



Titre: Recalage d'images du fond d'oeil pour la construction d'un atlas des structures anatomiques de la rétine
Title:

Auteur: Cédric Meurée
Author:

Date: 2013

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Meurée, C. (2013). Recalage d'images du fond d'oeil pour la construction d'un atlas des structures anatomiques de la rétine [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. <https://publications.polymtl.ca/1280/>
Citation:

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/1280/>
PolyPublie URL:

Directeurs de recherche: Farida Cheriet, & Timothée Faucon
Advisors:

Programme: Génie biomédical
Program:

UNIVERSITÉ DE MONTRÉAL

RECALAGE D'IMAGES DU FOND D'ŒIL POUR LA CONSTRUCTION D'UN ATLAS
DES STRUCTURES ANATOMIQUES DE LA RÉTINE

CÉDRIC MEURÉE
INSTITUT DE GÉNIE BIOMÉDICAL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE BIOMÉDICAL)
DÉCEMBRE 2013

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

RECALAGE D'IMAGES DU FOND D'ŒIL POUR LA CONSTRUCTION D'UN ATLAS
DES STRUCTURES ANATOMIQUES DE LA RÉTINE

présenté par : MEURÉE Cédric

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

Mme PÉRIÉ-CURNIER Delphine, Doct., présidente

Mme CHERIET Farida, Ph.D., membre et directrice de recherche

M. FAUCON Timothée, Ph.D., membre et codirecteur de recherche

M. BILODEAU Guillaume-Alexandre, Ph.D., membre

À ma famille et à mes amis

REMERCIEMENTS

J'aimerais exprimer ma reconnaissance envers toutes les personnes qui ont fait en sorte que le travail présenté dans ce mémoire puisse être réalisé.

Je tiens tout d'abord à remercier ma directrice de recherche, Madame Farida Cheriet, pour les opportunités qu'elle m'a données, sa confiance et ses conseils avisés qui ont permis la réalisation de ce projet.

Un grand merci également à Timothée Faucon et Housseem Ben Tahar qui m'ont permis de travailler au sein de l'équipe de Diagnos Inc. Je suis particulièrement reconnaissant à Timothée pour son suivi régulier de mes travaux et son aide tant sur les plans professionnel que personnel.

Je souhaite remercier les membres du jury, Madame Delphine Périé-Curnier et Monsieur Guillaume-Alexandre Bilodeau d'avoir accepté et pris le temps d'évaluer le travail réalisé dans le cadre de ce projet.

Merci à Lama Séoud pour son aide technique et pour avoir partagé avec moi l'espace des stagiaires chez Diagnos pendant ces deux années. Merci également aux membres du LIV4D : nos associés de recherche Philippe Debanné et Sébastien Grenier pour leur disponibilité, Cosmin Ludusan, Mahnaz Fasih, Argyrios Christodoulidis, Najat Chihab, Rola Harmouche, Mathias Adankon, Nesrine Zaglam, Hervé Lombaert, Mitchel Benovoy, Kondo-Claude Assi pour les bons moments passés ensemble et David Thivierge-Gaulin pour sa force de motivation et la visite du parc de la Chute-Montmorency. J'aimerais également remercier les membres de l'équipe de recherche sur l'imagerie de fond d'œil que sont Pierre Langlois, Thomas Hurtut, Valérie Béland et Jean-Philippe Daunais, ces deux derniers ayant effectué des études visant à poursuivre les travaux réalisés dans le cadre de mon projet de maîtrise.

Merci encore à l'équipe de Diagnos Inc., avec laquelle j'ai passé de très bons moments, et en particulier Paul Boivin et Stéphane Henry pour leur aide en informatique.

J'aimerais adresser mes remerciements à mes proches amis et à la famille de Centrale Marseille qui m'ont été d'un grand soutien depuis mon arrivée à Montréal.

Finalement, je souhaite remercier tout particulièrement ma famille, Thierry, Pascale et Fiona Meurée, qui m'ont soutenu et encouragé tout au long de mes études et de cette aventure canadienne.

RÉSUMÉ

Le développement de nouvelles modalités d'imagerie et les progrès réalisés en analyse et en archivage d'images ont fait évoluer la pratique de l'ophtalmologie ces dernières décennies. Ces avancées sont motivées par le fait que les images de la rétine, également appelées images de fond d'œil, sont les seules qui permettent d'observer le réseau vasculaire rétinien de manière non-invasive et d'obtenir des informations concernant un grand nombre de pathologies. La quantité d'images rétinienne acquises afin de suivre des patients atteints de différentes maladies est donc en constante augmentation, ce qui pose des problèmes dans la prise en charge de la population. De plus, l'évaluation qualitative de ces images par des experts est soumise à des problèmes de reproductibilité. De ces problématiques croît la nécessité de développer des méthodes automatiques d'analyse des images de fond d'œil et d'aide au diagnostic. Certaines des pathologies évaluées à l'aide de ces images affectant la structure du réseau vasculaire rétinien, la segmentation automatique de la vasculature est l'un des enjeux particulièrement importants aujourd'hui.

Diverses méthodes de segmentation de la vasculature rétinienne sont proposées dans la littérature, l'une des principales difficultés à surmonter étant de segmenter les vaisseaux sanguins les plus fins. Notre hypothèse est que l'utilisation d'atlas pourrait améliorer la détection de ces petits vaisseaux. Les atlas sont des modèles moyens de la répartition des structures au sein de populations qui, une fois alignés sur des images à segmenter, permettent de localiser les structures d'intérêt.

L'objectif général de ce projet est de proposer des méthodes de recalage automatique d'images de fonds d'œil afin de construire un atlas des principales structures anatomiques de la rétine. Puisque les images de fonds d'œil utilisées pour construire un tel atlas sont planes, leur contenu dépend de la configuration selon laquelle elles ont été imagées. Nous proposons donc aussi d'étudier la possibilité d'une reconstruction de la surface des rétines en trois dimensions afin d'obtenir un atlas 3D indépendant de cette configuration. Cette extension à la troisième dimension implique le recours à plusieurs images de chaque fond d'œil qui doivent être recalées afin de ne conserver que les déformations liées à la profondeur de la rétine.

Trois objectifs spécifiques sont donc à distinguer. Le premier consiste à proposer une méthode de recalage automatique intra-sujet pour permettre la reconstruction en trois dimensions. Le deuxième vise à adapter le recalage au cas de l'alignement d'images de sujets différents dans le but de construire l'atlas 2D, qui est utile dans le cas d'images acquises selon des configurations similaires. Le troisième objectif est d'étudier la possibilité de reconstruire

la surface des rétines en trois dimensions.

Afin de répondre au premier objectif qui est de proposer une méthode de recalage automatique intra-sujet, un algorithme évaluant les déformations quadratiques (d'ordre 2) entre les images a été développé. Il permet de prendre en compte les transformations introduites par le déplacement de l'œil par rapport aux caméras entre les acquisitions de plusieurs images et de considérer les déformations liées au fait que les images sont acquises par projection de rétines sphériques sur un plan. Cette méthode de recalage est évaluée en mesurant la différence de positionnement entre des points correspondants dans les images après recalage. La moyenne des distances mesurées est égale à 2.51 pixels, ce qui correspond à 0,040mm sur une rétine. Le recalage intra-sujet est ensuite évalué de manière qualitative en formant des mosaïques d'images qui permettent d'étendre la zone observée sur une seule image. Le recalage est également utilisé pour effectuer le suivi de pathologies chez des patients, certaines maladies pouvant entraîner l'apparition de taches ou d'une microvasculature qui évoluent au cours du temps.

L'algorithme de recalage est ensuite adapté au cas de l'alignement d'images de différents fonds d'œil. La principale difficulté étant qu'il n'y a que deux structures communes entre de telles images que sont la fovéa et le disque optique, des ellipses ont été placées sur les images afin de pouvoir évaluer des déformations globales. L'écart-type de la distribution des positions des centres du disque optique et de la fovéa après recalage par rapport à leur position moyenne est calculé afin d'évaluer quantitativement l'algorithme. Les valeurs obtenues sont de 35.71 pixels (0.573mm) pour la fovéa et 32.92 pixels (0.528mm) pour le disque optique. Afin d'évaluer qualitativement ce recalage, une étude de la répartition des principales structures anatomiques rétinienne est réalisée et une carte de la probabilité des pixels des images de fonds d'œil d'appartenir au réseau vasculaire est construite dans le but de faciliter l'initialisation des algorithmes de segmentation du réseau vasculaire.

Une réponse au troisième objectif qu'est celui d'étudier la possibilité de reconstruire les fonds d'œil en trois dimensions a été proposée. Celle-ci repose sur l'utilisation d'images stéréoscopiques des rétines et sur l'hypothèse selon laquelle la configuration de leur acquisition est semblable à une géométrie épipolaire. Un profil 3D de la surface de la rétine est alors obtenu à une transformation projective près par rapport à la réalité, les caméras n'ayant pas été calibrées.

Les principales propositions d'extension des travaux réalisés dans le cadre de ce projet concernent le développement d'une méthode de calibrage des caméras de fond d'œil, ce qui permettrait d'obtenir une reconstruction métrique de la surface de la rétine et de construire un atlas 3D qui rende compte de la répartition des structures anatomiques dans l'espace. Le suivi des recherches entreprises en ce qui concerne la segmentation du disque optique

pourrait permettre un meilleur positionnement de son centre, ce qui améliorerait la capacité d'alignement inter-sujets de l'algorithme de recalage.

ABSTRACT

The development of new imaging modalities and the progress in image analysis and archiving systems have changed the practice of ophthalmology for the last decades. These advances are driven by the fact that retinal images, also known as fundus images, are the only ones allowing a non-invasive observation of the human vasculature and getting information about many pathologies. The amount of retinal images acquired to examine patients with various diseases is constantly increasing, which poses problems in the management of the population. In addition, the qualitative evaluation of these images by experts is subject to problems of reproducibility. These issues increase the need to develop automatic retinal image analysis and diagnosis support. Some pathologies are affecting the structure of the retinal vasculature. Therefore, the automatic retinal vessels segmentation is a particularly important issue today.

Various methods have been proposed in the literature concerning retinal vessels segmentation. One of the key challenges is to segment the finest blood vessels. Our hypothesis is that the use of atlases could improve the detection of these small vessels. Atlases mean the distribution of anatomical structures within populations and, once aligned with images to segment, allow the localisation of the structures of interest.

The general objective of this project is to propose a method for automatic images registration to build an atlas of the principal anatomical structures of the retina. Since the fundus images used to build this atlas are two dimensional, their content depends on the configuration of their acquisition. We propose to study the possibility of reconstructing the retinal surface in three dimensions to obtain an atlas not depending on this configuration. This extension to the third dimension involves the use of multiple images of each retina which must be registered in order to keep only the deformations related to the depth of the retina.

Thus, three specific objectives are to distinguish. The first is to provide an automatic intra-subject registration method to allow the three-dimensional reconstruction. The second is to adapt the registration algorithm to the alignment of images of different subjects in order to build the 2D atlas, which is useful in the case of images acquired under similar configurations. The third objective is to study the possibility to reconstruct the retinal surface in three dimensions.

To meet the first objective, which is to propose a method to automatically register images of a single retina, an algorithm evaluating the quadratic deformations between the images has been developed. It allows taking into account transformations introduced by the movements of the eye relative to the camera between several acquisitions and considering deformations

related to the fact that the images are acquired by projecting spherical retinas on planes. This registration method is evaluated by measuring the difference in positioning between corresponding points after registration. The mean value of the calculated distances is equal to 2.51 pixels, which corresponds to $0,040mm$ on a retina. The intra-subject registration is also qualitatively evaluated by forming mosaics of images that extend the observed area on a single image. The registration is also used to follow diseases that may cause the appearance of stains or microvasculature over the time.

The registration algorithm is then adapted to the case of aligning images of different retinas. The main difficulty is that there are only two common structures between such images, which are the fovea and the optic disc. Ellipses were placed on the images in order to assess the overall deformation. The standard deviation of the distribution of the central positions of the optic disc and the fovea relative to their average positions after registration is calculated to quantitatively evaluate the algorithm. The obtained values are 35.71 pixels ($0.573mm$) for the fovea and 32.92 pixels ($0.528mm$) for the optic disc. To qualitatively evaluate the registration, a study of the distribution of the key anatomical structures is performed and a map of the likelihood of the pixels of being part of the vasculature is constructed for the purpose of making the initialization of retinal vessels segmentation algorithms easier.

A response to the third objective, which is to study the possibility of reconstructing the fundus in three dimensions, has been proposed. It is based on the use of stereoscopic images of the retinas and the assumption that the configuration of their acquisition is similar to an epipolar geometry. A 3D profile of the retinal surfaces is then obtained, differing from the reality by a projective transformation because the cameras have not been calibrated.

The main proposals for extending the work realized in the framework of this project include the development of a method for calibrating fundus cameras, which would provide a metric reconstruction of the surface of the retina and the building of a 3D atlas that would spatially reflect the distribution of the anatomical structures. Research projects regarding the segmentation of the optic disc could allow a better positioning of its center, which would improve the inter-subjects alignment.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	viii
TABLE DES MATIÈRES	x
LISTE DES TABLEAUX	xii
LISTE DES FIGURES	xiii
LISTE DES SIGLES ET ABRÉVIATIONS	xvi
CHAPITRE 1 INTRODUCTION	1
1.1 Organisation du mémoire	3
CHAPITRE 2 REVUE DE LITTÉRATURE	4
2.1 Anatomie	4
2.1.1 Structure de la rétine	4
2.1.2 Pathologies détectables à partir des images de fond d'œil	6
2.2 Systèmes d'imagerie de l'œil	10
2.2.1 Historique	10
2.2.2 Systèmes d'imagerie 2D	11
2.2.3 Systèmes d'imagerie 3D	13
2.3 Recalage d'images	16
2.3.1 Choix du type de transformations	17
2.3.2 Détection de structures d'intérêt	23
2.3.3 Mesures de similarité et critères d'optimisation	30
2.4 Reconstruction 3D de structures anatomiques	31
2.4.1 Géométrie épipolaire	32
2.4.2 Détermination de la configuration du système d'acquisition	34
2.4.3 Reconstruction 3D	35

2.5	Objectifs	41
CHAPITRE 3	Méthodologie	42
3.1	Recalage d'images de fond d'œil	42
3.1.1	Débruitage	43
3.1.2	Sélection de points d'intérêt	43
3.1.3	Estimation graduelle des transformations	45
3.2	Construction d'atlas 2D d'images de fonds d'œil	47
3.2.1	Sélection des structures	48
3.2.2	Recalage	52
3.3	Vers la construction d'un atlas 3D de la surface rétinienne	52
3.3.1	Images stéréoscopiques	52
3.3.2	Reconstruction dense	53
3.4	Méthode d'évaluation de l'approche proposée	55
CHAPITRE 4	Résultats et discussions	57
4.1	Recalage intra-sujet d'images de fond d'œil	57
4.1.1	Évaluation quantitative de l'algorithme de recalage intra-sujet	57
4.1.2	Évaluation qualitative de l'algorithme de recalage intra-sujet	62
4.2	Construction d'atlas 2D	64
4.2.1	Évaluation quantitative de l'algorithme de construction d'atlas 2D	64
4.2.2	Évaluation qualitative de l'algorithme de construction d'atlas 2D	66
4.3	Atlas 3D	71
CHAPITRE 5	CONCLUSION	74
RÉFÉRENCES	76

LISTE DES TABLEAUX

Tableau 4.1	Distances moyennes entre les paires de points correspondants après application de l'algorithme de recalage intra-sujet.	58
Tableau 4.2	Distances moyennes entre les paires de points correspondants après application du recalage quadratique à des images de dimensions différentes	60
Tableau 4.3	Évaluation de la similitude entre les distributions des distances entre des points sélectionnés manuellement ou avec OpenSURF	61
Tableau 4.4	Évaluation de l'écart-type des positions des centres de la fovéa et du disque optique après recalage par rapport à leur position moyenne. . .	64

LISTE DES FIGURES

Figure 2.1	Anatomie de l'œil humain	5
Figure 2.2	Image de fond d'œil	7
Figure 2.3	Image de papille. Le nerf optique est visible sous la forme d'une excavation physiologique dont le diamètre représente 6/10 de celui de la papille (Y. Lachkar and E. Sellem (2004))	7
Figure 2.4	Principe de fonctionnement des caméras de fond d'œil à illumination externe (a) et illumination interne (b), ((a) 1. œil; 2. miroir semi réfléchissant; 3., 5., 6., 8., 10. lentilles; 4. diaphragme; 7. image de la source; 9. obscurateur central; 11. source, (b) 1. œil; 2., 4., 5., 7., 8., 9., 11. lentilles; 3. miroir à trou central; 6. image de la source; 10. obscurateur central; 12. source) (DeHoog et Schwiegerling (2009)) . .	12
Figure 2.5	Schéma illustrant la méthode d'acquisition des images OCT (Fujimoto <i>et al.</i> (2000))	15
Figure 2.6	Différence de luminosité entre deux images de la rétine d'un même sujet et exemple de réflexion intense	17
Figure 2.7	Distances de quatre points colinéaires à une origine O. Les transformations projectives conservent les rapports de distances.	20
Figure 2.8	Géométrie de l'acquisition de deux images de la rétine d'un même sujet (Can <i>et al.</i> (2002))	22
Figure 2.9	Représentation d'une image transférée dans les espaces des échelles gaussiennes et des différences d'échelles gaussiennes. Un changement d'octave correspond à une modification de la taille de l'image d'un facteur 2. Un changement de sous-niveau implique une modification du paramètre σ (C. Evans (2009))	26
Figure 2.10	Comparaison d'un point d'intérêt avec ses voisins dans les niveaux $(o, s - 1)$, (o, s) et $(o, s + 1)$ (C. Evans (2009))	26
Figure 2.11	Calcul d'une aire sur une image en représentation intégrale (C. Evans (2009))	27
Figure 2.12	Approximation des Laplaciens de Gaussiennes (LoG) par des filtres de type <i>box filters</i> . Les LoG de la ligne supérieure L_{xx} , L_{yy} et L_{xy} ont respectivement pour approximations les filtres D_{xx} , D_{yy} et D_{xy} de la ligne inférieure (C. Evans (2009))	29

Figure 2.13	Construction des vecteurs descripteurs. Les réponses à des ondelettes de Haar sont calculées en 25 points de chacune des 16 régions qui constituent le voisinage du point d'intérêt (C. Evans (2009))	29
Figure 2.14	Modèle de caméra de type sténopé (F. Cheriet (2011a))	32
Figure 2.15	Géométrie épipolaire. O et O' sont les centres des caméras, p et p' les images du point P sur les plans H et H' , e et e' les épipoles et l et l' les lignes épipolaires (F. Cheriet (2011b))	33
Figure 2.16	Construction d'un graphe à partir d'une image (Droppelmann <i>et al.</i> (2010))	37
Figure 2.17	Coupe de graphe (Droppelmann <i>et al.</i> (2010))	37
Figure 2.18	Recherche des pixels correspondants le long des lignes épipolaires. Les mesures de similarité sont calculées pour une fenêtre centrée sur les points candidats dans une zone de recherche restreinte (Xu et Chutatape (2006))	40
Figure 3.1	Les images de fond d'œil sont sujettes à du bruit de hautes fréquences.	44
Figure 3.2	Application d'un filtrage de type fenêtre de Hann sur une image de fond d'œil	44
Figure 3.3	Mise en correspondance de points d'intérêt par l'algorithme OpenSURF (C. Evans (2009))	46
Figure 3.4	Exemple de recalage quadratique d'images de fond d'œil	48
Figure 3.5	Détection de la zone maculaire	50
Figure 3.6	Positionnement du centre de la fovéa	50
Figure 3.7	Détection du centre du disque optique	51
Figure 3.8	Construction d'une ellipse dont les foyers sont les centres du disque optique et de la fovéa	52
Figure 3.9	Recherche de points correspondants sur les ellipses construites	53
Figure 3.10	Images stéréoscopiques d'un fond d'œil	54
Figure 3.11	Images stéréoscopiques d'un fond d'œil après rectification épipolaire	54
Figure 4.1	Images de fond d'œil acquises suivant différents angles de vue.	58
Figure 4.2	Résultat obtenu après recalage quadratique des images de la figure 4.1	59
Figure 4.3	Recalage d'images de dimensions différentes.	60
Figure 4.4	Mosaique construite par recalage de trois images de fond d'œil	62
Figure 4.5	Recalage d'images d'un patient prises lors de différents examens. Des lésions sont apparues dans la zone encadrée en rouge entre les deux examens.	63

Figure 4.6	Position des centres de la fovéa (en rouge) et du disque optique (en vert) après recalage d'images de différents fonds d'œil. Les disques ont un rayon égal à l'écart-type relatif à la position des points après recalage par rapport à leur position moyenne.	65
Figure 4.7	Carte de la répartition des intensités sur les images de fond d'œil . .	67
Figure 4.8	Carte de la répartition moyenne des intensités sur les images de fond d'œil.	67
Figure 4.9	Carte des écarts-types locaux de la valeur des pixels du canal vert de l'atlas en deux dimensions.	68
Figure 4.10	Carte de la probabilité des pixels des images de fond d'œil d'appartenir au réseau vasculaire.	70
Figure 4.11	Carte des disparités entre deux images stéréoscopiques d'un fond d'œil.	72
Figure 4.12	Reconstruction 3D de fonds d'œil à un facteur près.	73

LISTE DES SIGLES ET ABRÉVIATIONS

ARIC	Atherosclerosis Risk in Communities Study
BMSE	Blue Mountain Eye Study
CSME	Clinical Significant Macular Edema
CT	Computed Tomography
DMLA	Dégénérescence Maculaire Liée à l'Âge
IRMf	Imagerie par Résonance Magnétique fonctionnelle
LoG	Laplacian of Gaussian
LSD	Line Segment Detector
NIH	National Institutes of Health
OCT	Optical Coherence Tomography
OMS	Organisation Mondiale de la Santé
RANSAC	RANdom SAMple Consensus
RGB	Red, Green, Blue
SGBM	Semi Global Block Matching
SIFT	Scale Invariant Feature Transform
SURF	Speeded Up Robust Features
SVD	Singular Value Decomposition

CHAPITRE 1

INTRODUCTION

La pratique de l’ophtalmologie a rapidement évolué au cours des dernières décennies. De nouvelles technologies ont été intégrées à l’instrumentation et aux modalités d’imagerie, ces dernières profitant également des progrès réalisés en traitement et analyse d’images.

Les praticiens sont amenés à acquérir une quantité de données croissante afin d’établir leurs diagnostics. Les images médicales constituent une part importante de ces données. L’analyse des images est sujette à des contraintes de temps et de reproductibilité. En effet, l’établissement des diagnostics demande parfois beaucoup de temps aux praticiens et il arrive que lorsqu’une image est analysée par deux cliniciens différents ou un même praticien à deux instants distincts, les diagnostics diffèrent.

Le développement d’outils d’analyse automatique des images médicales est donc un enjeu important, notamment en ce qui concerne les images de la surface de la rétine, appelées communément images de fonds d’œil. L’une des particularités de ces images est qu’elles sont les seules à permettre une observation directe de vaisseaux sanguins chez les êtres humains de manière non-invasive. Obtenir une segmentation automatique précise du réseau vasculaire est l’un des enjeux actuels de l’analyse des images de fond d’œil.

Les algorithmes de segmentation ont pour objectif de localiser des régions d’intérêt dans des images. Les premières méthodes de segmentation à avoir été proposées peuvent être divisées en deux catégories. La première consiste à déterminer les contours des objets d’intérêt et requiert l’utilisation de techniques de recherche de frontières telles que le filtre de Marr-Hildreth (Marr et Hildreth (1980)) ou le filtre de Canny (Canny (1986)). La deuxième approche repose sur la sélection de régions, par seuillage des intensités (Otsu (1979)) ou en utilisant des algorithmes de croissance de région tels que celui proposé par Brice et Fennema (1970). Ces méthodes ont par la suite été améliorées dans le but d’être adaptées à des applications spécifiques. Des travaux ont notamment été réalisés afin d’implémenter des solutions propres à la segmentation du réseau vasculaire rétinien. Les algorithmes de segmentation basés sur l’étude des intensités afin d’extraire le réseau vasculaire sont notamment utilisés (Zhang et Liu (2007), Wu *et al.* (2009)). Ils sont cependant très dépendants de la qualité des images et échouent souvent dans la recherche des vaisseaux les plus fins qui se situent souvent dans des régions à faible contraste. Les approches par apprentissage machine reposent sur la construction de classifieurs qui permettent de reconnaître les vaisseaux en utilisant des structures aptes à les distinguer. Des travaux sont actuellement réalisés afin de trouver des

structures adaptées aux vaisseaux les plus fins (Kovács et Hajdu (2011), Sofka et Stewart (2006)). Les modèles paramétriques de type contours actifs visent à segmenter les images en appliquant des forces sur des contours afin de les accoler aux frontières du réseau vasculaire (D.Rivest-Hénault *et al.* (2010)). Les principales difficultés de cette approche résident dans l’initialisation des contours et le choix des forces appliquées. Des méthodes de type suivi de vaisseaux ont également été étudiées (Xu et Luo (2010), Yedidya et Hartley (2008)). Elles consistent à suivre le réseau vasculaire en cherchant son orientation de manière itérative. Cette technique permet d’obtenir des informations sur la position des lignes centrales du réseau vasculaire et sur la largeur des vaisseaux. Elle est cependant sensible à des perturbations comme la présence de reflets ou une diminution du contraste dans les images. Des travaux ont également été réalisés afin de combiner plusieurs des techniques citées précédemment. Des algorithmes procédant à la recherche des branches principales du réseau vasculaire par analyse des intensités suivie d’un apprentissage machine afin de segmenter les vaisseaux les plus fins ont notamment été proposés (Cinsdikici et Aydin (2009)).

Une autre méthode de segmentation est basée sur l’utilisation d’atlas. Un atlas est un modèle construit par alignement d’images dans lesquelles les structures d’intérêt ont été pré-segmentées. Un modèle moyen de ces structures au sein d’une population est ainsi construit, et les variabilités inter-sujets sont évaluées au sein de cette population. L’atlas peut alors être aligné sur des images non segmentées afin d’y localiser les objets d’intérêt. Des atlas d’images médicales ont notamment été construits à partir d’images du cœur dans le cadre du projet *euHeart initiative* (Hoogendoorn (2013)) ou d’images cérébrales avec le projet *BigBrain* (Amunts *et al.* (2013)).

Seuls Lee *et al.* (2010) ont proposé une méthode de construction d’atlas d’images de fonds d’œil jusqu’à présent afin de fournir une carte d’intensités qui représente la répartition des trois principales structures anatomiques visibles sur les images rétinienne que sont le disque optique, la fovéa et les branches principales du réseau vasculaire. Cette carte d’intensités est obtenue à l’aide d’un alignement d’images rétinienne de sujets différents, cet alignement étant appelé recalage. Des points sont sélectionnés manuellement sur les images afin de déterminer les déformations à leur appliquer. L’une des principales limites de la méthode proposée par Lee *et al.* est que les déformations sont déterminées de manière semi-automatique, les points étant sélectionnés par un spécialiste des images de fonds d’œil. De plus, les déformations appliquées ne sont évaluées qu’autour des principales structures anatomiques. Il serait donc utile d’implémenter un algorithme de recalage automatique d’images de différents sujets tenant compte des particularités géométriques de la rétine. En outre, l’atlas construit par Lee *et al.* est en deux dimensions, ce qui implique une dépendance vis-à-vis de la configuration dans laquelle les images utilisées pour le constituer ont été acquises. Une extension de l’atlas

à la troisième dimension pourrait être considérée. Il faudrait alors développer un algorithme de recalage d'images intra-sujet dans le but de ne conserver que des déformations liées à la profondeur de la rétine entre différentes images. Le recalage intra-sujet pourrait alors être utilisé afin d'effectuer un suivi de pathologies évolutives chez des patients ou de construire des mosaïques d'images qui permettraient d'étendre les régions observées en juxtaposant des images acquises suivant différents angles de vue.

L'objectif de ce projet est de développer des algorithmes de recalage intra et inter-sujet et d'étudier la possibilité de reconstruire la surface des rétines en trois dimensions afin de proposer une solution concrète aux suggestions formulées précédemment et de construire un atlas qui puisse faciliter la segmentation du réseau vasculaire sur les images de fonds d'œil.

Pour ce faire, nous proposons de développer un algorithme de recalage qui permette de tenir compte des principales transformations induites entre les acquisitions de plusieurs images rétinienne. Ces déformations sont majoritairement dues au fait que les yeux peuvent avoir des positions différentes par rapport aux caméras entre plusieurs acquisitions, d'où l'introduction de translations, rotations et changements d'échelles ainsi qu'au fait que les images soient obtenues par projection de la structure sphérique qu'est la rétine sur un plan image, ce qui implique des transformations d'ordre quadratique (d'ordre 2). Cet algorithme est d'abord implémenté dans le cas d'un recalage d'images intra-sujet avant d'être adapté au recalage inter-sujet en vue de construire l'atlas. La principale difficulté à surmonter pour recalager des images de sujets différents est qu'elles n'ont que peu de structures communes. Une fois l'atlas formé, une analyse de la répartition des structures anatomiques est effectuée et une application visant à faciliter spécialement la segmentation du réseau vasculaire est proposée. Nous proposons finalement d'étudier la possibilité d'une extension de l'atlas à la troisième dimension en utilisant des images stéréoscopiques de fonds d'œil et en considérant que celles-ci sont acquises dans une configuration semblable à la géométrie épipolaire.

1.1 Organisation du mémoire

Ce mémoire suit une organisation classique. La démarche suivie dans le cadre du projet est présentée dans quatre chapitres. Le premier consiste en une revue de la littérature existante en ce qui concerne les étapes nécessaires à la construction d'un atlas d'images médicales en trois dimensions, à savoir un aperçu rapide des structures anatomiques impliquées, une présentation des méthodes de recalage d'images ainsi qu'une explication des techniques de reconstruction en trois dimensions à partir d'images stéréoscopiques. La méthode adoptée est décrite dans le deuxième chapitre. Les résultats obtenus sont rapportés dans le troisième chapitre, accompagnés de discussions. Une conclusion est proposée dans le dernier chapitre.

CHAPITRE 2

REVUE DE LITTÉRATURE

Ce chapitre de revue de littérature est constitué de quatre parties. La première fournit des renseignements en ce qui concerne l'anatomie de l'œil humain ainsi que de la rétine et décrit les principales pathologies qui peuvent les affecter tout en étant visibles sur les images de fond d'œil. La deuxième partie consiste en une présentation des systèmes d'imagerie de la rétine à travers un bref historique et une distinction entre les modalités d'imagerie en deux et trois dimensions. Suit une revue des connaissances relatives au recalage d'images en deux dimensions. Les informations relatives à la reconstruction en trois dimensions figurent quant à elles dans la quatrième partie.

2.1 Anatomie

Le présent projet repose sur l'exploitation d'images de fond d'œil, qui sont des photographies de la rétine. Cette section a pour objectif de présenter les caractéristiques anatomiques principales de l'œil humain et la manière dont il agit sur les rayons lumineux lors de l'acquisition des images. La figure 2.1 présente l'anatomie de l'œil de manière simplifiée ainsi que les différentes couches qui constituent la rétine, tissu qui nous intéresse particulièrement.

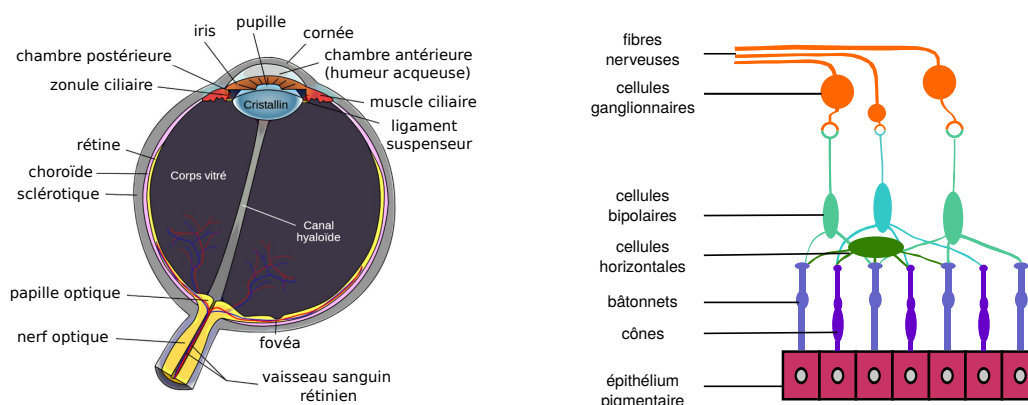
La partie visible depuis l'extérieur de l'œil est constituée de *la cornée* qui couvre environ 1/5 de sa surface et fait office de première lentille convergente, *la sclérotique*, membrane opaque et blanche, *l'iris*, diaphragme qui permet de contrôler la quantité de lumière qui entre dans l'œil en se contractant ou en se dilatant et de *la pupille*, orifice au centre de l'iris qui laisse passer la lumière.

Un rayon lumineux traverse successivement *la cornée*, la *chambre antérieure*, *la pupille* et *le cristallin* dont la forme peut se modifier de manière à bien focaliser la lumière sur la rétine après qu'elle ait voyagé à travers *le corps vitré*.

Une autre structure importante de l'œil est *la choroïde*. Celle-ci est située entre la sclérotique et le corps vitré. Elle est constituée de nombreux éléments vasculaires et nerveux et est essentielle en ce qui concerne l'irrigation de l'œil (Abramoff *et al.* (2010), Santallier *et al.* (2008)).

2.1.1 Structure de la rétine

La rétine est constituée de plusieurs couches (voir figure 2.1b). Elles sont ici présentées dans l'ordre de leur activation suite à la réception de lumière, qui va des couches les plus



(a) Anatomie de l'œil et couches rétiniennes (Marieb (1998)) (b) Détail des couches cellulaires rétiniennes (Abramoff *et al.* (2010))

Figure 2.1: Anatomie de l'œil humain

profondes, vers l'arrière de la tête, aux plus avancées. *L'épithélium pigmentaire* est rattaché à la choroïde. Son nom lui vient de la couleur rose-orangée que lui donnent les grains de mélanine qui s'y trouvent. Cette couche permet les échanges électrolytiques nécessaires à l'alimentation des cellules de la rétine. La vitamine A y est notamment stockée. Cette dernière permet de générer la *rhodopsine*, protéine qui active les cellules photoréceptrices en présence de lumière.

Ces *cellules photoréceptrices* sont en contact avec l'épithélium pigmentaire. Il en existe deux types : les *bâtonnets* et les *cônes*. Les premiers permettent d'évaluer la luminosité de la scène observée. Ils sont les seuls à s'activer en présence d'une faible intensité lumineuse, un certain seuil devant être atteint pour solliciter les cônes responsables de la perception des couleurs (il en existe trois chez l'homme, que l'on nomme communément rouge, bleu et vert car leur excitation maximale se situe à des longueurs d'ondes différentes).

Les *cellules horizontales* agissent en tant que cellules de contrôle en évaluant les signaux émis par les bâtonnets ou cônes dans un certain voisinage de manière à ajuster la perception des contrastes.

La *couche plexiforme externe* désigne la zone de jonction entre les synapses des cellules photoréceptrices et des *cellules bipolaires*. Celles-ci sont également reliées aux *cellules ganglionnaires*, la *couche plexiforme interne* constituant la zone de jonction entre leurs synapses.

Il existe deux types de cellules bipolaires : les cellules bipolaires de bâtonnets qui transmettent les signaux émis par plusieurs bâtonnets à une cellule ganglionnaire et les cellules bipolaires de cônes qui agissent de même avec les cônes. Les cellules ganglionnaires sont le point de départ des fibres optiques qui se rejoignent au niveau de *la papille* pour former le nerf optique qui transmet les informations visuelles au cerveau.

Sur les photographies de fond d'œil, trois structures sont à remarquer (voir figure 2.2) :

- *La papille* : zone circulaire généralement claire qui correspond au point d'arrivée du nerf optique et des principaux vaisseaux sanguins qui irriguent la rétine. Elle est également désignée par le terme de *tache aveugle* car aucun cône ou bâtonnet ne se trouve à cet endroit de la rétine.
- *La fovea* : zone où sont concentrés les cônes et un grand nombre de bâtonnets. Ce point est responsable d'une vision centrale nette et colorée.
- *Le réseau vasculaire* : le réseau vasculaire primaire correspond aux branches qui sortent de la tache aveugle avant de se ramifier en branches du réseau secondaire, tertiaire et éventuellement d'autres niveaux, le diamètre des vaisseaux diminuant après chaque ramification.

2.1.2 Pathologies détectables à partir des images de fond d'œil

La rétine est le seul tissu du corps humain qui permette une observation du réseau vasculaire de manière non-invasive. Cette particularité fait que les images du fond d'œil peuvent renseigner les praticiens sur de nombreuses pathologies. Il est possible d'y détecter des symptômes de maladies qui touchent essentiellement la rétine ainsi que d'autres qui influent sur plusieurs régions anatomiques voir l'ensemble du corps du patient.

Pathologies propres à la rétine

Glaucome Le glaucome consiste en une détérioration du nerf optique (détérioration des cellules ganglionnaires et de leurs axones), ce qui a pour conséquence une perte progressive de l'acuité visuelle. Cette détérioration est causée par une pression trop élevée à l'intérieur de l'œil. Le nerf optique se creuse et une diminution du rapport entre les diamètres de la papille et du nerf apparaît (ce rapport est appelé *rapport cup to disc* par les spécialistes). Les praticiens évaluent et suivent la valeur de ce rapport dans le temps afin de réaliser leur diagnostic. Le nerf optique se reconnaît sur les images de fond d'œil comme une zone plus claire à l'intérieur de la papille comme le montre la figure 2.3. Des opérations de dépressurisation ou des interventions chirurgicales sont nécessaires afin de ralentir le développement de la pathologie. Le glaucome se développe généralement autour de 40 ans et touche 4,5 millions de personnes dans le



Figure 2.2: Image de fond d'œil

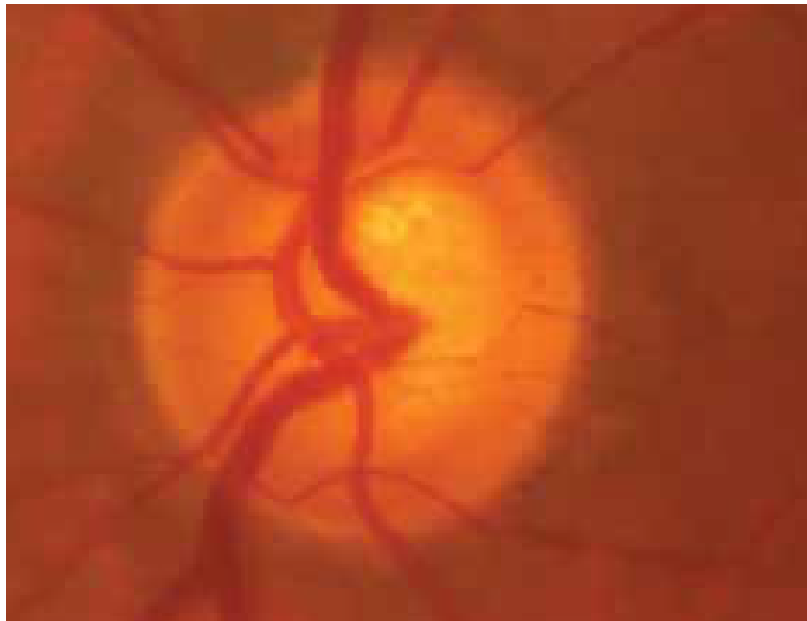


Figure 2.3: Image de papille. Le nerf optique est visible sous la forme d'une excavation physiologique dont le diamètre représente $\frac{6}{10}$ de celui de la papille (Y. Lachkar and E. Sellem (2004))

monde (15% des cas de cécité) d'après l'OMS (Organisation Mondiale de la Santé) (OMS (2013)).

Dégénérescence maculaire liée à l'âge (DMLA) La DMLA est la première cause de cécité dans les pays développés et la troisième dans le monde (8,7% des cas de cécité selon l'OMS). Elle apparaît généralement après 50 ans. Elle est caractérisée par une détérioration de l'épithélium pigmentaire produite par une déficience de la circulation sanguine dans le voisinage de la fovea. Cette détérioration provoque l'apparition de dépôts de couleur jaune à la surface de la rétine, appelés *drusen*. En résulte une perte progressive de la vision centrale. Le développement de cette maladie peut être ralenti par la prise de compléments alimentaires ou l'administration de régulateurs de la circulation sanguine.

Pathologies systémiques

Diabète Le diabète est un trouble métabolique qui a des conséquences sur diverses parties de l'organisme. Il se caractérise par une production ou une action défectueuse de l'insuline dans l'organisme. Le métabolisme des glucides, lipides et protéines est alors perturbé, ce qui implique une hyperglycémie chronique (on désigne par *hyperglycémie* une glycémie supérieure à 7mmol/L à jeûn). La prévalence mondiale du diabète est de 211 millions de personnes, essentiellement dans les pays développés. On distingue deux catégories de diabète :

- Le diabète de type 1 : il se manifeste dès l'enfance et est lié à une déficience de la sécrétion d'insuline par le pancréas. Des injections quotidiennes d'insuline doivent être réalisées durant toute la vie du patient.
- Le diabète de type 2 : il est lié principalement aux habitudes de vie (obésité, manque d'activité physique, mauvaise alimentation). Il représente environ 90% des cas de diabète dans le monde. Une modification des habitudes de vie ou des injections d'insuline permettent d'atténuer les effets de cette maladie.

L'hyperglycémie liée au diabète endommage les vaisseaux sanguins, les cellules nerveuses, les reins ainsi que les yeux, notamment par le biais d'une complication appelée rétinopathie diabétique.

Rétinopathie diabétique La rétinopathie diabétique résulte de troubles vasculaires rétiens comme des dilatations ou occlusions de vaisseaux qui peuvent provoquer des œdèmes maculaires diabétiques (la paroi des vaisseaux devient poreuse et du sang se répand dans la rétine) et des ischémies (apparition de néo-vaisseaux). On appelle CSME (Clinical Significant Macular Edema) une cécité provoquée par des œdèmes proches de la fovea. Un amincissement de cette zone y induit en effet une atrophie et une apoptose (mort cellulaire) des tissus. Cette

pathologie est la seconde cause de cécité dans les pays développés. 100% des diabétiques de type 1 développent cette complication dans les dix ans qui suivent les premiers symptômes de leur diabète et 33% des diabétiques de type 2 après vingt ans.

Un suivi régulier est indispensable afin de pouvoir effectuer un traitement au laser dès les premiers signes d'apparition de la rétinopathie. La photocoagulation est en effet l'intervention la plus rentable en terme de coût / efficacité (OMS (2013)). Lors de cette opération, un laser, généralement de type Argon, est focalisé sur la rétine à l'aide d'une lentille pour forcer la coagulation de vaisseaux desquels s'échappe du sang ou qui sont potentiellement dangereux. Ce traitement est également utilisé pour détruire des zones malades qui pourraient provoquer une diminution des facultés visuelles si elles n'étaient pas détruites.

Maladies cardiovasculaires Le fait que le réseau vasculaire soit observable sur les images de rétines a amené des équipes de recherche à se poser la question de leur utilisation dans la construction de méthodes de quantification objectives du risque de maladies cardiovasculaires. Ces méthodes pourraient remplacer ou être intégrées au *score de Framingham* qui vise à établir la probabilité qu'un sujet a de subir un accident cardiovasculaire au cours des dix prochaines années. L'étude *ARIC* (Atherosclerosis Risk in Communities Study) a permis d'établir un lien entre une diminution du ratio entre les diamètres des artères et des veines visibles sur les images de fond d'œil de 10 000 sujets et le risque d'incident cardiovasculaire, mais uniquement chez les femmes, aucun lien n'ayant pu être constaté chez les hommes entre ces deux propriétés. Ce résultat se rapproche des observations cliniques qui tendent à montrer que la microvasculature joue un rôle plus important dans les maladies cardiovasculaires chez les femmes que chez les hommes. La principale limite de l'étude ARIC est qu'elle ne précise pas si les ratios entre les diamètres des artères et des veines sont faibles car les artères ont un diamètre réduit ou parce que les veines ont un diamètre important. L'étude *BMES* (Blue Mountain Eye Study), réalisée sur un échantillon de 3654 personnes suivies sur une durée de 5 à 10 ans, montre une corrélation entre l'augmentation du diamètre des veines et le risque de subir un incident cardiovasculaire. Selon cette même étude, une diminution du diamètre des artères est un facteur de risque chez les femmes (Liew et Wang (2011)).

Le nombre d'études réalisées à ce sujet étant limité, l'utilisation des images de fond d'œil dans l'aide au diagnostic des maladies cardiovasculaires ne fait pas consensus aujourd'hui parmi les praticiens. Les recherches se poursuivent pour tenter de construire des mesures quantitatives qui puissent être intégrées au score de Framingham. En effet, les propositions qui ont été formulées jusqu'à présent ne semblent pas augmenter de manière significative la validité de ce score.

2.2 Systèmes d'imagerie de l'œil

Les pathologies mentionnées dans la section précédente ont pu être découvertes et étudiées grâce aux avancées technologiques qui ont permis le développement de modalités d'imagerie adaptées à l'observation de l'œil. Un historique de ces modalités est proposé dans la section suivante, ainsi qu'une description des principaux systèmes d'imagerie 2D et 3D utilisés aujourd'hui.

2.2.1 Historique

La première observation de la rétine d'un être vivant est attribuée à J. Mery (1645-1722). Celui-ci aurait constaté que les vaisseaux sanguins présents sur la rétine d'un chat plongé vivant dans l'eau étaient visibles depuis l'extérieur du récipient (Abramoff *et al.* (2010)).

Le premier appareil utilisable sur les êtres humains doit quant à lui son invention à J. E. Purkyne. Il s'agit de *l'ophtalmoscope*, encore utilisé de nos jours par les médecins bien que son design ait légèrement évolué. C. Babbage et H. von Helmholtz y ont en effet apporté des améliorations respectivement en 1845 et 1851. Un examen du fond d'œil à l'aide de cet appareil consiste à envoyer un faisceau lumineux d'une largeur adaptée en direction de la rétine du patient et à observer celle-ci à travers une lentille de focale adéquate. Aujourd'hui, les sources lumineuses sont introduites dans le manche de l'ophtalmoscope et le praticien peut choisir la lentille qui correspond à l'amétropie du patient de manière à pouvoir observer nettement la zone souhaitée.

La première photographie de la rétine a été réalisée en 1891 par l'ophtalmologiste O. Gerloff. L'acquisition de telles photographies a été encouragée car elle permet d'éviter aux praticiens de trop s'approcher du visage de leurs patients et de diminuer l'influence des contraintes de temps dans la réalisation des diagnostics qui peut ainsi être différée. Les appareils photographiques ont connu quelques évolutions avant que A. Gullstrand ne propose en 1910 un concept de caméra de fond d'œil tel qu'on les connaît aujourd'hui. Gullstrand reçu le prix Nobel de physiologie de 1911 pour cette invention.

Il est également possible d'étudier la rétine de manière fonctionnelle. L'imagerie angiographique est la plus répandue, le principal marqueur utilisé étant la fluorescéine. De plus en plus de questions se posent cependant quant à la nocivité et la rentabilité de cette modalité d'imagerie.

La découverte de pathologies telles que le glaucome ont par la suite suscité le besoin d'obtenir des informations en trois dimensions sur les rétines. L. Allen a ainsi proposé d'acquérir des images stéréoscopiques du fond d'œil en 1964. Les avancées réalisées dans le domaine de l'imagerie optique ont ensuite permis d'introduire l'imagerie confocale et la

technique OCT (Optical Coherence Tomography ou tomographie par cohérence optique) en ophtalmologie. La rétine a d'ailleurs été le premier tissu anatomique imagé à l'aide de la modalité OCT en 1993 (Swanson *et al.* (1993)).

Le développement de l'analyse d'images depuis les années 1950 et les progrès réalisés en informatique et en imagerie numérique ont conduit M. Matsui à entreprendre les premières tentatives de segmentation du réseau vasculaire rétinien en 1973 (Matsui *et al.* (1973)). Le nombre de publications concernant l'imagerie de fond d'œil est en constante augmentation depuis lors, et a connu un fort essor dans les années 1990.

2.2.2 Systèmes d'imagerie 2D

M.D. Abramoff définit l'imagerie de fond d'œil comme suit (Abramoff *et al.* (2010)) :

Définition 2.1. *Processus par lequel une représentation en deux dimensions des tissus semi-transparents en trois dimensions de la rétine est obtenue sur un plan image par réflexion de lumière.*

L'objectif de cette section est de présenter les modalités d'imagerie de la rétine en deux dimensions les plus courantes, à savoir la photographie de fond d'œil et l'imagerie angiographique. Ces deux techniques ont connu un important essor dans les années qui ont suivies leur découverte, la transparence des yeux étant une invitation à leur utilisation afin d'imager directement les vaisseaux sanguins de la rétine, chose qu'il n'est possible de réaliser que sur cette partie de l'organisme.

Photographie de fond d'œil

Imager la rétine implique de faire face aux difficultés suivantes :

- les rayons lumineux destinés à l'illumination de la rétine et ceux qu'elle réfléchit en destination du capteur partagent un même chemin optique à travers la pupille.
- la rétine est peu réfléchissante.
- des aberrations s'accumulent au cours du trajet des rayons lumineux dans l'œil, ce qui bruite les images.
- les réflexions par la cornée ou les constituants du système optique sont, si elles ne sont pas corrigées, plus importantes que la lumière réfléchie par la rétine.

On distingue deux catégories de caméras de fond d'œil :

- les caméras avec source d'éclairage interne
- les caméras à source externe.

Le principe général de fonctionnement de ces deux types de caméras est le même (DeHoog et Schwiegerling (2009), voir figure 2.4). La source lumineuse (11a, 12b) est imagée à travers

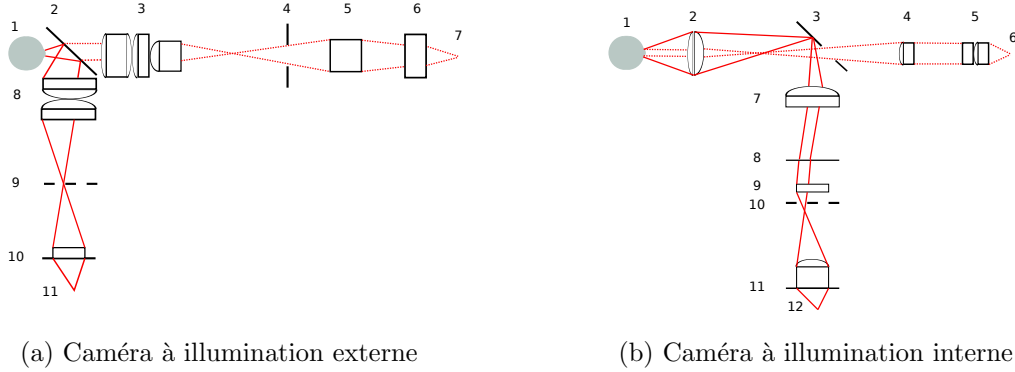


Figure 2.4: Principe de fonctionnement des caméras de fond d'œil à illumination externe (a) et illumination interne (b), ((a) 1. œil ; 2. miroir semi réfléchissant ; 3., 5., 6., 8., 10. lentilles ; 4. diaphragme ; 7. image de la source ; 9. obscurateur central ; 11. source, (b) 1. œil ; 2., 4., 5., 7., 8., 9., 11. lentilles ; 3. miroir à trou central ; 6. image de la source ; 10. obscurateur central ; 12. source) (DeHoog et Schwiegerling (2009))

la pupille afin d'éclairer la rétine. Dans le plan conjugué de la pupille le long du chemin emprunté par cette lumière d'éclairage, un obscurateur central (9a, 10b) est placé dans le but d'empêcher les réflexions au niveau de la cornée. Le système de formation des images est constitué d'un objectif (3a, 2b) ainsi que de lentilles de zoom (5a, 6a, 4b, 5b). Un système d'accomodation (7a, 6b) est positionné en fin de parcours de manière à corriger les erreurs de réfraction propres au sujet.

Les caméras sont catégorisées en fonction de la façon dont sont combinés les rayons qui servent à l'éclairage de la rétine et à son imagerie. Les caméras à illumination externe utilisent un miroir semi-réfléchissant pour séparer puis réunir les rayons alors que les caméras à illumination interne mettent en jeu un miroir avec un trou central pour réaliser les mêmes opérations. Les avantages principaux des caméras à illumination interne sont qu'elles permettent une plus grande ouverture de leurs constituants optiques, donc d'accéder à un plus grand champ de vision. Elles fournissent des images de qualité similaire aux autres caméras tout en ayant recours à moins de composants, donc se vendent à un prix inférieur. Les caméras à illumination externe permettent quant à elles une meilleure suppression des réflexions sur la cornée et les constituants optiques ainsi qu'une meilleure correction des dérives chromatiques. Selon le NIH (National Institutes of Health), les deux types de caméras produisent des images de qualité satisfaisante et peuvent être utilisés indifféremment par les praticiens.

Les photographies de fond d'œil utilisées dans le cadre de ce projet ont des tailles qui varient entre 2 *megapixels* et 18 *megapixels* selon les modèles de caméras utilisés. Les angles

d'ouverture s'étalent quant à eux entre 45° et 60° . La résolution (mesurée en pixels par degré dans le cas des caméras de fond d'œil, de l'ordre de 50 pixels/deg dans le cas de la caméra la plus utilisée pour la réalisation du projet) qu'il est possible d'obtenir avec cette modalité d'imagerie augmente lorsque la densité de pixels par unité d'angle imagée croît. Le fait que le système d'imagerie soit constitué en partie du système optique de l'œil des sujets implique toutefois une limite à cette résolution. Des équipes de recherche travaillent actuellement au développement de systèmes optiques plus sophistiqués, comme les miroirs déformables, dans le but d'augmenter la résolution des images de fond d'œil (Dubra *et al.* (2011)). Ces solutions, qui permettent d'atteindre des résolutions de l'ordre du micromètre, sont encore peu démocratisées aujourd'hui.

Imagerie angiographique

Le premier écrit rapportant une observation de la rétine par angiographie est attribué à Novotny et Alvis (1961). Leur méthode consiste à acquérir des photographies de fond d'œil à intervalles réguliers, la première étant une image de référence prise avant injection de fluoresceine et d'une solution saline dont l'objectif est d'augmenter la visibilité des vaisseaux sur les images suivantes. Les acquisitions se font ensuite toutes les $12s$ pendant $3min30s$, durée de la présence de la fluoresceine dans la circulation rétinienne. La longueur d'onde qui permet une activation maximale se situe dans le spectre visible, à $490nm$ et la lumière émise en retour par la fluoresceine se trouve dans le vert à $520nm$. Des filtres sont donc placés en avant du flash utilisé lors de l'acquisition des images et du capteur de la caméra afin d'obtenir un contraste maximal entre les vaisseaux et le fond des photographies. Dans le cas d'yeux sains, les auteurs ont pu distinguer des phases de remplissage séparées pour les veines et les artères ainsi que la formation de taches lors du remplissage des veines sur le reste du fond d'œil, liées selon eux à la circulation choroidale, plus lente. L'imagerie angiographique a permis la découverte du phénomène de néovascularisation associé à l'hypertension et à la rétinopathie diabétique. Cette modalité permet donc d'étudier la circulation sanguine dans la rétine au cours du temps, sa principale limite étant le fait qu'elle ne permette pas d'observer les éventuelles hémorragies ou œdèmes présents sur le fond d'œil.

2.2.3 Systèmes d'imagerie 3D

Des pathologies telles que le glaucome ont suscité le besoin de développer des modalités d'imagerie de la rétine en trois dimensions. L'OCT est utilisée par les praticiens, comme le laser à balayage toutefois moins répandu. Ces deux techniques sont présentées dans la section qui suit.

Tomographie à cohérence optique

L'OCT est une modalité d'imagerie en trois dimensions en temps réel. L'objectif de cette technique est de trouver la profondeur à laquelle se trouvent les interfaces entre différentes couches de tissus superposées.

Le principe est proche de celui de l'imagerie ultrasonore (Fujimoto *et al.* (2000)). Des ondes lumineuses sont émises en direction du tissu à imager et leur temps d'échos est mesuré afin de connaître la profondeur des interfaces. La subtilité réside dans le fait que la lumière se propage très rapidement et qu'il n'existe aucun dispositif capable de mesurer directement ces temps d'échos. Il est donc nécessaire de procéder à des mesures d'interférométrie en couplant un signal de référence au parcours connu à celui issu des réflexions dans les tissus à observer. Il s'agit donc d'un dispositif de type interféromètre de Michelson qui est présenté dans la figure 2.5.

La résolution en profondeur est inversement proportionnelle à la largeur de bande de la source lumineuse. En effet, plus la largeur de bande du faisceau émis par la source est élevée, plus la longueur de cohérence entre deux signaux couplés est réduite. Si Δz représente cette résolution, λ la longueur d'onde centrale de la source et $\Delta\lambda$ sa largeur de bande :

$$\Delta z = (2\ln 2/\Pi)(\lambda^2/\Delta\lambda) \quad (2.1)$$

La résolution transverse dépend quant à elle du système optique utilisé. Si f désigne la distance focale de la lentille équivalente de ce système et d la taille d'un point issu de la focalisation d'un rayon lumineux de longueur d'onde λ , alors la résolution transverse Δx est donnée par :

$$\Delta x = (4\lambda/\Pi)(f/d) \quad (2.2)$$

Le facteur d'atténuation des tissus augmente de plus avec la longueur d'onde. Les sources lumineuses utilisées en tomographie à cohérence optique émettent donc généralement des ondes dans le proche-infrarouge.

L'acquisition d'une image OCT s'effectue en mode B, c'est-à-dire que la profondeur des différents tissus est mesurée sur un segment transverse. Pour obtenir une image en trois dimensions, il est nécessaire de répéter ces acquisitions en deux dimensions et de procéder à une interpolation entre les profils voisins. Les appareils actuels sont capables de fournir des images 3D-isotropiques (qui ont une résolution identique dans chacune des trois dimensions, de l'ordre de la dizaine de micromètres) en un temps d'environ 2s.

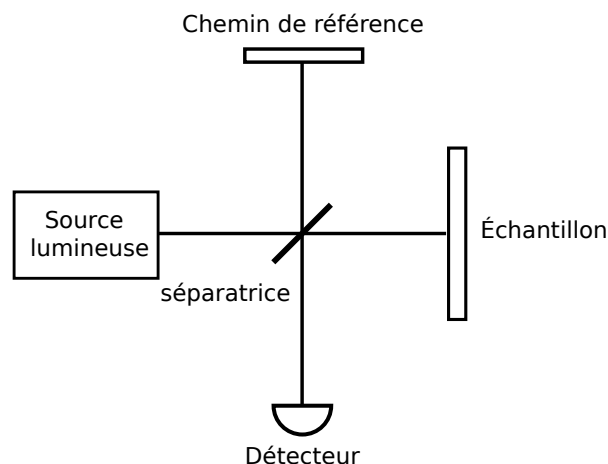


Figure 2.5: Schéma illustrant la méthode d'acquisition des images OCT (Fujimoto *et al.* (2000))

Ophtalmoscopie par laser à balayage

L'ophtalmoscopie par laser à balayage repose sur le principe de la microscopie confocale, à ceci près que des balayages axiaux et transverses y sont ajoutés. La source lumineuse est monochrome et les tissus sont observés point par point. Un sténopé est en effet placé au niveau du plan conjugué du plan focal image de l'objectif du système optique. Ceci permet de sélectionner la profondeur du point que l'on souhaite imager le long de l'axe optique et de supprimer la lumière parasite issue des points voisins. L'image en trois dimensions est obtenue par assemblage des images des points acquis. La profondeur accessible à l'aide de cette technologie est inférieure à celle atteinte par l'OCT. La résolution latérale (nombre de points acquis par unité de surface de l'objet imagé) est toutefois supérieure.

Ont été présentées dans cette section les principales modalités d'imagerie de la rétine en deux et trois dimensions. Les images obtenues à l'aide des instruments correspondants sont la plupart du temps soumises à des traitements avant d'être analysées par les praticiens. Puisque les zones observées sur les images rétinienne sont restreintes et que les sujets bougent entre plusieurs acquisitions, l'un des traitements régulièrement appliqués consiste en la superposition de différentes images. Les algorithmes de recalage qui permettent d'effectuer cette opération sont détaillés dans la section qui suit.

2.3 Recalage d'images

Les praticiens souhaitent souvent pouvoir suivre des pathologies évolutives chez leurs patients au cours du temps ou observer des zones plus étendues que celles qui sont visibles sur les images acquises à l'aide des modalités d'imagerie disponibles. Chaque modalité produit de plus des images qui contiennent des informations différentes. Dans le cas de l'imagerie cérébrale, la modalité CT (Computed Tomography) renseigne par exemple les cliniciens sur l'anatomie de leurs patients alors que l'IRMf (Imagerie par Résonance Magnétique fonctionnelle) leur fournit des informations d'ordre fonctionnelles. Il est ainsi utile de superposer plusieurs images, acquises par le biais d'une unique modalité ou non, afin d'augmenter la quantité d'informations disponibles en une seule observation.

Le recalage d'images permet de répondre à ces besoins. Il consiste à évaluer les transformations géométriques qui ont été introduites entre deux images d'une même scène, à considérer l'une de ces images comme référence et appliquer la transformation à la seconde afin de pouvoir les aligner.

En imagerie médicale, on distingue le *recalage monomodal* du *recalage multimodal*. Le premier implique de ne considérer que des images obtenues en utilisant la même modalité d'imagerie. Le second a pour objectif de trouver les transformations entre des images acquises grâce à différents dispositifs. Il est ainsi possible d'effectuer des recalages $2D/2D$, $3D/3D$ ou encore $2D/3D$ suivant les modalités d'imagerie considérées. Les images de fonds d'œil étant planes, seul un recalage de type $2D/2D$ a été nécessaire dans le cadre de ce projet. Nous nous limiterons donc à l'étude de ce cas dans la revue de littérature.

Deux grandes familles d'algorithmes de recalage sont à différencier. Les recalages *icôniques* reposent sur l'exploitation de l'intensité des pixels des images et les recalages *géométriques* visent à utiliser des structures géométriques caractéristiques (points, lignes, surfaces, etc...) dans le but de déterminer les déformations entre les images. Dans le cas des images de fond d'œil, la luminosité peut varier de manière très significative entre des images prises chez un sujet dépendamment des conditions variables de leur acquisition (changement de caméra entre deux acquisitions, différente luminosité dans la pièce, modification de la taille de la pupille, zones de réflexion intense). De plus, l'information contenue dans les images rétinienne est essentiellement géométrique, au niveau des structures que sont le disque optique, le réseau vasculaire et la fovéa, le fond de l'image étant peu texturé. Les recalages géométriques sont donc plus adaptés à ce type d'images.

Tout algorithme de recalage, qu'il soit icônique ou géométrique, est composé d'au moins trois parties principales :

- évaluation de la transformation

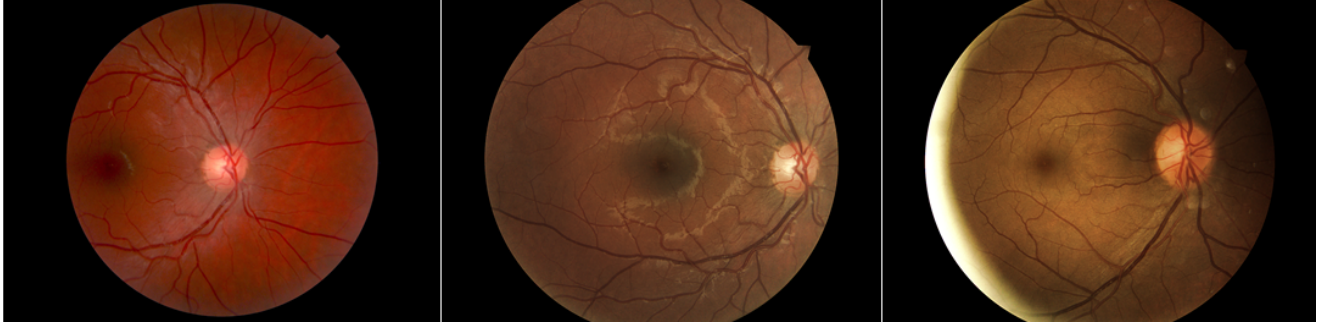


Figure 2.6: Différence de luminosité entre deux images de la rétine d'un même sujet et exemple de réflexion intense

- calcul d'une mesure de similarité
- optimisation

Ces aspects sont présentés plus en détail dans les paragraphes qui suivent.

2.3.1 Choix du type de transformations

Les images numériques sont souvent représentées sous la forme de tableaux ou de matrices dans lesquels chaque case correspond à un pixel. Ceci permet d'avoir recours aux calculs matriciels dans le but de caractériser les déformations géométriques entre les images.

Transformations rigides

On définit comme rigide une transformation composée de *translations* et de *rotations*. Si I_r désigne l'image de référence et I_d l'image à déformer, la correspondance entre la position d'un pixel qui se trouve à la position $X = (x, y)$ dans l'image I_r et son emplacement $X' = (x', y')$ dans l'image I_d s'établit de la manière suivante :

$$X' = RX + T \quad (2.3)$$

avec R une matrice de rotation et T une matrice de translation.

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (2.4)$$

$$T = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (2.5)$$

Il est alors intéressant d'introduire le concept de coordonnées homogènes, dans le but de

réunir rotation et translation dans une unique matrice. Les coordonnées des points deviennent alors $X = (x, y, 1)$ et $X' = (x', y', 1)$ et la matrice de transformations :

$$M = \begin{pmatrix} \cos \theta & -\sin \theta & t_x \\ \sin \theta & \cos \theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \quad (2.6)$$

Transformations affines

Sont appelées affines les transformations qui conservent le parallélisme, c'est-à-dire telles que si deux droites sont parallèles dans une image, elles le sont également après application d'une transformation affine. Des combinaisons de *translations*, *rotations*, *changements d'échelles* et *cisaillements* permettent d'obtenir de telles transformations. La représentation sous forme matricielle donne :

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2.7)$$

Les paramètres t_x et t_y sont relatifs aux translations tandis que les informations concernant les autres transformations sont contenues dans la sous-matrice

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Les rotations, changements d'échelle et cisaillements s'écrivent de manière matricielle respectivement comme suit :

$$R = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.8)$$

$$S = \begin{pmatrix} \Delta x & 0 & 0 \\ 0 & \Delta y & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & \alpha & 0 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Les déformations affines sont alors obtenues en effectuant le produit de ces différents éléments, produit qui n'est toutefois pas commutatif.

Transformations projectives

Les transformations projectives ont les propriétés de conserver la *colinéarité* (une droite reste une droite après transformation projective) et les *birapports de distance* entre des points colinéaires. Dans l'exemple de la figure 2.7, les rapports

$$\frac{\frac{\bar{AC}}{\bar{AD}}}{\frac{\bar{BC}}{\bar{BD}}} = \frac{\frac{A'C'}{A'D'}}{\frac{B'C'}{B'D'}} \quad (2.9)$$

ont été conservés dans le cadre d'une déformation projective.

Transformations élastiques (splines plaques minces)

La méthode de recalage basée sur les splines plaques minces consiste à considérer l'image à recaler comme flexible dans l'espace, à la manière d'une fine plaque de métal. L'objectif est de superposer exactement des points correspondants présents dans les deux images et de déformer l'image à recaler en minimisant l'énergie de flexion de la plaque de métal qui la représente (Bookstein (1989)).

Si n est le nombre de paires de points correspondants entre les deux images, les déformations s'écrivent mathématiquement comme suit :

$$u(x, y) = \sum_{k=1}^3 a_k \varphi_k + \sum_{i=1}^n \omega_i U(r) \quad (2.10)$$

avec

$$r = \sqrt{x^2 + y^2} \quad (2.11)$$

$$(\varphi_1, \varphi_2, \varphi_3) = (1, x, y) \quad (2.12)$$

et

$$U(r) = r^2 \log(r) \quad (2.13)$$

Les a_k correspondent aux coefficients de transformations affines et les ω_i sont des pondérations appliquées aux fonctions radiales centrées en chacun des points.

Les déformations u doivent minimiser l'énergie de flexion définie par :

$$\int \int \left(\left(\frac{\partial^2 u}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 u}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 u}{\partial y^2} \right)^2 \right) dx dy \quad (2.14)$$

Les équations ci-dessus font intervenir la distance de chaque pixel aux points qui ont été mis en correspondance. Les déformations locales ont ainsi moins d'influence sur le calcul des

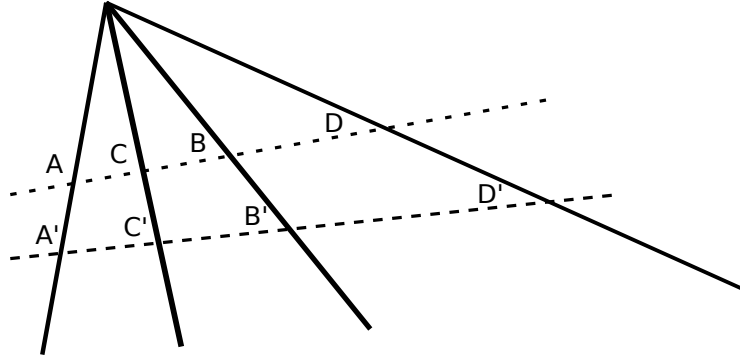


Figure 2.7: Distances de quatre points colinéaires à une origine O. Les transformations projectives conservent les rapports de distances.

transformations entre les images à recaler que dans le cas des méthodes exposées dans les paragraphes précédents qui ne calculent les déformations que de manière globale.

Recalage des images de fond d'œil

Les ophtalmologistes apprécient particulièrement l'application des méthodes de recalage aux images de fond d'œil. En effet, recaler des images de la rétine d'un même sujet prises selon plusieurs angles de vue leur permet d'élargir la zone observée. La possibilité de superposer des images toujours prises chez un même sujet mais à des instants différents favorise le suivi de l'évolution de pathologies ou l'analyse des résultats d'un traitement.

Can *et al.* (2002) proposent d'introduire un type de transformation particulier en ce qui concerne les images de rétines acquises par le biais de caméras de fond d'œil. La première hypothèse formulée par les auteurs est que l'œil peut être considéré comme une sphère. Dans le repère lié à la caméra, un point de coordonnées (X, Y, Z) situé sur la rétine a donc ses coordonnées reliées par une relation de la forme :

$$Z = A_1X^2 + A_2XY + A_3Y^2 + A_4X + A_5Y + A_6 \quad (2.15)$$

La seconde hypothèse stipule que les mouvements introduits entre l'acquisition de deux images peuvent être considérés comme étant des transformations rigides. Les translations sont causées par d'éventuels déplacements de la tête du sujet et les rotations par des mouvements de son œil. Ainsi, prendre deux photographies de la rétine d'un sujet peut être visualisé comme

sur la figure 2.8. C_p et C_q sont deux positions de la caméra à chacune desquelles sont associés un repère et un plan image. Bien que seul l'œil bouge entre l'acquisition de plusieurs images, nous avons en effet choisi d'inverser le problème en considérant que l'œil reste fixe mais que les caméras sont déplacées. Les deux configurations permettent pratiquement d'obtenir les mêmes images, celle impliquant un mouvement des caméras étant plus adaptée à une représentation mathématique. Si $\mathbf{P} = (X, Y, Z)$ et $\mathbf{Q} = (X', Y', Z')$ sont les positions d'un même point dans les repères respectivement associés à C_p et C_q , la relation entre ces deux jeux de coordonnées est de la forme :

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (2.16)$$

avec (r_{ij}) la matrice de rotation et $\begin{pmatrix} t_x & t_y & t_z \end{pmatrix}^t$ la matrice de translation.

Les images en deux dimensions sont obtenues par projection de la rétine des sujets sur un plan image. Soit M_p et M_q les matrices de projection associées respectivement aux caméras C_p et C_q . Les points $\mathbf{p} = (x, y)$ et $\mathbf{q} = (x', y')$ ont pour coordonnées homogènes :

$$\begin{pmatrix} wx \\ wy \\ w \end{pmatrix} = M_p \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} w'x' \\ w'y' \\ w' \end{pmatrix} = M_q \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} \quad (2.17)$$

L'équation 2.17 permet ainsi d'exprimer les coordonnées 3D en fonction des coordonnées 2D des projections. En remplaçant dans l'équation 2.15, on obtient :

$$Z = a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 \quad (2.18)$$

En considérant les matrices de projection de la forme :

$$M_p = \begin{pmatrix} \alpha_x & 0 & 0 & c_x \\ 0 & \alpha_y & 0 & c_y \\ 0 & 0 & 0 & s \end{pmatrix}$$

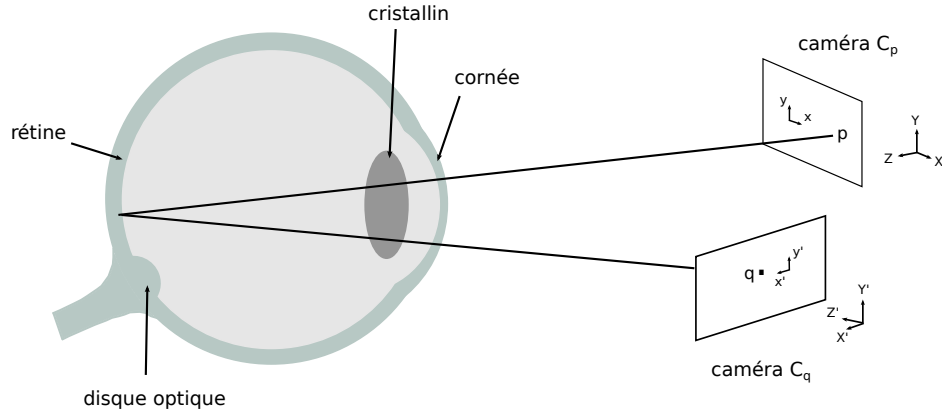


Figure 2.8: Géométrie de l'acquisition de deux images de la rétine d'un même sujet (Can *et al.* (2002))

$$M_q = \begin{pmatrix} \alpha'_x & 0 & 0 & c'_x \\ 0 & \alpha'_y & 0 & c'_y \\ 0 & 0 & 0 & s' \end{pmatrix} \quad (2.19)$$

on obtient :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\alpha_x X + c_x}{s} \\ \frac{\alpha_y Y + c_y}{s} \end{pmatrix}$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{\alpha'_x X' + c'_x}{s'} \\ \frac{\alpha'_y Y' + c'_y}{s'} \end{pmatrix} \quad (2.20)$$

Les coordonnées 3D s'expriment à l'aide des relations 2.15 et 2.20 :

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \frac{sx - c_x}{\alpha_x} \\ \frac{sy - c_y}{\alpha_y} \end{pmatrix}$$

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} r_{11} \frac{sx - c_x}{\alpha_x} + r_{12} \frac{sy - c_y}{\alpha_y} + r_{13}(a_1 x^2 + a_2 xy + a_3 y^2 + a_4 x + a_5 y + a_6) \\ r_{21} \frac{sx - c_x}{\alpha_x} + r_{22} \frac{sy - c_y}{\alpha_y} + r_{23}(a_1 x^2 + a_2 xy + a_3 y^2 + a_4 x + a_5 y + a_6) \end{pmatrix} \quad (2.21)$$

Finalement, la relation entre les coordonnées des points correspondants dans les deux

images est obtenue en utilisant 2.21 dans 2.20 :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \Theta \begin{pmatrix} x^2 \\ xy \\ y^2 \\ x \\ y \\ 1 \end{pmatrix} \quad (2.22)$$

Il s'agit donc de considérer les transformations entre les deux images comme étant d'ordre quadratique. Les déformations introduites par la projection des rétines sphériques sur un plan image sont ainsi prises en compte.

Application des transformations

Une fois les matrices de transformations calculées, il convient de déplacer chaque pixel de l'image à recaler afin de le rapprocher de sa position dans l'image de référence. Cette opération s'effectue en utilisant la méthode dite de *transformation inverse*. Celle-ci consiste à considérer que l'image à recaler remplit une grille de pixels vide lorsqu'elle est déformée. Plutôt que d'appliquer les transformations aux pixels de l'image à recaler, l'approche inverse consiste à y chercher le point dont un pixel de l'image cible se voit attribué la valeur. L'avantage de cette méthode est qu'elle est surjective, tous les pixels de l'image cible recevant une valeur. Il pourrait en effet arriver que, puisque les pixels ont des coordonnées entières, certains pixels de l'image cible ne correspondent à aucune destination de pixels de l'image de départ après transformation.

2.3.2 Détection de structures d'intérêt

Les recalages géométriques sont plus adaptés aux images de fond d'œil que les recalages icôniques. Il est donc nécessaire de trouver des structures d'intérêt afin d'évaluer les transformations entre deux images. Les points sont les éléments les plus utilisés dans les algorithmes de recalage. Les algorithmes de détection de points les plus rencontrés dans la littérature sont les suivants :

Détecteur de Harris Le détecteur de Harris repose sur l'utilisation de la fonction d'autocorrélation calculée sur une fenêtre W de l'image I considérée :

$$c(x, y) = \sum_W (I(x_i, y_i) - I(x_i + dx, y_i + dy))^2 \quad (2.23)$$

Si l'on effectue un développement de Taylor à l'ordre 1 de l'expression ci-dessus, on obtient :

$$c(x, y) = \begin{pmatrix} dx & dy \end{pmatrix} C(x, y) \begin{pmatrix} dx \\ dy \end{pmatrix} \quad (2.24)$$

avec

$$C(x, y) = \begin{pmatrix} \left(\frac{dI}{dx}\right)^2 & \frac{dI}{dx} \frac{dI}{dy} \\ \frac{dI}{dx} \frac{dI}{dy} & \left(\frac{dI}{dy}\right)^2 \end{pmatrix} \quad (2.25)$$

L'étude des valeurs propres de la matrice C permet de connaître les variations d'intensité dans la zone considérée de l'image. Si λ_1 et λ_2 sont les deux valeurs propres de la matrice C , on distingue trois situations :

- λ_1 et λ_2 petites : pas de variation d'intensité significative
- λ_1 ou λ_2 grande et l'autre valeur propre petite : variation d'intensité significative dans une direction donc présence d'un contour
- λ_1 et λ_2 grandes : variation d'intensité importante suivant deux directions donc point de bifurcation (Harris et Stephens (1988))

Il est ainsi possible d'appliquer un seuillage des valeurs propres et de ne prendre en compte que les points pour lesquels λ_1 et λ_2 sont supérieures au seuil. La principale limite de cet algorithme est de ne pas intégrer d'aspect multiéchelle. En effet, une fenêtre de même taille et des seuils identiques sont appliqués à tous les points de l'image.

Scale Invariant Feature Transform (SIFT) Comme présenté dans la figure 2.6, les images de fond d'œil sont de nature très hétérogène, dépendamment des conditions dans lesquelles elles sont acquises et des caméras utilisées. La taille des images peut notamment être différente d'une machine à l'autre. L'algorithme SIFT décrit dans ce paragraphe et développé par Lowe (1999) a la propriété d'être invariant par changement d'échelle, en plus d'être invariant par translation, rotation, projection affine et variation d'illumination (dans une certaine mesure). L'aspect le plus novateur de cet algorithme est la méthode mise en œuvre afin de le rendre insensible aux changements d'échelles. Lowe, pour ce faire, transfère les images dans deux espaces : *l'espace des échelles gaussiennes* et *l'espace des différences d'échelles gaussiennes*. Dans ces espaces, un point de l'image est défini par un triplet de coordonnées (x, y, σ) où x et y sont les coordonnées spatiales du point et σ sa coordonnée d'échelle (dans la suite de ce paragraphe, par souci d'allègement de la notation, les coordonnées spatiales seront désignées par la lettre $X = (x, y)$) (A. Vedaldi (2006)). L'espace des échelles gaussiennes est obtenu en effectuant l'opération suivante :

$$G(X, \sigma) = (g_\sigma * I)(X) \quad (2.26)$$

avec

$$g_{\sigma}(X) = \frac{1}{\sqrt{2\pi}\sigma} \exp -(x^2 + y^2)/2\sigma^2 \quad (2.27)$$

De cette manière, l'échelle de l'image est modifiée en introduisant un filtrage gaussien, présenté par Lindeberg (1994) comme étant la méthode la plus appropriée pour transférer les images dans l'espace des échelles. L'espace des échelles gaussiennes est divisé de la manière suivante :

- *les octaves*, chaque changement d'octave impliquant une division par deux de la taille de l'image
- *les sous-niveaux*, entre lesquels seule la valeur de l'écart-type σ est modifiée

Si l'on désigne par o l'indice des octaves et s l'indice des sous-niveaux, chaque position (o, s) dans l'espace des échelles gaussiennes est associée à une valeur de σ par la relation :

$$\sigma(o, s) = \sigma_0 2^{(o+s/s)} \quad (2.28)$$

avec S le nombre de sous-niveaux, qui est identique dans chaque octave. Lowe a fixé la valeur $\sigma_0 = 1,6 \cdot 2^{1/s}$.

Dans l'espace des différences d'échelles gaussiennes, l'image est transformée selon l'équation :

$$D(X, \sigma(o, s)) = G(X, \sigma(o, s + 1)) - G(X, \sigma(o, s)) \quad (2.29)$$

À chaque niveau (o, s) , les potentiels points d'intérêt sont ceux qui correspondent aux maxima locaux de $|D(., \sigma(o, s))|$ dans le voisinage constitué de ses 8 voisins directs dans les images des niveaux (o, s) , $(o, s - 1)$ et $(o, s + 1)$ (26 points au total, voir figure 2.10).

Une orientation est ensuite déterminée pour les points qui satisfont le critère précédent. Cette étape est effectuée en construisant l'histogramme des gradients associés à chaque point situé dans un voisinage du point sélectionné. Cet histogramme est pondéré par une fenêtre gaussienne d'écart-type 1.5σ . Si une valeur est supérieure à 80% de la valeur maximale obtenue dans l'histogramme, l'orientation correspondante est retenue et intégrée au vecteur descripteur associé au point d'intérêt. Une normalisation des vecteurs descripteurs est alors effectuée afin de rendre l'algorithme SIFT invariant par rotation.

La mise en correspondance des points sélectionnés dans deux images d'une même scène est réalisée en utilisant un algorithme de type *recherche du plus proche voisin* (Beis et Lowe (1997)).

Speeded Up Robust Features (SURF) L'algorithme SURF repose sur le même concept d'espace des échelles gaussiennes que SIFT, en y ajoutant la manipulation des images sous

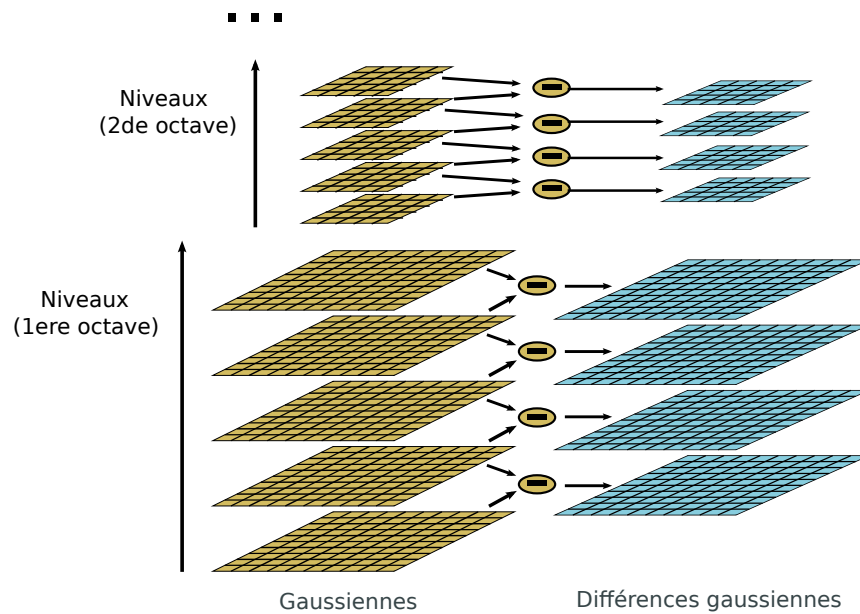


Figure 2.9: Représentation d’une image transférée dans les espaces des échelles gaussiennes et des différences d’échelles gaussiennes. Un changement d’octave correspond à une modification de la taille de l’image d’un facteur 2. Un changement de sous-niveau implique une modification du paramètre σ (C. Evans (2009))

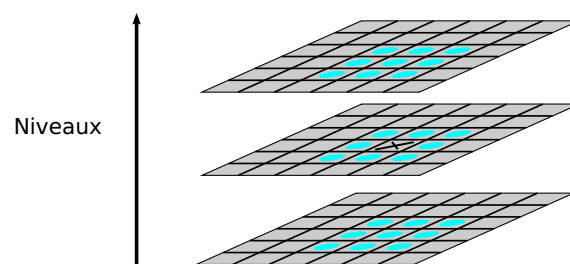


Figure 2.10: Comparaison d’un point d’intérêt avec ses voisins dans les niveaux $(o, s - 1)$, (o, s) et $(o, s + 1)$ (C. Evans (2009))

leur forme *intégrale* (Bay *et al.* (2006)). Si I est une image, l'image intégrale issue de I , appelée I_Σ est obtenue, si (x, y) représentent les coordonnées d'un point de ces images, par l'opération :

$$I_\Sigma(x, y) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y) \quad (2.30)$$

En d'autres termes, la valeur associée au point (x, y) dans l'image I_Σ est égale à la somme des valeurs des pixels situés dans le rectangle compris entre l'origine de l'image et le point en question. Ainsi , il s'agit d'effectuer le calcul suivant pour obtenir l'aire comprise entre les points A, B, C et D de la figure 2.11 :

$$S = I_\Sigma(x_A, y_A) + I_\Sigma(x_D, y_D) - I_\Sigma(x_B, y_B) - I_\Sigma(x_C, y_C) \quad (2.31)$$

Cette propriété permet de diminuer le temps de calcul de manière significative pour un grand nombre d'opérations.

Les algorithmes SIFT et SURF diffèrent en ce qui concerne la méthode de construction de l'espace des échelles gaussiennes. En effet, SIFT procède à une construction des niveaux successifs par multiplication des images des niveaux (o, s) par un filtre gaussien. Chaque image dépend donc de celle du niveau d'ordre inférieur. Seule la taille des filtres gaussiens est modifiée dans l'algorithme SURF et non celle de l'image. Ces filtres sont appliqués à l'image originale, ce qui implique une indépendance entre les images de tous les octaves et sous-niveaux et une réduction significative du temps de calcul.

Dans le cas de l'algorithme SURF, la sélection des points d'intérêt repose sur le calcul du Laplacien de Gaussienne (LoG, Laplacian of Gaussian) en chaque point de l'image. Ce calcul s'effectue en deux étapes. L'image I est tout d'abord convoluée avec une gaussienne dans chacun des niveaux de l'espace des échelles ($g = \text{gaussienne} * I$), l'écart-type σ dépendant

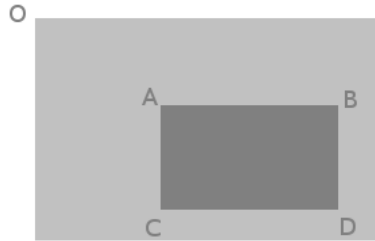


Figure 2.11: Calcul d'une aire sur une image en représentation intégrale (C. Evans (2009))

du niveau. Le Laplacien du résultat de cette convolution est alors calculé.

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (2.32)$$

où $L_{xx}(X, \sigma) = \frac{\partial^2 g(\sigma)}{\partial x^2}$, $L_{xy}(X, \sigma) = \frac{\partial^2 g(\sigma)}{\partial x \partial y}$ et $L_{yy}(X, \sigma) = \frac{\partial^2 g(\sigma)}{\partial y^2}$.

Afin de réduire le temps de calcul, des filtres de type *box filters* sont utilisés pour effectuer ces opérations. Les LoG sont approximés par les filtres de la figure 2.12 D_{xx} , D_{yy} et D_{xy} . Pour les filtres D_{xx} et D_{yy} , les pixels blancs ont pour valeur -1 , les gris 0 et les noirs 2 . Pour le filtre D_{xy} , les blancs sont associés à la valeur -1 , les gris 0 et les noirs $+1$. Bay *et al.* (2006) ont montré que de telles approximations n'affectent la précision des calculs que de manière négligeable mais qu'elles permettent d'obtenir une efficacité et un gain de temps de calcul très importants.

Les points pour lesquels le déterminant de la matrice H est positif sont candidats au titre de points d'intérêt car ceci signifie que les deux valeurs propres de H sont de même signe, donc qu'ils constituent des extrema locaux. Un processus tel que celui utilisé dans le cadre de l'algorithme SIFT permet de sélectionner les points qui sont des extrema au niveau (o, s) ainsi qu'aux niveaux $(o, s - 1)$ et $(o, s + 1)$ en les comparant à leurs 26 plus proches voisins. La position des points d'intérêt est affinée à l'aide d'une interpolation inspirée de Brown et Lowe (2002).

L'orientation associée aux points retenus est déterminée en calculant la réponse à des ondelettes de Haar de taille 4σ (σ dépendant du niveau associé au point en question) d'un ensemble de points situés dans un voisinage de rayon 6σ . Les ondelettes sont orientées selon des axes \vec{x} et \vec{y} . L'orientation relative à la réponse la plus importante est retenue.

Les vecteurs descripteurs SURF sont construits de la manière suivante : une fenêtre de taille 20σ dont le centre est le point d'intérêt est divisée en 16 carrés de dimensions identiques dans lesquels les réponses à deux ondelettes de Haar de taille 2σ sont calculées en 25 points différents. Avec (x, y) la position d'un point et (dx, dy) ses réponses aux ondelettes, le vecteur

$$V = \left[\sum dx, \sum dy, \sum |dx|, \sum |dy| \right] \quad (2.33)$$

est construit. L'assemblage de ces vecteurs obtenus dans les 16 régions fournit un vecteur descripteur de taille 64 pour chaque point d'intérêt, qui est invariant par rotation, changement d'échelle, variations de luminosité et de contraste après normalisation. La mise en correspondance des points est réalisée par calcul de la distance entre les vecteurs descripteurs, la distance euclidienne ou la distance de Mahalanobis pouvant être utilisées.

D'autres structures géométriques simples peuvent être exploitées dans le cadre de re-

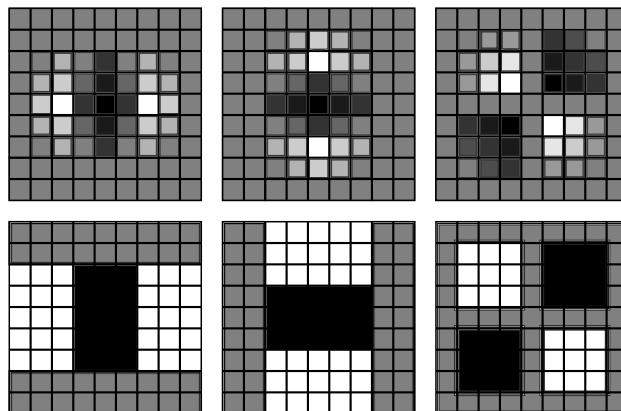


Figure 2.12: Approximation des Laplaciens de Gaussiennes (LoG) par des filtres de type *box filters*. Les LoG de la ligne supérieure L_{xx} , L_{yy} et L_{xy} ont respectivement pour approximations les filtres D_{xx} , D_{yy} et D_{xy} de la ligne inférieure (C. Evans (2009))

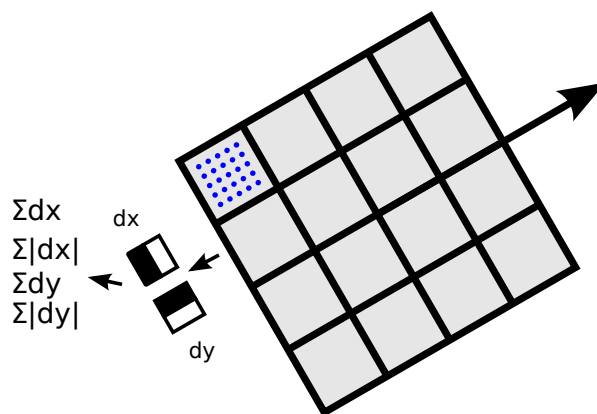


Figure 2.13: Construction des vecteurs descripteurs. Les réponses à des ondelettes de Haar sont calculées en 25 points de chacune des 16 régions qui constituent le voisinage du point d'intérêt (C. Evans (2009))

calages géométriques. Ji *et al.* (1999) citent notamment les segments, lignes, courbes, cercles et ellipses. Les transformées de Hough et l’algorithme LSD (Line Segment Detector, von Gioi *et al.* (2010)) permettent d’obtenir ces primitives.

2.3.3 Mesures de similarité et critères d’optimisation

La deuxième étape des algorithmes de recalage, après l’évaluation des transformations et leur application à l’image à déformer, consiste en une quantification de la similarité entre les images.

Dans le cas d’un recalage géométrique, il s’agit de mesurer la distance entre les primitives une fois la déformation appliquée. La moyenne des distances entre les structures mises en correspondance est l’une des mesures les plus largement utilisées. Si N est le nombre de points correspondants faisant partie des primitives et (x_i, y_i) , (x'_i, y'_i) leurs coordonnées, on a :

$$D = \frac{1}{N} \sum_{i=1}^N \sqrt{(x'_i - x_i)^2 + (y'_i - y_i)^2} \quad (2.34)$$

Le calcul de l’erreur quadratique moyenne (ou mesure RMS, Root Mean Square) entre la position de pixels correspondants est également pertinent dans le cas d’un recalage géométrique :

$$EQM = \sqrt{\sum_{i=1}^N \frac{(x'_i - x_i)^2 + (y'_i - y_i)^2}{N}} \quad (2.35)$$

L’optimisation constitue le lien qui joint la mesure de similarité entre les images et l’évaluation de la transformation. La méthode des *moindres carrés* est la plus répandue dans les algorithmes géométriques. Nous avons vu précédemment (2.3.1) que les transformations peuvent souvent s’écrire sous forme matricielle. Prenons l’exemple d’une transformation affine :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2.36)$$

Cette équation s'écrit de manière équivalente :

$$\begin{pmatrix} x' \\ y' \\ \dots \\ \dots \end{pmatrix} = \begin{pmatrix} x & y & 0 & 0 & 1 & 0 \\ 0 & 0 & x & y & 0 & 1 \\ & & & \dots & & \\ & & & \dots & & \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ t_x \\ t_y \end{pmatrix} \quad (2.37)$$

$$Am = b \quad (2.38)$$

les coordonnées d'un nombre de points supérieur ou égal à trois étant nécessaire à la détermination de la matrice m . L'application de la méthode des moindres carrés revient à résoudre :

$$m = [A^t A]^{-1} A^t b \quad (2.39)$$

Cette méthode permet de minimiser l'erreur quadratique moyenne calculée pour un ensemble de paires de points correspondants.

Dans cette section ont été présentées les étapes principales des algorithmes de recalage, à savoir la détermination des transformations, la mesure de la similarité entre les images et l'utilisation de critères d'optimisation. L'application de tels algorithmes permet d'aligner des images obtenues chez un même sujet ou peut servir à la construction d'atlas en recalant des images dans lesquelles des structures d'intérêt ont préalablement été mises en valeur ou segmentées. Les images de fonds d'œil résultant de la projection d'une rétine de forme sphérique sur un plan, leur contenu dépend de la configuration de leur acquisition. La section qui suit propose d'étudier les méthodes qui pourraient permettre leur reconstruction en trois dimensions, ce qui permettrait de s'affranchir de la dépendance vis-à-vis de la géométrie selon laquelle la surface des rétines est imagée par les caméras de fond d'œil.

2.4 Reconstruction 3D de structures anatomiques

Dans cette section est présentée une revue de la littérature concernant la reconstruction en trois dimensions de la surface de la rétine à partir d'images stéréoscopiques de fond d'œil. Les hypothèses communément formulées quant à la géométrie du problème sont décrites dans un premier temps, suivies d'une explication de la méthode utilisée afin d'en déterminer les paramètres et d'une présentation des techniques les plus fréquemment rencontrées afin de procéder à une reconstruction dense des zones imagées.

2.4.1 Géométrie épipolaire

Nous nous limiterons dans cette partie à l'étude de la géométrie de l'acquisition des images de fond d'œil. Celle-ci est généralement comparée à celle d'un système sténopé (voir figure 2.14). Dans la section 2.3.1, nous avons introduit l'équation de la projection des points 3D sur le plan image :

$$\begin{pmatrix} wx \\ wy \\ w \end{pmatrix} = M_p \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (2.40)$$

La matrice de projection M_p peut être décomposée par une factorisation de type QR (Fusiello (2000)) :

$$M_p = A[R|t] \quad (2.41)$$

avec A contenant les paramètres intrinsèques de la caméra :

$$A = \begin{pmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.42)$$

α_u et α_v sont les distances focales horizontales et verticales. R et t sont les matrices de rotation et translation, donc constituées des paramètres extrinsèques des caméras.

Le cas de l'acquisition de deux images dans la configuration présentée en 2.3.1 s'apparente à celui de la *géométrie épipolaire*, illustrée dans la figure 2.18. O et O' y représentent les centres des caméras (rappelons l'hypothèse selon laquelle seule la caméra est considérée comme mouvante lors de l'acquisition des images), p et p' les images d'un point de l'espace

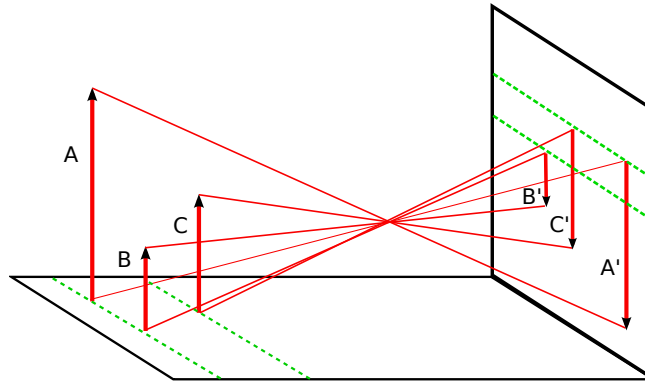


Figure 2.14: Modèle de caméra de type sténopé (F. Cheriet (2011a))

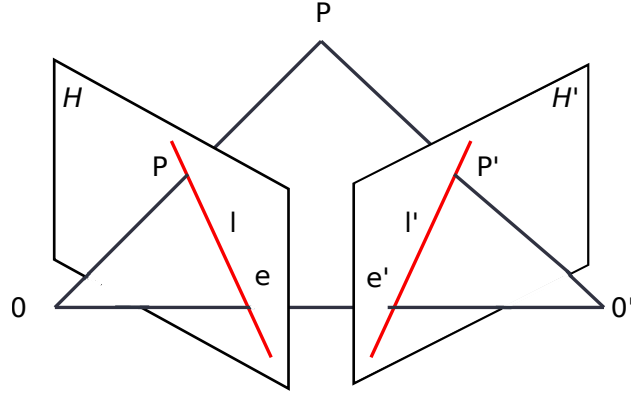


Figure 2.15: Géométrie épipolaire. O et O' sont les centres des caméras, p et p' les images du point P sur les plans H et H' , e et e' les épipoles et l et l' les lignes épipolaires (F. Cheriet (2011b))

P sur les plans images H et H' , e et e' les épipoles, c'est-à-dire les images des centres des caméras sur le plan image de l'autre caméra et l et l' les lignes épipolaires, l reliant p et e et l' passant par p' et e' . Les épipoles sont donc les seuls points de chaque plan image par lesquels passent toutes les lignes épipolaires.

Dans cette configuration, le repère associé à l'une des caméras peut être choisi comme repère d'origine. Dans ce cas, sa matrice de projection est :

$$M_p = A[I|0] \quad (2.43)$$

et la matrice de projection relative à la seconde caméra s'écrit :

$$M'_p = A'[R|t] \quad (2.44)$$

Avec \vec{x} et \vec{x}' les coordonnées des points 2D images du point 3D \vec{X} , on a :

$$w\vec{x} = [A|0]\vec{X} \quad (2.45)$$

$$w'\vec{x}' = A'[R|t]\vec{X} \quad (2.46)$$

d'où

$$w'\vec{x}' = wA'RA^{-1}\vec{x} + A't = wH_\infty\vec{x} + \vec{e}' \quad (2.47)$$

Cette équation signifie que le point correspondant de \vec{x} dans la deuxième image est situé sur la droite qui passe par les points \vec{e}' et $H_\infty\vec{x}$, avec H_∞ l'homographie du plan épipolaire (qui contient P , O et O') à l'infini. Il y a donc colinéarité entre \vec{x}' , \vec{e}' et $H_\infty\vec{x}$, ce qui s'exprime de

la manière suivante :

$$\vec{x}^t(\vec{e}' \wedge H_\infty \vec{x}) = 0 \quad (2.48)$$

ou encore :

$$\vec{x}^t F \vec{x} = 0 \quad (2.49)$$

où F est la matrice fondamentale. Le point \vec{x}' appartient à la ligne $F\vec{x}$ (appelée *ligne épipolaire*) et \vec{x} à la ligne $F^t\vec{x}'$. Les épipoles \vec{e} et \vec{e}' vérifient $F\vec{e} = F^t\vec{e}' = 0$. Le recours à la géométrie épipolaire permet de limiter à une unique ligne la recherche des points correspondants entre des images stéréoscopiques. Il est de plus possible d'appliquer une transformation dite *rectification épipolaire* aux images afin que toutes les lignes épipolaires soient orientées horizontalement et que les lignes correspondantes possèdent la même ordonnée.

2.4.2 Détermination de la configuration du système d'acquisition

Deux algorithmes sont fréquemment rencontrés dans la littérature en ce qui concerne la détermination de la matrice fondamentale :

- L'algorithme de *Hartley* (Hartley (1997)) : il s'agit de calculer la matrice F en résolvant l'équation $\vec{x}'^t F \vec{x} = 0$. F possède 8 inconnues. 8 paires de points correspondants sont donc nécessaires au minimum pour les déterminer, un nombre supérieur de paires permettant d'avoir recours à un système surdéterminé. L'équation évoquée précédemment est équivalente à :

$$\begin{pmatrix} x' & y' & 1 \end{pmatrix} = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

ou encore :

$$\begin{pmatrix} xx' & yx' & x' & xy' & yy' & y' & x & y & 1 \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} \quad (2.50)$$

qui est donc une équation du type $mf = 0$ avec f une matrice colonne qui contient les éléments de la matrice F . Cette équation est résolue en construisant la matrice A dont les lignes sont les matrices m correspondant à au moins 8 paires de points, puis

en procédant à une décomposition en valeurs singulières (communément appelée SVD, Singular Value Decomposition). Cette méthode consiste à effectuer la décomposition $A = UDV^t$ avec U et V des matrices unitaires et D une matrice diagonale dont tous les éléments sont positifs ou nuls. Afin de forcer la solution f à être de rang 2, la plus petite valeur propre de D est fixée à 0. La solution de l'équation est alors proportionnelle à la dernière colonne de la matrice V .

- L'algorithme *RANSAC* (RANdom SAmple Consensus) : dans cet algorithme, la matrice fondamentale est calculée en ne tenant compte que du nombre minimal (8) de paires de points nécessaires à la résolution de l'équation $x'Fx = 0$. Les paramètres obtenus sont testés sur un échantillon plus large afin de confirmer leur validité. L'opération est renouvelée jusqu'à ce qu'un résultat satisfaisant soit obtenu. Cette méthode a l'avantage d'être robuste vis-à-vis des mauvais appariements de points.

Une fois la matrice fondamentale calculée, il est possible de déterminer la position des deux épipoles dans les images ($F\vec{e} = F^t\vec{e}' = 0$). La rectification épipolaire est ensuite obtenue en appliquant la transformation H' , qui correspond à une projection à l'infini de l'épipole e' à l'image qui contient e' et H , choisie telle que la somme des distances $\sum d(Hx, H'x')$ soit minimisée dans l'image contenant e .

2.4.3 Reconstruction 3D

Une fois la rectification épipolaire effectuée, la recherche des pixels correspondants dans les deux images est facilitée car ils se trouvent sur une même ligne. Différents algorithmes ont été développés dans le but d'effectuer des reconstructions denses. À citer notamment ceux qui exploitent la théorie des graphes et les coupes de graphes (Boykov *et al.* (2001)), ainsi que ceux basés sur un recours aux mesures de similarité (Xu et Chutatape (2006)).

Une fois la mise en correspondance réalisée, il est possible de calculer la disparité relative à chaque paire de points définie par :

$$d = x_r - x_d \quad (2.51)$$

avec x_r et x_d les positions des pixels dans les deux images. La profondeur est reliée à la disparité par la relation :

$$z = \frac{bf}{d} \quad (2.52)$$

où b est la distance entre les centres des deux caméras et f leur distance focale (b et f sont également exprimés en pixels par souci d'homogénéité).

Coupes de graphes

Les graphes sont un outil employé en imagerie lorsque l'on cherche à conserver une cohérence au sein de certaines régions de pixels lors du traitement ou de l'analyse des images. Des fonctions d'énergie sont calculées sur les graphes dans le but d'assigner des étiquettes aux différentes régions par l'intermédiaire d'un procédé appelé coupe de graphes. Ceci revient dans notre cas à considérer la mise en correspondance des pixels des images stéréoscopiques comme un problème d'étiquetage de régions de pixels en fonction des disparités associées.

La construction d'un graphe à partir d'une image en deux dimensions se fait comme suit. Soit $G = \{\nu, \epsilon\}$ un graphe dont ν est l'ensemble des nœuds et ϵ l'ensemble des arêtes. Sur l'image de référence, soit N un ensemble de pixels p adjacents situés sur une même ligne. Les étapes de la construction du graphe G sont :

- $\forall p \in I_r$, ajouter p à ν .
- Ajouter les liaisons terminales f^0 et f^a au graphe G .
- $\forall p, q \in N$, ajouter $\{p, q\}$ à ϵ .
- $\forall p \in \nu$, ajouter $\{f^0, p\}$ et $\{f^a, p\}$ à ϵ .
- $\forall \{p, q\} \in \epsilon$, assigner le poids w à l'arête $\{p, q\}$.

Les poids correspondent aux énergies définies par :

$$E(f) = E_{donnee}(f) + E_{lissage}(f) + E_{occlusion}(f) \quad (2.53)$$

avec f la fonction d'étiquetage choisie, qui assigne à chaque pixel p une étiquette f_p , E_{donnee} qui évalue l'appartenance du pixel à un objet environnant, $E_{lissage}$ une énergie de lissage et $E_{occlusion}$ une pénalité attribuée aux pixels victimes d'occlusion. Les énergies $E(f)$ sont données comme valeurs aux arêtes du graphe. De cette manière, des pixels adjacents qui appartiennent à un même objet et sont associés à une disparité similaire sont liés dans le graphe par des arêtes de poids élevé. La coupe de graphe consiste à détecter les liaisons de faible poids de manière à séparer les objets et leur attribuer des étiquettes qui correspondent aux valeurs des disparités associées (Boykov *et al.* (2001)). Les arêtes à prendre en compte dans les coupes sont celles qui lient les pixels adjacents ainsi que celles qui lient chaque pixel aux liaisons terminales (voir figures 2.16 et 2.17).

Cette technique est très séduisante car elle permet de conserver une continuité au sein des objets. Elle nécessite toutefois un temps de calcul très élevé et limite l'observation des détails et faibles variations dans les disparités.

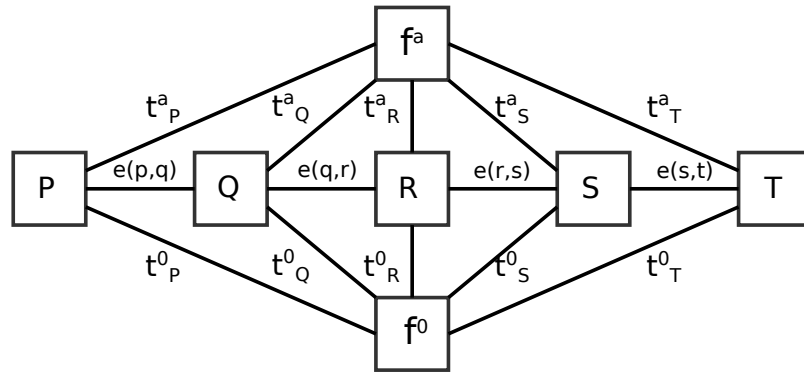


Figure 2.16: Construction d'un graphe à partir d'une image (Droppelmann *et al.* (2010))

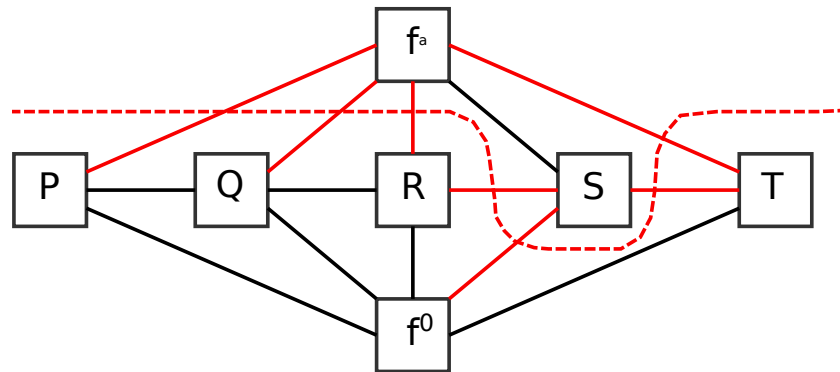


Figure 2.17: Coupe de graphe (Droppelmann *et al.* (2010))

Calcul de mesures de similarité

Une seconde méthode d'appariement des pixels correspondants repose sur le calcul de mesures de similarité entre les intensités de fenêtres de taille réduite, variant généralement entre $(7 * 7)$ et $(11 * 11)$, le long des lignes épipolaires.

Les mesures de similarité de ce type les plus courantes sont (Thivierge-Gaulin (2012)) :

- les mesures basées sur les *différences d'intensité pixel à pixel* : somme des différences absolues et somme des différences au carré, qui correspondent respectivement aux équations suivantes :

$$S_{DA}(I_r, I_d) = \frac{1}{N} \sum_{x,y} |I_r(x, y) - I_d(x, y)| \quad (2.54)$$

$$S_{DC}(I_r, I_d) = \frac{1}{N} \sum_{x,y} (I_r(x, y) - I_d(x, y))^2 \quad (2.55)$$

Ces deux mesures ne prennent pas en compte les variations d'intensité éventuelles entre les images à mettre en correspondance, qui sont fréquentes même dans le cas d'images acquises avec une même modalité d'imagerie.

- la *corrélation croisée normalisée*. Cette mesure repose sur l'évaluation de l'écart-type de l'intensité des images par rapport à leur moyenne. Il s'agit d'une mesure de similarité appropriée dans le cas d'images dont les intensités sont corrélées de manière linéaire.

$$CCN(I_r, I_d) = \frac{1}{N} \frac{\sum_{x,y} (I_r(x, y) - \bar{I}_r)(I_d(x, y) - \bar{I}_d)}{\sqrt{\sum_{x,y} (I_r(x, y) - \bar{I}_r)^2} \sqrt{\sum_{x,y} (I_d(x, y) - \bar{I}_d)^2}} \quad (2.56)$$

- l'*entropie conjointe* L'entropie est une grandeur qui permet de mesurer la quantité d'information contenue dans un échantillon. Si X est une variable aléatoire et x_i un événement ayant la probabilité $p(x_i)$ d'apparaître, l'entropie $H(X)$ s'obtient comme suit :

$$H(X) = - \sum_i p(x_i) \log(x_i) \quad (2.57)$$

Si l'on considère deux variables aléatoires X et Y , il est possible de définir les probabilités conjointes $p(x_i, y_i)$, qui correspondent à la probabilité d'obtenir un événement y_i sachant x_i . L'entropie conjointe correspondante est :

$$H(X, Y) = - \sum_{i,j} p(x_i, y_i) \log p(x_i, y_i) \quad (2.58)$$

Dans le cas d'une comparaison de deux fenêtres de pixels, les probabilités conjointes sont

calculées par l'intermédiaire de *l'histogramme conjoint normalisé*. Ce dernier consiste en un tableau dont les lignes et colonnes correspondent aux intensités dans les deux images, que nous appellerons I_r et I_d . Pour chaque point de coordonnées (x, y) , la case du tableau située au croisement des ligne et colonne qui correspondent aux intensités $I_r(x, y)$ et $I_d(x, y)$ voit sa valeur incrémentée de 1. Après normalisation, le tableau fournit donc les probabilités d'obtention des intensités de l'image I_d connaissant celles de l'image I_r . Avec HCN l'histogramme conjoint normalisé, l'entropie conjointe s'écrit :

$$H(I_r, I_d) = - \sum_{i,j} HCN(I_r, I_d) \log HCN(I_r, I_d) \quad (2.59)$$

- La valeur minimale de l'entropie est atteinte lorsque les fenêtres se correspondent, la quantité d'informations qu'apporte une fenêtre en plus de l'autre étant alors minimale.
- *l'information mutuelle normalisée* Si chacune des images contient peut d'informations, l'entropie conjointe a une valeur faible quel que soit leur degré de recouvrement des fenêtres. L'information mutuelle normalisée permet de pallier à ce problème.

$$IMN(I_r, I_d) = \frac{H(I_r) + H(I_d)}{2 \cdot H(I_r, I_d)} \quad (2.60)$$

Cette grandeur est maximale lorsque les fenêtres correspondent.

L'algorithme de mise en correspondance pixel à pixel connu sous le nom de *SGBM* (Semi Global Block Matching ou mise en correspondance de blocs de manière semi-globale) repose sur l'utilisation de l'information mutuelle et est largement employé. Il est constitué de trois étapes principales :

- Calcul d'une fonction de coût
- Assemblage des coûts
- Formation de la carte des disparités

Fonction de coût Si p_r désigne le point de l'image de référence dont le correspondant est recherché sur la deuxième image, l'ensemble des points potentiellement correspondants est constitué des pixels situés sur un certain intervalle de la ligne épipolaire sur la seconde image. Une fenêtre centrée en ces points leur est associée. L'information mutuelle entre ces fenêtres et une fenêtre de même taille centrée au niveau du pixel de l'image de référence est calculée pour chacun des candidats (voir figure 2.18). Le point pour lequel la valeur de l'information mutuelle est la plus proche de 1 est alors désigné comme correspondant du pixel de l'image de référence.

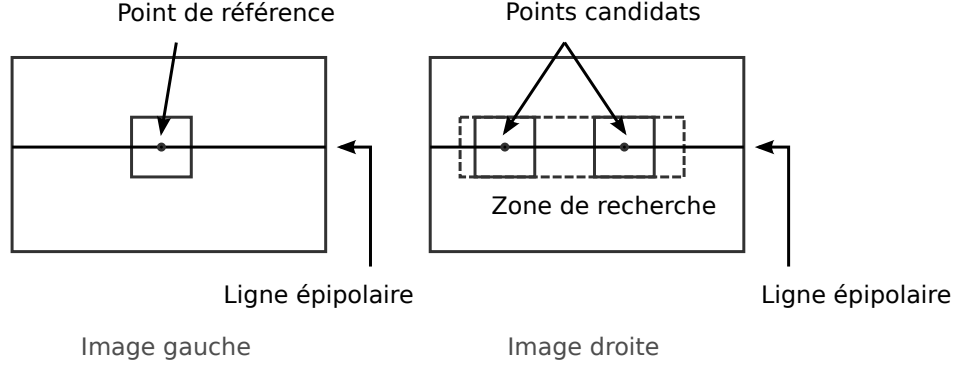


Figure 2.18: Recherche des pixels correspondants le long des lignes épipolaires. Les mesures de similarité sont calculées pour une fenêtre centrée sur les points candidats dans une zone de recherche restreinte (Xu et Chutatape (2006))

Assemblage des coûts Le calcul de l'information mutuelle constitue une première évaluation de l'éventualité que deux pixels soient correspondants. N'utiliser que cette mesure impliquerait cependant la présence de nombreuses erreurs d'appariement. Il est donc proposé dans l'algorithme SGBM de calculer, sur des blocs centrés au niveau de la position du pixel de l'image de référence, une énergie qui prenne en compte le calcul de l'information mutuelle effectué précédemment ainsi qu'un paramètre de lissage. Soit D_p la valeur de la disparité obtenue pour le point p de l'image de référence par calcul de l'information mutuelle et D_q celle associée à un point q compris dans le bloc. Une pénalité P_1 ou P_2 , avec $P_1 < P_2$, est affectée au point q en fonction de la valeur de la quantité $|D_p - D_q|$. Si cette valeur est faible, c'est-à-dire que les deux points sont associés à des valeurs proches de disparité, la pénalité P_1 est affectée au point q et P_2 dans le cas contraire. L'énergie obtenue est alors de la forme :

$$E(p) = IM(p, D_p) + \sum_{q \in \text{bloc}} P(p, q) \quad (2.61)$$

avec $P = P_1$ ou $P = P_2$ suivant l'écart des disparités associées aux points p et q et $IM(p, D_p)$ la valeur de l'information mutuelle obtenue pour les fenêtres centrées en p et en son correspondant sur la seconde image. L'objectif est alors de minimiser la fonction d'énergie totale de l'image des disparités en effectuant un lissage des valeurs des disparités au niveau des points qui augmentent le plus cette énergie totale.

Formation de la carte des disparités La carte finale des disparités est obtenue en approximant les valeurs de disparités obtenues par une fonction quadratique par morceaux, ce qui permet un lissage supplémentaire des écarts qui pourraient persister.

2.5 Objectifs

Les sections précédentes de ce chapitre ont permis dans un premier temps de présenter l'anatomie de l'œil humain et celle de la rétine qui nous intéresse particulièrement dans le cadre de ce projet. Une revue des systèmes d'imagerie du fond d'œil a ensuite été proposée, de manière à décrire les modalités qui permettent l'acquisition d'images en deux et trois dimensions, les caméras de fond d'œil étant le système d'imagerie grâce auquel ont été acquises les images utilisées pour réaliser le projet.

Les objectifs généraux du projet étant de proposer une méthode de recalage d'images de manière totalement automatique afin de construire un atlas des principales structures rétinienne et de reconstruire la surface du fond d'œil en trois dimensions, deux sections de la revue de littérature ont été consacrées à ces deux aspects, la première permettant une étude des algorithmes de recalage et la seconde s'intéressant aux méthodes de reconstruction en trois dimensions à partir d'images stéréoscopiques. Considérer des transformations quadratiques afin de réaliser le recalage des images de fond d'œil semble pertinent étant donnée la structure sphérique de la rétine. Le fait que des images stéréoscopiques permettent le recours aux propriétés de la géométrie épipolaire nous amène à la considérer comme la base de notre algorithme de reconstruction en trois dimensions dans la suite du rapport.

Les objectifs spécifiques sur lesquels est basée la méthodologie présentée dans le chapitre suivant sont dans un premier temps de proposer une méthode de recalage intra-sujet afin d'aligner des images d'une unique rétine, ce qui est notamment une étape indispensable lors de la reconstruction en trois dimensions. Le second objectif consiste à proposer une méthode de recalage inter-sujets afin de constituer un atlas d'images de fonds d'œil. Le troisième objectif a pour but de proposer une méthode de reconstruction de la surface rétinienne en trois dimensions.

CHAPITRE 3

Méthodologie

Ce chapitre est consacré à l'explication de la méthodologie employée dans le but de construire un atlas d'images de fonds d'œil et d'obtenir des informations en ce qui concerne leur profil en trois dimensions. Est présentée dans un premier temps la méthode adoptée afin de recalibrer les images rétinienne. Suit une description de la démarche originale suivie afin de construire l'atlas en deux dimensions. Les étapes de la reconstruction en trois dimensions sont présentées dans une troisième partie avant que ne soient expliqués les critères choisis afin d'évaluer les résultats obtenus.

3.1 Recalage d'images de fond d'œil

L'objectif du recalage des images de fond d'œil est de compenser les déformations qui existent entre les images et qui sont dues aux transformations rigides ou au caractère sphérique de la surface de la rétine. L'algorithme implémenté est de type progressif. Il s'agit en effet de déterminer des transformations d'ordre croissant qui améliorent successivement le modèle. Les premières déformations considérées, d'ordre 0, sont les translations suivant la longueur et la largeur des images :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.1)$$

Suit la recherche des transformations affines, d'ordre 1 :

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (3.2)$$

et des déformations quadratiques, d'ordre 2 (Can *et al.* (2002)) :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} & \Theta_{14} & \Theta_{15} & \Theta_{16} \\ \Theta_{21} & \Theta_{22} & \Theta_{23} & \Theta_{24} & \Theta_{25} & \Theta_{26} \end{pmatrix} \begin{pmatrix} x^2 \\ xy \\ y^2 \\ x \\ y \\ 1 \end{pmatrix} \quad (3.3)$$

Les images de fond d’œil sont dans un premier temps soumises à un débruitage, divers bruits étant introduits lors de leur acquisition. La sélection de points d’intérêts est ensuite à compléter afin de pouvoir calculer les matrices de transformations ci-dessus et de procéder à la déformation de l’image à recaler par approche inverse.

3.1.1 Débruitage

Les images de fond d’œil sont généralement très bruitées, comme le montre la figure 3.1. Les principales catégories de bruits introduits sont : un bruit de hautes fréquences lié aux milieux non homogènes traversés par les rayons lumineux avant d’atteindre le capteur de la caméra, un bruit de basses fréquences dû aux dérives d’éclairement, et le bruit introduit par la compression JPEG. Ces bruits impliquent que des pixels qui devraient avoir des intensités équivalentes dans deux images car représentant un même point de la rétine se voient affecter des valeurs différentes, ce qui pose problème lors de l’application d’algorithmes de mise en correspondance de pixels comme dans la section 4.3. Les outils utilisés afin de réduire l’influence des bruits sont d’une part un algorithme de correction de la dérive d’éclairement développé par *Diagnos Inc.* et d’autre part un filtre passe-bas de type fenêtre de Hann. Ce filtre permet d’homogénéiser les intensités de pixels voisins dont les intensités devraient être semblables tout en conservant des détails tels que la frontière des vaisseaux sanguins, qui seraient détériorés dans le cas d’une utilisation de filtres moyennneurs ou médians, ces filtres devant être de grande taille au regard des perturbations engendrées par le bruit de hautes fréquences.

3.1.2 Sélection de points d’intérêt

Les déformations entre les images à recaler sont évaluées sur un groupe de paires de points correspondants. Dans le cas d’images de fond d’œil d’une personne, l’algorithme de détection de points SURF, sous la forme de sa version libre OpenSURF (C. Evans (2009)) se révèle performant en ce qui concerne l’efficacité et le temps de traitement. Cet algorithme, invariant par translation, rotation, changement d’échelle, projections affines et variations d’illumination fournit en effet un nombre important de points correspondants (en moyenne 215) par couple d’images aussi bien en couleurs qu’en niveaux de gris (si les images sont en couleurs, l’algorithme s’exerce uniquement sur le canal vert de la représentation RGB (Red, Green, Blue), qui contient le plus d’informations). Le fait qu’il repose sur l’utilisation d’images intégrales lui permet d’être très rapide.

Quelques erreurs apparaissent parfois dans la mise en correspondance des points détectés par l’algorithme SURF. Afin de remédier à ce problème, un seuillage est appliqué sur la



Figure 3.1: Les images de fond d'œil sont sujettes à du bruit de hautes fréquences.



Figure 3.2: Application d'un filtrage de type fenêtre de Hann sur une image de fond d'œil

distance entre les points correspondants. Si cette distance est inférieure à la médiane des distances plus un écart type, le couple de points est conservé. Si la valeur seuil est dépassée, la paire n'est pas considérée dans la suite de l'algorithme. Le choix de cet intervalle a pour unique but de ne pas prendre en compte les appariements aberrants lors de l'évaluation des transformations et de limiter le nombre de paires de points considérées par souci de limitation du temps de calcul. D'autres seuils auraient pu être choisis, cette valeur nous semblant un bon compromis entre la proportion de points rejetés et le temps d'exécution de l'algorithme.

La figure 3.3 est un exemple de mise en correspondance de points entre deux images stéréoscopiques d'un même fond d'œil. Les points détectés sont répartis sur toute l'image, ce qui permet d'évaluer les transformations sur une large zone.

3.1.3 Estimation graduelle des transformations

L'algorithme de recalage proposé est de type progressif, afin de suivre la méthode proposée par Can *et al.* (2002). Les paragraphes qui suivent décrivent la manière dont sont évaluées successivement les translations, transformations affines et déformations quadratiques. Chaque transformation est appliquée aux paires de points correspondants avant d'évaluer la suivante à partir de ces mêmes points transformés.

Translations Les translations entre l'image de référence et l'image à déformer sont déterminées dans les directions de la hauteur et de la largeur, soit selon les axes des \vec{x} et des \vec{y} . La médiane des distances entre les points d'intérêt est calculée en nombre de pixels dans chacune des deux directions. L'image à recaler est ensuite déplacée selon les deux axes. Un tel procédé est applicable du fait que la zone des rétines visible sur les images de fond d'œil est considérée comme quasi plane et que des points d'intérêt sont obtenus sur l'ensemble de l'image.

Transformations affines Le second niveau de transformations étudié est celui des déformations affines. Sous forme matricielle, ces déformations s'écrivent comme présenté dans l'équation (3.2) :

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

La matrice de transformation contient 6 inconnues. La recherche de points correspondants avec l'algorithme OpenSURF fournit comme nous l'avons vu en moyenne 215 paires de points, ce qui est plus que suffisant pour résoudre ce système.

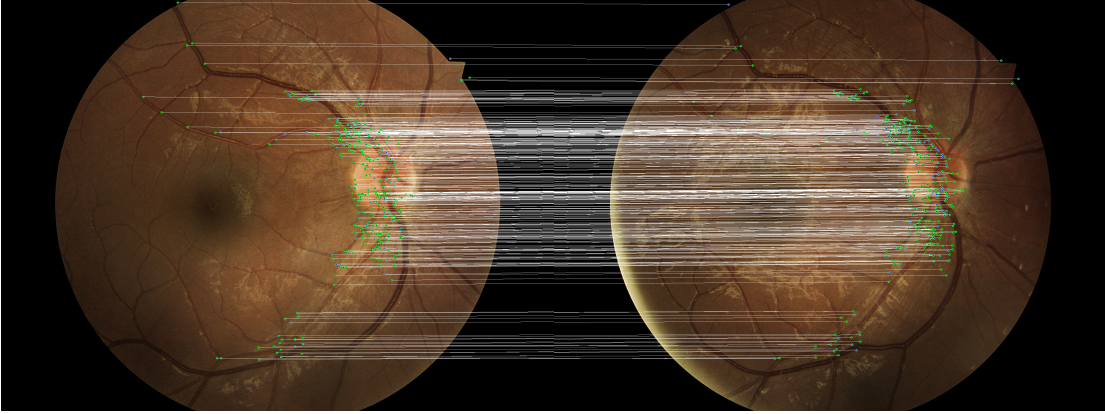


Figure 3.3: Mise en correspondance de points d'intérêt par l'algorithme OpenSURF (C. Evans (2009))

Le critère d'optimisation choisi est celui des moindres carrés. Nous avons vu dans la section 2.3.3 que cette méthode consiste à considérer l'équation (2.36) :

$$\begin{pmatrix} x'_1 \\ y'_1 \\ \dots \\ x'_i \\ y'_i \\ \dots \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ & & & \dots & & \\ x_i & y_i & 0 & 0 & 1 & 0 \\ 0 & 0 & x_i & y_i & 0 & 1 \\ & & & \dots & & \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ t_x \\ t_y \end{pmatrix}$$

$$Am = b$$

et à résoudre (2.39)

$$m = [A^t A]^{-1} A^t b$$

Une fois les paramètres de la matrice de transformation affine calculés, la méthode de recalage inverse est appliquée (chaque pixel de l'image vide de destination est associé à un pixel de l'image à recaler).

Transformations quadratiques Le troisième et dernier niveau de déformations pris en compte dans l'algorithme de recalage est celui des transformations quadratiques. La méthode adoptée dans le but de déterminer les paramètres de la transformation quadratique est identique à celle employée dans le cas de la transformation affine, à ceci près que les points d'intérêt sont désignés par leurs coordonnées quadratiques : $(x^2, xy, y^2, x, y, 1)$. Le problème

s'écrit donc de manière matricielle 3.3 :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} & \Theta_{14} & \Theta_{15} & \Theta_{16} \\ \Theta_{21} & \Theta_{22} & \Theta_{23} & \Theta_{24} & \Theta_{25} & \Theta_{26} \end{pmatrix} \begin{pmatrix} x^2 \\ xy \\ y^2 \\ x \\ y \\ 1 \end{pmatrix}$$

Cette équation est équivalente à :

$$\begin{pmatrix} x'_1 \\ y'_1 \\ \dots \\ x'_i \\ y'_i \\ \dots \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_i^2 & x_i y_i & y_i^2 & x_i & y_i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_i^2 & x_i y_i & y_i^2 & x_i & y_i & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \Theta_{11} \\ \Theta_{12} \\ \Theta_{13} \\ \Theta_{14} \\ \Theta_{15} \\ \Theta_{16} \\ \Theta_{21} \\ \Theta_{22} \\ \Theta_{23} \\ \Theta_{24} \\ \Theta_{25} \\ \Theta_{26} \end{pmatrix} \quad (3.4)$$

Ce système est résolu en faisant appel à la méthode des moindres carrés. Un recalage par méthode inverse est également effectué une fois que les paramètres de la transformation sont calculés.

L'image 3.4 illustre le recalage quadratique de deux images de fond d'œil. Cet exemple correspond à la superposition d'images d'une seule rétine. Appliqué à des photographies du fond d'œil de différents sujets, le recalage permet la construction d'un atlas, dont il est sujet dans la section suivante.

3.2 Construction d'atlas 2D d'images de fonds d'œil

Peu de travaux ont été réalisés en ce qui concerne la construction d'atlas d'images de fonds d'œil (Lee *et al.* (2010)), les praticiens étant habitués à établir leurs diagnostics à partir des images obtenues directement à l'aide des caméras de fond d'œil. Les apports principaux d'un atlas d'images rétinienne sont la possibilité d'étudier la répartition des principales structures



Figure 3.4: Exemple de recalage quadratique d'images de fond d'œil

anatomiques présentes sur la rétine et l'éventualité d'une facilitation de l'initialisation des algorithmes de segmentation du réseau vasculaire.

La construction d'un atlas implique de recaler des images de sujets différents de manière à les amener dans un référentiel commun. Il est également nécessaire ici de sélectionner des structures communes aux différentes images avant de pouvoir évaluer les transformations.

3.2.1 Sélection des structures

Lorsque l'objectif est de recaler des images issues de différents sujets, la méthode employée en 3.1 ne peut être appliquée comme telle. En effet, il est impossible de trouver des points qui possèdent un vecteur descripteur de type SURF semblable dans des images de fonds d'œil différents, ou si tel est le cas, il y a peu de chance que ces points soient réellement correspondants. D'après l'équation (2.22), 11 points sont nécessaires au minimum pour déterminer les paramètres d'une transformation quadratique. Ne considérer que les deux points qui sont les centres du disque optique et de la fovéa, présents sur toutes les images de fond d'œil, n'est donc pas suffisant et il est impossible de considérer ces deux structures dans leur globalité puisque leur taille et leur forme sont très variables d'un fond d'œil à l'autre. La solution proposée dans le cadre de ce projet est de construire une ellipse dont les deux foyers sont les centres de la fovéa et du disque optique et dont le grand axe est fonction de la distance entre

ces foyers. Une telle structure, fonction de la position du disque optique et de la fovéa ainsi que de la distance qui les sépare, permet en effet d'évaluer les déformations entre images de fonds d'œil de manière globale.

Localisation du centre de la fovéa

La détection de la fovéa est effectuée en utilisant un algorithme développé par *Diagnos Inc.* La première étape consiste à rechercher la zone maculaire, qui est la région qui entoure la fovéa. L'algorithme procède comme suit :

- Segmentation du réseau vasculaire à l'aide de l'algorithme de segmentation actuel de *Diagnos Inc.*
- Positionnement d'un cercle qui suit le réseau vasculaire au sens d'une minimisation des moindres carrés.
- Sélection du disque formé de la zone de l'image qui se trouve à l'intérieur du cercle (figure 3.5).

Le choix d'un disque comme structure à associer au réseau vasculaire s'explique par le fait que les vaisseaux sanguins entourent généralement la zone maculaire. Suite à ces opérations, la zone de recherche de la fovéa est réduite à environ 1/4 de la taille des images utilisées au cours du projet. Une recherche plus précise de la fovéa est ensuite menée. Elle est basée sur le raisonnement suivant : la fovéa se situe dans la grande majorité des cas dans une zone sombre dans laquelle la densité du réseau vasculaire est réduite. Cette étape de l'algorithme se compose des opérations suivantes :

- Sélection des pixels dont 95% des voisins situés dans un voisinage circulaire de taille 50 n'appartiennent pas au réseau vasculaire.
- Application d'un filtre moyenneur de taille 50 aux pixels qui remplissent le critère précédent.
- Placement du centre de la fovéa au niveau du pixel qui possède l'intensité la plus faible (figure 3.6).

Localisation du centre du disque optique

Le disque optique apparaît généralement comme une zone en forme de disque et d'intensité plus élevée que le reste de l'image, même lorsqu'il est affecté par des pathologies telles que le glaucome. Ces constatations étant faites, la méthode employée est la suivante :

- Convolution du canal vert de l'image considérée avec un disque d'intensité 255 et de diamètre $l/10$, avec l la longueur de l'image.
- Un seuillage est appliqué à l'image résultant de la convolution précédente. Les pixels



Figure 3.5: Détection de la zone maculaire



Figure 3.6: Positionnement du centre de la fovéa



Figure 3.7: Détection du centre du disque optique

de cette image dont la valeur après convolution est supérieure à la valeur maximale obtenue moins trois écarts-types sont conservés.

- Un étiquetage de surface permet de ne conserver que la région composée de pixels connexes et possédant l'aire la plus grande.
- Le centre de gravité de cette région est défini comme étant le centre du disque optique (figure 3.7).

Positionnement des ellipses

Une fois les centres du disque optique et de la fovéa placés sur les images à recaler, des ellipses dont ils sont les foyers sont tracées (figure 3.8). Dans le cas des images utilisées, le demi grand axe de l'ellipse associée à l'image de référence est fixé à la valeur de 1000 pixels. Les caméras de fond d'œil qui permettent d'obtenir les images ont un champ de vision de 45 degrés. Puisque les yeux des sujets dont sont issues les deux images considérées dans ce paragraphe peuvent avoir des dimensions légèrement différentes, la distance entre les centres du disque optique et de la fovéa peut varier. Afin de prendre en considération ces variations, si d_1 correspond à la distance entre les centres de la fovéa et du disque optique dans l'image de référence et d_2 la distance entre ces mêmes points dans l'image à recaler, le demi grand axe D de l'ellipse associée à l'image à transformer est calculé comme suit :

$$D = 1000 \cdot \frac{d_2}{d_1} \quad (3.5)$$



Figure 3.8: Construction d'une ellipse dont les foyers sont les centres du disque optique et de la fovéa

3.2.2 Recalage

Seules les images des ellipses sont prises en compte dans le recalage d'images de différents sujets et l'algorithme utilisé dans le cas du recalage d'images d'un seul patient leur est appliqué. Les images des ellipses sont en effet très similaires, ce qui permet à l'algorithme SURF de trouver des points correspondants exploitables (figure 3.9).

La superposition des images après recalage par rapport à une image de référence permet de former un atlas d'images de fonds d'œil en deux dimensions. La section suivante vise à présenter la démarche suivie afin d'introduire la troisième dimension au sein de cet atlas.

3.3 Vers la construction d'un atlas 3D de la surface rétinienne

Afin de s'affranchir de la dépendance vis-à-vis de la configuration de l'acquisition des images de fond d'œil, il est souhaitable de pouvoir obtenir une reconstruction en trois dimensions de la surface des rétines. Cette section décrit la méthodologie adoptée afin d'obtenir des informations sur le profil 3D de ces surfaces à partir d'images stéréoscopiques.

3.3.1 Images stéréoscopiques

Les fonds d'œil sont reconstruits en trois dimensions à partir d'images stéréoscopiques. Celles-ci sont acquises grâce au mode d'acquisition automatique incorporé dans les caméras de fond d'œil. Une image centrée entre la fovéa et le disque optique est prise dans un premier

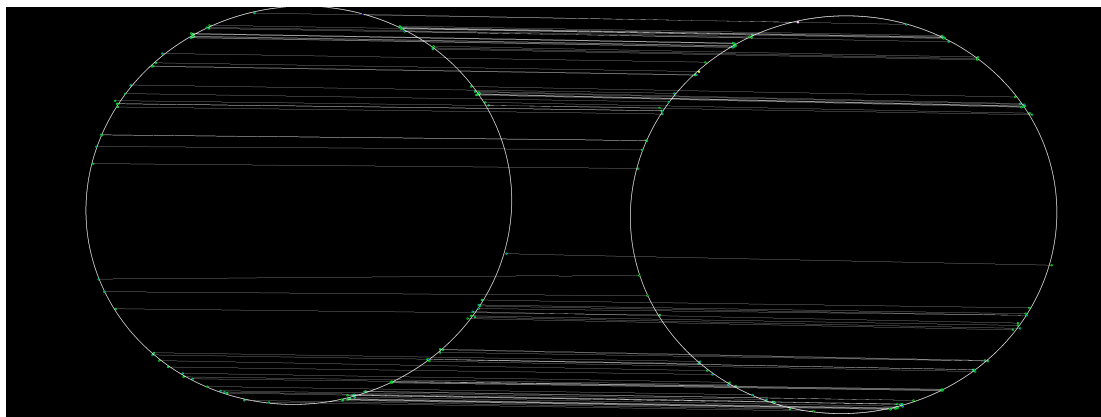


Figure 3.9: Recherche de points correspondants sur les ellipses construites

temps, suivie d'une image légèrement décalée. Le sujet est incité à fixer un point lumineux indiqué par la caméra afin que son œil soit bien orienté. Un exemple d'images stéréoscopiques est présenté en figure 3.10.

3.3.2 Reconstruction dense

Les hypothèses sur lesquelles repose le raisonnement suivi lors de la reconstruction en trois dimensions sont que les images sont acquises au travers d'un système d'imagerie de type sténopée (2.4.1) et que la géométrie du problème peut être supposée comme épipolaire.

Une étape préalable de recalage affine des images stéréoscopiques est intégrée à l'algorithme. Celle-ci permet de supprimer l'influence de déformations telles que des translations, rotations et changement d'échelle qui peuvent être introduits entre les deux images. Seules les déformations liées à la profondeur des fonds d'œil sont ainsi conservées.

Rectification épipolaire

Comme expliqué en 2.4.1, la matrice fondamentale permet de décrire la géométrie épipolaire. L'algorithme OpenSURF permet une fois de plus d'obtenir des paires de points correspondants qui servent au calcul de la matrice fondamentale. La méthode RANSAC (2.4.2), dont une implémentation est disponible dans la librairie OpenCV (http://opencv.willowgarage.com/documentation/cpp/camera_calibration_and_3d_reconstruction.html)

Connaissant la matrice fondamentale, il est possible de procéder à la rectification épipolaire des deux images stéréoscopiques (2.4.2). La librairie OpenCV fournit des fonctions qui permettent de réaliser cette opération (http://opencv.willowgarage.com/documentation/cpp/camera_calibration_and_3d_reconstruction.html)(figure 3.11).

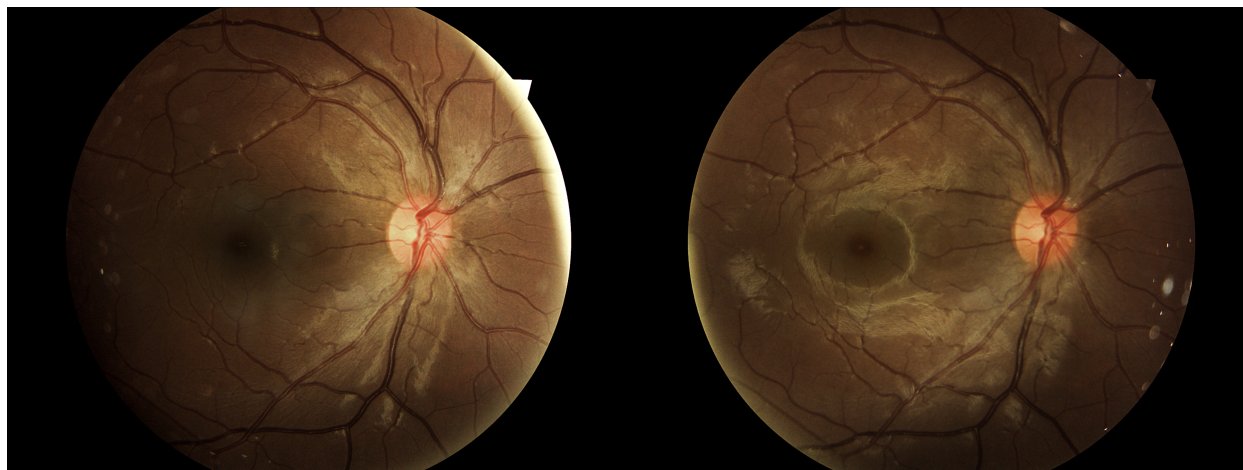


Figure 3.10: Images stéréoscopiques d'un fond d'œil

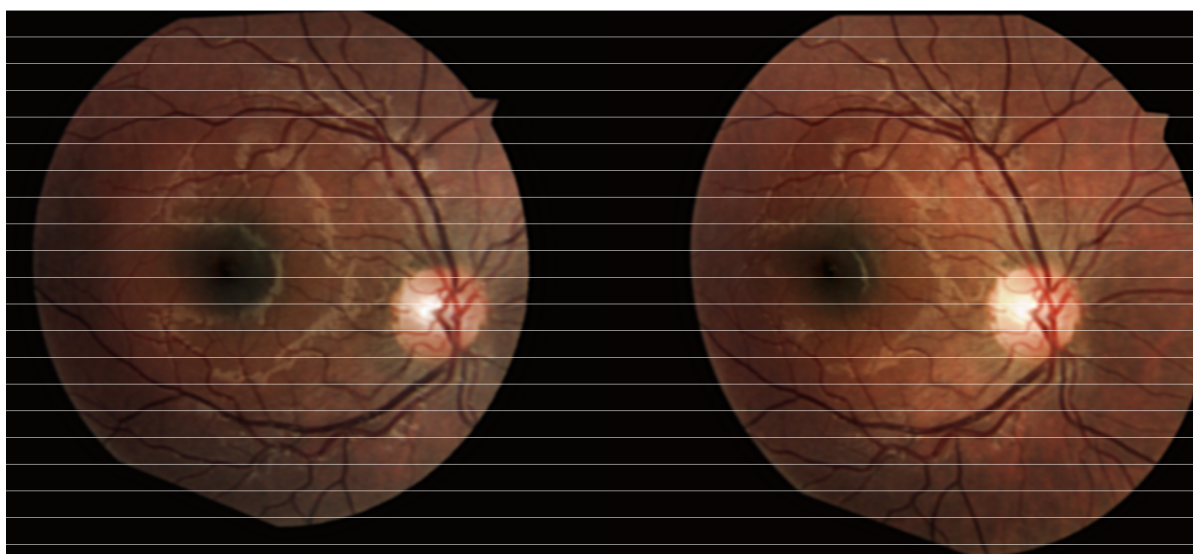


Figure 3.11: Images stéréoscopiques d'un fond d'œil après rectification épipolaire

Semi Global Block Matching

La rectification épipolaire permet de faciliter l'étape de la recherche des pixels correspondants entre les deux images stéréoscopiques, qui s'effectue par application de l'algorithme SGBM (2.4.3). La librairie OpenCV propose une implémentation de cet algorithme.

Une mise en correspondance dense des pixels est alors obtenue, ce qui permet d'établir la carte des disparités relative au fond d'œil étudié.

Un lissage supplémentaire des cartes des disparités est effectué au moyen d'un filtre médian de taille 3×3 et d'un filtre moyenneur de taille 5×5 de manière à prévenir l'influence d'erreurs apparues lors de leur formation.

3.4 Méthode d'évaluation de l'approche proposée

Les trois sections précédentes ont permis de décrire la méthodologie adoptée afin de recaler les images de fond d'œil, de construire des atlas en deux dimensions et d'obtenir le profil 3D de la surface de rétines. Cette nouvelle section a pour but de présenter les méthodes adoptées afin d'évaluer les algorithmes proposés pour vérifier l'atteinte de ces trois objectifs.

Recalage Puisque l'algorithme SURF fournit un nombre élevé de paires de points correspondants entre les images à recaler, la méthode proposée afin d'évaluer quantitativement l'algorithme de recalage consiste à diviser l'échantillon de paires de points en deux groupes. Le premier sert au calcul des matrices de transformations. Les déformations sont appliquées aux points du deuxième groupe qui appartiennent à l'image à recaler. Leur position d'arrivée idéale étant celle de leur correspondant dans l'image de référence, la différence entre cette position visée et la position réelle d'arrivée est évaluée en calculant la distance entre ces deux positions pour l'ensemble des paires de points du deuxième groupe. L'algorithme est ensuite évalué qualitativement, en testant sa capacité à produire des mosaïques constituées de multiples images et à permettre le suivi de pathologies évolutives.

Construction d'atlas 2D Les structures communes qu'il est possible de localiser dans toutes les images de fond d'œil sont le disque optique et la fovea. L'évaluation quantitative de la méthode de formation des atlas en deux dimensions consiste donc à calculer l'écart-type de la distribution des positions des centres de la fovéa et du disque optique par rapport à leur position moyenne après recalage. Une évaluation qualitative est également effectuée en analysant le recouvrement des structures anatomiques rétinienne suite à l'application de l'algorithme à différentes images et en réalisant une superposition de réseaux vasculaires pré-segmentés afin de construire une carte de la probabilité des pixels d'images de fonds d'œil

d'appartenir à la vasculature.

Atlas 3D Les reconstructions en trois dimensions de la surface des rétines différant à une transformation projective près de la réalité, seule une évaluation qualitative de la méthode de reconstruction proposée est effectuée en établissant un lien entre les profils 3D obtenus et les informations dont nous disposons en ce qui concerne l'anatomie de la rétine. Le chapitre qui suit présente les résultats obtenus en effectuant ces évaluations.

CHAPITRE 4

Résultats et discussions

Dans ce chapitre sont présentés les résultats et discussions qui découlent de l'application des méthodes d'évaluation exposées dans le chapitre précédent.

4.1 Recalage intra-sujet d'images de fond d'œil

Le premier algorithme évalué est celui du recalage des images de fonds d'œil intra-sujet. Est proposée une évaluation quantitative, suivie d'une analyse qualitative de l'algorithme.

4.1.1 Évaluation quantitative de l'algorithme de recalage intra-sujet

Les ensembles de points fournis par l'algorithme OpenSURF lors de la recherche de points correspondants ont été divisés en deux groupes, l'un servant au calcul des matrices de transformations et l'autre, contenant trente paires de points, étant conservé afin d'évaluer quantitativement la méthode de recalage proposée.

Le tableau 4.1 ressenne les valeurs de la distance moyenne entre les paires de points après recalage. Cette mesure correspond à la formule :

$$Distance = \sum_i \frac{\sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2}}{n} \quad (4.1)$$

avec $n = 30$ le nombre de paires de points utilisées pour réaliser les calculs.

Ce tableau permet de constater une réduction de l'écart entre la position des points correspondants lorsque l'ordre des transformations prises en compte augmente. La distance entre les points correspondants est en effet réduite avec l'ordre des transformations puisqu'elle diminue de 11.45 pixels (0.184mm) dans le cas des translations à 4.30 pixels (0.069mm) pour les transformations affines et 2.51 pixels (0.040mm) après déformations quadratiques.

La distance moyenne entre les paires de points correspondants obtenue suite à un recalage quadratique, égale à $40\mu m$, est égale à la moitié du diamètre des vaisseaux sanguins les plus fins observables sur les images rétinienues, dont le diamètre est d'environ $80\mu m$, les branches principales du réseau vasculaire pouvant avoir un diamètre atteignant $350\mu m$.

Tableau 4.1: Distances moyennes entre les paires de points correspondants après application de l'algorithme de recalage intra-sujet.

Images	Translation		Recalage affine		Recalage quadratique	
	pixels	mm	pixels	mm	pixels	mm
Paire 1	19.41	0.312	5.12	0.082	2.59	0.042
Paire 2	15.78	0.253	4.52	0.073	3.30	0.053
Paire 3	18.22	0.293	4.01	0.064	3.58	0.057
Paire 4	6.69	0.107	5.31	0.085	1.32	0.021
Paire 5	4.19	0.067	3.30	0.053	2.04	0.033
Paire 6	4.42	0.071	3.54	0.054	2.23	0.036
Moyenne	11.45	0.184	4.30	0.069	2.51	0.040



Figure 4.1: Images de fond d'œil acquises suivant différents angles de vue.



Figure 4.2: Résultat obtenu après recalage quadratique des images de la figure 4.1

Tableau 4.2: Distances moyennes entre les paires de points correspondants après application du recalage quadratique à des images de dimensions différentes

Image 1		Image 2		Distance	
Largeur	Longueur	Largeur	Longueur	pixels	mm
1958	2196	1944	2592	5.77	0.093
1958	2196	1944	2592	7.34	0.118
2448	3264	1944	2592	21.98	0.353
2448	3264	1944	2592	16.93	0.272



Figure 4.3: Recalage d'images de dimensions différentes.

Les images analysées dans le tableau 4.1 sont toutes de taille $2592 * 1944$. La précision du recalage diminue cependant lorsque les images sont de tailles différentes, comme le montrent la figure 4.3 et le tableau 4.2. Le centre des images, à proximité du disque optique, est en effet bien superposé car on y trouve le plus grand nombre de points d'intérêt. Le fait que l'algorithme se contente de transformations d'ordre 2 implique cependant une diminution de la précision du recalage avec l'éloignement par rapport au centre des images. Lorsque le rapport entre les dimensions des images est supérieur à 1.25, le recalage n'est plus applicable, la distance entre les points correspondants devenant trop hétérogène pour permettre un calcul satisfaisant des matrices de transformations.

Afin de vérifier la validité des opérations présentées ci-dessus, qui reposent sur l'utilisation de paires de points fournies par l'algorithme OpenSURF, nous avons sélectionné manuellement des points d'intérêt sur des couples d'images recalées et évalué la distance entre les positions pointées. Une comparaison des distances entre les points correspondants obtenues par la méthode automatique et par la méthode manuelle est alors effectuée afin de vérifier que les distributions sont bien de la même forme. Un test de Wilcoxon-Mann-Whitney permet d'effectuer cette comparaison. L'hypothèse nulle de ce test non-paramétrique, H_0 , consiste à considérer les deux distributions comme identiques. Les résultats obtenus pour deux paires d'images choisies aléatoirement sont présentés dans le tableau 4.3.

Les valeurs élevées de p-valeur obtenues, considérées comme telles car très supérieures à la valeur critique de 0.05, permettent de considérer les distributions comme étant de même type, ce qui conforte les résultats obtenus à l'aide de notre méthode d'évaluation quantitative.

Tableau 4.3: Évaluation de la similitude entre les distributions des distances entre des points sélectionnés manuellement ou avec OpenSURF

	Affine 1	Quadratique 1	Affine 2	Quadratique 2
p-valeur	0.4636	0.8289	0.9189	0.3476

4.1.2 Évaluation qualitative de l'algorithme de recalage intra-sujet

L'algorithme de recalage ayant été évalué qualitativement, deux évaluations quantitatives sont proposées, à savoir la construction de mosaïques suite au recalage de multiples images d'un fond d'œil et le suivi de pathologies évolutives au cours du temps.

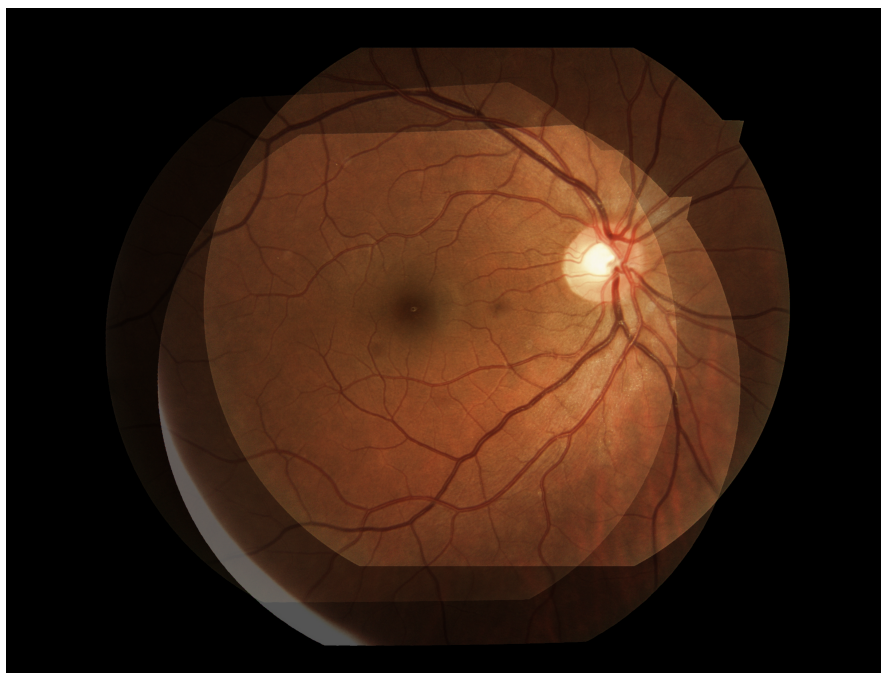


Figure 4.4: Mosaïque construite par recalage de trois images de fond d'œil

Construction de mosaïques Comme expliqué en 2.3.1, les praticiens peuvent être intéressés par le fait de pouvoir étendre la zone de la rétine observée sur une seule image en ayant recours au recalage de plusieurs images de fond d'œil. Le rendu obtenu est semblable à celui d'une mosaïque, terme utilisé par Can *et al.* (2002) pour désigner des images telles que celle de la figure 4.4, qui est issue de la superposition de trois images recalées à l'aide de l'algorithme développé dans le cadre du projet. L'algorithme implémenté permet donc aux praticiens d'accéder à une zone étendue de la rétine sur une image et de leur faire économiser le temps passé à explorer diverses images de fond d'œil afin d'effectuer un diagnostic.

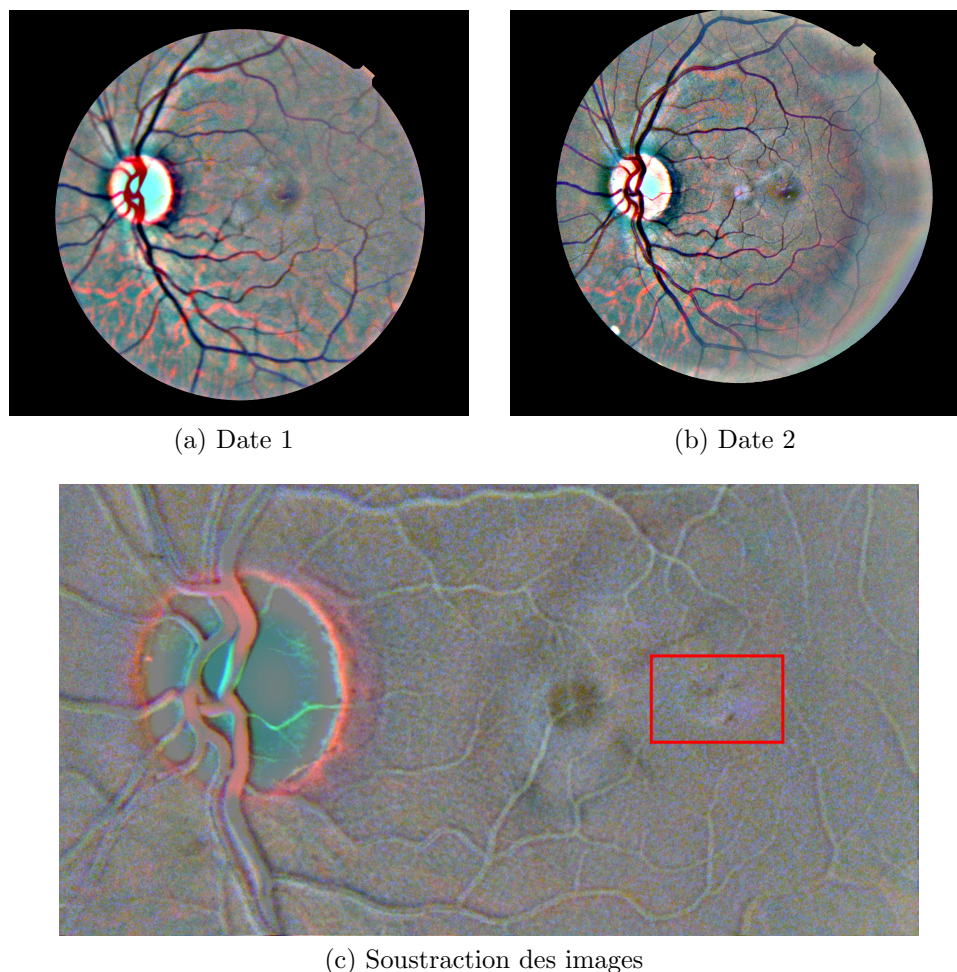


Figure 4.5: Recalage d'images d'un patient prises lors de différents examens. Des lésions sont apparues dans la zone encadrée en rouge entre les deux examens.

Suivi dans le temps Nous avons également souligné en 2.3.1 que les praticiens souhaitent pouvoir suivre leurs patients au cours du temps, notamment lorsque ces derniers sont appelés à subir des examens de fond d'œil réguliers. Le patient de la figure 4.5 a connu l'apparition de lésions sur sa rétine entre les deux examens dont sont issues les images a) et b) dans la zone encadrée en rouge dans l'image c), qui résulte de la soustraction des deux images précédentes. La soustraction des images recalées permet en effet d'homogénéiser l'image et de faire ressortir les différences entre des photographies prises à des instants différents. Les contours des vaisseaux sanguins peuvent encore être distingués ici malgré la soustraction des images, l'une d'entre elles étant légèrement floue.

Le recalage des images de la rétine d'un sujet ayant été évalué, nous allons dans la section suivante nous intéresser à l'application de l'algorithme à la construction d'atlas 2D d'images

de fonds d'œil.

4.2 Construction d'atlas 2D

Comme dans le cas de l'application de l'algorithme de recalage aux images d'un seul sujet, son utilisation dans le cadre de la construction d'atlas 2D est évaluée de manière quantitative avant qu'une analyse qualitative ne soit effectuée.

4.2.1 Évaluation quantitative de l'algorithme de construction d'atlas 2D

L'algorithme de construction d'atlas 2D est évalué quantitativement en analysant la position des centres de la fovéa et du disque optique après déformation des images à recaler. Ces deux points sont en effet les seuls présents sur l'ensemble des images impliquées dans la formation des atlas. Les résultats obtenus pour la base d'images utilisée au cours du projet sont présentés dans le tableau 4.4, sachant que les images de cette base sont de dimensions $1944 * 2592$. Les valeurs contenues dans ce tableau portent sur les écarts-types entre les positions des points dans chaque image et leur position moyenne. Le choix de cette méthode d'évaluation est lié au fait que l'on souhaite superposer du mieux possible les images en vue de la construction de l'atlas et que les centres de la fovéa et du disque optique sont les deux seuls points reliés aux structures qui se trouvent de manière certaine dans chaque image.

Les valeurs obtenues montrent que le recalage d'images de la rétine de personnes différentes permet de positionner les centres de la fovéa et du disque optique suivant une distribution dont l'écart-type est égal à 35.71 pixels ($0.573mm$) pour la fovéa et 32.92 pixels ($0.528mm$) pour le disque optique. La zone dans laquelle se trouvent les centres de la fovéa et du disque optique recalés est limitée puisque ces écarts-types correspondent à environ 1.5% de leur largeur. La figure 4.6 montre sur l'image de référence utilisée pour la construction de l'atlas les régions où se retrouvent les centres du disque optique (en vert) et de la fovéa (en rouge), le rayon des disques placés sur l'image étant égal aux valeurs d'écarts-types présentées dans le tableau 4.4.

Tableau 4.4: Évaluation de l'écart-type des positions des centres de la fovéa et du disque optique après recalage par rapport à leur position moyenne.

	Fovéa	Disque optique
Écart-type (pixels)	35.71	32.92
Écart-type (mm)	0.573	0.528



Figure 4.6: Position des centres de la fovéa (en rouge) et du disque optique (en vert) après recalage d'images de différents fonds d'œil. Les disques ont un rayon égal à l'écart-type relatif à la position des points après recalage par rapport à leur position moyenne.

4.2.2 Évaluation qualitative de l'algorithme de construction d'atlas 2D

L'algorithme de construction d'atlas est évalué ici de deux manières. Sa capacité à superposer les principales structures anatomiques est d'abord analysée en étudiant la répartition des intensités au sein de l'ensemble des images recalées. La capacité du modèle à faciliter la segmentation du réseau vasculaire est ensuite évaluée en procédant à la construction d'une carte de la probabilité des pixels des images de fond d'œil d'appartenir au réseau vasculaire.

Étude de la répartition des structures anatomiques Les intensités des images de fond d'œil peuvent être analysées à l'aide d'un atlas en deux dimensions. Le canal vert des images a été choisi car il s'agit de celui qui contient le plus d'informations en représentation RGB. Ce type d'étude est proposé par Lee *et al.* (2010), la principale différence avec le présent projet étant que Lee et al. ont eu recours à une sélection de points le long des branches principales du réseau vasculaire de la part d'un spécialiste, alors que notre étude est réalisée de manière totalement automatique. La figure 4.7 montre le résultat de la superposition des images de la base utilisée dans le cadre de la construction de l'Atlas. Les intensités ont été normalisées afin que toutes les images possèdent une répartition de leurs niveaux de gris de même moyenne et écart-type. Avec (μ, σ) la moyenne et l'écart-type de la distribution des intensités d'une image et (μ', σ') les valeurs choisies pour la construction de l'atlas, la formule utilisée afin d'adapter le niveau de gris de chaque pixel de l'image est :

$$I' = \frac{I - \mu}{\sigma} \cdot \sigma' + \mu' \quad (4.2)$$

Les figures 4.8 et 4.9 illustrent respectivement la répartition moyenne des intensités sur les images qui constituent l'atlas 2D calculée sur des fenêtres de taille $10 * 10$ et l'écart-type local de cette répartition pour des fenêtres de taille $5 * 5$. Sur la figure 4.8, la coordonnée z correspond à l'intensité moyenne des pixels. Le disque optique présente bien une intensité plus élevée que le reste de l'image alors que la région maculaire apparaît plus sombre. À constater également l'influence des réflexions sur le bord gauche de l'image. Une intensité anormalement élevée est obtenue dans cette zone lors de l'acquisition d'images dans le cas où le diamètre de la pupille des sujets est trop réduit. Sur la figure 4.9, plus l'écart-type local associé à un pixel est élevé, plus son niveau de gris est important. Les régions qui correspondent aux plus fortes densités de présence des vaisseaux sanguins sont plus sujettes aux variations d'intensité lumineuse que le reste des images. Elles correspondent donc à des valeurs d'écarts-types plus élevées.

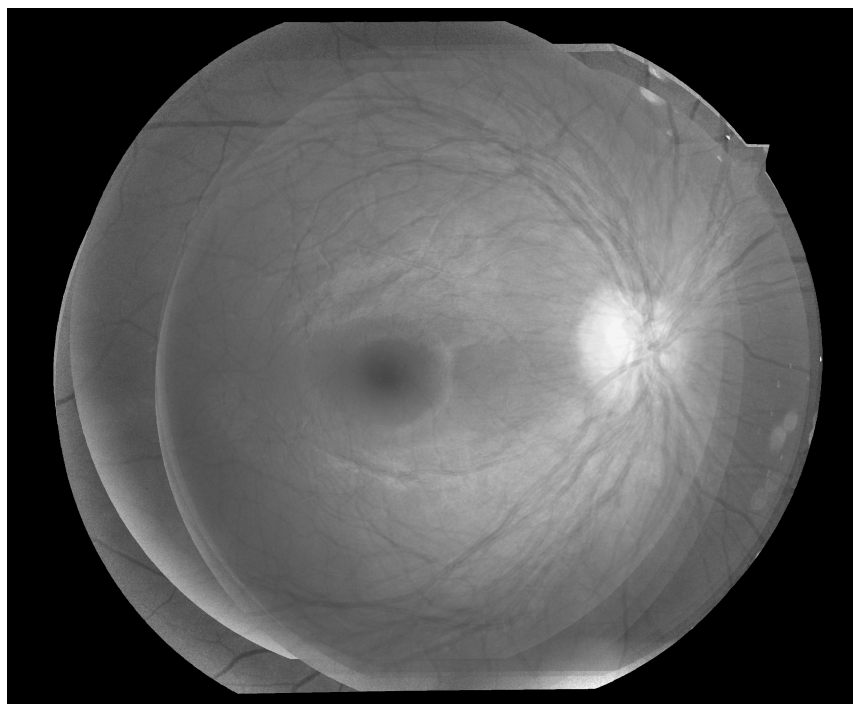


Figure 4.7: Carte de la répartition des intensités sur les images de fond d'œil

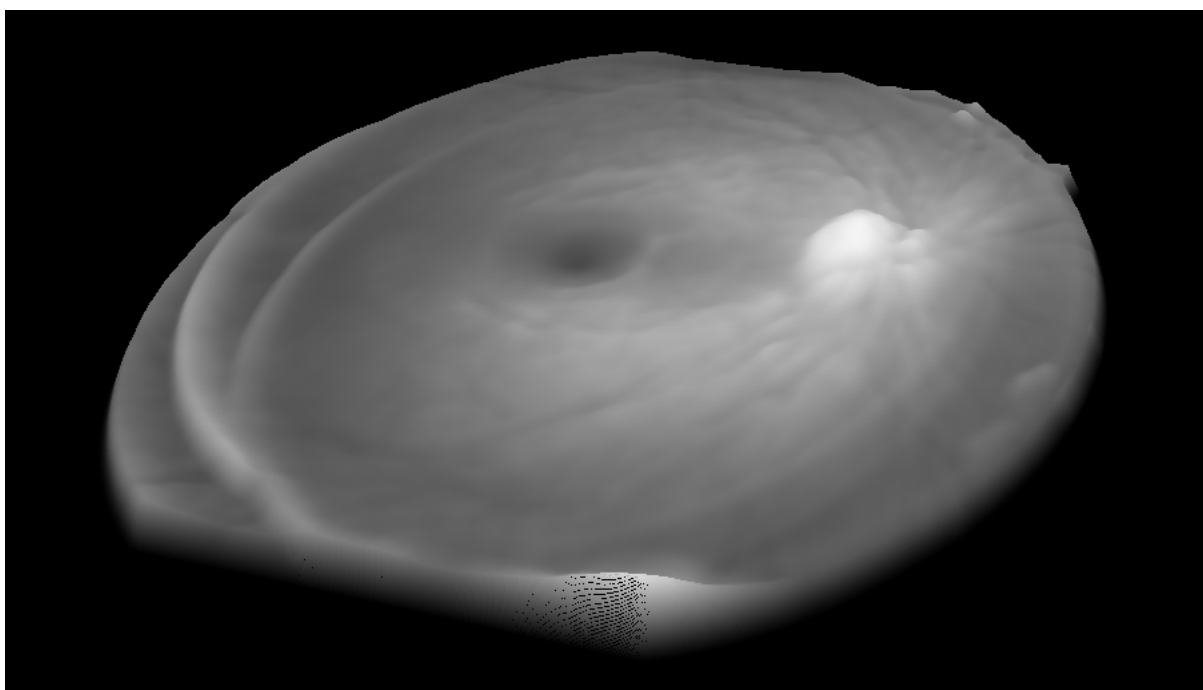


Figure 4.8: Carte de la répartition moyenne des intensités sur les images de fond d'œil.

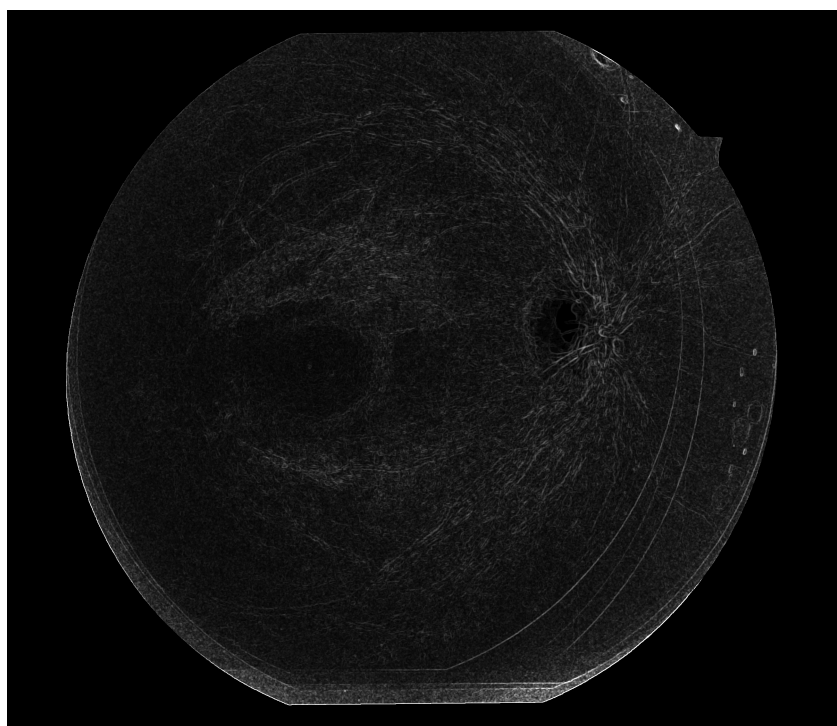


Figure 4.9: Carte des écarts-types locaux de la valeur des pixels du canal vert de l'atlas en deux dimensions.

Appartenance d'un pixel au réseau vasculaire L'objectif de la construction de l'atlas est à court terme de faciliter la segmentation du réseau vasculaire dans les images de fond d'œil. Ces images étant en deux dimensions, une étude de la possibilité d'obtenir de premiers résultats dans ces deux dimensions uniquement a été menée. Cette étude a conduit à la construction de la carte de la probabilité d'un pixel d'une image de fond d'œil d'appartenir au réseau vasculaire.

En effet, l'entreprise *Diagnos Inc.* possède ses propres algorithmes de segmentation du réseau vasculaire, qu'il conviendra d'améliorer suite à la construction de l'atlas. La précision de ces algorithmes est jugée satisfaisante aujourd'hui au regard des autres algorithmes comparables présents sur le marché ou dans la littérature. En segmentant la vasculature sur des images de différentes personnes et en amenant ces images dans un même repère à l'aide de l'algorithme de recalage, il est ainsi possible d'obtenir une carte de la probabilité de l'appartenance d'un pixel au réseau vasculaire sur l'ensemble des images de fond d'œil. En effet, un pixel associé à une intensité élevée sur cette carte se retrouve dans le réseau vasculaire d'un nombre important des images de la base, alors qu'un pixel de faible intensité ne fait que rarement partie de la vasculature. En effectuant un seuillage de cette carte, un atlas du réseau vasculaire est obtenu en deux dimensions, même si les petits vaisseaux présents autour de la zone maculaire ne sont pas considérés dans ce modèle puisqu'ils ont une probabilité moins élevée de se superposer suite au recalage des différentes images que les branches principales de la vasculature. L'atlas 2D obtenu à l'aide de la base d'images disponible pour la réalisation du projet est illustré sur la figure 4.10. Chacune des images correspond à un sujet différent. La seule condition appliquée quant à l'inclusion des images dans la base est que le disque optique, la fovéa et les branches principales du réseau vasculaire doivent appartenir au champ de vision. La limite de cet atlas est que chacune des images utilisées lors de sa construction est dépendante de la géométrie et des conditions de son acquisition.

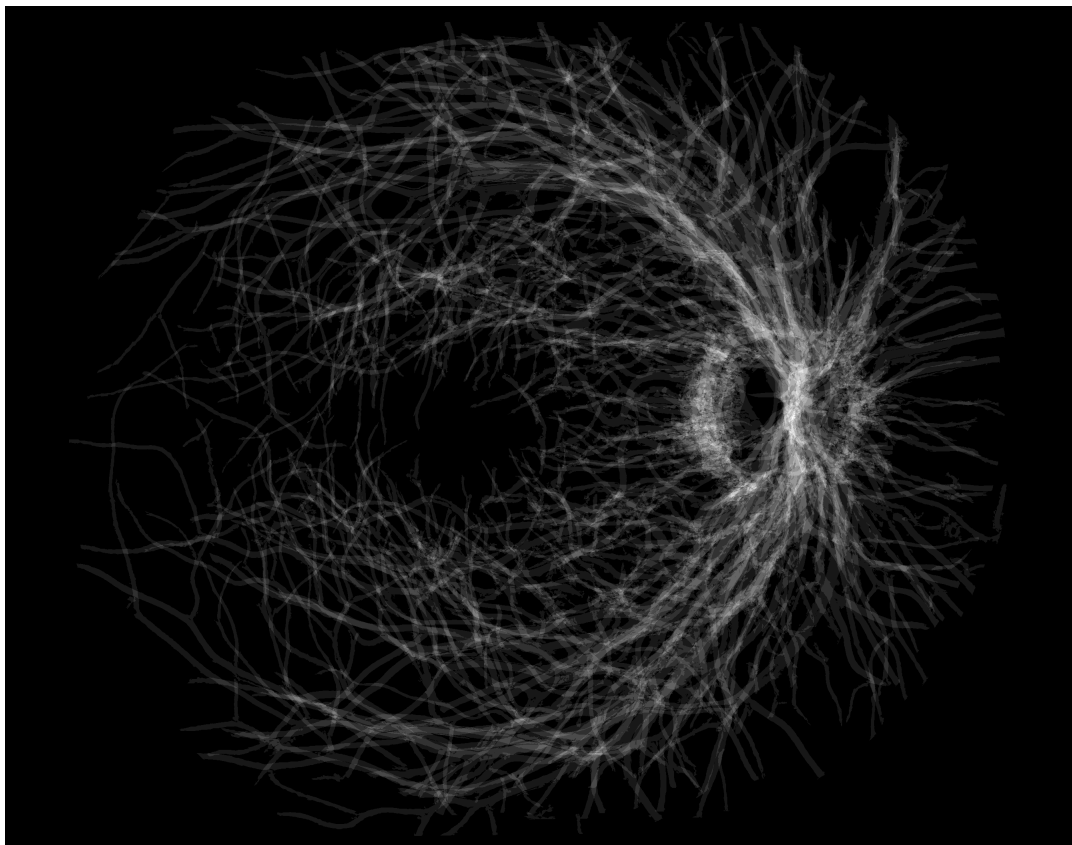


Figure 4.10: Carte de la probabilité des pixels des images de fond d'œil d'appartenir au réseau vasculaire.

4.3 Atlas 3D

Un exemple de carte des disparités entre les images stéréoscopiques d'une rétine, obtenue suivant la méthode SGBM présentée en 3.3.2, est exposé en figure 4.11. Les pixels dont l'intensité est la plus élevée correspondent aux zones de la rétine les plus proches des caméras lors de l'acquisition des images. La partie droite de l'image 4.11 permet également de constater que les réflexions introduisent des perturbations importantes dans le calcul des cartes de disparités.

En associant la valeur de chacun des pixels de cette image à la profondeur correspondante (à un facteur près car il ne s'agit pas d'une reconstruction métrique), et en ajustant l'image couleur de référence sur ce profil 3D, un résultat de la forme de ceux de la figure 4.12 est obtenu.

Utiliser la carte des disparités comme support de la reconstruction 3D signifie que les fonds d'œil sont obtenus avec des profondeurs qui diffèrent d'une transformation projective par rapport à la réalité.

Puisque les reconstructions ne sont pas à l'échelle, il n'est possible que d'apporter une analyse qualitative des reconstructions effectuées grâce à l'algorithme. Sur les trois images de la figure 4.12, on retrouve une élévation de la surface de la rétine par rapport à l'arrière du globe oculaire au niveau de la bordure du disque optique. Elle correspond à la sortie du disque des vaisseaux sanguins qui irriguent la rétine et des fibres nerveuses. Une dépression est également à signaler sur chacune des images au niveau du centre du disque optique. C'est par cette cavité que les tissus nerveux et les vaisseaux sanguins entrent dans la rétine.

En ce qui concerne la zone maculaire, le *clivus*, qui forme un bourrelet autour de la *fovéola* est également visible sur les trois images. Le clivus est la zone qui délimite la rétine vascularisée de la fovéola, zone amincie dont la couche plexiforme interne, les cellules ganglionnaires et les fibres nerveuses sont absentes. C'est dans la fovéola que se concentrent la grande majorité des cônes. La jonction entre les photorécepteurs et les cellules ganglionnaires environnantes se fait par l'intermédiaire de la couche plexiforme externe.

Puisque les images de fond d'œil utilisées en entrée de l'algorithme sont très bruitées et que des erreurs peuvent s'introduire dans l'étape de la mise en correspondance dense entre les images stéréoscopiques, d'importants lissages ont été appliqués. Un premier filtrage a en effet été effectué sur les images originales dans le but d'atténuer le bruit de haute fréquence. Des filtres passe-bas et moyenneurs ont également été appliqués aux cartes des disparités afin d'atténuer l'impact d'erreurs apparues lors de la mise en correspondance dense des images stéréoscopiques. Seul un profil 3D lissé de la surface des fonds d'œil est donc obtenu, l'algorithme ne permettant pas de distinguer de fines structures telles que les vaisseaux sanguins.

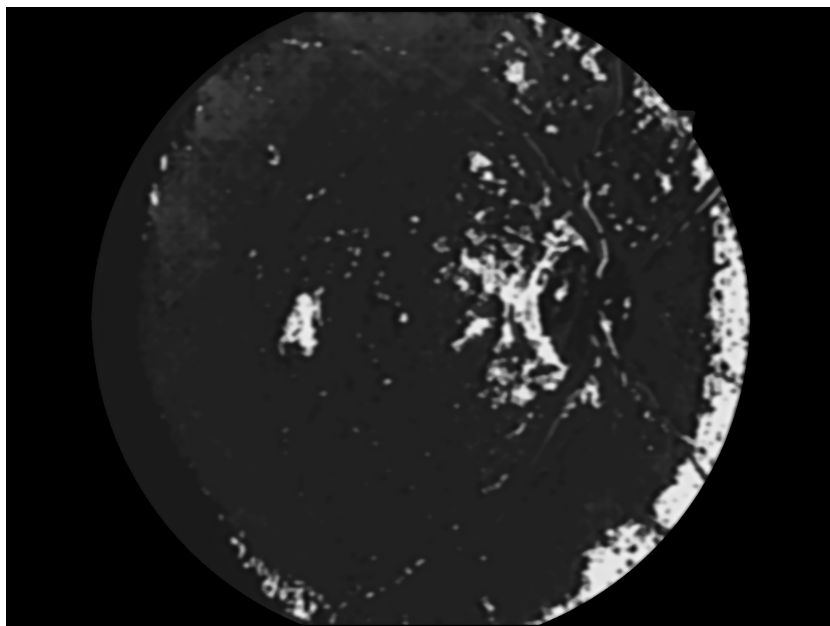
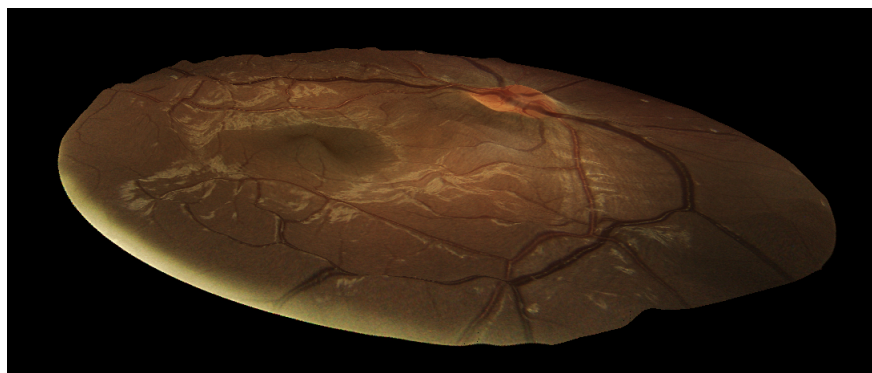
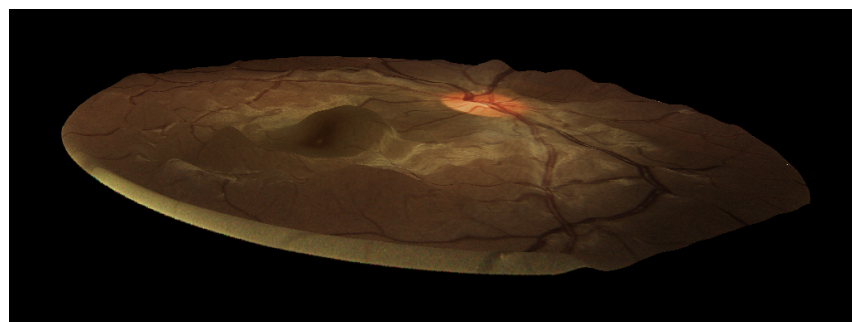


Figure 4.11: Carte des disparités entre deux images stéréoscopiques d'un fond d'œil.

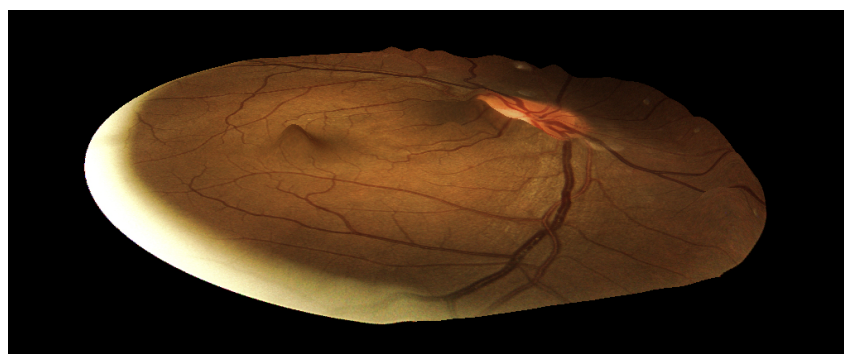
Les trois parties précédentes ont permis une évaluation des algorithmes développés dans le cadre de ce projet de maîtrise en proposant des analyses quantitatives et qualitatives des algorithmes de recalage d'images de fond d'œil et de construction d'atlas ainsi qu'une évaluation qualitative de la reconstruction en trois dimensions de la surface de la rétine.



(a) Sujet a



(b) Sujet b



(c) Sujet c

Figure 4.12: Reconstruction 3D de fonds d'œil à un facteur près.

CHAPITRE 5

CONCLUSION

Le principal objectif de ce projet était de développer un algorithme de recalage d'images de fonds d'œil dans le but de construire un atlas de structures anatomiques de la rétine. Un tel atlas vise à analyser la répartition des principales structures anatomiques et à faciliter l'initialisation des algorithmes de segmentation, en particulier du réseau vasculaire. Le contenu des images en deux dimensions dépendant de la géométrie de leur acquisition, il était également nécessaire d'étudier la possibilité de reconstruire la surface de la rétine en trois dimensions.

Le recalage d'ordre 2 des images de fond d'œil développé tient compte des transformations rigides et du caractère sphérique des rétines, ce qui en fait un recalage adapté à cette structure anatomique. Le fait d'avoir construit l'algorithme de manière progressive en ce qui concerne l'ordre des déformations appliquées aux images à recaler permet de l'utiliser jusqu'à l'ordre 2 dans le cas de recalages en deux dimensions pour la construction de mosaïques ou d'atlas et de l'intégrer jusqu'à l'ordre 1 dans l'algorithme de reconstruction en trois dimensions afin de ne conserver que les déformations liées à la profondeur dans les images stéréoscopiques.

D'abord appliqué au cas d'un recalage intra-sujet, l'algorithme de recalage permet d'obtenir une moyenne des distances entre 30 points correspondants entre les images recalées égale à 2.51 pixels, ce qui correspond à $0.040mm$ sur la surface d'une rétine. Des mosaïques constituées de multiples images rétinienne ont ainsi pu être formées et des patients suivis au cours du temps afin d'examiner la progression de pathologies évolutives.

L'atlas 2D est construit en positionnant des ellipses sur les images, ce qui permet de pallier au fait qu'il n'y ait que peu de similitudes entre les images de différents fonds d'œil. Les positions des centres de la fovéa et du disque optique des images recalées ont été comparées afin d'évaluer la méthode de construction d'atlas. L'écart-type de distribution de ces points par rapport à leur position moyenne est égal à 35.71 pixels ($0.573mm$) pour la fovéa et 32.92 pixels ($0.528mm$) pour le disque optique, ces écarts-types correspondant à environ 1.5% de la largeur des images utilisées.

La construction des ellipses repose principalement sur la détermination des centres de la fovéa et du disque optique. La méthode proposée afin de déterminer le centre de ce dernier permet bien de superposer convenablement le disque de chacune des images, l'écart-type de la répartition de ces points par rapport à leur moyenne étant égal à 32.92 pixels ($0.528mm$). Des travaux sont toutefois en cours, notamment chez Diagnos Inc., afin d'améliorer la détection

des contours du disque optique. La détermination du centre du disque pourrait ainsi être effectuée en cherchant le barycentre de la zone considérée comme le représentant sur les images de fond d'œil.

L'algorithme de construction d'atlas 2D a permis d'effectuer une étude de la répartition des principales structures rétinienne et de former une carte de la probabilité des pixels des images de fonds d'œil d'appartenir au réseau vasculaire. Ces deux études seraient certainement améliorées si plus d'images de sujets sains étaient recalées à l'aide de l'algorithme de construction d'atlas. La carte de la probabilité des pixels des images de fond d'œil d'appartenir au réseau vasculaire et celles qui permettent l'analyse de la répartition des structures anatomiques sur les rétines seraient plus pertinentes si elles étaient formées à partir d'une population plus nombreuse. Nous avons l'intention d'acquérir un grand nombre d'images de sujets sains à l'École Polytechnique afin d'affiner nos résultats sur une base d'images plus conséquente. Un dossier présenté au comité d'éthique de l'école a récemment été approuvé en vue de procéder à ces acquisitions.

En ce qui concerne la reconstruction en trois dimensions, les profils obtenus font ressortir les principales structures impliquées dans le relief des rétines que sont le disque optique et la fovéa. Les fines structures telles que les vaisseaux sanguins ne peuvent cependant être distinguées de la surface des rétines car d'importants lissages sont appliqués.

La reconstruction est obtenue à une transformation projective près, car les caméras ne sont pas calibrées et que leurs paramètres intrinsèques sont inconnus. Des élèves du LIV4D travaillent toutefois au calibrage des caméras. L'objectif est d'utiliser un objet de calibrage, qui est dans ce cas un modèle d'œil factice, d'obtenir une image en trois dimensions des couches proches de la surface par le biais de la modalité d'imagerie OCT, de sélectionner des points d'intérêt sur cette image ainsi que sur des images stéréoscopiques de ce même objet acquises grâce aux caméras de fond d'œil et de procéder à leur calibrage. Une fois les caméras calibrées, il sera possible de construire un atlas de la surface de la rétine en trois dimensions qui permette d'étudier la répartition des structures anatomiques visibles sur les fonds d'œil dans l'espace.

RÉFÉRENCES

- A. VEDALDI (2006). An Implementation of SIFT Detector and Descriptor. University of california.
- ABRAMOFF, M. D., GARVIN, M. K. et SONKA, M. (2010). Retinal Imaging and Image Analysis. *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169–208.
- AMUNTS, K., LEPAGE, C. et BORGEAT, L. (2013). BigBrain : an Ultrahigh-Resolution 3D Human Brain Model. *Science*, 340, pp. 1472–1475.
- BAY, H., TUYTELAARS, T. et GOOL, L. V. (2006). SURF : Speeded Up Robust Features. *European Conference on Computer Vision*, 1, pp. 404–417.
- BEIS, J. et LOWE, D. (1997). Shape Indexing Using Approximate Nearest-Neighbour Search in High-Dimensional Spaces. *Conference on Computer Vision and Pattern Recognition*, pp. 1000–1006.
- BOOKSTEIN, F. (1989). Principal Warps : Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, pp. 567–585.
- BOYKOV, Y., VEKSLER, O. et ZABIH, R. (2001). Fast Approximate Energy Minimisation via Graph Cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, pp. 1222–1239.
- BRICE, C. et FENNEMA, C. (1970). Scene analysis using regions. *Artificial Intelligence Journal*, 1, pp. 205–226.
- BROWN, M. et LOWE, D. (2002). Invariant Features from Interest Point Groups. *British Machine Vision Conference*, 1, pp. 656–665.
- C. EVANS (2009). Notes on the OpenSURF Library. University of Bristol.
- CAN, A., STEWART, C. V., ROYSAM, B. et TANENBAUM, H. L. (2002). A Feature-Based, Robust, Hierarchical Algorithm for Registering Pairs of Images of the Curved Human Retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 347–364.
- CANNY, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, pp. 679–698.
- CINSDIKICI, M. et AYDIN, D. (2009). Detection of Blood Vessels in Ophthalmoscope Images Using mf/ant (Matched Filter / ant Colony) Algorithm. *Computer Methods and Programs in Biomedicine*, 96, pp. 85–95.

- DEHOOG, E. et SCHWIEGERLING, J. (2009). Fundus Camera Systems : a Comparative Analysis. *Circulation*, 48, pp. 221–228.
- D.RIVEST-HÉNAULT, CHERIET, M., DESCHÊNES, S. et LAPIERRE, C. (2010). Length Increasing Active Contour for the Segmentation of Small Blood Vessels. *IEEE 20th International Conference on Pattern Recognition*, pp. 2796–2799.
- DROPPELMANN, S., HUETING, M., LATOUR, S. et VAN DER VEEN, M. (2010). Stereo Vision using the OpenCV Library.
- DUBRA, A., SULAI, Y., NORRIS, J., COOPER, R., DUBIS, A., WILLIAMS, D. et CARROLL, J. (2011). Noninvasive Imaging of the Human Rod Photoreceptor Mosaic Using a Confocal Adaptive Optics Scanning Ophthalmoscope. *Biomedical Optics Express*, 2, pp. 1864–1876.
- F. CHERIET (2011a). Calibrage de Systèmes à Rayons-X.
- F. CHERIET (2011b). Mise en correspondance de primitives.
- FUJIMOTO, J., PITRIS, C., BOPPART, S. et BREZINSKI, M. (2000). Optical Coherence Tomography : An Emerging Technology for Biomedical Imaging and Optical Biopsy. *Neoplasia*, 2, pp. 9–25.
- FUSIELLO, A. (2000). Uncalibrated Euclidean Reconstruction : A Review. *Image and Vision Computing*, 18, pp. 555–563.
- HARRIS, C. et STEPHENS, M. (1988). A combined corner and edge detector. *Alvey Vision Conference*, pp. 147–152.
- HARTLEY, R. (1997). In Defense of the Eight-Point Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, pp. 580–593.
- HOOGENDOORN, C. (2013). A High-Resolution Atlas and Statistical Model of the Human Heart from Multislice CT. *IEEE Transactions on Medical Imaging*, 32, pp. 28–44.
- JI, Q., COSTA, M., HARALICK, R. et SHAPIRO, L. (1999). An Integrated Linear Technique for Pose Estimation from Different Geometric Features. *International Journal of Pattern Recognition and Artificial Intelligence*, 13, pp. 705–733.
- KOVÁCS, G. et HAJDU, A. (2011). Extraction of Vascular System in Retina Images Using Averaged One-Dependence Estimators and Orientation Estimation in Hidden Markov random Fields. *IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pp. 693–696.
- LEE, S., ABRAMOFF, M. et REINHARDT, J. (2010). Retinal Atlas Statistics from Color Fundus Images. *SPIE Medical Imaging*, 7623, pp. 762310–1–762310–9.

- LIEW, G. et WANG, J. (2011). Retinal Vascular Signs : A Window to the Heart ? *Revista Espanola de Cardiologia*, 64, pp. 515–521.
- LINDEBERG, T. (1994). Scale-Space Theory : A Basic Tool for Analysing Structures at Different Scales. *Journal of Applied Statistics*, 21, pp. 224–270.
- LOWE, D. (1999). Object Recognition from Local Scale-Invariant Features. *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, 2, pp. 1150–1157.
- MARIEB, E. (1998). Anatomie et physiologie humaines.
- MARR, D. et HILDRETH, E. (1980). Theory of Edge Detection. *Proceedings of the Royal Society*, 207, pp. 187–217.
- MATSUI, M., TASHIRO, T., MATSUMOTO, K. et YAMAMOTO, S. (1973). A Study on Automatic and Quantitative Diagnosis of Fundus Photographs. *Nippon Ganka Gakkai Zasshi*, 77, pp. 907–918.
- NOVOTNY, H. et ALVIS, D. (1961). A Method of Photographing Fluorescence in Circulating Blood in the Human Retina. *Circulation*, 24, pp. 82–86.
- OMS (2013). Prévention de la cécité et des déficiences visuelles.
- OTSU, N. (1979). A Threshold Selection Method from Gray Level Histograms. *IEEE Transactions on Systems Man Cybernet*, 9, pp. 62–66.
- SANTALLIER, M., PÉCHEREAU, J. et PÉCHEREAU, A. (2008). *Anatomie pour les écoles d’orthoptie, v1.0*. Alain et Jocelyne Péchereau.
- SOFKA, M. et STEWART, C. (2006). Retinal Vessel Centerline Extraction Using Multiscale Matched Filters, Confidence and Edge Measures. *IEEE Transactions on Medical Imaging*, 25, pp. 1531–1546.
- SWANSON, E., IZATT, J., HEE, M., HUANG, D., LIN, C., SCHUMAN, J., PULIAFITO, C. et FUJIMOTO, J. (1993). In Vivo Retinal Imaging by Optical Coherence Tomography. *Arch. Ophthalmol*, vol. 18, pp. 1864–1866.
- THIVIERGE-GAULIN, D. (2012). *Mesh-Derived Image Partition for 3D-2D Registration in Image-Guided Interventions*. Thèse de doctorat, École Polytechnique de Montréal.
- VON GIOI, R. G., JAKUBOWICZ, J., MOREL, J. et G.RANDALL (2010). LSD : A Fast Line Segment Detector with a False Detection Control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32, pp. 722–732.
- WU, C., DERWENT, J. et STANCHEV, P. (2009). Retinal Vessel Radius Estimation and a Vessel Center Line Segmentation Method Based on Ridge Descriptors. *Journal of Signal Pocessing Systems*, 55, pp. 91–102.

- XU, J. et CHUTATAPE, O. (2006). Auto-Adjusted 3-D Optic Disk Viewing from Low-Resolution Stereo Fundus Images. *Computers in Biology and Medicine*, 36, pp. 921–940.
- XU, L. et LUO, S. (2010). A Novel Method for Blood Vessel Detection from Retinal Images. *Biomedical Engineering Online*, 9, pp. 14.
- Y. LACHKAR AND E. SELLEM (2004). La Papille Glaucomateuse en Pratique.
- YEDIDYA, T. et HARTLEY, R. (2008). Tracking of Blood Vessels in Retinal Images Using Kalman Filter. *IEEE Digital Image Computing, : Techniques and Applications*, pp. 52–58.
- ZHANG, M. et LIU, J. (2007). Directional Local Contrast Based Blood Vessel Detection in Retinal Images. *IEEE International Conference on Image Processing*, 4, pp. 317–320.