sont obtenus en fonction de la mesure de l'illumination, de la transformation utilisée et de la valeur du fond.

Pour déterminer les changements d'illumination il faut extraire les variations d'éclairage de l'image. Pour cela, il y a deux possibilités : soit nous utilisons une mire qui est une page blanche de référence, soit nous faisons l'hypothèse que le fond du document est similaire à la mire.

Lorsque nous utilisons une mire, les mesures vont permettre d'identifier les pixels du fond ainsi que tous les défauts d'illumination ne permettant pas d'avoir une mire parfaitement

homogène, par contre, lorsque nous utilisons le fond du document il n'est pas certain que la valeur du fond soit blanc et le fond n'est connu que partiellement.

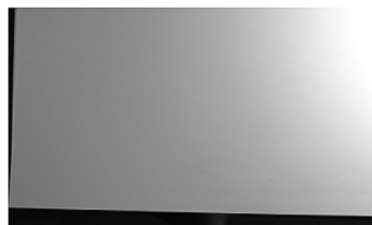
Pour appliquer la correction, nous pouvons choisir différentes transformations utilisées pour gérer l'illumination. Il existe différentes approches, mais nous nous concentrerons sur les opérateurs permettant de résoudre l'équation :

$$R(x, y) = I(x, y)/L(x, y) \quad (2.14)$$

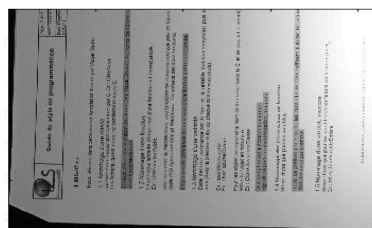
2.2.2.1 La transformation

L'opérateur de l'équation 2.2 est la division, celui-ci a une influence sur la transformation des pixels qui ne sont pas des pixels de fond. Il faut par conséquent utiliser un opérateur respectant des propriétés d'illumination.

Nous proposons de comparer la soustraction et la division pour corriger les défauts d'illumination. Pour cela, nous prenons deux numérisations : la première I_f est une page blanche (cf. la figure 2.8.a) servant de référence et la seconde I (cf. la figure 2.8.b) est la numérisation d'un document. La résolution du problème consiste à considérer $L(x, y)$ comment étant la page blanche et $I(x, y)$ la numérisation du document.



a) page blanche numérisé servant de référence



b) document numérisé dans les mêmes conditions

FIGURE 2.8: Mise en place d'une procédure de suppression de fond à l'aide d'une référence

L'opérateur $-$ est utilisé pour faire une suppression de fond, nous prenons la numérisation que nous soustrayons la page blanche, c'est-à-dire que l'image corrigée est obtenue par :

$$I_c = I - C_- \quad (2.15)$$

où

$$C_- = I_f \quad (2.16)$$

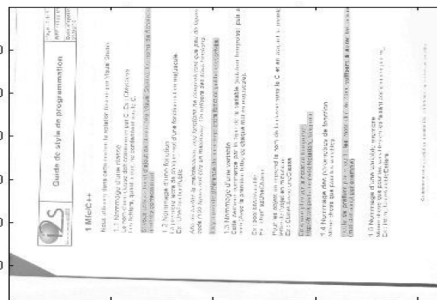
Nous obtenons le résultat à la figure 2.9.a. L'opérateur \times consiste à calculer la carte de correction en prenant la page de référence et une valeur de fond déterminée par l'utilisateur, en l'occurrence nous avons choisi 245 qui correspond aux valeurs du blanc de la mire colorimétrique utilisée (voir chapitre 4). Nous appliquons les coefficients pour chaque pixel sur la numérisation. L'image corrigée est alors :

$$I_c = I * C_\times \quad (2.17)$$

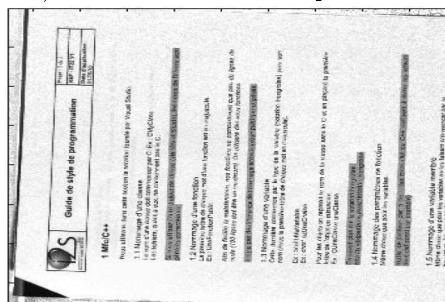
où

$$C_\times(x, y) = \frac{I_f(x, y)}{245} \quad (2.18)$$

Nous obtenons le résultat à la figure 2.9.b.



a) Correction basée sur l'opérateur –



b) Correction basée sur l'opérateur \times

FIGURE 2.9: Résultats avec les opérateurs euclidiens

Nous constatons que l'opérateur $-$ donne un mauvais résultat alors que l'opérateur \times est plus convaincant car il conserve le non-fond et le fond converge vers une valeur uniforme. Toutefois, le fond est très bruité mais il a une bonne dynamique _le surlignage ressort plutôt bien.

Nous proposerons maintenant deux opérateurs particuliers qui sont des opérateurs sont souvent utilisés pour les problématiques d'illumination : l'opérateur gamma et l'opérateur Logarithmic Image Processing (LIP).

La Gamma correction

Il est important de rappeler que la relation qu'il existe entre l'intensité lumineuse et la sensibilité humaine est basée sur les deux lois psycho-visuelles de Weber [Hecht, 1924] et de Fechner [Hecht, 1924]. Ces lois ont permis d'établir la relation entre l'intensité perçue et la luminosité qui est proportionnelle au logarithmique en base 10.

Ce processus est utilisé dans la création de mire et permet d'établir une relation entre le patch de niveau de gris et la sensation perçue par l'œil qui peut-être estimé par la fonction puissance gamma. Le gamma caractérise alors le contrôle de la lumière faite par l'acquisition jusqu'à l'affichage de l'image sur l'écran.

La gamma correction est généralement utilisée pour compenser les défauts de luminance des écrans de télévision. De ce fait, nous utiliserons la gamma correction pour corriger les défauts d'illumination.

Soit V_e la valeur d'entrée comprise dans l'intervalle $[0, 1]$, soit V_s la valeur de sortie comprise dans le même intervalle et γ la puissance. La transformation gamma s'écrit :

$$V_s = V_e^\gamma \quad (2.19)$$

L'application classique du gamma est généralement utilisée dans une transformation globale. C'est une transformation qui va rehausser les niveaux de gris, mais ne n'augmentera pas nécessairement beaucoup les niveaux de blanc.

Nous notons que la valeur γ est calculée, pour toutes les coordonnées, en faisant le ratio des logarithmiques du fond théorique sur le fond mesuré, ainsi nous avons :

$$C_\gamma(x, y) = \frac{\log(I_f(x, y))}{\log(245)} \quad (2.20)$$

Ainsi, la correction d'éclairage s'applique de la manière suivante :

$$I_c = I^{C_\gamma} \quad (2.21)$$

Les opérateurs du modèle Logarithmic Image Processing

Les auteurs J.-C. Pinoli et M. Jourlin. ont traité le sujet à plusieurs reprises ([[Jourlin and Pinoli, 1995](#)], [[Mayet et al., 1996](#)], [[Pinoli, 1997](#)] et [[Pinoli and Debayle, 2006](#)]). Le LIP a été introduit en examinant le problème de l'addition des intensités de deux images, que ce soit en lumière transmise ou dans le cadre de la perception visuelle humaine. Le modèle LIP est ainsi basé sur les lois physique et respecte les deux lois psycho-visuelles.

Cet opérateur est pertinent car il nous permet de supprimer une image de variation d'illumination à notre acquisition courante en respectant ces différentes lois.

Dans l'article [[Mayet et al., 1996](#)], l'intensité d'une image (noté $I(x, y)$) est représentée par sa fonction en tons de gris associée (noté $f(x, y)$). Une telle fonction est définie sur un domaine spatial notée D et prend ses valeurs dans l'intervalle réel borné $[0, 256[$. Dans le cadre de l'imagerie à lumière transmise une fonction en tons de gris $f(x, y)$ est considérée comme un filtre de l'intensité lumineuse et est défini par la relation suivante :

$$f(x, y) = 256 \left(1 - \frac{I(x, y)}{I_0(x, y)} \right) \quad (2.22)$$

où $I_0(x, y)$ et $I(x, y)$ désignent respectivement l'intensité incident et l'intensité transmise. I_0 est la valeur du blanc ou si elle est inconnue 256. D'après cette définition, une fonction en tons de gris $f(x, y)$ représente l'opacité du milieu absorbant considéré, puisque le rapport $\frac{I}{I_0}$ n'est autre que la réflectance de ce milieu. Les limites de l'intervalle $[0, 256[$ des tons de gris ont alors la signification physique suivante : la valeur 0 correspond à un point totalement transparent et la valeur 256 à un point totalement opaque. Les lois du modèle LIP sont alors définies de la façon suivante :

Soit $f(x, y)$ et $g(x, y)$ deux fonctions de tons de gris et λ un nombre réel positif.

$$f(x, y) \oplus g(x, y) = f(x, y) + g(x, y) - \frac{f(x, y)g(x, y)}{256} \quad (2.23)$$

$$\lambda \odot f(x, y) = 256 - 256 \left(1 - \frac{f(x, y)}{256} \right)^\lambda \quad (2.24)$$

$$f(x, y) \ominus g(x, y) = 256 \left(\frac{f(x, y) - g(x, y)}{256 - g(x, y)} \right) \quad (2.25)$$

L'application de l'opérateur LIP que nous comparons aux classiques : $-$ et \times . Pour cela, nous reprendrons les exemples précédemment utilisés (cf. la figure 2.8) et nous y appliquons les opérateurs LIP.

L'opérateur \ominus est calculée de la manière suivante :

$$I_c = I \ominus I_f \quad (2.26)$$

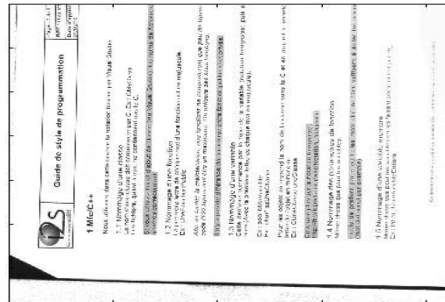
,et l'opérateur \ominus est calculée ainsi :

$$I_c = C_{\ominus} \odot I \quad (2.27)$$

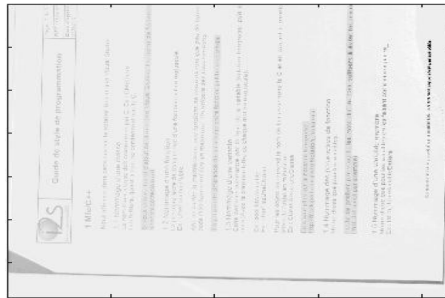
où

$$C_{\ominus}(x, y) = \frac{\log\left(\frac{M-f_{f_t}}{M}\right)}{\log\left(1 - \frac{f_{I_f}(x,y)}{M}\right)} \quad (2.28)$$

avec, f_{f_t} la valeur de fond souhaité en tons de gris, et f_{I_f} l'image de la page blanche en tons de gris.



a) Correction basé sur l'opérateur \ominus



b) Correction basé sur l'opérateur \odot

FIGURE 2.10: Mise en place d'une procédure de suppression de fond à partir de l'opérateur LIP

L'opérateur \ominus a une dynamique intéressante quand nous observons l'atténuation du bruit du fond comparée aux autres images et les noirs sont plus accentués. Par contre, nous distinguons une différence au niveau des surlignages, fait qui n'apparaît pas dans l'opérateur \times . Ce phénomène s'explique par la définition de \ominus dont le calcul accentue la différence entre le fond (g dans l'équation 2.25) et la valeur de l'encre.

Pour exemple, une opacité de 55, soit un pixel dont le fond est de 200, nous obtiendrons un pixel 105 à 135 après correction et un pixel de 55 à un pixel de 71. Nous avons une augmentation de plus en plus grande lorsque la valeur se rapproche de la valeur du fond. L'opérateur \odot donne un résultat très moyen où la dynamique de l'image est très faible.

Nous proposons une comparaison avec l'opérateur multiplicatif, la gamma et la multiplication du LIP dans la figure 2.11.

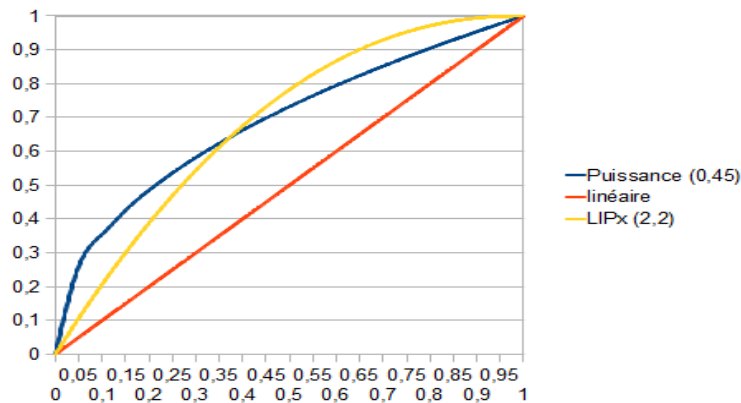


FIGURE 2.11: Comparaison des opérateurs \times , γ et \odot

Comme nous l'avons déjà noté, le gamma et le LIP sont relativement similaires au niveau des gains sur les gris étant donné qu'ils sont basés sur la même loi. Toutefois, nous avons vu que l'opérateur multiplicatif donne de résultats intéressants sur le rendu de l'image malgré que les gains aux niveaux des gris soient linéaires.

Il existe aussi d'autres approches basées sur le contenu de l'image comme par exemple l'égalisation de l'histogramme ou encore le contrôle de la dynamique de l'image. Ces méthodes se résument, dans le cas de la numérisation, à faire en sorte que les gris foncés deviennent noirs et les gris clairs deviennent blancs. Cela va avoir pour conséquence d'atténuer l'effet d'ombre sur le document et d'augmenter le contraste du texte. Cette transformation se généralise et elle est appelée la courbe en "S" à cause de sa forme.

2.2.2.2 La binarisation dans le contexte des défauts d'illumination

Lorsque nous avons introduit notre problématique, nous avons mis en évidence que l'utilisation d'une mire blanche pour calibrer un scanner matriciel est complexe à mettre en place à cause des variations d'illumination. Il est possible d'appliquer ce type de protocole mais il doit être fait à chaque modification de l'éclairage. La numérisation étant industrielle, ou le nombre de numérisation à l'heure est un critère non-négligeable, il n'est pas possible d'effectuer la procédure trop souvent. Par conséquent, utiliser une approche basée sur la récupération du fond pour chaque image est plus pertinent.

Le fond étant a priori homogène pour une page, cette information nous apporte une solution possible dans le contexte de la numérisation. En partant de cette hypothèse, il est alors possible de considérer que les variations du fond soit dû à des défauts d'illumination. Notre problématique principale est de pouvoir segmenter le fond du document et

de le reconstruire en totalité afin de récupérer les variations d'intensité lumineuse. Les deux méthodes de segmentation les plus logiques pour récupérer le fond du document sont la segmentation par approche de régions et la segmentation basée sur le seuillage des pixels en fonction de leur intensité. A partir de la récupération du fond partiel, nous pouvons établir un profil 1D pour un scanner linéaire ou 2D pour un scanner matriciel pour créer la carte de correction.

Notre objectif est alors de déterminer si un pixel appartient au fond ou non. La segmentation la plus intuitive pour notre problème est la binarisation. La binarisation est une méthode qui classe les pixels en deux classes : les pixels du fond et les pixels de l'encre, plus généralement du non-fond. Les algorithmes de binarisation se divisent en deux parties : les approches qui travaillent sur l'image globale et les approches qui travaillent localement sur des parties de l'image. Le livre de P. Stathis [Stathis et al., 2008] propose un état de l'art complet sur les méthodes de binarisation.

Dans le cas des méthodes globales, une partie des algorithmes [Rosenfeld and De La Torre, 1983] [Sezan, 1990] se basent sur la forme de l'histogramme des niveaux de gris. Ils recherchent deux pics et un minimum entre les deux. D'autres algorithmes [Otsu, 1975], [Kittler and Illingworth, 1985] définissent la séparation des deux pics en définissant l'histogramme comme le mélange de deux gaussiennes. Ils existent d'autres méthodes qui utilisent l'entropie de la distribution tels que [Kapur et al., 1985], [Li and Lee, 1993], [Pal, 1996] et [Li and Tam, 1998]. Malgré ces approches, il est connu qu'un problème d'illumination crée des erreurs de classification.

Nous présentons dans les figures 2.12 et 2.13, des images accompagnés de leurs histogrammes globaux respectifs.

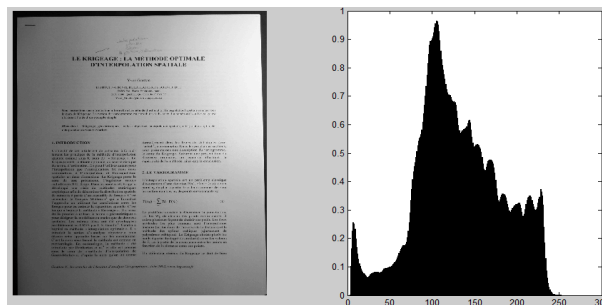


FIGURE 2.12: Premier exemple de problème d'éclairage associés à son histogramme

Dans la première image, nous avons deux pics : le premier (10) qui correspond à l'extérieur de la page et le second (100) qui représente le fond. Nous voyons bien la similitude entre

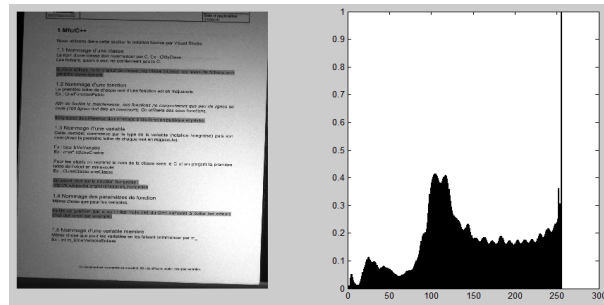


FIGURE 2.13: Deuxième exemple de problème d'éclairage associés à son histogramme

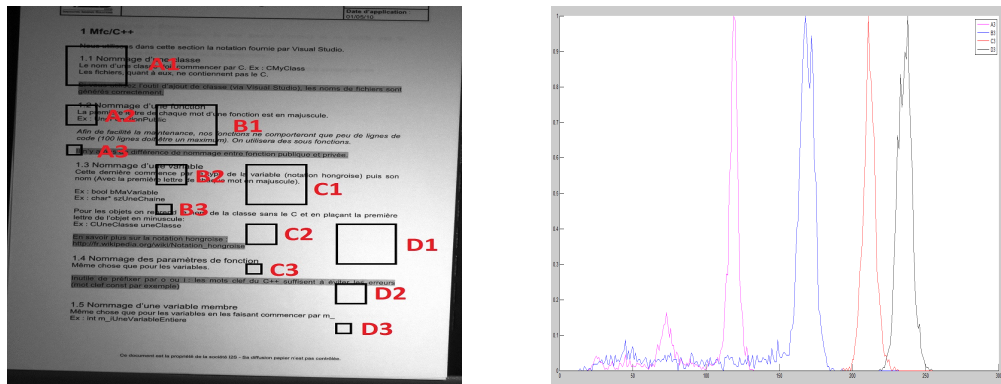
le dégradé du fond et l'histogramme. Par contre, il est impossible de connaître le bon seuil permettant de récupérer l'information du texte car il n'est pas représenté par un pic.

Dans la seconde image, nous avons quatre labélisations : le fond, le texte, les surlignages et l'extérieur du livre. Nous observons quatre pics : 10, 30, 110 et 250. Les pics à 250 et à 110 appartiennent à la même catégorie de pixel étant donné que le fond est composé de pixels saturés et de pixels de gris. Nous voyons encore une fois le noir du support et les surlignages mais l'apparence du texte n'apparaît pas et ne peut apparaître pour une telle approche. Les variations d'illumination sont importantes dans notre contexte et l'analyse de l'histogramme ne permet pas de segmenter les pixels. La binarisation du fond du document par une approche globale n'est pas concluante.

L'autre famille est dit adaptative. Cette approche consiste à calculer un seuil de classification sur tous les pixels de l'image. Pour déterminer le seuil du pixel, l'analyse est faite à partir d'une fenêtre centrée sur le pixel courant. Les algorithmes de [Sauvola and Pietikäinen, 2000], [Niblack, 1985] et de [Bernsen, 1986] utilisent des indicateurs statistiques (moyenne et écart-type). Ces approches sont plus robustes aux changements d'illumination, mais en contre-partie la fenêtre courante n'a pas systématiquement les deux types de pixels (fond et encre).

Effectivement, l'information de la valeur du pixel est influencée par les variations d'illumination, par conséquent, le pixel va augmenter ou diminuer de valeur en fonction de l'intensité de l'illumination en cette coordonnée. L'approche globale ne prend pas en compte la position puisque elle prend la totalité de la distribution de l'image pour déterminer un seuil. L'analyse statistique faite localement va alors atténuer les variations de l'illumination. La précision du seuil sera moins pertinente mais le seuil sera peu influencé par ces défauts

L'approche locale est plus pertinente car la variation d'illumination est plus faible sur des zones plus petite. Nous observons des histogrammes locaux dans l'image 2.14.



a) Image d'un document à fond blanc sur-exposée b) Répartitions des régions A3, B3, C3 et D3

FIGURE 2.14: Histogrammes locaux

Les histogrammes sont calculés à partir des fenêtres données. Nous constatons que la variation d'illumination est plus faible lorsque la région devient plus petite, la binarisation est alors plus pertinente.

Il est important de prendre en compte que l'échelle choisi va influencer sur la précision du résultat.

Nous avons pris les répartitions des fenêtres notées *A3*, *B3*, *C3* et *D3* car ce sont les plus précise dans notre exemple ; la première possède 3 labels (encre/surlignage/fond), la seconde possède 2 labels (encre/fond) et les deux dernières n'ont qu'un seul label (fond). Nous remarquons que l'analyse de l'historgramme est beaucoup plus pertinente car *A3* possède deux pics nets (110 et 75) et un pic à 25. L'historgramme *B3* possède un pic net à 170 et un pic à 50, et les deux derniers ont chacun un pic net à 210 et 245. Nous constatons que l'approche locale permet de calculer un seuil plus précis mais pose un problème de nombre de classe.

Nous pouvons considérer plusieurs cas possibles : soit l'historgramme dispose d'une classe alors il s'agit du fond ; soit il y en a deux, cela signifie que nous sommes dans le cas fond/non-fond, nous supprimerons alors les pixels du non-fond ; soit il y a plus de deux classes, il est possible que nous soyons dans un cas où la fenêtre est composée de deux fonds, ce cas ne peut alors pas être traité.

Le choix de la binarisation adaptative est pertinente pour distinguer le fond du document dans des conditions d'éclairage ambiant car l'analyse de l'historgramme met en évidence les pixels du fond. La variation de la taille de la fenêtre permet d'affiner le résultat dans la mesure où plus la taille diminue moins la région est soumise à l'influence de l'éclairage. Ce qui fait que la problématique consiste à connaître la taille de la fenêtre pour optimiser notre évaluation.

Une fois les pixels du fond détecté, il faut déterminer la totalité du fond du document afin d'y appliquer la correction pour tous les pixels. Il n'est pas question ici de reconstruire une information mais, de par les propriétés d'éclairage, il suffit simplement de lisser les informations. Par contre, il est indispensable de distinguer l'existence de plusieurs fonds (ou d'erreur) car il n'est pas possible de lisser directement les informations illuminations de deux fonds de couleurs différentes.

Nous verrons dans la partie suivante comment effectuer ces opérations.

2.3 Correction d'éclairage par une carte du fond

Nous rappelons que l'objectif est de proposer une correction d'éclairage pour une image donnée. Nous rajouterons que la méthode a des contraintes de temps et qu'elle ne doit pas générer d'artefact.

Notre objectif est de concevoir une correction d'éclairage qui a pour but de compenser les défauts de luminosité d'un scanner matriciel sans contrôle de la lumière. L'utilisation d'une page blanche comme une mire de référence montre ses limites. En effet, la forme du livre ou encore les reflets possibles sur la scène vont faire varier les défauts d'illumination. Lorsque nous appliquons la correction, les gains seront incorrectes et feront apparaître des saturations ou des variations en luminance aberrantes. Il faut alors corriger l'éclairage en fonction de l'image pour éviter ces artefacts.

Nous avons vu, dans la première partie, deux approches permettant de faire de la correction d'éclairage. La première s'applique principalement aux images naturelles alors que la seconde approche est essentiellement basée sur la segmentation du fond du document. Cette dernière nous paraît évidemment plus logique à utiliser et nous allons ainsi la développer. La carte de correction d'éclairage est déterminée en fonction du fond du document. Par conséquent, l'information d'illumination du fond va paramétrer la correction. La précision de la correction ne doit pas faire apparaître d'artefact, cela signifie que la segmentation doit avoir une précision suffisante et son temps de d'exécution doit être faible. Les images que nous cherchons à corriger l'éclairage ont au minimum 9 M de pixel (soit supérieur à $3500p \times 2500p$).

Dans le cas de la numérisation d'un document n'ayant pas de problème d'éclairage, nous pouvons définir que si p est un pixel du fond F , et I l'image numérisée, alors

$$\forall p \in F, I(p) = cte \tag{2.29}$$

La problématique du défaut d'éclairage est que l'ensemble des pixels du fond ne sont plus identiques.

L'éclairage ambiant est a priori un ensemble de source de lumière. Son impact sur le document est souvent sous la forme d'une atténuation ou accentuation de la quantité de lumière qui fait diminuer ou augmenter la valeur des pixels. La conversion de l'appareil va systématiquement chercher à les convertir par une transformation linéaire dans l'intervalle $[0, 255]$ en fonction des distributions des photons sur les cellules des capteurs. La variation d'illumination est alors continue. La correction d'éclairage que nous proposons va consister à créer une carte de luminosité en se basant sur l'image courante. Cette carte permettra de corriger l'éclairage ambiant en supposant que le ou les fond(s) sont

homogènes.

Les différentes étapes sont décrites dans la figure 2.15. Nous avons ainsi une image numérisée de livre ancien en entrée pour aboutir à une image dont l'éclairage a été corrigé en sortie.

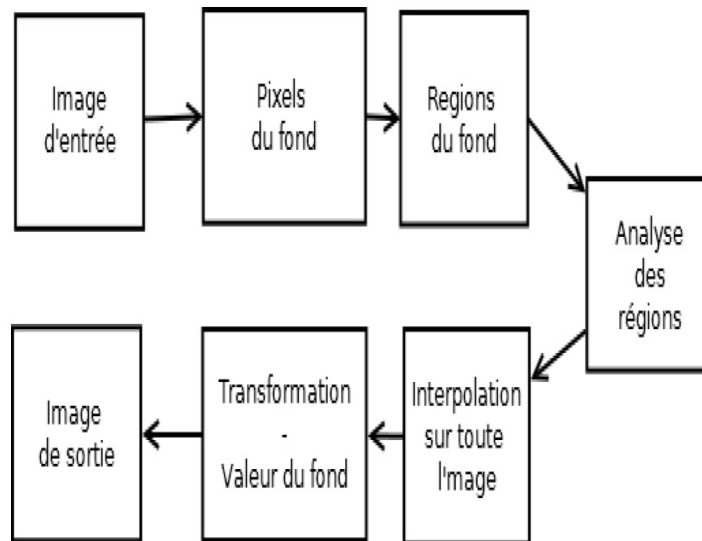


FIGURE 2.15: Étapes de la correction d'éclairage par une carte de fond dans le cas

Nous suggérons une alternative qui se positionne au niveau du pixel et détermine si le pixel courant est un pixel du fond en fonction de ses informations locales. Nous nous servirons de la propagation de la lumière pour déterminer comment caractériser un pixel du fond. Étant donnée que la variation d'éclairage influe en fonction de la distance (cf. loi en carré inverse), il est pertinent de regrouper les pixels plats pour identifier les régions invariantes. Par exemple, deux pixels de même valeur suggèrent qu'il n'y a pas de différence de luminosité entre les deux points. Nous chercherons donc à regrouper ces pixels en fonction de leur position spatiale et de leur valeur en segmentant ces pixels. Les étapes "pixels du fond" et "régions du fond" donnent la segmentation du fond du document.

L'étape "Analyse des régions" consiste à comparer les différentes régions obtenues dans le but d'identifier le ou les différents fonds. L'information du fond étant partiel, nous ferons, dans l'étape "interpolation sur toute l'image", un remplissage du fond pour déterminer la totalité de la carte de correction.

Et nous avons dans l'étape "Transformation - Valeur du fond", les éléments nous permettant d'appliquer notre transformation, nous choisirons d'utiliser l'opérateur multiplicatif

qui linéarisera l'information de luminosité, et nous pourrons ensuite appliquer une correction gamma si besoin est, comme il est fait pour la calibration couleur. Nous proposerons aussi certaines alternatives pour estimer la valeur du fond du document et nous montrons l'interface du logiciel pour l'application de la fonction.

La création de la carte de fond va se baser sur des points caractéristiques de l'image que nous identifions comme appartenant aux fonds. Nous verrons comment les identifier dans une première partie.

Ensuite, nous allons segmenter en région l'ensemble des points ayant des caractéristiques similaires et exclure les erreurs d'identifications.

Puis, nous verrons dans une troisième partie une heuristique qui comparera les différentes régions.

Enfin, nous présenterons, dans une dernière partie, l'application de la méthode.

2.3.1 Les pixels caractéristiques du fond d'un document :

La segmentation du fond du document, que nous proposons, se base sur des pixels caractéristiques du fond que nous appellerons pixels "plats". Au lieu de labéliser tous les pixels de l'image en fond/non fond à partir des méthodes de seuillage adaptative, nous allons labéliser uniquement les pixels du fond en se basant sur deux propriétés que nous expliquerons.

Les pixels "plats"

Dans le cas d'une image n'ayant pas de défaut d'éclairage, un pixel est un pixel "plat" lorsque les pixels autour de lui ont une valeur identique ou quasi-identique à celui-ci.

La distribution statistiques dans de "petites" zones de l'image, nous permet de différencier entre une zone homogène et une zone non-homogène dans un contexte de problème d'éclairage ou non. Effectivement, une zone homogène a une distribution avec une variance très faible, voir nulle, et si nous nous situons sur un objet ou à moitié sur un objet, la variance sera élevée. Nous pouvons faire la différence entre les deux cas.

Soit le carré du voisinage centré au pixel p de coordonnée (i, j) de côté n :

$$C_n(p_{(i,j)}) = I(n - i : n + i, n - j : n + j)$$

Lorsque l'on prend n proche de zéro, la différence entre les pixels de C_n est nulle dans une zone homogène. Dans le cas d'un défaut d'éclairage, l'illumination varie peu sur une petite distance, mais nous pouvons alors établir une tolérance. Cette tolérance est calculée en fonction de la variance, et elle caractérise le défaut d'éclairage en fonction de la taille du carré.

La première propriété, que nous posons pour définir des points caractéristiques, est qu'ils doivent avoir une variance quasi-nulle. La taille du carré n peut être interprétée comme un critère de précision ou un critère d'échelle. Selon cette propriété, nous pouvons déterminer que la valeur du pixel du fond est précis en fonction de la variance des pixels du carré. La relation se formule de la manière suivante :

$$p \in P \Leftrightarrow \text{Var}(C_n(p)) \leq \delta_p \quad (2.30)$$

avec p un pixel de l'image, un pixel plat noté P , $C_n(p)$ le carré centré en p de taille $2n + 1$, et δ_p le critère de validation.

Le défaut d'éclairage est linéaire dans une petite région, ce qui aura pour conséquence qu'une partie sera sur-exposée et l'autre sous-exposée systématiquement, et de plus, l'une compensera l'autre (cf. figure 2.16). Lorsque la fenêtre sera dans une zone homogène ayant des défauts d'illumination, le pixel central sera la moyenne du carré.

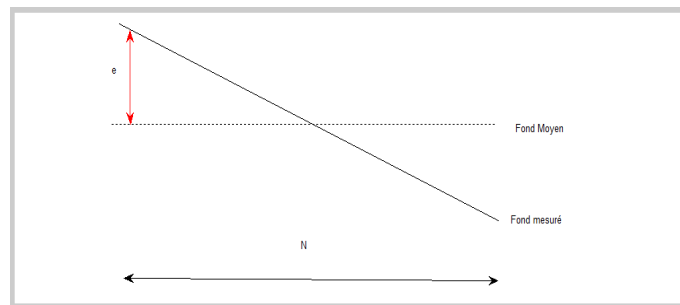


FIGURE 2.16: Schéma d'une coupe 1D d'un fond

Dans le cas d'une image de document ayant un problème d'éclairage, un pixel est un pixel "plat" lorsque la moyenne de la fenêtre carrée centrée sur lui est égale à ce pixel.

A partir de cette information, nous pouvons établir un modèle de calcul pour déterminer les pixels plats d'une image et nous pouvons évaluer δ_p plus simplement.

Pour déterminer l'ensemble des pixels "plats" d'une image, nous suggérons la seconde propriété qui consiste à calculer la différence entre l'image originale, et son image moyennée par un filtre moyenneur de taille $2n + 1$.

$$p \in P \Leftrightarrow |I(p) - I_M(p)| \leq \delta_p \quad (2.31)$$

avec p la position spatiale de coordonnées (i, j) , I l'image originale, I_M l'image moyennée par un filtre de taille $2n + 1$ avec n petit et δ_p le critère de validation proche de 0. En pratique, nous avons utilisé pour un noyau de taille 7×7 et une erreur de 0.25.

Le calcul d'un pixel plat doit satisfaire en premier lieu la seconde propriété (voir eq. 2.31) pour une question de performance, puis dans un second temps, la première propriété (voir eq. 2.30).

Afin d'éviter des cas particuliers où la moyenne du carré est égale au pixel central, nous utiliserons le calcul d'écart-type pour exclure les fonctions par paliers qui permet de seuiller la différence entre le fond théorique et le fond mesuré. Ce second critère permet aussi de seuiller l'angle d'inclinaison que nous tolérons.

Observons dans la figure 2.17, deux cas d'images où nous mettons en évidence l'utilisation des pixels plats.

Nous avons numérisé deux images dans des conditions d'éclairage similaire : l'une blanche (d), l'autre du texte sur fond blanc (a). Dans les images (b) et (e), nous observons les pixels plats sélectionnés. Nous observons une certaine similitude entre les deux images. Les images (c) et (d) présente le résultat d'un plan déterminé par les moindres carrés en se basant sur les points plats obtenus.

La différence en valeur absolue entre le résultat de l'image f) et l'image original d) est faible (inférieur à 5 dans l'ensemble). L'erreur entre le fond de l'image a) et l'image c) est relativement faible, mais il y a toutefois des imprécisions aux alentours de 15. Ces erreurs s'expliquent par des points plats au niveau de la zone en haut à gauche où il y a un fond noir, ainsi que certains pixels du textes qui ont été sélectionnés. Les erreurs sont faibles car le fond est majoritaire.

Les pixels d'une lettre peuvent être considérés comme un pixel plat parce que la taille du filtre est inférieure à la taille de la lettre. effectivement, la taille optimal du filtre varie en fonction du DPI de l'image qui n'est pas toujours connu. Toutefois, il existe d'autres cas qui peuvent générer des erreurs de labélisation même en maîtrisant ce paramètre :

- des effets visuels, qui génèrent un effet d'ombre artificiel ;

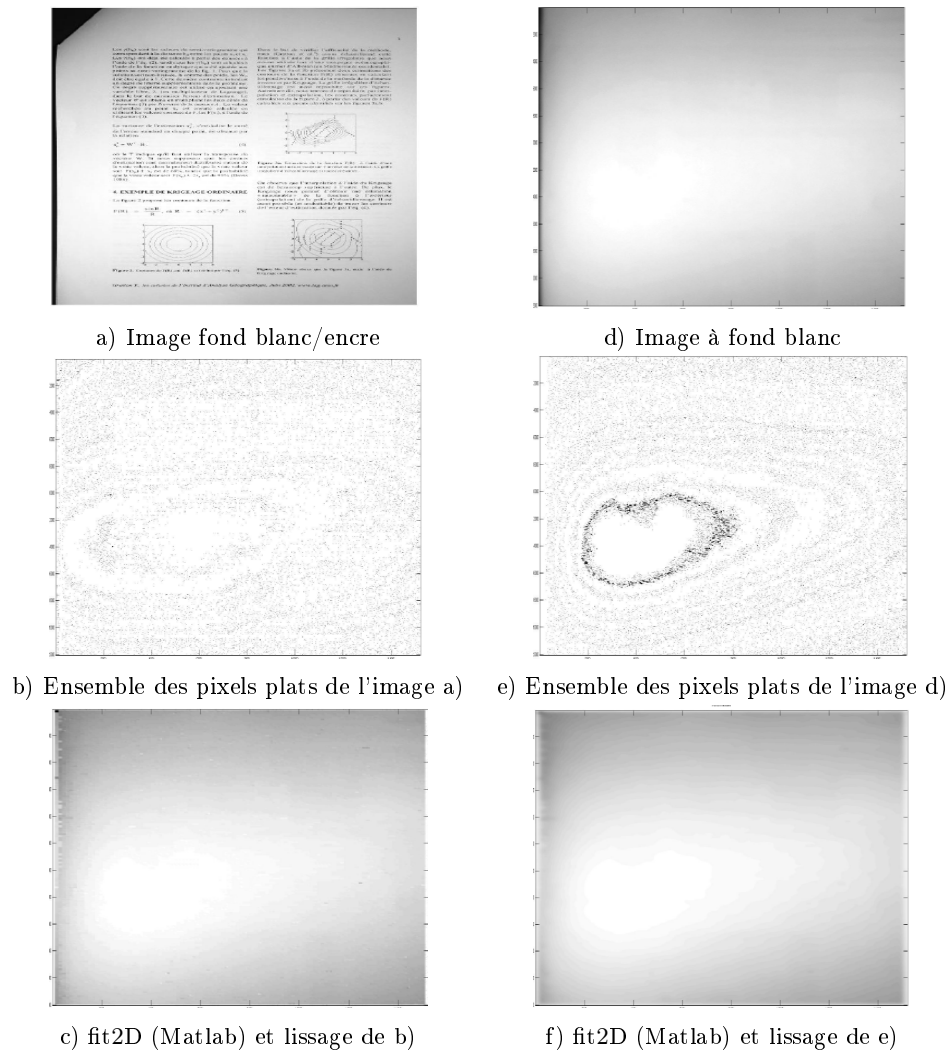


FIGURE 2.17: Exemple de pixels plats

- des fonds colorés pour mettre en valeur le texte, ce qui nous amène à avoir plusieurs fonds ;
- des surlignages pour les mêmes raisons ;
- des images.

Lorsque nous appliquons une fonction du type `fit2D`, nous considérons que tous les pixels sont sur le même plan, or cette hypothèse n'est pas toujours vraie. Par exemple, dans la figure précédente, nous remarquons que la pliure du document en haut à gauche fait apparaître le support du scanner qui est noir, cela va avoir pour conséquence de modifier le calcul du plan dans cette zone. Nous généralisons ce problème sur des numérisations de livre récents possédant différents fonds, pour répondre à ce problème il faut pouvoir segmenter les différents fonds. De plus, cette segmentation servira à exclure les erreurs. Nous notons que la sélection de ces points représente un sous-échantillonnage particulier de l'image. Nous obtenons donc une image représentée par une matrice creuse composée de pixels plats, où nous connaissons la position et la valeur en luminance de chacun.

2.3.2 Segmentation

Notre objectif est de réunir les pixels "plats" ayant des caractéristiques communes pour estimer les variations de luminance.

C'est une segmentation de type "split and merge". Il produit automatiquement une partition initiale en régions petites (split), qui vont ensuite croître en se regroupant (merge). La structure de données a été nommée par R. Finkel et J.L. Bentley en 1974 (cf. [Finkel and Bentley, 1974]).

La partition initiale (split) est réalisée en divisant récursivement l'image en régions lorsqu'un certain critère d'homogénéité n'est pas satisfait.

Lors de cette phase, le graphe d'adjacence, ou region adjacency graph (RAG), est créé : à chaque région est associée un sommet de graphe, et des arêtes relient les sommets correspondants à deux régions qui se touchent (par ex : R est divisée si $\sigma_R > \text{seuil}$). La phase de regroupement (merge) utilise le RAG pour modifier la partition initiale : pour chaque sommet R du RAG, on cherche s'il existe un sommet R' voisin dans le RAG et de valeur suffisamment proche, et si c'est le cas, on les fusionne (par ex : R et R' sont fusionnées si $|\mu_R - \mu_{R'}| < \text{seuil}$).

2.3.2.1 Split

Nous rappelons qu'à cette étape de la méthode, nous avons un ensemble de points, nous avons leurs positions et leurs valeurs en luminance, ils sont donc définis dans l'ensemble $[0 \ 255]$.

La spécificité de notre split se base sur le calcul de la variance de la région puisque nous recherchons à rassembler les ensembles de points du fond formant un bloc possédant des informations en luminance similaire.

Donc, R est divisé si et seulement si

$$\sigma_R > \delta_Q \tag{2.32}$$

En pratique, nous avons choisi $\delta_Q = 5$. De plus nous avons seuillé les quadrants de taille 7×7 car il correspond à la taille du filtre des points plats et le bloc est trop petit.

Une fois que tous les quadrants sont caractérisés, l'information du fond est divisée en fonction du quadtree. Afin de reconstruire l'ensemble des fonds du document, nous allons, dans une prochaine étape, expliquer l'étape de merge.

2.3.2.2 Merge

L'information du fond du document est divisée en plusieurs sous-régions, l'objectif est maintenant de recomposer l'ensemble du fond à partir de ces informations.

Pour fusionner les différentes régions, nous nous positionnerons dans une problématique de graphe des régions adjacentes. Ce graphe est construit à partir des régions obtenues. Une région est déterminée par un noeud. Nous conservons la moyenne en luminance du bloc ainsi que de la position centrale de celui-ci. Le principe de la fusion consiste à réunir les différents noeuds adjacents en fonction d'un critère de similarité.

Nous considérons que le fond d'une image est un ensemble de pixels continus. De la même manière, nous allons déterminer notre critère de similarité en fonction de la définition de continuité des blocs. Nous avons en pratique, $|(x - a)| = d$ avec x et a les centres des quadrees, et $|f(x) - f(a)| < \epsilon$ la différence de luminance des moyennes des deux blocs, et le paramètre ϵ . La distance est un facteur qui influe sur la continuité des quadrants. Effectivement, si les zones du quadtree sont grandes, alors la moyenne est moins pertinente. Il faut donc définir ϵ en fonction de d .

Le filtre pour déterminer un pixel plat est de taille 7x7, nous pensons qu'un pixel peut représenter la moyenne d'un bloc de pixel de taille 7x7. Si nous avons deux blocs adjacents, nous considérons qu'ils sont continus si ils ont une différence inférieure ou égale à 1. La distance maximale des pixels appartenant à ces deux blocs est d'environ 15 pixels. Par conséquent, nous avons 1 de différence pour 15 pixels.

Nous obtenons alors $\epsilon_{final} = \epsilon + \frac{d}{15}$, cela signifie que plus d est grand, plus on élargit notre critère de décision, ϵ est le critère de continuité et 15 est le critère de dilatation.

Soit Q_i et Q_j deux quadrants du quadree 4-connexes.

$$Q_i \in Z_k, Q_j \in Z_k, \text{ssi} |\mu_{Q_i} - \mu_{Q_j}| < \epsilon + \frac{d}{15} \quad (2.33)$$

, avec μ_Q la moyenne du quadrant.

Nous déterminons l'ensemble des fonds d'une image à partir d'un regroupement de région. Notre croissance de région opère de manière itérative. Lorsque deux régions fusionnent, nous ne fusionnons pas les informations, le critère est basé sur la continuité d'un bloc du quadtree par rapport à un autre que nous avons expliqué précédemment. Pour déterminer l'ensemble des fonds, nous parcourons tous les quadrants indéterminés auxquels nous appliquons notre accroissement de région. L'ensemble des fonds de l'image sont déterminés lorsque tous les blocs ont été itérés.

Nous notons que notre critère de distance est linéaire alors que la loi physique nous impose la loi en carré inverse. Ce point s'explique car nous nous focalisons uniquement sur

les courtes distances des quadrants. Dans le cas des grands quadrants, cela signifie qu'il n'y a pas de problème d'éclairage, mais que l'origine du défaut est dû à la forme du livre.

Selon la structure du fond du document, nous constatons que la méthode peut diverger. Par exemple, dans les cas de documents récents, nous observons des transitions invisibles entre les images et le fond du document. Nous constatons aussi que le défaut d'éclairage peut fusionner les différents fonds du document. Nous n'avons pas trouvé de solution à ce problème.

Nous présentons le résultat d'une segmentation dans les figures 2.18 et 2.19 pour le cas d'un journal. Nous remarquons que pour ce cas, la région 5 se situe dans une région saturée (que nous observons aussi sur la fréquence de points plats dans la région) et le critère ϵ influe beaucoup sur sa fusion avec la région 1.

Nous proposons d'autres exemples en Annexe 6.7, 6.9 et 6.10.

2.3.3 Analyse du fond

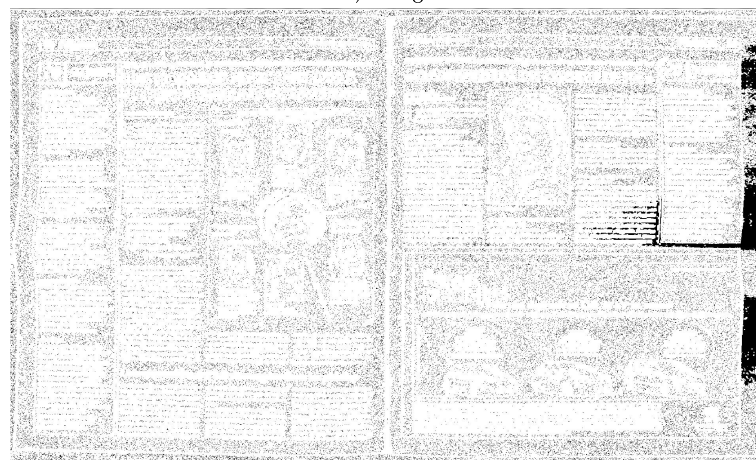
Notre analyse a pour but de déterminer l'ensemble du fond à partir des régions obtenues suites à l'étape du merge. Notre problématique se situe dans le cadre de livre ancien, mais nous souhaitons aussi que cela puisse être applicable dans les cas de livre récent. Nous rappelons alors les caractérisations pour chacun des cas.

Dans le cadre du **livre ancien**, le fond est unique, a priori blanc et il est présent dans toute la page. Ces trois propriétés sont assez évidente, toutefois, nous gardons à l'esprit que l'état du livre n'est jamais parfait, ce qui a pour conséquence d'avoir des dégradations telles que des pliures, des tâches, de la transparence, ou encore des pages jaunies. Dans l'ensemble, la méthode ne détecte pas les défauts de hautes fréquences mais elle prend en compte l'information colorimétrique et certaines tâches du fond. Nous remarquons aussi que la reliure est très marquée (au point d'avoir systématiquement une séparation nette entre les deux pages) et qu'un livre ancien possède alors deux pages et donc deux fonds identiques. Dans le cas des pages uniques, la détection ne pose pas de problème.

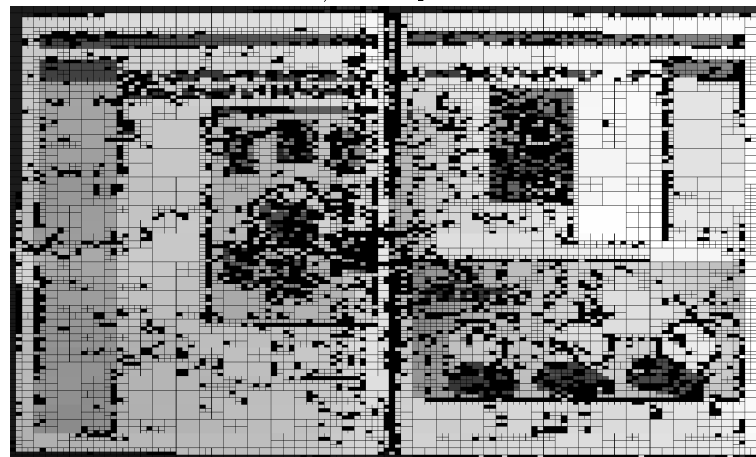
Dans le cadre des **documents nouveaux**, il y a plusieurs fonds, généralement, nous avons un fond blanc et des encadrements colorés. Les transitions ne sont pas systématiquement marquées au point d'utiliser des images pour découper deux zones. Nous partons du principe qu'il y a un fond "principal". Ce fond est le papier utilisé. Celui-ci est généralement blanc mais pour palier à différents cas complexes nous considérons qu'il est le plus clair et que les encadrés sont plus foncés et plus petits. Ce fond sera alors notre



a) image



b) Points plats



c) quadtree représenté par la moyenne du bloc

FIGURE 2.18: Résultat du split

référentiel. Ce fond est présent dans les extrémités de la page, même si nous observons parfois la présence d'image ou des encadrés qui découpent la page en deux. En ce qui concerne la forme de la reliure, elle est importante et elle découpe d'une manière significative les deux pages à cause de reflets importants. Dans le cas de livre récent, le fond

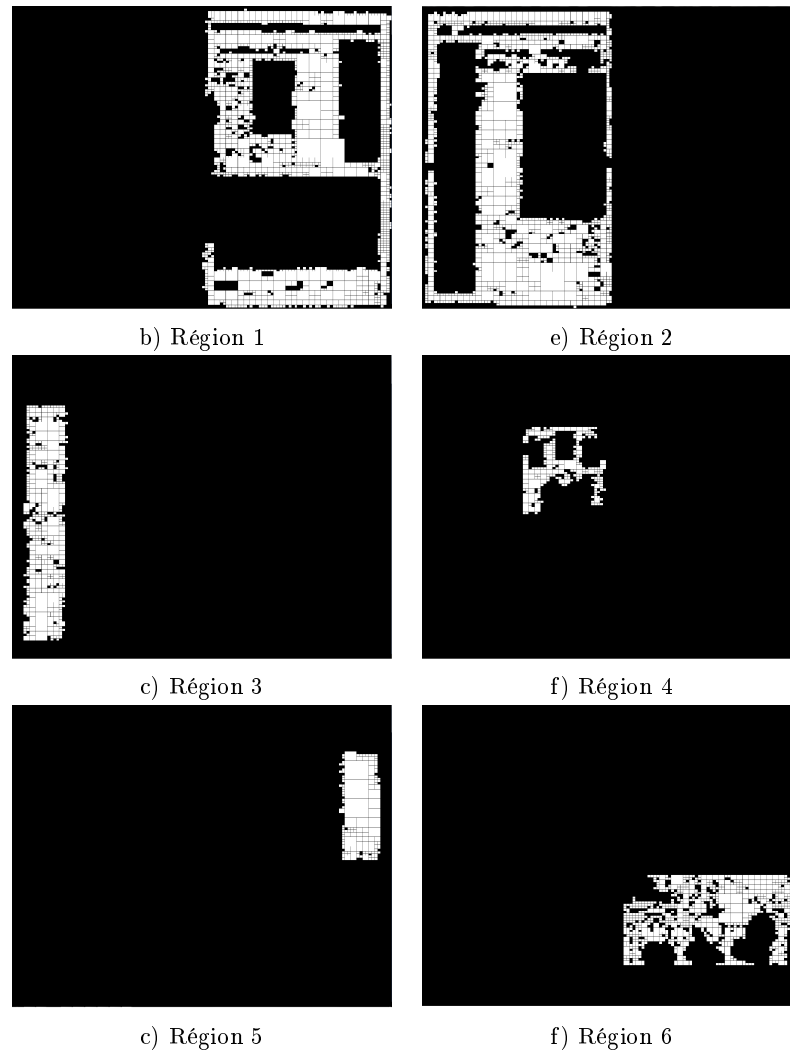


FIGURE 2.19: Les différentes régions continues

des deux pages n'est pas systématiquement le même. Dans le cas d'une page unique, ce dernier présente un problème de saturation dans une zone de l'image.

Le résultat obtenu dans l'étape précédente ne donne pas systématiquement une seule région du fond du document. Dans les deux cas de figure, et particulièrement dans le cas du livre récent, il est pertinent d'analyser chacune des régions obtenues pour en déduire l'ensemble des régions du fond du document qui ont été découpé par les éléments du document.

Nous allons donner les caractéristiques et les détailler pour chacun des points :

La couleur du fond : Pour chacune des régions obtenues après segmentation, nous observons la moyenne et la valeur maximale de la zone. Nous déterminerons à partir de l'image d'origine si la région est composée de pixels de niveaux de gris ou de couleur. Nous utilisons ces informations car le fond est la région la plus grande, et dans ce cas

là, la région est susceptible d'être la plus dégradée, ce qui fait que la différence entre la moyenne et le maximum est une information pertinente. Ce critère permet aussi de seuiller les régions ayant un maximum trop bas car nous jugeons que les informations d'illuminations ne sont pas pertinentes.

Concernant l'analyse colorimétrique de la région, nous considérons que le fond d'une image en niveaux de gris est prioritaire sur un fond d'image colorée.

La répartition spatiale du fond : Le fond est la région la plus grande et/ou qui est présente sur les extrémités du document. Pour valider la taille de la région, nous faisons le calcul d'aire de la boîte englobante de toutes les régions. Nous seuillons les régions dont les aires sont inférieures à une certaine valeur (proportionnelle à la taille de l'image). Et pour valider la répartition de la région, nous calculons la distance entre les quatre coins de l'image et le point le plus proche de la région courante.

A partir de ces différentes analyses, nous sommes capables d'effectuer un tri de régions.

Afin de déterminer si nous sommes sur une page simple ou sur une double page, nous utilisons un algorithme de détection de reliure basé sur les gradients au milieu de l'image. Dans le cas de la détection d'une reliure, nous analyserons la page de gauche et de droite afin de déterminer si elles ont le même fond. Pour cela, nous prendrons les régions centrales de chacune des pages et calculons les histogrammes des trois canaux pour établir un triplet de référence. L'écart des triplets conclura si le fond est commun aux deux pages.

Toutefois, notre méthode ne diffère pas dans les deux cas.

Le fond est défini par au moins deux régions : Les scanners sont majoritairement utilisés pour numériser des livres. Par conséquent, nous nous basons sur le fait que la reliure du livre découpe en deux régions distinctes le fond. Nous traiterons de manière itérative toutes les régions en fonction du classement obtenu. A partir de tous les points évoqués ci-dessus, nous considérons que notre première région est la candidate pour être notre fond de référence. La reconstruction du fond est considéré comme complet lorsque les contraintes spatiales sont validés par le fond courant ou si il n'y a pas d'autre région à tester.

Dans le cas contraire, nous comparons sur deux points la fusion de la région du fond. Le premier point est spatial, il faut que les régions soient complémentaires, ce qui, en pratique, consiste à comparer les boîtes englobantes de la région du fond courant et de la région candidate. Pour valider ce critère, il faut que la région courante soit à l'extérieure de la boîte englobante du fond.

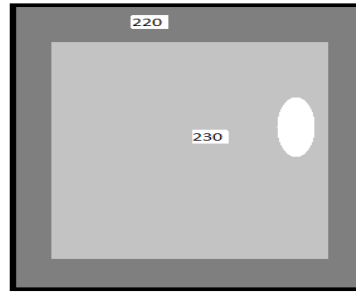


FIGURE 2.20: Cas où la région centrale est supérieure en luminance car elle est sur-exposée par la lumière

Le second point est de comparer les triplets colorimétriques de la région. Effectivement, il est nécessaire qu'il existe un coefficient k , tel que le triplet (r_f, v_f, b_f) du fond courant, et le triplet (r_c, v_c, b_c) de la région courante, puisse avoir la relation suivante : $(r_f, v_f, b_f) = k(r_c, v_c, b_c)$.

Si notre région courante valide ces deux critères alors nous fusionnons la région courante et le fond. Et nous mettons à jour les informations du nouveau fond, ainsi que notre liste triée de régions.

Une fois que nous avons le fond, nous analysons à nouveau les régions pour récupérer les fonds secondaires de type encadrés, etc. S'il existe des régions non-identifiées, nous faisons à nouveau notre démarche itérative à partir de la région classée la plus haute. Effectivement, nous considérons qu'un encadré peut être découpé en deux par une image. Lorsque ce nouveau fond est établi, nous remettons à l'échelle en fonction du fond de référence afin de récupérer les informations d'illumination. Pour cela, nous nous baserons de nouveau sur la valeur maximale que nous comparerons au point le plus proche du fond principal.

Les cas des livres anciens ne sont généralement pas aussi complexe, nous détectons une seule grande région pour une page simple et deux grandes régions complémentaires pour une double page. Le cas des livres récents propose des cas plus complexe que nous illustrons dans les figures 2.20, 2.21 et 2.22 sur lesquelles nous nous sommes appuyés pour établir nos critères.

La figure 2.20 représente le cas d'une image sur un fond blanc dans un contexte d'éclairage ambiant. Admettons que notre méthode de segmentation nous suggère 2 zones de niveaux de gris, nous observons que la zone du milieu est plus clair que la zone qui l'encapsule, or ceci n'est généralement pas le cas. La zone dont la luminance est la plus élevée se situe bien souvent à l'endroit de la zone de sur-exposition. Par conséquent, nous jugeons plus pertinent de prendre la zone étant le plus proche des extérieurs de la page.

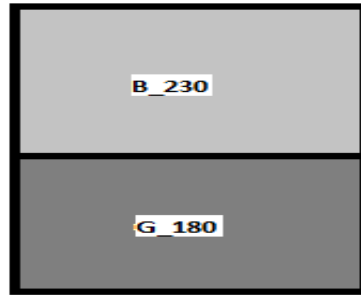


FIGURE 2.21: Cas où la région supérieure est bleu avec une luminance de 230 et la région inférieure est grise avec une luminance de 180

Le second cas met en évidence que la combinaison de différents points est nécessaire pour évaluer le fond. Dans la figure 2.21, nous observons une page coupée en deux, la partie supérieure est colorée (Bleu) et sa luminance est égale à 230 ; la partie inférieure est en niveau de gris et sa moyenne est de 180. Ce cas nous a apparus lors de la numérisation d'un journal dans un contexte d'éclairage ambiant assez faible, nous remarquons que la différence en luminance est élevée. Effectivement, il s'avère que la réflectance de l'encre utilisée est beaucoup plus forte par rapport à la réflectance du papier. Nous avons alors un cas où il n'est pas possible de conclure quelle partie détermine le fond en fonction des informations spatiales. Nous privilégierons donc les régions en niveaux de gris pour déterminer le fond malgré. Toutefois, nous traiterons la partie supérieure après, mais nous le réajusterons en fonction de la partie inférieure.

Dans le dernier exemple, nous illustrons un cas plus classique des documents récents ; nous avons quatre zones : la 1^{er} zone a une moyenne de 230, la seconde de 210, la troisième de 190 et la dernière de 50. Les trois premières zones ont une moyenne relativement proche mais la seconde zone est beaucoup plus grande, celle-ci sera alors considérée comme le fond principal. Il ne sera plus nécessaire de rechercher un nouveau fond car ce fond recouvre l'ensemble de la page.

La deuxième partie de la méthode va chercher à combler les trous éventuels par les autres zones. Si celle-ci sont suffisamment grandes et supérieure à un seuil alors nous les prendrons en compte. En l'occurrence, la zone à 50 sera ignorée car sa moyenne est trop faible. Les cas de la première et la troisième régions seront pris en compte. A priori, notre fond de référence est censé être blanc, ce qui signifie que c'est le fond le plus clair, il est alors impossible qu'un autre fond puisse être plus clair, le fond de 230 est alors un problème. La logique voudrait que ce fond soit diminué de 10 pour que son plan soit remis au fond de référence. Mais ceci n'a pas vraiment de sens. La solution que nous proposons est alors basée sur la valeur des pixels les plus proches de la zone. Nous constatons que le fond principal n'est pas homogène est qu'il faut respecter au mieux les différences de luminosité. Le point du fond principal le plus proche de la zone et celui

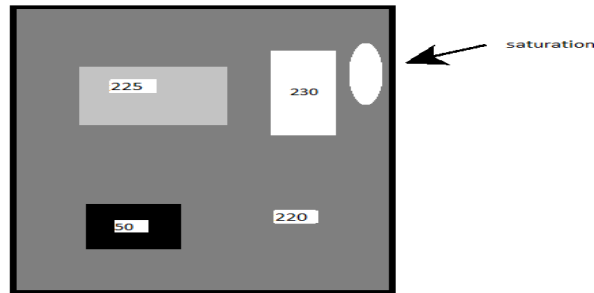


FIGURE 2.22: Exemple de problème de sélection dans un cas de saturation de luminosité

de la région courante permettra de déterminer la différence entre les deux régions. Ce résultat permettra de remettre à niveau les deux plans.

La suite de la méthode fera de même pour les cas valides suivants.

A la fin de cette étape d'analyse, nous avons pu déterminer le fond du livre et nous avons remis à niveau d'éventuels fonds secondaires. Toutefois, ces zones sont généralement incomplètes à cause d'images ou de zones trop petite, et il est nécessaire de déterminer ces zones par des méthodes de reconstruction d'image.

2.3.4 Reconstruction du fond

La dernière étape de notre traitement consiste à trouver les informations manquantes du fond du livre afin d'obtenir notre carte du fond, puis d'y appliquer notre transformation.

2.3.4.1 Reliure du livre

Nous constatons de manière générale que la détermination du fond pose essentiellement des soucis au niveau de la reliure. La forme de la reliure est un vrai problème car les défauts d'illumination de cette zone sont très importantes. Par conséquent, nous ne pouvons pas la segmenter. A partir du moment où nous avons connaissance du fond du document, nous jugeons qu'une analyse spécifique de la reliure s'opère en suivant. Nous notons que dans le cas de livre ancien, c'est véritablement une problématique car les pages se distinguent nettement. Nous pensons alors qu'il est nécessaire de corriger spécifiquement cette région du livre qui est très complexe.

Dans le cas des livres anciens, nous avons des reliures usées, voir des pages qui se détachent ou encore une zone où la page jaunie de plus en plus. Pour les livres récents, en

plus d'être une zone sensible aux variations d'illumination, il y a des images ou encore du texte coupés a cause de la forme du livre.

Nous proposons un échantillon de reliures dans les figures 2.23 et 2.24.

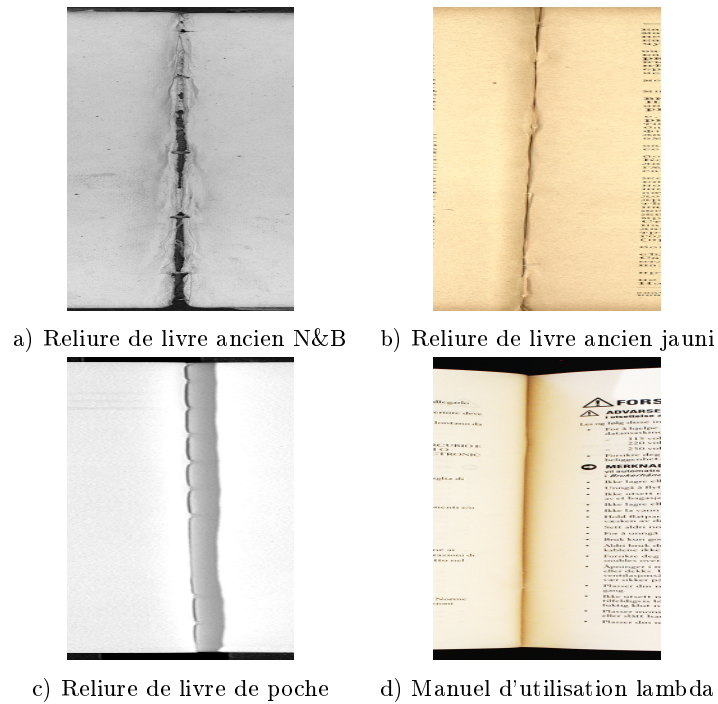


FIGURE 2.23: Reliures de livre ancien

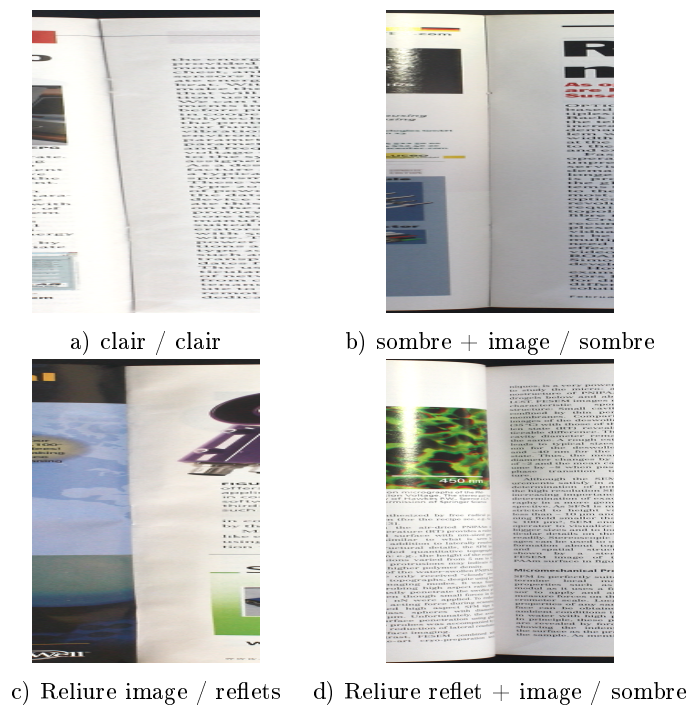


FIGURE 2.24: Reliures de livre nouveau

La première démarque les reliures plutôt anciennes ou de papiers de moyenne qualité, et dans le second, des reliures de magazines, qui ont plutôt tendance à générer des reflets

et à avoir des images.

Nous notons que les livres sont positionnés de telle sorte que la reliure soit verticale et au milieu du livre. Nous déterminons avec plus ou moins de précision la reliure. Nous observons sur les figures que les reliures sont différentes les unes des autres.

Notre analyse de la reliure fait des profils sur les lignes comprises dans un intervalle $I = [M - D; M + D]$ avec M le milieu du document, et D la distance d'intervalle. Cette distance D est a priori la distance de la marge du livre. Pour l'estimer nous avons fait une analogie entre la construction classique de représentation typographique d'une page de format A4 et le nombre de pixel. Nous pouvons alors observer les représentations 1–D équivalentes dans les figures 2.25 et 2.26.

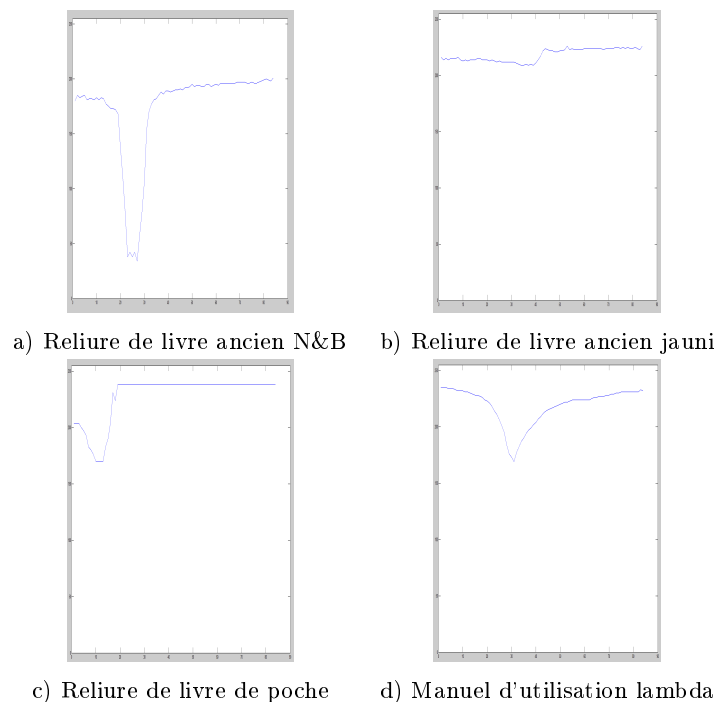


FIGURE 2.25: Courbe de reliure de livre ancien

Nous constatons que la reliure est souvent en forme de "V" ou un changement de niveau est important. Nous remarquons que les images et le texte sont parfois présents, nous traiterons les informations des coupes différemment des deux côtés de la reliure pour n'obtenir que les valeurs du fond. Cette information pourra être comparée à l'approximation du fond que nous avons fait. Nous notons que les deux informations se complètent et nous permet, par exemple dans le cas de la figure 2.26.c, de ne pas déterminer la reliure du côté de l'image. Nous notons pour finir que cette analyse n'est parfois pas suffisante car la quantité de lumière varie en fonction du haut et du bas du document.

L'analyse de cette reliure permet d'avoir une information de luminosité dans une zone

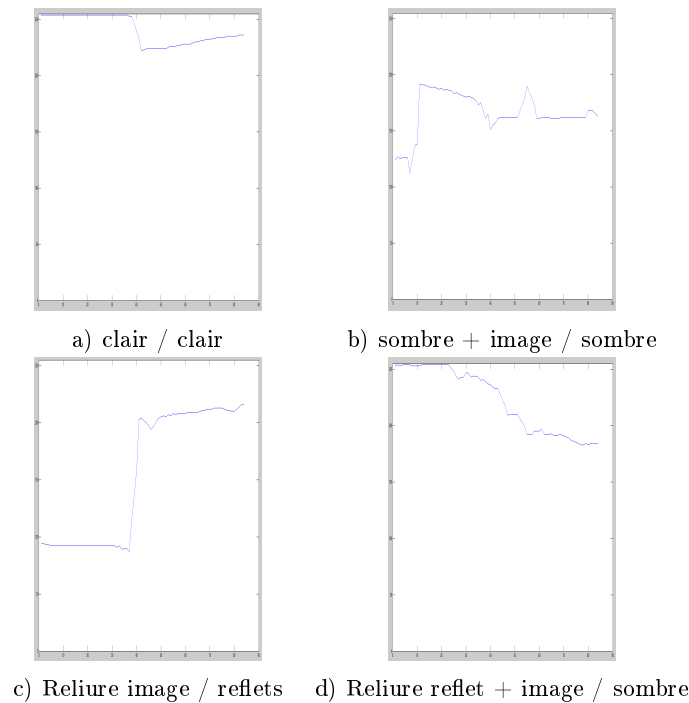


FIGURE 2.26: Courbe de reliure de livre nouveau

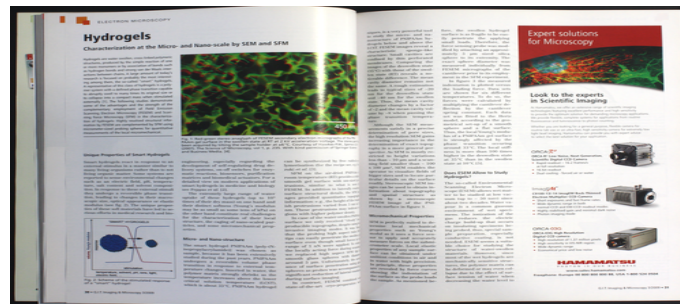
fortement discontinue, mais il reste d'autres zones inconnues dont nous n'avons pas d'information.

2.3.4.2 Reconstruction du fond du document

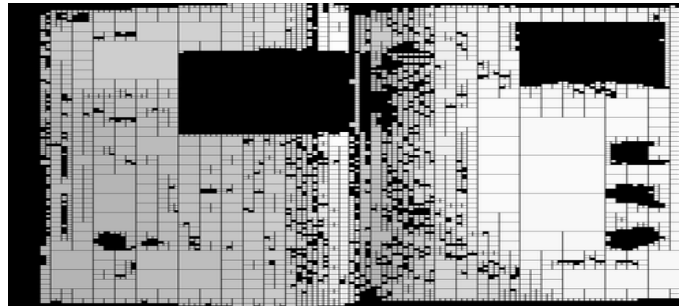
A partir des régions du fond, nous reprenons les valeurs de luminance obtenues lors du quadtree pour reconstruire le fond partiel. Nous rajoutons les informations de la reliure lorsque le fond est défini. Pour illustrer, observons la figure 2.27.

Afin de remplir les trous du fond, nous nous servons des méthodes de détection des blobs de petite taille. Nous remplissons ces zones en utilisant la moyenne à proximité du blob courant.

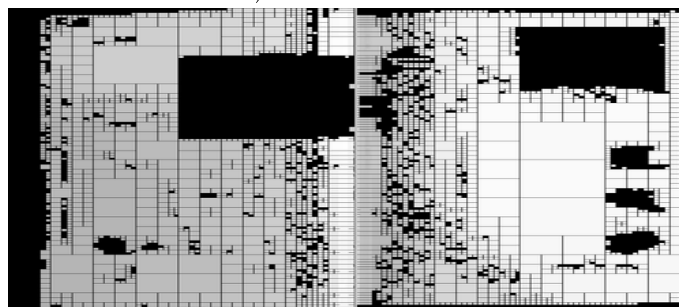
Pour les régions plus grandes, nous utilisons une méthode d'inpainting basique, qui privilégiera les informations verticales. Ce point se justifie simplement dans l'utilisation des appareils de numérisation. Effectivement, les appareils sont souvent posés contre un mur avec un individu face à la machine, par conséquent les sources de lumières sont souvent sur les côtés. L'éclairage sur le côté met davantage en valeur la forme du livre. Nous finirons par un lissage classique pour gommer les aberrations. Nous observons ces deux points dans la figure 2.28.



a) Numérisation originale



b) fond de référence



c) ajout au niveau de la reliure

FIGURE 2.27: Reconstruction partiel du fond

Nous notons qu'il arrive parfois que de grandes zones sont parfois incomplètes, pour conserver la dynamique de la transformation nous y appliquerons une constante équivalente à la moyenne du fond principal.

Notre résultat est alors la carte de fond. A ce stade, notre méthode a donc reconstruit le fond du document. Pour mettre en application la correction d'éclairage, nous avons besoin de déterminer qu'elle est la valeur de référence et qu'elle est l'opérateur que nous utiliserons.

2.3.5 Application

Une fois la carte de fond générée, le calcul des coefficients de correction peut être calculé. Nous avons vu avec les différents opérateurs dans la section 1.2.2 que nous pouvons utiliser l'opérateur additif du LIP, la puissance gamma et la multiplication. Mais, pour

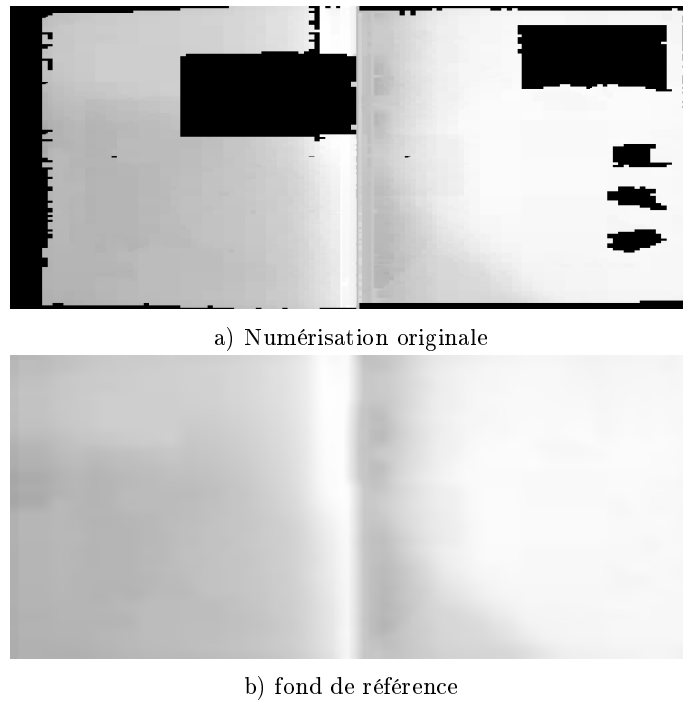


FIGURE 2.28: Reconstruction partielle du fond

une question de performance, nous utilisons la multiplication classique, d'autant plus que nous avons supposé à plusieurs reprises la linéarité du défaut d'illumination.

Le temps d'exécution est variable en fonction de la complexité de l'image, c'est-à-dire la facilité à découper l'image en quadtree. Généralement, toutes les images de livre ancien sont considérées comme "simple" ainsi qu'une partie des images de document récent. Pour un ordinateur performant (i7, gtx2), notre algorithme mono-tread s'exécute en moins de 1 seconde, alors qu'une image complexe se déroule aux alentours de 2 secondes pour des images de 9 M de pixels. Les résultats de performance sont intéressants car l'amélioration de la méthode se fera sur le quadtree qui gagnera beaucoup en temps d'exécution en passant à du multi-threads.

Afin de calculer la carte des gains, il faut déterminer un fond de référence. Ce dernier ne peut être déterminé dans le cadre d'un problème d'éclairage au vu de la variation du fond. Pour cela, nous proposons à l'utilisateur de calculer une valeur de référence en fonction de la carte ou qu'il en détermine une pour que l'ensemble des numérisations soient homogènes.

Notre calcul du fond de référence a différentes approches. Le premier se base sur la valeur maximum du fond. En pratique, nous nous apercevons que les documents sont saturés à un faible pourcentage (admettons 5% des pixels totaux de l'image). Par conséquent,

nous calculons la valeur du fond en enlevant les x pixels les plus clairs. Ainsi, la valeur du maximum va être ajustée comme étant la valeur que l'on souhaite du fond du document.

L'autre approche consiste à calculer la valeur médian du fond.

La premier valeur va avoir pour conséquence d'augmenter tous les pixels alors que le second va trouver un équilibre. Toutefois, ce dernier peut donner un résultat un peu trop sombre. De ce fait, nous avons proposer la moyenne des deux valeurs.



a) Numérisation originale



b) Correction à partir du fond blanc



b) Correction à partir du fond blanc et de l'encadré

FIGURE 2.29: Utilisation de la correction d'éclairage pour enlever une ombre humaine

L'ensemble de ces informations a été mis dans une interface que nous proposons dans la figure 6.11 dans l'annexe 6.4 concernant les logiciels développés durant la thèse.



a) Numérisation originale



b) Correction

FIGURE 2.30: Utilisation de la correction d'éclairage

2.3.6 Conclusion

La correction d'éclairage développée crée elle-même sa propre carte de correction et elle applique une transformation. Cette fonction peut-être paramétrée via une interface.

Pour permettre la segmentation du fond de document ancien et récent, Notre méthode se fait en quatre étapes.

Tout d'abord, nous avons cherché à labeliser des pixels caractéristiques du fond. Ces derniers sont détecté à partir des informations locales : la moyenne du carrée centrée au pixel courant doit avoir la même valeur que ce pixel ; et statistiques : le carrée doit avoir un écart-type inférieure à une valeur seuil.

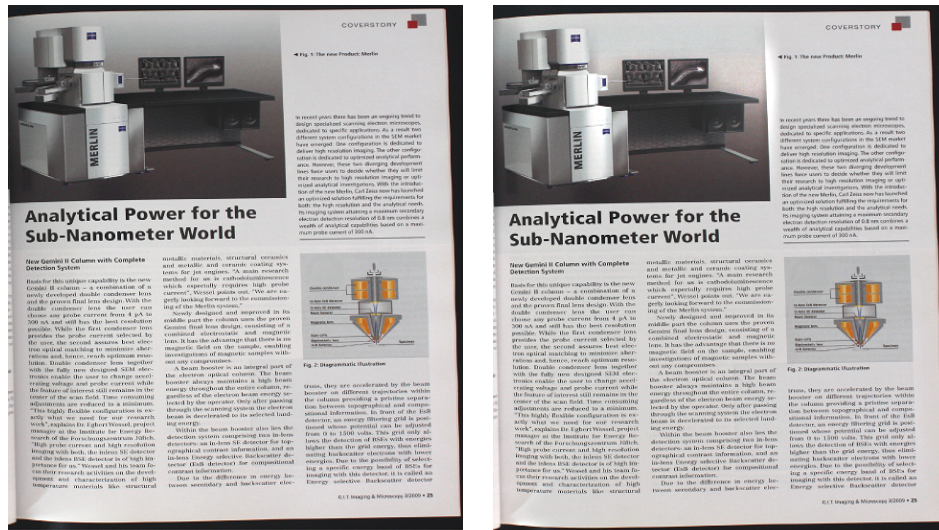
Afin de résoudre des problèmes de labélisation et de distinguer des fonds de couleurs différentes, nous avons choisis de faire un split et un merge. L'écart-type des pixels validés est notre critère du split et le merge est basé sur un critère de continuité.

La troisième étape consiste à identifier chacune des régions obtenues afin de reconstruire le fond partiellement.

La reconstruction du fond de l'image est ainsi la dernière étape de notre segmentation. Nous obtenons ainsi une carte du fond à partir de l'image courante.

Afin d'appliquer la correction d'éclairage, nous avons utilisé l'opérateur multiplicatif afin de répondre au critère de temps d'exécution de l'entreprise et une méthode d'estimation de la couleur du fond. Ce premier critère de performance reste correcte dans la mesure où la méthode n'est pas multi-threadée.

La méthode montre des difficultés pour les images qui ont un dégradé particulier. Nous pouvons le voir dans l'image 2.31. Nous avons proposé une alternative en découpant le graphe obtenu par le merge, mais les résultats ne sont pas convaincants.



a) Numérisation originale

b) Correction

FIGURE 2.31: Nous observons dans la correction une aberration de la correction d'éclairage au niveau de l'image et son dégradé artificiel (voir annexe 6.1.5)

Dans l'ensemble, nos résultats sont convaincants et nous proposons un échantillon en annexe (cf. 6.1.1) ainsi que des corrections que nous avons utilisé comme référence de l'état de l'art .

2.4 Conclusion

Nous retenons de ce chapitre qu'il y a un moyen d'identifier les problèmes d'illumination en se basant sur la décomposition luminance-réfectance de l'image. Il existe des transformations (filtre homomorphique, Retinex) ou encore des méthodes de labelisations des pixels (intrinsic images). Toutefois, les images de documents ont des caractéristiques plus simples que les images naturelles. De ce fait, nous nous sommes concentrés sur la labélisation des pixels du fond (binarisation - fond/non-fond) et des opérateurs mathématiques liés aux problèmes d'éclairage (gamma, LIP, multiplication).

La méthode que nous proposons est spécifique aux documents. Elle se base sur une analyse plus complexe des pixels de l'image, cette analyse aboutit à la sélection de pixels du fond. Toutefois, notre critère engendre des erreurs et il nous faut utiliser un quadtree pour identifier les régions du fond. Afin de reconstruire la totalité du fond, ces régions sont analysées pour identifier les bonnes régions que nous prendrons en référence. Nous procédons pour finir à des méthodes de remplissage d'image.

Dans la mesure où il n'est pas possible de déterminer avec précision la valeur vraie du fond, nous avons proposé une interface pour l'utilisateur lui permettant de piloter ou de trouver une valeur du fond en fonction de certains paramètres. L'interface propose aussi un moyen de modifier la précision, et donc le temps d'exécution de la correction.

Nous observons que la méthode a quelques failles. Ces failles sont généralement dues à un merge trop abusif. Donc, cela peut remettre en question le critère du merge ou d'identifier une approche complémentaire permettant d'éviter ce type d'erreur.

Chapitre 3

Flou de mise au point

3.1 Introduction

Le flou de mise au point d'un scanner est un défaut qui est engendré par un mauvais réglage de l'appareil optique. Ce défaut dégrade de manière significative la numérisation et dans un soucis de qualité de numérisation, il est impératif de détecter ce type d'anomalie. De plus, les acquisitions aux moyens des scanners doivent répondre aux contraintes industrielles dont celui de la netteté.

Nous proposerons, dans ce chapitre, une méthode d'évaluation du flou de numérisation. L'objectif de cette étude est de quantifier le flou en analysant le contenu de l'image après acquisition sans mire de référence. Notre finalité est de pouvoir se comparer aux normes industrielles concernant la netteté de numérisation et de pouvoir l'utiliser dans la chaîne modulaire vu dans l'introduction générale.

La première partie de ce chapitre expliquera la problématique de netteté dans le cas de la numérisation. Nous commencerons par expliquer l'appareil optique pour comprendre ses mécanismes et ses caractéristiques. Nous verrons la fonction d'étalement du point qui caractérise les défauts de flou de l'appareil optique.

La deuxième partie se focalisera sur les contraintes industrielles qui requièrent une mire de référence, nous détaillerons certains critères correspondant aux mesures de la netteté de l'image et nous verrons que la fonction de transfert du module est essentielle pour faire ces mesures.

La troisième partie du chapitre propose des outils de mesure du flou de numérisation. Nous présenterons divers approches en se basant sur la structure du document ancien et nous les combinerons afin d'améliorer leur précision. Et nous terminerons par une comparaison entre la mesure obtenue par les normes et notre mesure de netteté.

3.2 La mise au point

Pour expliquer le fonctionnement de la mise au point, nous reprendrons certains éléments d'optique géométriques. L'ensemble de ces éléments constitue un système optique. Dans notre cas, et dans les cas expliquant l'usage d'un appareil photographique, notre système optique est caractérisé par une lentille convergente.

Nous considérons que le système est centré, c'est-à-dire que la lentille possède un axe optique (il existe un point lumineux qui n'est pas dévié par la lentille). Le foyer principale image (que l'on note F') est le point conjugué d'un objet placé à l'infini sur l'axe optique ; autrement dit c'est le point de convergence d'un faisceau parallèle à l'axe optique. Certains rayons ont des particularités dans le cas des lentilles convergentes. Trois rayons doivent être considérés, ces rayons sont présentés dans la figure 3.1. Le rayon 1 est parallèle à l'axe optique, il ressort sous forme d'un rayon passant par le foyer image. Le rayon 2 passe par le centre optique et il n'est pas dévié. Et le dernier rayon passe par le foyer objet et ressort du système optique en un rayon parallèle à l'axe optique.

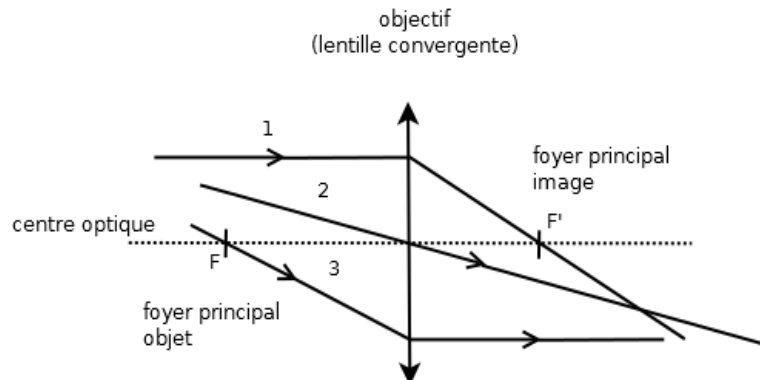


FIGURE 3.1: Rayons particuliers dans le cas d'une lentille convergente

La **mise au point** consiste à régler un appareil optique de telle façon que la convergence de la lentille fait passer les rayons par le foyer principal image. Dans le cas des rayons provenant de l'infini, la convergence de la lentille va systématiquement faire passer le rayon lumineux par le foyer principal image car les rayons sont considérés parallèles à l'axe optique. Alors que dans le cas d'un rayon provenant d'une distance plus faible, en l'occurrence de l'objet à numériser, cela nécessite un réglage. Le scanner est dans ce dernier cas. Nous représentons la mise au point dans la figure 3.2 où notre foyer principal image se confond avec le capteur.

En pratique, la capture d'une information lumineuse à travers la lentille n'est pas parfaitement représentée à cause d'un étalement (fonction d'étalement, PSF). Par conséquent, pour déterminer si un point est représenté nettement, nous comparerons l'étalement de ce point à un cercle de référence. Ce cercle est appelé le cercle de confusion (cf. figure

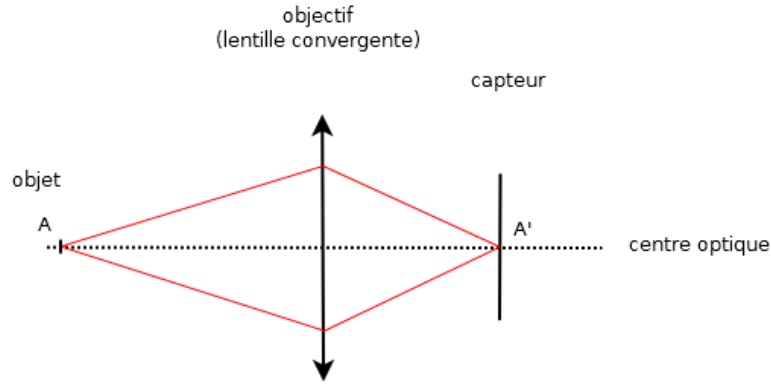


FIGURE 3.2: Mise au point

3.3). Il est construit en fonction d'un angle d'incidence passant par le centre optique d'une erreur ϵ . Nous avons alors un diamètre c qui prend en compte la distance entre l'objectif et le capteur, c'est-à-dire la distance focale.

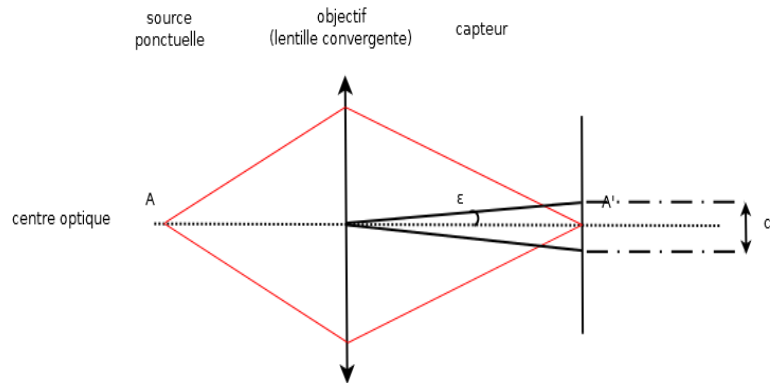


FIGURE 3.3: cercle de confusion

L'imprécision de la netteté est corrélée avec le diamètre du cercle de confusion. Nous énumérons les éléments qui ont une influence sur ce cercle et sur la création de ce point.

La **distance focale** est la distance entre la lentille et le foyer image. C'est un paramètre important sur la mise au point puisqu'il contrôle la taille du cercle de confusion. Nous remarquons bien dans la figure 3.4 l'étalement du point sur le capteur.

Le **champ angulaire** est l'angle de vision nette de la caméra. L'angle de vision θ est calculé par la relation suivante :

$$\tan\theta \approx \frac{l}{2f} \quad (3.1)$$

avec l la diagonale du capteur (qui est a priori rectangulaire) et la distance focale f . L'angle joue sur la convergence de la lentille. Si nous modifions cette convergence alors nous modifions le foyer principale image. Nous observons, dans la figure 3.5, que la

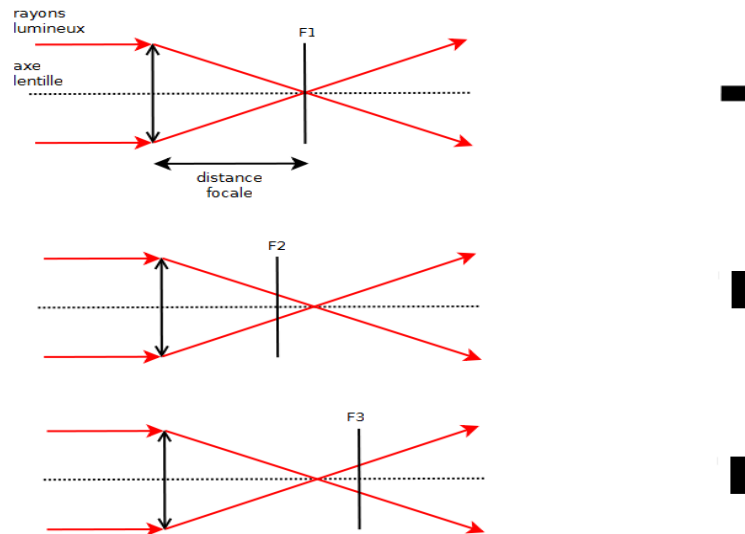


FIGURE 3.4: Focale - Ici, nous pouvons voir que la modification de la focale impacte directement l'information sur le capteur par un étalement du point.

convergence des rayons lumineux modifie l'information sur le capteur qui devient très vite aberrante.

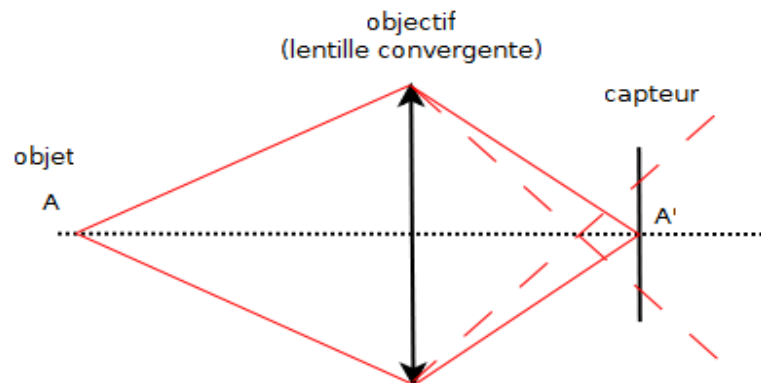


FIGURE 3.5: La convergence des rayons de la lentille. Nous pouvons voir que la convergence des rayons lumineux modifie la position du foyer principal image, par conséquent, l'information du point sur le capteur s'étaie.

L'**ouverture** est caractérisée par le diamètre de l'entrée de l'objectif. Il est géré par un mécanisme permettant de modifier la quantité de la lumière reçue par le capteur. Ce procédé a une conséquence directe sur la formation du point puisque plus il y a de lumière, plus l'étalement du point sera faible à cause de la quantité importante du point. Mais, plus il y a de lumière, moins la conversion numérique est précise.

La **profondeur de champ** est la zone de netteté de l'objet ou de la scène, on la caractérise par les points extrêmes nettes de l'axe optique vu sur le capteur. Si nous reprenons la figure 3.3, nous pouvons estimer un espace net à partir du diamètre c du cercle de

confusion. Nous avons représenté dans la figure 3.6 les rayons extrêmes inclus dans le cercle de confusion.

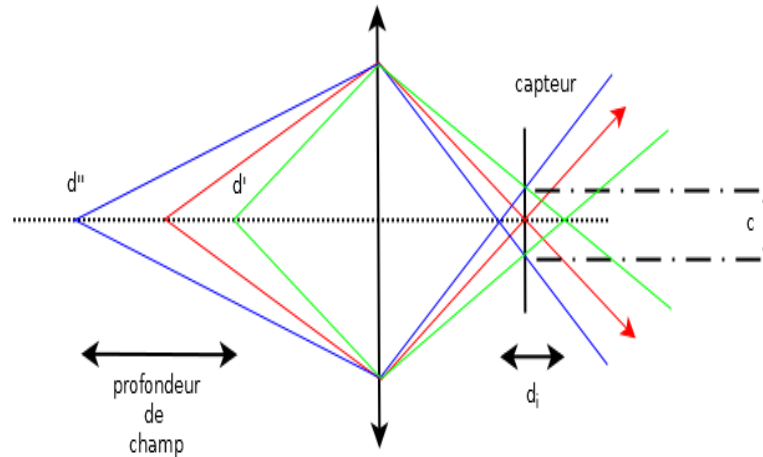


FIGURE 3.6: La profondeur de champ

Le calcul de **la distance optimale de mise au point** permet de trouver la distance de mise au point en fonction de la distance de l'objet le plus proche que nous souhaitons net (soit d') et le plus loin (soit d'') qui s'exprime par la relation suivante :

$$\frac{1}{d'} + \frac{1}{d''} \approx \frac{2}{d} \quad (3.2)$$

La profondeur de champ peut s'exprimer par la relation suivante :

$$d' d'' \approx \frac{2cN \cdot O}{f^2} d^2 \quad (3.3)$$

A partir de cette relation, nous pouvons en déduire que pour augmenter la profondeur de champ, soit $d' d''$, nous pouvons augmenter l'ouverture et la distance optimale de mise au point et que la profondeur de champ diminue quand la focale augmente.

Dans le cas pratique, nous pouvons représenter la profondeur de champ dans la figure 3.7. Afin d'obtenir une plus grande souplesse, la profondeur de champ doit être grande. Cela permet de garder une numérisation nette dans le cas d'un livre. La profondeur de champ peut être interprétée comme une tolérance de netteté puisque l'épaisseur du livre modifie la distance entre l'objet et la caméra.

Nous remarquons que les scanners linéaires (voir la figure 1.2 de 1) et matriciels (voir la figure 1.3 de 1) ont un angle de vision différent pour que l'image puisse être nette sur la

totalité de la zone à numériser. En théorie le scanner linéaire est susceptible d'être hors champ sur le haut et le bas de la numérisation, tandis que le scanner matriciel l'est sur les quatre extrémités.

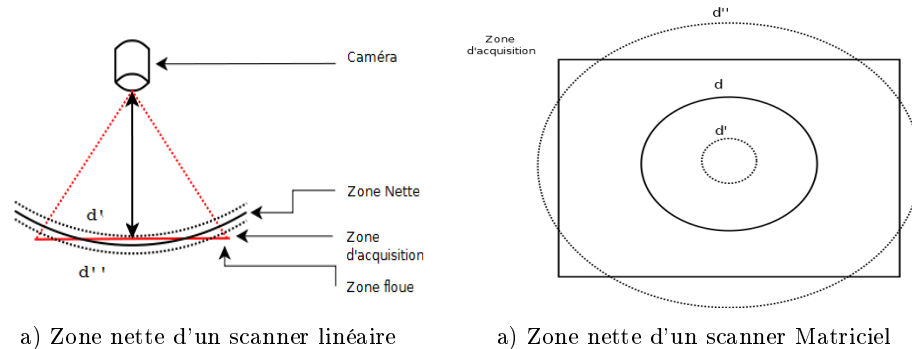


FIGURE 3.7: Zone nette d'un scanner

Nous avons énuméré les différents composants de l'appareil optique permettant de faire une acquisition nette. Ce réglage demande de jouer sur plusieurs paramètres en fonction de la scène à capturer. Son usage devient alors contraignant car il faut régler en permanence son matériel. Par conséquent, un dispositif est mis en place pour effectuer la mise au point automatique (autofocus). Il existe deux types d'autofocus : actif et passif.

Les méthodes actives consistent à émettre un signal, puis de calculer la distance entre l'appareil et le sujet en fonction du temps mis par le signal pour faire l'aller-retour. Ce signal est émis par infrarouge ou par ultrason, mais les deux sont sujets à des interactions avec l'environnement entraînant des erreurs.

Les méthodes passives sont basées sur le contraste. Cela consiste à faire varier l'appareil optique et de déterminer la meilleure combinaison en fonction du critère à maximiser ou à minimiser ayant pour but d'obtenir l'image la plus nette possible. L'autofocus dit à mesure de contraste cherche à maximiser la différence entre la luminosité la plus claire et la plus foncée. L'autofocus à contraste de phase se base sur le principe du télémètre qui sépare l'information lumineuse en deux. La séparation de la luminosité permet de se mettre dans un problème similaire à la figure de la focale (cf. la figure 3.4) en les faisant à nouveau converger sur un plan, par conséquent, la mise au point est faite lorsque les deux rayons minimisent une certaine taille.

Dans le cas de la numérisation, l'usage d'un autofocus est généralement utilisé pour un scanner matriciel, tandis que le scanner linéaire est réglé manuellement. De plus, pour une question de productivité certains utilisateurs enlève l'autofocus pour numériser plus vite.

Nous allons maintenant aborder dans la section suivante la fonction d'étalement du point et expliquer ses caractéristiques pour un mauvais réglage de l'acquisition.

3.2.1 Fonction d'étalement du point

Nous pouvons trouver un état de l'art complet sur le sujet dans la troisième version du "The image processing handbook" [Russ, 2010]. La création d'une image est une projection d'une scène en trois dimensions (3-D) sur un plan. La projection se fait sur le capteur par l'appareil optique. Cette information est ensuite enregistrée par le capteur. Et ce dernier va convertir le signal reçu en une image numérique. Soit $g(x, y)$ l'image produite, on peut modéliser cette projection par la formule suivante :

$$g(x, y) = s \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y; s, t) f(s, t) ds dt \right) + n(x, y) \quad (3.4)$$

Dans la formule ci-dessus, $g(x, y)$ correspond à l'image mesurée, et $f(s, t)$ est utilisé pour définir l'image idéale, qui est la projection parfaite de la scène. Le système de coordonnées n'est donc pas le même entre l'image idéale et l'image mesurée. $h(x, y; s, t)$ est une réponse impulsionnelle 2-D du système linéaire flouté que l'on peut interpréter comme les erreurs de projection dues à l'optique aboutissant aux différences entre l'image mesurée et l'image idéale. s est la réponse du capteur. La contribution du bruit est traduite ici par un procédé additif aléatoire dont les corrélations statistiques ne sont pas liées à l'image mais au défaut de l'appareil d'acquisition.

Si la réponse impulsionnelle est stationnaire sur l'image, elle est définie par une fonction où seul les arguments sont les différences $x-s$ et $y-t$ (voir l'équation 3.5).

$$g(x, y) = s \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x - s, y - t) f(s, t) ds dt \right) + n(x, y) \quad (3.5)$$

Dans ce cas, la formule 3.6 se traduit par :

$$g(x, y) = s \{h(x, y) * f(x, y)\} + n(x, y) \quad (3.6)$$

où $*$ est la convolution 2-D, où (x, y) sont les coordonnées du capteur.

On obtient la formule dans un espace discret en échantillonnant.

$$g(i, j) = s \{h(i, j) * f(i, j)\} + n(i, j) \quad (3.7)$$

où, ici, $*$ est la convolution 2-D discrète et (i, j) sont les coordonnées de l'image. Bien souvent, le capteur non-linéaire est par convention, négligé (ou linéarisé). Lorsque cette non-linéarité est ignorée, 3.8 est réduite à un modèle de convolution linéaire :

$$g(i, j) = h(i, j) * f(i, j) + n(i, j) \quad (3.8)$$

Pour la suite, nous allons utiliser la notation simplifiée suivante 3.9 :

$$g = h * f + n \quad (3.9)$$

On peut décrire et/ou interpréter l'équation 3.9 en se basant sur le défaut optique que nous avons schématisé à la figure 3.4. Lorsque la calibration optique est parfaite alors les rayons lumineux forment un point. Lorsqu'il y a une erreur, le point "s'étale". On définit h comme étant la fonction d'étalement du point (Point Spread Function en anglais), qui est une réponse impulsionnelle liée aux défauts optiques.

Pour obtenir une approximation de la PSF, la solution proposée dans [Russ, 2010] consiste à utiliser une mire basée sur la fonction sinusoïdale. Nous rappelons que la transformée de Fourier de la fonction $\sin(w_0 t)$ est deux impulsions de Dirac qui dépendent du coefficient w_0 du sinus. Par conséquent, lorsque nous faisons une acquisition de ce type de mire, nous pouvons calculer sa transformée de Fourier et la comparer avec son résultat théorique. Lorsque l'image perd en netteté, les points mesurés dans l'espace de Fourier présentent une multiplication des points théoriques par une gaussienne

Nous pouvons considérer alors que la fonction d'étalement du point a des propriétés de flou de type gaussien. Nous modéliserons h , une fonction gaussienne et nous simulerons un flou de mise au point par l'utilisation d'un tel filtre tout au long du sujet. Par conséquent, l'objectif est de mesurer et d'interpréter les paramètres de la gaussienne, afin de connaître la qualité de numérisation. En résumé, l'objectif de la détection du flou de focus sera de mesurer les caractéristiques de h , un filtre gaussien.

3.3 Numérisation industrielle

La diversité du matériel de numérisation est devenue très importante : caméras, méthodes de numérisation, capteurs. Les différents appareils vont obtenir des images très différentes. Le but est d'obtenir une image similaire d'un appareil à un autre et qu'elle puisse converger vers une acquisition la plus précise possible. Depuis 1997, des normes ont été créées sur les appareils d'acquisition, ces normes établissent des caractéristiques à respecter en fonction de la mesure spécifique. Ils existent une vingtaine de métriques. Cette section va ainsi énumérer quelques métriques relatives à la netteté de numérisation.

Ces normes évoluent au fur et à mesure du temps et elles sont la référence pour quantifier la qualité des outils numériques. Nous utiliserons le guide de qualité d'image [FADGI, 2010] (et [Initiative, 2010]) dans lequel la FADGI (Federal Agencies Digitization Guidelines Initiative) impose un standard de qualité numérique pour la numérisation. Il est intéressant d'étudier ces différentes métriques afin de mieux cerner la problématique de la quantification de la netteté d'un appareil ainsi que de voir quelles sont les mesures standards (cf figure 3.8).

Ce guide donne aussi les méthodes standards pour l'utilisation des mires. C'est une référence technique pour les opérateurs professionnels. On trouve des normes pour tous les types d'acquisitions (documents, documents manuscrits, photographies, photographies aériennes, etc).

Nous retrouvons les mesures utilisées pour détecter les défauts des scanners. Ces mesures se basent sur un document à numériser dont on connaît parfaitement la représentation (une mire) pour régler l'appareil. Les mires contiennent une information spécifique aux critères à quantifier. Il existe des mires pour régler la couleur, la géométrie, la résolution, etc.

Les mesures en corrélation avec la netteté se déduisent des critères évaluant la bonne représentation du signal. On retrouve ces métriques dans la catégorie des réponses spatiales fréquentielles. On peut trouver un tableau récapitulatif dans [FADGI, 2010] que nous avons repris dans la figure 3.8. Nous détaillons en suivant les mesures concernant la netteté.

- SIGNAL -

SFR - Spatial Frequency Response – (ISO 12233, ISO 16067-1, ISO 16067-2, ISO 15524) MTF – Modulation Transfer Function	
definition : A spatial frequency descriptor of an imaging system's ability to maintain the relative contrast of input stimuli	
Engineering Metric	
Derivative Metrics	<p>Sampling Rate Definition: The reciprocal of the center-to-center distance between closest adjacent pixels. The number of samples per unit. - Megapixels - Dots per inch (dpi) - Pixels per inch (ppi) - Sampling frequency</p>
Related Descriptive term	<p>Resolution Definition: An imaging system's ability to resolve finely spaced detail. The level of spatial detail that can be resolved in an image - Blurred - Soft - Sharp - In/Out of focus - Spherical aberration - Spatial detail</p>
Evaluation Criteria (units)	<p>Sharpening Definition: Amplification of the SFR by means of image processing to achieve sharper appearing images - Oversharpening (halong, garish edges) - Snap - Edgy, Sharp, Crisp - Edge enhancement - Unsharp masking</p>
	<p>Acutance Definition: An objective SFR based metric that is used as a correlate to perceived image sharpness. - Sharp</p>
Possible failure causes	<p>Flare Definition: a skirt or wide spreading of light. - Low contrast - Hazy - Ghosting - Veiling flare - Glare - Integrating cavity effect (ICE)</p>
	<p>Depth of Focus Definition: The distance along the optical axis that remains within acceptable focus. - Depth of field - Circle of confusion - Focus tolerance - Hyperfocal distance</p>
	<p>Resolution - 10% sampling efficiency based on Luminance SFR units: (unit less) - Min/Max 10% spatial frequency limits of Luminance SFR units: dpi, cycles/mm - Min/Max 50% spatial frequency limits of Luminance SFR units: dpi, cycles/mm</p>
	<p>Sharpening - Max SFR gain units: % SFR response - Sign of SFR slope units: positive/negative slope value</p>
	<p>Acutance - Area under the SFR as weighted by an appropriately chosen visual contrast function. units: TBD</p>
	<p>Flare - % Flare - units: (unit less)</p>
	<p>Depth of Focus - Distance along the optical axis that remains in acceptable focus units: inches, mm.</p>
	<p>Flare - Dirty lens - Light source directed into lens - Poor quality lens - Stray light</p>
	<p>Acutance - Optical performance exceeds sampling rate</p>
	<p>Sharpening - Over aggressive sharpening settings - Insufficient signal to amplify - Thinking that if a little is good then more must be better.</p>
	<p>Resolution - Poor (auto) focus - Poor optics - Poor choice of aperture stop - Mechanical vibration - Over aggressive noise control</p>

FIGURE 3.8: Tableau récapitulatif des normes lié à la représentation du signal

3.3.1 Sampling Rate

Cette étape consiste à mesurer l'échantillonnage de l'image. Pour cela, on utilise une mire dont on connaît les caractéristiques géométriques (en l'occurrence, la distance entre deux objets) : on obtient alors une mesure en dpi (1 inch \approx 24.5 mm). Nous rappelons que la résolution optique est déterminée par le nombre de cellules du capteur et la résolution de l'image est déterminée par le nombre de pixels entre deux référentiels géométriques. Il ne faut donc pas les confondre étant donné qu'il ne peut avoir plus de cellules que de pixels. On considère que plus l'échantillonnage est grand, plus la caméra doit être susceptible de faire la distinction entre les éléments.

3.3.2 Resolution

Pour expliquer ce critère de resolution, nous nous baserons sur les normes [ISO 12233, 2014], [ISO 16067-1, 2003], [ISO 16067-2, 2004] et [ISO 21550, 2004]. La *resolution* est la capacité à décrire les détails de l'image pour un échantillonnage donné. Une mesure visuelle est donnée sur la mire 3.9 en observant la lisibilité des motifs horizontaux et verticaux de plus en plus serrés. La mesure quantitative est faite par la fonction de transfert de modulation (FTM). L'approximation de la FTM se calcule en faisant le ratio entre la courbe théorique et la courbe mesurée. La fonction peut être aussi exprimée en calculant le contraste pour chaque fréquence :

$$M(f) = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (3.10)$$

avec I la luminance de la fréquence f . La FTM permet de mesurer l'atténuation des ondes sinusoïdale en fréquence spatiale (voir la figure 3.10) et permet de quantifier automatiquement la résolution. La FTM reprend l'analyse faite de la fonction d'étalement pour chaque fréquence de la fonction sinus.

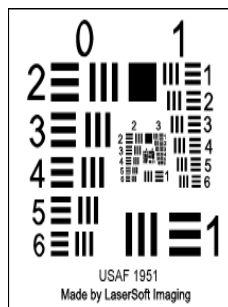


FIGURE 3.9: Mire de résolution - Sampling Rate de SilverFast

La FTM a deux points caractéristiques : le point dit à 50% et à 10%, c'est-à-dire que l'on regarde pour quelles valeurs de f , $M(f)$ est égale à 0.5 et à 0.1. Le point à 50% est la

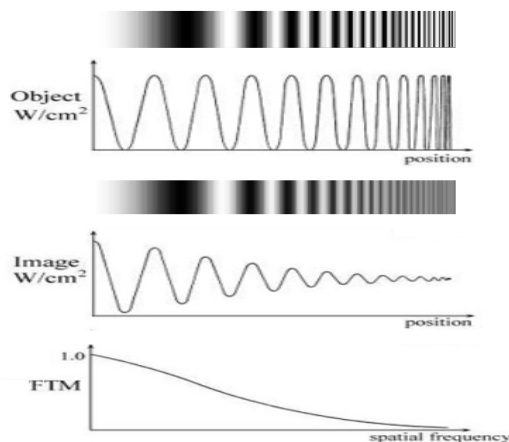


FIGURE 3.10: Fonction de transfert de modulation. De haut en bas, nous avons représenté la mire sinusoïdale avec le profil associé. Ensuite, nous avons l'observation théorique d'une caméra et le profil associé. Nous constatons que le premier cycle est bien représenté puis le signal converge à cause de la diffraction optique. La dernière courbe est la FTM.

représentation psycho-visuelle de la netteté perçue. Plus ce point sera élevé plus l'image apparaîtra nette. Toutefois, le modèle parfait de la FTM est défini en fonction d'une diffraction théorique, il existe alors une courbe théorique en fonction des caractéristiques de la lentille et par conséquent, ce point à 50 % est seuillé. S'il existe un point supérieur à cette courbe alors des traitements d'images ont été appliqués. L'autre point est le seuil à 10%, il correspond à la capacité du matériel à distinguer les détails.

3.3.3 Sharpening

Pour expliquer le critère sharpening, nous nous baserons sur les normes [ISO 12233, 2014], [ISO 16067-1, 2003] et [ISO 16067-2, 2004].

Le *sharpening* est une extension de la *resolution* puisqu'elle mesure la FTM spécifiquement sur les contours. La FADGI suggère comme démarche de calculer une nouvelle image en appliquant un filtre passe-haut. La comparaison des deux FTM donne une mesure du *sharpening*. Effectivement, si l'image traitée devient nette comparée à l'image d'origine, alors le ratio entre les courbes augmentera. La mesure FADGI est basée sur le gain maximum entre les deux FTM qui quantifie le critère du *sharpening*.

3.3.4 Acutance

L'acutance est une notion similaire à celle du *sharpening*. La définition donnée par l'ISO [6196–5] 1987 est : "l'acutance est une mesure objective de la capacité matérielle photographique de montrer une ligne de démarcation entre deux zones adjacentes composées d'une zone claire et d'une zone sombre". En d'autres termes, l'acutance est la mesure

du contraste dans un contour (voir figure 3.11 et 3.12). Ce phénomène n'a pas vocation d'augmenter le contraste global de l'image car ceci est un défaut mais il a pour vocation d'améliorer la distinction des détails aux niveaux des moyennes fréquences. L'acutance cherche à quantifier la distinction des fréquences sensibles à l'œil qui se situe au niveau des moyennes fréquences. Elle est ainsi calculée à partir de la FTM.

Les fréquences sensibles à l'œil humain sont définies par une fonction appelée le Contrast Sensitivity Function [Barten, 1990]. La construction de cette fonction a été faite à partir de la capacité à distinguer des lignes (dont la fréquence varie) selon l'angle de vision. L'objectif de la mesure est de vérifier que les fréquences sensibles à l'œil sont bien représentées. Pour cela, on utilise le calcul de la fonction 3.11 donnée par la documentation d'Imatest, logiciel de mesure de qualité d'image [Imatest, 2013].

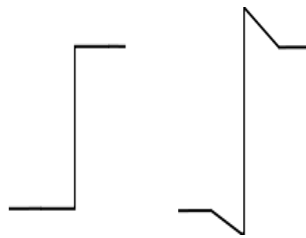


FIGURE 3.11: Exemples de deux profils de contours. Nous observons que le second profil présente un contraste plus important au niveau des contours.



FIGURE 3.12: Exemple d'images mettant en évidence l'acutance. La première image est l'image originale sans traitement. La seconde et la troisième image sont les résultats de l'image avec un filtre passe-haut faible et fort. On a une sensation de netteté qui est de plus en plus marquée.

$$Acutance = \frac{\int_0^{\infty} FTM_L(\nu)CSF(\nu)}{\int_0^{\infty} CSF(\nu)} \quad (3.11)$$

avec

$$CSF(\nu) = a\nu^c \frac{\exp(-b\nu)}{K}, a = 75, b = 0.2, c = 0.8, K = 34.05 \quad (3.12)$$

Nous notons qu'il n'existe pas de mire et d'ISO permettant de faire le calcul directement. Actuellement, le calcul de l'acutance est basé sur une mire de résolution et il est nécessaire de convertir la fréquence spatiale f (cycles / mm) en fréquence angulaire ν (cycles / degré). Afin de faire cette opération, on utilise la fonction suivante :

$$\nu = f * \left(\frac{\Pi}{180} \right) * D * PPi_{effectif} \quad (3.13)$$

avec D la distance de mise au point (en mm) et PPi le nombre de pixel par unité de distance (en mm). L'objectif de la CSF est de pondérer les fréquences sensibles à l'œil. Cette pondération renforcera les caractéristiques de la FTM comme nous le voyons dans la figure 3.13. Nous notons que la FADGI utilise le calcul de l'aire de ce résultat pour quantifier l'acutance ce qui signifie que la capacité séparatrice du capteur est aussi prise en compte dans le calcul.

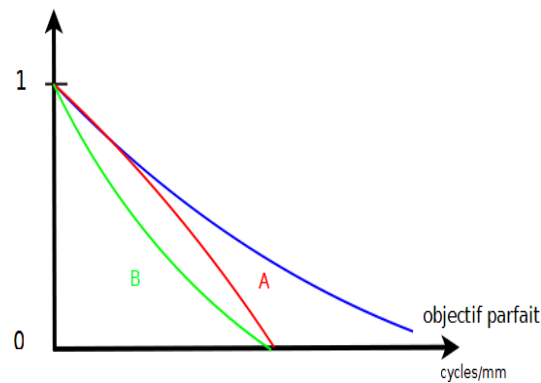


FIGURE 3.13: Exemple de FTM avec des courbes ayant des acutances différentes. Nous observons que la courbe A et B ont la même capacité à différencier les détails, mais A conserve un contraste plus élevé.

3.3.5 Conclusion

Les normes mesurent la fidélité de représentation de l'appareil. Ici, nous nous sommes focalisés sur la quantification de la netteté. Nous retenons que cette quantification est faite à partir de plusieurs mesures basées sur les propriétés de la FTM. Nous en concluons que la mesure de netteté fait sur une mire demande un nombre important de critère pour la quantifier : la résolution de l'image, la résolution effective, le sharpening et l'acutance

Nous constatons qu'elles sont toutes basées sur la FTM et qu'il est nécessaire d'utiliser une mire spécifique. Dans le cadre d'une mesure du flou sans référence, il va falloir investiguer différentes approches que la FTM.

3.4 Mesure de netteté sur une image unique sans référence

Le flou de focus est un défaut optique assimilable à un « filtrage passe-bas » de type gaussien. Notre mesure du flou dans une image se basera sur l'estimation des paramètres du modèle gaussien (moyenne, écart type) dans un premier temps, puis nous réitérons notre estimation sur des images numérisées ayant une mire de référence dans un second temps. La présence de la mire va nous permettre de calculer la FTM ce qui va permettre de comparer notre estimateur et les descripteurs utilisés par les normes industrielles des scanners.

La mesure du flou se fait sur l'analyse des contours et de leur étalement. Dans le cadre du document, la forte proportion de texte nous permet de disposer d'une information de contours très importante. Dans une première partie, nous citerons des références de l'analyse des contours. Puis, dans la seconde partie, nous détaillerons notre méthode de catégorisation du flou. Notre approche se base sur un ensemble de descripteurs qui participent à la caractérisation des contours et de leur étalement. Nous reprendrons certains descripteurs vu dans la première partie. Cet ensemble de descripteurs sera utilisé dans un modèle « d'apprentissage supervisé » dont le résultat sera une catégorisation de la force du flou.

3.4.1 Étalement de contour

L'approche que nous proposons s'appuie sur la mesure de la qualité des contours sans référence. Dans le domaine de la compression d'une image, certaines des caractéristiques prises en compte pour évaluer la qualité de la compression sont basées sur la qualité de reproduction des contours et donc sur la connaissance de l'image d'origine. Même si les conditions sont différentes, certains des descripteurs utilisés pour qualifier la qualité des contours peuvent être retenus dans notre contexte. L'article de Marziliano [Marziliano et al., 2004] utilise une métrique basée sur l'étude des épaisseurs de contours verticaux, dont l'épaisseur est calculée en fonction de la distance en nombre de pixels séparant le minimum local du maximum local. La mesure finale est une moyenne calculée sur tous les contours détectés. Chung et al [Chung et al., 2004] ont établi une métrique sur le calcul du contraste, le module du gradient et l'épaisseur de contour. La méthode calculant l'épaisseur de contour est une version améliorée de Marziliano. La méthode propose alors une fusion des informations en pondérant les résultats en fonction du contraste local. Ce qui a pour conséquence de favoriser le résultat du gradient lorsqu'il y a beaucoup de contraste et de favoriser l'épaisseur du contour dans le cas contraire. Une amélioration "adaptative" a été proposée dans l'article de Narvekar et Karam [Narvekar and Karam, 2009] où la mesure du contour peut ne pas être prise en

compte. Les auteurs se basent sur l'étalement des contours défini par Marziliano, mais la solution qu'ils proposent est d'utiliser les informations locales de l'image en découpant l'image en bloc de même taille. Dans le bloc courant, une évaluation du contraste et du nombre de pixel de contour est faite. Ces informations évaluent la pertinence des contours inclus dans ce même bloc. De cette manière, les contours sont sélectionnés et pondérés en fonction de l'information locale.

Dans le domaine de l'amélioration d'image, Bae et Durand [Bae and Durand, 2007] définissent une carte du flou afin d'accentuer ou d'atténuer le flou présent dans certaines zones de l'image. Cette carte est construite en évaluant l'erreur de profondeur de champ en chaque point. Pour calculer cette information au niveau des pixels, la méthode évalue les informations de netteté au niveau des hautes fréquences en se basant sur l'article de Elder et Zucker [Elder and Zucker, 1998], puis, ils utilisent une méthode permettant d'étendre l'estimation du flou sur l'ensemble de l'image. Les auteurs de la méthode d'estimation du flou, Elder et Zucker, proposent d'évaluer différents niveaux de flou. Ils proposent de l'estimer en utilisant un gradient lissé par une gaussienne et la dérivée du résultat. En pratique, nous pouvons lissé le filtre dérivatif (vertical et horizontal) avec une gaussienne puis l'appliquer sur l'image ou encore appliquer directement la dérivée en x et en y d'une gaussienne qui s'écrit alors :

$$g^x(x, y, \sigma) = \frac{-x}{2\pi\sigma^4} \exp^{-(x^2+y^2)/2\sigma^2} \quad (3.14)$$

et

$$g^y(x, y, \sigma) = \frac{-y}{2\pi\sigma^4} \exp^{-(x^2+y^2)/2\sigma^2} \quad (3.15)$$

La particularité de leur approche est qu'ils font varier les paramètres de la gaussienne pour déterminer le niveau de flou du contour. Les auteurs expliquent qu'ils cherchent le seuil minimal le plus adéquate pour identifier le niveau du flou appliqué au contour en fonction des différents résultats. La combinaison de ces deux informations permet d'affiner la précision du niveau et ce point particulier est alors la quantification du flou en ce pixel.

L'article [Tai and Brown, 2009] suggère l'utilisation du contraste local de l'image pour obtenir les informations de netteté. L'analyse des contours proposée par Zhuo et Sim [Zhuo and Sim, 2011] est originale : pour chaque pixel de contour, le flou est calculé comme le rapport entre le gradient de l'image initiale et le gradient de l'image floutée par une gaussienne connue au niveau des contours comme le montre la figure 3.14.

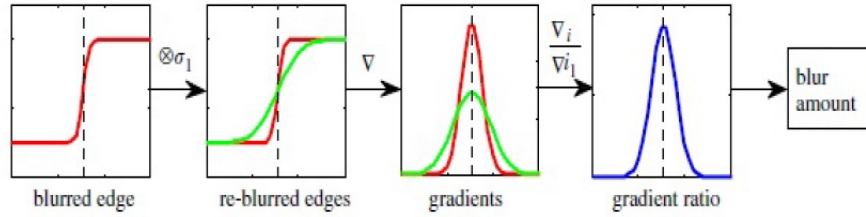


FIGURE 3.14: L'approche du defocus blur estimation [Zhuo and Sim, 2011] : ici, \otimes est le produit de convolution, Δ est l'opérateur du gradient et la ligne en pointillé est la position du contour.

D'autres démarches se basent sur une classification des contours pour en mesurer leur qualité. Par exemple, l'article de Tong et al [Tong et al., 2004] propose une méthode basée sur la structure des contours. A l'aide de la transformée en ondelettes de Haar de l'image, ils classifient les différents contours selon leurs caractéristiques à différentes échelles. Ainsi, ils déterminent une mesure du flou en se basant sur le nombre de "contours nets" et de "contours flous". Le même auteur, dans [Tong et al., 2005], propose un modèle d'apprentissage supervisé basé sur la dégradation des contours selon le type de compression. Le descripteur utilisé est la répartition locale des pixels de contour. D'autres auteurs utilisent la classification pour divers objectifs, l'article de Ke et al [Ke et al., 2006] suggère de quantifier la qualité de l'image en fonction de plusieurs descripteurs tels que la distribution spatiale des contours, la distribution des couleurs, l'analyse de l'histogramme des niveaux de gris de l'image ainsi qu'une analyse dans l'espace de Fourier.

Liu et al [Liu et al., 2008] proposent d'utiliser la classification pour segmenter les zones de flous de l'image. Cet article vise à caractériser le flou d'une région de l'image, l'image est ainsi découpée en bloc. Pour chacun des blocs, un ensemble de descripteur sont calculés, ces descripteurs ont la particularité d'identifier le flou de mise au point et/ou le flou de bougé. Les auteurs exploitent ces résultats par un classifieur bayésien pour déterminer le type de flou dans le bloc.

Un des descripteurs utilisé par Liu est construit à partir d'une analyse fréquentielle, ce type de descripteur permet de mettre en évidence certains défauts optiques (flou de bougé, flou de focus) grâce aux informations des contours dans une région de l'image. Les travaux de D. J. Field [Field et al., 1987] et [Field and Brady, 1997] ont mis en avant l'importance du coefficient directeur de la pente de la densité spectrale de puissance pour caractériser une image.

La densité spectrale de puissance est la puissance du signal pour représenter toutes les fréquences. La transformée de Fourier peut s'exprimer en coordonnée polaire plutôt que cartésien, ce qui donne :

$$F(u) = |F(u)| \cdot e^{i\varphi(u)} \quad (3.16)$$

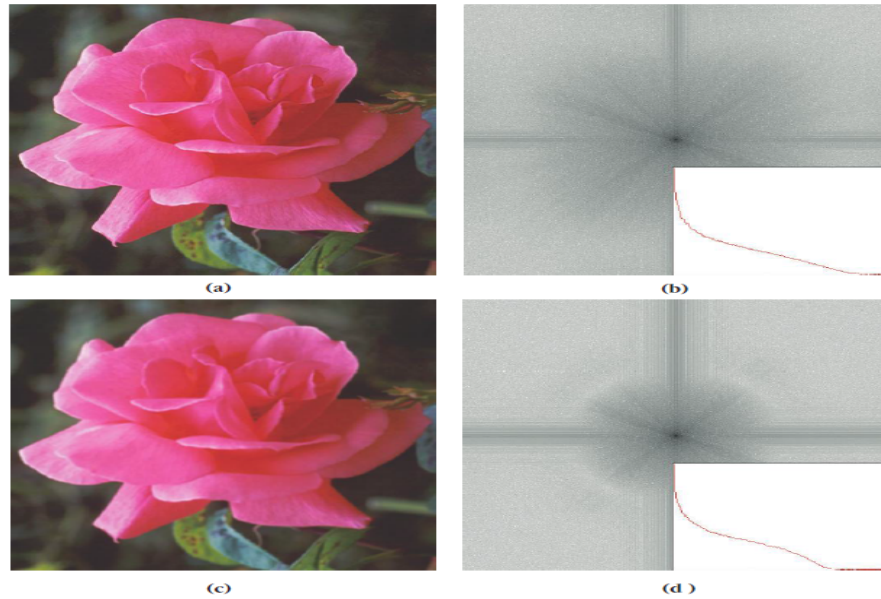


FIGURE 3.15: Nous avons l'image (a) nette et l'image (c) floue. Les images (b) et (d) représentent les transformées de Fourier et leur puissance spectrale de puissance associées aux image (a) et (c). [Russ, 2010]

où $|F|$ est l'amplitude et φ est la phase. Le carré de l'amplitude $|F(u)|^2$ est la densité spectrale de puissance de $f(x)$. Nous pouvons visualiser son comportement dans la figure 3.15.

Nous observons que la courbe (d) de la figure 3.15 est plus plate aux niveaux des hautes fréquences que la courbe (b) a cause du flou de l'image. Ainsi le coefficient directeur est une information donnant une information de netteté de l'image. Liu réutilise cette approche pour estimer la netteté.

Parmi les approches basées sur la transformée de Fourier, on notera la méthode spécifique au flou de mouvement d'une scène supposée statique et artificielle, proposée par Lelégard [Lelegard et al., 2010]. Leur méthode consiste à binariser la transformée de Fourier et à récupérer la forme du spectre qui dans le cas d'un flou de bougé fait apparaître une direction (cf. figure 3.16). Par contre, il n'est pas spécifié dans l'article l'impact du flou sur la méthode. Nous proposons de l'étudier par la suite dans nos conditions.

Nous utiliserons par la suite les descripteurs sélectionnés pour caractériser les informations de contours. Nous nous placerons dans un contexte d'images non-naturelles (documents anciens) contenant un flou uniforme gaussien (défaut de mise au point lors de la numérisation). Nous proposerons d'estimer les caractéristiques du flou à l'aide d'un modèle d'apprentissage supervisé qui se base sur un ensemble de descripteurs participant à la caractérisation du flou.

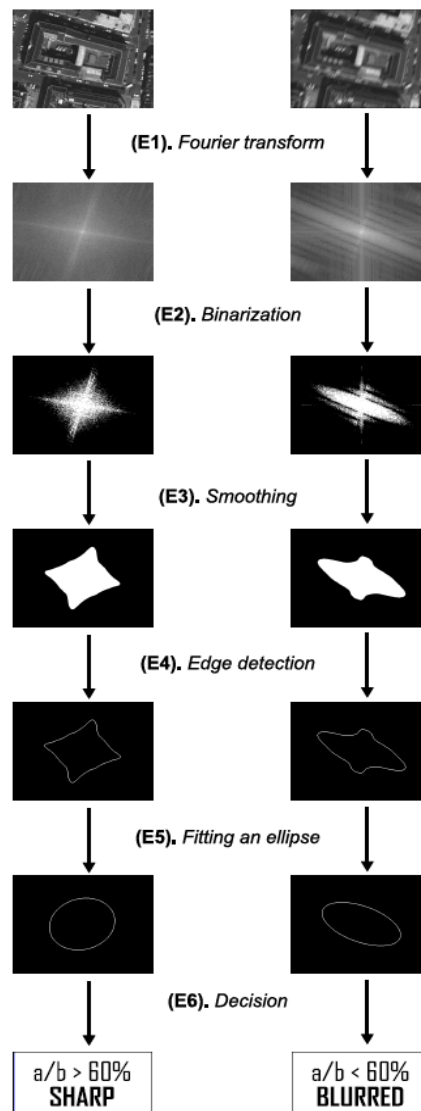


FIGURE 3.16: Schématisation de la binarisation faite dans l'article [Lelegard et al., 2010]

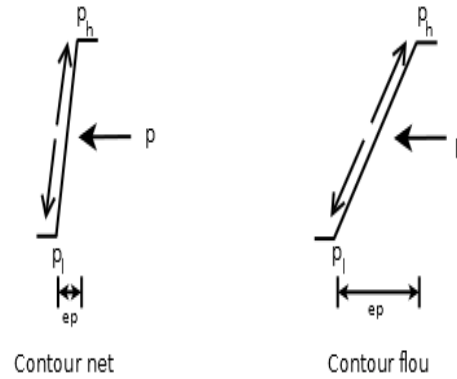


FIGURE 3.17: Profils de contour 1D

3.4.2 Descripteurs

L'étude de la section précédente montre qu'il existe de nombreux descripteurs pour déterminer la netteté (étalement) d'un contour. Chaque descripteur cherche à évaluer une des caractéristiques d'un contour tels que son amplitude, son épaisseur, sa dégradation en fonction d'un flou, ses caractéristiques fréquentielles, etc. La mesure de ces spécificités permet d'évaluer la netteté. L'ensemble de nos descripteurs sont calculés pour des images en niveaux de gris. Nous présentons ci-dessous les cinq descripteurs qui nous semblent les plus pertinents :

– **D_1 le maximum du gradient :**

Le descripteur le plus simple est le maximum du gradient. Le gradient de l'image mesure la transition entre deux pixels.

Afin de normaliser cette mesure selon les différentes images, nous diviserons par 255 (ou le nombre de niveaux maximal requis). Soit L l'image en niveaux de gris, on écrit alors :

$$D_1 = \frac{\max_{p \in L} (|\nabla p|)}{255} \quad (3.17)$$

, avec ∇ l'opérateur gradient.

– **D_2 l'histogramme des épaisseurs des contours :**

Ce deuxième descripteur se base sur une première segmentation des contours de l'image. La détection est obtenue par un seuillage global du gradient normalisé. La normalisation est faite à partir du maximum du gradient obtenu de l'image à la différence de D_1 .

Pour tous les pixels de contours P_c , nous calculons l'épaisseur de contour.

Pour définir l'épaisseur d'un contour, nous le parcourons dans les deux sens selon la direction du gradient (cf. figure 3.17). Le parcours s'arrête lorsque la dérivée est

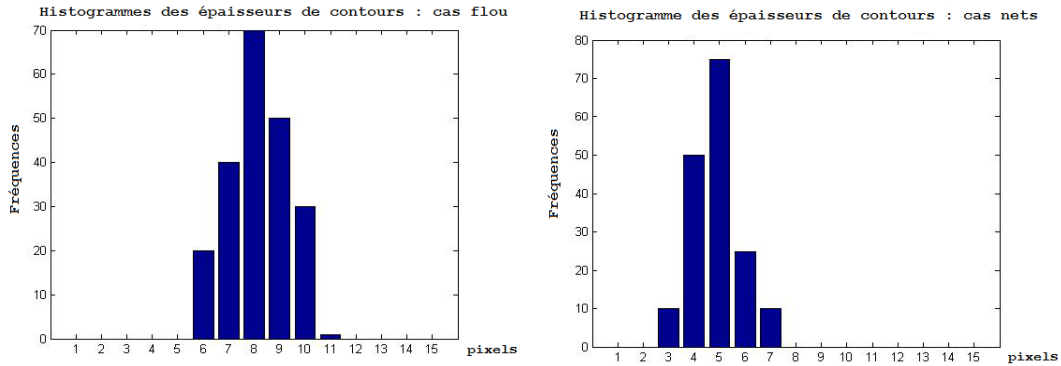


FIGURE 3.18: Histogrammes des épaisseurs de contours

inférieure à une valeur seuil proche de zéro correspondants aux point P_l et P_h . Soit $Ep(P_c)$ le nombre total de pixels d'un contour dont le pixel de départ est P_c . La mesure de l'épaisseur d'un contour est une mesure locale, nous proposons d'utiliser un histogramme H_{ep} pour analyser l'ensemble des contours sur toute l'image. L'histogramme sera construit à partir d'intervalle uniforme en N classes représentant l'épaisseur du contour. Le nombre de classe dépend de la résolution de l'image. Par exemple, pour une image en 300 Dpi l'épaisseur des contours varie entre 2 et 10 pixels. La contribution des $Ep(P_c)$ sera pondérée par l'amplitude du gradient du point P_c courant. Nous considérons qu'un contour plus fort permet d'obtenir de manière plus précise la tangente du contour. De ce fait, nous utiliserons le module du gradient (normalisé) de P_c comme valeur de pondération. Le résultat obtenu sera la moyenne de cet histogramme qui s'écrit :

$$D_2 = \frac{\sum_{i=1}^N |H_{ep}(i)| \cdot i}{\sum |H_{ep}(i)|} \quad (3.18)$$

– **D_3 la catégorisation de Haar :**

La méthode des catégorisations de contours est basée sur l'article de Tong et al [Tong et al., 2004].

Leur approche se base sur une classification des contours en utilisant les ondelettes de Haar. L'algorithme se découpe en deux parties, la première est la segmentation des contours (1) et la seconde classeifie les contours (2).

1. La segmentation est faite sur les trois niveaux de la transformée en ondelette de Haar de l'image. Les auteurs proposent de calculer une carte des contours basés sur les blocs des hautes fréquences à savoir LH, HL et HH (cf. [Meyer, 1990]) et que nous pouvons le voir à la figure 3.19.

LL ₃	HL ₃	HL ₂	HL ₁ : Horizontal Detail
LH ₃	HH ₃		
LH ₂		HH ₂	
LH ₁ : Vertical Detail			HH ₁ : Diagonal Detail

FIGURE 3.19: Pyramide d'ondelette sur trois niveaux

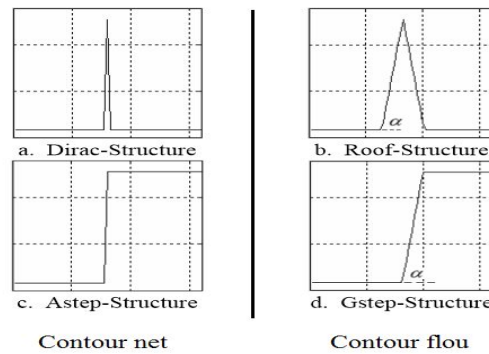


FIGURE 3.20: Les différents types de structure de contour

Le calcul proposé par les auteurs est le suivant :

$$Emap_i(k, l) = \sqrt{LH_i^2 + HL_i^2 + HH_i^2} \quad (i = 1, 2, 3) \quad (3.19)$$

où, i le niveau de décomposition de l'ondelette et (k, l) les coordonnées du pixel sur le niveau i . On calculera un maximum sur une fenêtre centrée sur le pixel (k, l) pour tous les $Emap$. La fenêtre est de plus en plus petite au fur et à mesure que l'échelle augmente. Le pixel courant sera labélisé comme point de contour lorsque la valeur obtenue dépassera une valeur seuil.

2. A partir des contours identifiés, les auteurs proposent quatre catégories de contours qui sont appelés structures (cf figure 3.20). Ils définissent les structures en 'Dirac' et en 'Astep' pour les contours nets, et les structure en 'Roof' et en 'Gstep' pour les contours flous. Ils définissent un ensemble de règle permettant de les classifier. Les règles considèrent les valeurs des coefficients à différentes échelles d'ondelettes.

La mesure de netteté est le ratio du nombre de contours nets sur le nombre total de contours, soit :

$$D_3 = \frac{\#Astep-struct + \#Dirac-struct}{\#\text{contour}} \quad (3.20)$$

– **D_4 l'estimation de netteté de l'autofocus (Defocus Blur Estimation) :**

Ce descripteur est basé sur la première partie de l'article de [Zhuo and Sim, 2011] qui estime la dégradation d'un contour par un filtre gaussien.

Soit I l'image initiale, que nous lissons par un filtre gaussien d'écart-type σ pour obtenir l'image lissée L . Nous détectons les pixels de contour P_c à partir de l'image initiale. La détection est faite à partir d'un seuillage global du gradient normalisé, comme nous l'avons fait pour le descripteur $D2$. Pour tous les pixels de contours P_c , nous calculons le ratio entre les deux contours par la formule suivante :

$$R(P_c) = \frac{\|\nabla L(P_c)\|}{\|\nabla I(P_c)\|} \quad (3.21)$$

Le résultat du ratio sera inférieure ou égale à 1, car $\|\nabla I\| > \|\nabla L\|$ lorsque $\|\nabla I\| > \delta$ avec δ proche de 1. Pour obtenir la mesure du flou, nous moyennons ces valeurs puisque l'erreur de mise au point est quasi uniforme sur l'ensemble de la numérisation. Ce descripteur s'écrit alors :

$$D_4 = \frac{\sum_{c=1}^n R(p_c)}{n} \quad (3.22)$$

avec n le nombre de pixel de contours.

– **D_5 la binarisation de Fourier :**

Ce dernier descripteur est obtenu par une approche similaire à celle proposée par Lelégard et al [Lelegard et al., 2010] (cf. la figure 3.16).

Nous calculons le module de la transformée de Fourier de l'image. Nous binarisons en prenant comme seuil la moyenne. Nous supprimons les éléments isolés par des outils de filtrage morphologiques afin d'avoir deux éléments distincts. Nous avons centré les hautes fréquences. Nous obtenons alors les basses et hautes fréquence comme on peut le voir dans la figure 3.21. Nous constatons dans l'ensemble d'image que la binarisation du module de la transformée de Fourier évolue en fonction du flou.

Nous constatons que les images de texte produisent une binarisation proche d'un cercle. Lorsque nous avons un élément dans l'image qui n'est pas du texte, nous avons un agrandissement du cercle. Notre évaluation va alors schématiser le problème de la manière de la figure 3.22, nous représentons les deux labels après binarisation et nous les caractérisons en fonction du cercle des hautes fréquences. La mesure consiste

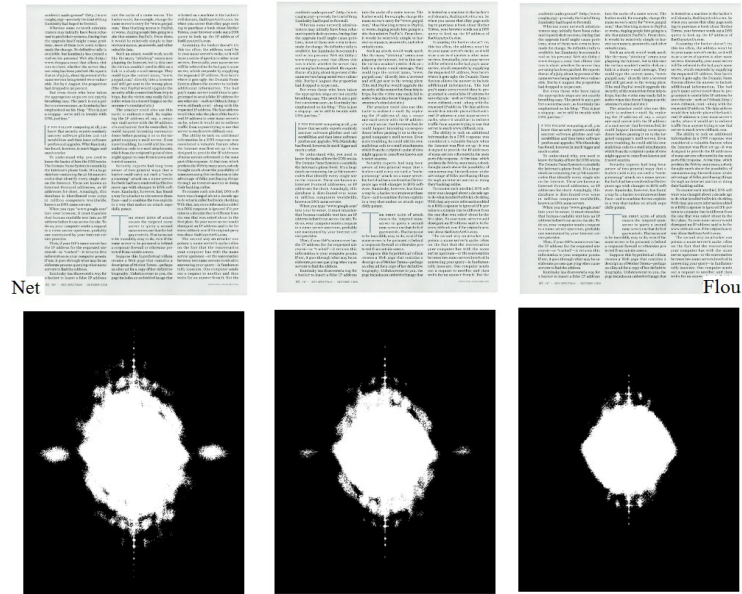


FIGURE 3.21: Exemple d'évolution du flou sur une même image et les binarisations dans l'espace de Fourier. Nous constatons que les hautes fréquences, au centre de l'image, diminuent lorsque le flou augmente.

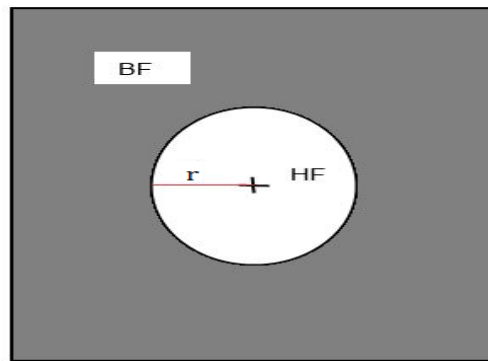


FIGURE 3.22: Binarisation dans l'espace de Fourier après filtrage

à déterminer r , c'est-à-dire le rayon du cercle inclus dans l'espace des pixels étiqueté comme hautes fréquences (blanc).

Nous calculons la distance r entre le point le plus proche des basses fréquences et le centre des hautes fréquences. Cette distance est normalisée par la taille de la diagonale de l'image. Nos expériences nous ont mis en évidence qu'un flou important sur une image composée de texte faisait perdre une trop grande quantité de hautes fréquences. La binarisation n'a alors plus de sens.

Nous obtenons la formule suivante :

$$D_5 = r \tag{3.23}$$

3.4.3 Résultats et analyses

Notre objectif est de mesurer la pertinence de chacun des descripteurs. Pour analyser ces descripteurs, nous observerons, dans un premier temps, l'évolution de la courbe des descripteurs en fonction de la puissance d'un flou gaussien. Le flou est représenté par l'écart-type σ de la gaussienne. Dans un second temps, nous mesurerons les erreurs d'estimations sur un échantillon d'images où la corrélation entre la valeur du descripteur et la valeur de l'écart-type est faite par une régression polynomiale. Cette étude permettra d'évaluer la pertinence des ces descripteurs dans le contexte du livre ancien.

A partir d'un ensemble composé de 100 images de même résolution (300 dpi), nous construisons une base de donnée d'images floues. Cette base de donnée est obtenue en appliquant un filtre gaussien dont nous faisons varier σ_{th} de 0 (pour l'image nette), à 1.8 par pas de 0.2. Une valeur de 1.8 rend l'image inexploitable.

Dans le but de définir l'évolution du descripteur en fonction du flou, nous calculons les descripteurs pour toutes les images. Nous moyennons chacune des valeurs obtenues pour chaque valeur σ . La figure 3.23 présente les résultats.

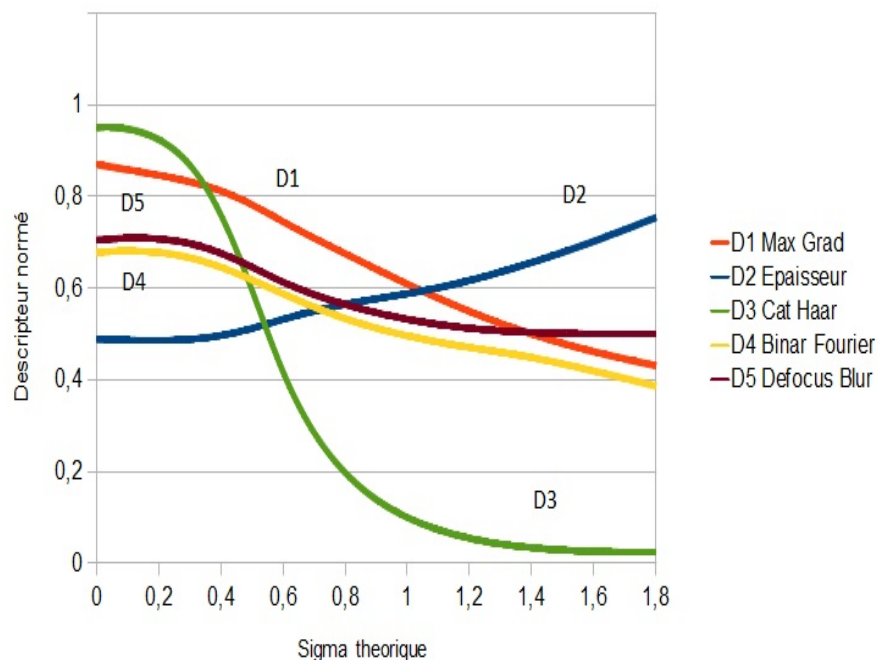


FIGURE 3.23: Représentation des valeurs normalisées des descripteurs en fonction de σ_{th}

Le descripteur idéal est, a priori, une droite passant par les valeurs extremums (1 et 0) pour les bornes de l'intervalle de σ . Une telle fonction montrerait que la fonction minimise au mieux les erreurs d'estimations, ainsi le descripteur serait de la forme $0 + 0,5\sigma_{th}$ ou

$1 - 0.5\sigma_{th}$. Aucun descripteur ne suit une telle fonction, que se soit dans l'intervalle de définition de la fonction que la valeur du coefficient directeur. Toutefois, nous souhaitons avoir une distinction notable pour les valeurs 0.6 et 1.2 puisque la séparation entre une image nette et une image floue se situe dans cet intervalle (cf. figure 3.24).

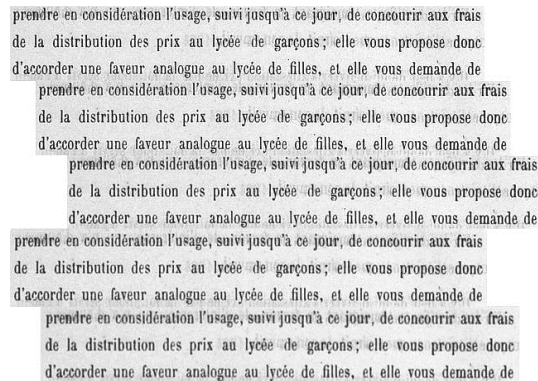


FIGURE 3.24: Représentation de 5 mêmes régions d'une image lissée par un filtre gaussienne de noyau 7×7 et de σ variant d'un pas de 0.4. Le morceau d'image du haut est supposée nette, la troisième image correspondant à un filtrage de σ de 0.8 et elle est considérée comme floue par une organisation telle que la Bibliothèque Nationale de France, tandis qu'une personne non spécialisée dans le domaine va davantage considérer qu'une image est floue pour un σ allant de 0.8 et 1.2 soit entre la troisième et quatrième image.

D'un point de vue général, les descripteurs n'arrivent pas à récupérer des informations permettant de différencier les σ inférieures à 0.4, les courbes sont relativement constantes pour ces valeurs. Ensuite, l'évolution des courbes sont toutes différentes, elles ne sont pas linéaires et aucunes n'ont un coefficient directeur proche de 0.5 ou -0.5 . Les courbes de D1, D2, et D4 semblent avoir un comportement linéaire sur l'intervalle $[0.4, 1.8]$, tandis que D5 n'est linéaire que sur l'intervalle $[0.4, 1]$, puis la courbe devient constante. La courbe D3, la catégorisation de Haar, a une forte décroissance sur l'intervalle $[0.4, 0.8]$ et devient invariante sur le reste de l'intervalle.

Si l'on se focalise sur la zone $[0.6, 1.2]$, nous nous apercevons que la courbe du gradient max a le coefficient directeur le plus discriminant (coefficient directeur de 0.3), tandis que le Defocus Blur a le plus faible (coefficient directeur de 0.1).

Les comportements de chaque descripteur sont ensuite modélisés par des régression de polynômes de degré 3. Nous estimons alors un σ en fonction de la donnée de chaque descripteur. Les erreurs d'estimation sont présentées dans la figure 3.25.

Les courbes oscillent entre 0.2 et 0.6 sur notre intervalle d'intérêt. Les images prises sont de sources différentes étant donné la position du problème (qualité du livre et machine variable mais même résolution pour les tests) ce qui peut expliquer le manque de précision des descripteurs pour les valeurs σ_{th} inférieure à 0.4. En soit, une erreur de 0.4 pour un σ_{th} nul est acceptable.

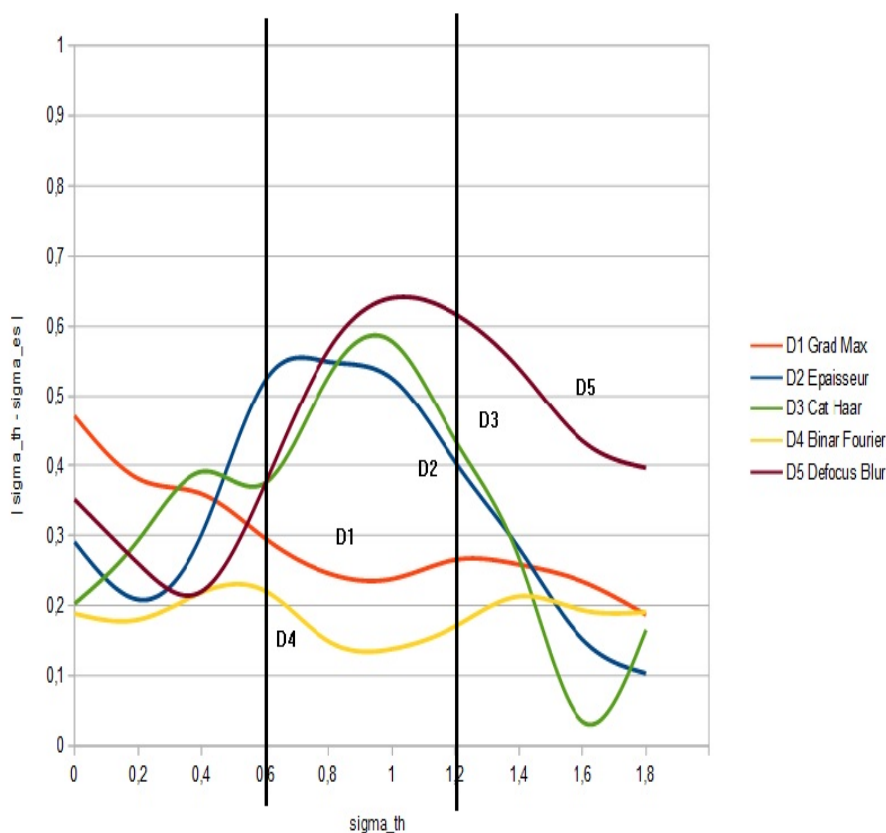


FIGURE 3.25: Représentation des erreurs d'estimation de σ en fonction du σ théorique

Nous remarquons que deux descripteurs semblent cohérent : D1 le gradient max dont la moyenne est aux alentours de 0.3 et D4 le descripteur basé sur la transformée de Fourier avec une erreur moyenne de 0.2.

Les trois autres courbes ont des erreurs qui augmentent pour atteindre un maximum d'environ 0.55 pour un σ_{th} de 1. Aucun de ces estimateurs ne peuvent être considérés comme bon. Les descripteurs D2 et D5 ont un comportement similaire à la courbe de la figure 3.23 lorsque nous analysons différents flous à partir d'une même image, l'évolution de la valeur du descripteur est correcte, c'est-à-dire qu'il augmente lorsque le flou augmente. Mais la valeur initiale est trop variable d'une image à une autre. Nous avons un problème de constance des valeurs de référence ici.

La courbe D3 est trop discriminante et la fonction semble être une fonction par morceau. Cette fonction se décompose entre trois parties, une première constante, une seconde linéaire et une troisième qui est de nouveau constante. Le deuxième morceau de la fonction est la fonction qui mesure le passage d'une image nette à une image floue, et ce morceau mesure la puissance du flou. Le problème que pose le descripteur est que cette fonction a un coefficient directeur très important. Par conséquent, l'estimation de σ n'est

pas possible, sa courbe d'erreur nous confirme bien que le descripteur passe de net à flou trop rapidement.

Procédons maintenant à une analyse de tous les descripteurs, en se basant sur les différentes méthodes, hypothèses et résultats.

La courbe D4, qui est la représentation de la binarisation de Fourier est la plus précise. L'erreur qui apparaît est obtenue lorsque la numérisation contient des images et lorsqu'il y a peu d'information dans le document. La courbe D1, le maximum du gradient, obtient des résultats pertinents, il est le descripteur le plus comparable aux méthodes d'autofocus des caméras. Mais, la motivation de la problématique permet à la fois de valider la pertinence de l'autofocus mais aussi de corriger lors de ces défaillances, par conséquent, on considère que ce descripteur ne peut être suffisant tout seul.

L'épaisseur de contour a un intervalle de définition trop petit, on peut voir sur la courbe qu'un contour net est d'environ 5 pixels, tandis qu'un contour flou de $\sigma_{th} = 1$ est d'environ 6 pixels. L'erreur d'estimation de l'épaisseur des contours engendre une erreur trop importante. Le modèle basé sur la catégorisation de Haar est trop discriminante, nous obtenons un bon estimateur net/flou, mais la quantification du flou est quasi impossible. L'approche "Defocus Blur" induit des difficultés pour estimer la puissance du flou. Il est capable de déterminer si une image est plus ou moins floue par rapport à une autre image, mais lorsque deux images sont relativement flou, le descripteur se trompe davantage.

Notre objectif est de fusionner les informations de tous les descripteurs par un modèle d'apprentissage supervisé. Nous allons donc évaluer cette méthode et estimer la pertinence de chacun des descripteurs dans ce modèle.

3.4.4 Apprentissage supervisé

Nous avons pu voir dans la sous-section 3.4.1 que les auteurs établissent parfois des règles en fonction de certains critères pour affiner la précision de la qualité de netteté. Nous suggérons ici de faire de l'apprentissage supervisé.

Notre base d'entraînement est composée de l'ensemble des images utilisées dans la section précédente, c'est-à-dire 100 images nettes de document numérisé. Nous utilisons un flou gaussien d'écart-type σ pour déterminer 9 autres classes en faisant varier σ de 0.2 à 1.8 par pas de 0.2. Nous considérons que les images d'origines ont σ égale à 0. Toutes les images sont définies par cinq descripteurs. Et l'image est étiquetée par sa valeur σ . Notre base de test est construite en prenant 20 images nettes de document numérisé différentes de la base de donnée. Toutes les images sont floutées de la même manière. Nous obtenons ainsi 200 images à prédire.

Nous proposons une visualisation de trois descripteurs à la figure 3.26. Cette visualisation confirme notre approche, nous distinguons une séparation d'information. Nous observons qu'il y a une certaine évolution globale qui nous permet de distinguer la valeur sigma, même si les clusters n'ont pas la même taille et que le pas choisi est assez grand (0.6 au lieu de 0.2)

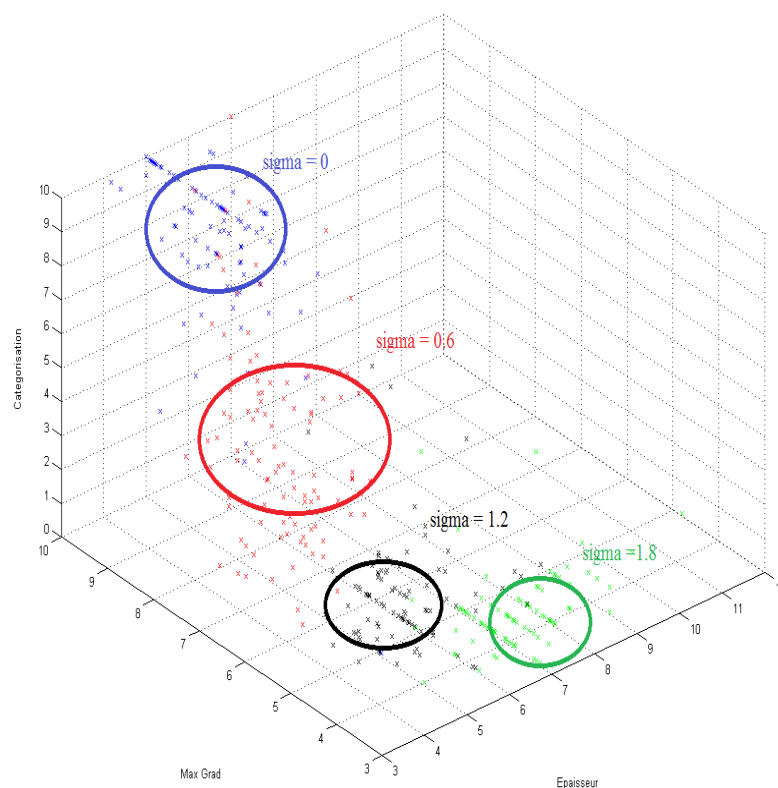


FIGURE 3.26: Représentation de trois descripteurs et la formation de cluster pouvant permettre l'estimation de σ

Nous procédons à l'estimation des images provenant de notre base de test. Puis, nous moyennons les valeurs d'estimations pour chaque force du flou. La figure 3.27 montre les erreurs moyennes de prédiction de cette base de test faite par la méthode des k-plus proches voisins (ici, $k = 20$), et par machine à vecteurs de support (SVM) où l'hyperplan est déterminé par un polynôme d'ordre 2.

Les résultats d'apprentissages sont nettement meilleurs que les résultats précédents. L'erreur moyenne est définie entre 0.1 et 0.2 sur l'ensemble de l'intervalle. L'erreur est plus faible pour la méthode des k-nn comparée au SVM. Le modèle est basé sur un pas de sigma de 0.2. Notre estimation est inférieure à ce pas, elle est donc très satisfaisante.

L'approche des k plus proches voisins semblent effectivement la plus pertinente car, comme nous l'avons vu dans la figure 3.26, les clusters se fusionnent et la séparation des

plans peut être imprécise. L'usage du vote majoritaire peut alors rectifier ce type d'approximation dans notre cas. Toutefois, dans les deux approches, nous constatons que l'erreur moyenne est de 0.15, ce qui est inférieure à notre pas de la puissance du flou.

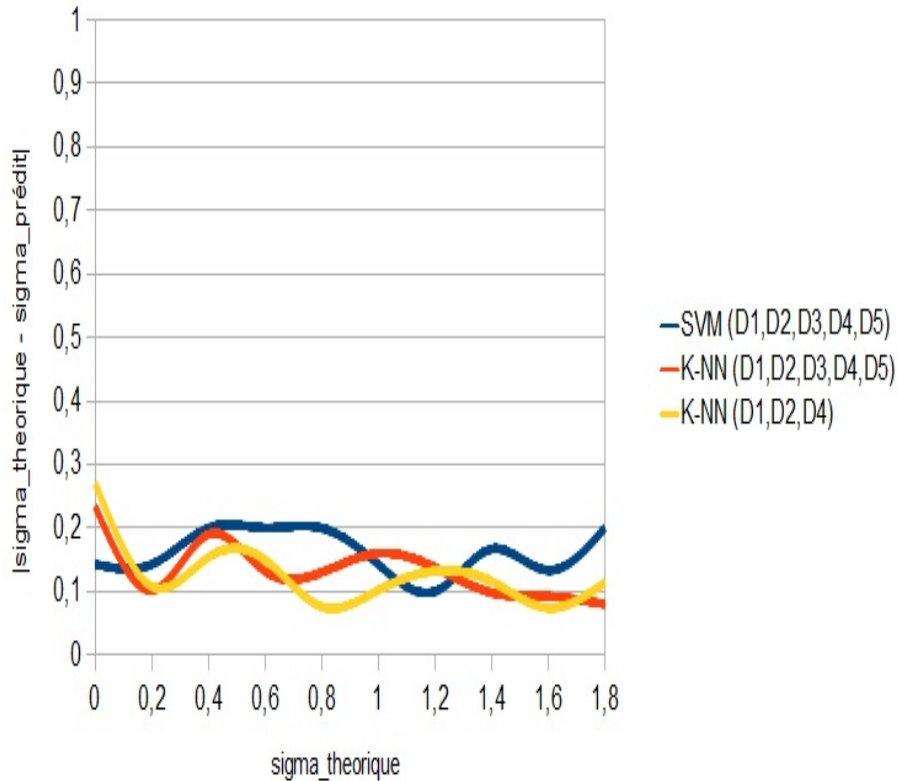


FIGURE 3.27: Représentation des erreurs d'estimation de σ obtenu par apprentissage supervisé en fonction du σ théorique

Nous proposons ensuite d'analyser la contribution de chacun des descripteurs dans l'estimation de sigma. Effectivement, nous avons vu à la figure 3.26 qu'il y a trois descripteurs "mauvais" (defocus blur estimation, épaisseur des contours, catégorisation de Haar) et deux "bons" (binarisation de Fourier, gradient max). Il nous paraît pertinent de connaître la contribution de chacun des descripteurs pour déterminer une estimation.

Pour effectuer cette opération, nous utilisons le modèle des moindres carrés pour résoudre le système $Ax = b$, avec A la matrice $N \times 5$, qui représente l'ensemble des descripteurs (5) de toutes les images (N), et b le vecteur de taille N , qui représente les σ_{th} de chacune des images. Ce modèle nous donne les résultats obtenus dans le tableau suivant :

D1 Max Grad	D2 Épaisseur	D3 Haar	D4 Fourier	D5 Defocus blur
0.27	0.19	-0.04	0.60	-0.01

A partir de ce résultat, on observe que deux descripteurs (D3 et D5) sont négligés. Il est vrai que la catégorisation de Haar et le Defocus Blur estimation ne sont pas les plus pertinents comme nous l'avons vu dans la section précédente. A l'inverse, le descripteur basé sur la binarisation de Fourier a une contribution importante pour aboutir à l'estimateur final.

Ce modèle régressif met en évidence que l'apprentissage est plus optimal en utilisant uniquement D1 (le gradient max), D2 (l'épaisseur des contours) et D4 (la binarisation de Fourier). La courbe K-NN(D1,D2,D4) montre de meilleur résultat que K-NN(D1,D2,D3,D4,D5) (cf. figure 3.27). La moyenne d'erreur est de 0.1 dans l'intervalle [0.6, 1.2].

3.4.5 Résultats expérimentaux

Dans cette partie, nous allons créer du flou à partir de la caméra et utiliser une mire pour le quantifier. Cette démarche va nous permettre de valider d'une part, que notre estimation est toujours robuste dans les cas pratique, et d'autre part, de faire une corrélation entre la FTM (mesure objective de netteté faite sur une mire) et notre estimateur.

Nous avons vu, dans la sous-section 3.2, l'ensemble des éléments permettant d'influencer la netteté de l'image. Nous rappelons que cet ensemble crée un équilibre permettant la formation d'une image nette, et que si un élément varie, nous allons créer un flou. Nous considérons quatre éléments principaux : le nombre d'ouverture, la lentille, la distance de mise au point et la focale (la distance lentille/capteur). Par soucis de simplicité technique, nous allons uniquement modifier la distance de mise au point et la distance focale.

Afin d'obtenir une vérité terrain, nous avons systématiquement mis une mire pouvant calculer la FTM (cf. la section 3.3) lors de la numérisation, comme on peut le voir dans l'image 3.28.

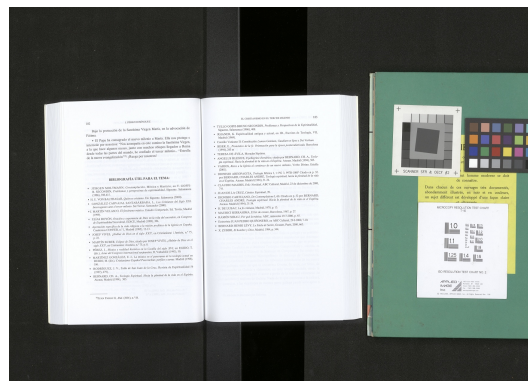


FIGURE 3.28: Exemple d'image d'acquisition. Dans la partie droite de l'image, nous avons mis une mire FTM, une mire couleur et une mire de résolution.

TABLE 3.1: Tableau récapitulatif des calculs des FTM en fonction de la modification des distance de mise au point

mm	0	2	4	6	8	10	12	14	16	18
FTM	100	100	100	99	97	85.5	77.5	69	63.5	59

TABLE 3.2: Tableau récapitulatif des calculs des FTM en fonction de la modification de la focale

mm	30	35	40	45	50	55	60
FTM	20	23	47.5	100	34.5	19.5	15

Le protocole de test a consisté à régler l'appareil selon la norme FADGI. Les différents réglages sont nos paramètres de références. Pour créer une base de donnée, nous allons faire varier uniquement la distance de mise au point dans un premier temps, puis dans un second temps la distance focale.

L'acquisition va se faire sur deux livres différents où nous prenons cinq double-pages de document ancien. Nous allons ainsi faire 10 numérisations pour chaque distance de mise au point et de distance focale. La distance de mise au point varie de $0mm$ à $18mm$ par un pas de $2mm$, 0 correspond à la position initiale de la caméra lors du réglage et nous diminuons la distance caméra/objet. La distance focale varie de $3mm$ à $6mm$ pour un pas de $0.5mm$, où la distance de $4.5mm$ est la position de calibration. Nous avons représenté l'évaluation de la FTM au point 10% dans les tableaux 3.1 et 3.2.

Nous remarquons dans ces tableaux que la FTM est de 100% pour le cas d'une modification de la distance de mise au point de $0mm$, et une FTM de 100% pour le cas d'une distance focale de $4.5mm$. Nous observons à l'aide de ces tableaux que la puissance du flou n'est pas la même entre les deux, effectivement la FTM varie doucement dans le cas de la distance de mise au point tandis qu'elle varie fortement dans le cas de la focale et ce dernier est symétrique.

Nous rajoutons que la FADGI a donné des références de notation comme le montre la figure 3.29. De plus, nous pouvons estimer que pour une résolution de 300 dpi, une image ayant un score de 80% est nette pour l'œil humain.

Maintenant, observons les courbes obtenues à partir des 200 numérisations que l'on retrouve dans les figures 3.31 et 3.30, nous séparons les images dont le flou a été créée par une modification de la distance de mise au point et une modification de la focale et nous avons ajouté la mesure de la FTM.

Nous constatons que l'évolution des descripteurs ont dans l'ensemble un comportement analogue au cas du flou gaussien.

Performance level	Sampling efficiency
****	> 90 %
***	> 85 %
**	> 75 %
*	< 75 %

FIGURE 3.29: Évaluation de la FTM selon la FADGI

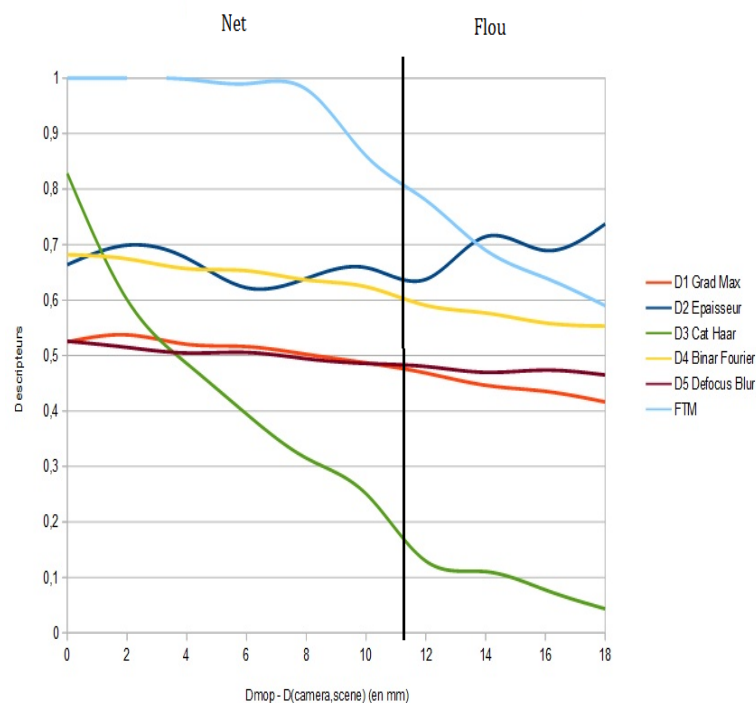


FIGURE 3.30: Évolution des descripteurs en fonction de la modification de la distance de mise au point

Dans le premier graphe, nous avons délimité la séparation net/flou quand la FTM est égale à 0.8. Nous observons aussi que la courbe de la FTM commence à décroître pour une distance de 8mm, alors que les descripteurs varient bien avant cela. Nous avons les trois courbes D1, D4 et D5 qui sont relativement similaire, même si D5 paraît constante. À l'inverse les deux méthodes basées sur les contours ont un comportement bien différents de ce que nous avons observé pour les cas de flou artificiel. L'épaisseur des contours a une courbe ondulatoire et la catégorisation de haar ne semble pas être par morceau.

Pour les résultats de la focale (cf. 3.31), la FTM varie fortement et nous avons délimité les parties de la FTM considérant l'image nette ou floue. Nous avons un extremum au point

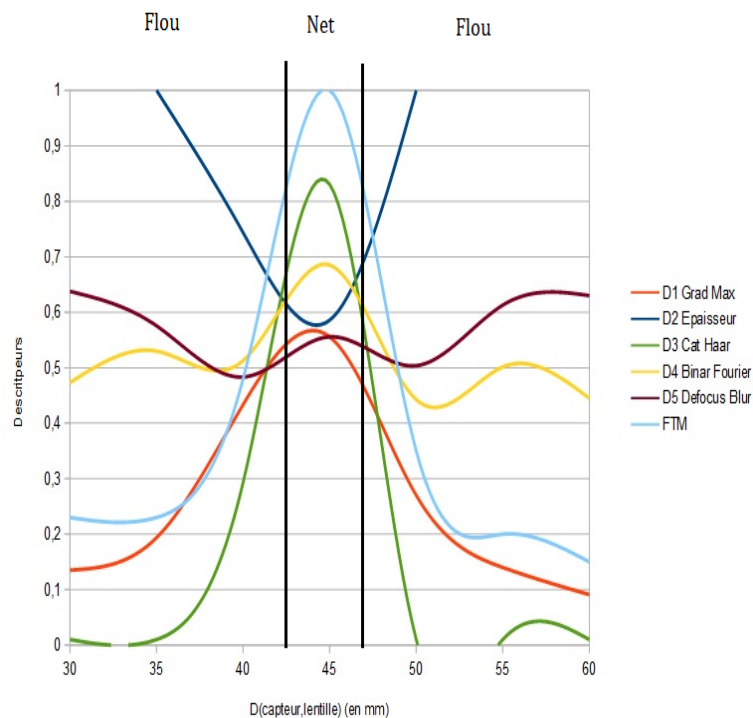


FIGURE 3.31: Évolution des descripteurs en fonction de la modification de la focale

net puis les courbes divergent de manière symétrique. Le comportement des descripteurs sont dans l'ensemble très bon et sont similaire à ce que nous avons vu avec le cas du flou gaussien. Nous observons que lorsque le flou est important les descripteurs $D4$ et $D5$ sont aberrants.

Pour en revenir sur le comportement des descripteurs de contour du premier graphe, la catégorisation de Haar découpe la séparation net/flou de l'erreur focale à 0.65 alors que dans le cas de l'erreur de mise au point l'erreur focale se situe à 0.15 et l'épaisseur de contour n'est pas du tout exploitable concernant l'erreur de mise au point alors qu'il est bon lors de la modification de la focale.

Dans la sous-section précédente, nous avons établi que les descripteurs Maximum du gradient ($D1$), l'analyse des épaisseurs de contours ($D2$) et la méthode basée sur la binarisation de Fourier ($D4$) proposaient la meilleure combinaison de prédiction de donnée, or la remarque précédente remet en question l'usage de l'analyse des épaisseurs de contours.

Nous constatons que les résultats sont toujours très moyen pour la méthode basée sur le defocus blur estimation ($D5$).

Nous voyons aussi que la catégorisation de Haar est pertinente ($D3$). Alors que dans les images lissées par un filtre gaussienne, nous obtenons des résultats trop discriminant, ici,

nous constatons une pente beaucoup plus faible. Au vu des résultats, il semble pertinent de prendre en référence ce dernier descripteur.

Afin d'être le plus robuste possible, l'utilisation de tous les descripteurs à l'exception du defocus blur estimation est le plus pertinent pour estimer la netteté de l'image. Nous avons donc utilisé notre estimateur KNN et nous obtenons les courbes moyennes dans les figures 3.33 et 3.32. Nous avons représenté dans ces courbes les valeurs normalisées, ainsi la courbe KNN donne un résultat entre 0 et 1 plutôt qu'entre 0 et 1.8. Nous avons rajouté dans ces courbes l'interprétation que nous faisons en fonction de la valeur obtenue. Nous rappelons que nous n'avons pas défini de valeur limite permettant de séparer une image nette d'une image floue, mais que cette valeur se situe entre 0.6 et 1.2 selon l'utilisateur.

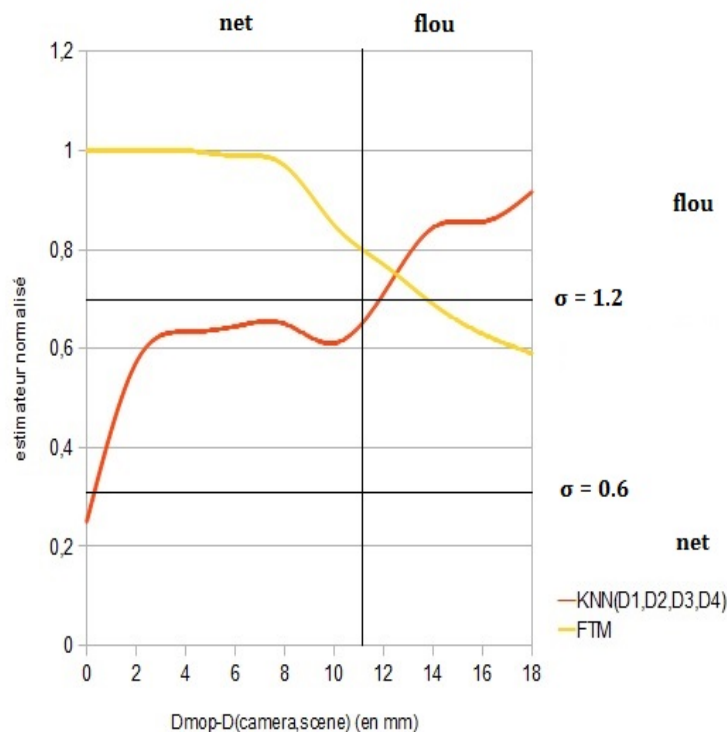


FIGURE 3.32: Évolution de la FTM et de notre estimateur en fonction de la modification de la distance de mise au point

Dans ce premier graphique, nous pouvons voir que le point initial est très net, puis la courbe d'estimation monte très vite vers 0.6 (soit $\sigma = 1.08$), elle reste constante puis elle redescend légèrement pour ensuite diverger totalement. Mise à part le point initial la courbe semble être constante tout comme la courbe de la FTM. Par contre, nous avons une image plus nette aux alentours de 10mm qui semble être un point particulier puisqu'il correspond au seuil des 80% de la FTM.

Le comportement du descripteur K-nn est bon puisqu'il a un comportement assez proche de la FTM.

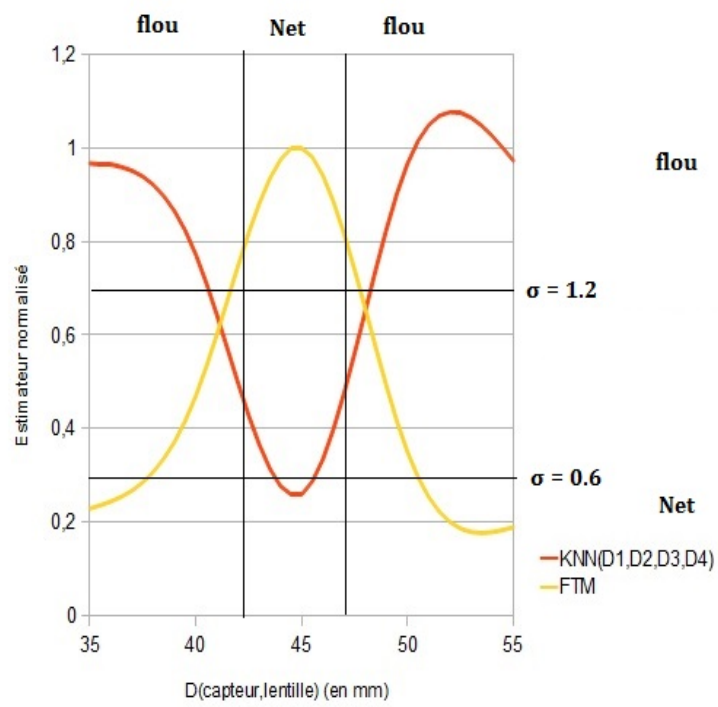


FIGURE 3.33: Évolution de la FTM et de notre estimateur en fonction de la modification de la focale

3.5 Conclusion

Dans ce chapitre, nous avons dressé l'ensemble des éléments permettant d'obtenir l'image la plus nette possible que cela soit d'un point de vue matériel ou des mécanismes automatique de calibration, nous avons évoqué les moyens de quantifier la représentation d'un signal dans le cadre de la numérisation de livre ancien. Notre problématique consistait à évaluer la netteté d'une image de document ancien sans référence. Pour cela nous avons proposé d'étudier différentes approches pour analyser le nombre important de contours présents dans le texte et gravure, caractérisant les livres anciens.

Nous avons choisi de caractériser ces informations par :

1. un opérateur simple tel que le gradient
2. une analyse locale tel que l'épaisseur des contours
3. une évaluation de la dégradation locale sur les contours à l'aide de la méthode de defocus blur estimation
4. une catégorisation des contours dans l'espace des ondelettes de Haar
5. une analyse des représentations fréquentielles par une binarisation dans l'espace de Fourier.

Nous avons pu constater des similitudes entre les caractérisations, mais la précision n'est pas pertinente à l'exception de la binarisation de Fourier et le maximum du gradient. Au vu de l'ensemble des descripteurs, nous nous sommes demandés si il était possible de combiner les différentes informations afin d'améliorer encore la précision de notre estimateur. La solution proposée dans ce sujet est d'utiliser un apprentissage supervisé en se basant sur les différentes catégorisations comme descripteurs globaux de qualité. Le résultat est très pertinent puisque nous arrivons à obtenir une erreur inférieure au pas du modèle.

Nous proposons ensuite une étude sur des cas réels qui met en évidence que les descripteurs de netteté ont un comportement applicable aux cas pratique. Nos résultats montrent que notre descripteur a un comportement analogue à la FTM, mesure normée du flou à partir d'une mire de référence. Cette corrélation démontre que notre estimateur est applicable dans les cas pratique et qu'elle est pertinente.

La principale perspective que nous envisageons est de faire une chaîne combinant l'analyse et la correction de l'image. L'estimation de la force du flou permet de déterminer un traitement d'image adéquate pour corriger ce défaut. En fonction de la force du flou, nous pouvons déterminer si il faut utiliser une méthode de déconvolution (prenant en paramètre la force) ou de reconstruction, même si pour ce dernier cas, il paraît plus logique de numériser à nouveau une image qui a perdu trop d'information.

L'autre perspective est de continuer à développer chacune des deux étapes, c'est-à-dire

d'évaluer de nouveaux descripteurs et d'adapter la méthode d'apprentissage supervisé en fonction de leur pertinence.

Chapitre 4

La calibration couleur

4.1 Introduction

Le produit que propose la société i2S à travers leur scanner de livre ancien est le premier maillon de la chaîne numérique. En effet, l'acquisition photographique crée une image numérique. Cette image est ensuite envoyée vers n'importe quel autre appareil pour être visualiser comme un écran d'ordinateur ou encore pour être imprimé. L'image est une information qui est doit être visualisable quelques soient le support que nous utilisons. Tous les appareils d'acquisition ou d'affichage ont des caractéristiques différentes (sensibilité, précision, ...), par conséquent les couleurs sont très différentes d'un appareil à l'autre. La gestion de la couleur a pour objectif de restituer au mieux la couleur d'origine permettant de passer d'un appareil à un autre. Ce travail a pour objectif de transférer les informations couleurs pour que l'objet que nous voyons puisse être identique sur n'importe quel écran et sur n'importe quelle impression.

Les conversions colorimétriques d'un appareil à un autre sont faites à partir d'un réglage spécifique de l'appareil que l'on appelle la calibration couleur. Ce réglage se base sur la conversion de l'appareil et d'une espace de couleur de référence. Cet espace de couleur va permettre de faire la transition colorimétrique entre l'appareil de création de l'image et l'appareil de visualisation ou d'impression de l'image.

Dans ce chapitre, nous nous focaliserons sur la calibration couleur dans le cadre de la numérisation. Nous rappelons que le processus de numérisation doit s'effectuer dans des conditions d'éclairage maîtrisées c'est-à-dire que l'éclairage doit être fixe et homogène sur l'ensemble de la zone à capter. La lumière est un facteur important dans la création de la couleur et elle ne doit pas varier une fois la calibration faite. De même, tous les appareils ne captent pas les mêmes informations colorimétriques à cause des fonctionnements différents des capteurs. L'information captée est ensuite encoder pour donner une

couleur en fonction de la sensibilité du capteur. Cette couleur est sujette à une erreur de mesure qui doit être corrigée par l'étape de calibration. La calibration couleur est alors faite systématiquement pour chaque scanner (et donc caméra) et pour un éclairage donné.

L'étape de calibration se fait à partir d'une mire colorimétrique dont les couleurs sont connues. L'acquisition d'une telle mire permet d'observer les aberrations colorimétriques.

Ces mires sont référencées via des organisations qui norment la couleur sur tous les appareils. L'entreprise étant fabricant de scanner est contrainte de respecter les normes où les résultats permettant d'évaluer leur matériel. Il existe deux normes : Métamorphose et FADGI. Ces deux normes coexistent, car elles s'appliquent selon la zone géographique. Nous y trouvons de nombreuses informations sur la calibration couleur dont les principaux éléments sont : la mire, les erreurs de précision tolérées, la précision des niveaux de gris.

Nous verrons dans une première partie la problématique de la gestion de la couleur plus en détail ; nous présenterons les différents espaces colorimétriques utilisés ainsi que les mesures de l'erreur colorimétrique ; nous préciserons les spécificités du contexte industriel et détaillerons les contraintes données par les normes ; nous finirons cette partie par le processus de calibration à mettre en place. Dans la deuxième partie, nous décomposerons les différentes étapes de la transformation permettant de faire la calibration, ainsi que l'importance du choix de la mire ; nous terminerons cette partie par plusieurs tests en combinant mires et transformations.

4.2 La gestion de la couleur

La gestion de la couleur s'avère être un problème global qui ne concerne pas uniquement l'affichage de l'image sur un écran, elle concerne également les appareils d'acquisition et les appareils d'impression. La figure 4.1 résume la problématique. L'image est manipulée par trois types d'appareils : les appareils d'acquisition comme les appareils photos, le scanner ; les appareils de travail comme les écrans d'ordinateurs ; et les appareils d'impression. Ces trois types d'appareils caractérisent la chaîne graphique. La gestion de la couleur s'applique pour tous les types d'appareils et son objectif est de faire que la visualisation de l'image numérique soit la même pour l'oeil.

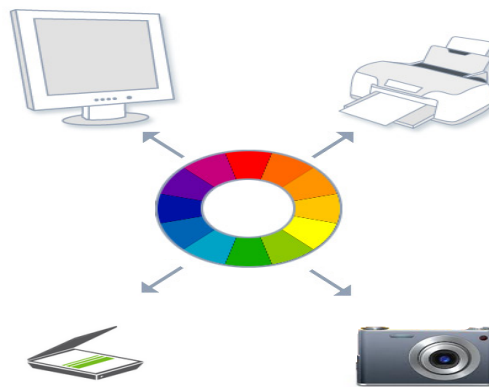


FIGURE 4.1: L'une des problématiques de la gestion de la couleur est de pouvoir communiquer la couleur entre les différents outils graphiques.

Le comité international de la couleur (International Color Consortium, ICC), constitué en 1993, est un consortium de plusieurs industriels qui a créé un système de gestion de couleur universel et invisible pour l'utilisateur.

Ainsi, un espace de couleur de référence a été choisi. Cet espace couleur est l'espace $L^*a^*b^*$. Il est l'espace permettant d'établir une relation entre tous les profils. Il est aussi appelé le PCS (Profil Connection Space). Cet espace a été choisi par sa capacité à mesurer les différences de couleurs en accord à la vision humaine.

Le comité a ainsi créé un format de fichier `.icc`, permettant de transmettre les informations couleurs dans un standard défini (cf. [ISO 15076-1 :2010, 2010]). Ce format de fichier peut être lu par n'importe quel appareil informatique, comme nous le représentons dans la figure 4.2. Il a pour but de communiquer l'information colorimétrique permettant son utilisation c'est-à-dire sa description de l'espace colorimétrique de l'appareil et sa conversion vers ou à partir de l'espace colorimétrique de référence. Ce changement d'espace couleur est généralement décrit par une table de correspondance.

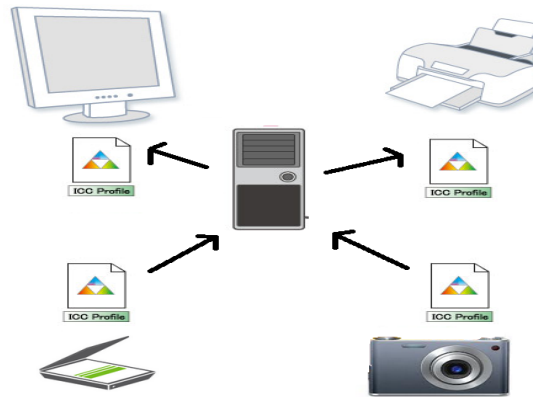


FIGURE 4.2: L'utilisation d'un fichier icc permet la conversion et il est invisible pour l'utilisateur.

Un fichier icc est un fichier regroupant différentes informations (voir la figure 4.3). Nous pouvons le découper en deux parties. La première est similaire à un fichier d'en-tête qui regroupe toutes les informations définissant la conversion. Selon le type d'appareil, l'utilisation du fichier et donc son contenu sera différent. La deuxième partie du fichier consiste à déterminer la transformation entre l'espace d'entrée et de sortie. Il peut y contenir les transformations gamma et les matrices de conversions pour les conversions dites classiques.

Le fichier icc utilise une table de conversion (LUT) dans le cas inverse. Elle définit les conversions spécifiques et détaillées. La structure du fichier permet de connaître l'espace colorimétrique d'entrée et de sortie. Par conséquent, on connaît le nombre d'éléments d'entrée et de sortie de la LUT en fonction des espaces choisis. La LUT est toutefois sous-échantillonnée pour une question d'espace et de performance, et il est standard de diviser l'espace par 33, ce qui nous donne des pas de 8 pour l'espace RVB.

Nous notons aussi qu'il peut y avoir plusieurs tables contenues dans cette partie, certains standards proposent la conversion dans les deux sens symbolisés par *AtoB* et *BtoA*, dans notre cas, *AtoB* correspond au passage de *RVB* à $L^*a^*b^*$ et *BtoA* au passage de $L^*a^*b^*$ à *RVB*.

Pour résumer, lorsque nous affichons une image sur un écran à partir d'une numérisation, l'image est convertie plusieurs fois. Elle est construite par la caméra dans un espace couleur *RVB* spécifique à la machine. Puis, l'image est convertie en $L^*a^*b^*$ selon la norme icc. Et pour finir, l'image est de nouveau convertie dans un espace spécifique pour l'affichage des écrans. Si, par exemple, l'affichage de notre écran est précis pour un format *RVB* du type *AdobeRGB*, alors nous utiliserons un second profil icc permettant de passer du $L^*a^*b^*$ au *RVB* de l'écran.

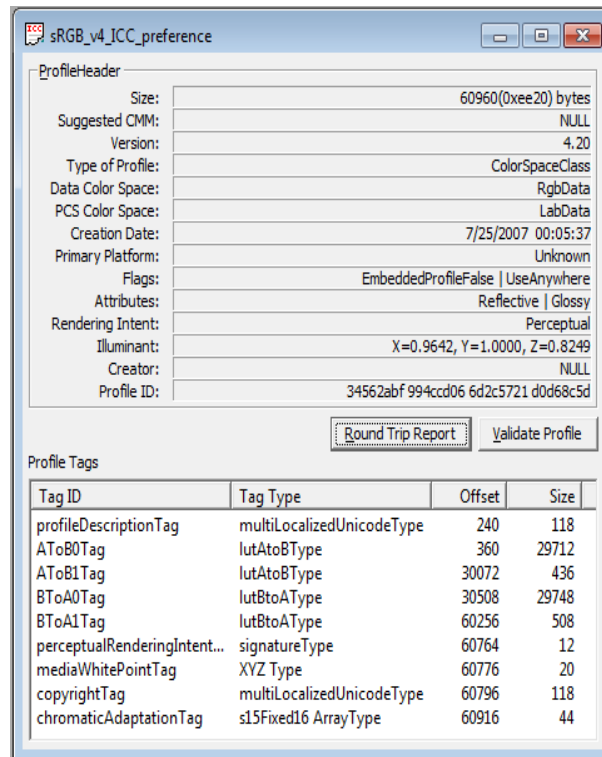


FIGURE 4.3: Un fichier icc se découpe en deux parties. La partie supérieure donne les informations d'utilisation. Principalement, le type de profile indique le sens de conversion entre le "Data Color Space" et "PCS Color Space". La deuxième partie donne les informations de conversion. Dans le fichier icc présenté, nous avons les luts de conversion explicite AtoB0 et les informations de conversion avec AtoB1 et les tags mediaWhitePointTag et chromaticAdaptationTag.

4.2.1 La gestion de la couleur des appareils d'acquisition - le scanner

Dans la chaîne graphique, le scanner fait partie des appareils d'acquisitions. Ils transforment l'information physique en information numérique. Le format utilisé est l'espace de couleur RVB . Afin de répondre aux contraintes de la norme ICC, il faut pouvoir convertir les informations en $L^*a^*b^*$. Il faut noter que les caractéristiques des espaces de couleurs standardisés ont des contraintes d'éclairage à respecter très strictes (cf. la commission international de l'éclairage (CIE)), irréalisable en pratique. Nous verrons alors la conversion d'un espace RVB à $L^*a^*b^*$.

Nous verrons aussi comment il est possible d'évaluer la distance entre deux couleurs en utilisant l'espace de couleur $L^*a^*b^*$. Le simple fait de numériser une mire de couleur dont on connaît la valeur des "patches" (région homogène de la mire représentant une couleur), nous permet d'évaluer la caméra en comparant les valeurs obtenues aux valeurs de référence. Nous séparerons les mesures des distances de la littérature scientifique et les contraintes industrielles.

4.2.2 Les différents espaces de couleur utilisés

Nous avons pu voir dans la partie précédente que la gestion de la couleur prend pour référence l'espace colorimétrique $L^*a^*b^*$, ce dernier a été choisi parce qu'il différencie les couleurs en fonction de la sensibilité de l'œil. Les espaces colorimétriques sont standardisés par la commission internationale de l'éclairage (CIE). Nous proposons de donner la transformation qui existe entre l'espace RVB et $L^*a^*b^*$. Il s'avère que la transformation de RVB en $L^*a^*b^*$ passe par un espace intermédiaire qui est l'espace XYZ . Nous allons détailler ces trois espaces et les expliquer.

4.2.2.1 CIE RVB

Le premier espace RVB créé date de 1931. Il est basé sur les lois de Grassmann qui proposent de définir un espace de couleur comme un espace vectoriel. L'ensemble des éléments de cet espace vectoriel est l'ensemble des couleurs visibles par l'homme. Cette espace est de dimension de 3. Chaque dimension détermine une couleur. Les trois couleurs sont indépendantes pour valider les propriétés d'un espace vectoriel. La loi additive va alors définir l'ensemble des couleurs possibles. Et la loi multiplicative va déterminer l'intensité lumineuse.

Le CIE RVB est basé sur les trois couleurs primaires rouge, vert et bleu. Ces trois couleurs sont indépendantes. Le triplet colorimétrique RVB est proche des cônes humains. Effectivement, les cônes humains sont sensibles à trois types de longueurs d'ondes : les faibles, les moyennes et les grandes équivalentes aux couleurs bleu, verte et rouge.

Ces trois variables définissent des coefficients permettant de créer la lumière blanche. Pour déterminer la lumière blanche, il est nécessaire, dans un premier temps, de définir le rouge, le vert et le bleu utilisés. Ces informations sont mesurables par leurs longueurs d'onde. Et dans un second temps, la pondération des coefficients de chacune des couleurs permet d'obtenir le blanc voulu. Nous pouvons alors caractériser n'importe quel espace de type RVB en fonction de ces paramètres. L'équation 4.1 illustre ce système.

$$W \equiv \alpha_W R_{[625-740]nm} + \beta_W V_{[520-565]nm} + \gamma_W B_{[450-500]nm} \quad (4.1)$$

L'espace CIE RVB 1931 est caractérisé par les valeurs données dans le tableau suivant :

Ces caractéristiques vont déterminer les paramètres de l'espace CIE RVB 1931, et lorsque nous allons définir une couleur quelconque Co , celle-ci sera donnée par un triplet de coordonnées sans unité. Nous notons que le modèle permet de retrouver la luminance d'une couleur quelconque Co par l'équation 4.2

Couleur	Rouge	Vert	Bleu	Blanc
Longueur d'onde	700.0 nm	546.1 nm	435.8 nm	Equi-énergétique
Coef. de luminance	1.00 $cd.m^{-2}$	4.59 $cd.m^{-2}$	0.06 $cd.m^{-2}$	5.65 $cd.m^{-2}$
Pourcentage	0.17697	0.81240	0.01063	100 %

TABLE 4.1: Couleurs primaires et blanc de référence du CIE RVB 1931

$$L_{\{X\}} = 0.17697R_{Co} + 0.81240V_{Co} + 0.01063B_{Co} \quad (4.2)$$

avec R_{Co} , V_{Co} et B_{Co} les coordonnées dans l'espace CIE RVB 1931 de Co . Les coefficients de luminance relative permettent de calculer la luminance pour toutes les couleurs appartenant à CIE RVB. Nous pouvons la visualiser par l'axe achromatique à la figure 4.4 représentant le cube colorimétrique.

Nous notons, d'un point de vue technique, que nous utilisons une transformation gamma pour corriger les défauts d'affichage des écrans à tube cathodique. Celle-ci est généralement de 2.2. Pour distinguer l'utilisation d'une transformation gamma dans l'espace RVB ou non, nous utilisons la notation $R'V'B'$. L'espace $R'V'B'$ est l'espace couleur de l'image qui n'a pas de transformation gamma. Si un gamma est appliqué à l'image, il faut alors l'enlever pour pouvoir convertir dans un autre format de couleur. Cela signifie que pour afficher une image, nous faisons implicitement le passage de $R_0V_0B_0$ à RVB .

Pour déterminer les couleurs possibles de CIE RVB 1931, nous pouvons définir cet ensemble par l'équation 4.3. Les coordonnées trichromatiques indiquent les proportions de chacun des primaires. Le résultat sera le plan unité que l'on peut visualiser dans la figure 4.4, ce plan s'appelle le triangle de Maxwell.

$$R_{Co} + V_{Co} + B_{Co} = 1 \quad (4.3)$$

Le triangle de Maxwell est simplifié en n'utilisant que deux coordonnées, la troisième étant déduite des deux premières. De cette manière, nous pouvons projeter le plan unité sur le plan orthonormé (OR, OG) que l'on représente dans la figure 4.5.

Nous avons rajouté les informations de longueur d'onde et l'ensemble des couleurs visibles (Spectrum locus) dans cette figure. Ce changement de repère montre que la définition de l'espace de couleur RVB est incomplet.

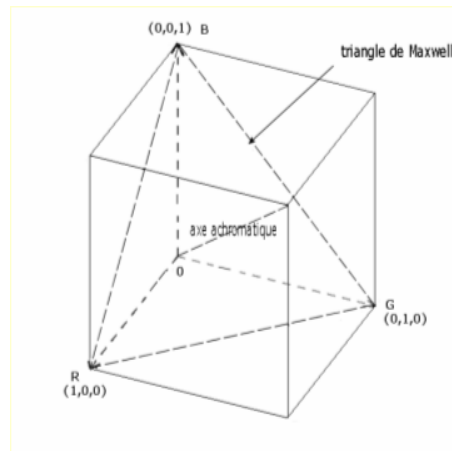


FIGURE 4.4: Triangle des couleurs dans le repère $(0, \vec{R}, \vec{V}, \vec{B})$ [pierre.nerzic, 2008].

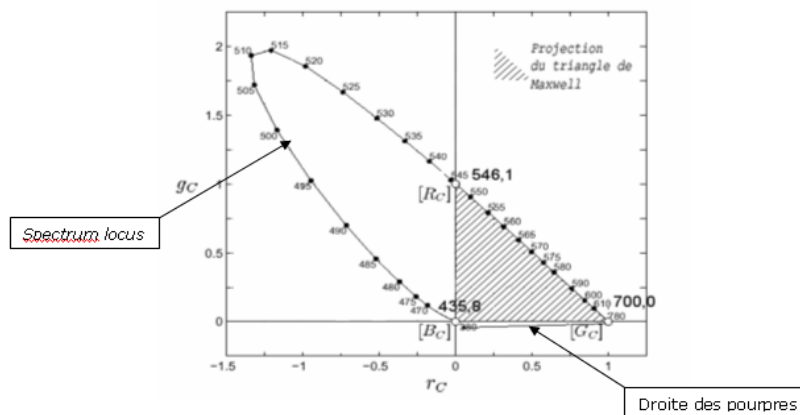


FIGURE 4.5: Diagramme Chromatique de Wright - Projection du Triangle des couleurs dans le repère (OR, OG) [pierre.nerzic, 2008].

Le Cie RVB propose un modèle spécifique pour l’œil humain permettant de jouer sur l’intensité lumineuse et les informations chromatiques. Malheureusement, le panel de couleurs de RVB est très inférieur aux couleurs visibles. Le système RVB ne remplit son rôle que partiellement.

4.2.2.2 CIE XYZ 1931

Le format CIE XYZ est ainsi une extension du format CIE RGB qui a pour vocation de décrire l’ensemble des couleurs visibles par l’œil humain. Pour comprendre la conception de ce format de couleur, nous reprendrons la figure 4.5, et nous pouvons nous rendre compte que pour représenter l’ensemble des couleurs, il faut changer le repère. Nous voyons dans cette figure que les couleurs primaires sont des représentations de longueurs d’ondes. Nous pouvons représenter un nouveau repère par un modèle physique. Nous avons les formules 4.4, 4.5 et 4.6, où $f(\lambda)$ est la densité spectrale d’une couleur et $\bar{r}, \bar{v},$

\bar{b} sont les distributions spectrales des couleurs primaires (ou fonctions colorimétriques) utilisées dans le modèle CIE RVB 1931. Ces fonctions sont définies par la Commission internationale de l'éclairage et nous les représentons dans la figure 4.6.

$$R_c = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda) \bar{r}(\lambda) \quad (4.4)$$

$$V_c = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda) \bar{v}(\lambda) \quad (4.5)$$

$$B_c = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda) \bar{b}(\lambda) \quad (4.6)$$

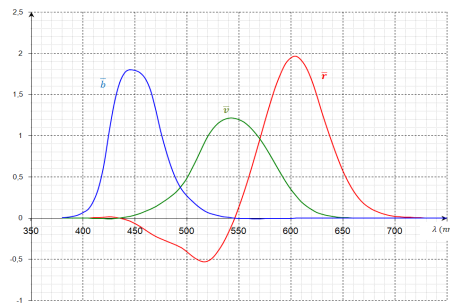


FIGURE 4.6: Représentation des Fonctions colorimétriques RVB [Wik, 2014].

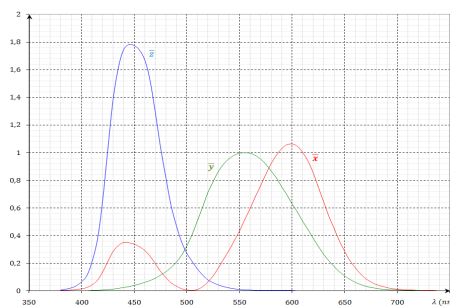


FIGURE 4.7: Représentation des fonctions colorimétriques XYZ [Wik, 2014].

L'espace CIE XYZ a la même base que l'espace CIE RVB. Les fonctions colorimétriques XYZ sont définies dans la figure 4.7 et elles ont la particularité d'être positives. Cet espace couleur est un espace trichromatique où l'on définit les couleurs en X, Y et Z de la manière suivante :

$$X_c = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda) \bar{x}(\lambda) \quad (4.7)$$

$$Y_c = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda) \bar{y}(\lambda) \quad (4.8)$$

$$Z_c = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda) \bar{z}(\lambda) \quad (4.9)$$

Le passage de l'espace CIE RVB 1931 au CIE XYZ 1931 est fait à partir d'un changement de repère, par le calcul matriciel suivant :

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} 0.49 & 0.31 & 0.2 \\ 0.17 & 0.81 & 0.01 \\ 0.00 & 0.01 & 0.99 \end{pmatrix} \cdot \begin{pmatrix} \bar{r} \\ \bar{g} \\ \bar{b} \end{pmatrix} \quad (4.10)$$

Nous avons alors un espace définissant l'ensemble des couleurs visibles par l'oeil humain comme représenté sur la figure 4.8 :

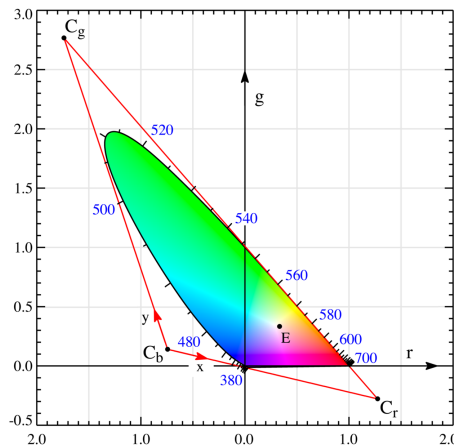


FIGURE 4.8: Diagramme des couleurs RVB et XYZ [Wik, 2014].

Nous voyons dans cette figure le repère (O, \vec{r}, \vec{g}) où il y a le triangle de Maxwell qui représente les couleurs RVB. Les points Cr , Cg et Cb représentent le nouvel espace XYZ, dans le repère (Cb, \vec{x}, \vec{y}) , où le triangle rouge montre l'espace de définition. Nous

constatons bien que l'ensemble des couleurs visibles par l'œil est inclus dans ce nouvel espace.

Le Cie XYZ est une extension de l'espace couleur Cie RVB, il définit l'ensemble des couleurs visibles en utilisant un changement de repère de RVB. Du fait qu'il englobe toute l'information colorimétrique, cet espace permet de transformer l'ensemble des couleurs en fonction du blanc de référence en faisant de nouveau un changement de repère. Toutefois, l'affichage des couleurs XYZ pose des complications sur un appareil car l'ensemble des couleurs possibles est beaucoup trop importante.

4.2.2.3 CIE $L^*a^*b^*$

L'espace des couleurs CIE $L^*a^*b^*$ est défini à partir des trois variables L^* , a^* et b^* , que l'on définit de la manière suivante :

$$L^* = 116f\left(\frac{Y}{Y_n}\right) - 16 \quad (4.11)$$

$$a^* = 500 \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right] \quad (4.12)$$

$$b^* = 200 \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right] \quad (4.13)$$

avec

$$f(\alpha) = \begin{cases} \alpha^{\frac{1}{3}} & \alpha \geq 0.008856 \\ 7.787\alpha + \frac{16}{116} & , \text{sinon} \end{cases} \quad (4.14)$$

Les valeurs trichromatiques X_n , Y_n et Z_n sont les valeurs définissant le blanc de référence. Pour donner un exemple, nous avons le blanc de l'illuminant standard D_{50} qui a pour valeur les résultats des calculs suivants :

$$X_n = \int_{\lambda_{min}}^{\lambda_{max}} l_{D_{50}}(\lambda) \bar{x}(\lambda) = 96.42 \quad (4.15)$$

$$Y_n = \int_{\lambda_{min}}^{\lambda_{max}} l_{D_{50}}(\lambda) \bar{y}(\lambda) = 100.0 \quad (4.16)$$

$$Z_n = \int_{\lambda_{min}}^{\lambda_{max}} l_{D_{50}}(\lambda) \bar{z}(\lambda) = 82.49 \quad (4.17)$$

L^* représente la luminance, qui est un indice de luminosité allant de 0 à 100, a^* représente la composante chromatique rouge-vert et b^* représente la composante chromatique jaune-bleu. Nous pouvons représenter ces trois valeurs dans la figure suivante :

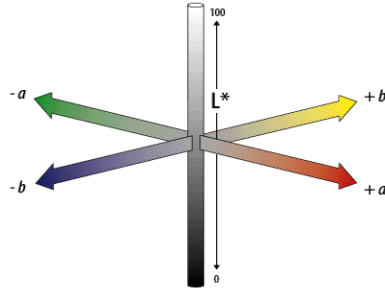


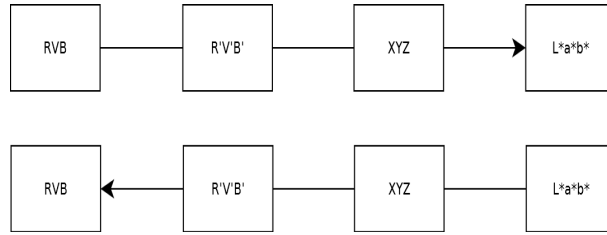
FIGURE 4.9: Représentation graphique de l'espace $L^*a^*b^*$ [Wik, 2014]

Dans cet espace, un point peut être également repéré par ses coordonnées cylindriques L^*, C^*, H^* où L^* est synonyme de luminosité (ou luminance), C^* représente la saturation et H^* représente la teinte. Ces deux variables se calculent à l'aide des formules suivantes :

$$C_{ab}^* = \sqrt{a^{*2} + b^{*2}} \quad (4.18)$$

$$H_{ab}^* = \arctan\left(\frac{b^*}{a^*}\right) \quad (4.19)$$

L'espace $L^*a^*b^*$ est donc construit selon deux axes : l'un la luminance et l'autre l'information chromatique. Cet archétype de représentation se veut comparable à la vision humaine. Effectivement, l'information sensorielle est acquise par la rétine composée de cellules photo-sensibles : les bâtonnets (s'adaptent à l'information de luminosité) et les cônes (différencient les couleurs). Par conséquent, la différence de deux couleurs dans cet espace est pertinente car il se calque sur le modèle humaine. Cela signifie que si deux couleurs sont éloignées dans l'espace, elles le sont aussi pour l'œil humain.

FIGURE 4.10: Conversion d'une image en RGB en $L^*a^*b^*$.

Les trois espaces couleurs que nous avons cités sont les principaux utilisés dans la gestion de la couleur. Nous pouvons résumer la conversion d'une image par le schéma 4.10. Nous rappelons que l'acquisition de l'image est faite dans un espace RGB et qu'elle doit être convertie dans l'espace de connexion ($L^*a^*b^*$). Nous avons pu voir que le modèle RGB est incomplet mais, en pratique, il reste l'espace de couleur de référence pour l'affichage sur un écran.

Ils existent plusieurs définitions permettant de déterminer la distance de deux couleurs correspondant à la vision humaine que nous développons dans la sous-section suivante.

4.2.3 La différence de deux couleurs : ΔE

La différence de deux couleurs permet de mesurer la pertinence du calibrage. En effet, la calibration couleur consiste à faire une transformation de RGB vers $L^*a^*b^*$. La mire colorimétrique est en $L^*a^*b^*$. Par conséquent, nous pouvons comparer la valeur théorique de la mesure. La comparaison se fait à partir du calcul de distance dans $L^*a^*b^*$ qui permet d'interpréter la différence visuelle entre les deux couleurs : le ΔE . Le calcul du ΔE a plusieurs définitions validées par le CIE (Commission international de l'éclairage) que nous énumérons en suivant.

La première est la distance euclidienne dans l'espace $L^*a^*b^*$, elle a été utilisée en 1976 par le CIE comme étant la référence de la différence de couleur. L'utilisation intuitive de la différence des trois informations colorimétriques semble être la plus pertinente.

Soit deux couleurs définies par $Co_1 = [L_1^*, a_1^*, b_1^*]$ et $Co_2 = [L_2^*, a_2^*, b_2^*]$. Soit,

$$\Delta L = L_1^* - L_2^* \quad (4.20)$$

$$\Delta a = a_1^* - a_2^* \quad (4.21)$$

$$\Delta b = b_1^* - b_2^* \quad (4.22)$$

Et donc, nous avons :

$$\Delta E76(C_{o1}, C_{o2}) = \sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2} \quad (4.23)$$

Le calcul du ΔE CMC $l : c$ (Color Measurement Committee lightness : chroma) est un calcul proposé par Clarke & al dans l'article [Clarke et al., 1984] en 1984. Il a la spécificité de prendre en compte les paramètres $l : c$ qui sont la luminance et la chrominance. Soit,

$$\Delta C = C_{ab2}^* - C_{ab1}^* \quad (4.24)$$

$$\Delta H = H_{ab2}^* - H_{ab1}^* \quad (4.25)$$

La mesure est définie à l'équation suivante :

$$\Delta E_{CMC}(C_{o1}, C_{o2}) = \sqrt{\left(\frac{\Delta L}{l.S_{L-CMC}}\right)^2 + \left(\frac{\Delta C}{c.S_{C-CMC}}\right)^2 + \left(\frac{\Delta H}{S_{H-CMC}}\right)^2} \quad (4.26)$$

Voir l'annexe 6.3.1 pour la définition de : S_{L-CMC} , S_{C-CMC} et S_{H-CMC} .

Une nouvelle mesure des différences de couleurs est utilisée par le CIE [Pub, 1995], notée $\Delta E94$, elle est calculée par une moyenne des carrés pondérés de la différence de luminance, de la différence de saturation et de la différence de teinte (redéfinie par la norme). Soit,

$$\Delta H94 = \Delta a + \Delta b - \Delta C \quad (4.27)$$

Ce calcul est défini par la formule suivante :

$$\Delta E_{94}(C_{o1}, C_{o2}) = \sqrt{\left(\frac{\Delta L}{k_{L-94}S_{L-94}}\right)^2 + \left(\frac{\Delta C}{k_{C-94}S_{C-94}}\right)^2 + \left(\frac{\Delta H_{94}}{k_{H-94}S_{H-94}}\right)^2} \quad (4.28)$$

Voir l'annexe 6.3.2 pour les définitions de : k_{L-94} , S_{L-94} , k_{C-94} , S_{C-94} , k_{H-94} et S_{H-94} .

La dernière mesure acceptée par le CIE est le modèle publié dans le rapport technique [Pub, 2001], notée ΔE_{00} . Il a été commenté dans quelques publications (cf. [Melgosa, 2000] [Luo et al., 2001] [Sharma et al., 2005]) et il reste à ce jour la mesure officielle de la différence de couleur.

Pour deux couleurs, notées $C_{o1} = (L_1^*a_1^*b_1^*)$ et $C_{o2} = (L_2^*a_2^*b_2^*)$ la différence de couleur est obtenue par les calculs suivants :

$$\Delta E_{00}(C_{o1}, C_{o2}) = \sqrt{\left(\frac{\Delta L}{K_{L-00}S_{L-00}}\right)^2 + \left(\frac{\Delta C}{K_{C-00}S_{C-00}}\right)^2 + \left(\frac{\Delta H_{00}}{K_{H-00}S_{H-00}}\right)^2 + R_T \left(\frac{\Delta C}{K_{C-00}S_{C-00}}\right) \left(\frac{\Delta H_{00}}{K_{H-00}S_{H-00}}\right)} \quad (4.29)$$

Voir annexe 6.3.3 pour les définitions de : ΔH_{00} , k_{L-00} , S_{L-00} , k_{C-00} , S_{C-00} , k_{H-00} et S_{H-00} .

L'évolution du ΔE reste proche de la formule du ΔE_{76} mais le calcul applique des coefficients de pondération sur des régions de couleur en fonction de la sensibilité de l'oeil. L'interprétation de cette métrique est importante pour pouvoir la mettre en application. Il est alors indispensable d'établir des valeurs de référence pour comprendre le résultat. Des publications traitant du sujet ont établi des règles. Les premiers travaux traitent de l'interprétation du ΔE_{76} . Kang et al [Kang, 1997] suggèrent que l'oeil commence à faire la différence (just noticeable difference) pour un $\Delta E = 1$, alors que l'article [Mahy et al., 1994] suggère une valeur supérieure : 2.3. Finalement, l'évaluation des scanners, établie par Abrardo et al dans [Abrardo et al., 1996] suggèrent de classer l'erreur en plusieurs catégories.

Lorsque le ΔE_{76} est inclus dans l'intervalle :

- [0, 1] : c'est la limite de la perception.
- [1, 3] : c'est de très bonne précision.
- [3, 6] : la précision est tolérable.
- [6, ∞ [: c'est insuffisant.

La norme basée sur le ΔE_{CMC} est une exception parce que l'interprétation du résultat dépend des paramètres l et c choisis. Les coefficients usuels sont $CMC(2 : 1)$ pour définir

la tolérance, et le $CMC(1;1)$ pour la distinction des couleurs. Les calculs donnent un résultat que l'on seuille afin d'interpréter la précision des couleurs.

La norme de 94 est importante car elle reprend notamment la fabrication de la mire et les méthodes de mesures. L'article [McDonald and Smith, 1995] énumère les différents points pour la création de la mire. Il propose ainsi de mesurer avec un angle d'incident fixe (2°) et une illumination fixe (en l'occurrence $D65$ pour la norme qu'il propose et $D50$ pour les valeurs de référence actuel), la mire doit comporter un niveau de gris pour $L^* = 50$ équivalent à une valeur fixe (a priori 128) servant d'un nouveau point de repère. Les couleurs sont représentées par des "patches" qui sont des régions homogènes (carré ou rectangle) et facilement détectable.

Nous notons qu'aujourd'hui, le protocole qu'il a suggéré est toujours respecté. La mire de référence est ainsi mesurée pour un illuminant et un champ de vision fixes. Ces informations sont données par le fournisseur. Il est important de noter que les valeurs théoriques sont différentes pour une même mire en fonction de la technologie utilisée, il est donc important de systématiquement se reporter aux données du fournisseur.

Le ΔE_{94} doit être inférieur à 5 pour valider la précision selon l'article cité précédemment. L'interprétation est ainsi variable selon les auteurs et le ΔE utilisé.

4.2.4 Les contraintes industrielles

L'objectif des normes est de faire en sorte que deux appareils différents puissent obtenir une image similaire. Pour cela, il faut que l'utilisateur ne puisse faire la différence. Par conséquent, il y a divers contraintes de mesures à valider. Les normes sur lesquelles nous basons sont spécifiques à la numérisation manuscrite ou impression.

Nous rappelons qu'ils existent deux normes que nous allons prendre en référence Métamorfoze et FADGI. La norme Metamorfoze (cf. [van Dormolen, 2008]) est une norme d'origine européenne (Pays-Bas), qui date de 2008. La norme FADGI (cf. [Initiative, 2010]) est une norme américaine, qui date de 2010.

4.2.4.1 Metamorfoze

Les mesures colorimétrique sont faites sur la *Colorchecker SG 140 patches*, la *Colorchecker classic 24 patches*, la *Kodak Gray Scale Q-13 (22 patches)* et les deux mires conçues pour la norme : *universal test target (utt)* et la *Scanner reference chart*. Nous pouvons découper les tests d'analyse de la couleur en deux parties, la premier concerne la luminance et la seconde la chromaticité.

La luminance :

- Le premier critère est la précision du patch blanc (balance des blancs) et du patch noir. L'objectif est d'avoir un même intervalle de définition quelque soit le matériel d'acquisition pour une image qui est définie dans un espace RVB classique (en l'occurrence le EciRGBv2 ou Adobe RGB qui ont un Gamma de 1.8 et 2.2 respectivement). La tolérance est de de 5 niveaux par rapport à la valeur de référence.
- Le second est la neutralité des niveaux de gris et la méthode suggérée pour la quantifier est la suivante :
soit ΔC^* tel que

$$\Delta C_{ab(i)}^* = \sqrt{a_{s(i)}^{*2} + b_{s(i)}^{*2}} - \sqrt{a_{ref(i)}^{*2} + b_{ref(i)}^{*2}} \quad (4.30)$$

avec s pour l'échantillon et ref pour la référence. L'indice i représente le numéro du patch des niveaux de gris. Il faut que :

$$C_{ab(i)}^* \leq 2 \quad \forall i \quad (4.31)$$

- Le troisième est la précision de chaque patch, la mesure est basée sur le calcul de la tolérance d'exposition définie par :

$$\Delta L_i^* = L_{s(i)}^* - L_{ref(i)}^* \quad (4.32)$$

avec i le numéro du patch de niveau de gris, s la mesure et ref la référence. Il faut que :

$$|\Delta L_i^*| \leq 2 \quad (4.33)$$

- Le quatrième critère est basé sur le gain modulation défini par :

$$g = \frac{L_{s(i)}^* - L_{s(i+1)}^*}{L_{ref(i)}^* - L_{ref(i+1)}^*} \quad (4.34)$$

avec les mêmes notations que précédemment. Il faut que :

$$g \leq 1.05 \quad (4.35)$$

La couleur :

La mesure des différences de couleur utilisée par cette norme est basée sur le calcul du ΔE_{76} . Et, il faut que :

$$\begin{aligned} Mean \Delta E &\leq 4 \\ Max \Delta E &\leq 10 \end{aligned}$$

4.2.4.2 FADGI

La documentation de la FADGI est très exhaustive puisqu'elle est à la fois un document récapitulant les mesures mais aussi un document détaillant les protocoles de numérisation pour tous les types de documents. La partie concernant la reproduction de la couleur est très exhaustive. Nous pouvons reprendre le tableau de cette documentation concernant notre sujet à la figure 4.11. Ce tableau représente l'ensemble des critères permettant de mesurer la réponse du capteur jusqu'à la valeur du pixel. Les deux premières colonnes mesurent la précision du capteur en faisant varier les conditions d'éclairage. Les deux dernières colonnes mesurent les valeurs obtenues des pixels : la balance des blancs et la précision colorimétrique après calibration. La FADGI conseille d'utiliser les mires Kodak pour faire les tests, soit les mire Q-13 (cf. figure 4.21), Q-14 (cf. figure 4.17) et la Q-60 (similaire à la it8 de la figure 4.18 fait par Kodak).

De la même manière que Metamorfoze, l'évaluation de la luminance est beaucoup plus important que la simple mesure des distances globales, l'évaluation est aussi faite sur l'image dans un espace *RVB* standard. La FADGI a la particularité de donner un système d'évaluation sous forme d'étoile.

Part II - Evaluation and Quality Control of Digital Imaging
- SIGNAL -

Engineering Metrics	OECF – Opto Electronic Conversion Function (ISO 14545) TTF – Tone Transfer Function TRC – Tone Reproduction Curve <i>definition : Average large area digital response of an electronic imaging device to light stimuli</i>			
Derivative Metrics	Sensitivity (ISO 12232) <i>definition: The reciprocal of the amount of light necessary to achieve a desired output response.</i>	Tone and Exposure <i>definition: characteristic behavior of large area digital output response (count value) to spectrally neutral input stimuli (gray patch)</i>	White Balance/Neutrality <i>definition: equivalence of large area color channel output responses to a range of spectrally neutral input stimuli</i>	Color Encoding/Rendering Accuracy <i>definition: The difference between selected physically measured input colors and their intended output rendering from a given color space.</i>
Related descriptive term	- Responsivity - Speed - Exposure Index (EI)	- Too dark/light - Under/over exposed - No shadow/highlight detail - Clipping - Contrast - Exposure Accuracy	- Color cast - Gray balance	- Over/under saturated colors - Color balance is wrong - Memory colors are not correct - Color Accuracy - Color Saturation
Evaluation Criteria (Units) ● Primary ○ Secondary	●: Saturation based speed <i>units: TBD</i> ▼: Noise based speed <i>units: TBD</i> ○: Exposure Index, Standard Output Sensitivity	●: Average, median, maximum or RMS deviation from aim for neutral patches of interest. <i>units: Count Values, ΔL*, Density, F-stops</i> ▼: Deviation from a reference OECF gamma value <i>units: gamma (unitless)</i>	●: Average, median, maximum, or RMS deviation from aim between color channels (R-G, R-B, G-B) for neutral patches of interest. <i>Units (●): Count Values, ΔE_{a1b*}, ΔE_{uv*}</i> <i>Units (▼): Delta C, Delta H</i>	●: Average, median, maximum, or RMS deviation from aim for chromatic patches of interest <i>Units (●): Count Values, Delta E (ΔE), Delta E (ΔE_{a1b*})</i> <i>Units (▼): Delta C, Delta H</i>
Possible failure causes	- Inefficient imaging detector	-Auto-contrast failures -Inappropriate black/white point calibration. - Wrong gamma selection or tone aim	- Poor auto-white balance algorithm - Bad white /black point calibration - Sparse gray patch balancing - Color Balance - Strongly colored environmental surround	- Color profile tweaked for preference - Wrong color profile intent - Wrong color profile chosen/embedded - Color profile assumptions inconsistent with practice (i.e. lighting quality, gamma, intent, etc.) - Environmental : highly chromatic color surround/clothing

FIGURE 4.11: Tableau des normes de la FADGI concernant la reproduction de la couleur [Initiative, 2010]

La luminance :

L'évaluation de la luminance est basée sur la norme ISO 14545 – Opto Electronic Conversion Function. Ses deux critères sont les suivants :

- L'erreur de la balance des blancs qui est évaluée en se basant sur la moyenne des différences entre les canaux vert et bleu, d'une part, et les canaux vert et rouge, d'autre part, sur tous les patches de niveaux de gris. La mesure se récapitule dans le tableau suivant :

Performance level	Objectif	Tolerance - équivalence en 8-bit (appliquer la différence entre VB et VR pour tous les patches neutres)
4	0	±3 niveaux
3	0	±4 niveaux
2	0	±6 niveaux
1	0	±8 niveaux

TABLE 4.2: L'erreur de la balance des blancs

- La précision des patches de niveaux de gris qui est effectuée sur tous les canaux comme le récapitule le tableau suivant :

Performance level	Objectif	Tolerance - équivalence en 8-bit (appliquer pour tous les niveaux de densité et pour chaque couleur)
4	0	±3 niveaux
3	0	±6 niveaux
2	0	±9 niveaux
1	0	< 9 niveaux

TABLE 4.3: L'erreur des niveaux de gris

Le document propose une visualisation de l'erreur pour une courbe de gamma de 2.2 que nous reprenons dans la figure 4.12.

La couleur : Les contraintes sur les couleurs sont basées sur la mesure du ΔE_{00} , nous notons que les critères sont proche de l'article de Abrado et al que nous avons vu dans la sous-section précédente. Nous récapitulons les critères dans le tableau 4.4. Nous observons dans ce tableau qu'il est coupé en deux parties. La différence entre les deux est la variable L^* qui est prise en compte dans la première partie mais pas dans la seconde. Le choix B s'explique par le fait que la mesure de la luminance justifie en partie la précision de la couleur. Les premiers critères sont basés sur la pertinence de L , il n'est alors pas utile de le re-mesurer à nouveau ce paramètre. Dans les deux cas, nous avons une évaluation en fonction du maximum et de la moyenne.

D'après ces deux normes, nous constatons que la précision des niveaux de gris est considérée comme plus importante que la précision couleur. Le nombre de contraintes et le

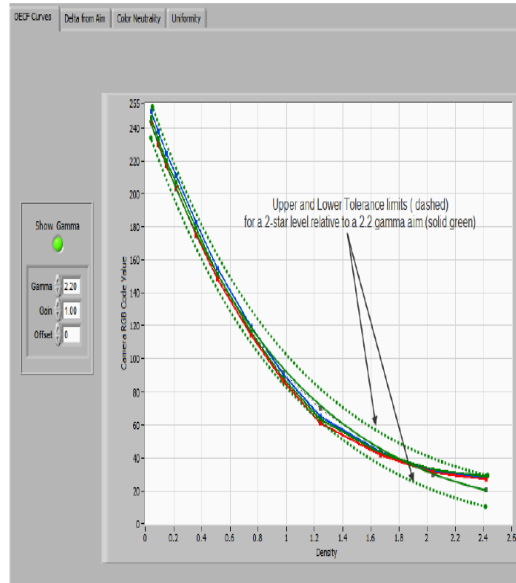


FIGURE 4.12: Mesure de la courbe des niveaux des gris [Initiative, 2010]. Nous avons en entrée (abscisse) la densité physique des patches de niveaux de gris et en sortie la valeur RVB au format standard. La valeur mesurée doit se situer entre les deux courbes en pointillés.

Performance Level	Tolerance (choose option A or B)			
	A) $\Delta E00 L^*a^*b^*$		B) $\Delta E00 a^*b^*$	
	max	avg.	max	avg.
4	< 6	< 3	< 3	< 2
3	< 10	< 5	< 5	< 3
2	< 15	< 10	< 8	< 6
1	> 15	> 10	> 8	> 6

TABLE 4.4: Tableau représentant les évaluations de la couleur en fonction du $\Delta E00$ [Initiative, 2010]

choix possible de décorrélérer la couleur de la luminance, pour la FADGI, mettent en évidence que la luminance est importante.

Pour mesurer ce critère, ces variables sont mesurées dans un icc sortie, il nous faut alors être précis dans notre conversion en $L^*a^*b^*$ vu qu'a priori, la conversion $L^*a^*b^*$ en RVB n'augmente pas l'erreur. Nous distinguerons alors le ΔE pour les valeurs de niveaux de gris et pour les valeurs couleurs.

4.2.5 Processus de calibration d'un matériel de numérisation

Nous allons définir l'ensemble des étapes permettant de calibrer un scanner. Il faut rappeler que l'étape de calibrage colorimétrique s'effectue après avoir fait la mise au point de l'appareil et régler l'éclairage. L'éclairage, quand à lui, est choisi selon l'utilisateur et ne doit pas être changé. Nous rappelons que la couleur est dépendante de la source lumineuse.

Le deuxième contrôle consiste à régler les niveaux de blancs et de noir pour avoir un intervalle de définition des trois canaux. Cette étape s'appelle la balance des blancs. La balance des blancs consiste à étalonner l'appareil afin qu'il puisse qu'il puisse s'adapter au contexte d'éclairage. La valeur blanche est ainsi réglée à l'aide d'une surface blanche (ou bien de la mire colorimétrique). Le cas de la numérisation s'apprête à régler la valeur du noir aussi, présente sur la mire. De cette manière, l'espace couleur est définie par deux extremums. Nous rappelons que le système de conversion d'une caméra est l'application d'une fonction linéaire et que ce réglage réajuste les paramètres de la fonction. Nous notons aussi que les références blanc et noir sont systématiquement présents dans les mires et, par conséquent, ces valeurs sont normées.

Ensuite, nous pouvons utiliser une mire colorimétrique pour calibrer le scanner. Les mires colorimétriques sont créées par des fabricants spécialistes où des mesures théoriques sont fournies pour chacun des patches de la mire.

Le premier travail à mettre en place va consister à détecter et identifier tous les patches de la mire afin de récupérer les valeurs et de définir la mire numériser. A partir de cela, nous allons avoir un ensemble de valeurs RVB .

Le processus de calibration va permettre de définir une transformation spécifique à l'appareil et l'éclairage utilisé vers l'espace de référence. Il faut pouvoir trouver une transformation qui compensera les erreurs de l'appareil et de les convertir en $L^*a^*b^*$.

Nous remarquons que l'image créée par l'appareil utilise un espace colorimétrique RVB dont l'ensemble des définitions des couleurs lui est spécifique (cf. figure 4.13) et qu'il est très différent d'un espace RVB standard (cf. figure 4.14).

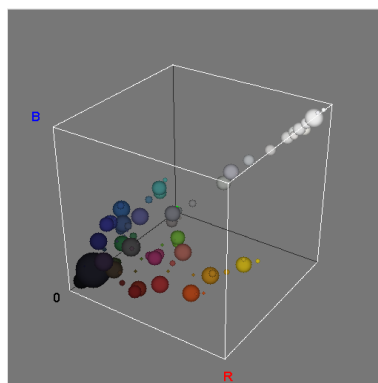


FIGURE 4.13: Nous avons représenté la répartition colorimétrique de la *colorchecker classic* dans l'espace RVB de l'appareil et nous le comparons à celle d'un format standard dans la figure 4.14

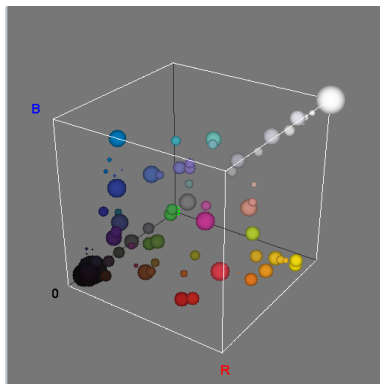


FIGURE 4.14: Représentation de la répartition colorimétrique de la *colorchecker classic* dans l'espace RVB standard

Une fois que nous avons déterminé une conversion RVB vers $L^*a^*b^*$, nous utilisons les données $L^*a^*b^*$ de la mire (cf. l'exemple de la figure 4.16). Ainsi, nous pourrions calculer la différence de deux couleurs et évaluer la précision colorimétrique. Pour exemple, l'application naïve de la conversion classique donne un delta E aux alentours de 5.5.

Nous résumons notre problématique dans le schéma 4.15.

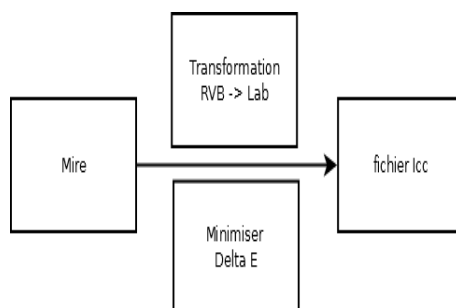


FIGURE 4.15: Schéma qui résume la problématique de la calibration couleur

Notre prochaine section présentera la solution à notre problématique.



The data below is intended to be an average measurement of all ColorChecker Charts. ColorChecker Charts should be replaced every two years as fading of the colors will alter the values of the chart.

No.	Number	sRGB			CIE L*a*b*			Munsell Notation Hue Value / Chroma
		R	G	B	L*	a*	b*	
1.	dark skin	115	82	68	37.986	13.555	14.059	3 YR 3.7 / 3.2
2.	light skin	194	150	130	65.711	18.13	17.81	2.2 YR 6.47 / 4.1
3.	blue sky	98	122	157	49.927	-4.88	-21.925	4.3 PB 4.95 / 5.5
4.	foliage	87	108	67	43.139	-13.095	21.905	6.7 GY 4.2 / 4.1
5.	blue flower	133	128	177	55.112	8.844	-25.399	9.7 PB 5.47 / 6.7
6.	bluish green	103	189	170	70.719	-33.397	-0.199	2.5 BG 7 / 6
7.	orange	214	126	44	62.661	36.067	57.096	5 YR 6 / 11
8.	purplish blue	80	91	166	40.02	10.41	-45.964	7.5 PB 4 / 10.7
9.	moderate red	193	90	99	51.124	48.239	16.248	2.5 R 5 / 10
10.	purple	94	60	108	30.325	22.976	-21.587	5 P 3 / 7
11.	yellow green	157	188	64	72.532	-23.709	57.255	5 GY 7.1 / 9.1
12.	orange yellow	224	163	46	71.941	19.363	67.857	10 YR 7 / 10.5
13.	blue	56	61	150	28.778	14.179	-50.297	7.5 PB 2.9 / 12.7
14.	green	70	148	73	55.261	-38.342	31.37	0.25 G 5.4 / 8.65
15.	red	175	54	60	42.101	53.378	28.19	5 R 4 / 12
16.	yellow	231	199	31	81.733	4.039	79.819	5 Y 8 / 11.1
17.	magenta	187	86	149	51.935	49.986	-14.574	2.5 RP 5 / 12
18.	cyan	8	133	161	51.038	-28.631	-28.638	5 B 5 / 8
19.	white (.05*)	243	243	242	96.539	-0.425	1.186	N 9.5 /
20.	neutral 8 (.23*)	200	200	200	81.257	-0.638	-0.335	N 8 /
21.	neutral 6.5 (.44*)	160	160	160	66.766	-0.734	-0.504	N 6.5 /
22.	neutral 5 (.70*)	122	122	121	50.867	-0.153	-0.27	N 5 /
23.	neutral 3.5 (1.05*)	85	85	85	35.656	-0.421	-1.231	N 3.5 /
24.	black (1.50*)	52	52	52	20.461	-0.079	-0.973	N 2 /

Cie L*a*b* values use Illuminant D50 2 degree observer sRGB values for illuminate D65.

FIGURE 4.16: Charte colorimétrique de la *colorchecker classic*

4.3 La calibration colorimétrique d'un scanner

L'objectif de la calibration couleur consiste à corriger les erreurs de mesure de l'appareil d'acquisition.

La méthode classique permettant de corriger les données mesurées est d'utiliser un modèle de résolution linéaire tel que la méthode des moindres carrés. Pour résoudre ce problème, il faut déterminer un modèle mathématique décrivant au mieux le problème nous permettant d'atteindre nos objectifs (c'est-à-dire un modèle se basant sur la conversion de RVB à $L^*a^*b^*$ vu dans la sous-section 4.2.2).

Le problème de la calibration se formalise de la manière suivante. Soit X un ensemble de départs qui représente les mesures des couleurs de la mire captée par la caméra et Y un ensemble d'arrivées défini par les mesures théoriques correspondantes. X et Y ont le même nombre d'éléments N . Soit f une application bijective tel que :

$$f : X \rightarrow Y, \text{ telle que } f(x) = y \text{ avec } x \in X \text{ et } y \in Y \quad (4.36)$$

et doit minimiser la somme des résidus S , tel que :

$$S = \sum_{i=1}^N \Delta E(y_i, f(x)) \quad (4.37)$$

La transformation est nécessairement bijective parce que les patches représentent a priori des couleurs différentes et si deux patches ont la même couleur (métamérisme), le résultat ne peut être correcte.

Le premier point que nous traiterons est le choix des patches. En effet, il existe un grand nombre de mire proposant des couleurs différentes. Nous observerons ce modèle qui propose différents points de référence et qui va conditionner la résolution de notre système. Le deuxième point que nous allons aborder est l'évaluation du gamma dans l'ensemble de départ RVB machine. Ce critère a une importance sur le résultat final car il permet d'homogénéiser toutes les luminances.

Nous étudierons ensuite les différents modèles mathématiques pour résoudre le problème des moindres carrés qui joue essentiellement sur le degré du polynôme. Nous verrons aussi une méthode non-linéaire permettant de résoudre ce type de problématique.

Et enfin, notre dernier point traitera du calcul du ΔE : dans un premier temps, nous évoquerons les erreurs de calcul du ΔE que l'on retrouve dans les logiciels d'analyse de

référence, puis dans un second temps, nous confronterons les résultats de notre méthode aux logiciels de calibration couleur existants.

4.3.1 Les caractéristiques de la mire de calibration

Nos couleurs de référence sont obtenues à partir de la mire. La mire donne à disposition un certain nombre de mesure à faire. En particulier, la mire influence la transformation en fonction du nombre de patch et de l'écart entre les couleurs des patchs. Nous proposons une analyse sur le choix de patchs de couleur et, dans un second temps, le choix des mires que nous avons choisis de travailler dans le contexte de la numérisation.

Tout d'abord, il existe un grand nombre de mires car elles correspondent souvent à l'évolution technologique. Les premières mires de référence sont la mire *Colorchecker classic* de 1998 (voir la figure 4.21) et la mire *Kodak* de 1997.

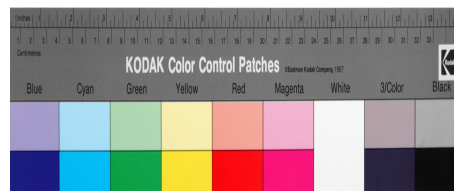


FIGURE 4.17: Mire de Kodak

Ces deux mires proposent des informations de niveaux de gris accompagnées des couleurs primaires et secondaires, ainsi que des couleurs "semi-claires" et "semi-foncées". La raison était que la faible qualité des appareils de l'époque ne permettait pas de différencier l'ensemble des couleurs. Par conséquent, le nombre de patch était faible et caractéristique de l'espace de définition de la couleur la plus simple.

Aujourd'hui, nous constatons qu'il existe une multitude de mires de couleurs. Différentes sociétés proposent des mires avec plus de patchs, par exemple la mire AGFA *it8* (cf. 4.18) qui propose environ 280 patchs, alors que d'autres essaient de trouver un juste équilibre entre la quantité et la pertinence du choix des couleurs comme la mire *colorchecker SG140* (cf. 4.18). Nous notons que la mire kodak Q-60 a été remplacé par la *it8* d'AGFA.

Nous voyons que les deux exemples de mires citées sont basés sur deux concepts différents.

La première mire est un exemple d'utilisation de la répartition exhaustive (environ 250 patchs) des couleurs combinées avec la variation de la luminance. Nous pouvons remarquer la présence des patchs cyan, magenta et jaune qui permettent de faire une transformation éventuelle de *RVB* à *CMYK* (espace couleur d'impression). Ces couleurs sont caractéristiques de l'impression. Nous constatons que d'autres mires composées de



FIGURE 4.18: Deux autres exemples de mières couleur. La mire Agfa *it8.7/2* en haut et la mire colorchecker SG en bas.

beaucoup de patches comme les mières de Munsell. Munsell est à l'origine de l'espace colorimétrique TVC (teinte, valeur (luminosité), chromaticité). Les patches de la mire sont construits à partir d'un échantillonnage particulier d'un espace de couleur prédéfini.

L'autre approche propose d'étudier une répartition géométrique plus pertinente des couleurs. L'analyse de la luminance est a priori faite sur les patches de niveaux de gris, il n'est pas nécessaire de traiter à nouveau cette information. Les couleurs utilisées sont prises dans le spectre visible de l'œil. L'ensemble des couleurs sélectionnées est a priori l'ensemble des informations représentant au mieux cet espace.

Nous observons aussi que le spectre visible peut être découpé en régions de couleurs. Nous pouvons remarquer que la mire peut avoir plus de patches sur certaines régions de couleur afin de gagner en précision (indépendamment de sa taille), comme la mire *colorchecker SG* (cf. figure 4.18) sur les tons proches de la couleur de peau (*D7* à *J8*),

les bleus (*B5* à *D8*) ou les verts (*C9* à *L9*).

La mire de type *it8* n'est pas réellement utilisée dans le cadre de la numérisation mais davantage dans le contexte de la photo parce qu'elle propose un gros volume de couleur différente et particulièrement présentes dans les images naturelles. L'ensemble complet des couleurs n'est pas pertinent dans le contexte de la numérisation de livre ancien. Les mires que nous prendrons en référence sont les mires *colorchecker classic* et *colorchecker SG 140*. Nous notons que la combinaison *Q-13* et *Q-14* sont très proche de la *colorchecker classic* mais la qualité du papier rentre en compte. Effectivement, il existe trois catégories de papier : brillant, semi-brillant et mat. Les mire *kodak* sont plus brillantes que les mires *colorchecker*. De ce fait, les mires *colorchecker* paraissent plus pertinente pour numériser des documents anciens.

4.3.2 Le gamma du scanner

La représentation d'une acquisition est faite dans l'espace *RVB* machine. Nous avons pu voir dans la partie 4.2.2.1 que l'espace *RVB* doit respecter un certain nombres de propriétés. La première est basée sur l'utilisation de trois variables indépendantes. La seconde est la définition du blanc qui est définie par les composants *R*, *V* et *B*. Pour chaque triplet *R*, *V* et *B*, nous avons un nouveau système et donc un nouvel espace de définition.

La calibration colorimétrique propose de contrôler les deux propriétés citées. L'une des étapes consiste à fixer les valeurs des noirs et des blancs en faisant la balance des blancs. Le réglage du gamma permet de consolider l'indépendance des canaux *RVB* en se basant sur la luminance de la mire. Par exemple, si un patch de niveau de gris est défini par une valeur *R* différente des canaux *G* et *B*, alors cela signifie que *G* et *B* compensent l'erreur de *R*. Les variables ne sont pas indépendantes, nous pouvons retrouver l'expression de "linéarisation de *RVB*" concernant l'étalonnage du gamma.

La gestion des niveaux de gris et des noirs est faite par l'application de la transformation gamma. Nous rappelons que la transformation gamma est une transformation basée sur la fonction puissance qu nous avons vu 2.19.

Cette transformation est systématique dans un espace *RVB* standard (usuellement 1.8 et 2.2). Typiquement, nos écrans de télévision ont un gamma de 2.2. Ce gamma est historique car il corrige l'affichage des télévisions à tube cathodique. On l'appelle le "decoding gamma" et son inverse l'"encoding gamma" (cf. figure 4.19), en référence à leur position dans la chaîne. En pratique, ce gamma 2.2 est devenu un modèle pour la

création des mires de niveaux de gris. En pratique, ce gamma 2.2 est devenu un modèle pour la création des mires de niveaux de gris. Effectivement, la création d'une mire avec un gamma identique à l'écran permet de comparer visuellement le résultat.

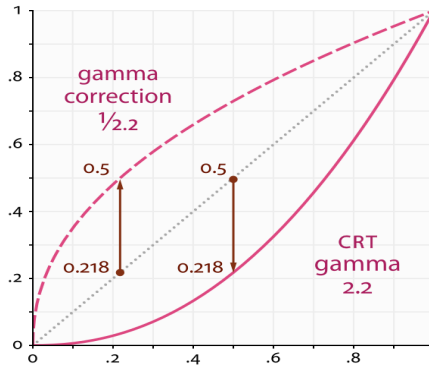


FIGURE 4.19: Représentation de la fonction gamma avec deux exemples caractéristiques de la représentation couleur

Pour estimer le gamma, il nous faut simplement déterminer la valeur γ de la gamma correction à partir des patches de niveaux de gris.

Il existe plusieurs estimateurs du gamma. La méthode évidente est d'utiliser le logarithme permettant de faire apparaître la puissance. Nous pouvons formuler le problème de la manière suivante.

D'après la transformation gamma, nous écrivons

$$\log(V_s) = \gamma \log(V_e) \quad (4.38)$$

L'estimation de γ peut alors être obtenue par le ratio suivant :

$$\gamma_{estimate} = \frac{\sum_i^n \gamma_i}{i} \quad (4.39)$$

, avec $n \geq 1$ le nombre de patch et

$$\gamma_i = \frac{\log(V_{s_i})}{\log(V_{e_i})} \quad (4.40)$$

Une autre approche se distingue : c'est l'estimateur proposé par Bruce Lindbloom (voir [Lindbloom, 2010]) qui se base sur le modèle des moindres carrés et il obtient l'équation suivante :

$$\gamma_{estimate} = \frac{\sum_i^n \log V_{s_i} \log V_{e_i}}{\sum_i^n \log V_{e_i}^2} \quad (4.41)$$

, avec n le nombre de patch, V_s la mesure relative de la luminance (définie dans l'intervalle $[0, 1]$) et V_e la valeur mesurée (définie dans l'intervalle $[0, 1]$).

Nous notons que lorsque nous avons une description détaillée des niveaux de gris, nous pouvons ajuster une courbe définissant le gamma localement sur l'ensemble de l'intervalle.

En pratique, la conception de mire se base sur la densité de niveau de gris. Effectivement, la mire la plus basique est composée de trois patchs de niveaux de gris (cf. figure 4.20) : le blanc (proche de $L^* = 100$), le noir (proche de $L^* = 0$) et le gris "parfait" représentant $L^* = 50$. Ces patchs sont représentés en fonction de la densité de noir dans une région donnée. Il n'est pas possible de produire physiquement une matière totalement noire ni totalement blanche. Nous avons un triplet de valeurs de référence comme nous le montrons dans le tableau 4.5.

Patch	Densité	Réflectance	Luminance
Blanc	-0.05	0.8913	95.63 %
Gris	-0.7	0.1995	51.78 %
Noir	-1.5	0.0316	20.68 %

TABLE 4.5: Tableau représentant un triplet Blanc/Gris/Noir basé sur les densités données par une mire ColorChecker que l'on peut voir dans la figure 4.20



FIGURE 4.20: Mire de référence Blanc/gris/Noir fabriqué par X-rite

Les différentes valeurs obtenues sont représentées en luminance dans ce tableau. Mais, on peut interpréter les valeurs de luminance sous forme de pourcentage sur les canaux rouge, vert et bleu. Le point blanc, par exemple, est défini à 95.63%. Dans une image définie sur 8 bits, nous allons avoir pour le blanc, $R = V = B = 0.9563 * 255 = 243.86$. Nous rappelons que la norme *Metamorfoze* impose une précision de plus ou moins 5 comme nous l'avons vu dans la sous-section 4.2.4.

Si nous prenons d'autres mires usuelles (cf. figure 4.21), nous retrouvons un nombre de patches plus important. Ce nombre oscille entre 6 et 20 patches.



FIGURE 4.21: Deux exemples de mires. La mire de gauche est la Colorchecker Classic de X-rite et la mire de droite est la Q13 de Kodak.

Selon la documentation de Metamorfoze (cf. [van Dormolen, 2008]), la mire de référence est normée en ce qui concerne la conversion entre la densité d'un patch de niveau de gris et sa valeur en luminance. Cela nous permet de connaître la valeur des patches.

Pour résoudre ce problème, nous avons choisi d'utiliser l'estimateur de Bruce Lindbloom, soit l'équation 4.41 pour les mires de référence qui sont composées d'un nombre restreint de patch de niveaux de gris (*colorchecker classic* et *SG140*), et nous proposons d'utiliser une courbe d'ajustement lorsque nous avons un grand nombre de patch (la mire *Q-13*, *it-8*).

La première étape de la calibration couleur va alors consister à estimer le gamma mesuré par le matériel d'acquisition. Le résultat de cette transformation nous permet d'avoir une image d'illumination neutre qui a pour notation $R'V'B'$.

4.3.3 Estimation par un modèle linéaire et non-linéaire pour la transformation de RVB à $L^*a^*b^*$

La transformation consiste à déterminer une conversion entre les valeurs mesurées et les valeurs originales (cf. figure 4.22). Après avoir réajuster le gamma, nous pouvons calculer la transformation de $R'V'B'$ à Lab .

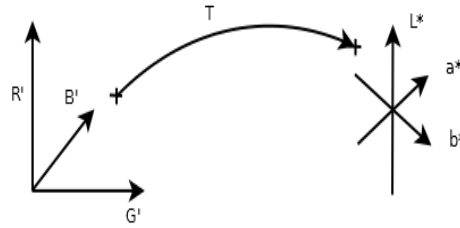


FIGURE 4.22: Transformation de l'espace $R'V'B'$ à l'espace de référence $L^*a^*b^*$

Nous nous situons dans l'étape qui consiste à calculer le passage de $R'V'B'$ en $L^*a^*b^*$. Le changement d'espace peut-être aussi calculé de $R'V'B'$ vers XYZ mais ce dernier doit à nouveau faire une conversion, il paraît évident de passer directement à l'espace $L^*a^*b^*$. Le changement d'espace d'un espace vers lui-même donne des résultats moins précis car la transformation de $R'V'B'$ à $L^*a^*b^*$ (ou XYZ) combine les trois variables R, V et B pour déterminer L^* , a^* ou b^* alors que dans les autres cas il n'y aura qu'une seule variable.

Les travaux concernant cette partie sont essentiellement basés sur la minimisation d'une fonction de conversion comme nous l'avons introduit au début de la section. Nous retrouvons des articles utilisant la régression polynomiale et d'autres approches non-linéaire avec notamment l'utilisation d'un réseau de neurone pour effectuer la régression.

Nous citons les articles suivants en ce qui concerne la régression polynomiale : [Hardeberg et al., 1996], [Kang, 1997], [Hardeberg, 2001] et [Yilmaz et al., 2004]. Les différents articles se distinguent par les différentes mires utilisées. Ils mettent en évidence que la méthode donne des résultats convaincants mais que le degré du polynôme influe sur la précision des couleurs. Par exemple, dans l'article de Hardeberg et al [Hardeberg et al., 1996], ils mettent en avant que l'utilisation d'un polynôme de degré 3 est plus pertinent qu'un polynôme de degré 2 pour le cas de la mire it8 qui est une mire composée de 290 patches. Par la suite, nous retrouvons dans sa thèse [Hardeberg, 2001] que la précision de la transformation est très influencée par la correction gamma. Ils proposent aussi une application supplémentaire d'un gamma de 1/3 sur ses coordonnées RVB lui permettant d'obtenir un ΔE moyen de 1.5.

L'article de Yilmaz et al [Yilmaz et al., 2004] propose, dans un premier temps, un bilan de l'état de l'art sur la calibration couleur d'un point de vue général et énumère des articles proposant une régression polynomiale pour résoudre la problématique, dans un second temps, les auteurs proposent leur approche et donnent leur résultat. Ils suggèrent une régression polynomiale dont le degré est supérieure à 3 pour la mire it8 et sur différents appareils d'acquisition. Ils constatent que la régression polynomiale de degré 4 est plus précise. Effectivement, un polynôme de degré 4 a un nombre important de coefficients à déterminer par conséquent lorsque le nombre de mesure est importante la transformation sera plus précise.

Nous en concluons que le degré du polynôme doit varier en fonction du nombre de patch

utilisé ; plus il y a de patch, plus le degré du polynôme doit être important. Comme nous pouvons connaître le nombre de patch en fonction de de la mire, chaque mire va posséder une transformation spécifique.

Posons le problème de la minimisation des moindres carrés. Soit y_i les valeurs colorimétriques théoriques, x_i les valeurs mesurées par la caméra et $f(x_i)$ les valeurs calculées, $i = [0..m]$ avec m le nombre de valeurs et f un polynôme de degré n ou $n + 1 > m$. L'expression de la méthode des moindres carrés est :

$$MC = \sum_{i=0}^m (y_i - f(x_i))^2 \quad (4.42)$$

La minimisation de cette somme au carré est atteinte pour :

$$\nabla(MC) = 0 \quad (4.43)$$

La solution de cette équation est obtenue par un système à $n + 1$ équations à résoudre. Ce système se résout lorsque l'on pose le problème de la manière suivante : Soit $f(x) = a_0 + a_1x + \dots + a_nx^n$, ($a_k \in R$; $k = 0, \dots, n$), alors on peut noter les matrices de la manière suivante :

$$y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}, a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}, V = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{pmatrix} \quad (4.44)$$

Le système s'écrit :

$$V^t V a = V^t y \quad (4.45)$$

La solution de cette équation pour un vecteur a , dans la mesure où V^tV est une matrice inversible, s'écrit :

$$a = (V^tV)^{-1}V^ty = V^-y \quad (4.46)$$

où V^- est la pseudo-inverse de V .

En pratique, nous avons trois fonctions à déterminer correspondant aux valeurs L^* , a^* et b^* à partir des trois variables indépendantes R , V et B de l'acquisition. Soit l_{coef} , a_{coef} et b_{coef} , les coefficients du polynômes, L , a et b les valeurs théoriques des patches et la matrice V_n^- obtenue par le calcul de la pseudo-inverse V_n qui est composée des valeurs R , V et B et de degré n . Nous obtenons alors les équations suivantes :

$$L_{coef} = V_n^-L, \quad a_{coef} = V_n^-a, \quad b_{coef} = V_n^-b \quad (4.47)$$

D'après notre modèle, nous définissons la fonction $f(x)$ en fonction de R , V et B . Pour un polynôme de degré 1, nous avons :

$$f(x) = a_0 + a_1R + a_2V + a_3B \quad (4.48)$$

Pour un polynôme de degré 2, nous avons :

$$f(x) = a_0 + a_1R + a_2V + a_3B + a_4R^2 + a_5RG + a_6RB + a_7G^2 + a_8GB + a_9B^2 \quad (4.49)$$

et pour un polynôme de degré 3, nous avons :

$$\begin{aligned} f(x) = & a_0 + a_1R + a_2V + a_3B + a_4R^2 + a_5RG + a_6RB + a_7G^2 + a_8GB + a_9B^2 \\ & + a_{10}R^3 + a_{11}R^2G + a_{12}R^2B + a_{13}RG^2 + a_{14}RGB + a_{15}RB^2 + a_{16}G^3 \\ & + a_{17}G^2B + a_{18}GB^2 + a_{19}B^3 \end{aligned} \quad (4.50)$$

Ce modèle change la forme de la matrice V qui ne prend en compte qu'une seule variable dans l'équation 4.44, il faut alors passer à 3 variables R' , V' , B' .

L'autre démarche que nous notons est basée sur l'article de Vrhel et Trussell. Les auteurs de [Vrhel and Trussell, 1999] proposent d'utiliser une méthode régressive à partir d'un réseau de neurones artificielles. Les auteurs utilisent l'algorithme de Levenberg-Marquardt

qui est une solution au problème de minimisation d'une fonction étant donné que le problème se pose de manière similaire à la méthode des moindres carrés.

Soit Y un modèle non-linéaire tel que $y = f(\theta, x)$ avec les données x_i et Y_i , $i = 1, \dots, m$ où x_i est de dimension k et θ est de dimension N . La méthode de Levenberg-Marquardt recherche θ^* la solution de θ localement minimisant :

$$g(\theta) = \sum_i^m (Y_i - f(\theta, x_i))^2 \quad (4.51)$$

La solution se trouve en appliquant l'équation suivante de manière itérative.

$$\theta_{j+1} = \theta_j - (J'J + \lambda D)^{-1} J'(Y - f(\theta, x)) \quad (4.52)$$

Où Y est le vecteur $m \times 1$ contenant Y_1, \dots, Y_m ; X est la matrice $m \times k$ contenant X_1, \dots, X_m ; J est la matrice jacobienne $m \times n$ pour $f(\theta, x)$ en respectant θ ; et D est la matrice diagonale $n \times n$ pour ajuster les facteurs d'échelle.

La mire de référence que les auteurs utilisent possède 264 patches et suggèrent une approche non-linéaire. Le modèle utilisé (ou la fonction d'activation) est une fonction sigmoïde (cf. 4.53), avec un nombre de couches cachés variant de 5 à 25.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4.53)$$

L'approche linéaire et non-linéaire sont des choix importants pour la résolution du problème. En effet, le problème de visualisation n'a pas été abordé. La résolution non-linéaire engendre des transitions couleurs parfois importante qui deviennent visible. Par conséquent, une telle approche ne peut pas être utilisée sur des mires possédant un faible nombre de couleur car les transitions colorimétriques seront trop important, le résultat du ΔE sera très satisfaisant mais le rendu visuel posera problème. Par contre, lorsqu'il y a beaucoup de patches alors (comme il est le cas dans l'article présenté) une telle approche semble être la plus pertinente.

Dans l'ensemble, nous constatons que le problème de la calibration couleur est un problème de minimisation qui peuvent être améliorable par des méthodes ou algorithmes plus complexe. Nous n'avons pas identifier d'article proposant une minimisation avec des

contraintes en fonction du delta E 00. Le problème est industriel et n'apparaît donc pas dans la littérature.

Nous remarquons aussi que l'ensemble des articles se base sur la mire it8. Or, nous avons vu que cette mire n'est pas une référence dans la numérisation de livre ancien. De plus, notre mire de référence n'est pas fixé, il faut alors trouver une solution en fonction de la mire.

4.3.4 L'évaluation de la calibration

Une fois que nous avons trouvé la transformation gamma et notre fonction selon la mire choisie, nous pouvons alors créer notre fichier de calibration. La création du fichier .icc possède l'information permettant de convertir RVB en $L^*a^*b^*$.

Cette étape est effectuée par une table de correspondance. C'est une structure de données qui a pour but de remplacer le calcul. La table va être créée à partir de tous les cas possibles d'entrée RVB. Nous rappelons que la conversion de RVB vers $L^*a^*b^*$ se fait en deux étapes. La première étape consiste à passer de RVB à $R'V'B'$ par la transformation gamma (cf. partie 4.3.2). La seconde étape convertit $R'V'B'$ en $L^*a^*b^*$ à partir de la fonction obtenue par la régression polynomiale (cf. partie 4.3.3). Grâce à ces opérations, nous avons les correspondances entre les triplets RVB et les triplets $L^*a^*b^*$.

L'évaluation de la transformation se calculera à partir des outils de la partie 4.2.3. Cependant, une fois l'image convertie dans l'espace de couleur de référence, il faut la convertir à nouveau dans un format standard RVB car nos écrans ne peuvent afficher du $L^*a^*b^*$ (sans faire à nouveau une conversion). Nous pouvons résumer ces conversions dans le schéma de la figure 4.23.

De ce fait, les logiciels existants se basent sur l'image en RVB des appareils d'affichages. Nous avons utilisé plusieurs logiciels d'analyses et nous constatons que les résultats sont différents, que se soit entre les différents logiciels et le calcul évident des distances obtenues en $L^*a^*b^*$. Nous avons alors étudié principalement le logiciel *Imatest* qui dispose d'une documentation accessible (cf. [Imatest, 2013]) ce qui va nous permettre de comprendre les divergences des résultats. Nous pouvons remarquer que les logiciels utilisent trois profils icc pour calculer le ΔE_{00} , comme on peut le montrer dans la figure 4.24, ce qui nous paraît illogique dans la mesure où une seule suffirait.

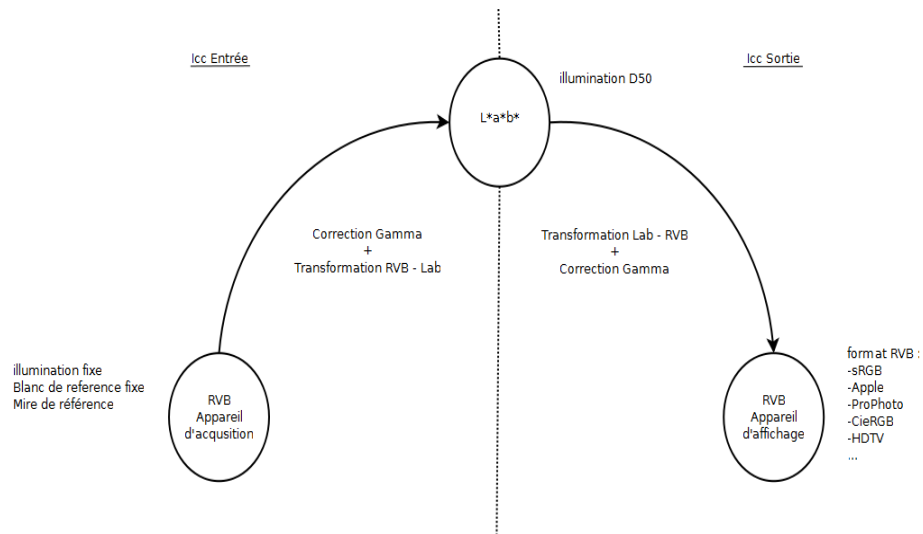


FIGURE 4.23: Application de deux fichiers .icc pour obtenir un visuel dans un espace RVB normé

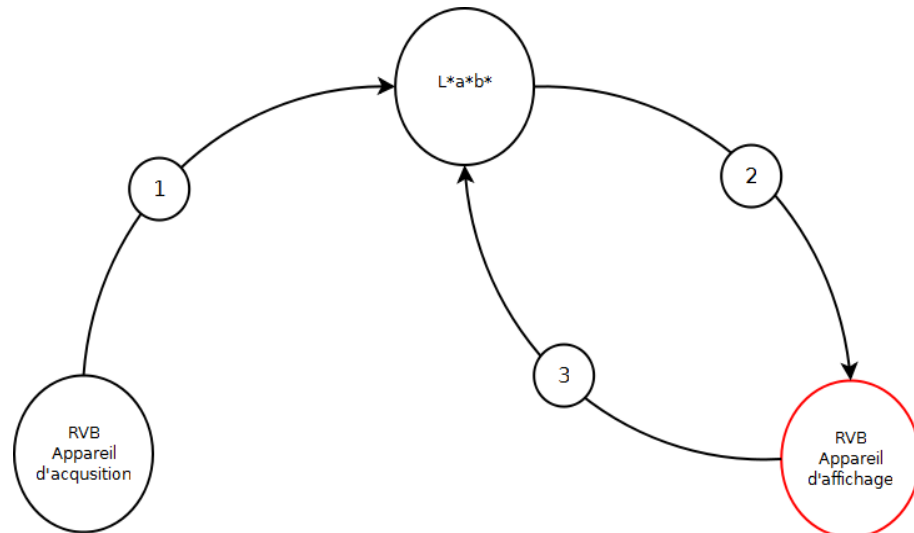


FIGURE 4.24: Les logiciels d'analyses d'images utilisent nécessairement trois profils ICC alors qu'un seul est nécessaire pour faire le calcul du ΔE qui se fait dans l'espace $L^*a^*b^*$

4.3.5 Analyse des divergences de résultat avec le logiciel Imatest (référence de l'évaluation colorimétrique)

Imatest propose une analyse complète de la mire *colorchecker classic* (cf. 4.21). L'image qu'il traite est a priori dans un format *RVB* standard. Le logiciel fait une estimation du gamma en fonction des patches de niveaux de gris, il convertit ensuite l'image en $L^*a^*b^*$ et il calcule le ΔE_{00} . Analysons ces trois étapes.

L'estimation du gamma est une étape classique de l'estimation de ce type de donnée. A ce propos, nous n'avons pas noté de différence ou d'erreur significative, mise à part la

notation sous forme d'encoding gamma.

La seconde étape est la conversion de sRGB (format *RVB* standard utilisé) qui est "incorrecte". Il est nécessaire que la conversion $L^*a^*b^* \rightarrow RVB \rightarrow L^*a^*b^*$ ne génère pas ou peu d'erreur. Nous observons que le fichier icc officiel de sRGB propose une conversion respectant cette condition (cf. le tableau dans l'annexe 6.2.2), puisqu'il génère une erreur de 0.18 pour le calcul du ΔE_{00} sur l'ensemble des patches, toutefois les autres formats ne proposent pas une aussi bonne précision que Adobe1998, Apple et EciRGB.

Nous constatons que pour un triplet $L^*a^*b^*$ lorsque nous appliquons les ICCs 2 et 3 de la figure 4.24, il y a des différences entre la valeur $L^*a^*b^*$ de départ et d'arrivée. Pour les quantifier, nous calculons en fonction des patches de référence de la *colorchecker classic* les conversions obtenues par le logiciel *Imatest*, la conversion de *Bruce Lindbloom* (notée comme étant une référence d'*Imatest*) et le profil icc officiel de sRGB (cf. [color.org, 2013]). Nous obtenons alors les courbes dans la figure 4.25 (avec les détails dans l'annexe 6.2.1) et les résultats dans le tableau 4.6.

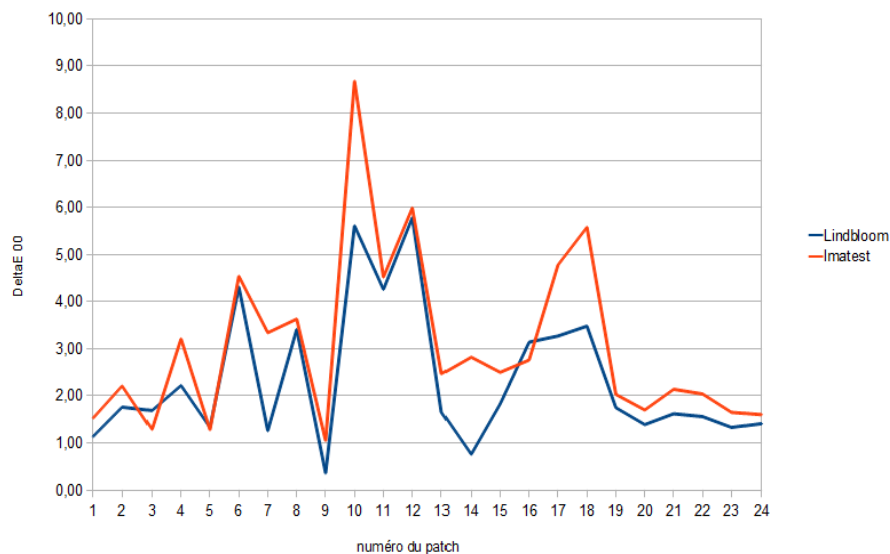


FIGURE 4.25: Représentation en deltaE des écarts de conversions à partir des patches de la *colorchecker classic*

	$\Delta E_{00}(\text{sRGB}, \text{Lindbloom})$	$\Delta E_{00}(\text{sRGB}, \text{Imatest})$
Moyenne	2.35	3.06

TABLE 4.6: Tableau des moyennes des ΔE_{00} entre les différentes conversions

Nous constatons avec cette série de résultats que la différence des multiples conversions d'un format standard est trop importante. Cela rend totalement inefficace l'évaluation de

la couleur. Le problème que pose ce genre de conversion est que le format supposé standard va toujours s'adapter à l'évolution technologique et qu'il y aura systématiquement un problème d'obsolescence.

Ce problème a été traité plus en profondeur par D. Pascale (cf. [Pascale, 2003] et [Pascale, 2006]). Ces articles énumèrent les erreurs qu'ils peuvent exister sur les imprécisions colorimétriques des mires et des erreurs de conversions possibles.

Le dernier point que nous souhaitons aborder est le problème de la mire. Nous constatons que les sociétés de fabrication de mire font évoluer leur technologies permettant de "colorer" leur patch que se soit en terme de coût de production et/ou de productivité. Ce phénomène engendre des différences de couleur de patch dans la même mire. Lorsque nous calculons le ΔE nous prenons en référence la documentation officielle (cf. [Pascale, 2014]), qui nous donne en référence les données de la colorchecker 2005. L'article [Pascale, 2006] fait une comparaison des différentes mires selon l'année d'édition. L'un des tableaux met en évidence un ΔE_{00} moyenne est de 1.23 entre la *colorchecker* 1976 et la *colorchecker* 2005, cette différence est a priori normale au vu des différentes années d'édition (l'œil ne fait pas de différence entre les deux mires). Nous avons choisi tout de même de comparer les différences proposées par le logiciel *Imatest* et la *colorchecker* 2005. Nous obtenons les résultats dans les deux tableaux dans l'annexe 6.2.3 puisque *Imatest* propose deux versions de mire. Les résultats des tableaux sont obtenus dans le tableau 4.7.

	ΔE_{00} (New Imatest,officiel)	ΔE_{00} (old Imatest,officiel)
Moyenne	1.52	2.48

TABLE 4.7: Tableau des moyennes des ΔE_{00} entre les différentes mires

La différence entre la mire officielle et les mires proposées par le logiciel est très élevé. Nous remarquons aussi qu'un ΔE de 2.48 met en évidence que l'oeil fait une distinction quasi nette entre les deux mires mais il n'est pas possible de s'en rendre compte visuellement. L'erreur d'estimation est ainsi due aux mauvais choix des référence de la mire.

Nous finirons notre analyse sur le fait que les erreurs engendrées dans les différentes étapes ne sont pas additives, par conséquent il est difficile de remarquer ce genre d'erreur. Malgré que le logiciel propose une méthodologie bonne, le ΔE n'a pas de sens à cause des imprécisions cités.

Nous calculerons, par la suite, le delta E à partir des données obtenues en $L^*a^*b^*$ en utilisant un unique fichier icc.

4.3.6 Résultats et analyses sur le $\Delta E00$

Notre analyse sur le ΔE va alors se faire uniquement à partir du fichier icc 1 tel que nous l'avons représenté dans la figure 4.24. Nous pouvons à partir de l'image de l'appareil d'acquisition, obtenir la valeur du patch en RVB et en lisant la table de correspondance du fichier icc de connaître son équivalence dans l'espace de couleur $L^*a^*b^*$. Par ce principe, nous pouvons alors évaluer le ΔE quelques soient la transformation qui a été utilisée pour créer la table de correspondance.

4.3.6.1 Analyse du degré du polynôme pour la résolution du problème linéaire par la méthode des moindres carrés

Nos travaux se sont focalisés sur les deux mires *colorchecker classic* et *colorchecker SG 140*. Nous allons nous comparons à deux logiciels *Coca* et *ProfileMaker* les plus référencés pour faire leur profil de calibration. Par conséquent, nous nous comparons à ces deux transformations inconnues. *Coca* est un logiciel de calibration couleur avec une version d'essai sur quelques mires notamment la colorchecker (pour plus d'informations voir la documentation de [of Australia, 2013]). *ProfileMaker* est un logiciel créé par la société fabriquant les mires colorchecker de la société *X-Rite*. Le tableau 4.8 présente les résultats sur la mire *colorchecker classic* et le tableau 4.9 présente les résultats de la *colorchecker SG140*.

Ces tableaux retracent les méthodes de transformations et les résultats du $\Delta E00$. Nous avons noté "LSE", la transformation des moindres carrées avec le degré du polynôme, "NN" l'utilisation d'un réseau de neurones et l'ajout de "+ Lum" pour signifier que nous avons utilisé une mire *Q13* pour calibrer le gamma. En ce qui concerne le $\Delta E00$, nous avons récapitulé la moyenne et le maximum parce que ces deux informations sont les critères de la norme FADGI (voir le tableau 4.4).

Dans le premier tableau, nous montrons dans un premier temps la pertinence de contrôler la luminance. Nous observons que pour chaque exemple nous gagnons en moyenne 0.8 de $\Delta E00$. De ce fait, nous utiliserons systématiquement ce contrôle par la suite.

Nous remarquons que la différence de résultat entre le polynôme de degré 2 et 3 est

échantillon	méthode	moyenne	maximum
1	LSE 2	3.11	6.08
	LSE 2 + Lum.	1.98	6.54
	LSE 3 + Lum.	4.5	8.05
	NN + Lum.	3.44	7.75
	CoCa	1.78	2.13
	Profile Maker	4.64	12.01
2	LSE2	2.7	5.36
	LSE 2 + Lum.	2.09	3.84
	LSE 3 + Lum.	4.36	8.7
	NN + Lum.	4.49	8.9
	CoCa	1.78	2.16
	Profile Maker	4.85	13.38
3	LSE2	1.91	3.6
	LSE 2 + Lum.	1.43	3.94
	LSE 3 + Lum.	xxx	xxx
	NN + Lum.	4.39	10.18
	CoCa	1.8	2.3
	Profile Maker	4.5	10.45
4	LSE2	2.23	4.17
	LSE 2 + Lum.	1.42	3.86
	LSE 3 + Lum.	2.7	8.2
	NN + Lum.	4.5	9.37
	CoCa	1.83	2.4
	Profile Maker	4.7	10.11
5	LSE2	2.02	3.38
	LSE 2 + Lum.	1.64	3.45
	LSE 3 + Lum.	5.88	11.17
	NN + Lum.	3.7	8
	CoCa	1.83	2.3
	Profile Maker	4.76	9.97
6	LSE2	1.76	3.64
	LSE 2 + Lum.	1.68	3.64
	LSE 3 + Lum.	xxx	xxx
	NN + Lum.	3.9	8.5
	CoCa	1.81	2.27
	Profile Maker	4.5	10.21

TABLE 4.8: Calcul du ΔE_{00} à partir de la mire *colorchecker classic*

grande, les résultats sont variables au point d'avoir des résultats aberrants (> 10 de ΔE_{00} représenté par "xxx") pour le polynôme le plus grand. Les résultats obtenus par le réseau de neurones sont au dessus de la norme et donc inexploitable. Nous notons que les résultats pour les autres mires sont elles aussi supérieures à 3.

La méthode proposée par le logiciel *CoCa* donne des résultats excellents : la moyenne est dans la norme et, surtout le ΔE maximum est très bas comparé aux autres méthodes. Le logiciel *Profile Maker* donne une précision assez surprenante sur la *colorchecker classic*, le ΔE est très élevé mais la visualisation de la mire montre des noirs proche de 5 alors que

les normes imposent des noirs aux alentours de 30. Les valeurs théorique ne concordent pas aux valeurs de la norme. En effet, les noirs sont des valeurs que nous maîtrisons avant calibration durant l'étape de balances des blancs nous les fixons. De ce fait, les valeurs noirs sont proches des valeurs de la mire et la transformation ne devrait pas modifier cette valeur. La transformation de Profile Maker modifie alors beaucoup la valeur. Nous en déduisons que les valeurs de référence ne sont pas celles utilisées. En conséquence, nous pouvons juger de la pertinence de la méthode utilisée par le logiciel *Profile Maker*.

échantillon	méthode	moyenne	maximum
1	LSE 2 + Lum.	2.65	7.94
	LSE3 + Lum.	2.38	7.23
	CoCa	xxx	xxx
	Profile Maker	2.06	5.16
2	LSE 2 + Lum.	2.63	9.4
	LSE3 + Lum.	2.66	11.7
	CoCa	xxx	xxx
	Profile Maker	1.98	5.9
3	LSE 2 + Lum.	2.14	11.74
	LSE3 + Lum.	2.02	10.05
	CoCa	xxx	xxx
	Profile Maker	2.02	5.93
4	LSE 2 + Lum.	2.25	12.5
	LSE3 + Lum.	2.3	8.71
	CoCa	xxx	xxx
	Profile Maker	2.18	8.76
5	LSE 2 + Lum.	2.09	9.3
	LSE3 + Lum.	2	6.11
	CoCa	xxx	xxx
	Profile Maker	1.87	4.26
6	LSE 2 + Lum.	2.1	7.06
	LSE3 + Lum.	2.05	4.23
	CoCa	xxx	xxx
	Profile Maker	1.84	4.83

TABLE 4.9: Résultats du ΔE_{00} à partir de la mire *colorchecker SG 140*

Dans le second tableau, nous donnons les résultats associés à la mire *colorchecker SG 140*. Nous rappelons que cette mire possède un nombre de patches plus important et qu'il y a des patches de noir ayant des densités très basse, ce qui peut avoir une influence sur la méthode de *Profile Maker*. Nous constatons que les méthodes ont des résultats différents. La méthode des moindres carrés du polynôme du degré 2 donne des résultats inférieure au polynôme de degré 3, effectivement, la moyenne est assez peu différente mais l'écart maximum est inférieure pour un polynôme de degré 3.

En ce qui concerne les logiciels, *CoCa* donne des résultats aberrants car il n'est pas possible de calibrer ce type de mire avec la version à disposition. *Profile Maker* donne

dans ce cas des résultats pertinents. La moyenne est aux alentours de 2 et la valeur maximal est souvent la plus faible des transformations.

De ces deux tableaux, nous en tirons notre première constatation à savoir que le contrôle de la luminance est primordiale pour obtenir des résultats optimaux. Par conséquent, la procédure de calibration est une étape de contrôle de la luminance à partir d'une mire composée d'un grand nombre de patches de niveaux de gris. Plus la mire est composée de niveaux de gris, plus nous serons précis dans l'étalonnage du gamma. Toutefois, utiliser la mire Q-13 pour calibrer le gamma et utiliser une autre mire pour calibrer la couleur ne nous paraît pas pertinent même si nous n'avons pas relevé d'incohérence.

Notre deuxième constatation est que les résultats obtenus à partir des méthodes des moindres carrés sont proches des résultats commerciaux mais que la différence se fait au niveau des ΔE maximums.

Et pour finir, l'utilisation du degré a bien un impact sur les résultats. Par conséquent, nous utilisons un polynôme de degré 2 pour les mires d'un faible nombre de patches et un polynôme de degré 3 pour les mires ayant un nombre de patches plus importants.

4.3.6.2 Analyse sur les coefficients de pondération de la méthode des moindres carrés pondérées

Nous rappelons que les normes industrielles ont des contraintes sur la moyenne du delta E, le maximum du delta E et aussi sur les niveaux de gris. Par conséquent, nous proposons de faire de nouveaux tests pour parvenir à valider l'ensemble des contraintes.

Pour modifier les ΔE maximums, nous proposons d'utiliser des coefficients de pondération dans le calcul de la transformation basée sur les moindres carrées. En partant du constat que les niveaux de gris ont été calibrés, nous émettons l'hypothèse que ces couleurs ont davantage d'importance et qu'il est nécessaire de les conserver. Nous suggérons ainsi une première approche naïve qui consiste à pondérer les patches de niveaux de gris par le coefficient 2 et les autres par le coefficient 0.5. Nos test ont été fait sur la mire *colorchecker* qui possède 24 patches. De ce fait, les coefficients 2 et 0.5 sont susceptibles d'être pertinents mais sur des mires possédant beaucoup plus de patches, ces coefficients peuvent être incorrects. Alors, nous suggérons une autre approche qui ne soit pas basé sur des coefficients arbitraires.

Notre seconde approche cherche à estimer un coefficient en proportion de l'erreur de la couleur. Nous convertissons les patches *RVB* en $L^*a^*b^*$ en passant par l'espace *XYZ*. Il nous permet d'établir les différentes distances entre les valeurs théoriques et les valeurs

calculées. Nous faisons une équivalence entre ces distances et des valeurs de pondération. Cette approche transformera davantage les patches dont la différence est plus grande que les plus proches. En l'occurrence, certains patches de couleur ont toujours des imprécisions à cause des défauts de l'illuminant utilisé.

L'intérêt ici est d'observer l'utilisation complémentaire de ces deux approches.

éch.	méthode	moyenne	maximum	écart-type	moy. (Gris)
1	sans pondération	1.76	5.87	1.2	1.84
	Gris(2)/Non Gris(0.5)	1.98	6.5	1.6	2.19
	erreur	1.81	6.05	1.37	2.18
2	sans pondération	2.12	5.01	1.06	3.32
	Gris(2)/Non Gris(0.5)	1.99	3.58	0.6	2.67
	erreur	2.25	6.23	1.4	3.62
3	sans pondération	1.45	4.53	1.05	1.76
	Gris(2)/Non Gris(0.5)	1.3	3.8	0.78	0.89
	erreur	1.6	5.1	1.2	2.2
4	sans pondération	1.53	4.33	0.9	1.98
	Gris(2)/Non Gris(0.5)	1.37	3.8	0.78	1
	erreur	1.6	4.8	1.23	4.8
5	sans pondération	1.6	3.43	0.7	1.67
	Gris(2)/Non Gris(0.5)	1.54	3.5	0.64	1.01
	erreur	1.72	4.2	1.02	2.02
6	sans pondération	1.58	3.25	0.7	2.01
	Gris(2)/Non Gris(0.5)	1.68	3.6	0.84	1.56
	erreur	1.82	4.35	1.39	3.17

TABLE 4.10: Utilisation de la pondération dans la méthode des moindres carrés permettant d'obtenir une minimisation de l'erreur en ΔE_{00} sur les patches de niveaux de gris

Nous observons les résultats dans le tableau 4.10 où nous comparons sans pondération et avec pondération. La première approche donne les résultats attendus, c'est-à-dire que nous avons des erreurs sur les niveaux de gris inférieure à la méthode sans pondération, par contre il n'était pas prévisible que le ΔE max diminue sur l'ensemble des cas.

La deuxième approche doit équilibrer le système, c'est-à-dire que la moyenne varie faiblement mais cela diminue la valeur maximale. En d'autres termes, les niveaux de gris perdent de leur précision aux profits des couleurs. Cela se vérifie dans les résultats, les niveaux des gris augmentent mais la valeur maximale augmente. Le problème est que cette valeur est un niveau de gris, cela signifie que la fonction mise en place est incorrecte. Il est intéressant de remarquer que le système est parfois pondéré de lui-même à cause des patches de couleur qui sont très proches les uns des autres. La mire SG140 possède plusieurs séries de patches définissant une région de couleur spécifique.

4.4 Conclusion

Dans ce chapitre, nous avons pu voir une partie importante des outils de la gestion de la couleur normée par l'International Color Consortium. Pour mettre en place un tel système, un espace de couleur de référence a été choisi pour sa capacité à mesurer les différences colorimétriques de manière objective, cette espace colorimétrique est l'espace $L^*a^*b^*$. Afin que tous les appareils puissent afficher les mêmes couleurs, il faut les calibrer en se basant sur une mire de référence.

La calibration est une transformation qui se décompose en deux parties. La première consiste à inverser le gamma en se basant sur les référence de niveaux de gris. La seconde partie consiste à convertir directement les informations de l'espace RVB dans l'espace $L^*a^*b^*$. Nous concluons que le modèle de transformation est la régression polynomiale minimisée par la méthode des moindres carrés. Nous avons pu mettre en évidence que le choix de la mire est non-négligeable dans l'étape de calibration puisqu'il n'existe pas de transformation parfaite pour toutes les mires. Malgré que la méthode des moindres carrés soit la plus pertinente, nous observons que le degré du polynôme joue sur la précision de la calibration.

Nos résultats sont équivalents par rapport les logiciels *CoCa* et *Profile Maker* mais la gestion du maximum est beaucoup mieux gérée par les logiciels de référence. Cette différence a une conséquence directe sur l'évaluation des normes. En effet, la norme FADGI impose que la valeur maximal du ΔE soit inférieure à 6 pour obtenir la meilleure note.

Nous avons mis en évidence la difficulté à évaluer la précision colorimétrique à partir d'une image dans un espace RVB spécifique. Les formats standards RVB créent des erreurs de précision colorimétrique. Ces erreurs rendent le calcul du ΔE incorrecte ne permettant pas d'avoir une analyse pertinente. Nous proposons de calculer le ΔE à partir de l'image d'acquisition avec uniquement le profil icc d'entrée.

Le contexte de notre problématique, qui consistait à obtenir les niveaux de gris les plus précis possibles, nous impose d'être rigoureusement précis sur les patches de niveaux de gris. Pour cela, nous utilisons de nombreux patches de niveaux de gris pour calibrer au mieux le gamma. Nous constatons aussi que l'utilisation des coefficients de pondération dans le modèle des moindres carrés nous permet de conserver notre bonne précision mais que nous sommes limités dans notre modèle de contrainte.

Les perspectives qui s'offrent à nous sont basés sur les méthodes de résolution linéaire ou non-linéaire. Concernant les méthodes linéaires, la fonction de pondération doit être

amélioré. Le deuxième point concerne l'espace d'arrivée de la transformation, la transformation passe de l'espace $R'V'B'$ vers $L^*a^*b^*$, or le changement de l'espace de couleur de XYZ vers $L^*a^*b^*$ utilise une fonction puissance. Cela implique que le polynôme de degré n ne modélise pas parfaitement le changement d'espace. Concernant les méthodes non-linéaires, nous avons vu au cours de la thèse qu'ils existent des fonctions respectant des lois psycho-visuelles comme les fonctions logarithme et exponentielle par conséquent, il s'avère intéressant de modéliser ces fonctions. De manière plus large, une étude exhaustive peut être faite sur d'autres méthodes linéaire/non-linéaires permettant la résolution de ce type de problème.

Chapitre 5

Conclusion

L'objectif de la société i2S est de mettre en place un "scanner sans éclairage" ainsi que de proposer des modules qui ont pour but de favoriser la mise en valeur des documents numérisés. Pour cela, il nous faut améliorer la qualité de numérisation. Cette thèse s'emploie à résoudre notamment trois paramètres liés à cet objectif : la correction d'éclairage, la mesure de netteté de numérisation et la calibration colorimétrique.

Le premier chapitre énumère différents outils de traitement d'images permettant de corriger les défauts d'illumination. Nous avons observé qu'ils existent deux approches. La première est la décomposition en luminance-réfectance de l'image consistant à supprimer l'information de luminance (filtre homomorphe [Jain, 1989], le Retinex ([Land, 1964] la labélisation des pixels ([Tappen et al., 2005])). L'autre approche est la détection du fond d'un document a priori homogène. La méthode la plus pertinente pour détecter le fond est la binarisation adaptative ([Sauvola and Pietikäinen, 2000]) qui va labéliser les pixels en fond/non fond. Une fois les pixels identifiés, il faut appliquer un coefficient de correction que nous avons discuté.

Nous proposons une alternative pour identifier les pixels du fond en prenant en compte les défauts d'illumination. Pour cela, nous sélectionnons nos pixels en se basant sur deux propriétés. La première compare le pixel courant à la moyenne de la fenêtre carrée centrée sur lui. Les deux valeurs obtenues doivent être approximativement égales. La seconde est un seuil d'écart-type sur la fenêtre qui nous permet d'exclure des cas particuliers. Nous segmentons les pixels par une méthode de split and merge. Cela nous permet d'exclure les erreurs et d'identifier différents fonds si il y en a plusieurs. La suite de l'algorithme va remplir les zones de l'image qui ne sont pas définies, puis d'y appliquer la transformation. Le gain que nous avons choisi est calculé par la multiplication classique et la valeur du

fond à cibler. Cette dernière peut être choisie par l'utilisateur ou nous lui proposons d'estimer une valeur en fonction de l'image du fond trouvée via une interface.

Ces résultats sont visibles dans l'annexe. La fonction a été rajoutée dans le produit Limb et va être rajoutée sur le moyen terme dans le "scanner sans éclairage". Le temps d'exécution de l'application et la transformation sont correctes. Toutefois le temps d'exécution ne peut être utilisé dans un scanner car elle est aussi longue que l'acquisition. L'amélioration se situe au niveau technique afin de réduire le temps d'exécution, il faudra alors le coder afin que la fonction puisse tourner sur les processeurs graphique ou encore d'optimiser l'algorithme. Nous pouvons par exemple remettre en cause l'usage de la segmentation pour tous les documents.

Le deuxième chapitre traite de la netteté de numérisation. Nous avons évoqué les mesures faites par les normes industrielles et nous avons conclu que ces critères sont nombreux et qu'ils se réfèrent à la fonction du transfert du module (FTM).

Notre analyse du flou quantifie la netteté d'une image, sans repère et après acquisition. Pour mettre en place notre solution, nous avons utilisé certaines méthodes complémentaires vues dans l'état de l'art. Ces descripteurs sont :

- le maximum du gradient
- la catégorisation des ondelettes de Haar (cf. [Tong et al., 2004]) qui classifie les contours nets et les contours flous
- l'estimation de netteté de l'autofocus (cf. [Zhuo and Sim, 2011]), c'est une mesure de la dégradation d'un lissage gaussien.

De notre côté, nous avons proposé :

- l'histogramme des épaisseurs des contours qui évalue la netteté des contours sur l'ensemble de l'image
- la binarisation de Fourier où nous proposons de quantifier les hautes fréquences par rapport à un seuil dans l'espace de Fourier.

A partir de cet ensemble de descripteur, nous utilisons un apprentissage supervisé, construit sur un lissage gaussien, qui permet d'évaluer la netteté. Les résultats sont convaincants et nous obtenons une bonne estimation. Afin de faire le lien entre les normes et notre estimation, nous avons évalué la mesure obtenue sur un ensemble de numérisation avec une mire permettant de calculer la FTM. Nos résultats sont plutôt concluants dans la mesure où le comportement de notre estimateur suit le comportement de la FTM. L'état de l'art propose un nombre important de méthode d'analyse et nous avons constaté que parmi les descripteurs utilisés, certains n'étaient pas suffisamment précis. Il serait intéressant de comparer les performances par catégorie de descripteur afin de trouver un équilibre entre performance et précision.

Ces travaux ont été présentés à la conférence du CIFED 2014. Certains descripteurs sont utilisés dans le logiciel LIMB mais la partie sur l'apprentissage supervisé est encore en développement à cause de temps de performance trop long.

Concernant ce projet, l'étape d'optimisation doit être davantage travailler. Il n'est pas possible d'avoir une méthode d'analyse aussi longue que la numérisation.

Le dernier chapitre aborde les problématiques de calibration colorimétrique du scanner. L'objectif est de créer un fichier ".icc" qui permet de convertir l'information RVB de la caméra dans l'espace de couleur de référence $L^*a^*b^*$.

Le problème de l'entreprise est que les contraintes sur la précision couleur (le ΔE) sont importantes, nous avons énuméré les contraintes industrielles de numérisation liées à la couleur. Nous en concluons que la représentation des niveaux de gris est importante.

La calibration se découpe en deux étapes. La première consiste à enlever le gamma de la mire colorimétrique permettant la calibration. Pour optimiser cette opération, nous utilisons une mire annexe possédant un plus grand nombre de patch. L'estimation du gamma est logiquement plus précise.

La seconde étape est la transformation de $R'V'B'$ (ie. sans gamma) à $L^*a^*b^*$. La méthode la plus pertinente est d'utiliser la méthode des moindres carrés afin de corriger les erreurs de mesure. toutefois, le modèle varie en fonction du nombre de patch de la mire. Concernant nos deux mires de référence : la mire colorchecker classique est modélisé par un polynôme de degré 2 et la mire colorchecker SG 140 est modélisé par un polynôme de degré 3. Les résultats obtenus sont relativement proche par rapport aux logiciels de référence (Coca et Profile Maker) mais la différence se situe au niveau du ΔE maximal. Ce critère fait que notre approche est moins bonne concernant le critère de la FADGI.

Les pistes sont assez nombreuses pour palier à ce problème. Nous avons abordé la solution de la pondération du modèle des moindres carrés, les résultats sont encourageants mais ne sont pas aboutit. Nous pensons aussi que notre problème peut-être posé de manière différente en utilisant un modèle mathématique différent, notamment en utilisant une fonction logarithmique qui valide l'aspect psycho-visuel des transitions de couleur dans le cas d'un modèle non-linéaire. Un autre argument conteste le modèle choisi car le changement d'espace de couleur XYZ à $L^*a^*b^*$ utilisé par une fonction puissance qui conteste le modèle mathématique pour résoudre le problème.

Le projet de calibration couleur est toujours en développement. L'application faite propose une partie création d'icc (détection, sélection des patches, plusieurs modèles sont

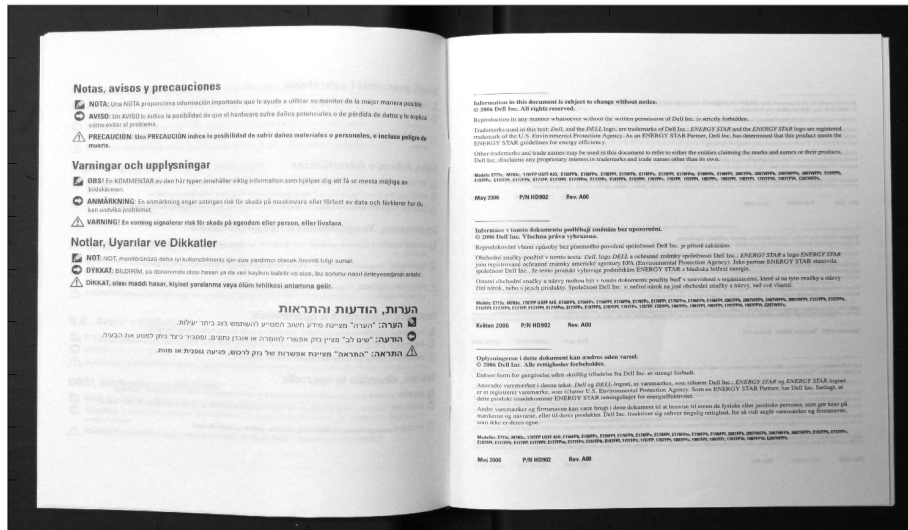
implémentés en fonction du gamma et du modèle mathématique de la transformation) et une partie analyse (gamma, ΔE).

Chapitre 6

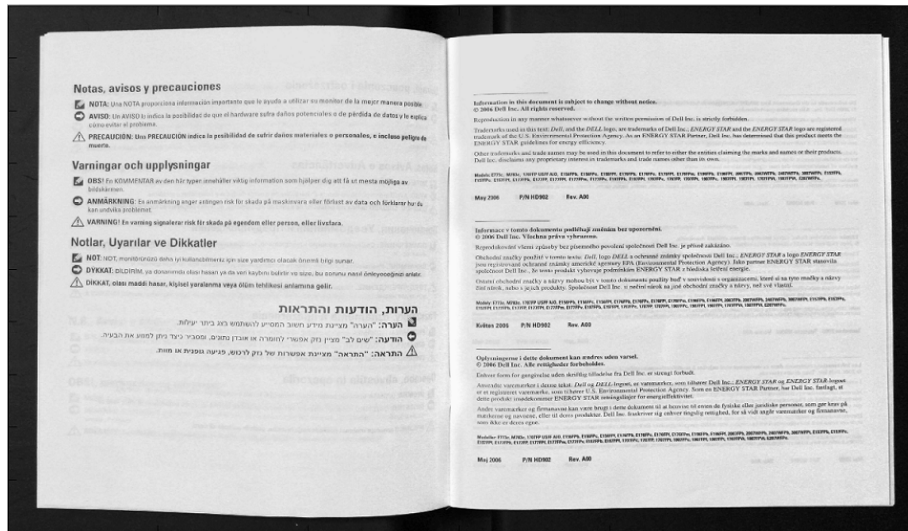
Annexes

6.1 Exemples de la correction d'éclairage

6.1.1 Exemples utilisés dans le chapitre

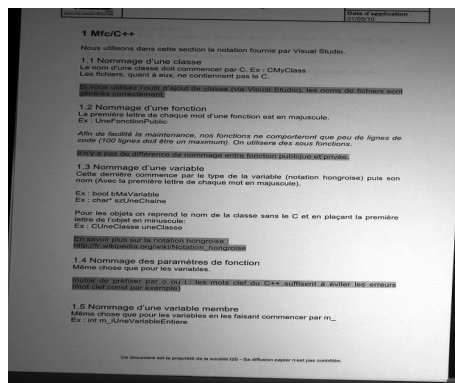


a) Numérisation originale

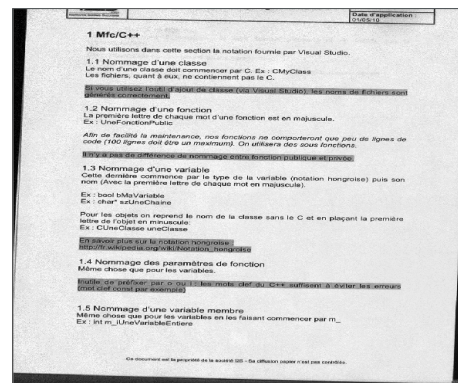


b) Correction

FIGURE 6.1: Correction du manuel d'utilisation utilisé dans les illustrations pour les méthodes de l'état de l'art



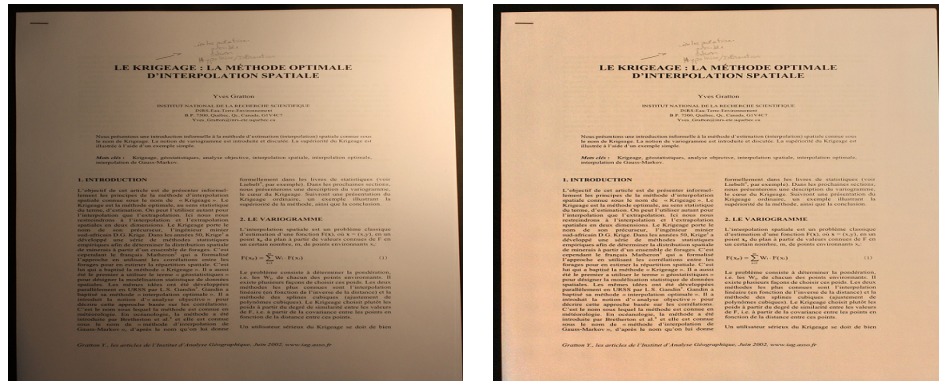
a) Numérisation originale



b) Correction

FIGURE 6.2: Correction de la page à forte variation d'illumination

6.1.2 Exemples

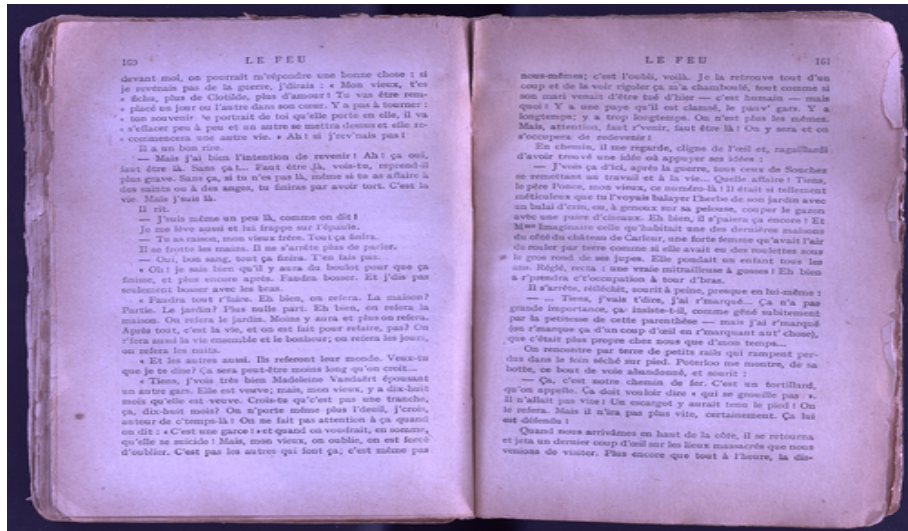


a) Numérisation originale

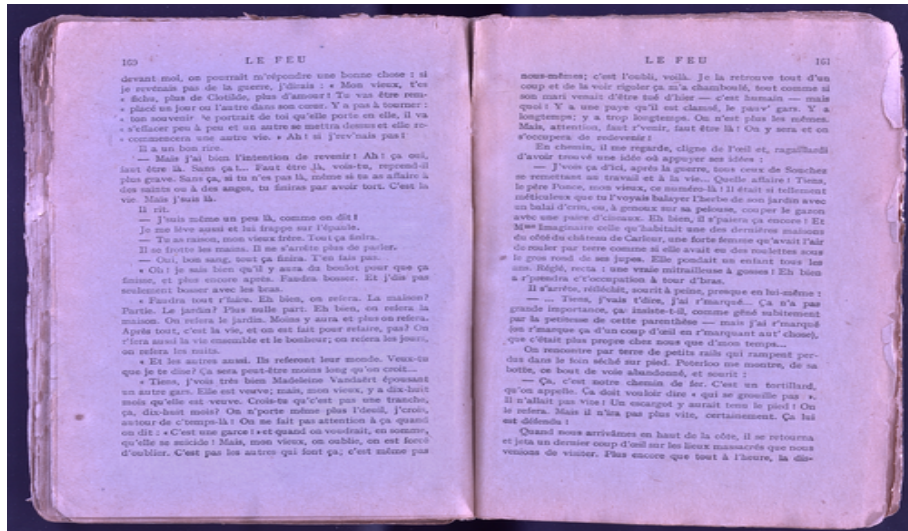
b) Correction

FIGURE 6.3: Correction d'une page à forte variation d'illumination

6.1.3 Exemple d'un livre ancien



a) Numérisation originale



b) Correction

FIGURE 6.4: Correction d'un livre ancien ayant des taches sur le fond du livre

6.1.4 Exemple de livres récents



a) Numérisation originale



b) Correction

FIGURE 6.5: Correction d'un livre récent



a) Numérisation originale



b) Correction

FIGURE 6.6: Deuxième correction d'un livre récent

6.1.5 Exemple détaillé d'une erreur de correction d'éclairage

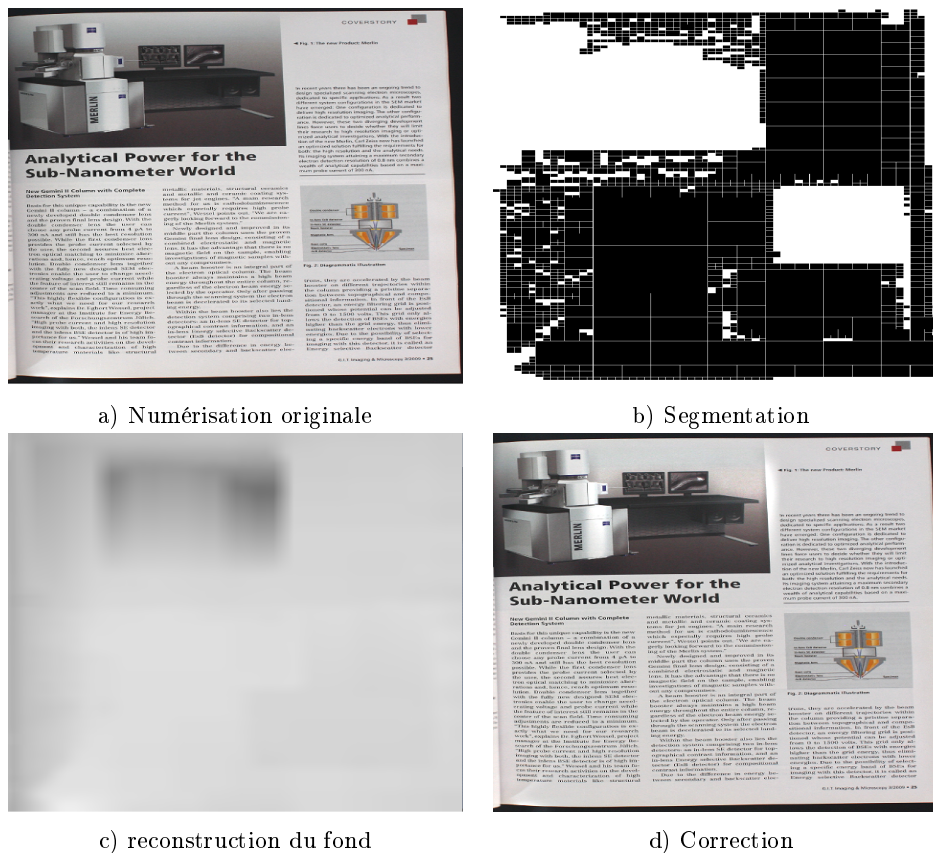
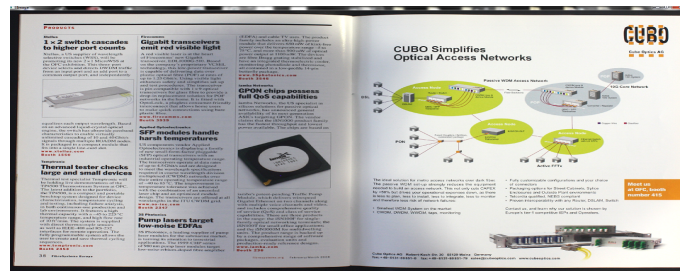
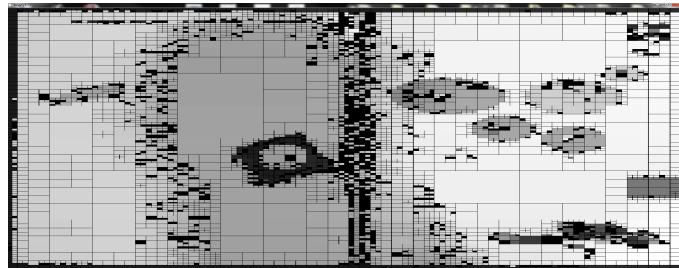


FIGURE 6.7: Exemple détaillé d'une erreur de segmentation du fond [2.3.6]

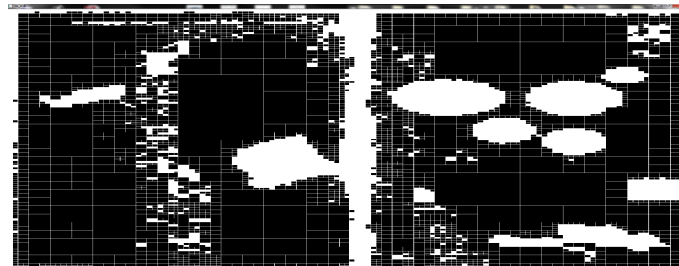
6.1.6 Exemple détaillé d'une suppression d'ombre



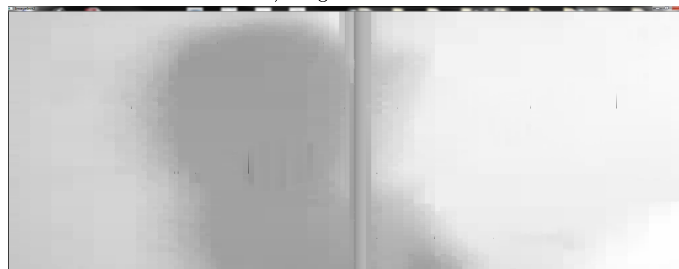
a) Numérisation originale



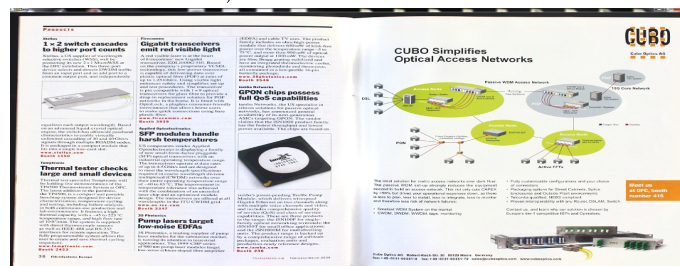
b) Quadtree



b) Segmentation



a) reconstruction du fond



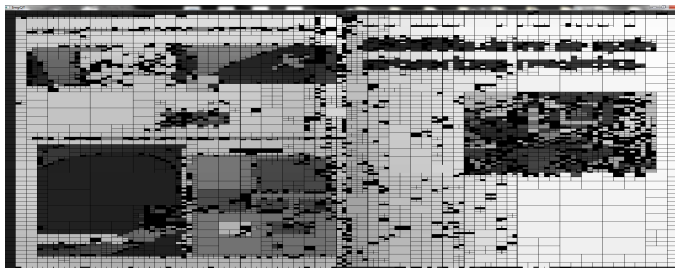
b) Correction

FIGURE 6.8: Exemple détaillé d'une suppression de fond [2.3.5]

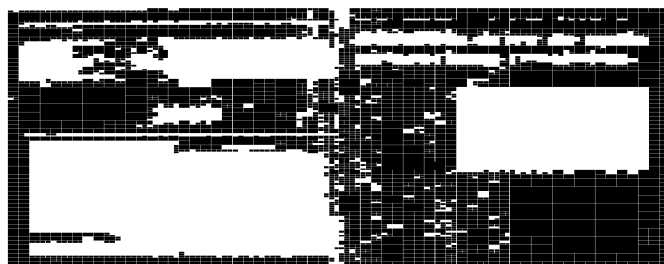
6.1.7 Exemple détaillé de correction d'éclairage de livre récent



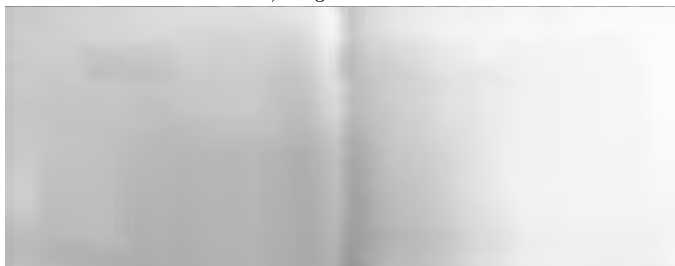
a) Numérisation originale



b) Quadtree



b) Segmentation



a) reconstruction du fond

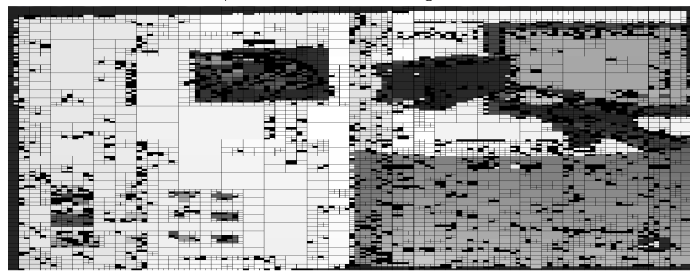


b) Correction

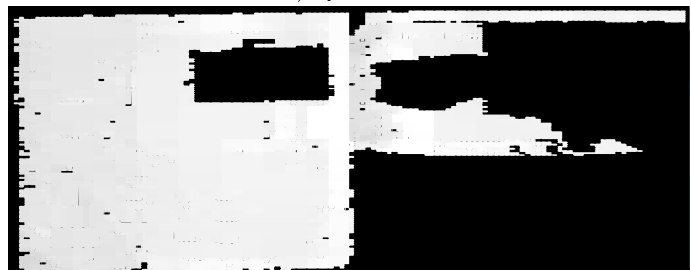
FIGURE 6.9: Exemple détaillé d'une suppression de fond



a) Numérisation originale



b) Quadtree



b) Segmentation



a) reconstruction du fond



b) Correction

FIGURE 6.10: Exemple détaillé d'une suppression de fond [2.4]

6.2 Tableau des résultats concernant la calibration couleur

6.2.1 Comparaison des conversions $sRGB$ [4.3.5]

RVB			fichier icc officiel			conversion <i>Lindbloom</i>			conversion <i>Imatest</i>		
0.49	0.32	0.26	40.53	17.74	17.5	39.39	17.32	16.51	39.25	16.2	16.17
0.77	0.54	0.49	64.65	22.34	18.35	62.93	21.38	16.49	62.66	20.24	16.01
0.4	0.49	0.59	50.57	-2.33	-17.91	51.4	-3.72	-17.17	51.63	-1.82	-16.87
0.37	0.46	0.23	44.7	-13.13	27.36	46.21	-16.11	29.16	46.17	-18.48	29.38
0.55	0.51	0.68	56.94	11.87	-19.98	56.46	11.38	-21.9	56.63	13.51	-21.75
0.5	0.79	0.67	71.49	-25.58	4.07	75.61	-29.53	8.14	75.65	-30.17	8.47
0.87	0.47	0.18	62.98	39.63	61.08	61.85	36.86	56.26	61.21	34.48	55.5
0.32	0.37	0.66	40.94	11.7	-41.41	41.71	11.68	-42.3	42.27	16.61	-41.63
0.79	0.32	0.35	52.7	48.38	23.12	51.27	48.89	20.71	50.74	48.2	19.61
0.38	0.22	0.42	31.46	26.1	-18.02	29.97	26.32	-22.64	30.25	28.34	-22.74
0.67	0.77	0.25	72.9	-17.37	61.06	75.54	-22.71	60.17	75.26	-27.09	60.76
0.87	0.65	0.2	73.38	20.58	71.52	72.15	14.3	64.11	71.6	10.77	63.99
0.19	0.29	0.55	30.37	9.06	-42.59	31.92	8.87	-41.61	32.56	14.06	-40.82
0.34	0.62	0.27	55.77	-36.19	35.37	59.32	-36.99	38.46	59.31	-40.41	39.12
0.72	0.2	0.22	43.05	54.38	31.54	43.04	54.27	30.61	42.39	53.58	29.25
0.9	0.78	0.18	81.09	9.85	83.04	80.7	0.53	72.8	80.21	-3.82	73.11
0.78	0.32	0.61	53.12	53.36	-9.37	52.47	52.86	-16.02	52.2	54.23	-16.66
0.28	0.58	0.67	52.78	-20.02	-26	57.62	-19.77	-19.39	57.92	-17.13	-18.88
0.96	0.97	0.95	96.71	-0.15	2.76	96.82	-1.32	2.2	96.8	-1.53	2.19
0.79	0.79	0.77	80.98	0.59	2.83	81.22	-0.31	2.35	81.19	-0.53	2.33
0.65	0.64	0.61	67.35	0.79	4.32	67.29	-0.32	3.96	67.24	-0.68	3.95
0.51	0.49	0.46	52.67	2.55	4.4	52.33	1.42	3.84	52.28	1.08	3.72
0.35	0.33	0.32	37.08	2.3	3.07	36.56	1.39	2.59	36.48	1.16	2.53
0.2	0.18	0.18	21.33	2.5	1.97	19.77	1.88	1.5	19.64	1.73	1.46

TABLE 6.1: Conversion de RVB en $L^*a^*b^*$ selon le modèle standard

$\Delta E00(\text{sRGB}, \text{Lindbloom})$	$\Delta E00(\text{sRGB}, \text{Imatest})$	$\Delta E00(\text{Lindbloom}, \text{Imatest})$
1.14	1.53	0.71
1.76	2.21	0.66
1.69	1.29	1.9
2.22	3.21	1.27
1.33	1.28	1.61
4.31	4.54	0.32
1.26	3.34	2.88
3.41	3.63	1.18
0.36	1.05	0.78
5.61	8.68	2.54
4.26	4.52	0.52
5.78	5.99	1.6
1.65	2.47	1.12
0.76	2.82	2.7
1.83	2.5	0.71
3.14	2.76	1.11
3.27	4.78	1.85
3.48	5.58	2.25
1.75	2.03	0.29
1.39	1.7	0.32
1.62	2.14	0.51
1.56	2.04	0.48
1.33	1.65	0.33
1.41	1.6	0.22

TABLE 6.2: Tableau des calculs du $\Delta E00$ entre les différentes conversions

6.2.2 L'icc *sRGB* officiel [4.3.5]

$L^*a^*b^*$			$L^*a^*b^* \rightarrow RVB$			$RVB \rightarrow L^*a^*b^*$			$\Delta E00$
38.58	17.45	16.86	0.49	0.32	0.26	38.49	17.59	17	0.12
65.15	21.91	17.77	0.77	0.54	0.49	65.13	22.08	17.69	0.12
49.31	-2.4	-19.63	0.4	0.49	0.59	49.38	-2.5	-19.6	0.12
42.65	-13.38	26.55	0.37	0.46	0.23	42.74	-13.51	26.69	0.12
55.85	11.76	-22.99	0.55	0.51	0.68	55.85	11.6	-23.13	0.16
71.72	-31.22	2.75	0.5	0.79	0.67	71.69	-31.32	3.05	0.21
62.06	38.58	60.19	0.87	0.47	0.18	62.2	38.88	60.67	0.18
39.76	11.3	-43.23	0.32	0.37	0.66	39.87	11.04	-43.25	0.17
52.5	47.98	22.38	0.79	0.32	0.35	52.4	47.9	22.55	0.13
30.52	25.46	-18.73	0.38	0.21	0.42	30.47	25.15	-18.91	0.27
72.55	-19.35	60.55	0.67	0.77	0.25	72.49	-19.52	60.47	0.1
73.61	18.96	73.18	0.87	0.65	0.2	73.48	19.1	73.02	0.15
29.48	10.58	-44.73	0.19	0.29	0.55	29.58	10.5	-44.74	0.09
54.87	-37.44	33.83	0.34	0.62	0.27	54.89	-37.28	33.62	0.08
41.33	54.05	30.32	0.72	0.2	0.22	41.34	54.24	30.64	0.12
81.29	8.4	81.32	0.9	0.78	0.18	81.11	8.35	81.31	0.13
53.2	52.12	-11.06	0.78	0.32	0.61	53.15	52.21	-10.86	0.15
51.85	-20.42	-28.49	0.28	0.59	0.67	51.92	-20.37	-28.36	0.09
96.29	0.21	1.8	0.96	0.97	0.95	96.2	0.13	1.9	0.16
81.76	0.02	2.36	0.79	0.79	0.77	81.77	-0.16	2.56	0.32
67.45	0.17	3.15	0.65	0.64	0.61	67.5	-0.03	3.06	0.3
51.2	2.3	3.13	0.51	0.49	0.46	51.24	2.16	3.29	0.25
35.59	2.51	2.04	0.35	0.33	0.32	35.56	2.37	2.27	0.29
21.3	2.99	2.42	0.2	0.18	0.18	21.33	2.76	2.74	0.42
Erreur moyenne :									0.18

TABLE 6.3: Tableau des calculs du $\Delta E00$ entre les différentes conversions de *sRGB.icc*

6.2.3 Comparaison en delta E des mires old et New d'imatest avec la référence de la mire utilisée [4.3.5]

Colorchecker 2005			Imatest new chart			$\Delta E00$
37.54	12.02	13.33	38.08	12.09	14.39	1.22
65.2	14.82	17.54	66.38	13.22	17.14	2.2
50.37	-1.57	-21.43	51.06	0.38	-22.06	2.81
43.13	-14.63	22.12	43.3	-16.52	21.46	0.95
55.34	11.45	-25.29	56.36	12.84	-25.29	1.84
71.36	-32.72	1.64	71.6	-30.71	1.17	1.35
61.37	32.89	55.16	61.7	27.54	58.23	1.66
40.71	16.91	-45.09	41.22	17.95	-43.16	3.42
49.86	45.93	13.88	51.57	43	14.75	1.73
30.15	24.92	-22.61	30.77	25.74	-23.38	0.98
72.44	-27.46	58.47	72.58	-28.17	58.61	1.51
70.92	15.58	66.54	71.6	12.45	66.58	2.26
29.62	21.43	-49.03	29.73	28.47	-51.99	3.68
55.64	-40.76	33.27	55.65	-41.52	33.73	1
40.55	49.97	25.46	41.22	50.85	25.86	1.86
80.98	-1.04	80.03	81.35	-4.13	79.25	2.95
51.01	49.88	-16.93	51.57	48.98	-15.96	1.79
52.12	-24.61	-26.18	51.57	-19.85	-24	2.13
96.54	-0.69	1.35	96	-0.06	0.06	0.4
81.27	-0.61	-0.24	81.35	-0.05	0.06	0.11
66.79	-0.65	-0.43	66.67	-0.04	0.05	0.13
50.87	-0.06	-0.25	51.57	-0.03	0.04	0.14
35.68	-0.22	-1.21	35.99	-0.03	0.03	0.29
20.48	0.05	-0.97	20.54	-0.02	0.02	0.19
Erreur moyenne :						1.52

TABLE 6.4: Calcul du $\Delta E00$ en fonction de la mire de référence d'*imatest* (new) et de la référence actuelle

6.3 Détails des calculs du deltaE

6.3.1 ΔE CMC [4.2.3]

Soit

$$S_L = \begin{cases} 0.511 & L_1^* < 16 \\ \frac{0.040975L_1^*}{1+0.01765L_1^*} & L_1^* \geq 16 \end{cases}, S_C = \frac{0.0638C_1^*}{1+0.0131C_1^*} + 0.638$$

$$S_H = S_C(FT + 1 - F), F = \sqrt{\frac{C_1^{*4}}{C_1^{*4} + 1900}}$$

Colorchecker 2005			Imatest old chart			$\Delta E00$
37.54	12.02	13.33	37.99	13.56	14.06	1.24
65.2	14.82	17.54	65.71	18.13	17.81	3.31
50.37	-1.57	-21.43	49.93	-4.88	-21.93	4.54
43.13	-14.63	22.12	43.14	-13.1	21.91	2.21
55.34	11.45	-25.29	55.11	8.84	-25.4	3
71.36	-32.72	1.64	70.72	-33.4	-0.2	1.6
61.37	32.89	55.16	62.66	36.07	57.1	4.6
40.71	16.91	-45.09	40.02	10.41	-45.96	4.34
49.86	45.93	13.88	51.12	48.24	16.25	1.78
30.15	24.92	-22.61	30.33	22.98	-21.59	1.42
72.44	-27.46	58.47	72.53	-23.71	57.26	1.79
70.92	15.58	66.54	71.94	19.36	67.86	4.07
29.62	21.43	-49.03	28.78	14.18	-50.3	6.51
55.64	-40.76	33.27	55.26	-38.34	31.37	1.25
40.55	49.97	25.46	42.1	53.38	28.13	1.26
80.98	-1.04	80.03	81.73	4.04	79.82	4.63
51.01	49.88	-16.93	51.94	49.99	-14.57	1.16
52.12	-24.61	-26.18	51.04	-28.63	-28.64	4.22
96.54	-0.69	1.35	96.54	-0.43	1.19	1.26
81.27	-0.61	-0.24	81.26	-0.64	-0.34	0.95
66.79	-0.65	-0.43	66.77	-0.73	-0.5	1.14
50.87	-0.06	-0.25	50.87	-0.15	-0.27	0.78
35.68	-0.22	-1.21	35.66	-0.42	-1.23	1.38
20.48	0.05	-0.97	20.46	-0.08	-0.97	0.97
Erreur moyenne :						2.48

TABLE 6.5: Calcul du $\Delta E00$ en fonction de la mire de référence d'*imatest* (old) et de la référence actuelle

$$T = \begin{cases} 0.56 + |0.2 \cos(h_1 + 168^\circ)| & 164^\circ \leq h_1 \leq 345^\circ \\ 0.36 + |0.4 \cos(h_1 + 35^\circ)| & \text{sinon} \end{cases}$$

6.3.2 $\Delta E 94$ [4.2.3]

Soit

$$C_1 = \sqrt{a_1^2 + b_1^2}, C_2 = \sqrt{a_2^2 + b_2^2}, \Delta a = a_1 - a_2, \Delta b = b_1 - b_2$$

$$S_L = 1, S_C = 1 + K_1 C_1, S_H = 1 + K_2 C_1$$

$$K_L = \begin{cases} 1 & \text{par défaut} \\ 2 & \text{textiles} \end{cases}, K_C = 1 \text{ par défaut}, K_H = 1 \text{ par défaut}$$

$$K_1 = \begin{cases} 0.045 & \text{arts graphiques} \\ 0.048 & \text{textiles} \end{cases}, K_2 = \begin{cases} 0.015 & \text{arts graphiques} \\ 0.014 & \text{textiles} \end{cases}$$

6.3.3 ΔE 00 [4.2.3]

Soit

$$\overline{C_{ab}^*} = \frac{C_{1,ab}^* + C_{2,ab}^*}{2} \quad (6.1)$$

$$G = 0.5 \left(1 - \left(1 - \sqrt{\frac{\overline{C_{ab}^*}^7}{\overline{C_{ab}^*}^7 + 25^7}} \right) \right) \quad (6.2)$$

$$a'_i = (1 + G)a_i^* , C'_i = \sqrt{(a'_i)^2 + (b_i^*)^2} \quad (6.3)$$

$$h' = \begin{cases} 0 & b_i^* = a'_i \\ \arctan(b_i^*, a'_i) & \text{sinon} \end{cases} \quad i = 1, 2 \quad (6.4)$$

Nous pouvons ensuite calculer,

$$\Delta h' = \begin{cases} 0 & C'_1 C'_2 = 0 \\ h'_2 - h'_1 & C'_1 C'_2 \neq 0; |h'_2 - h'_1| \leq 180^\circ \\ (h'_2 - h'_1) - 360 & C'_1 C'_2 \neq 0; (h'_2 - h'_1) > 180^\circ \\ (h'_2 - h'_1) + 360 & C'_1 C'_2 \neq 0; (h'_2 - h'_1) < -180^\circ \end{cases} \quad (6.5)$$

$$\Delta H' = 2\sqrt{C'_1 C'_2} \sin\left(\frac{\Delta h'}{2}\right) \quad (6.6)$$

$$\overline{L'} = (L_1^* + L_2^*)/2 , \overline{C'} = (C'_1 + C'_2)/2 \quad (6.7)$$

$$h' = \begin{cases} \frac{h'_1 + h'_2}{2} & |h'_1 - h'_2| \leq 180^\circ; C'_1 C'_2 \neq 0 \\ \frac{h'_1 + h'_2 + 360}{2} & |h'_1 - h'_2| > 180^\circ; (h'_1 + h'_2) < 360^\circ; C'_1 C'_2 \neq 0 \\ \frac{h'_1 + h'_2 - 360}{2} & |h'_1 - h'_2| > 180^\circ; (h'_1 + h'_2) \geq 360^\circ; C'_1 C'_2 \neq 0 \\ h'_1 + h'_2 & C'_1 C'_2 = 0 \end{cases} \quad (6.8)$$

$$T = 1 - 0.17 \cos(\overline{h'} - 30^\circ) + 0.24 \cos(2\overline{h'}) + 0.32 \cos(3\overline{h'} + 6^\circ) - 0.2 \cos(4\overline{h'} - 63^\circ) \quad (6.9)$$

$$\Delta\Theta = 30 \exp\left(-\left[\frac{\overline{h'} - 275^\circ}{25}\right]^2\right) \quad (6.10)$$

$$R_C = 2\sqrt{\frac{\overline{C'}^7}{\overline{C'}^7 + 25^7}}, S_L = 1 + \frac{0.015(\overline{L'} - 50)^2}{\sqrt{20 + (\overline{L'} - 50)^2}} \quad (6.11)$$

$$S_C = 1 + 0.045\overline{C'}, S_H = 1 + 0.015\overline{C'}T, R_T = -\sin(2\Delta\Theta)R_c \quad (6.12)$$

6.4 Les logiciels

Dans ce chapitre, nous allons présenter l'ensemble des logiciels qui ont été produit durant ces trois années.

6.4.1 La correction d'illumination

Le premier logiciel est celui de la correction d'illumination dont nous présentons l'interface dans la figure 6.11.

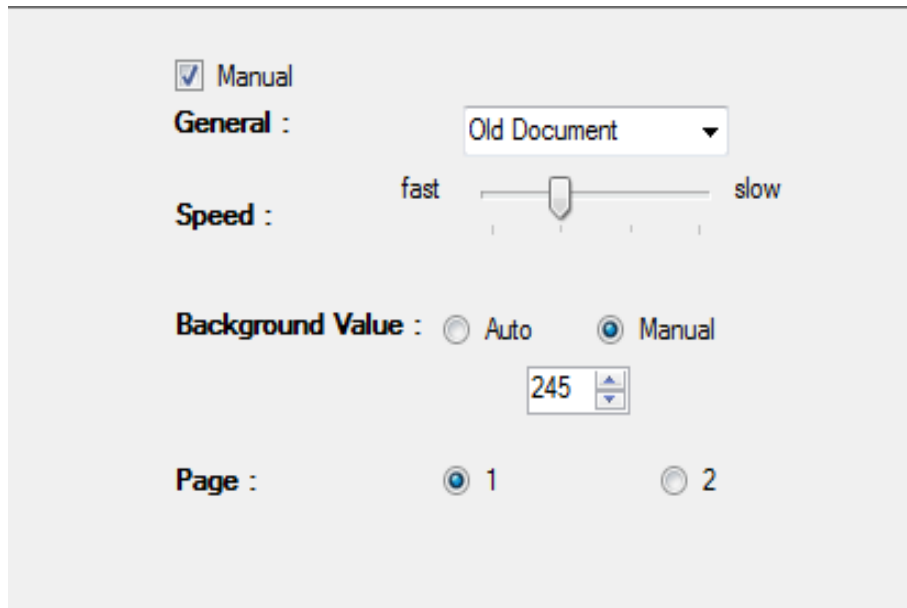


FIGURE 6.11: Interface de la correction d'éclairage

Le logiciel se veut d'avoir une fonctionnalité automatique et manuelle. Concernant la fonctionnalité manuelle, nous demandons à l'utilisateur de renseigner différents paramètres assez simple. L'argument "general" a trois arguments : Document ancien, Document récent, Suppression d'ombre. La méthode Document ancien va corriger uniquement les régions de niveaux de gris que nous avons vu dans la section analyse du fond. La méthode Document nouveau va appliquer l'analyse du fond tel que nous l'avons définie. Les paramètres de segmentation de la méthode document ancien seront supérieure aux

paramètres de la méthode document récent car la quantité de bruit est non-négligeable. Et enfin, la méthode suppression d'ombre va modifier le critère de continuité que nous avons définie dans l'étape du merge. Ce critère va fusionner abusivement les régions. Nous pouvons voir à l'annexe [6.8](#) un exemple de correction d'ombre.

L'argument "speed" régule les critères de redimensionnement de l'image et la taille du filtre utilisée pour le calcul des points caractéristiques. Comme nous l'avons dit dans le premier chapitre, la complexité de l'image va déterminer le temps d'exécution de la méthode, l'objectif est de jouer sur cet aspect.

L'argument "Background Value" va déterminer la valeur du fond. Nous pouvons soit donner une valeur en niveaux de gris, soit utiliser une approche automatique qui reprend les méthodes citées dans le premier chapitre.

Le dernier argument "Page" demande à l'utilisateur de renseigner le nombre de page.

La méthode automatique va paramétrer les critères en fonction de la taille de l'image, la détection de la reliure et en imposant certains paramètres.

6.4.2 Le flou de mise au point

L'interface du deuxième logiciel est présentée dans la figure 6.12

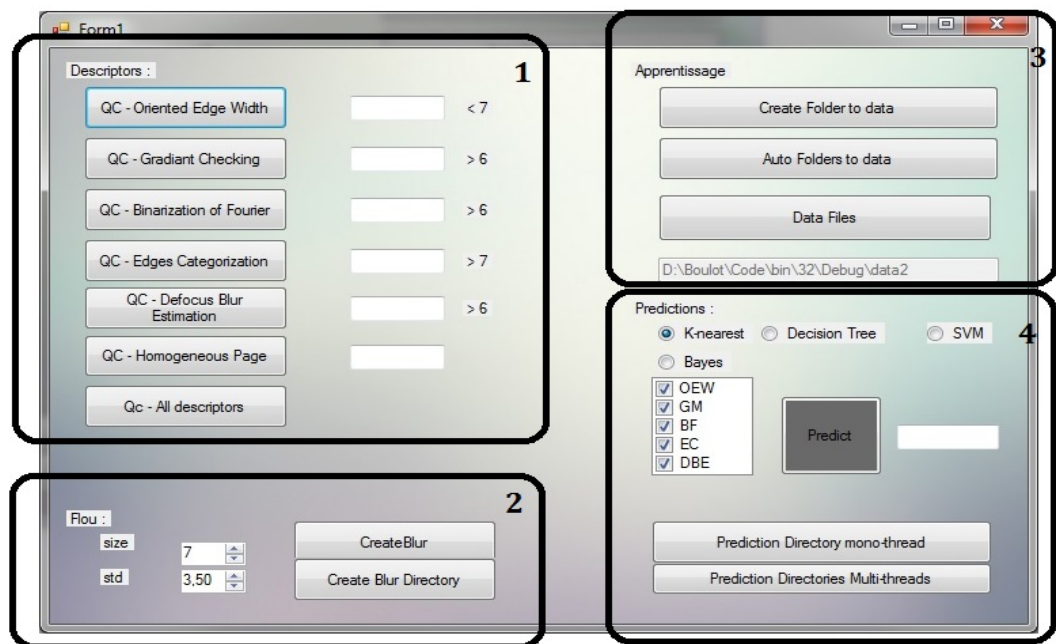


FIGURE 6.12: Interface de la mesure du flou de mise au point

Le logiciel se découpe en quatre partie.

La première est la partie de gauche où nous avons tous les descripteurs étudiés. Nous avons normalisé le résultat entre 0 et 10 et donné un indicateur naïf à côté. Nous retrouvons dans l'ordre l'épaisseur des contours, le maximum du gradient, la binarisation de fourier, la catégorisation des contours à partir des ondelettes de haar et le defocus blur estimation. Nous avons rajouté aussi une fonctionnalité permettant de détecter les pages homogènes qui est un contrôle fait systématiquement dans le logiciel limb pour écarter les images sans information avant la mesure de qualité de netteté.

La partie 2 permet de créer des images floues. Nous pouvons paramétrer le filtre puis créer une nouvelle image en fonction des ces paramètres ou l'appliquer sur toutes les images d'un dossier.

La partie 3 est composée des trois boutons. Le premier bouton sert à calculer les descripteurs de toutes les images d'un dossier et de l'écrire dans un fichier texte, le second bouton va faire la même chose pour tous les dossiers inclus dans le dossier sélectionné, cela va créer un nouveau répertoire avec un fichier texte pour chaque sous-dossier traité.

Le dernier bouton va récupérer le chemin du dossier de la base de donnée pour pouvoir apprendre la base.

La partie en bas à droite est la partie de prédiction. Nous pouvons sélectionner la méthode d'apprentissage : les k plus proches voisins, les arbres de décision, les machines à vecteurs de support ou la méthode de Bayes. Nous pouvons cocher ou décocher le ou les descripteurs que nous souhaitons prendre en compte. Les trois boutons servent à mettre en application la fonction. Predict le fait sur une image et les autres sur des dossiers.

6.4.3 La calibration couleur

Le logiciel de la calibration couleur est en plusieurs étapes. Il est prédestiné à un usage technique et l'objectif est de faire évoluer l'utilisateur pas à pas. L'interface de départ est représentée dans la figure 6.13. Dans la partie de gauche, nous avons mis les différents pas à faire de l'image de calibration jusqu'à l'application des ICCs ; en bas, nous avons rajouté les calculs du delta E.

Dans la partie de droite, soit l'étape d'Image de calibration, il y a un bouton pour sélectionner l'image à calibrer puis un bouton pour régler le gamma et un dernier bouton pour détecter les patches.

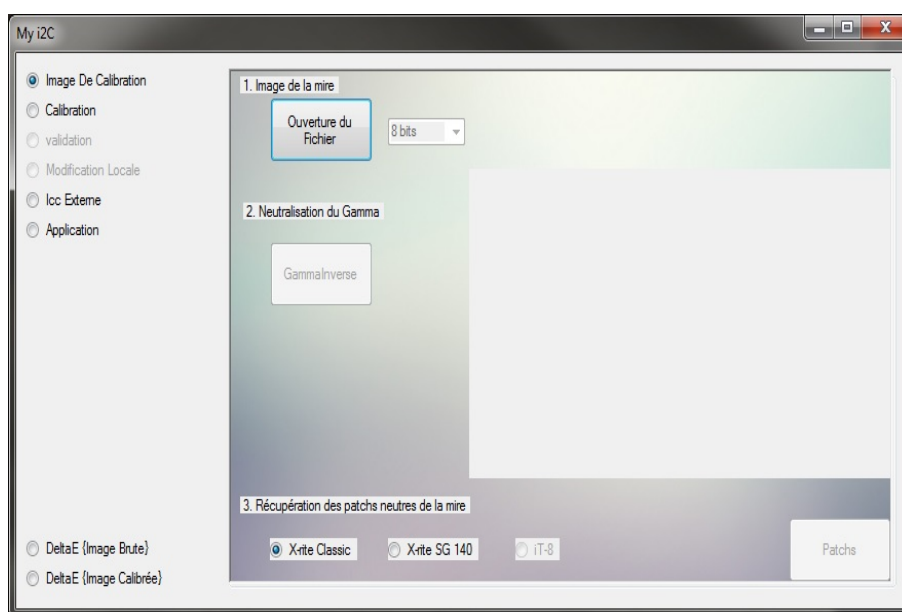


FIGURE 6.13: Interface de départ du logiciel de calibration couleur

Le réglage du gamma nous amène sur un autre interface qui propose plusieurs possibilités pour corriger le gamma (cf. la figure 6.14). L'option 1 est l'option automatique et idéale pour l'utilisateur (qui n'est pas implémenté) ; l'option 2 consiste à récupérer une mire de type Kodak Q-13 possédant 20 patches de niveaux de gris, si la mire a des valeurs spécifique il est possible de choisir un fichier de référence. L'option 3 consiste à calculer le gamma en fonction des patches de niveaux de gris sur les colorchecker classic et SG 140. La dernière option consiste à appliquer un gamma directement connaissant a priori l'information.

Lorsque nous choisissons l'option 2 ou 3, il apparaît une interface permettant de sélectionner la région d'intérêt de l'image. Ce viewer, que l'on voit à la figure 6.15, permet de sélectionner la mire voulue. Nous pouvons sélectionner une région en créant un rectangle avec la souris puis affiner la région en utilisant le bouton semi-automatique qui va trouver



FIGURE 6.14: Interface de gestion du Gamma

les croix blanches de référence qu'il y a sur les deux mires colorchecker, ou en modifiant localement les coins du rectangle manuellement. Le bouton warp va transformer la région sélectionnée en un rectangle, et enfin, le bouton mire va extraire les patches de la mire. Si aucune région n'a été sélectionnée une détection automatique va se mettre en place.

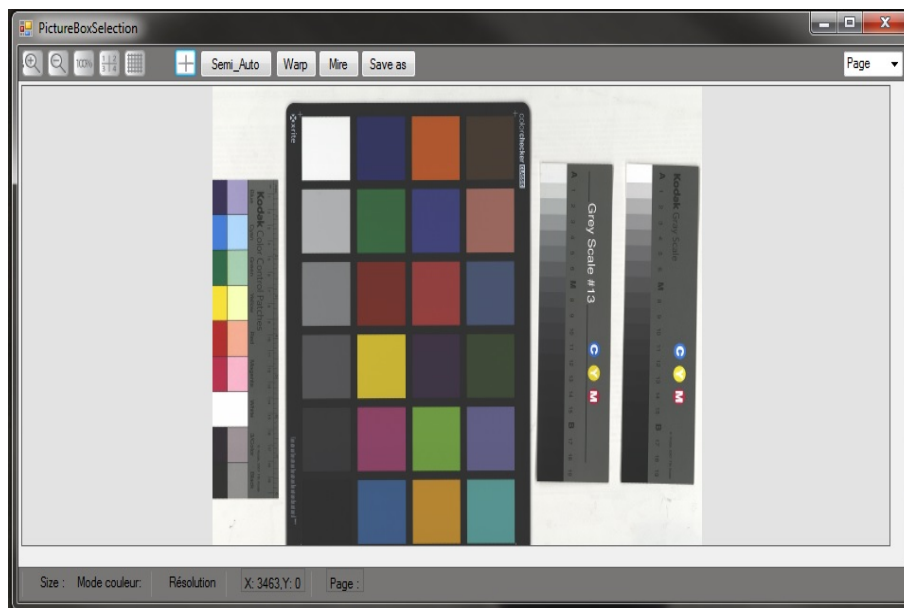


FIGURE 6.15: Viewer

Pour l'étape du gamma, nous présentons la courbe des niveaux de gris et la nouvelle courbe après correction Gamma, comme nous le voyons à la figure 6.16.

Nous nous retrouvons ensuite dans l'interface de départ, et nous pouvons identifier les patches de la mire. Nous repassons par le viewer et réitérons la sélection de la mire.

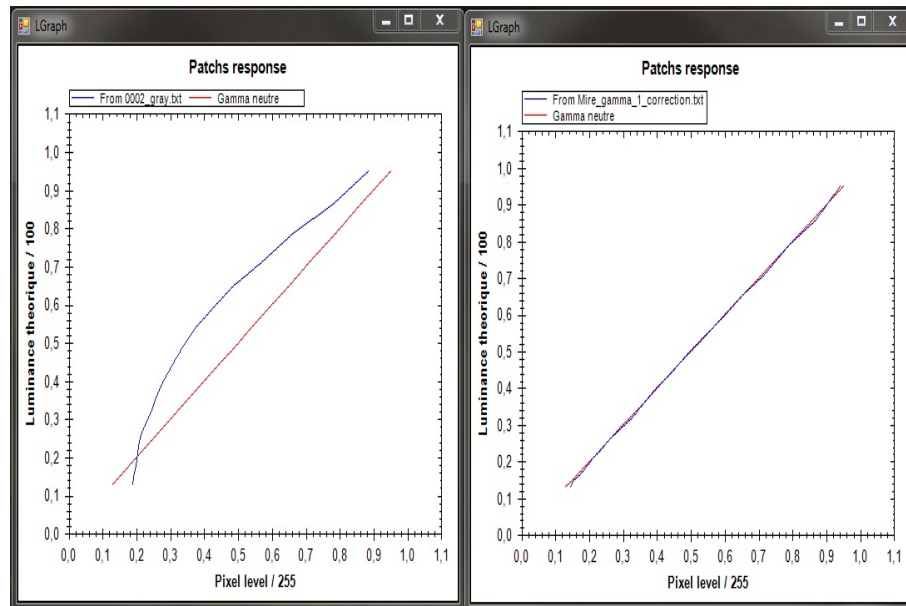


FIGURE 6.16: Courbes des niveaux de gris

Dans l'étape de Calibration (cf. figure figure 6.17), nous faisons un récapitulatif des informations, c'est-à-dire les mesures $R'V'B'$ (ie, sans gamma) et les valeurs $L^*a^*b^*$. Nous pouvons éditer, ajouter ou supprimer des mesures.

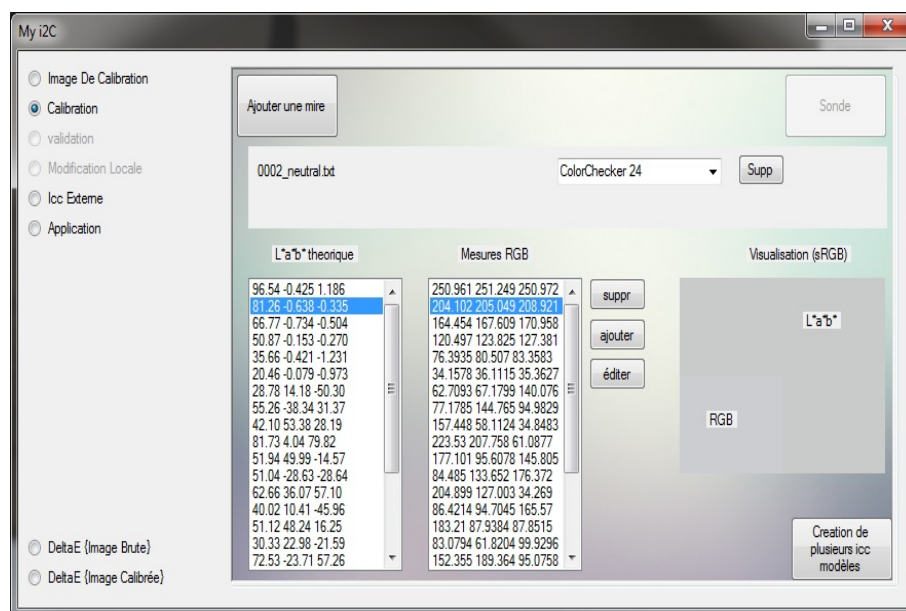


FIGURE 6.17: Valeurs de calibration

Nous avons mis une visualisation de la couleur, et nous pouvons constater que le gris clair est très proche du clair théorique. Le bouton en bas à droite va mettre en place un scrip calculant plusieurs modèles de calibration. Nous avons vu que la transformation est dépendante de la mire par conséquent nous proposons plusieurs transformations et nous

donnons les résultats de calibrations. Nous faisons passer à l'étape suivante et présentons les résultats dans une nouvelle interface 6.18.

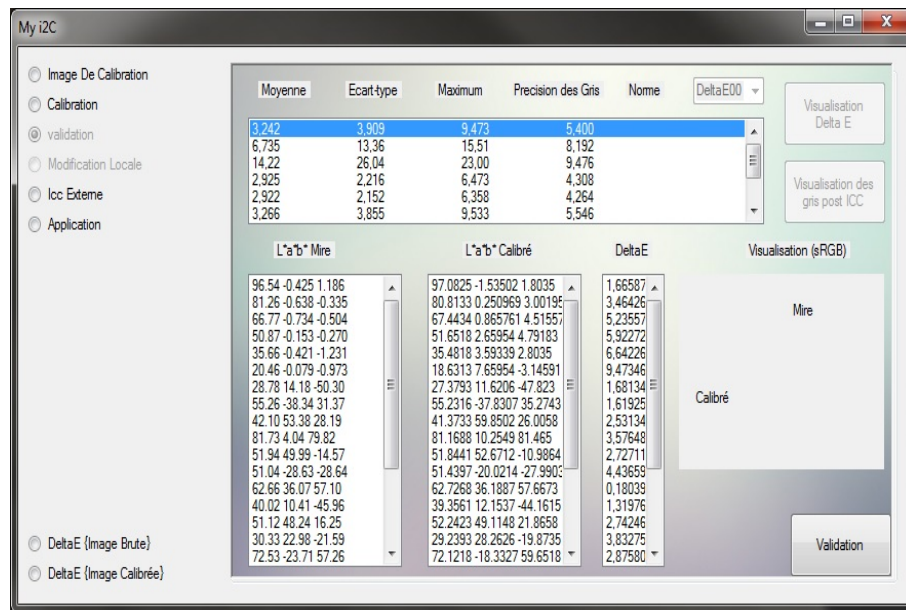


FIGURE 6.18: Résultats des ICCs selon plusieurs modèles

Dans cette interface, nous affichons les résultats obtenus en appliquant l'ICC de calibration en fonction de la moyenne, écart-type, Maximum et la précision des niveaux de Gris pour un $\Delta E00$. Il n'est a priori pas utile pour l'utilisateur de connaître la méthode mais simplement de trouver le meilleur compromis entre la moyenne, le maximum et les gris. Une fois le résultat choisi en sélectionnant, nous validons et nous obtenons une étape de modification locale (voir figure 6.19) de la LUT qui est toujours en développement.

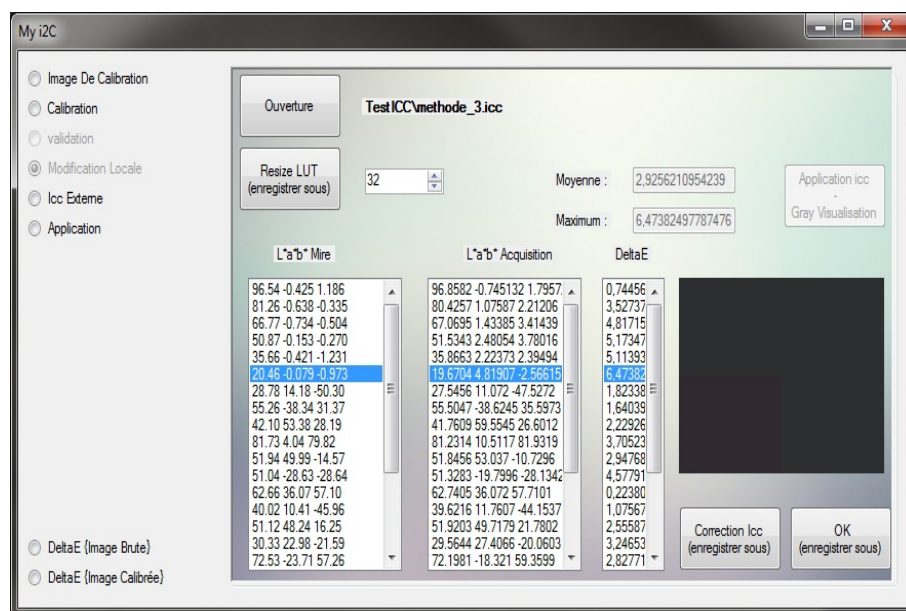


FIGURE 6.19: Modification locale d'une LUT

L'étape suivante, l'étape Icc Externe, permet de créer des ICCs en fonction des informations de la Commission International de l'Éclairage ou des informations connues des formats standards. L'étape d'application consiste à prendre une image et à y appliquer l'ICC que nous venons de créer à partir de la mire et l'ICC standard ou créer dans l'étape Icc Externe.

L'autre partie du soft consiste à analyser une mire. Nous avons affiché deux étapes : l'étape où nous calculons le ΔE avec l'image de départ (dit Brute) et l'ICC de l'appareil, c'est-à-dire l'ICC que nous avons créer à partir de la mire ou à partir d'un autre logiciel ; l'autre étape consiste à calculer le δE à partir de l'image en format RVB standard (dit Calibrée). Comme nous le voyons dans la figure 6.20, nous proposons de choisir l'ICC du format standard permettant de retourner dans l'espace $L^*a^*b^*$, le bouton "calcul" permet de sélectionner la mire avec le viewer.

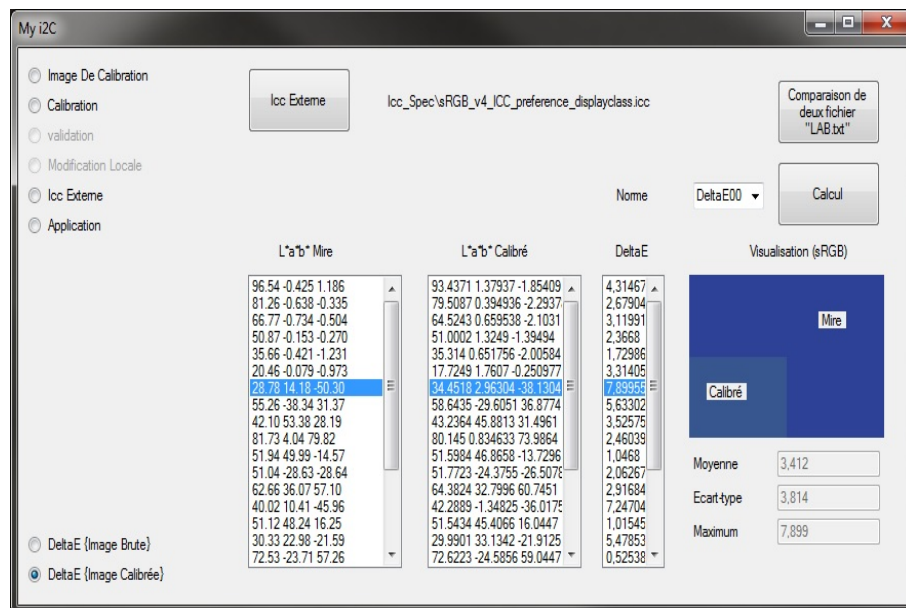


FIGURE 6.20: Calcul du Delta E

Bibliographie

- [Abrardo et al., 1996] Abrardo, A., Cappellini, V., Cappellini, M., and Mecocci, A. (1996). Art-works colour calibration by using the vasari scanner. In *Color and Imaging Conference*, pages 94–96. Society for Imaging Science and Technology.
- [Bae and Durand, 2007] Bae, S. and Durand, F. (2007). Defocus magnification. In *Computer Graphics Forum*, volume 26, pages 571–579. Wiley Online Library.
- [Barten, 1990] Barten, P. G. (1990). Evaluation of subjective image quality with the square-root integral method. *JOSA A*, 7(10) :2024–2031.
- [Bernsen, 1986] Bernsen, J. (1986). Dynamic thresholding of grey-level images. In *International conference on pattern recognition*, pages 1251–1255.
- [Chung et al., 2004] Chung, Y.-C., Wang, J.-M., Bailey, R. R., Chen, S.-W., and Chang, S.-L. (2004). A non-parametric blur measure based on edge analysis for image processing applications. In *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, volume 1, pages 356–360. IEEE.
- [Clarke et al., 1984] Clarke, F., McDonald, R., and Rigg, B. (1984). Modification to the jpc79 colour-difference formula. *Journal of the Society of Dyers and Colourists*, 100(4) :128–132.
- [color.org, 2013] color.org (2013). srgbv4iccpreference.icc.
- [Elder and Zucker, 1998] Elder, J. H. and Zucker, S. W. (1998). Local scale control for edge detection and blur estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(7) :699–716.
- [FADGI, 2010] FADGI (2010). *Digital Imaging Framework*. <http://www.digitizationguidelines.gov>.
- [Field and Brady, 1997] Field, D. J. and Brady, N. (1997). Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision research*, 37(23) :3367–3383.
- [Field et al., 1987] Field, D. J. et al. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12) :2379–2394.

- [Finkel and Bentley, 1974] Finkel, R. A. and Bentley, J. L. (1974). Quad trees a data structure for retrieval on composite keys. *Acta informatica*, 4(1) :1–9.
- [Hardeberg, 2001] Hardeberg, J. Y. (2001). *Acquisition and reproduction of color images : colorimetric and multispectral approaches*. Universal-Publishers.
- [Hardeberg et al., 1996] Hardeberg, J. Y., Schmitt, F., Tastl, I., Brettel, H., and Crettez, J.-P. (1996). Color management for color facsimile. In *Color and Imaging Conference*, pages 108–113. Society for Imaging Science and Technology.
- [Hecht, 1924] Hecht, S. (1924). The visual discrimination of intensity and the weber-fechner law. *The Journal of general physiology*, 7(2) :235–267.
- [Hurlbert, 1989] Hurlbert, A. (1989). The computation of color.
- [Imatest, 2013] Imatest, L. (2013). Imatest 3.10.
- [Initiative, 2010] Initiative, F. A. D. (2010). Technical guidelines for digitizing cultural heritage materials : Creation of raster image master files.
- [ISO 12233, 2014] ISO 12233 (2014). Photography - Electronic still-picture cameras - Resolution and spatial frequency responses.
- [ISO 15076-1 :2010, 2010] ISO 15076-1 :2010 (2010). Image technology colour management – Architecture, profile format and data structure.
- [ISO 16067-1, 2003] ISO 16067-1 (2003). Photography – Spatial resolution measurements of electronic scanners for photographic images – Part 1 : Scanners for reflective media.
- [ISO 16067-2, 2004] ISO 16067-2 (2004). Photography – Electronic scanners for photographic images – Spatial resolution measurements – Part 2 : Film scanners.
- [ISO 21550, 2004] ISO 21550 (2004). Photography – Electronic scanners for photographic images – Dynamic range measurements.
- [Jain, 1989] Jain, A. K. (1989). *Fundamentals of digital image processing*. Prentice-Hall, Inc.
- [Jobson et al., 1997] Jobson, D. J., Rahman, Z.-U., and Woodell, G. A. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. *Image Processing, IEEE Transactions on*, 6(7) :965–976.
- [Jourlin and Pinoli, 1995] Jourlin, M. and Pinoli, J.-C. (1995). Image dynamic range enhancement and stabilization in the context of the logarithmic image processing model. *Signal processing*, 41(2) :225–237.
- [Kang, 1997] Kang, H. R. (1997). *Color Technology for electronic imaging devices*. SPIE press.
- [Kapur et al., 1985] Kapur, J., Sahoo, P. K., and Wong, A. K. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Computer vision, graphics, and image processing*, 29(3) :273–285.

- [Ke et al., 2006] Ke, Y., Tang, X., and Jing, F. (2006). The design of high-level features for photo quality assessment. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 419–426. IEEE.
- [Kittler and Illingworth, 1985] Kittler, J. and Illingworth, J. (1985). On threshold selection using clustering criteria. *Systems, Man and Cybernetics, IEEE Transactions on*, 15(5) :652–655.
- [Land, 1964] Land, E. H. (1964). The retinex. In *Ciba Foundation Symposium-Colour Vision : Physiology and Experimental Psychology*, pages 217–227. Wiley Online Library.
- [Lelegard et al., 2010] Lelegard, L., Bredif, M., Vallet, B., and Boldo, D. (2010). Motion blur detection in aerial images shot with channel-dependent exposure time.
- [Li and Tam, 1998] Li, C. and Tam, P. K.-S. (1998). An iterative algorithm for minimum cross entropy thresholding. *Pattern Recognition Letters*, 19(8) :771–776.
- [Li and Lee, 1993] Li, C. H. and Lee, C. (1993). Minimum cross entropy thresholding. *Pattern Recognition*, 26(4) :617–625.
- [Lindbloom, 2010] Lindbloom, B. J. (2010). Useful color equations.
- [Liu et al., 2008] Liu, R., Li, Z., and Jia, J. (2008). Image partial blur detection and classification. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE.
- [Luo et al., 2001] Luo, M. R., Cui, G., and Rigg, B. (2001). The development of the cie 2000 colour-difference formula : Ciede2000. *Color Research & Application*, 26(5) :340–350.
- [Mahy et al., 1994] Mahy, M., Eycken, L., and Oosterlinck, A. (1994). Evaluation of uniform color spaces developed after the adoption of cielab and cieluv. *Color Research & Application*, 19(2) :105–121.
- [Marziliano et al., 2004] Marziliano, P., Dufaux, F., Winkler, S., and Ebrahimi, T. (2004). Perceptual blur and ringing metrics : application to jpeg2000. *Signal Processing : Image Communication*, 19(2) :163–172.
- [Mayet et al., 1996] Mayet, F., Pinoli, J.-C., and Jourlin, M. (1996). Justifications physiques et applications du modèle lip pour le traitement des images obtenues en lumière transmise. *Traitement du signal*, 13(3).
- [McDonald and Smith, 1995] McDonald, R. and Smith, K. J. (1995). Cie94-a new colour-difference formula*. *Journal of the Society of Dyers and Colourists*, 111(12) :376–379.
- [Melgosa, 2000] Melgosa, M. (2000). Testing cielab-based color-difference formulas. *Color Research & Application*, 25(1) :49–55.
- [Meyer, 1990] Meyer, Y. (1990). *Ondelettes, filtres miroirs en quadrature et traitement numérique de l'image*. Springer.

- [Narvekar and Karam, 2009] Narvekar, N. D. and Karam, L. J. (2009). A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection. In *Quality of Multimedia Experience, 2009. QoMEX 2009. International Workshop on*, pages 87–91. IEEE.
- [Niblack, 1985] Niblack, W. (1985). *An introduction to digital image processing*. Strandberg Publishing Company.
- [of Australia, 2013] of Australia, A. S. L. N. L. (2013). Icc color camera calibrator.
- [Otsu, 1975] Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285-296) :23–27.
- [Pal, 1996] Pal, N. R. (1996). On minimum cross-entropy thresholding. *Pattern Recognition*, 29(4) :575–580.
- [Pascale, 2003] Pascale, D. (2003). A review of rgb color spaces... from xyy to r'g'b'. *Babel Color*.
- [Pascale, 2006] Pascale, D. (2006). Rgb coordinates of the macbeth colorchecker. *The BabelColor Company*, pages 1–16.
- [Pascale, 2014] Pascale, D. (2014). Colorchecker classic analysis.
- [pierre.nerzic, 2008] pierre.nerzic (2008). cours - module outils et matériels pour l'imagerie numérique.
- [Pinoli, 1997] Pinoli, J.-C. (1997). A general comparative study of the multiplicative homomorphic, log-ratio and logarithmic image processing approaches. *Signal Processing*, 58(1) :11–45.
- [Pinoli and Debayle, 2006] Pinoli, J.-C. and Debayle, J. (2006). Logarithmic adaptive neighborhood image processing (lanip) : introduction, connections to human brightness perception, and application issues. *EURASIP Journal on Advances in Signal Processing*, 2007.
- [Pub, 1995] Pub, C. (1995). 116-1995. industrial colour-difference evaluation. *Vienna : CIE Central Bureau*.
- [Pub, 2001] Pub, C. (2001). 142-2001.improvement to industrial colour-difference evaluation. *Vienna : CIE Central Bureau*.
- [Rahman et al., 1996] Rahman, Z.-u., Jobson, D. J., and Woodell, G. A. (1996). Multiscale retinex for color rendition and dynamic range compression. In *SPIE's 1996 International Symposium on Optical Science, Engineering, and Instrumentation*, pages 183–191. International Society for Optics and Photonics.
- [Rahman et al., 2004] Rahman, Z.-u., Jobson, D. J., and Woodell, G. A. (2004). Retinex processing for automatic image enhancement. *Journal of Electronic Imaging*, 13(1) :100–110.

- [Rahman et al., 1997] Rahman, Z.-u., Woodell, G. A., Jobson, D. J., et al. (1997). A comparison of the multiscale retinex with other image enhancement techniques. In *IS AND T ANNUAL CONFERENCE*, pages 426–431. Citeseer.
- [Rosenfeld and De La Torre, 1983] Rosenfeld, A. and De La Torre, P. (1983). Histogram concavity analysis as an aid in threshold selection. *Systems, Man and Cybernetics, IEEE Transactions on*, 13(2) :231–235.
- [Russ, 2010] Russ, J. C. (2010). *The image processing handbook*. CRC press.
- [Sauvola and Pietikäinen, 2000] Sauvola, J. and Pietikäinen, M. (2000). Adaptive document image binarization. *Pattern recognition*, 33(2) :225–236.
- [Sezan, 1990] Sezan, M. I. (1990). A peak detection algorithm and its application to histogram-based image data reduction. *Computer vision, graphics, and image processing*, 49(1) :36–51.
- [Sharma et al., 2005] Sharma, G., Wu, W., and Dalal, E. N. (2005). The ciede2000 color-difference formula : Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1) :21–30.
- [Stathis et al., 2008] Stathis, P., Kavallieratou, E., and Papamarkos, N. (2008). An evaluation survey of binarization algorithms on historical documents. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE.
- [Tai and Brown, 2009] Tai, Y.-W. and Brown, M. S. (2009). Single image defocus map estimation using local contrast prior. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1797–1800. IEEE.
- [Tappen et al., 2005] Tappen, M. F., Freeman, W. T., and Adelson, E. H. (2005). Recovering intrinsic images from a single image. *IEEE transactions on pattern analysis and machine intelligence*, 27(9) :1459.
- [Tieu and Viola, 2004] Tieu, K. and Viola, P. (2004). Boosting image retrieval. *International Journal of Computer Vision*, 56(1-2) :17–36.
- [Tong et al., 2004] Tong, H., Li, M., Zhang, H., and Zhang, C. (2004). Blur detection for digital images using wavelet transform. In *Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on*, volume 1, pages 17–20. IEEE.
- [Tong et al., 2005] Tong, H., Li, M., Zhang, H.-J., Zhang, C., He, J., and Ma, W.-Y. (2005). Learning no-reference quality metric by examples. In *Multimedia Modelling Conference, 2005. MMM 2005. Proceedings of the 11th International*, pages 247–254. IEEE.
- [van Dormolen, 2008] van Dormolen, H. (2008). Metamorfoze preservation imaging guidelines. In *Archiving Conference*, pages 162–165. Society for Imaging Science and Technology.

- [Vrhel and Trussell, 1999] Vrhel, M. and Trussell, H. (1999). Color scanner calibration via a neural network. In *Acoustics, Speech, and Signal Processing, 1999. Proceedings. Vol 6, 1999 IEEE International Conference on*, pages 3465–3468. IEEE.
- [Weiss, 2001] Weiss, Y. (2001). Deriving intrinsic images from image sequences. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 68–75. IEEE.
- [Wik, 2014] Wik (2014). Wikipédia - l'encyclopédie libre.
- [Yilmaz et al., 2004] Yilmaz, I., Bildirici, I. O., Yakar, M., and Yildiz, F. (2004). Color calibration of scanners using polynomial transformation. In *XXth ISPRS Congress Commission V, Istanbul, Turkey*, pages 890–896.
- [Zhuo and Sim, 2011] Zhuo, S. and Sim, T. (2011). Defocus map estimation from a single image. *Pattern Recognition*, 44(9) :1852–1858.