

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



UNIVERSITE DJILLALI LIABES DE SIDI-BEL-ABBES
Faculté de Génie Electrique
Département d'Electronique

Thèse présentée par :

Abdelkrim GHAZ

Pour l'obtention du diplôme de :

Doctorat en Sciences
Spécialité: Electronique
Option: Traitement du signal

Intitulé de la thèse :

**EXTRACTION DE CONTOURS MULTIRÉSOLUTIONS PAR LES
TRANSFORMÉES GÉOMÉTRIQUES POUR LE MOSAIQUAGE DES
IMAGES MEDICALES**

Présentée devant le jury composé de :

Nasreddine TALEB	Professeur (U.D.L. Sidi Bel-Abbès)	Président
Abdennacer BOUNOUA	Professeur (U.D.L. Sidi Bel-Abbès)	Rapporteur
Kidiyo KPALMA	Professeur (INSA de Rennes)	Co-directeur
Abdelhafid BESSAID	Professeur (U.A.B. Tlemcen)	Examineur
Youcef BENTOUTOU	Dir de Recherche (C.D.S. Oran)	Examineur
Kamel BELLOULA	Professeur (U.D.L. Sidi Bel-Abbès)	Examineur

À Mes chers Parents

À ma Femme

À mes enfants: Cheïmaa, Kawther, et Mohammed

À toute ma Famille

Abdeskrim ghaz

Remerciements

Je remercie mon directeur de thèse, M. Abdennacer Bounoua, Professeur à l'Université Djillali Liabes de Sidi-Bel-Abbes, pour m'avoir encadré avec compétence et disponibilité tout au long de ma thèse. Ses connaissances étendues en traitement d'images ont été un support et une inspiration pour mon travail.

Je suis très reconnaissant à mon Co-directeur de thèse, M. Kidiyo KPALMA, Professeur des universités à l'INSA de Rennes (France) pour m'avoir accueilli au sein du laboratoire groupe image et télédétection et de m'avoir encadré tout au long de cette thèse. Je le remercie aussi pour ses conseils et de m'avoir consacré son temps à chaque fois que j'avais besoin de lui. Ses qualités à la fois en tant que personne et en tant que chercheur me serviront toujours comme un exemple.

Surtout, je suis profondément redevable à M. Nasreddine Taleb, Professeur à l'Université Djillali Liabes de Sidi-Bel-Abbes (Algérie) pour son aide lors de la rédaction de mon article scientifique, pour ces encouragements scientifiques et humains et d'avoir accepté de présider ce jury, je lui doit toute ma gratitude.

Je tiens à exprimer ma gratitude au comité d'examen de la thèse:

Abdelhafid BESSAID, Professeur à Abou Bekr Belkaid Université de Tlemcen (Algérie). Kamel BELLOULA, Professeur à l'Université Djillali Liabes de Sidi-Bel-Abbes (Algérie), Youcef BENTOUTOU Directeur de Recherche en C.D.S. Oran (Algérie), Pour leur temps et pour leurs remarques constructives.

Je voudrais remercier toutes les personnes qui, d'une façon ou d'une autre, m'ont soutenu dans mes recherches et pendant mes stages effectués à Rennes, en France.

Un grand merci à tous mes collègues de l'Université Djillali Liabes de Sidi Bel-Abbes en particulier ceux de la faculté de Génie Electrique.

Je voudrais remercier tous mes amis, pour leurs soutiens et leurs encouragements.

Merci enfin à ma femme et mon adorable fils et mes gentilles filles. À mes chers parents, sœurs et frères, pour m'avoir soutenu si fortement pendant toute la durée de ma thèse. Merci à toute ma famille.

Résumé

L'utilisation des images panoramiques est largement répandue dans différents domaines tels que la robotique, la vision par ordinateur et la surveillance. Ces images sont également utilisées en médecine pour éliminer la redondance de l'information contenue dans les clichés des angiographies rétiniennes et de l'endoscopie vésicale. Ainsi elles apportent une aide énorme pour l'ophtalmologiste et l'urologue dans le diagnostic des pathologies de la rétine et de la vessie et permettre le suivi de l'évolution de ces pathologies et de leurs traitements.

L'objectif, affiché, de ce travail consiste à l'élargissement du champ visuelle et la création d'une vue globale de la rétine et de la vessie.

Pour parvenir à ce but; dans la première grande phase, on applique une décomposition multirésolution **NSCT** sur l'image originale pour rehausser les contours, dans la deuxième grande phase, on extrait les points clés à l'aide de la technique **SIFT** pour le calcul de la transformée servant au recalage et à l'assemblage des différentes parties de l'image panoramique.

L'extraction des points clés en rehaussant l'image originale par la **NSCT** et l'application de **SIFT** ont permis de donner des points d'intérêt plus stables et de diminuer les erreurs de correspondance entre les points des différentes portions d'images à mosaïquer même en présence de fort bruit. Aussi, on a pu corriger les sutures et la différence d'exposition lumineuse entre les images en utilisant le '**pyramid Blending**' et améliorer la qualité visuelle.

Les résultats de mosaïquage obtenus par l'application de l'ensemble des algorithmes développés sur les images angiographiques et cystoscopiques, ont donné des vues panoramiques élargies avec un temps de calcul assez rapide, ce qui offre aux cliniciens un outil d'aide au diagnostic assez fiable.

Mots clé : SIFT, NSCT, pyramid Blending, rehaussement de contours, recalage.

ملخص

إنّ استعمال الصور البانورامية منتشر في مختلف المجالات مثل الروبوتات ، الرؤية عن طريق الكمبيوتر و المراقبة ، كما تستخدم في المجال الطبي لإزالة المعلومات المتكررة في صور أوعية الشبكية و نسيج المثانة و ذلك لمساعدة طبيب العيون و أخصائي أمراض الكلى للكشف عن أمراض الشبكية و المثانة و مراقبة العلاج. لبلوغ هدفنا المتمثل في الحصول على صورة إجمالية لشبكة العين و المثانة قمنا في أول الأمر بتحليل الصورة عن طريق تحويلات الصور متعددة-السلم و متعددة - الاتجاهات NSCT لتعزيز معالم الصور.

في الخطوة الثانية تم استخراج النقاط الأساسية المثيرة للاهتمام عن طريق التحويل SIFT تستعمل هذه النقاط لحساب التحويل الملائم لجمع الصور و تركيبها في نظام احداثي موحد.

لقد تم إستخراج النقاط الرئيسية عن طريق تعزيز محيطها بالتحويل NSCT ثم تطبيق التحويل SIFT الذي أعطى نقاطا أكثر استقرارا و صحّح الخلل في التطابق بين أجزاء الصور التي نريد تركيبها و ه ذه الخوارزمية فعالة حتى بوجود تشويشات كما قمنا بتصحيح الحدود الفاصلة بين أجزاء الصور و الإختلاف في التعرض للضوء باستخدام الخلط الهرمي pyramid Blending للحصول على صورة عالية الجودة.

بشكل عام إن النتائج التي تحصلنا عليها عن طريق التحويلات للحصول على صورة شاملة لشبكية العين و المثانة سمحت باختزال وقت الحسابات مما يتيح للأطباء استغلال هذه النتائج أثناء الفحص الكلينيكي للمريض و بالتالي يساعد في الكشف عن الأمراض و متابعة العلاج.

Abstract

The use of panoramic images is widely used in various fields such as robotics, computer vision and surveillance. It is in particular used in medicine to eliminate redundant information in retinal angiography and vesical endoscopy and help the ophthalmologist and urologist to better detect retinal and bladder pathologies and better follow up the evolution of the pathology and subsequent treatments.

To achieve our goal and create a global view of the retina and of the bladder, In a first step, we have applied an NSCT multiresolution decomposition on the original image to enhance the contours.

In a second step, we have extracted the key points using **SIFT** for calculating the registration transform used to assembly the different parts of the panoramic image.

The extraction of the key points, when enhancing the original image using **NSCT** and applying **SIFT**, has given more stable points of interest and helped reducing mismatches between points of different portions of the images, to be mosaicked, even in the presence of higher noise. In addition, the developed algorithms permit to correct seams as well as the difference in light exposure between images using the **pyramid Blending** and to improve visual quality.

The obtained mosaicking results, by applying all mosaicking processes upon retinal angiography and cystoscopy, are satisfactory in sense of fairly fast computation time for clinicians to use these results in the clinical examination of the patient.

Sommaire

Introduction générale	1
Chapitre 1: Intérêt du mosaïquage pour les images médicales	
1.1. Introduction.....	4
1.2. Anatomie de l'œil.....	5
1.3. Les examens complémentaires en ophtalmologie.....	8
1.3.1. Echographie oculaire.....	8
1.3.2. L'Angiographie à la fluorescéine rétinienne.....	9
1.3.2.1. Interprétation de l'angiographie à la fluorescéine.....	10
1.3.3. Angiographie numérisée ICG.....	11
1.3.4. Tomographie en cohérence optique OCT.....	11
1.4. Les pathologies de la rétine.....	12
1.4.1. Rétinopathie diabétique.....	12
1.4.2. La Dégénérescence Maculaire Liée à l'Age DMLA.....	13
1.4.3. Importance du test d'Amsler pour le diagnostic de la DLMA.....	14
1.5. L'endoscopie vésicale et le cancer de la vessie.....	15
1.5.1. Caractéristiques des tissus urothéliaux.....	16
1.5.2. Types et classification des cancers de la vessie.....	16
1.6. L'endoscopie vésicale et les problèmes scientifiques.....	18
1.6.1. La qualité des images cystoscopiques.....	18
1.6.2. Contraintes cliniques influençant le mosaïquage des images cystoscopiques.....	21
1.7. Conclusion.....	22
Chapitre 2: Extraction des points d'intérêt	
2.1. Introduction.....	23
2.2. Définition de point d'intérêt.....	24
2.3. Méthodes de détection des points d'intérêt.....	24
2.3.1. Méthodes à base de modèles.....	25
2.3.2. Méthodes à base géométrique.....	25
2.3.2.1. Méthodes basées sur les contours.....	25
2.3.2.2. Méthodes basées sur l'intensité.....	25

2.4. Quelques exemples de détecteurs des points d'intérêt	26
2.4.1. Détecteur de Moravec (1977).....	26
2.4.2. Détecteur de Harris (1988).....	28
2.4.3. Principe de fonctionnement et les étapes du détecteur SIFT.....	31
2.4.3.1. Détection d'extrema dans l'espace des échelles.....	31
2.4.3.2. Espace-échelle de détection des extremas.....	32
2.4.3.3. Localisation des points clés.....	34
2.4.3.4. Affectation d'orientation aux points d'intérêt.....	35
2.4.4. Détecteur de Points d'intérêt Basé sur la NSCT	37
2.5. Évaluation des détecteurs étudiés	41
2.6. La mise en correspondance.....	46
2.6.1. Difficultés et contraintes de la mise en correspondance	46
2.6.1.1. Difficultés de la mise en correspondance.....	46
2.6.1.2. Contrainte de la mise en correspondance.....	47
2.6.2. Les méthodes de mise en correspondance.....	48
2.6.2.1. Méthodes locales	48
2.7. Conclusion.....	51
Chapitre 3: Recalage et Evaluation de l'algorithme NSCT-SIFT	
3.1. Introduction.....	52
3.2. Principe de recalage.....	53
3.2.1. Approches iconique et géométriques.....	54
3.2.2. Méthodes utilisant les relations spatiales entre les primitives.....	55
3.3. Estimation du modèle de transformation.....	56
3.3.1. Transformation affine.....	56
3.3.2. Transformation projective.....	57
3.3.3. Interpolation des niveaux de gris des images transformées.....	58
3.4. Méthode proposée basée sur NSCT SIFT.....	59
3.4.1. Calcul des coefficients de pondération	62
3.5. Détection des points d'intérêt.....	63
3.5.1. Recalage des images après extraction des points de contrôles.....	65
3.6. Application de l'algorithme NSCT- SIFT sur les images médicales.....	71

3.6.1. Précision de recalage.....	71
3.6.2. Appariement des points d'intérêt.....	72
3.7. Conclusion.....	75
Chapitre 4 : Mosaïquage et blending	
4.1. Introduction.....	76
4.2. Acquisition et prétraitement des images.....	76
4.3. Algorithme du mosaïquage.....	76
4.3.1. La sélection de l'image référence.....	78
4.3.2. Différentes méthodes du mélange (blending).....	79
4.3.3. Sélection des sutures.....	80
4.4. Poisson blending.....	81
4.5. Correction des sutures par pyramid blending.....	82
4.6. Résultats du mosaïquage des images angiographiques de la rétine.....	83
4.7. Résultats du mosaïquage des images cystoscopiques.....	86
4.8. La correction d'exposition.....	89
4.9. Conclusion.....	90
Conclusion générale.....	91
Références bibliographiques.....	93

Liste des illustrations

Figure 1.1 : Echographie Anatomie de l'œil humain	5
Figure 1.2 : (a) coupe de l'œil ; (b) Schéma simplifié de l'architecture de la rétine	7
Figure 1.3: oculaire	8
Figure 1.4 : Angiographie à la fluorescéine rétinienne en trois couleurs.....	9
Figure 1.5 : Angiographie à la fluorescéine normale	11
Figure 1.6: Examen OCT de la rétine.....	12
Figure 1.7 : Rétinopathie diabétique	13
Figure 1.8 : Dégénérescence Maculaire Liée à l'Age.....	14
Figure 1.9: Test d'Amsler : (a) test de DMLA négatif (vision normale), (b) test de DMLA positif (vision anormale).....	15
Figure 1.10 : (a) L'appareil urinaire. (b) Représentation des couches constituantes de la paroi de la vessie.	16
Figure 1.11 : Différents degrés d'invasion d'une tumeur de vessie.	17
Figure 1.12 : Schéma illustrant le trajet de l'endoscope pour la construction d'images panoramiques de l'intérieur de la vessie.	18
Figure 1.13: Les deux types de cystoscope utilisés en clinique.	19
Figure 1.14 : Quelques exemples d'images pouvant être extraites d'une séquence vidéo cystoscopique.	20
Figure 2.1 : Différents type de coin : (a) jonction en "L", (b) jonction en "V ",.....	24
Figure 2.2 : Les différentes situations considérées par le détecteur de Moravec.	27
Figure 2.3 : Les points détectés avec une réponse forte sur les contours.....	28
Figure 2.4 : La détection de coins pour Harris.	29
Figure 2.5: Répartition des valeurs de R.	30
Figure 2.6 : Construction de la pyramide avec différentes échelles.....	32
Figure 2.7 : Illustration de calcul de la différence gaussienne DOG	33
Figure 2.8 : Histogramme d'orientation des points d'intérêt.....	36
Figure 2.9 : Détections des points d'intérêt par SIFT	37
Figure 2.10 : La transformée en contourlets non sous-échantillonnée. (a) Schéma bloque de l'implantation de la NSCT à l'aide d'un banc de filtre pyramidal directionnel non sous-échantillonné. (b) le partitionnement fréquentiel résultant, où le nombre de directions augmente avec la fréquence.....	38
Figure 2.11 : La transformée en contourlets non sous-échantillonnée de l'image "Zoneplate". L'image est décomposée en trois niveaux pyramidaux, qui sont ensuite décomposés en 2, 4, et 8	

sous-bandes directionnelles. Les petits coefficients sont représentés en noir alors que les grands coefficients sont représentés en blanc. (a) L'image originale. (b) Sous-bande passe-bas. (c), (d), (e) Sous-bandes directionnelles passe-bande.....	39
Figure 2.12 : Extraction des points d'intérêt utilisant différents détecteurs.....	42
Figure 2.13 : Différence de détections des points entre SIFT et Harris.	43
Figure 2.14 : Extraction des points d'intérêt de l'image synthétique 2 contenant plusieurs formes de jonctions: (a) Le détecteur de Moravec (b) Le détecteur de Harris $\sigma=1.5$ (c) Le détecteur à base de NSCT (d) Le détecteur de SIFT.	44
Figure 2.15 : Exemple de mise en correspondance de deux pixels.	46
Figure 2.16 : La contrainte d'unicité: dans la figure nous pouvons voir les correspondances des pixels en bleu qui respectent l'unicité alors que celles en rouge la ne respectent pas.....	47
Figure 2.17: Contrainte d'ordre - nous pouvons voir deux correspondances, l'une en bleu qui respecte la contrainte de l'ordre et l'autre en rouge qui ne la respecte pas.	48
Figure 2.18: Principe de la corrélation : recherche du point issu de la seconde image présentant la meilleure ressemblance.	49
Figure 2.19 : Exemple de mesures de corrélation entre le point x_1 et les points x_2 présents dans la zone de recherche.	49
Figure 2.20 : Correspondance des points d'intérêt entre l'image d'angiographie rétinienne et son homologue recalé de 7 degré.....	50
Figure 3.1 : Images angiographiques de la rétine (a) Rétine saine (b) Rétine pathologique.....	53
Figure 3.2 : Illustration de l'approche directe IR est l'image originale et ITr est l'image transformée. T est la transformation appliquée aux coordonnées des images.....	58
Figure 3.3 : Estimation de l'intensité X_c à partir des intensités des pixels voisins par interpolation bilinéaire.....	59
Figure 3.4 : Diagramme en blocs de l'algorithme proposé NSCT –SIFT pour l'extraction des points de contrôle.....	62
Figure 3.5 : Choix du coefficient et de la méthode de rehaussement des contours.....	63
Figure 3.6 : Comparaison des trois détecteurs selon le nombre et la précision de localisation des points de contrôle.....	64
Figure 3.7 : Images recalées LENA et BOAT (a) Utilisant la méthode SIFT, (PSNR = 34.1389) (b) Utilisant la méthode NSCT-SIFT, (PSNR = 38.5915) (c) Utilisant la méthode SIFT méthode, (PSNR = 41.3492) (d) Utilisant la méthode NSCT-SIFT méthode, (PSNR = 42.8109).....	66
Figure 3.8 : Image recalée superposée à l'image (a) Avec la méthode SIFT (b) Avec la méthode NSCT-SIFT.....	67
Figure 3.9 : Comparaison du PSNR après recalage entre les transformée SIIFT et NSCT-SIFT....	67

Figure 3.10 : Image IKONOS recalée utilisant (a) La méthode SIFT, PSNR =35.1609 (b) La méthode NSCT-SIFT, PSNR =38.5511.....	69
Figure 3.11 : La détection et l'appariement des points d'intérêt sur l'image référence et l'image à recaler avec le détecteur SIFT pour deux seuils différents.....	73
Figure 3.12 : La détection et l'appariement des points d'intérêt sur l'image référence et l'image à recaler avec le détecteur NSCT-SIFT pour deux seuils différents.....	74
Figure 4.1 : Schéma bloc du mosaïquage.....	77
Figure 4.2 : Relation des images à mosaïquer selon la surface de chevauchement.....	78
Figure 4.3 : Création du masque pour limiter les points à l'intérieur du cercle image.....	80
Figure 4.4 : Assemblage des masques des images à mosaïquer.....	80
Figure 4.5 : Deux mosaïques de l'angiographie rétinienne avec sutures visibles.....	81
Figure 4.6 : Le masque de la région commune.....	81
Figure 4.7 : Résultat de poisson blending obtenu après mosaïquage.....	82
Figure 4.8 : Différents niveaux du 'pyramid blending'.....	83
Figure 4.9 : Rétinopathie diabétique non proliférante sévère.....	84
Figure 4.10 : Œdème maculaire tête flèche sur logette cystoïde.....	84
Figure 4.11 Mosaïque formé par 8 images montrant une dilatation diffuse et tortuosité vasculaire veineuse.....	85
Figure 4.12 : Rétinopathie diabétique proliférante modérée.....	85
Figure 4.13 : 27 images cystoscopiques sont utilisées pour la construction ce panorama avec une résolution de 288X567 en utilisant une transformation affine.....	86
Figure 4.14 : 27 images cystoscopiques sont utilisées pour la construction ce panorama avec une résolution de 296X583 en utilisant une transformation projective.....	87
Figure 4.15 : 70 images forment ce panorama de la vessie avec une résolution de 802X517.....	87
Figure 4.16 : Mosaïque formée par 300 images cystoscopiques avec une résolution de 947x428.....	88
Figure 4.17 : Mosaïque formée par 353 images cystoscopiques avec la même résolution de 947x428.....	88
Figure 4.18: Les histogrammes des trois canaux (a: canal rouge, b: canal vert; c: canal bleu) avant correction d'exposition lumineuse et après correction (d: canal rouge, e: canal vert; f: canal bleu).....	89
Figure 4.19 : Correction de l'exposition lumineuse en utilisant l'égalisation adaptative de l'histogramme.....	90

Liste des tableaux

Tableau 2.1 : Tableau comparatif des détecteurs de points d'intérêt selon leur nombres, leur localisations et leur temps de calcul pour l'image damier	42
Tableau 2.2 : Tableau comparatif des détecteurs de points d'intérêt selon leur nombres, leur localisations et leur temps de calcul pour l'image synthétique contenant différents formes de jonctions.....	45
Tableau 3.1: Les valeurs des paramètres d'évaluation du recalage (RMSE, CC et le temps de calcul) pour différents nombres de niveaux de décomposition de l'image Ikonos orientée de 10°.....	60
Tableau 3.2: Les valeurs des paramètres d'évaluation du recalage (RMSE, CC et le temps de calcul) pour différents nombres de directions de la transformée NSCT appliquée à l'image Ikonos orientée de 10°.....	61
Tableau 3.3 : Comparaison des détecteurs SIFT et NSCT avec la méthode proposée NSCT-SIFT selon le pourcentage des faux points détectés.....	65
Tableau 3.4 : Les différents résultats analytiques des méthodes de recalages.....	70
Tableau 3.5 : Les résultats comparatifs entre recalages affines basé sur SIFT et NSCT-SIFT.....	71
Tableau 3.6 :Les résultats comparatifs entre recalages projectifs basé sur SIFT et NSCT-SIFT	71

Introduction générale

Introduction générale

Le mosaïquage d'images est l'assemblage d'images prises suivant des angles de vue différents et ramenées à un même repère. Un panorama est une représentation d'une mosaïque d'images, permettant de voir l'intégralité d'une scène en collant plusieurs images ayant des régions communes.

Louis Lumière dépose le brevet du Photorama en décembre 1900. Son approche a permis de prendre en une seule prise de vue l'élargissement du champ visuel de la scène puis la projection intégrale de ce panorama sur un cylindre de 360°.

Les développements des techniques de visualisation des dernières années et notamment les travaux de Chen [1] et de Szeliski [2] ont permis la création et l'exploitation de panoramas virtuels sur des ordinateurs personnels. Ils existent actuellement, plusieurs logiciels commerciaux permettant de créer facilement un panorama en quelques secondes; on cite, entre autres, le logiciel PhotoStitch. D'autres logiciels sont disponibles sur internet, en version open source, comme Panorama Factory, Panorama Composer, AutoStitch ou PanoramaTools.

Les applications sont diverses; Chen est l'un des premiers à utiliser un panorama cylindrique à 360° pour représenter une scène et permettre la navigation en temps réel dans une application de réalité virtuelle.

On utilise le mosaïquage dans plusieurs domaines: médical, spatial et robotique pour avoir des prises de vue plus large mais il reste toujours le problème des sutures après le collage des différentes parties des images.

Notre travail s'intéresse d'une part à l'application du mosaïquage dans deux grandes spécialités à savoir l'ophtalmologie et la cystoscopie vésicale qui est un examen fondamental en urologie, afin de faciliter le diagnostic et le traitement de certaines pathologies telles que la rétinopathie diabétique et les tumeurs vésicales. D'autre part, il consiste à l'élimination ou la diminution des sutures résultantes de l'assemblage par l'application du *blending* à multi résolution.

Pour illustrer l'importance du mosaïquage dans ces deux disciplines, il fallait décrire l'intérêt de cette technique pour les cliniciens:

- Pour les ophtalmologues, l'inspection de la vascularisation du fond d'œil peut indiquer des signes d'hypertension, de diabète, d'artériosclérose et de maladie cardiovasculaire. La segmentation

des vaisseaux rétiniens est une étape principale vers l'analyse automatisée de la rétine afin de recalculer les images et détecter les pathologies, dans certains travaux de recherche. L'évaluation automatisée de la morphologie vasculaire rétinienne peut être utilisée comme un outil de diagnostic pour le dépistage précoce de la rétinopathie diabétique et les hémorragies. Le mosaïquage des images rétiniennes a un intérêt dans la détection des changements de l'aspect de la rétine, et permet aussi le suivi et le repérage spatial par rapport aux structures anatomiques tel que le disque optique offrant ainsi un grand aide aux ophtalmologistes que ce soit pour le traitement ou le suivi post opératoire.

- Pour les urologues, La cystoscopie est actuellement l'examen clinique de référence qui permet de détecter et de diagnostiquer des lésions cancéreuses, le support photographique et le support vidéo sont communément utilisés. Sur une photographie, seule une petite région d'une zone d'intérêt restera enregistrée. Sur un support vidéo, une région plus grande le sera, mais avec une grande quantité de données redondantes. De plus, la recherche d'une zone particulière d'intérêt sur la vidéo (enregistrée à un instant donné non connu dans la séquence) peut prendre beaucoup de temps pour la retrouver. Donc un support visuel contenant des zones plus larges de la vessie et sans redondance de données permettrait au clinicien d'observer, dans le même coup, des zones de tissu plus étendues. Voir des zones élargies permettrait de faciliter le repérage de lésions satellitaires autour d'une lésion principale. Ce support visuel peut être obtenu par des méthodes de mosaïquage d'images.

Le mémoire est divisé en deux grandes parties: La première partie, englobe les deux premiers chapitres, décrivant l'état de l'art du sujet et le fondement mathématique utilisé; la deuxième partie, couvre les chapitres trois et quatre, est consacré au développement des algorithmes et l'évaluation.

Le mémoire est organisé comme suit:

Le premier chapitre décrit l'intérêt de mosaïquage des images et son domaine d'utilisation surtout en ophtalmologie et en urologie.

Le deuxième chapitre présente une synthèse des différentes méthodes de l'extraction des points d'intérêt des images à mosaïquer.

Le chapitre 3 détaille les principes de recalage et donne un aperçu sur ses différentes transformations utilisées, au même temps, il présente les tests menés pour évaluer et valider notre méthode de mosaïquage basée sur la combinaison NSCT-SIFT. On teste dans un premier temps la précision et la robustesse de la méthode de recalage en présence du bruit. Puis, on évalue la précision de la transformée de recalage calculée à partir des points extraits par notre méthode.

Le chapitre 4 décrit la méthode d'assemblage des images angiographiques et cystoscopiques acquises pendant des examens cliniques standard pour la construction d'une vue panoramique élargie, puis on entame les différentes méthodes pour éliminer les sutures.

On finalise ce document par une conclusion qui met en avant nos contributions et des perspectives.

Chapitre 1

Intérêt du mosaïquage pour les images médicales

1.1. Introduction

L'imagerie médicale grâce à l'apparition des outils d'acquisition et de traitement de l'image de plus en plus évolués, représente un domaine de recherche privilégié des techniques de l'analyse numérique d'images.

L'imagerie médicale permet de représenter visuellement et d'une façon intelligible l'information à caractère médical. Cette problématique s'inscrit plus globalement dans le cadre de l'image scientifique et technique, l'objectif est en effet de pouvoir représenter sous un format relativement simple et compacte une grande quantité d'informations issues d'une multitude de mesures acquises selon un mode bien défini. L'image obtenue peut être traitée informatiquement pour obtenir par exemple [3-a]:

- Une reconstruction tridimensionnelle ou bidimensionnelle d'un organe ou d'un tissu.
- Un film ou une animation montrant l'évolution ou les mouvements d'un organe au cours du temps.
- Une imagerie quantitative qui représente les valeurs mesurées pour certains paramètres biologiques dans un volume donné.

Dans cette thèse, nous nous intéressons en particulier à l'imagerie rétinienne et la cystoscopie vésicale. En effet, avec l'arrivée des systèmes d'angiographie rétinienne numérisée, il est possible de traiter et d'analyser de manière automatique et précise les images du fond de l'œil pour un diagnostic précoce des pathologies rétiniennes. De la même façon, la cystoscopie vésicale sert au dépistage d'une lésion cancéreuse de la vessie. Elle constitue l'examen de référence en utilisant un endoscope souple ou rigide, elle permet de voir la tumeur vésicale, d'en déterminer sa localisation, d'en préciser ses caractères morphologiques et d'explorer le reste de la vessie à la recherche de l'existence d'une ou plusieurs autres tumeurs.

On trouve très peu de travaux sur le mosaïquage de la paroi interne de la vessie. Les travaux de Behrens et al, en 2009 [4] se sont intéressés à la construction d'images panoramiques de fluorescence de la paroi interne de la vessie qui améliore le contraste des structures vasculaires, mais l'examen le plus fréquent appliqué c'est la cystoscopie à lumière blanche et dans ce cas aucune primitive image n'est détectable d'une manière systématique et robuste. Les niveaux de gris des pixels représentent alors la seule information disponible pour le recalage d'images cystoscopiques.

Les travaux de recherche actuels visent à développer des outils d'aide au diagnostic qui soient non invasifs, de lecture rapide, réalisables et interprétables par l'urologue ou l'ophtalmologue. Mais la recherche d'une image spécifique dans un enregistrement au sein d'une séquence complète d'images

(25 images par seconde) peut s'avérer fastidieuse pour chaque patient. Dans le cas de l'angiographie rétinienne on dispose des images fixes avec la même résolution et distance focale, ce qui n'est pas le cas de la cystoscopie où le clinicien réalise l'examen en orientant son appareil et en localisant les zones qu'il observe. Par contre, lorsqu'il regarde une vidéo enregistrée, il n'a pas cette information d'orientation du cystoscope. D'autre part, le champ de vision limité du cystoscope ne permet pas d'estimer de façon précise l'extension d'une lésion cancéreuse ni la distribution spatiale des éventuelles lésions multifocales.

La localisation des lésions sur un schéma réalisé par le clinicien est hautement dépendante de son expérience.

Dans les deux cas, il apparaît qu'un système combinant les informations spatiales d'images globales (images panoramiques) s'avère nécessaire et peut contribuer à améliorer la détection de lésions néoplasiques de la vessie et les pathologies de l'œil. Elle devrait aussi améliorer la surveillance après l'intervention et le suivi d'un traitement.

Dans un premier temps, nous allons nous intéresser à l'œil et ses pathologies et ses examens complémentaires puis on passe à la cystoscopie.

1.2. Anatomie de l'œil

La vision, chez l'homme est assurée par une chaîne complexe dont l'œil est le capteur. Ce capteur perçoit la lumière et la transforme en messages nerveux à destination du cerveau qui en assure le traitement.

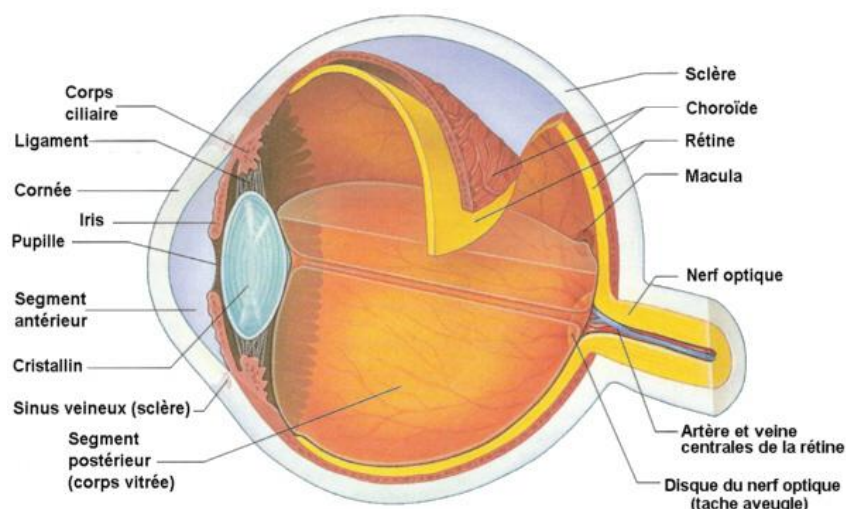


Figure 1.1 : Anatomie de l'œil humain¹

¹Source: <http://pta.nbed.nb.ca/bio/Bio%2053421/Module%205/picts/oeil.jpg> [Anatomie et physiologie humaines](#) (p.548)

L'œil est le moyen de la perception visuelle du monde à l'aide de la partie visible du rayonnement électromagnétique. Elle recouvre l'ensemble des mécanismes physiologiques et psychologiques par lesquels la lumière émise ou réfléchiée par l'environnement détermine les détails des représentations sensorielles, comme les formes, les couleurs, les textures, le mouvement, la distance et le relief [3-b]. L'œil est une extension du cerveau avec lequel il est connecté par environ un million de fibres nerveuses qui sont réunies dans le nerf optique. Il est constitué par les éléments suivants :

- **La cornée**

C'est la partie antérieure transparente du globe oculaire, elle représente l'élément réfractif de l'œil qui sert à transmettre la lumière au cristallin et à la rétine. La cornée est très innervée donc très sensible. Son rôle est de protéger l'avant du globe oculaire.

- **L'iris**

C'est un diaphragme circulaire se réglant automatiquement suivant la quantité de lumière pénétrante dans l'œil. L'iris est responsable de la couleur de l'œil.

- **Le cristallin**

Est responsable de l'accommodation, est une lentille organique avasculaire biconvexe qui permet de focaliser le stimulus grâce à sa capacité à modifier sa courbure sous l'action du muscle ciliaire. Avec l'âge la perte de l'élasticité du cristallin est responsable de la presbytie.

- **Le corps vitré**

C'est un liquide de remplissage. Il représente les 4/5 du volume de l'œil, et est le premier constituant de l'œil. Son rôle est de maintenir la rigidité du globe oculaire, et de maintenir la rétine en place bien collée contre le fond du globe oculaire.

- **La papille**

La papille est une zone circulaire correspondant à la naissance du nerf optique, elle se situe sur la rétine, à l'endroit où se réunissent les fibres optiques provenant des cellules ganglionnaires nerveuses de la rétine. La papille mesure environ 1,5 mm de diamètre, et apparaît très illuminée sur les clichés de couleur.

- **La rétine**

Est la couche sensorielle de l'œil qui permet la vision. Cette zone tapisse le fond de l'œil et est parcourue par des vaisseaux sanguins qui émergent au niveau de la papille optique où l'on distingue l'artère et la veine centrale qui se divisent elle-même en artères et veines de diamètre plus faible pour vasculariser les cellules qui permettent la vision. La rétine est le lieu de traduction du message lumineux en signaux nerveux. Il s'agit d'un tissu neuronal très

fin, de 100 à 500 μm d'épaisseur, organisé en plusieurs couches de cellules. Elle peut être considérée comme un prolongement direct de cerveau dans l'œil. La rétine contient deux types de cellules photosensibles, les cônes et les bâtonnets (voir Fig 1.2.b) qui n'ont pas la même sensibilité spectrale :

- Les bâtonnets permettent la vision nocturne.
- les cônes permettent la vision diurne.

Après avoir traversé les différents éléments cités ci-dessus, le stimulus lumineux arrive sur une zone photosensible localisée au fond de l'œil, la rétine où viennent se projeter les images de la scène que nous observons.

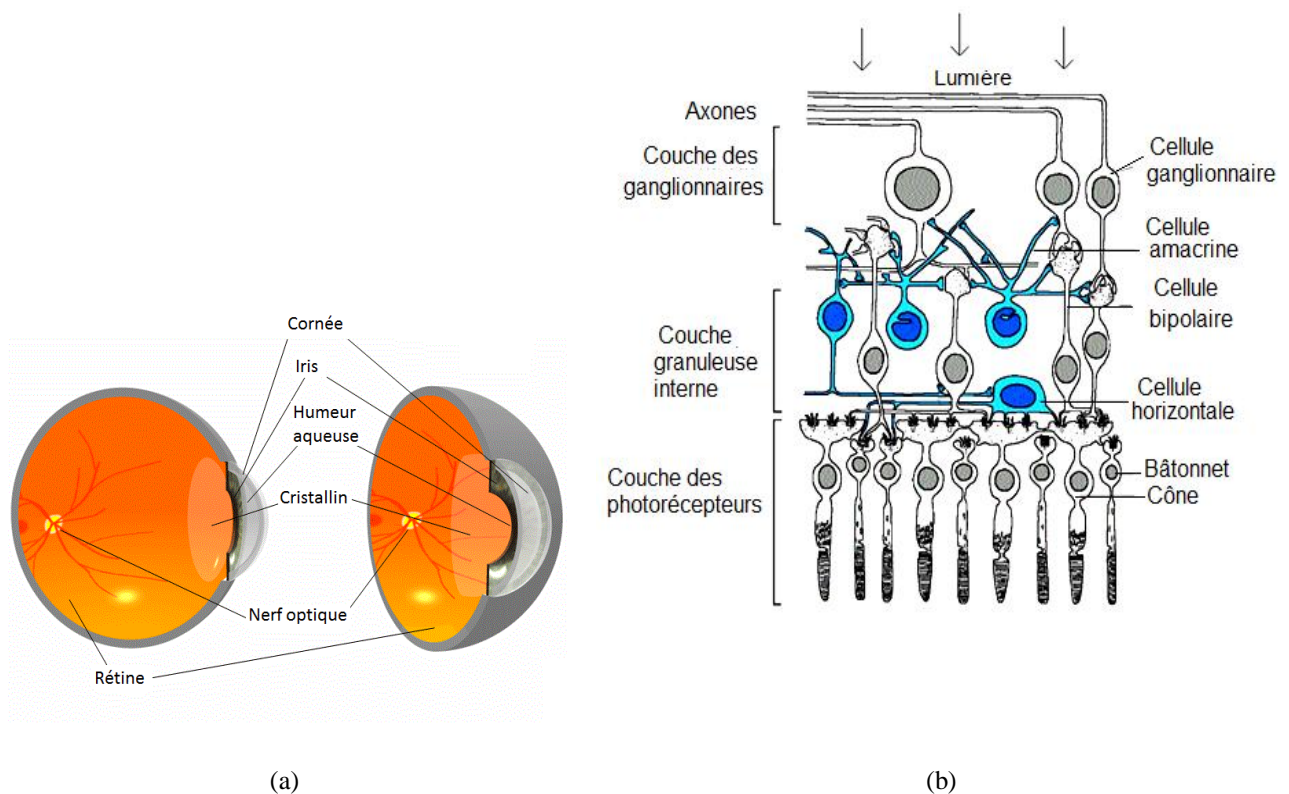


Figure 1.2 : (a) coupe de l'œil² ; (b) Schéma simplifié de l'architecture de la rétine³ .

- La macula

La macula appelée aussi "tache jaune" est le centre de la rétine située au fond de l'œil dépourvue de capillaires sanguins. La macula est uniquement composée de cônes qui directement exposés à la lumière permettent la précision de vision, et contribuent à la qualité de vision au centre du champ visuel.

On trouve au centre de la macula une petite zone de 1,5 mm de diamètre, appelée la

² Source: <https://www.gatinel.com/wp-content/uploads/2009/12/vue-3d-de-loeil.gif>

³ Source: <http://www.bioinformatics.org/oeil-couleur/dossier/images/retine.png>

fovéa, de forme conique et creuse.

1.3. Les examens complémentaires en ophtalmologie

Les examens complémentaires de l'œil permettent d'obtenir des informations anatomiques et ou les pathologies affectant l'œil.

Pour le diagnostic des différentes pathologies rétinienne on peut citer les clichés suivants:

1.3.1. Echographie oculaire

Est une technique d'imagerie employant des ultrasons. Elle permet l'exploration de l'œil même si celui-ci a perdu sa transparence en cas d'hémorragie intraoculaire, l'examen se fait à travers un gel qui est déposé sur les paupières. Elle est généralement indiquée dans les baisses d'acuité visuelle (74 % des cas). Elle peut détecter certaines pathologies comme le décollement rétinien, hémorragie du Vitré [3] (voir figure 1.3) et des rétinoblastomes.

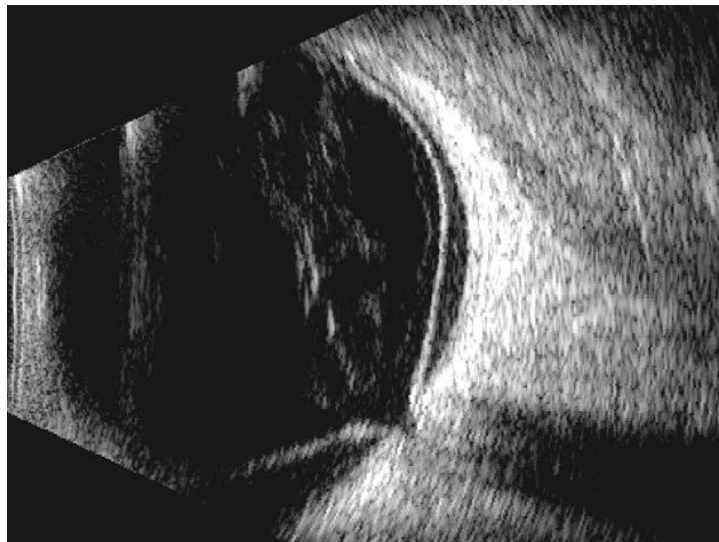


Figure 1.3 : Echographie oculaire⁴

Montrant un décollement de rétine avec hémorragie intravitréenne et décollement postérieur du vitré. Le décollement de rétine apparaît comme une membrane très écho-gène, partant de la papille et doublant la paroi oculaire.

⁴ Source: "Échographie de l'œil et de l'orbite" EMC-Ophtalmologie 2 page 17, 2005

1.3.2. L'Angiographie à la fluorescéine rétinienne

L'angiographie à la fluorescéine représente l'examen de référence pour les pathologies rétiniennes (voir Figure.1.4). Cet examen consiste à prendre des photographies de fond d'œil après avoir injecté un colorant fluorescent dans une veine. Le colorant est transporté par le courant sanguin dans le réseau artériel et veineux de la rétine. C'est un moyen irremplaçable pour connaître l'état des vaisseaux de fond d'œil et principalement de la rétine [6]. L'angiographie est indiquée pour :

- Établir le diagnostic de certaines maladies de l'œil.
- Guider un traitement notamment quand une photo coagulation au laser est nécessaire.
- Ainsi, les indications les plus fréquentes de l'angiographie concernent la rétinopathie diabétique et les dégénérescences de la rétine.

L'Angiographie à la fluorescéine consiste, dans un premier temps, en la réalisation de clichés de la rétine monochromatiques, en couleur verte, rouge et bleue.

- Le cliché en lumière verte permet de visualiser au mieux les vaisseaux rétiniens.
- Le cliché en lumière bleue permet de visualiser les couches les plus superficielles de la rétine et l'aire maculaire.
- Le cliché en lumière rouge donne accès aux couches plus profondes de la rétine et permet de visualiser les vaisseaux choroïdiens.

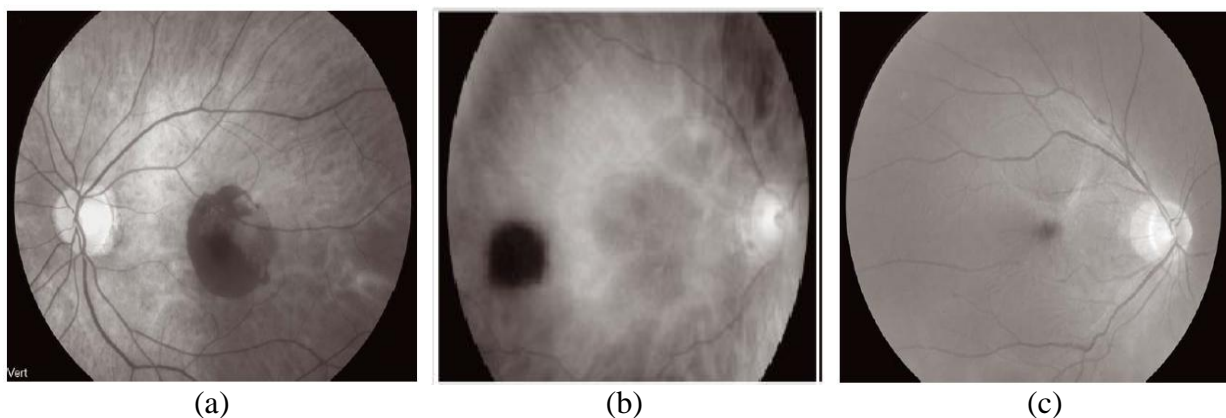


Figure 1.4 : Angiographie à la fluorescéine rétinienne en trois couleurs⁵

⁵ Source: "Rétinophotographies, clichés en autofluorescence et angiographies: technique et interprétation" page 2-3 EMC Ophtalmologie Elsevier Masson 2011.

- (a) Cliché en lumière verte montrant une hémorragie sous-rétinienne maculaire avec une bonne visibilité des vaisseaux rétiniens.
- (b) Cliché en lumière rouge. Nævus choroïdien temporo-maculaire avec des limites bien visualisées.
- (c) Cliché en lumière bleue. Membrane épimaculaire d'aspect blanchâtre.

Le colorant injecté est éliminé au cours des 24 à 48 heures qui suivent l'examen. La numérisation rend les résultats accessibles de façon instantanée et supprime le délai d'attente pour le développement des photos permettant ainsi, d'adapter le traitement plus rapidement. Cet examen est indiqué pour déceler certaines pathologies comme la dégénérescence maculaire liée à l'âge (DMLA), la rétinopathie diabétique et autres maladies de la rétine et de la choroïde.

En deuxième étape, des clichés sont réalisés après l'injection de la fluorescéine 10% en injection intraveineuse.

On distingue classiquement différents clichés de la séquence angiographique reliés au temps artériel et veineux [7]:

- Stade Précoce : 0mn 15 secs à 0mn 40 secs ;
- Stade Intermédiaire : 1mn 30 secs à 2mn 30 secs ;
- Stade Tardif : 4mn 00 secs à 5mn 00 secs.

1.3.2.1. Interprétation de l'angiographie à la fluorescéine

Avant toute interprétation de l'angiographie à la fluorescéine, il faut tout d'abord savoir lire un cliché normal de la rétine.

- Angiographie normale

Deux régions du fond d'œil présentent des particularités : la macula et la tête du nerf optique. La fovéola est le centre de la macula. Cet endroit est avasculaire et paraît sombre, couleur accentué par le pigment xanthophylle.

La tête du nerf optique, ou papille optique, est irriguée par deux systèmes circulatoires différents : le système rétinien et le système des artères ciliaires postérieures. Différentes anastomoses existent entre ces deux systèmes [8-a] voir figure 1.5

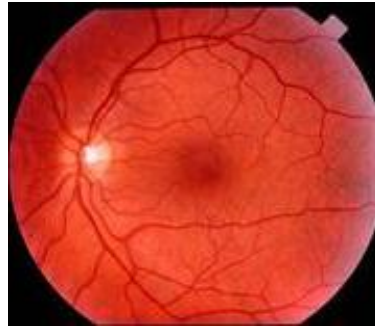


Figure 1.5 : Angiographie à la fluorescéine normale⁶

- Angiographie pathologique

En cas de pathologie, la fluorescéine va diffuser dans certains espaces anormaux et être responsable d'hypo- ou d'hyperfluorescence anormale. En fonction du siège de cette diffusion et de son aspect, on va pouvoir déterminer son origine et ainsi établir le diagnostic de la pathologie responsable [8-b], cette section sera détaillée lorsqu'on va aborder les pathologies les plus sévères de la rétine.

1.3.3. Angiographie numérisée ICG⁷

Il s'agit de la même technique que pour l'angiographie en fluorescence, mais le colorant utilisé est le vert d'indocyanine, les photos étant prises avec une caméra infrarouge pour indiquer la DMLA, certaines maladies de la rétine et de la choroïde. Elles permettent de percevoir des couches plus profondes.

1.3.4. Tomographie en cohérence optique OCT⁸

C'est un examen non invasif, sans contact, sans injection et de haute résolution qui dure quelques minutes et donne une vision bi ou tridimensionnelle des structures rétinienne, en utilisant un faisceau laser infrarouge de faible puissance. L'examen donne, en temps réel, une image microscopique en coupe des couches profondes de la rétine. Il permet de visualiser et individualiser, en temps réel, les différentes couches de la chorio-rétine, un œdème maculaire, des kystes sous maculaires, un décollement séreux de l'épithélium pigmentaire, des drusens. Il permet à l'ophtalmologiste de mesurer l'épaississement rétinien dû à l'œdème et de mieux adapter le suivi et le traitement [9-a].

⁶ Source: Service d'ophtalmologie CHU de sidi Bel Abbes

⁷ Indocyanine green

⁸ Optical coherence tomography.

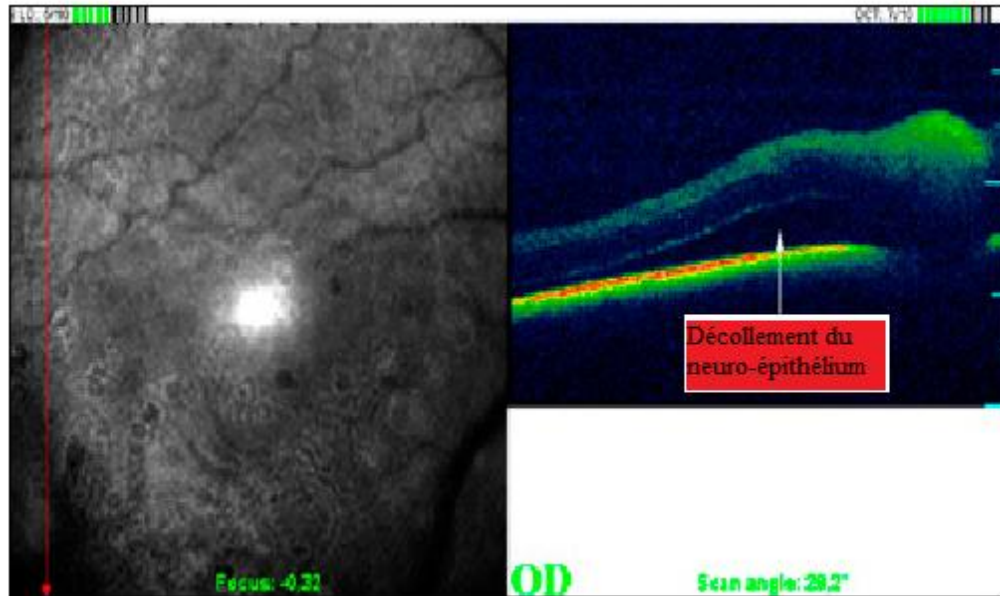


Figure 1.6 : Examen OCT de la rétine⁹

L'OCT, en complément de la fluoro-angiographie permet d'établir un diagnostic lésionnel précoce et la mise en place d'un traitement à la photo-coagulation au laser infra rouge.

1.4. Les pathologies de la rétine

L'œil humaine peut être affecté par plusieurs pathologies provoquant des troubles de la vision. On peut citer en particulier la rétinopathie diabétique et la DMLA qui sont les principales causes de malvoyance dans notre pays. De ce fait un diagnostic précoce de ces deux pathologies s'avère très important et se base principalement sur l'analyse des différents types de clichés de la rétine à la recherche des modifications des structures anatomiques.

1.4.1. Rétinopathie diabétique

La rétinopathie diabétique est la complication la plus grave du diabète. Elle est, le plus souvent, asymptomatique avant l'apparition d'une gêne visuelle centrale ou périphérique.

La collaboration entre médecin généraliste et ophtalmologiste s'avère nécessaire pour éviter la cécité, le généraliste essaye de préserver un équilibre parfait et permanent de la glycémie de son patient, tandis que l'ophtalmologiste se charge de surveiller l'évolution de la maladie en étudiant les différents clichés de l'angiographie rétinienne, afin d'intervenir le plus vite que possible. Cette affection se caractérise par l'atteinte des capillaires chorio-rétiens et l'apparition des œdèmes rétinien le plus souvent maculaire qui va être responsable de baisse de vision. L'autre

⁹ Source: https://www.chvsm.com/images/contenu/ophtalmologie/oct-slo_5.jpg

caractéristique de la maladie est la survenue des occlusions des capillaires entraînant une ischémie rétinienne responsable de l'apparition des néovaisseaux extrêmement fragiles qui auront tendance à avancer dans le vitré et à saigner. On distingue trois formes successives en fonction des lésions : non proliférante, préproliférante et proliférante. Toutes trois sont divisées selon l'angiographie en : minime, modérée et sévère [9-b].

- **La rétinopathie non proliférante** : Est le type le moins grave à affecter la vue, mais il doit quand même être régulièrement surveillé par l'ophtalmologiste.
- **La rétinopathie préproliférante** : Entraîne le gonflement de la rétine et des hémorragies, peuvent apparaître et commencer à obscurcir la vue.
- **La rétinopathie proliférante** : Se produit lorsque des hémorragies importantes apparaissent et provoquent une perte soudaine de la vision, le patient commence à voir flou en raison de ces hémorragies.

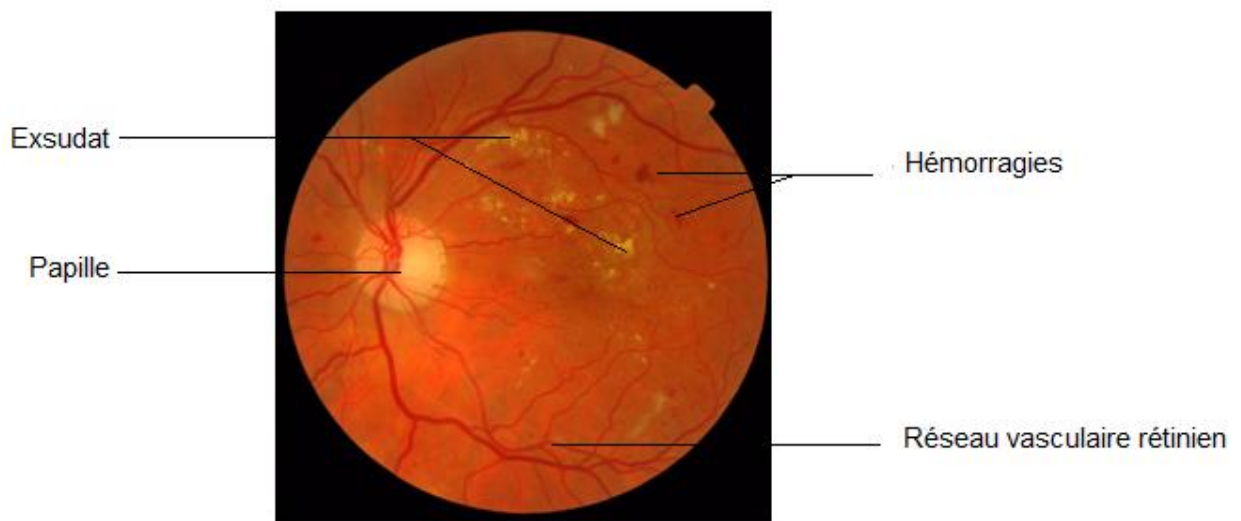


Figure 1.7 : Rétinopathie diabétique¹⁰

1.4.2. La Dégénérescence Maculaire Liée à l'Age DMLA

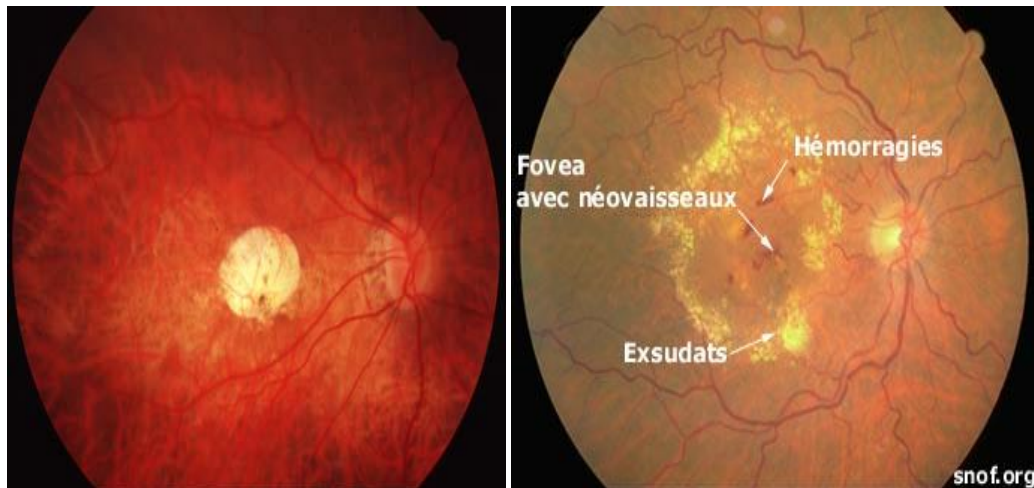
Elle débute de plus en plus tôt, parfois vers 50 ans ; l'évolution sans traitement se fait vers la cécité. La dégénérescence maculaire du sujet âgé est une affection binoculaire, les signes prémonitoires se traduisent par l'apparition de petites taches jaunâtres ou "drusens" pérимaculaires, qui apparaissent petit à petit au pôle postérieur de la rétine. La maladie peut rester silencieuse longtemps; c'est à ce

¹⁰ Source: Service d'ophtalmologie CHU de sidi Bel Abbes

stade silencieux qu'il convient de la découvrir afin de surveiller et traiter dès les premiers signes de menace ou de souffrance maculaire.

Par la suite, peuvent apparaître les trois formes de la maladie qui sont [9-c]:

- La première est celle qui présente des drusens péri-maculaires ou juxtamaculaires, sans baisse visuelle et d'évolution lente. Le test d'Amsler en autocontrôle s'avère nécessaire et important.
- La deuxième est la forme dite sèche, atrophique. Cette forme est aussi d'évolution lente; il n'y a aucun traitement actuellement. Le médecin peut prescrire des vasodilatateurs, et des suppléments alimentaires tels que la carotène et des vitamines C et E, omega 3.
- La troisième, la forme œdémateuse ou humide est la plus grave, la plus rapidement évolutive, mais la seule qui peut être traitée et stoppée. Le premier signe, c'est la visibilité de "drusens" et l'existence d'hémorragies centrales à l'angiographie fluorescéinique à ce stade, il faut traiter en urgence.



a) DMLA sèche ou atrophique¹¹

b) DMLA humide ou exsudative¹²

Figure 1.8 : Dégénérescence Maculaire Liée à l'Age.

1.4.3. Importance du test d'Amsler pour le diagnostic de la DLMA

Le test d'Amsler permet de mettre en évidence les déformations et parfois un scotome central débutant. C'est un signe certain de l'atteinte de la macula qui matérialise le déficit central et révèle la tache ou le flou central ou les déformations des lignes caractéristiques des

¹¹ Source: <http://www.snof.org/images/encyclopedie/dmlaexsud.jpg>

¹² Source: <http://www.snof.org/images/encyclopedie/dmlaatrodrusen.jpg>

métamorphoses qui deviennent ondulantes en cas de DMLA. Ce test peut être effectué par le patient pour autocontrôle de son champ visuel central, à domicile, au cas où l'ophtalmologiste aurait détecté la présence de drusens paracentraux de la macula.

Le test se fait en fixant le point central noir, en monoculaire, à une distance de 30 cm, avec une éventuelle correction, le patient doit voir nettes toutes les lignes des carreaux, y compris les angles du test. S'il n'y parvient pas, on lui demande de dessiner, tout en fixant le centre, la tache qui cache un ou plusieurs des carreaux ou les déformations des lignes voir figure 1.9.

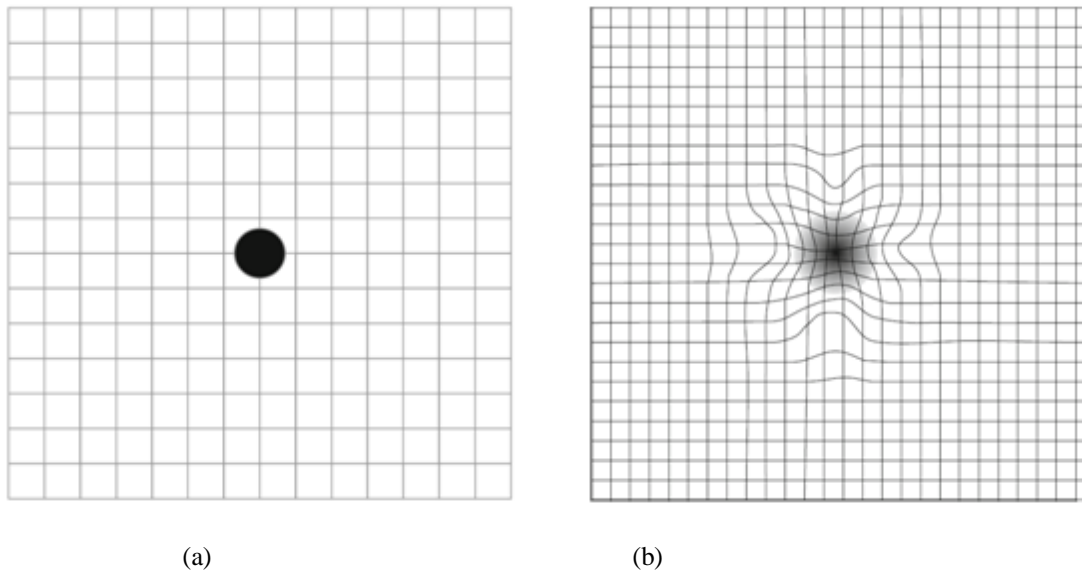


Figure 1.9 : Test d'Amsler¹³ : (a) test de DMLA négatif (vision normale), (b) test de DMLA positif (vision anormale).

1.5. L'endoscopie vésicale et le cancer de la vessie

Le cancer de la vessie est le 7^{ème} cancer le plus répandu dans le monde, soit 3,2% de l'ensemble des cancers, il est diagnostiqué ou traité dans le monde chez 2,7 millions de personnes chaque année [10]. En Algérie, selon le CIRC (Centre International de Recherche sur le Cancer), le cancer de la vessie représente 8,4% des cancers masculins et occupe par conséquent le troisième rang chez l'homme. Il reste rare chez la femme et représente 1,2% seulement de l'ensemble des cancers féminins.

Par ailleurs, l'âge moyen de survenue est de 60 à 68 ans.

L'endoscopie vésicale représente l'examen clé pour le dépistage et le diagnostic du cancer vésicale, elle permet de fournir à l'urologue des images sous forme de vidéo nécessitant un temps énorme

¹³ Source: Marie-José bot-escluse "Abord clinique en ophtalmologie" pp 31 et 32 Springer-Verlag France, Paris 2013.

pour bien les exploiter et trouver la lésion que le médecin cherche, pour faciliter cette tâche l'assemblage de ces images se trouve nécessaire permettant ainsi au clinicien de gagner du temps et de se repérer parmi les structures anatomiques de la vessie, qu'on va essayer de les détailler dans le paragraphe suivant.

1.5.1. Caractéristiques des tissus urothéliaux

La vessie est un réservoir naturel de forme sphérique dans lequel s'accumule l'urine produite par les reins (figure 1.10(a)). Elle a une paroi qui mesure 0,5 cm et qui est constituée de deux tuniques: Une externe c'est le tissu musculaire qui se contracte pour évacuer l'urine; l'autre interne, qui est en contact direct avec l'urine est la muqueuse. Du point de vue histologique, la paroi de la vessie est constituée de plusieurs couches. Une couche de cellules urothéliales recouvre l'intérieur de la vessie appelée urothélium transitionnel reposant sur une membrane basale. Sous l'urothélium se trouve une fine couche de tissu conjonctif appelée chorion. La couche suivante, plus profonde, est une zone de tissu musculaire appelée musculeuse. Enfin, une couche de tissu adipeux sépare la vessie des autres organes. La figure 1.10 (b) représente les différentes couches constituant la paroi de la vessie.

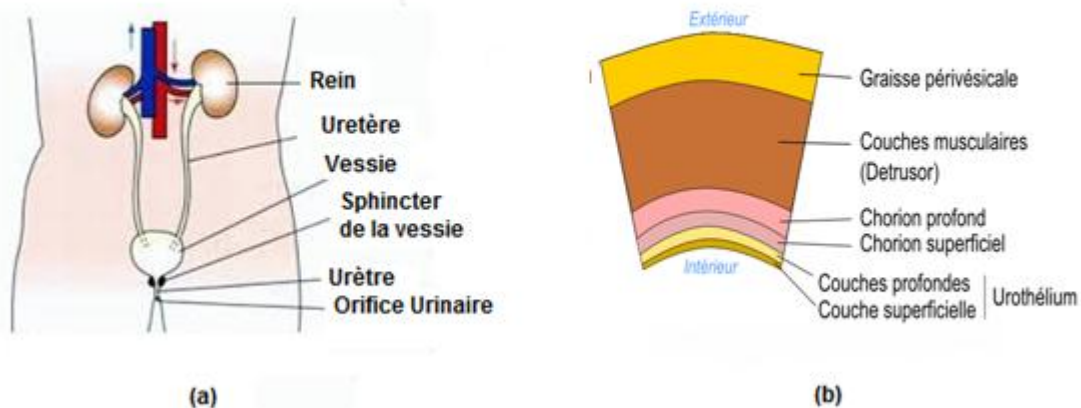


Figure 1.10 : (a) L'appareil urinaire. (b) Représentation des couches constituant la paroi de la vessie¹⁴.

1.5.2. Types et classification des cancers de la vessie

Histologiquement, deux types de tumeurs peuvent être distingués: Les tumeurs papillaires et les lésions planes.

Les lésions superficielles, papillomes

¹⁴ Source: http://urovar.fr/_medias/images/vessie%20nm.png

- Non infiltrantes: Ces tumeurs n'infiltrent pas la lamina propria (tissu conjonctif lâche situé sous les épithéliums qui tapissent notamment la muqueuse vésicale), laissant la membrane basale intacte.
- Infiltrantes: Leurs extensions s'enfoncent plus ou moins profondément dans la paroi de la vessie, jusqu'à la couche musculuse (qui contient les fibres musculaires de la vessie) et au delà. L'appréciation du degré d'extension permet de déterminer le stade (figure 1.11).

Les lésions tumorales « planes »

Carcinome in-situ (CIS) est une prolifération néoplasique strictement limitée au revêtement urothélial. Il peut être isolé ou plus fréquemment associé à une véritable tumeur vésicale [11]. Le CIS est une tumeur plate, non papillaire, située dans l'épaisseur même de la muqueuse. Le CIS est un cancer épithélial résistant et agressif avec un potentiel évolutif. Il peut être multifocal et diffus. Le risque de la récurrence et de la progression reste élevé chez les patients avec des lésions de CIS dû au faux diagnostic du CIS qui peut passer inaperçu à l'examen cystoscopique habituel.

Le T de la classification TNM, Tumor Node Metastasis en anglais, classe les tumeurs par rapport à leur degré de pénétration dans la paroi de la vessie (figure 1.11).

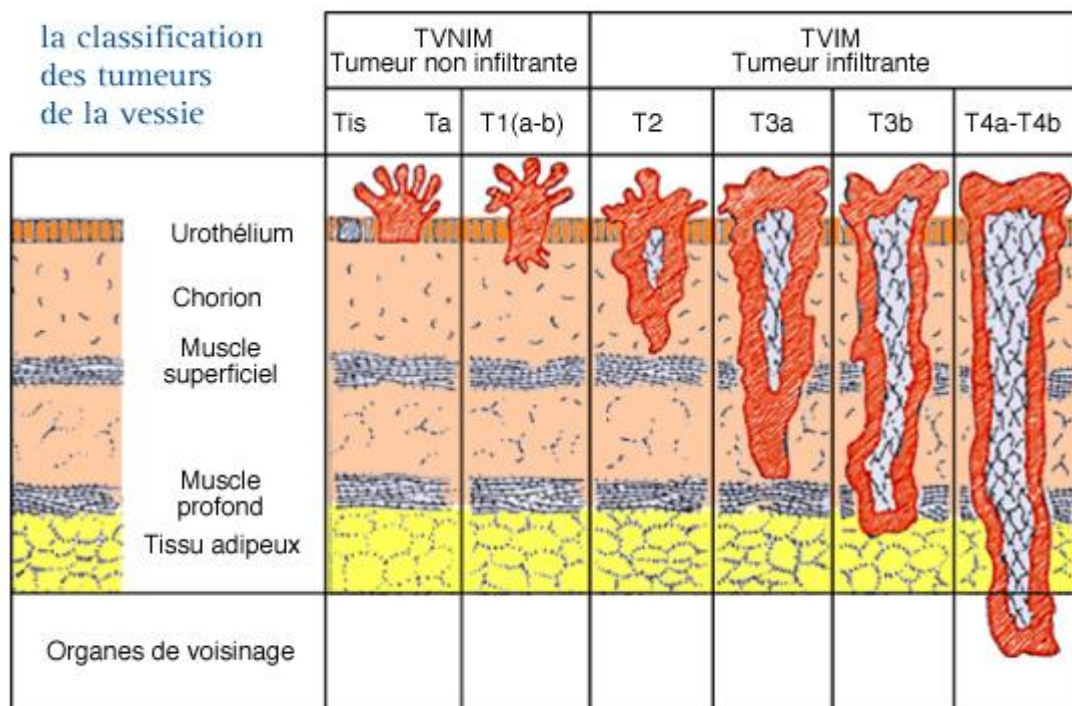


Figure 1.11 : Différents degrés d'invasion d'une tumeur de vessie¹⁵.

Ta Tumeur papillaire non-invasive (confinée à la muqueuse).

¹⁵ Source: http://www.arcagy.org/infocancer/img/84_popup_types-de-tumeurs-vessie.jpg

Tis Carcinoma in situ (Tumeur plane résistante, confinée à l'urothélium).

T1 La tumeur franchit la lamina propria.

T2 La tumeur envahit le muscle de la vessie.

T3 La tumeur envahit le tissu adipeux périvésical.

T4 La tumeur envahit un organe voisin (prostate, utérus, vagin, paroi pelvienne ou la paroi abdominale).

1.6. L'endoscopie vésicale et les problèmes scientifiques

Dans cette section nous présentons les difficultés rencontrées lors de la cystoscopie pour parvenir à la reconstruction d'une image mosaïquée avec un champ visuel élargi voir figure 1.12.

1.6.1. La qualité des images cystoscopiques

Les caractéristiques des images cystoscopiques dépendent de l'instrument utilisé pour acquérir les images (le cystoscope) et de la qualité de l'examen cystoscopique réalisé par le clinicien.

Effets produits par l'optique du cystoscope

Le cystoscope est un instrument fin qui, inséré au travers de l'urètre, permet d'observer les parois internes de la vessie. Il existe deux types de cystoscopes qui peuvent être utilisés pour un examen clinique : les cystoscopes rigides et les souples.

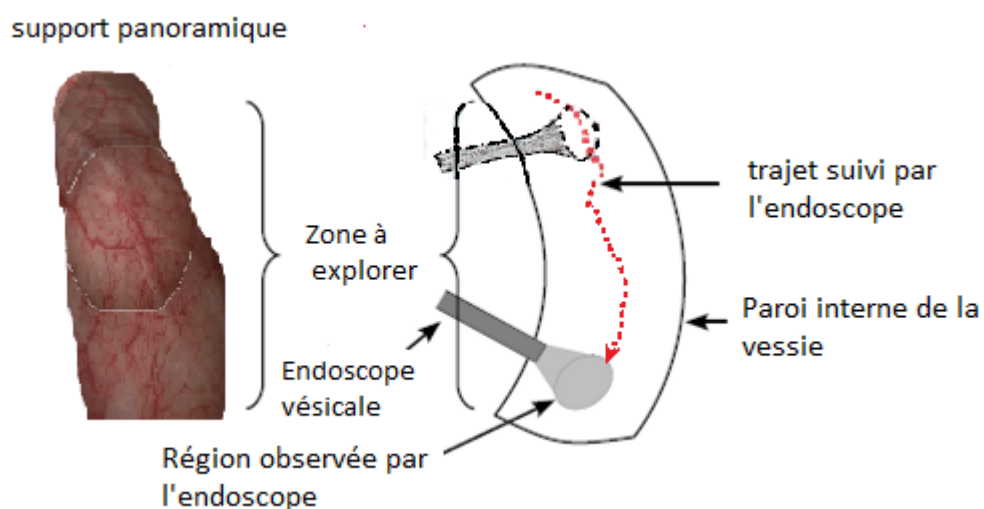


Figure 1.12 : Schéma illustrant le trajet de l'endoscope pour la construction d'images panoramiques de l'intérieur de la vessie¹⁶.

¹⁶ Source: http://www.arcagy.org/infocancer/img/84_popup_types-de-tumeurs-vessie.jpg

Le cystoscope rigide est un dispositif composé d'un fin cathéter métallique creux, rectiligne. L'extrémité proximale comporte deux robinets l'un pour l'arrivée du liquide d'irrigation, l'autre pour la sortie.

L'introduction de la chemise est facilitée par l'utilisation d'un mandrin interne. Lors du bon positionnement dans la vessie, il est remplacé par le système optique qui transmet les images reçues par l'objectif situé dans la partie distale de l'instrument. Le diamètre externe de la gaine varie entre 15 et 25 Ch (1 unité Charrière = 1/3 mm).

Le cystoscope souple Contrairement au cystoscope rigide, le fibroscope est un appareil monobloc constitué de deux parties une gaine principale et une poignée de commande. C'est un dispositif qui utilise de longs faisceaux de fibres optiques pour transmettre la lumière et les images de la scène observée. La figure 1.13 montre des exemples des ces dispositifs [12].



Figure 1.13 : Les deux types de cystoscope utilisés en clinique¹⁷.

(a) Cystoscope rigide Karl Storz 27005BA. (b) Cystoscope souple Olympus EndoEYE

Variation de la qualité des images acquises par le cystoscope

certaines contraintes et conditions influent sur la qualité d'image obtenue parmi elles , on peut noter le degré de l'illumination, la distance du cystoscope par rapport à la paroi de la vessie qui n'est pas fixe ce qui donne des images avec différentes résolutions, les mouvements brusques de l'urologue lors de l'exploration qui peuvent engendrer un flou sur les images, les courbures de la vessie dûes à sa morphologie anatomique peuvent donner des ombres de relief sur certaines

¹⁷ Source: "Cystoscopie" Progrès en urologie vol 20, page 823, 2010

images et finalement les bulles d'air qui peuvent être un obstacle entre le cystoscope et la paroi à explorer. La figure (1.14) donne des exemples d'images montrant ces contraintes.

Du fait de cette variabilité problématique, quelques séquences d'images peuvent avoir des structures très visibles tandis que d'autres images en auront très peu, donc pour construire notre image panoramique et tester l'efficacité de l'extraction des points de contrôle selon l'algorithme proposé en chapitre 3, on a du faire un choix subjectif en prenant que les images qui ont les mêmes caractéristiques tel que la distance du cystoscope par rapport à la paroi interne de la vessie, en tenant compte de certaines critères qui seront abordés dans le paragraphe suivant

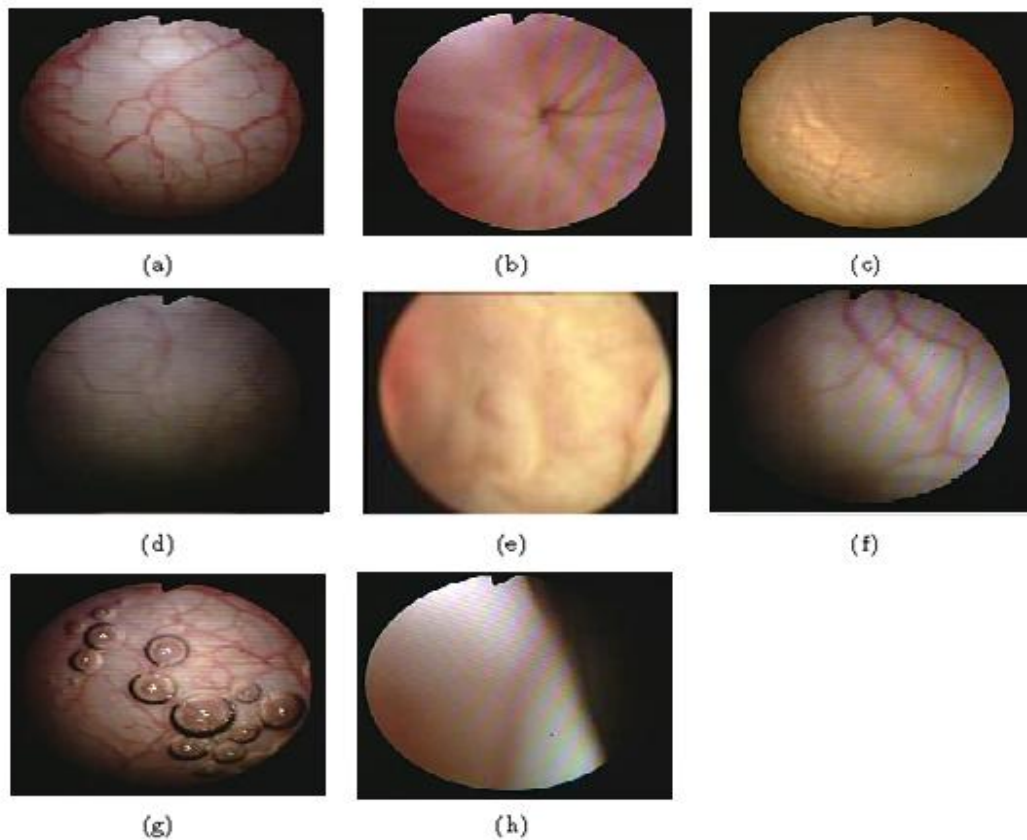


Figure 1.14 : Quelques exemples d'images pouvant être extraites d'une séquence vidéo cystoscopique¹⁸.

- (a) Image avec une grande quantité de vaisseaux visibles. (b) Cystoscope qui passe près de l'orifice de la vessie. (c) Image floue dû au mouvement rapide du cystoscope.
- (d) Image sombre avec faible illumination. (e) Image avec très peu de structures visibles. (f) Image avec cystoscope à courte distance focale et avec différence d'illumination. (g) Image présentant des bulles d'air. (h) Image avec une très faible illumination.

¹⁸ Source: support vidéo extrait de www.drbadarn.com

1.6.2. Contraintes cliniques influençant le mosaïquage des images cystoscopiques

Informations disponibles

La localisation exacte du cystoscope à l'intérieur de la vessie est inconnue et il n'y a aucune technique utilisant le cystoscope rigide ou souple traditionnel permettant de situer le cystoscope lors de l'exploration clinique. Même les repères anatomiques tels que les orifices des uretères ou l'urètre ne sont pas la plupart du temps visible pour guider la construction d'images panoramiques. La seule information disponible pour construire les images panoramiques sont les images acquises par le cystoscope enregistré sur le support vidéo.

Zone de chevauchement entre les images cystoscopiques

Pour construire notre image mosaïquée, il faut que les images extraites à partir du support vidéo, possèdent une zone d'intersection commune entre elles. La vitesse des mouvements du cystoscope joue un rôle important pour que cette zone soit assez grande lors de l'extraction des images cystoscopiques du support vidéo, car on doit tenir compte que ces images sont acquises à une vitesse de 25 images par seconde. Une vitesse de déplacement du cystoscope maximale peut générer une zone de chevauchement minimale ou aucune zone entre les images endoscopiques, ce qui rend leurs assemblage impossible, d'un autre côté une vitesse de déplacement du cystoscope lente peut générer des images semblables, donc la seule solution envisagée pour l'assemblage des images est le choix subjectif en gardant que les images présentant des changements d'informations de la partie de la vessie explorée avec zone d'intersection acceptable.

Qualité de la construction des images panoramiques

Le support visuel proposé doit être de bonne qualité pour fournir à l'urologue les informations nécessaires lui permettant d'effectuer des observations post opératoires, et le suivi de l'évolution de la pathologie en faisant des comparaisons entre différents examens.

1.7. Conclusion

Dans ce chapitre, nous avons présenté l'intérêt des images en panorama pour l'exploration des images médicales que ce soit celles issues de l'angiographie ou de la cystoscopie. Nous avons également présenté les contraintes auxquelles est soumise une telle représentation.

La représentation en couleur des images panoramiques doit être privilégiée et les artefacts introduits par les différences d'illumination entre les bords des images doivent être minimisés. Les images panoramiques devront avoir la même résolution que les images originales pour conserver le niveau de détail.

Le support panoramique unique permet d'observer une zone de tissu plus étendue qui pourrait aider à localiser des zones de pathologies différentes en une seule observation ce qui permet au clinicien de gagner beaucoup de temps pour le diagnostic et le suivi des maladies.

Chapitre 2

Extraction des points d'intérêt

2.1. Introduction

L'extraction des informations caractéristiques contenues dans une ou plusieurs images, est un problème majeur en traitement et analyse d'image et en vision par ordinateur. La représentation que notre œil se fait d'une droite ou d'un coin n'est pas nécessairement facile à interpréter du point de vue du signal. Ainsi plusieurs solutions ont été proposées pour extraire des informations ou les formes faciles à interpréter qui ne peuvent être distinguées par l'œil. Alors, chercher à extraire de l'image les primitives points encore appelés, points d'intérêt ou points de contrôle s'avèrent les plus pertinents. Ces points d'intérêt permettent de caractériser les endroits où le signal est riche en information, et ainsi de réduire la description de l'image en un nombre de points bien défini.

Un point d'intérêt est un point d'une image qui diffère de ses voisins immédiats, en terme de certaines propriétés spécifiques de l'image telles que l'intensité, la couleur, l'orientation, la texture, la courbure, etc... La détection de points d'intérêt est l'un des thèmes les plus importants du domaine de la vision par ordinateur, car ces points constituent des caractéristiques utiles dans les problèmes d'appariement d'images, de reconnaissance d'objets [13].

Ils sont également utilisés dans différentes applications : la robotique, l'indexation ou la reconnaissance d'objets (on peut extraire d'une image une suite de points caractéristiques d'un objet afin de pouvoir l'indexer dans une base de données), le suivi (ou tracking des objets dans une séquence d'images), ou encore l'imagerie médicale où les points d'intérêt correspondent à des points anatomiques particuliers. Pour cette raison, de nombreux travaux ont été dédiés à la détection de différents types de points d'intérêt. Les plus utilisés de ces détecteurs, on peut citer le détecteur de Moravec, Harris, SIFT (*Scale Invariant Feature Transform*), SURF (*Speeded-Up Robust Features Principe*), Tuytelaars et al [14] et le détecteur basé sur NSCT (NonsubSampled Contourlet Transform), sont des détecteurs très populaires.

Dans ce chapitre nous allons définir les détecteurs des points d'intérêt puis étudier les différentes approches pour extraire ces points. Ensuite une étude comparative entre certains détecteurs sur des images test sera faite.

2.2. Définition de point d'intérêt

Les points d'intérêt correspondent à des changements bidimensionnels significatifs de l'intensité se produisant dans plus d'une seule orientation prédominante généralement ce sont des discontinuités des niveaux de gris comme qui peuvent être provoqués selon L.Rosenthaler et F. Heitger [15] par:

(a) une perte de contraste (*e.g.* fin de ligne), (b) intersection de deux ou plusieurs bords d'orientations différentes (*e.g.* jonction, vertex), (c) changements continus d'orientation (courbure). La figure 2.1 montre quelques variétés de ces points d'intérêt, sous forme de coins (jonctions en L), les jonctions en X, V, T et les jonctions flèches et damier. Ils peuvent également apparaître lors d'une modification de la structure, de la texture ou de la géométrie de l'image.

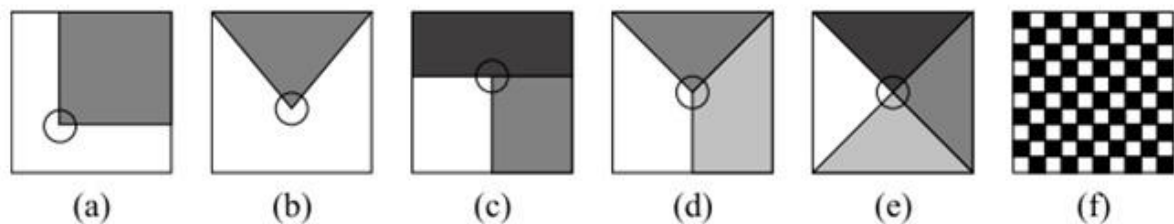


Figure 2.1 : Différents type de coin¹⁹ : (a) jonction en "L", (b) jonction en "V", (c) jonction en "T", (d) jonction en "Y", (e) jonction en "X", (f) jonction en "damier".

2.3. Méthodes de détection des points d'intérêt

Il existe plusieurs types de détecteurs de points d'intérêt qui sont décrits dans la littérature par exemple les détecteurs de coins et de jonctions : Moravec [16], Harris [17], Smith et al [18], les détecteurs basés sur les ondelettes Wu [19], les détecteurs basés sur la différence gaussienne DOG Lowe [20], les détecteurs couleurs Gouet [21].

Le choix d'un détecteur des points d'intérêt repose essentiellement sur l'utilisation souhaitée. On va essayer de décrire les plus répandus et les évaluer par de images tests afin de juger leurs efficacités et leurs précisions et leurs fiabilités, on distingue trois catégories de détecteurs :

- Les détecteurs utilisant des modèles paramétriques.
- Les détecteurs s'appuyant sur une détection de contours.
- Les détecteurs exploitant les niveaux de gris.

¹⁹ Source: Thèse Descripteurs 2D et 2D+t de points d'intérêt pour des appariements robustes page 7.

2.3.1. Méthodes à base de modèles

Cette méthode sert à une identification des points d'intérêt dans l'image en approximant le signal observé (signal image) par un modèle théorique (mathématique) des points d'intérêt considérés. Cette catégorie n'est utilisable que pour des types précis de points d'intérêt comme les coins. Par exemple les jonctions en «L». De plus, la performance de ce type de détecteurs n'est pas satisfaisante en raison du coût élevé de calcul. On peut citer quelques détecteurs de cette approche par exemple : Rohr [22] [23] qui utilise un modèle paramétrique permettant de différencier les jonctions de plusieurs lignes.

2.3.2. Méthodes à base géométrique

Dans la littérature on trouve deux classes pour ces méthodes. L'une est basé sur les contours, c'est-à-dire à partir le long du contour les endroits avec une courbure maximale ou un point d'inflexion sont recherchés. Et l'autre extrait le point d'intérêt directement à partir du signal de niveaux de gris (l'intensité).

2.3.2.1. Méthodes basées sur les contours

Le principe de ce détecteur est de rechercher les points de courbures maximales sur contour des images ou d'effectuer une approximation polygonale afin de déduire des points particuliers (intersections de contours, inflexion, etc). Ce sont les premières méthodes à bases de géométrie, elles ont été développées vers les années 80.

On commence par appliquer un détecteur de contours dans l'image puis on cherche sur ces contours les points d'inflexion, les points où la courbure est localement maximale ou encore les points d'intersections de plusieurs contours (les jonctions). on peut citer par exemple le détecteur de Medioni et al [24], et de Schmid [25].

2.3.2.2. Méthodes basées sur l'intensité

Ces méthodes sont basées sur la dérivée des niveaux de gris pour repérer les points où l'intensité varie fortement dans une ou plusieurs directions, elles permettent aussi de calculer pour chaque pixel une « valeur d'intérêt », appelée réponse, qui prend en compte les valeurs des pixels du voisinage.

Nous nous sommes intéressés à ces dernières méthodes qui sont les plus utilisées dans les travaux existants. Un des premiers détecteurs de cette catégorie est celui de Moravec [16]. On trouve également dans cette catégorie le célèbre détecteur de Harris [17] que nous détaillerons plus

particulièrement. Ainsi que David Lowe [20] qui utilise des déférences de gaussiennes pour trouver les points clés.

2.4. Quelques exemples de détecteurs des points d'intérêt

Les détecteurs des points de contrôle doivent tenir compte des changements de points de vue (rotation, déformations locales), de l'échelle, la luminosité des clichés et du bruit. La nature des points d'intérêt détectés dépend du type de détecteur utilisé : détecteurs de coins, d'intersections, de contours, de points isolés ou de zones homogènes. Nous présentons les détecteurs couramment utilisés comme suite :

2.4.1. Détecteur de Moravec (1977)

Moravec propose l'un des tout premiers détecteurs de points d'intérêt et s'intéresse à la fonction d'intensité I , son détecteur est connu comme étant un détecteur de coins. En 1977, il propose d'utiliser un détecteur qui se base sur le calcul de la fonction d'auto-corrélation, cette fonction utilisée pour détecter les coins et la somme du carré de la différence du signal avec lui-même décalé spatialement. Le principe est de calculer les variations d'intensité lumineuse dans quatre directions parallèles aux lignes et aux colonnes de l'image et de mesurer ensuite les différences d'intensité entre une fenêtre rectangulaire autour d'un pixel qui se déplace dans les diverses directions. Plus précisément on considère la fonction :

$$E(x, y) = \sum_{u,v} w(x, y) |I(x + u, y + u) - I(u, v)|^2 \quad (2.1)$$

- w spécifie la fenêtre/voisinage considérée (valeur 1 à l'intérieur de la fenêtre et 0 à l'extérieur);
- $I(u, v)$ est l'intensité au pixel (u, v) .
- $E(x, y)$ représente la moyenne du changement d'intensité lorsque la fenêtre est déplacée de (x, y) .

En appliquant cette fonction dans les trois situations principales suivantes voir la figure 2.2 ci-dessous, on obtient :

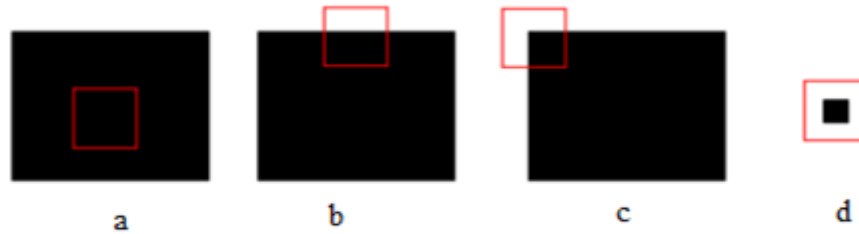


Figure 2.2 : Les différentes situations considérées par le détecteur de Moravec²⁰.

- a) L'intensité est approximativement constante dans la zone image considérée : la fonction $E(x, y) \approx 0$
- b) Contour : $E(x, y) \approx 0$ pour des déplacements le long du contour et $E(x, y) > 0$ pour des déplacements perpendiculaires au contour.
- c) Coin : La zone image considérée contient un coin ou un point isolé : la fonction E prendra de fortes valeurs dans toutes les directions. $E(x, y) > 0$ pour tout déplacement $(x, y) \neq (0, 0)$.
- d) Coin isolé.

Moravec définit un point d'intérêt comme étant un point où il y a une large variation de l'intensité dans certaine direction. Lorsque le minimum de ces quatre différences est supérieur à un seuil qui est 0, ceci indique la présence d'un point d'intérêt. Le détecteur de Moravec souffre de nombreuses limitations telles que : La réponse de ce détecteur est anisotropique en raison du caractère discret des directions de déplacements que l'on peut effectuer. La réponse est aussi bruitée en raison du voisinage rectangulaire considéré. Enfin, le détecteur de Moravec répond de manière trop forte aux contours en raison du fait que seul le minimum des différences est pris en compte de chaque pixel, la figure 2.3 illustre ce phénomène montrant des points d'intérêt suivants les contours du toit de la maison et les arrêtes de la la boite en forme d'hexagone.

²⁰ Source: <http://dept-info.labri.fr/~vialard/Traitement/cours/cours4.pdf>

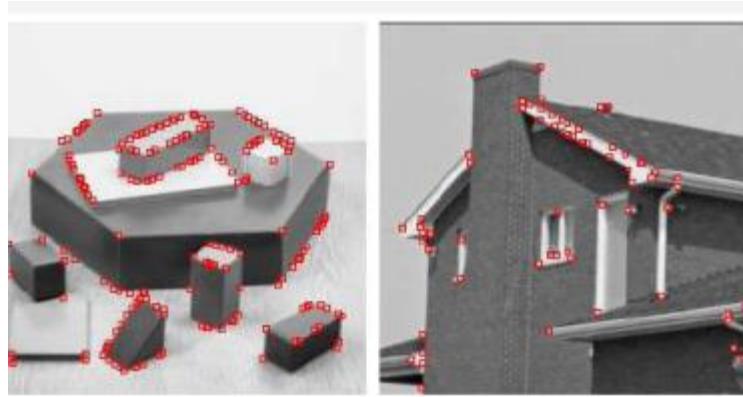


Figure 2.3 : Les points détectés avec une réponse forte sur les contours²¹.

2.4.2. Détecteur de Harris (1988)

Le détecteur de Harris et Stephen (Harris et Stephens, 1988), est certainement le plus connu et utilisé ; il est considéré comme robuste et fiable et a un côté universel, dans le sens où il fonctionne sur un large spectre d'images . En identifiant certaines limitations du détecteur de Moravec et en les corrigeant, ils en ont déduit un détecteur de coins très utile.

Ce détecteur a été conçu sur le même principe du détecteur de Moravec. L'idée de base de ce détecteur est d'utiliser la fonction d'auto-corrélation pour déterminer les endroits où le signal change dans deux directions simultanément. La version précise de ce détecteur a été évaluée dans Schmid [26], Schmid [25], dans laquelle le gradient de l'image est calculé en utilisant les dérivées d'une gaussienne plutôt qu'avec un simple masque discret. Comme le bruit contenu dans une image peut perturber considérablement le calcul de ses dérivés, il faut soit faire un lissage de l'image avant de calculer ses dérivées soit effectuer un lissage de l'opérateur de dérivation. Un moyen simple pour stabiliser les calculs des dérivées est donc d'utiliser les dérivées d'une fonction de lissage. La gaussienne est la plus communément choisie pour une telle fonction. La figure 2.4 est la même que celle de la figure 2.3 mais avec une bonne détection par rapport au détecteur Moravec, on peut remarquer que les points d'intérêt qui se trouvaient sur les contours du toit de la maison et les arrêtes de la boîte en forme d'hexagone ont disparu.

²¹ Source: http://read.pudn.com/downloads187/sourcecode/graph/texture_mapping/877328/house__.jpg

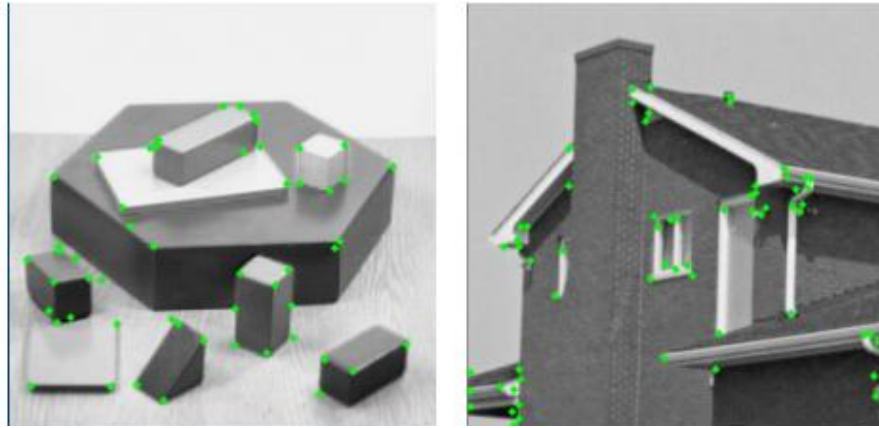


Figure 2.4 : La détection de coins pour Harris.

La fonction d'auto-corrélation locale mesure les changements locaux du signal. Cette mesure est obtenue comme suit : à l'aide de développement de Taylor de la fonction d'intensité I , En calculant les dérivées premières des points de l'image en ligne et en colonne :

$$I(x + u, y + v) = I(u, v) + x \frac{\delta I}{\delta x} + y \frac{\delta I}{\delta y} + o(x^2, y^2) \quad (2.2)$$

D'où :

$$E(x, y) = \sum_{u,v} w(u, v) \left[x \frac{\delta I}{\delta x} + y \frac{\delta I}{\delta y} + o(x^2, y^2) \right]^2 \quad (2.3)$$

Le terme $o(x^2, y^2)$ est négligeable pour les petits déplacements, nous aurons trois matrices avec les dérivées aux carrés et leurs multiplications ce sont :

- $A = \frac{\delta I^2}{\delta x} * w$
- $B = \frac{\delta I^2}{\delta y} * w$
- $C = \left(\frac{\delta I}{\delta x} \frac{\delta I}{\delta y} \right) * w$

On obtient l'expression analytique suivante :

$$E(x, y) = Ax^2 + 2C_{xy} + By^2 \quad (2.4)$$

La réponse du détecteur de Moravec est bruitée en raison de la fenêtre rectangulaire. Pour améliorer cela, Harris et Stephen proposent d'utiliser un filtre passe-bas gaussien pour lisser les trois matrices qu'on a vues:

$$w(u, v) = \exp\left(-\frac{(u^2+v^2)}{2\sigma^2}\right) \quad (2.5)$$

Avec : $M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$

Où M est une matrice (2 x 2) calculée à partir des dérivées de l'image.

Les valeurs propres de M sont les courbures principales de la fonction d'auto corrélation. Deux valeurs suffisamment grandes indiquent la présence d'un point d'intérêt. Pour ne pas avoir à calculer ces valeurs, Harris et Stephen proposent d'utiliser une mesure R faisant intervenir le déterminant et la trace de la matrice pour détecter les coins dans une image :

$$R = \text{Det}(M) - k\text{Trace}(M)^2 \quad (2.6)$$

Avec $\text{Det}(M) = AB - C^2 = \lambda_1 \cdot \lambda_2$ et $k\text{Trace}(M)^2 = A + B = \lambda_1 + \lambda_2$

Où k est un paramètre permettant de combiner l'information du contour donné par la trace de M et k est compris entre 0.04 et 0.06 mais ne dépasse pas 0.25.

Les valeurs prises par R peuvent se répartir en quatre régions :

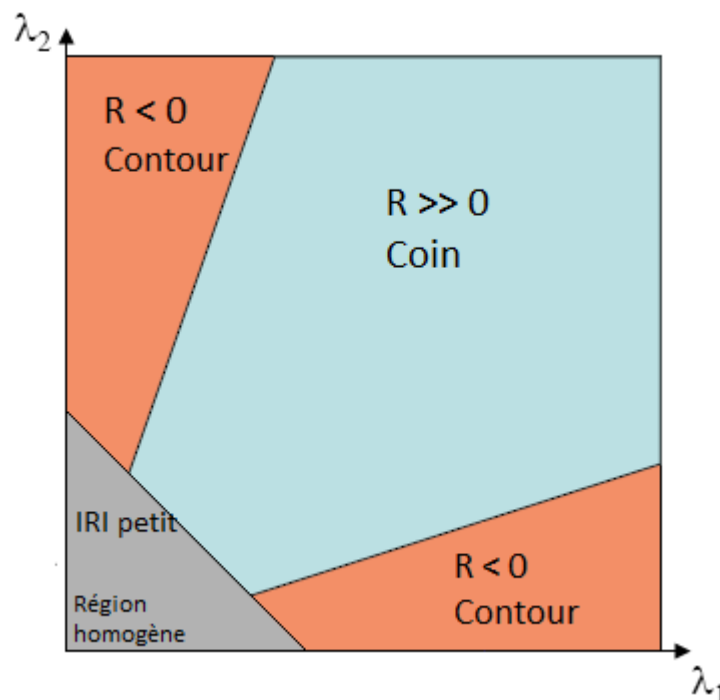


Figure 2.5 : Répartition des valeurs de R^{22} .

²² Source: http://www.csie.ntu.edu.tw/~cyy/courses/vfx/07spring/lectures/handouts/lec04_feature_4up.pdf

Comme le montre la figure 2.5, une fois la valeur de R calculée en chaque point, on effectue un seuillage pour ne garder que les valeurs les plus élevés de R . tel que : $R < seuil$.

Les valeurs de R sont positives au voisinage d'un coin, négatives au voisinage du contour et faibles en de région d'intensité constant

2.4.3. Principe de fonctionnement et les étapes du détecteur SIFT

La transformée des caractéristiques visuelles invariante à l'échelle en anglais " The Scale-invariant feature transform" (SIFT) a été publiée par David Lowe en 1999 [20] [27] pour détecter et décrire les caractéristiques locales des images en vision par ordinateur. Il a été largement utilisé dans la reconnaissance d'objets, la cartographie robotique et la navigation, la reconnaissance des gestes, le suivi vidéo et la correspondance.

L'algorithme SIFT utilise une technique de filtrage séquentiel qui sélectionne progressivement des points de contrôle pour un examen plus approfondi. Il a été conçu pour extraire des caractéristiques invariantes à la rotation, au changement d'échelle et à la variation de la luminance des images, ces points peuvent être utilisés pour effectuer une correspondance fiable du même objet ou scène entre différentes images.

L'algorithme SIFT génère un ensemble important de points de contrôle avec des descripteurs pour chaque point clé en utilisant des gradients d'image locaux dans le voisinage pour chaque échelle sélectionnée. La transformée SIFT se compose de différentes étapes pour l'extraction des points clés, ces étapes seront détaillées dans la section suivante.

2.4.3.1. Détection d'extrema dans l'espace des échelles

L'espace des échelles est une théorie qui s'appuie sur le fait que dans le monde réel les objets ont des propriétés qui ne peuvent être mise en évidence qu'à une échelle particulière.

Dans le cadre de l'algorithme des SIFT, l'outil de représentation du signal sur plusieurs échelles utilisé est une pyramide de gaussienne. Un "gradient noté L donné par l'équation 2.7 est le résultat de la convolution d'une image I par un filtre gaussien G de paramètre σ . Cette convolution permet de lisser l'image originale de telle sorte que les détails trop petits, sont éliminés et les contours sont atténués

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2.7)$$

Pour construire la pyramide :

- Les images de mêmes dimensions subissant un lissage avec un σ variant avec une progression géométrique, et donc avec une variation du gradient, constituent une octave.

- Passer d'une octave à une autre revient à doubler le paramètre σ voir figure 2.6. L'image utilisée pour créer l'octave suivante est donc l'image 2σ dont les dimensions sont divisées par 2.

On obtient un lissage d'image sur plusieurs échelles et à des dimensions différentes selon l'équation

$$L(x, y, k\sigma) = G(x, y, k\sigma) * I(x, y) \quad (2.8)$$

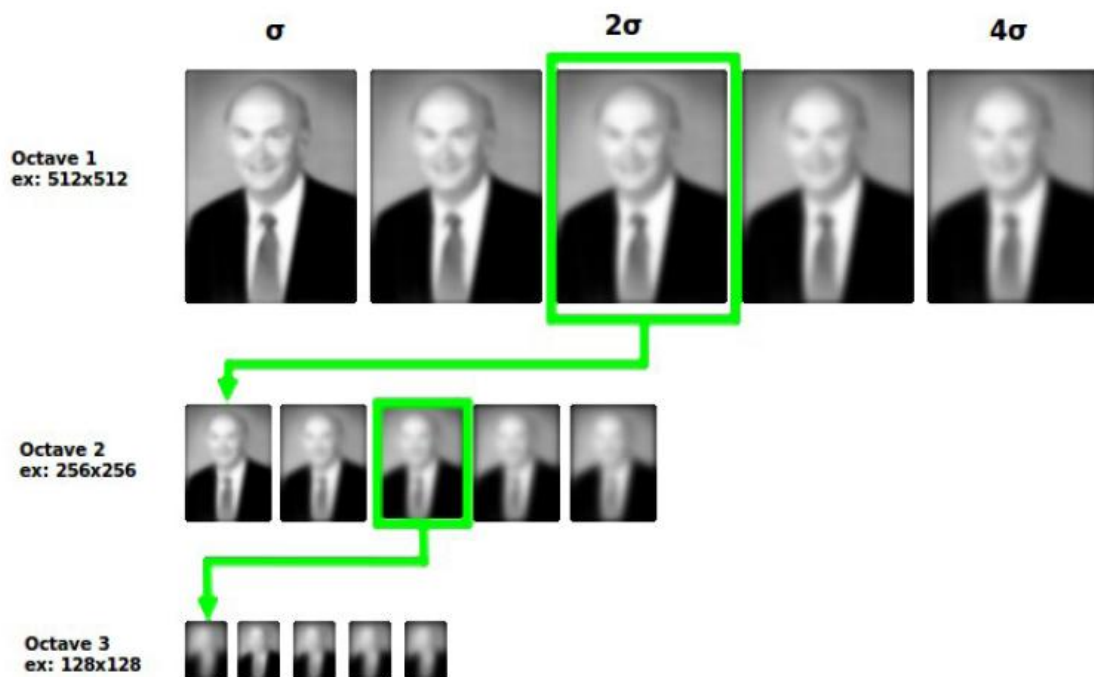


Figure 2.6 : Construction de la pyramide avec différentes échelles²³

2.4.3.2. Espace-échelle de détection des extremas

L'idée consiste ensuite à faire une différence de gaussiennes (DoG), entre deux images consécutives d'une même octave dans la pyramide de gaussienne pour obtenir une pyramide de DoG :

$$D(x, y, \sigma) = L(x, y, k_i \sigma) - L(x, y, k_j \sigma) \quad (2.9)$$

où k est un nombre constant afin d'obtenir un nombre fixe d'images lissées par octave, et de garantir que nous aurons le même nombre de DoG par octave. La différence de deux images

²³ Source: https://www.googlegroups.com/site/poublangsift/1-algorithme-des-sift/espace-des-echelles-et-extraction-de-caracteristiques/construction-de-l-espace-des-echelles/pyr_gauss.png

consécutives lissées par un filtre gaussien constitue une bonne approximation du LoG (Laplacian of Gaussian), un filtre passe-haut, utilisé en traitement d'image pour la détection de contours.

Par conséquent, une image d'un DOG entre les échelles $k_i\sigma$ et $k_j\sigma$ est tout simplement la différence des images floues gaussiennes à des échelles $k_i\sigma$ et $k_j\sigma$ voir figure 2.7.

Une fois les images de DoG ont été obtenues, les points clés sont identifiés sur plusieurs échelles (« scale invariant »), c'est à dire sur différents niveaux de résolution. Les points d'intérêt recherchés constituent les extrema locaux des images des DoG à travers les différentes échelles. Chaque pixel des images des DoG est alors comparé à ses 26 voisins :

- 8 voisins dans la même échelle.
- 9 voisins sur les deux échelles voisines qui l'encadrent.

Le pixel est un extremum local, s'il est supérieur ou inférieur à ses voisins alors il est sélectionné en tant que point candidat, Cette étape nous a permis de détecter les points d'intérêt candidats avec une approche multirésolution. Pour Lowe, cela permet de justifier l'invariance des futurs descripteurs à la mise à l'échelle.

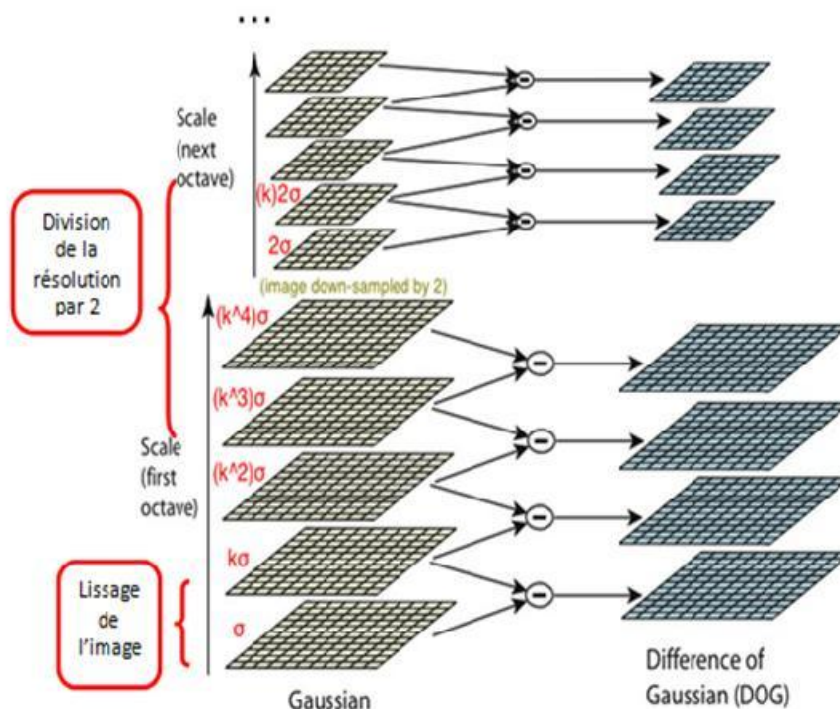


Figure 2.7 : Illustration de calcul de la différence gaussienne DOG²⁴

²⁴ Source: <http://www.aishack.in/static/img/tut/sift-dog-idea.jpg>

2.4.3.3. Localisation des points clés

L'étape suivante de la procédure d'extraction de caractéristiques consiste à localiser de façon précise les points d'intérêt (PI). Afin d'aboutir à ce but, David Lowe propose d'affiner les données acquises en décrivant « l'environnement » de chacun des PI, ce qui permet le rejet des points instables.

En effet, en effectuant une interpolation des coordonnées des points où se trouvent les extremums. Lowe utilise un développement de Taylor à l'ordre 2 au point candidat, de la fonction DOG $D(x)$ avec $x = (x, y, \sigma)^T$ où x est un point candidat sélectionné dans l'étape précédente.

$$D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (2.10)$$

La localisation de l'extremum \hat{x} réel est déterminée en prenant la dérivée de cette fonction par rapport à x en 0.

Pour déterminer si la position du point candidat est exacte, on calcule $\hat{x} - x$ dans les 3 dimensions. On regarde si la position obtenue est stable et si la localisation obtenue par l'interpolation est suffisamment proche de celle trouvée au départ :

- si $\hat{x} - x > +0.5$ dans n'importe quelle dimension alors on réévalue l'interpolation au point $x+1$
- si $\hat{x} - x < -0.5$ dans n'importe quelle dimension alors on réévalue l'interpolation au point $x-1$

Lowe propose de rejeter les points d'intérêt à faible contraste en calculant $D(x)$ en point \hat{x}

Si $|D(\hat{x})| < 0.03$, alors \hat{x} est considéré comme un point à faible contraste et il est éliminé de l'ensemble des points d'intérêt.

Il est également mieux d'éliminer les points de contours, sur lesquels la DoG a de fortes réponses en se basant sur le critère de la courbure de ces points le long du contour qui peut-être calculée grâce à la matrice Hessienne H donnée par l'équation (2.11) (la matrice des dérivées partielles secondes) à la position et à l'échelle du PI.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (2.11)$$

Les valeurs propres de cette matrice nous donnent des informations sur la courbure principale de D car elles sont proportionnelles avec celle-ci

On suppose que λ_1 est la plus grande valeur propre et λ_2 la plus petite. On peut calculer le rapport $r=\lambda_1/\lambda_2$ et $\lambda_1=r*\lambda_2$. On compare le nombre R donné par la formule (2.12) avec un seuil $s= 10$ qui est fixé par David Lowe

$$R = \frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\lambda_1+\lambda_2)^2}{\lambda_1\lambda_2} = \frac{(r\lambda_2+\lambda_2)^2}{r\lambda_2} = \frac{(r+1)^2}{r} \quad (2.12)$$

Ce calcul permet d'éliminer nombre important de points PI dont le rapport entre les deux courbures principales est plus grand que 10. Les points d'intérêt restants doivent avoir des caractéristiques robustes.

2.4.3.4. Affectation d'orientation aux points d'intérêt

Dans cette étape, on attribue à chaque point d'intérêt une ou plusieurs orientations en utilisant la direction des gradients des voisins directes de ce point. Pour cela on parcourt tous les pixels de toutes les images gaussiennes à toutes les octaves et on leur affecte une orientation et une norme.

On détermine leur norme $m(x, y)$ et leur orientation $\theta(x, y)$ en se servant des différences finies :

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad (2.13)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (2.14)$$

Afin de déterminer l'orientation d'un point d'intérêt, un histogramme figure 2.8 d'orientation de gradient est calculé par rapport au voisinage de ce point en utilisant l'image gaussienne dont l'échelle est la plus proche de l'échelle du PI considéré.

La taille du voisinage qui va permettre de d'établir l'histogramme des orientations, est choisi par Lowe en considérant une fenêtre gaussienne dont la taille est égale à 1.5 fois le facteur d'échelle. Chaque élément ajouté à l'histogramme est également pondéré par sa norme pour que l'importance de sa contribution soit juste.

Les caractéristiques des points "PI" sont définies maintenant par (x, y, σ, θ) . Cet histogramme dispose de 36 graduations pour les 360°. Les pics permettent de mettre en évidence des orientations

à prendre en compte pour caractériser les PI. L'orientation dominante au voisinage d'un PI est caractérisée par les pics de l'histogramme. Pour augmenter la quantité d'information disponible, Lowe propose de se servir des pics de l'histogramme atteignant au moins 80 % du pic dominant pour créer de nouveaux PI. Ceci contribue plus à la stabilité de l'algorithme.

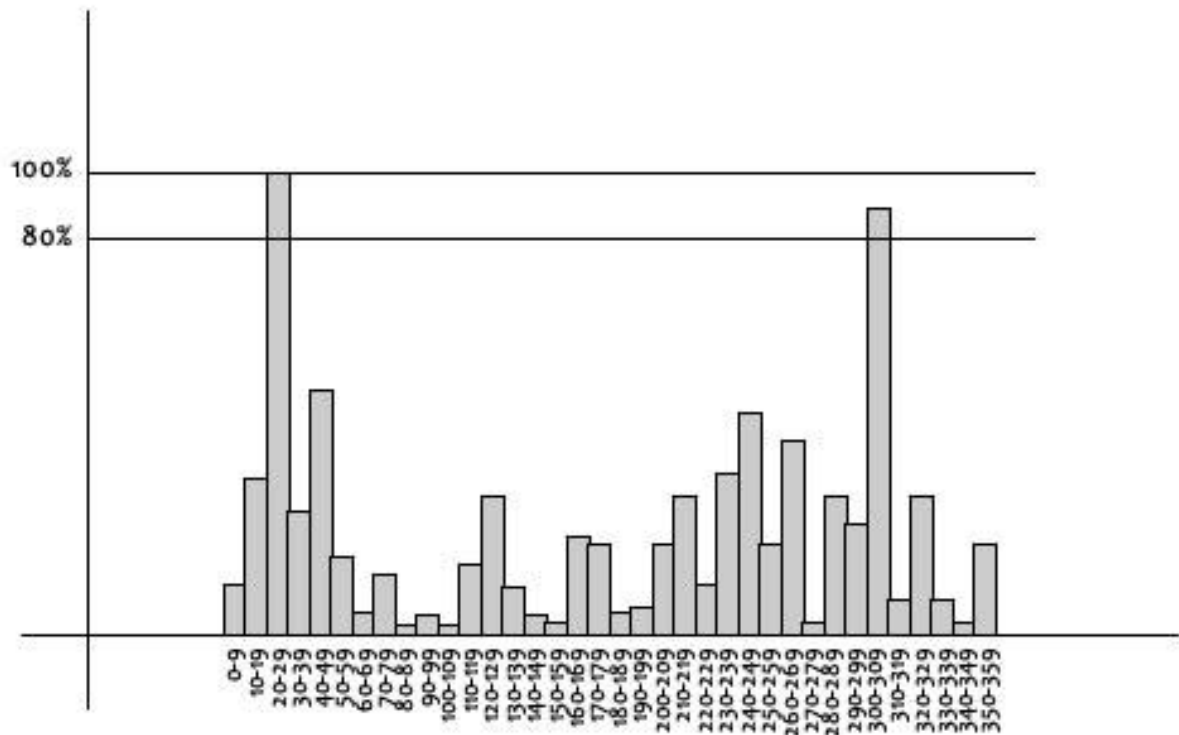


Figure 2.8 : Histogramme d'orientation des points d'intérêt²⁵

La figure 2.9 illustre un exemple de la détection des points d'intérêt en utilisant la transformée SIFT qui donne 3000 points avec différentes échelles et orientations

²⁵ Source: <http://www.aishack.in/static/img/tut/sift-orientation-histogram.jpg>



Figure 2.9 : Détections des points d'intérêt par SIFT

2.4.4. Détecteur de Points d'intérêt Basé sur la NSCT

La transformée en Contourlets non sous échantillonnée, NSCT (NonsubSampled Contourlet Transform) Cunha et al [28], est une version invariante par translation de la transformée en Contourlets. L'invariance par translation de la NSCT est assurée par une implantation d'une structure pyramidale non sous-échantillonnée et des bancs de filtres directionnels non sous-échantillonnés. La NSCT possède des filtres avec une sélectivité fréquentielle meilleure ce qui assure une meilleure décomposition en sous-bandes.

La Figure 2.10 illustre le principe de la NSCT. La construction proposée est basée sur une structure pyramidale et des bancs de filtres directionnels non-échantillonnés. Il en résulte une décomposition d'image flexible, multiéchelle, multidirectionnelle et invariante par translation, qui peut être efficacement mise en œuvre via l'algorithme à trous [29].

La propriété multiéchelle de la NSCT est obtenue à partir d'une structure de filtrage invariante par translation qui effectue une décomposition en sous-bandes similaire à celle de la pyramide Laplacienne. Ceci est obtenu par des bancs de filtres 2D non sous-échantillonnés à deux canaux. La décomposition directionnelle invariante par translation repose sur un banc de filtres directionnels non sous-échantillonnés, NSDFB.

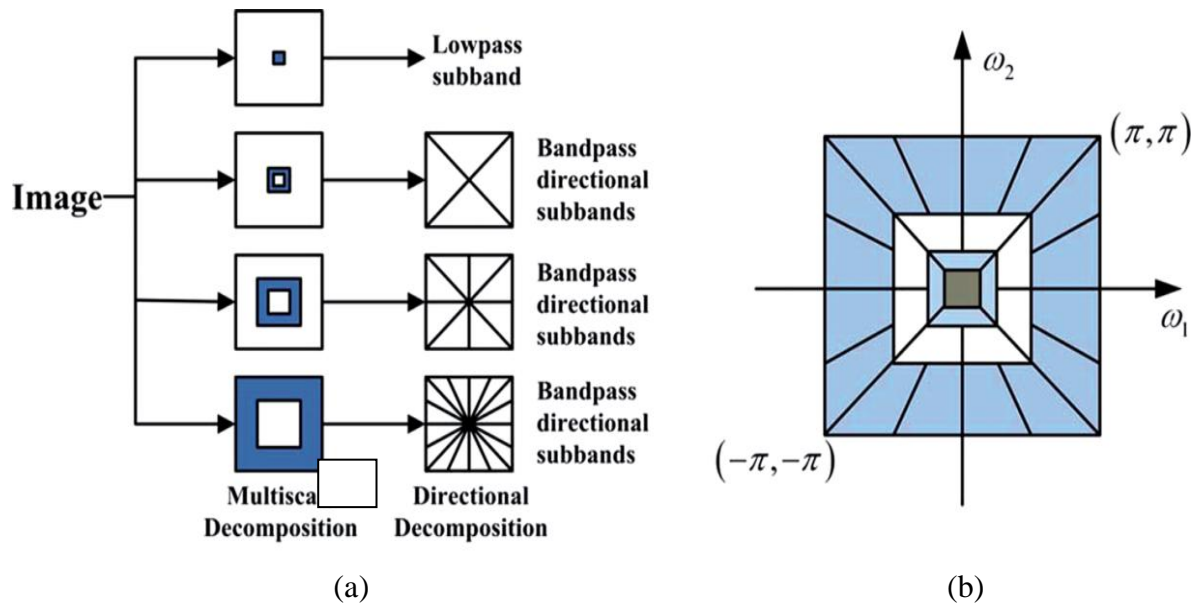


Figure 2.10 : La transformée en contourlets non sous-échantillonnée²⁶. (a) Schéma bloqué de l'implantation de la NSCT à l'aide d'un banc de filtre pyramidal directionnel non sous-échantillonné. (b) le partitionnement fréquentiel résultant, où le nombre de directions augmente avec la fréquence.

La Figure 2.11 montre un exemple de la décomposition en NSCT de l'image standard 'Zoneplate' (256×256 pixels) après l'exécution du programme NSCT demo sous matlab . Trois niveaux NSP (Non Subsampled Pyramid) et 8-4-2 sous-bandes directionnelles NSDFB ont été utilisés. L'image originale et les sous-bandes résultantes ont la même taille. Comparée à la transformée en contourlets, la NSCT est une décomposition d'image entièrement invariante par translation, multiéchelle, et multidirection, qui a une meilleure sélectivité fréquentielle et une implémentation rapide. La NSCT s'est avérée être très efficace dans le débruitage et l'amélioration d'images Cunha [28].

Serief et Bentoutou [30] proposent un algorithme basé sur la NSCT pour l'extraction des points d'intérêt utilisés pour le recalage des images.

²⁶ Source: http://opticalengineering.spiedigitallibrary.org/data/Journals/OPTICE/23431/067010_1_1.png

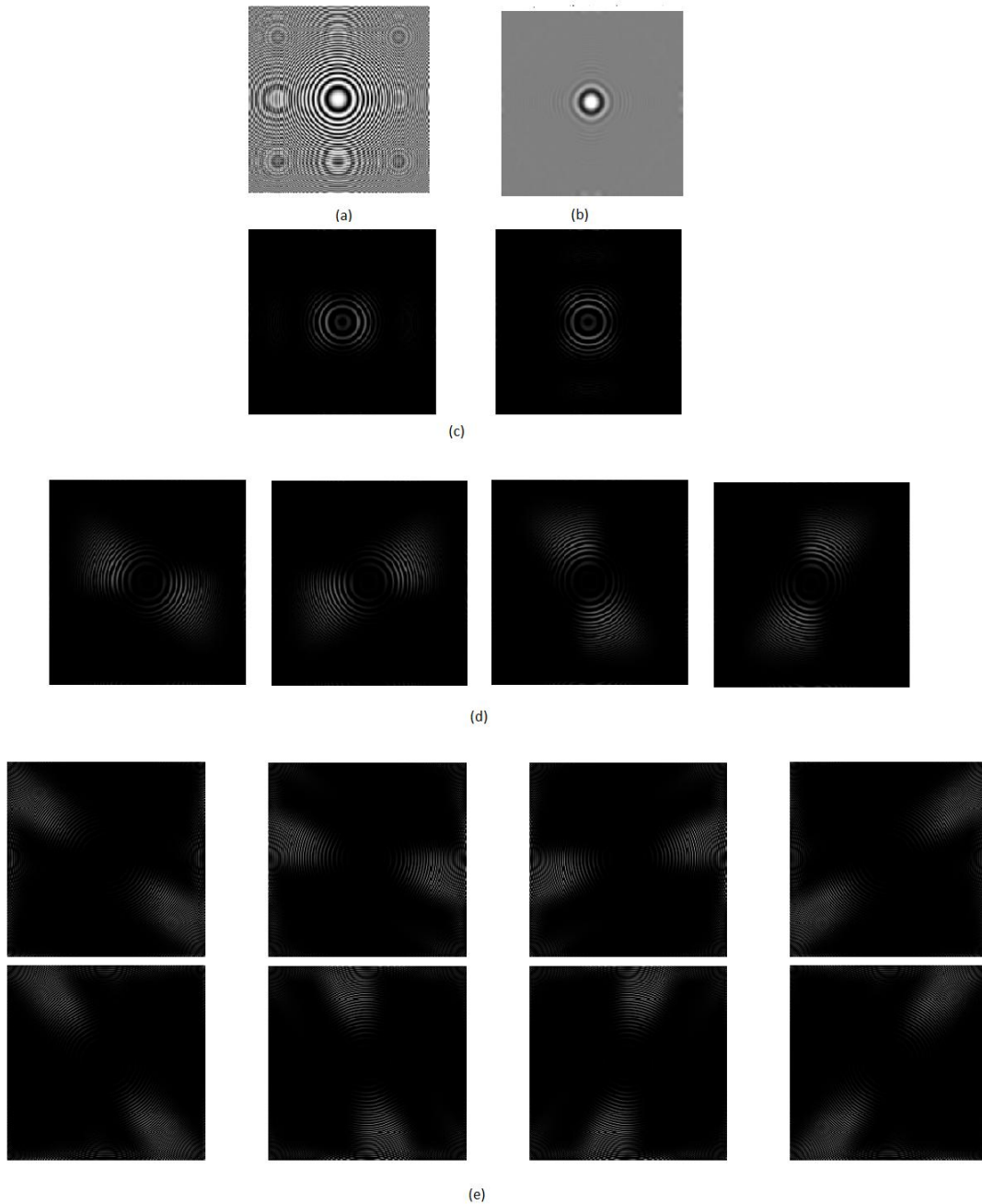


Figure 2.11 : La transformée en contourlets non sous-échantillonnée de l'image "Zoneplate". L'image est décomposée en trois niveaux pyramidaux, qui sont ensuite décomposés en 2, 4, et 8 sous-bandes directionnelles. Les petits coefficients sont représentés en noir alors que les grands coefficients sont représentés en blanc. (a) L'image originale. (b) Sous-bande passe-bas. (c), (d), (e) Sous-bandes directionnelles passe-bande.

Algorithme d'extraction des points d'intérêt basé sur NSCT

Serief et al considèrent que les points d'intérêt sont des caractéristiques locales d'une image. Elles appartiennent généralement à des structures de tailles différentes. Un moyen efficace pour détecter de telles caractéristiques est d'avoir une décomposition multi-échelle de l'image inspiré du modèle d'interaction d'échelle (scale-interaction) conçu par Manjunath et Chellappa [31]. En outre, une représentation directionnelle de l'image fournit une information intuitive et utile pour décrire la forme des structures 2D.

La NSCT ne fournit pas seulement une analyse multiéchelle, mais aussi une représentation géométrique directionnelle. Elle est invariante par translation telle que chaque pixel des sous-bandes, résultant de la décomposition, correspond à son homologue de l'image originale dans la même position. L'information géométrique peut donc être recueillie pixel par pixel à partir des coefficients NSCT. En raison de cette analyse, Serief constate qu'il est intéressant d'appliquer la NSCT dans la détection des points d'intérêt.

La méthode de détection de points d'intérêt proposée par Serief se compose des trois étapes suivantes:

1. Tout d'abord, l'image d'entrée est décomposée à l'aide de la NSCT le long des différentes résolutions directionnelles pour différentes échelles pour l'extraction des caractéristiques directionnelles de premier niveau.
2. La deuxième étape consiste en les interactions d'échelles en calculant, pour chaque résolution directionnelle, la différence entre les caractéristiques obtenues à deux échelles différentes. Cela se justifie par le fait que la NSCT est invariante par translation.
3. La dernière étape sert à localiser ces caractéristiques, en calculant, à chaque emplacement de pixel, la somme des carrés des magnitudes de toutes les sous-bandes différence obtenues. Il en résulte une carte de caractéristiques qui met en relief les caractéristiques de l'image originale. on applique un seuillage et une suppression des non maxima sur la carte des caractéristiques obtenue afin de détecter les points d'intérêt. Ce seuil est choisi pour être proportionnel à la moyenne et l'écart-type de la carte des caractéristiques.

Les différentes étapes de l'algorithme sont décrites succécivement ci-dessous:

Etape 1) Calculer les coefficients NSCT de l'image d'entrée pour N niveaux et L sous-bandes directionnelles.

Etape 2) Calculer la différence entre chaque sous-bande à un niveau et la sous-bande correspondante à un autre niveau. L sous-bandes différence seront obtenues à la fin.

Etape 3) A chaque emplacement de pixel, calculer la somme des carrés des magnitudes de toutes les sous-bandes différence obtenues. Ces points sont appelés «maxima des coefficients NSCT». Un seuillage est ensuite appliqué à l'image des maxima afin d'éliminer les points non significatifs. Un point est enregistré si

$$\text{NSCT maxima} > Th$$

Où $Th = c(\sigma + \mu)$, c étant un paramètre dont la valeur est définie par l'utilisateur, et σ et μ sont l'écart-type et la moyenne de la carte des caractéristiques respectivement.

Etape 4) Appliquer la suppression des non-maximas aux résultats obtenus à l'étape 3. La suppression des non-maximas est une technique de post-traitement simple et efficace en traitement d'images. Elle utilise une fenêtre de taille, $w \times w$, balayant tous les pixels de l'image. Les points d'intérêt finaux sont sélectionnés comme les points qui sont les maximums locaux à l'intérieur de cette fenêtre

2.5. Évaluation des détecteurs étudiés

Pour évaluer les performances des détecteurs de point d'intérêt étudiés dans les paragraphes précédentes, une étude comparative est établie des détecteurs de Moravec, Harris, SIFT et le détecteur basé sur NSCT selon les critères suivants:

- Détection: le détecteur de coins doit détecter même les coins imperceptible, tout en étant robuste au bruit, et éviter les faux points
- Localisation: les coins doivent être détectés le plus près possible de leur véritable emplacement.
- Coût calculatoire: la vitesse est un critère fondamentale pour des applications à temps réel, cependant les progrès technologiques ont pu résoudre ce problème et permettent aujourd'hui de réduire le temps de calcul.

Tout d'abord, on a choisi les plus simples des jonctions sous forme de damier. La figure 2.12 nous montre les coins détectés avec les différents détecteurs. Pour réaliser cette comparaison on a du

utiliser une détection subjective afin de localiser exactement les points de contrôle et les dénombrer et on a trouvé 86 points d'intérêt pour l'image "Damier". On a constaté que le détecteur de Moravec a donné une réponse très forte pour les contours donnant ainsi à une multitude de points sur le côté haut de l'image damier

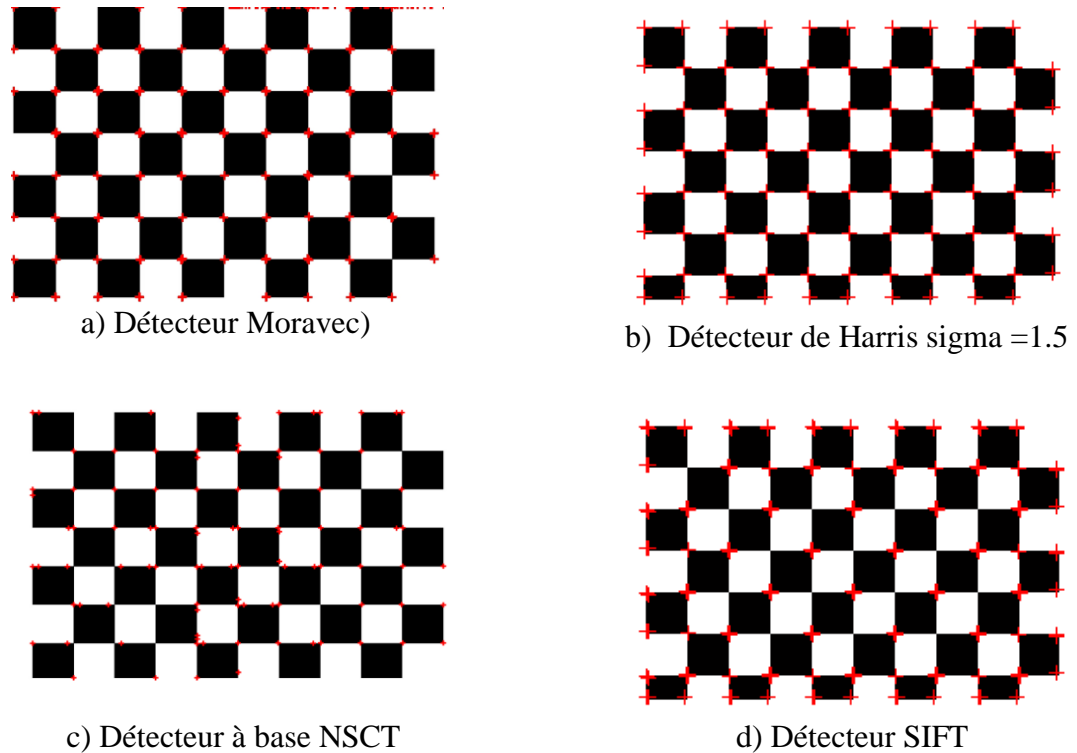


Figure 2.12 : Extraction des points d'intérêt utilisant différents détecteurs.

Détecteurs proposés	Nombres de points détectés	Nombres des vrais points détectés	Pourcentage des points détectés	Temps de calcul en secondes
Moravec	77	77	89.53 %	4.36
Harris sigma=1.5	86	69	80.23 %	9.60
NSCT	94	59	68.60 %	13.83
SIFT	77	59	68.60%	15.59

Tableau 2.1 : Tableau comparatif des détecteurs de points d'intérêt selon leur nombres, leur localisations et leur temps de calcul pour l'image damier

D'après le tableau 2.1, on peut remarquer que le détecteur de Moravec a donné un pourcentage élevé (89.53 %) des points d'intérêt trouvés avec une bonne localisation, mais il représente une réponse forte pour les contours et a engendré des faux points qui peuvent présenter un problème lors de l'appariement et le calcul de la transformé de recalage, le détecteur de Harris vient en deuxième position avec une valeur de 80.23 %. Cependant, on doit rappeler que l'image test représente seulement des coins sous forme d'angle à 90 degrés c'est pour cela on a jugé d'utiliser une autre image test jouissant d'une variété de formes de points d'intérêt.

On a constaté aussi que certains points de contrôle ont été détectés par NSCT et non pas par SIFT et vice versa, la figure 2.13 illustre cette différence de détections des points entre SIFT et NSCT.

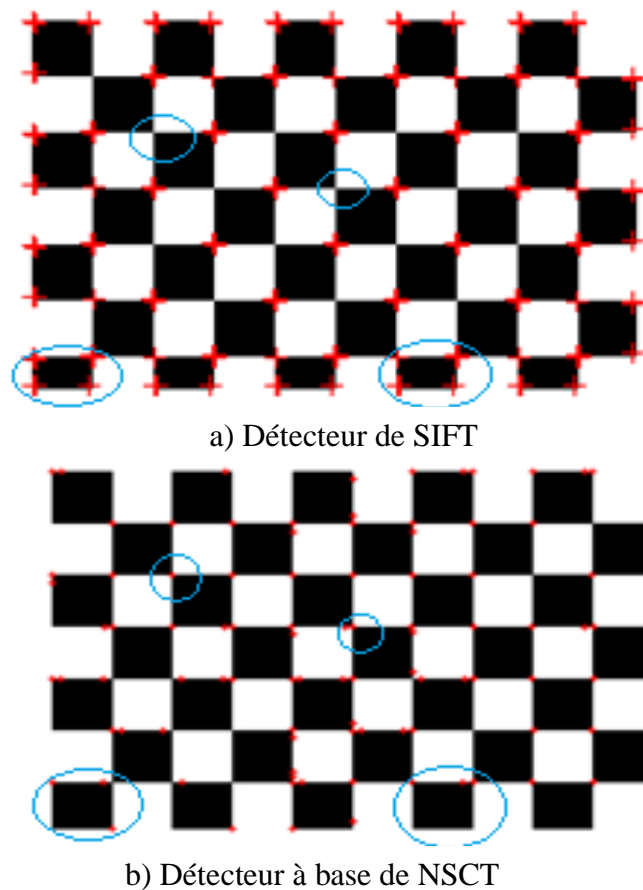


Figure 2.13 : Différence de détections des points entre SIFT et Harris.

Pour pousser un peu plus loin notre évaluation, nous avons proposé une image synthétique, qui contient plusieurs types de coins incluant : les jonctions en L, en Y, en X, et les jonctions-flèches. La détection subjective a révélé 79 points d'intérêt Les réponses visuelles de chaque détecteur

sont illustrées sur la Figure 2.14 montrant que le détecteur de Moravec a trouvé un nombre minimal de points d'intérêt et qu'il n'a détecté aucun point dans la zone de variation des niveaux de gris (partie à gauche de l'image de la figure 2.14. a), Alors que le détecteur à base de NSCT ainsi que SIFT ont pu localiser des points clé dans cette zone avec une meilleure précision pour le détecteur basé sur NSCT.

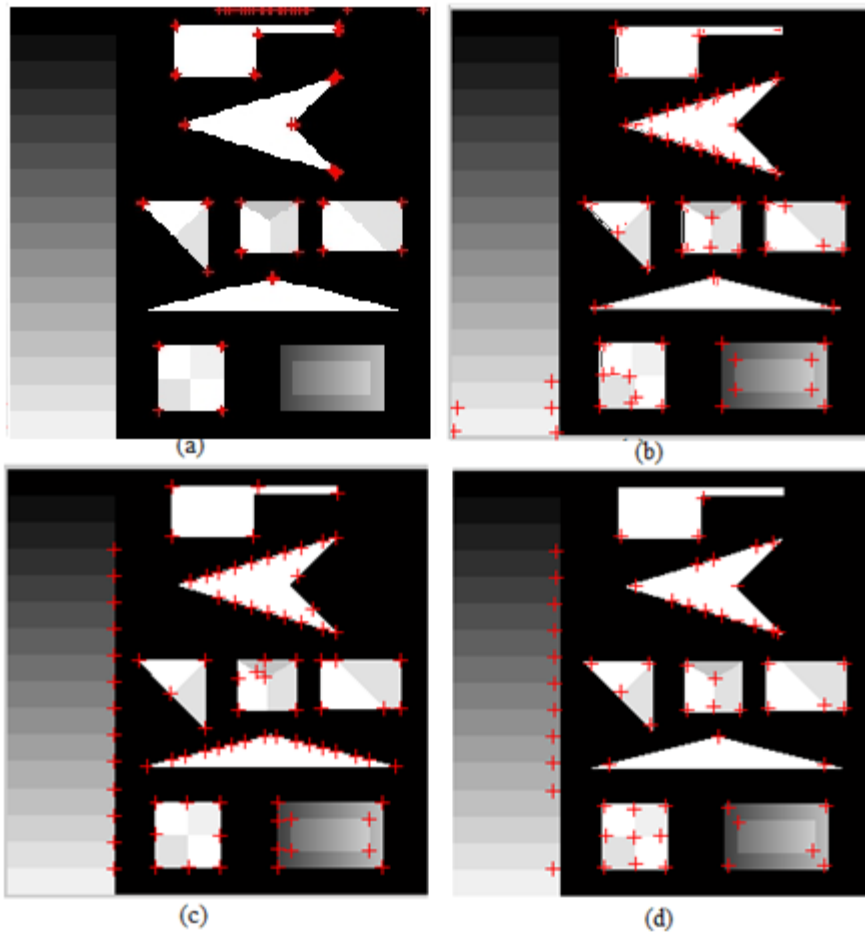


Figure 2.14 : Extraction des points d'intérêt de l'image synthétique 2 contenant plusieurs formes de jonctions: (a) Le détecteur de Moravec (b) Le détecteur de Harris $\sigma=1.5$ (c) Le détecteur à base de NSCT (d) Le détecteur de SIFT.

Détecteurs proposés	Nombres de points détectés	Nombres des vrais points détectés	Pourcentage des points détectés	Temps de calcul en secondes
Moravec	28	26	32.91 %	5.20
Harris sigma=1.5	62	39	49.36 %	6.73
NSCT	89	52	65.82 %	24.19
SIFT	60	40	50.63 %	16.11

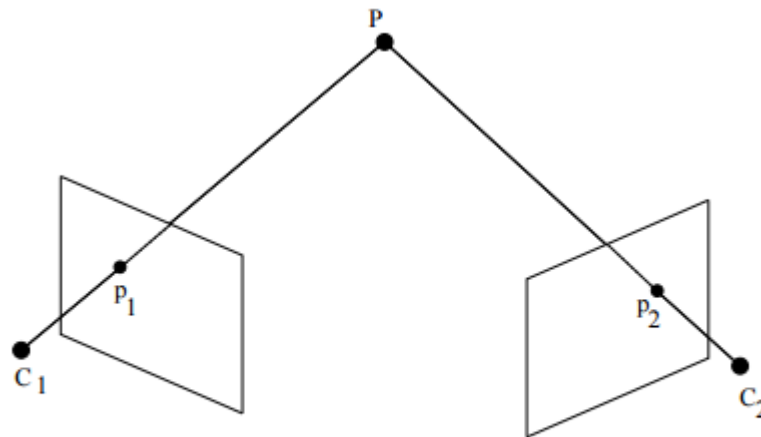
Tableau 2.2 : Tableau comparatif des détecteurs de points d'intérêt selon leur nombres, leur localisations et leur temps de calcul pour l'image synthétique contenant différents formes de jonctions.

Le tableau 2.2 contient les statistiques des différents détecteurs, on peut remarquer en comparant le nombre de ces points, que le détecteur à base de NSCT donne un pourcentage de 65.82% ce qui représente le meilleur résultat par rapport aux autres détecteurs en se référant aux points de contrôle subjectivement détecté (79 points), puis vient en deuxième position le détecteur SIFT avec un pourcentage de 50.63 %. Le temps de calcul des différents détecteurs ne varie que de quelques secondes, ce qui ne donne pas une grande importance à ce critère pour le bon choix du détecteur des points de contrôles, surtout avec les progrès et l'évolution des calculateurs récents.

La comparaison statistique et visuelle des points clé détectés nous a permis de conclure que le détecteur SIFT et celui à base de NSCT offrent les meilleurs résultats du point de vue efficacité et précision de localisation et que l'un peut compléter l'autre puisque certains points ont été détecté par SIFT et d'autres par NSCT.

2.6. La mise en correspondance

L'appariement des points d'intérêt entre les images est une étape importante pour le recalage, cet appariement doit tenir compte des changements de points de vue (rotation, déformations locales), de l'échelle et de la luminosité des images prises de différents angles. Soit deux images I_1 et I_2 d'une scène, il s'agit de déterminer le point clé, pour l'image I_1 , qui lui correspond dans l'image I_2 et éventuellement dans d'autres images voir figure 2.15. La mise en correspondance des primitives est un problème fondamental de la vision par ordinateur.



Deux pixels p_1 et p_2 correspondants.

Figure 2.15 : Exemple de mise en correspondance de deux pixels.

2.6.1. Difficultés et contraintes de la mise en correspondance

2.6.1.1. Difficultés de la mise en correspondance

Les différents problèmes souvent rencontrés par les méthodes de mise en correspondance sont:

- Le bruit : Un problème commun à la vision par ordinateur et au traitement d'images, le bruit affecte les niveaux de gris ou la couleur des pixels. Pour limiter les conséquences de ce problème, on peut appliquer un filtre dans l'image. Il existe différents filtres adaptés à chaque type de bruit.
- Les occultations : Il s'agit des zones qui ne sont visibles que dans une image et qui sont donc occultées dans l'autre image. Cela pose le problème de la détermination des correspondants des pixels occultés notamment pour les méthodes locales, car dans ce cas l'hypothèse de départ sur les voisinages respectifs n'est plus respectée.

- discontinuités de profondeur : Au niveau d'un contour séparant deux objets de profondeurs différentes, les voisinages ne sont pas les mêmes sur deux images. Par conséquent, les méthodes locales basées sur le voisinage du pixel considéré montrent une certaine faiblesse.
- Les raccourcissements : Lorsque les surfaces des objets de la scène sont très inclinées par rapport à l'un des plans image, un pixel d'une image peut correspondre à plusieurs pixels de l'autre image.
- Les zones homogènes : Dans les zones peu texturées, les niveaux de gris des pixels sont uniformes ou quasi-uniformes. Il est donc plus difficile de distinguer le bon correspondant de ses voisins en utilisant une mesure de corrélation.

2.6.1.2. Contrainte de la mise en correspondance

- Contrainte d'unicité : Lorsque l'on veut mettre en correspondance deux images, il est fréquent d'imposer la contrainte d'unicité. En effet un point d'une image doit correspondre à un point unique dans la deuxième image et inversement. Cette contrainte, permet également de réduire le coût des calculs car si un pixel de l'image de droite a été attribué à un pixel de l'image gauche, alors, il n'est plus nécessaire de le prendre en compte lorsque l'on cherche les correspondants des pixels restants à traiter (figure 2.16).

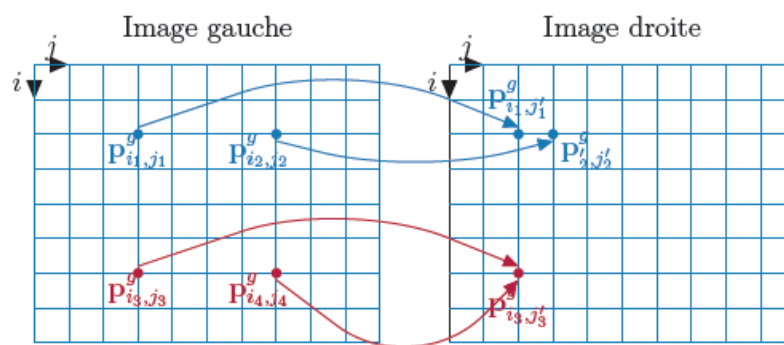


Figure 2.16 : La contrainte d'unicité²⁷: dans la figure nous pouvons voir les correspondances des pixels en bleu qui respectent l'unicité alors que celles en rouge la ne respectent pas

- Contrainte de visibilité : De très nombreuses techniques ont été proposées pour prendre en compte le problème des occultations. L'une d'elles, proposé par Sun et al [32], introduit la

²⁷ Source: Thèse "Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d'appariements de points d'intérêt et sondage de régions" page 39 par Guillaume Gales

contrainte suivante : un pixel occulté ne doit pas avoir de correspondant et un pixel non-occulté doit avoir au moins un correspondant.

- Contrainte de symétrie : Toujours pour résoudre le problème des occultations, certains auteurs Szeliski, Bobick [33] [34] proposent un modèle symétrique et calculent les cartes de disparité gauche et droite puis les comparent afin de localiser les zones occultées.
- Contrainte d'ordre : L'ordre des pixels de l'image gauche le long d'une droite épipolaire doit être le même que celui de leurs correspondants (figure 2.17).

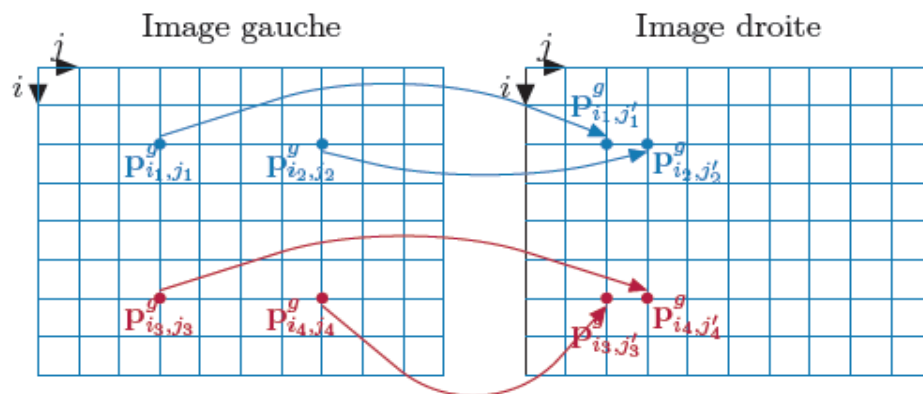


Figure 2.17 : Contrainte d'ordre²⁸ - nous pouvons voir deux correspondances, l'une en bleu qui respecte la contrainte de l'ordre et l'autre en rouge qui ne la respecte pas.

2.6.2. Les méthodes de mise en correspondance

Il existe un grand nombre de méthodes d'appariement de points d'intérêt, notre analyse des méthodes d'appariements se limite donc à la mise en correspondance de ces derniers. Le principe est d'interpréter les informations du point d'intérêt et de son voisinage proche pour extraire le couple présentant la meilleure ressemblance. Nous nous sommes contentés d'étudier la méthode locale.

2.6.2.1. Méthodes locales

- Méthode par corrélation : Les méthodes par corrélation sont principalement utilisées dans l'analyse de l'information d'intensités pour la mise en correspondance. Le principe est de déterminer, pour le voisinage d'un point x_1 de la première image, la corrélation maximale (distance minimale) avec un voisinage issu de la seconde image. Ce calcul permet donc

²⁸ Source: Thèse "Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d'appariements de points d'intérêt et sondage de régions" page 39 par Guillaume Gales

d'extraire le point x_2 formant ainsi le couple (x_1, x_2) présentant la meilleure ressemblance au sens de la corrélation. La figure 2.18, résume une telle mise en correspondance.

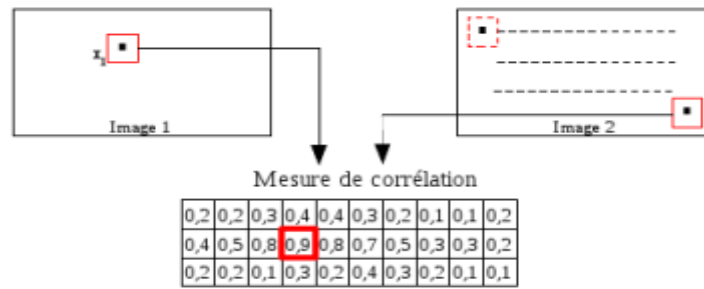


Figure 2.18 : Principe de la corrélation : recherche du point issu de la seconde image présentant la meilleure ressemblance.

Afin d'optimiser cette méthode, une estimation de la position de x_2 peut être introduite. Nous déterminons alors les mesures de corrélation à l'intérieur d'une zone de recherche, et non plus sur l'image entière. La figure 2.19 donne un aperçu de la zone de recherche.

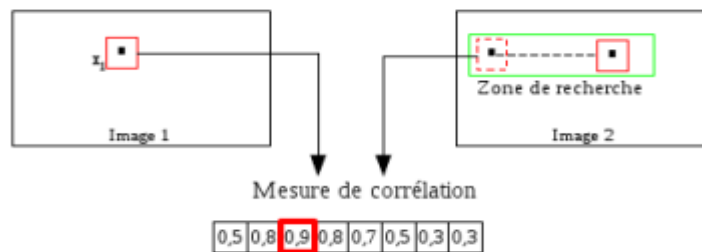


Figure 2.19 : Exemple de mesures de corrélation entre le point x_1 et les points x_2 présents dans la zone de recherche.

La figure 2.20 illustre un exemple de mise en correspondance de deux images de l'angiographie rétinienne en utilisant l'algorithme SIFT montrant l'échelle d'appartenance des points d'intérêt et leurs orientations.

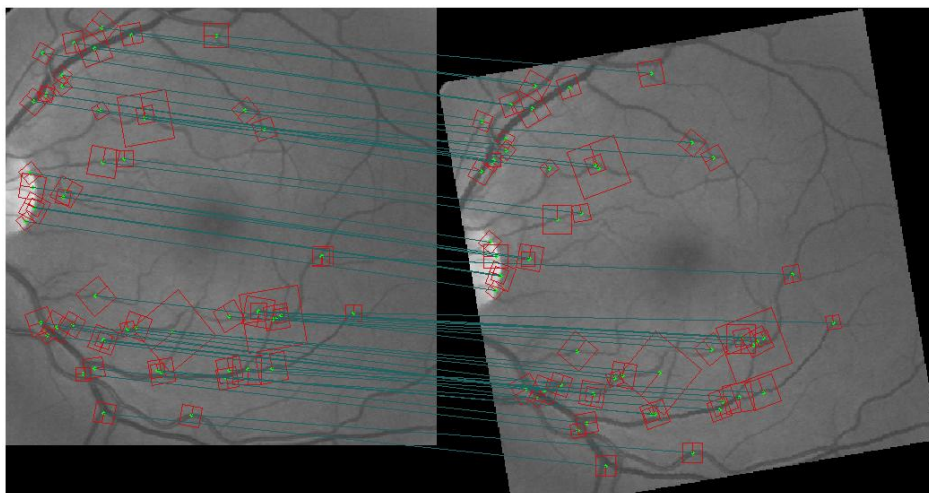


Figure 2.20 : Correspondance des points d'intérêt entre l'image d'angiographie rétinienne et son homologue recalé de 7 degré.

2.7. Conclusion

La première étapes qui nous permet de réaliser la reconstruction panoramique des images est :

La détection et la correspondance des points d'intérêt. Donc l'extraction des points joue un rôle important pour le calcul de la transformée de recalage, la précision d'extraction des points de contrôle influe beaucoup sur la précision de leur appariement .

Une étude des principes de fonctionnement de certains détecteurs a été faite avec une comparaison de nombres des points détectés ainsi que la précision de localisations, nous a permis de juger que le détecteur à base de NSCT et celui de David Lowe (SIFT) offrent les meilleurs résultats, ce qui nous a donné l'idée de combiner ces deux transformées pour concevoir un détecteur à base de NSCT et SIFT. Cet algorithme proposé sera détaillé dans le chapitre suivant.

Chapitre 3

Recalage et Evaluation de l'algorithme NSCT-SIFT

3.1. Introduction

Le recalage des images est une étape essentielle pour la création d'une vue panoramique qui consiste à superposer deux ou plusieurs images de la même scène prises à partir de différents angles, à des instants différents et/ou par des capteurs différents selon qu'il s'agit d'un recalage monomodale ou multimodale.

Il existe plusieurs algorithmes de recalage qu'on peut pas tous les dénombrer. Une étude comparative donnant plus de détail sur ces algorithmes a été établie par Kratika Sharma et al [35], ainsi que Zitova et Flusser[36].

Le recalage d'image peut être appliquer dans le domaine médical dans plusieurs applications telles que la fusion des images médicales pour faciliter le diagnostic, la radiothérapie pour le positionnement du patient lors de la cure thérapeutique.

Plusieurs méthodes de recalage ont été proposées dans ce domaine. Can et al [37] ont utilisé l'arbre de bifurcation vasculaire pour la détection des points d'intérêts sur la surface de la rétine et la création d'une vue panoramique de l'angiographie rétinienne.

La transformée SIFT reste parmi les méthodes qui ont prouvé leur efficacité pour la détection et la mise en correspondance des points de contrôle. Elle a apporté l'invariance à la rotation, aux changements d'échelle, d'angle d'observation et aux différences des expositions de luminosité. Cependant lors de l'application de la transformée SIFT pour l'extraction des points clés pour établir la correspondance entre les différents points des images de l'angiographie rétinienne et/ou les images de la cystoscopie, on a constaté des erreurs de correspondance donnant ainsi des transformées de recalage imprécises.

On propose dans notre algorithme de rehausser les contours de l'image par la transformée contourlet NSCT puis d'appliquer la transformée SIFT (Scale Invariant Feature Transform) pour pour essayer de détecter les points d'intérêt efficacement et aussi corriger les erreurs de correspondance.

Cette méthode diffère des autres techniques, car elle n'a pas besoin de rechercher les bifurcations vasculaires de la rétine ou des images cystoscopiques qui sont parfois difficiles à détecter dans certaines pathologies telles que les tumeurs [38] [39]. La figure 3.1 donne un aperçu de la rétine saine face à une rétine avec une surface pathologique cachant les structures vasculaires.

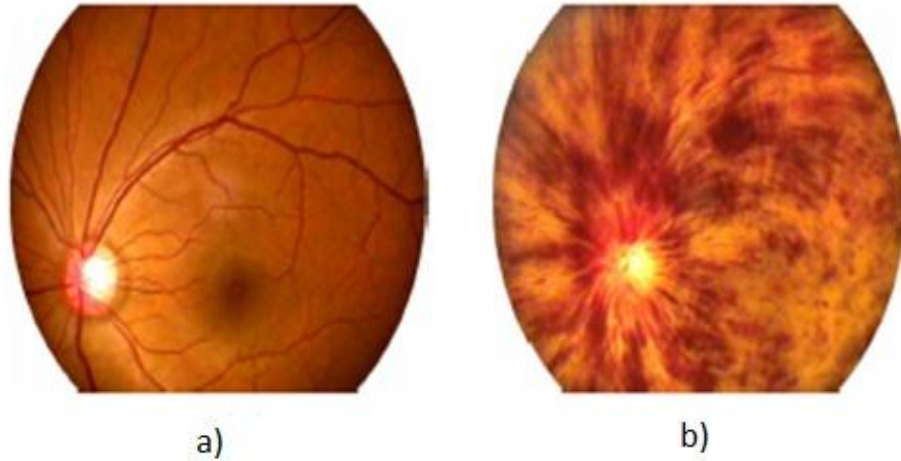


Figure 3.1 : Images angiographiques de la rétine²⁹

a) Rétine saine b) Rétine pathologique

3.2. Principe du recalage

On peut définir le recalage des images comme étant la transformation géométrique qui ramène deux images contenant des informations communes dans un même repère et cela en se basant sur les points d'intérêts trouvés par les détecteurs usuels déjà traités en chapitre précédent.

Le processus de recalage de deux images peut se formaliser de la façon suivante :

$$\tilde{T} = \operatorname{argopt} \sum_x \operatorname{sim} [g(I_R(X), \Omega(g(I_D(T(X)))))] \quad (3.1)$$

I_D est l'image à recalcer avec I_R qui est l'image de référence

\tilde{T} est la transformation qui ramène les coordonnées des pixels de l'image I_D dans le système de coordonnées de l'image I_R .

Les valeurs des paramètres de cette transformation sont actualisées par une méthode d'optimisation, "argopt", qui cherche à maximiser ou minimiser une mesure de similarité sim calculée à partir de primitives images détectées par une fonction g , ou directement avec les niveaux de gris des images; dans ce dernier cas, aucune fonction g n'est appliquée. Cet algorithme d'optimisation cherche les valeurs des paramètres de la transformation dans un espace de recherche donné par T et Ω est la méthode d'interpolation associée à la transformation T qui sera appliquée à l'image I_D .

²⁹ Source: Jupeng et al "A Robust Feature-Based Method for Mosaic of the Curved Human Color Retinal Images"

Il existe plusieurs méthodes de recalage:

- 1-Méthode à base d'intensité;
- 2- Méthode à base de transformées;
- 3- Méthode à base de caractéristique de haut niveau;
- 4- Méthode à base de caractéristique bas niveau.

Cette dernière méthode utilise les primitives géométrique pour l'alignement optimal des images à recaler, dans notre cas on va s'appuyer sur cette technique afin d'évaluer la performance de l'algorithme proposé pour l'extraction des points de contrôles.

3.2.1. Approches iconiques et géométriques

L'objectif de ces approches est de trouver la correspondance entre les points deux à deux, une fois qu'ils sont déjà été détectés, , il existe deux grandes approches pour trouver cet appariement.

Approches iconiques

Les méthodes iconiques, ou méthodes de corrélation fusionnent l'étape de détection de primitives avec celle de mise en correspondance. Ces méthodes traitent les images directement sans tenter de détecter les objets saillants. Des fenêtres de tailles prédéfinis sont utilisées pour établir la correspondance. Pour comparer les deux images une mesure de similarité est utilisée telle que la corrélation croisée Normalisée (CNN)

$$CNN = \frac{\sum_X [R_I(x) - \hat{I}_R][I_D(T(x)) - \hat{I}_D]}{\sqrt{\sum_X [I_R(x) - \hat{I}_R]^2 [I_D(T(x)) - \hat{I}_D]^2}} \quad (3.2)$$

Ou

$$\hat{I}_R = \frac{1}{N} \sum_X I_R(x) \text{ et } \hat{I}_D = \frac{1}{N} \sum_X I_D(T(x)) \quad (3.3)$$

Sont les moyennes des images et N est le nombre de pixels dans les images.

Le résultat de la corrélation croisée normalisée est toujours compris entre -1 et 1. Cette mesure de similarité est souvent utilisée pour aligner des images dont lien est décrit par des translations 2D, ou des variations légères de rotation et de changement d'échelle.

Les inconvénients des méthodes iconiques sont: la sensibilité des intensités au bruit et à la variation de la luminosité. Les fenêtres rectangulaires, qui sont souvent utilisées, s'adapte aux images qui diffèrent localement et seulement par une translation. Si les images présentent des déformations avec des transformations complexes, ce type de fenêtres ne peut pas couvrir la même partie de la scène dans l'image de référence et l'image capturée.

Approche géométrique

Les méthodes géométriques sont utilisées quand la structure locale de l'image est plus significative que ses intensités. Ils permettent de recalibrer des images de nature complètement différentes (comme les photos aériennes et les images médicales) et peut traiter les distorsions complexes entre les images. Leur utilisation est recommandée si les images contiennent des objets suffisamment distinctifs et facilement détectables.

3.2.2. Méthodes utilisant les relations spatiales entre les primitives

Ces méthodes sont généralement appliquées dans le cas où les primitives détectées sont ambiguës ou si leurs voisins sont déformés localement. On exploite l'information concernant la distance entre les points de contrôle (PCs) et leur répartition dans l'espace. Les points les plus proches sont donc considérés en correspondance. Les primitives de type ligne ou région sont appariées au moyen de la minimisation de la distance généralisée entre eux. Une meilleure mesure de correspondance existe aussi qui est la transformée de distance séquentielle avec la moyenne des moindres carrés. La technique la plus connue dans ce type de méthodes est l'algorithme du plus proche point itératif (ICP- Iterative Closest Point- en anglais) introduit par Besl et McKay [40].

Description de l'algorithme ICP

L'algorithme de recalage ICP est basé sur l'appariement explicite itéré entre les entités géométriques (feature based registration), Le recalage itératif basé sur la recherche du point le plus proche est une des techniques les plus couramment utilisées pour aligner deux nuages de points en 2D ou dans l'espace 3D. Ce type d'algorithme itératif alterne les étapes d'appariement de données et l'estimation de la transformée de recalage. Le principe consiste à déterminer, pour chaque point p du premier ensemble S_1 , le point le plus proche de p dans le second ensemble S_2 , au sens d'une distance euclidienne. À partir de ces appariements, une transformation rigide minimisant un critère des moindres carrés est calculée et appliquée à S_1 .

Ce processus alterné est itéré, au bout de chaque itération l'algorithme fournit une liste de points appariés et une estimation de la transformation de repère entre les ensembles de points à recalibrer. Cette transformation est utilisée, pour l'itération suivante pour la mise à jour la liste des points appariés. Ces derniers serviront, à leur tour, pour calculer une nouvelle estimation de la transformation. Ces étapes sont répétées jusqu'à la convergence de l'algorithme. Celui ci converge lorsque l'erreur résiduelle de distance entre les points appariés est inférieure à un seuil donné.

Cet algorithme permet d'apparier plusieurs formes, De nombreuses variantes ont été proposées pour améliorer l'algorithme ICP introduit par Besl et Mc Kay. Le lecteur trouvera dans [41] une

classification et une comparaison expérimentale de ces différentes solutions, on peut distinguer trois types d'algorithme, selon les primitives géométriques utilisées:

- Algorithme ICP standard : appariement de deux nuages de points.
- Algorithme ICL4: appariement des primitives géométriques de type courbes, lignes ou contours.
- Algorithme ICS5: appariement des surfaces dans les constructions 3D.

Dans notre technique de mosaïquage, on utilise l'algorithme ICP standard qui va appairer les ensembles de points de contrôle trouvés par la méthode proposée NSCT-SIFT. On peut résumer les étapes de l'algorithme ICP comme suit:

- Soit T une estimation de la transformation initiale ;
- Soit S un seuil de convergence;
- Répéter, pour $k=1 \dots k \text{ max}$, ou jusqu'à ce que le critère soit atteint;
- Trouver un ensemble de paires de points les plus proches dans S_1 et S_2 ;
- Estimer la transformation qui minimise le critère de distance ;
- Appliquer la transformation de minimisation à tous les points S_1 ;

Parfois la distance géométrique euclidienne "classique" n'est, plus suffisante pour établir des appariements relativement corrects et nécessaires à la convergence de l'ICP. Ainsi, les données photométriques constituent une information supplémentaire intéressante pour améliorer la qualité des appariements. Dans ce cas, les points sont classés, en premier lieu, selon leurs distances photométriques, basées sur la différence de couleur, et qui ne doivent pas dépasser un certain seuil, ensuite parmi ces points qu'on appelle compatibles on calcule les points les plus proches en utilisant une distance géométrique.

Dans notre technique de mosaïquage, on utilise l'algorithme ICP standard qui va appairer les ensembles de points de contrôle trouvés par la méthode proposée NSCT-SIFT

3.3. Estimation du modèle de transformation

La définition d'une fonction de transformation doit transformer l'image capturée à l'échelle de l'image référence afin de pouvoir aligner les points de contrôles. Donc, le problème est comment choisir le modèle de la fonction de transfert et l'estimation de ses paramètres. Dans notre cas on s'intéresse aux transformations globales qui transforment tous les points de contrôle de l'image. Ils existent plusieurs modèles globaux, on cite parmi eux les suivants:

3.3.1. Transformation affine

Pour cette transformation, le changement d'échelle est asymétrique et donc, un effet de cisaillement apparaît. La transformation est définie par la matrice de taille 2×3 et donc à 6 degrés de liberté :

$$X' = [A \ t] \tilde{X} = \begin{bmatrix} a_{00} & a_{01} & t_x \\ a_{10} & a_{11} & t_y \end{bmatrix} \tilde{X} \quad (3.4)$$

$\tilde{X}=(x, y, 1)^T$ est un vecteur 2D en coordonnées homogènes

La matrice A peut être représentée par une combinaison de matrices de rotation et une matrice de facteurs d'échelle asymétriques avec les valeurs λ_1 et λ_2 [42] :

$$A = R(\theta)R(-\varphi)DR(\varphi) \quad (3.5)$$

et

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Cette transformation préserve le parallélisme et les rapports de longueurs entre les lignes ainsi que les rapports de surfaces. Ce type de transformation est très intéressant dans le cas de recalage intra-patient multimodale. Il est aussi utilisé dans le mosaïquage des images dans le cas où l'appareil est loin de la scène, et cette dernière est plate

3.3.2. Transformation projective

Une transformation projective est une application entre deux espaces projectifs qui préserve la structure projective. Cette transformation est aussi appelée transformation perspective. C'est un modèle de transformation qui implique des mouvements de translation, de rotations, de changements d'échelle, elle peut être exprimée comme suit:

$$X' = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \tilde{X} \quad (3.6)$$

où H représente la matrice de transformation projective et peut être décomposé en une chaîne de transformations [42].

$$H = H_s H_A H_p = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} I & t \\ V^T & V \end{bmatrix} = \begin{bmatrix} A & t \\ V^T & V \end{bmatrix} \quad (3.7)$$

$$A=sRK+tv^T \text{ et } K \text{ est une matrice triangulaire dont } \det(K)=1.$$

R représente matrice de rotation, s est un facteur d'échelle, t est le vecteur de translation et $v=(v_1, v_2)^T$ contient les valeurs (h_{20}, h_{21}) de la matrice de transformation projective .

Ce modèle décrit exactement la déformation d'une scène plate photographiée par un appareil avec sténopé, dont l'axe optique n'est pas perpendiculaire sur la scène. Il est généralement utilisé dans le mosaïquage des images.

3.3.3. Interpolation des niveaux de gris des images transformées

La transformation des coordonnées d'origines discrètes (entières) peut donner des valeurs non-entières dans le résultat du calcul. La solution est d'affecter le résultat au pixel le plus proche ne résout pas le problème de la construction de l'image résultat; en effet certains pixels auront deux valeurs différentes, alors que d'autres n'auront aucune valeur.

Il y a deux méthodes pour résoudre ce problème, la méthode directe qui présente des inconvénients car elle fait apparaître des discontinuités et des trous de niveaux de gris dans l'image transformée. On préfère utiliser la méthode inverse est cette méthode, une fois l'extension de l'image transformée connue, on applique la transformation inverse T^{-1} à toutes les coordonnées entières des pixels de l'image transformée pour calculer, pour chaque pixel, le niveau de gris correspondant à cette position, dans l'image originale [43].

Comme il est probable que cette position ne soit pas entière, il faut alors calculer la valeur de ce pixel par interpolation des valeurs des pixels voisins. Le principe de la méthode inverse est montrée dans la figure 3.2. Cette méthode garantit l'absence de trous dans l'image transformée.

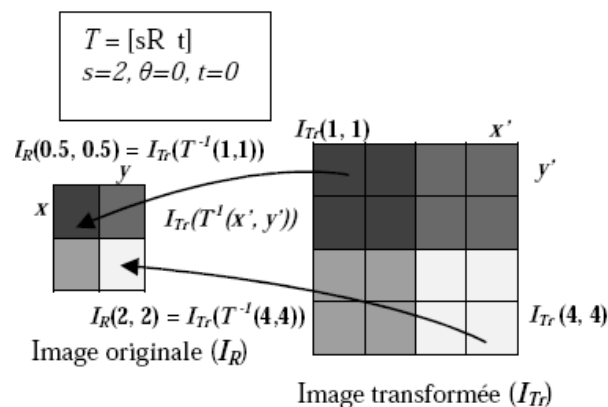


Figure 3.2 : Illustration de l'approche directe I_R est l'image originale et I_{Tr} est l'image transformée. T est la transformation appliquée aux coordonnées des images.

La transformation prise comme exemple ici est une similarité qui va multiplier par deux la taille de l'image ($s = 2$)³⁰.

Pour calculer la valeur du pixel dans l'image transformée, une méthode d'interpolation qui prend en compte les valeurs d'intensité des pixels dans le voisinage du pixel correspondant dans l'image source est utilisée, les plus connues étant les interpolations par le plus proche voisin, bilinéaire et bicubique. Généralement l'interpolation bilinéaire est simple à calculer et donne de bons résultats dans l'ensemble.

³⁰ Thèse Construction rapide d'images panoramiques applicables à l'exploration cystoscopique et à l'endoscopie de fluorescence en cancérologie" de Yahir HERNANDEZ MIER 2007

Interpolation bilinéaire

L'intensité dans un point est déterminée à partir de la somme pondérée des intensités aux quatre pixels les plus proches, par l'équation suivante

$$X_c = (1 - b) \cdot (1 - a) \cdot X_{i,j} + a \cdot (1 - b) \cdot X_{i+1,j} + b \cdot (1 - a) \cdot X_{i,j+1} + a \cdot b \cdot X_{i+1,j+1} \quad (3.8)$$

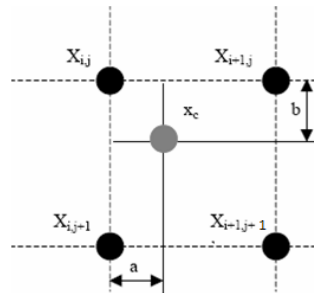


Figure 3.3 : Estimation de l'intensité X_c à partir des intensités des pixels voisins par interpolation bilinéaire.

3.4. Méthode proposée basée sur NSCT SIFT

Dans notre algorithme de recalage qui est l'étape clé du mosaïquage, on a combiné la NSCT servant au rehaussement des contours [20], puis on a appliqué la transformée SIFT pour extraire les points de contrôle avec une grande précision. Autres méthodes combinent la transformée SIFT avec la transformée contourlets pour affiner les choix des points à appairer servants au recalage des images.

Un débat est soulevé autour du choix judicieux de nombre des niveaux et des directions de la décomposition contourlets NSCT, d'une part certaines études optent pour le choix de quatre (4) niveaux de décomposition avec quatre (4) directions [44], d'autres décrivent que le changement du nombre d'orientations dans le même niveau de la décomposition NSCT a un faible effet à la fois sur le spectre et la qualité spatiale. La qualité spectrale et spatiale des images affilées PAN dépendent principalement du nombre de niveaux de décomposition. Trop peu de niveaux de décomposition entraînent une mauvaise qualité spatiale, alors que des niveaux excessifs de la décomposition entraînent une mauvaise qualité spectrale.

Mohamed R et al [45] stipule que en augmentant le nombre de niveaux de décomposition au-delà de 3 provoque plus de distorsion spectrale et spatiale et nécessite plus de temps de calcul et de mémoire et que la meilleure qualité spatiale est obtenue dans le cas de 2 niveaux de décomposition et 4 orientations dans chaque niveau.

De notre part; une phase préliminaire consistait sur un travail de simulation basé sur une approche empirique peut nous amener vers le choix adéquat des niveaux de décomposition et le nombre directions.

La première étape consistait à faire varier le nombre de niveaux de décomposition pour un nombre de directions fixe et qui vaut quatre appliquée à l'image Ikonos orientée de 10° (voir tableau 3.1). Le tableau 3.1 liste les valeurs des paramètres d'évaluation du recalage à savoir le RMSE, le coefficient de corrélation CC (voir les équations 3.9 et 3.10) et le temps de calcul. Les résultats obtenus montrent la meilleure performance du choix du nombre de niveaux de décomposition qui est égale à trois (faible RMSE et valeur importante de CC). Bien évidemment, le temps de calcul est plus important que celui de la décomposition à deux niveaux mais le choix du nombre de trois niveaux demeure un meilleur compromis.

Avec le coefficient de corrélation CC entre deux images et la racine carrée de l'erreur moyenne donné par :

$$CC = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} ((X_{ij} - m_x) \times (Y_{ij} - m_y))}{\sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (X_{ij} - m_x)^2 \times \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (Y_{ij} - m_y)^2}} \quad (3.9)$$

$$RMSE = \sqrt{\frac{1}{N \times N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (X_{ij} - Y_{ij})^2} \quad (3.10)$$

Où X et Y représentent des images avec la taille NxN et X_{ij} et Y_{ij} représentent respectivement les valeurs des pixels des images X et Y.

Nombres de niveaux de décomposition	RMSE	CC	Temps de calcul (secondes)
2	0.0618	0.9318	66.477
3	0.0393	0.9731	68.449
4	0.0628	0.9224	79.153
5	0.0473	0.9606	93.387

Tableau 3.1: Les valeurs des paramètres d'évaluation du recalage (RMSE, CC et le temps de calcul) pour différents nombres de niveaux de décomposition de l'image Ikonos orientée de 10° .

Une fois que le niveau de décomposition est fixé, on a cherché à déterminer le nombre de directions optimal pour trouver la transformée du recalage en suivant la même procédure qu'auparavant voir tableau 3.2. Les résultats obtenus montrent la meilleure performance du choix

du nombre de directions de la transformée NSCT qui est égale à quatre (de même faible RMSE et valeur importante de CC).

Nombre de directions	RMSE	CC	Temps de calcul (secondes)
2	0.1257	0.7170	39.417
4	0.0393	0.9731	68.449
6	0.0580	0.9401	167.288
8	0.0823	0.8776	372.629

Tableau 3.2: Les valeurs des paramètres d'évaluation du recalage (RMSE, CC et le temps de calcul) pour différents nombres de directions de la transformée NSCT appliquée à l'image Ikonos orientée de 10°

De ce qui précède on a opté pour le choix du couple formé de trois niveaux de décomposition et quatre directions pour la transformée NSCT le long de notre travail

La méthode d'extraction des points clé proposée par A. ghaz et al [46] est résumée selon les étapes suivantes :

1. On applique une décomposition NSCT sur l'image angiographique en considérant trois (3) niveaux de décomposition et quatre (4) directions.
2. On calcule le maximum entre les quatre (4) sous-bandes selon l'équation (3.11), cette opération est appelée "maxima edge".

$$S_{ij} = \arg \max(|\rho|)_{\rho \in [B_1^n(i,j), B_2^n(i,j), B_3^n(i,j), B_4^n(i,j)]} \quad (3.11)$$

Avec B_k représentant les différentes directions des sous-bandes

$K=1, 2, 3, 4$.

n : indique le niveau d'échelle.

3. A chaque niveau, la sortie de « maxima edge » est pondérée respectivement par α , β et γ alors que l'image originale est pondérée par Θ , de telle sorte que la condition $\alpha+\beta+\gamma+\Theta=1$ soit respectée.
4. Le résultat de la carte des contours est rajouté à l'image originale pour rehausser les contours de celle-ci.

5. Enfin, la transformée SIFT est appliquée sur l'image rehaussée pour extraire les points de contrôle.

La figure 3.4 illustre le diagramme en blocs de l'algorithme NSCT-SIFT avec une décomposition à 3 niveaux et 4 directions.

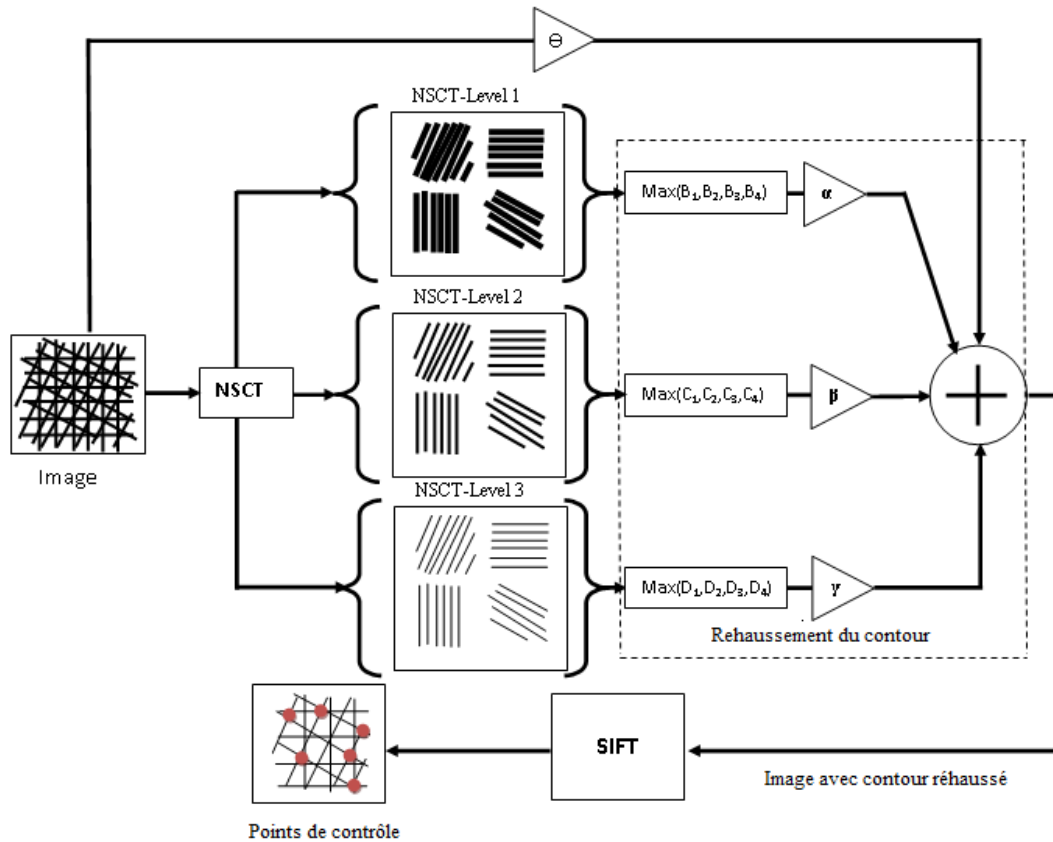


Figure 3.4 : Diagramme en blocs de l'algorithme proposé NSCT –SIFT pour l'extraction des points de contrôle

3.4.1. Calcul des coefficients de pondération

Le calcul des coefficients de pondération α , β et γ sont obtenus d'une façon expérimentale pour les trois méthodes de rehaussement des contours à savoir la sommation de tous les niveaux de décomposition, prendre le maximum en valeur absolue des différents sous bandes et la dernière méthode on calcule toujours le maximum entre les sous bandes en tenant compte des valeurs négatives.

On a recalé une image d'un angle de 7 degrés et essayé de trouver la transformée T de recalage pour ramener les deux images dans un même repère. On a calculé le coefficient de corrélation entre ces deux images en changeant à chaque fois les valeurs des coefficients de pondération de chaque sous bande, après plusieurs itérations on a trouvé:

$$\alpha = c1*(1 - \Theta), \beta = c2*(1 - \Theta) \text{ et } \gamma = c3*(1 - \Theta)$$

Une fois ces valeurs fixés, il reste à trouver le coefficient Θ propre à la pondération de l'image originale la courbe expérimentale de la figure 3.5 montre que la meilleure valeur de Θ est 0.6 et concerne la méthode maximum des sous bandes ou le coefficient de corrélation reste assez stable par rapport aux autres méthodes.

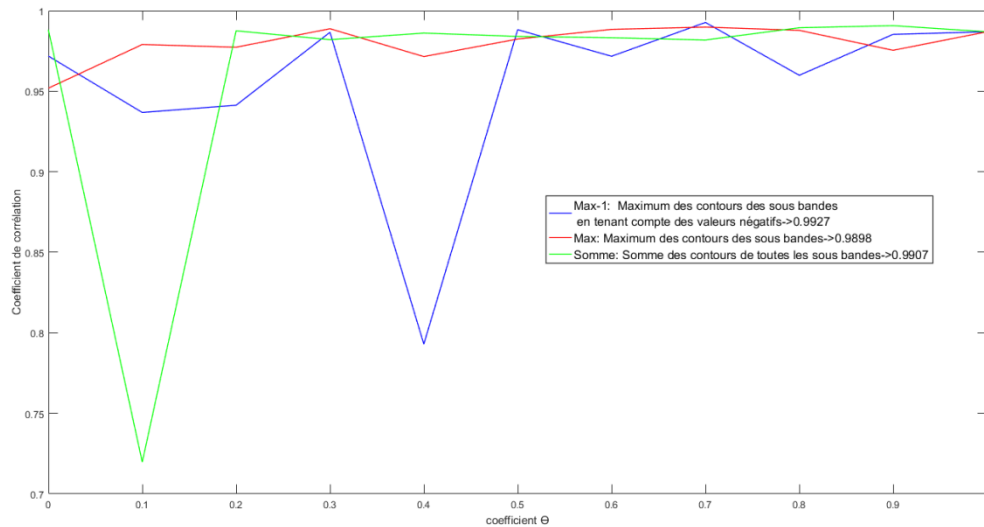


Figure 3.5 : Choix du coefficient et de la méthode de rehaussement des contours

3.5. Détection des points d'intérêt

Après avoir appliqué notre algorithme sur les images tests utilisées dans le chapitre 2, on a pu obtenir des points d'intérêts plus stables avec des localisations plus précises et diminuer ainsi les erreurs de correspondance entre les points des différentes portions des images à mosaïquer, la figure 3.6 montre les points d'intérêt des images tests comparées à ceux trouvés par le détecteur basé sur NSCT proposé par Serrief et al [30] et ceux trouvés par le détecteur SIFT de David Lowe [27].

On peut remarquer d'une part sur l'image test damier que certains points qui n'étaient détectés ni par le détecteur NSCT ni par le détecteur SIFT sont extraits par notre méthode basé sur NSCT et SIFT, d'autre part l'image test 2 montre la réduction des faux points détectés sur les deux triangles de l'image d'une façon considérable lorsque on a utilisé le détecteur NSCT-SIFT. En plus le tableau 3.3 dressé donne un pourcentage des faux points détectés plus bas par rapport aux deux autres détecteurs.

On remarque que le détecteur NSCT-SIFT donne une bonne précision de localisation des points d'intérêt par rapport au détecteur SIFT cela peut être visualisé sur l'image test 2 de la figure 3.6.c au niveau de la partie gauche représentant une dégradation des niveaux de gris, on voit bien que les

points de contrôle se trouvent sur les coins des rectangles alors qu'elles sont largement décalés sur l'image test 2 de la figure 3.6.b (points extraits par le détecteur SIFT).

On s'est intéressé à comparer rien que les points détectés par le détecteur NSCT-SIFT et le détecteur SIFT, car ceux détectés par la transformée NSCT donnent des faux points qui ne représentent ni des jonctions en T ou X,V,L,Y ce qui augmente la probabilité de la fausse correspondance qui est une étape essentielle pour le calcul de la transformée de recalage.

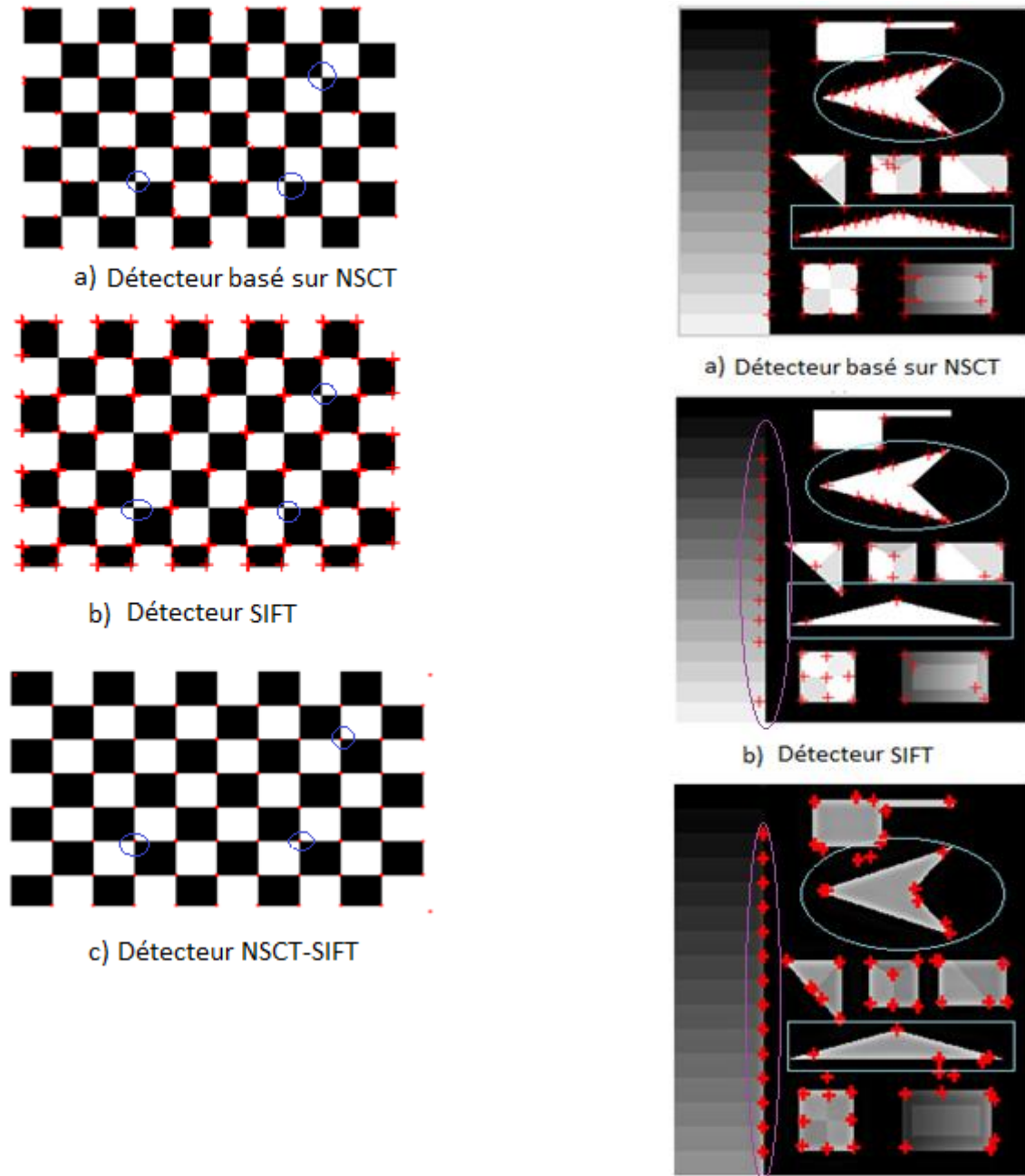


Figure 3.6 : Comparaison des trois détecteurs selon le nombre et la précision de localisation des points de contrôle

	Nombres de points détectés par NSCT		Nombres des points détectés par SIFT		Nombres des points détectés par NSCT- SIFT	
	faux+vrai	vrai	faux+vrai	vrai	faux+vrai	vrai
Image test 1	94	59	77	59	69	69
Image test 2	89	52	60	40	68	52
Faux points clé image test 1 en %	37.23 %		23.37 %		0 %	
Faux points clé image test 2 en %	41.57 %		33.33 %		23.52 %	

Tableau 3.3 : Comparaison des détecteurs SIFT et NSCT avec la méthode proposée NSCT-SIFT selon le pourcentage des faux points détectés.

On désigne par faux points, les points que le détecteur les déclare, par erreur, comme points d'intérêt (faux positif) et par les vrais points, ceux qui représentent vraiment des points de contrôle (vrai positif).

3.5.1. Recalage des images après extraction des points de contrôles

Des essais expérimentaux ont été réalisés en utilisant les images Lena et bateaux. Dans ces premières expériences, les images transformées ou les images d'entrées à corriger sont tout simplement les images de références en rotation avec un angle de 7 degrés Les images recalées sont ré-échantillonnées en utilisant une interpolation bilinéaire.

La figure 3.7 illustre les images recalées obtenues après calcul de la transformation géométrique en utilisant le détecteur basé sur NSCT-SIFT et le détecteur SIFT. La paire d'images à comparer est représenté avec une taille de 256×256 pixels.

Il est à noter que la méthode NSCT-SIFT donne un meilleur PSNR par rapport à la méthode SIFT après recalage.

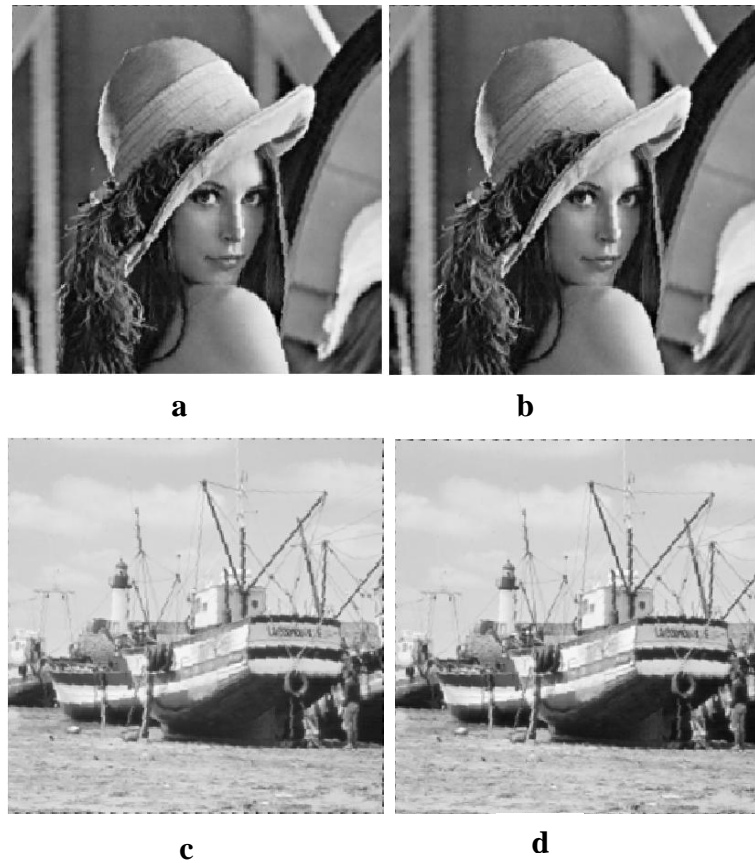


Figure 3.7 : Images recalées LENA et BOAT **(a)** Utilisant la méthode SIFT, (PSNR = 34.1389) **(b)** Utilisant la méthode NSCT-SIFT, (PSNR = 38.5915) **(c)** Utilisant la méthode SIFT, (PSNR = 41.3492) **(d)** Utilisant la méthode NSCT-SIFT méthode, (PSNR = 42.8109)

Pour illustrer la robustesse et la précision de notre méthode de recalage on a appliqué une rotation de 10 degrés sur l'image référence avec un bruit gaussien [47] [48] de valeur moyenne 0 et de variance 0.02. Les résultats expérimentaux montrent que les points extraits par la méthode NSCT-SIFT donnent un meilleur recalage par rapport à ceux obtenus par la méthode SIFT ou on peut voir la duplication de certaines structures après superposition de l'image originale avec celle recalée comme il est illustré sur la figure 3.8

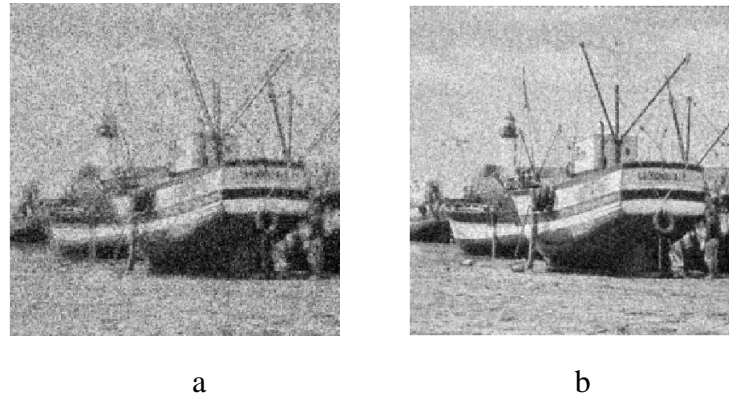


Figure 3.8 : Image recalée superposée à l'image

a) Avec la méthode SIFT b) Avec la méthode NSCT-SIFT.

Puis on a testé notre méthode avec des différentes variances de bruit, pour chaque valeur de la variance, on prend 10 itérations et à chaque fois on calcule le PSNR puis on prend la moyenne des PSNRs trouvés. La figure 3.9 représente les valeurs de PSNR obtenues pour différentes valeurs de la variance utilisant les méthodes SIFT et NSCT-SIFT (notre méthode), cette figure montre sans aucune ambiguïté que l'algorithme proposé est plus précis et plus efficace que l'algorithme SIFT. Le meilleur PSNR est toujours obtenu pour l'algorithme NSCT-SIFT proposé.

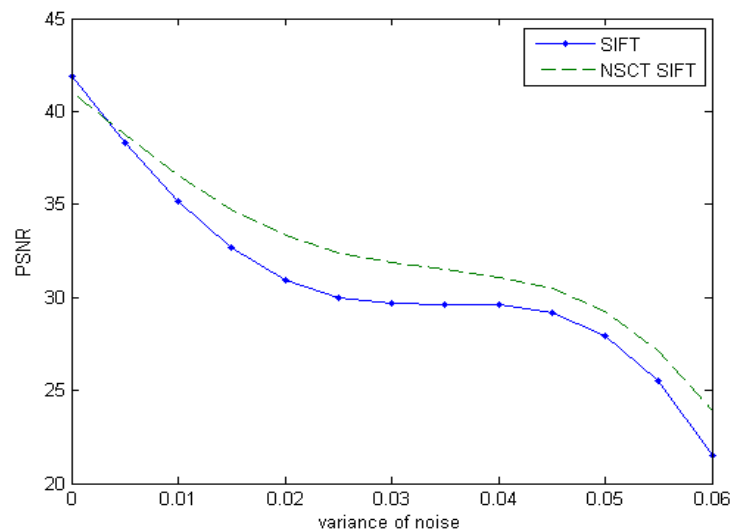


Figure 3.9 : Comparaison du PSNR après recalage entre les transformées SIIFT et NSCT-SIFT

Pour tester l'efficacité de notre technique nous l'avons comparée avec celles proposées par Meskine et al [49] pour différentes images satellitaires comme SPOT et IKONOS et aussi image radar, les paires d'images comparées sont montrées avec une taille de 256 x 256 et sont simplement recalées de 10 degrés.

Dans le travail de Meskine et al [49], elle propose d'utiliser deux techniques pour accomplir le recalage à décomposition multi-résolution des images afin de réduire et optimiser de l'espace de recherche des algorithmes génétiques (GAs). La décomposition multi-résolution de l'image emploie deux transformations : ondelettes et NSCT. Lors de l'utilisation de la transformée en ondelettes, l'image de référence et une image d'entrée ou transformée qui doivent être corrigées sont tout d'abord décomposés en utilisant une décomposition multi-résolution par ondelettes. Pour chaque niveau de décomposition, trois images sont obtenus, à savoir LL, LH et HL. La sous-bande HH inclut le bruit haute fréquence qui affecte l'image correspondant et n'est donc pas utile pour le recalage. Les couples de sous-bande correspondant sur les deux images sont comparées à l'aide de GAs. Le but est de maximiser le coefficient de corrélation entre les deux images, afin de trouver les meilleurs paramètres de la transformation (R, X, Y). A chaque niveau de décomposition, la recherche se concentre dans l'intervalle autour de la "meilleure" transformation trouvée au niveau précédent et s'affine au niveau suivant travaillant par itération de niveau le plus profond de décomposition (où la taille de l'image est la plus petite) au niveau supérieur de décomposition, c'est à dire allant de grossière à fine résolution spatiale. En d'autres termes, les paramètres trouvés dans le niveau "L" servent à estimer le nouvel espace de recherche de GAs de niveau "L-1" et à minimiser la taille de la population afin de réduire l'espace de recherche. L'exactitude de cette recherche augmente lors du passage de la résolution grossière à une résolution fine. Le processus de recalage se termine lorsque les critères de correspondance est optimisé au plus haut niveau de résolution. À résolution fine, l'image corrigée est reconstruite.

La deuxième décomposition multi-résolution utilisée par Meskine est la NSCT ou elle propose deux méthodes de recalage. Dans ces deux méthodes, elle procède de la même façon comme celle utilisant la transformée en ondelettes, mais elle applique une décomposition NSCT qui a plus de directions ou de sous-bandes. Dans la première méthode proposée NSCT1, les deux images à comparer sont décomposées en plusieurs niveaux avec des directions différentes. Les résultats de la décomposition sont une sous-bande LL, qui est l'approximation de l'image originale et des sous-bandes avec différentes directions. A chaque niveau de décomposition, le rapport de corrélation entre les images de sous-bandes correspondantes des images de référence et l'entrée est calculé successivement et maximisé à l'aide de GAs. Elle corrige les sous-bandes de l'image d'entrée avec les paramètres de transformation optimale (R, X, Y) trouvées lors du processus d'exécution à ce niveau. Ces paramètres obtenus sont utilisés pour affiner l'espace de recherche du niveau inférieur. le passage d'un niveau à l'autre se fait selon deux critères : concentré l'espace de recherche autour des valeurs optimales trouvées dans le niveau précédent, puis adapter au mieux la taille de la population en la réduisant au minimum. De cette façon, la complexité du processus de raffinement

à travers différents niveaux est vraiment réduite, tandis que la précision du recalage est augmentée. Il s'agit en effet d'un GA adaptable. Enfin, l'image corrigée est reconstruit en pleine résolution avec toutes les différentes sous-bandes corrigées.

Une deuxième méthode NSCT2 consiste à appliquer le processus de GAs sur seulement la sous-bande LL. A chaque niveau de décomposition, les paramètres trouvées par GAs sont utilisés pour corriger les directions ou les sous bandes de ce niveau. Le processus de changement de niveau se fait de la même manière que celui de la dynamique et l'adaptabilité dans la première méthode. En pleine résolution, elle reconstruit l'image recalée avec les images de la sous-bande corrigée de toutes les directions.

Il est à noter que la méthode NSCT-SIFT donne le meilleur PSNR par rapport à celui obtenu par la méthode SIFT après recalage comme illustré par la figure 3.10



Figure 3.10 : Image IKONOS recalée utilisant

a) La méthode SIFT, PSNR =35.1609 **b)** La méthode NSCT-SIFT, PSNR =38.5511

Les différents résultats analytiques des méthodes de recalages sur toutes les séries d'images sont récapitulés dans le tableau 3.4. Deux mesures sont considérées pour déterminer la précision du recalage, le coefficient de corrélation et la racine de l'erreur quadratique moyenne. Les résultats sont comparés à ceux obtenus dans [49].

		Techniques				
		Wavelet	NSCT1	NSCT2	SIFT	NSCT-SIFT
SPOT	C.C	0.8298	0.9863	0.9840	0.9908	0.9914
	RMSE	0.0462	0.0137	0.0147	0.0111	0.0108
IKONOS	C.C	0.6935	0.9562	0.9468	0.9627	0.9731
	RMSE	0.1245	0.0504	0.0549	0.0461	0.0393
SAR	CC	0.5622	0.9474	0.9294	0.9517	0.9569
	RMSE	0.1549	0.0608	0.0703	0.0520	0.0491

Tableau 3.4 : Les différents résultats analytiques des méthodes de recalages

Le tableau 3.4 montre clairement que les deux méthodes proposées SIFT et NSCT-SIFT donnent de meilleurs résultats en termes de corrélation et RMSE comparées à celles utilisant les ondelettes, NSCT1 et NSCT2 proposées par Meskine.

3.6. Application de l'algorithme NSCT- SIFT sur les images médicales

3.6.1. Précision de recalage

Après avoir testé des images satellitaires et des images bruités, on appliqué notre algorithme sur des images cystoscopiques ayants subi des différentes rotations et qu'on a essayé de les recalcr avec des transformations affines et projectives, Les différents résultats des méthodes de recalages pour la transformée affine et projective sur les images sont enregistré respectivement dans les tableaux 3.5 et 3.6, en prenant comme critère de mesure de la qualité de recalage le PSNR entre l'image référence et celle recalée.

Angle de rotation en degré	Points d'intérêt détectés par SIFT		Points d'intérêt détectés par NSCT-SIFT	
	Recalage affine " angle de rotation trouvée"	PSNR dB	Recalage affine " angle de rotation trouvée"	PSNR dB
10	9.911	47.5760	9.940	51.8276
20	20.273	53.3249	19.998	53.5889
30	29.851	51.9443	29.940	56.6915
40	39.981	54.7984	39.995	55.5104
50	49.908	48.3071	49.865	52.5132

Tableau 3.5 : Les résultats comparatifs entre recalages affines basé sur SIFT et NSCT-SIFT

Angle de rotation en degré	Points d'intérêt détectés par SIFT		Points d'intérêt détectés par NSCT-SIFT	
	Recalage projectif "angle de rotation trouvée"	PSNR dB	Recalage projectif "angle de rotation trouvée"	PSNR dB
10	10.078	47.5760	10.028	50.5787
20	19.780	53.3249	20.007	53.4376
30	29.946	51.9443	29.9874	54.7088
40	40.012	54.7984	39.9230	55.4518
50	50.064	48.3071	50.023	53.5312

Tableau 3.6 : Les résultats comparatifs entre recalages projectifs basé sur SIFT et NSCT-SIFT

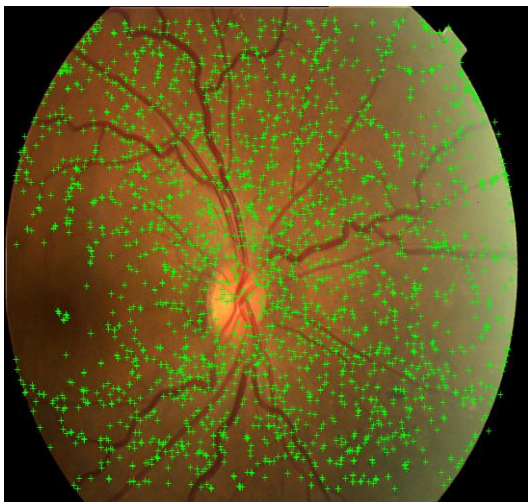
Les deux tableaux 3.5 et 3.6 montrent que le PSNR est toujours meilleur pour les recalages utilisant des points d'intérêts extraits par la méthode NSCT-SIFT, que ce soit pour la transformation affine ou projective et que les angles récupérés sont plus proches des angles initialement subits par les images d'origines, on peut citer comme exemple pour l'angle de rotation

20 degrés, on a trouvé 19.78 degrés en utilisant un recalage projectif basé sur SIFT et 20.007 degrés en utilisant un recalage projectif basé sur NSCT- SIFT, cette précision de recalage, nous permettra de construire une vue panoramique de la paroi interne de la vessie sans décalage trop visible de l'arbre vasculaire irrigant le tissu vésical.

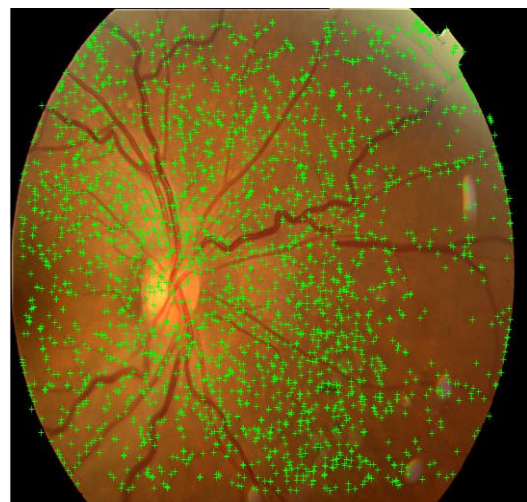
3.6.2. Appariement des points d'intérêt

L'appariement des points d'intérêt consiste à trouver une correspondance entre les mêmes points de deux images, représentant des zones communes, en comparant la distance euclidienne entre ces points à un seuil donné. Deux points sont appariés si la distance euclidienne est inférieure au seuil.

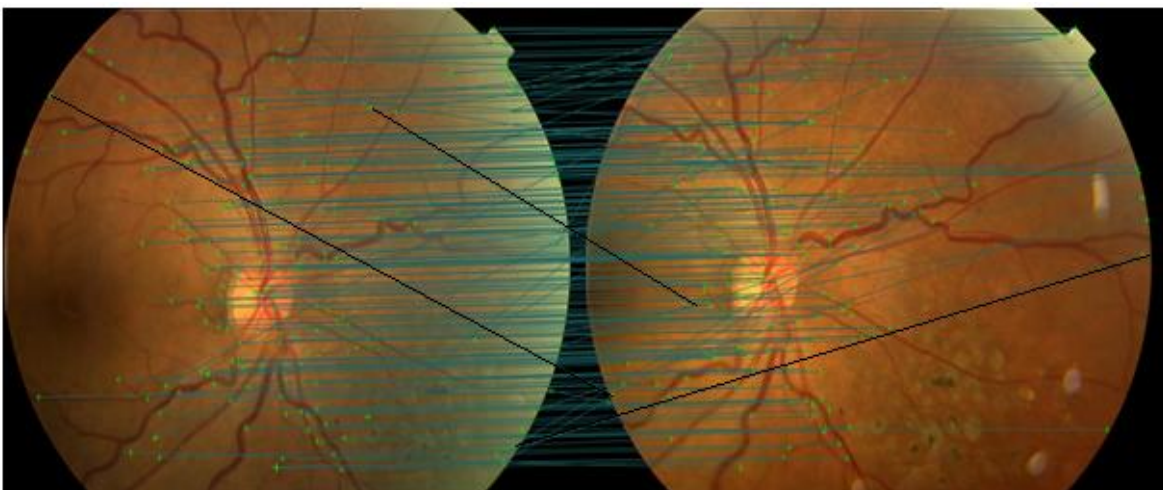
Plus le seuil est grand, plus on obtiendra un nombre élevé de points appariés vrais mais en revanche ceci risque d'augmenter la probabilité de la fausse correspondance des paires. C'est bien ce que l'on peut observer sur la figure 3.11 ci-dessous pour les seuils de 0,85 et 0,7, l'extraction des points à l'aide de l'algorithme SIFT donne 2471 points d'intérêt pour l'image de référence et 2703 points d'intérêt pour l'image à recaler.



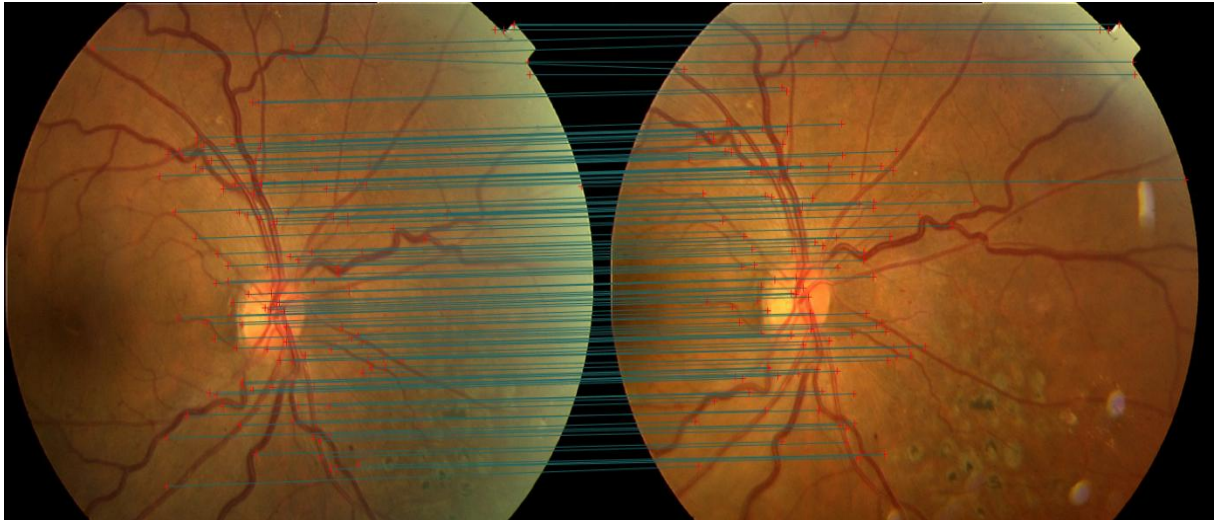
a) Image référence avec 2471 points détectés



b) Image à recaler avec 2703 points détectés



c) Appariement de 202 points d'intérêt avec un Seuil=0,85



d) Appariement de 173 points d'intérêt avec un Seuil=0,7

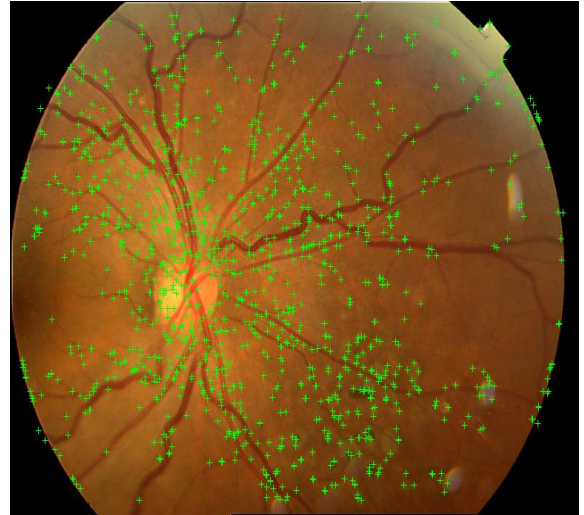
Figure 3.11 : La détection et l'appariement des points d'intérêt sur l'image référence et l'image à recalcer avec le détecteur SIFT pour deux seuils différents.

On remarque que lorsque la valeur du seuil est trop grande, la validité des paires de points devient discutable et on peut d'ailleurs observer pour le seuil de 0,85 que certains points appariés sont même complètement faux voir figure 3.11.c, alors que notre algorithme basé sur NSCT et SIFT nous a permis de réduire le nombre des points détecté ce qui a joué en faveur de la diminution de l'erreur d'appariement, on a eu que 2 faux points appariés sur 174 points pour la méthode NSCT-SIFT par rapport à 19 faux points appariés sur 202 points avec un seuil qui vaut 0,85 pour le détecteur SIFT voir figure 3.12.

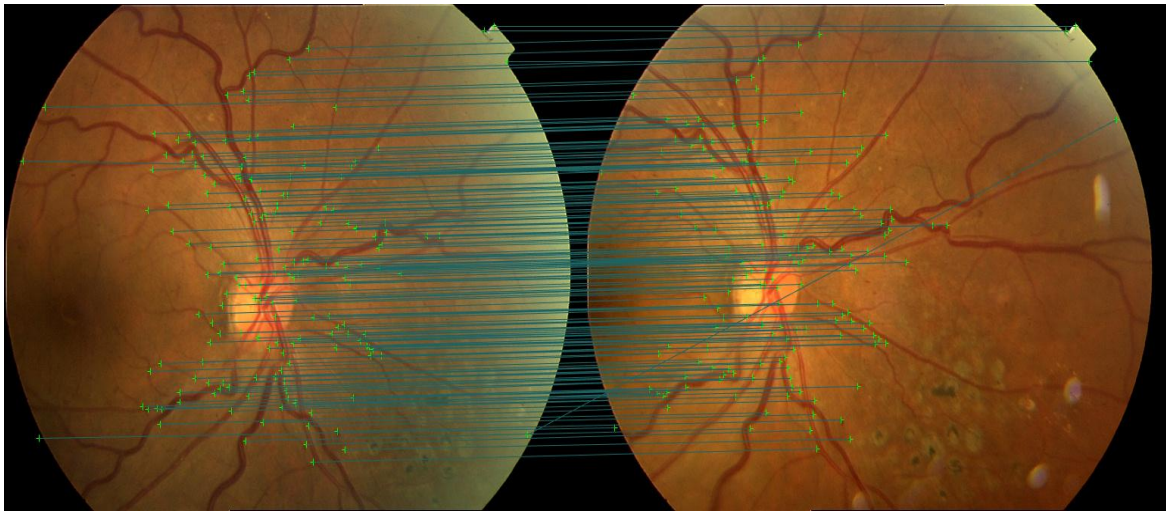
Lorsque on a pris un seuil de 0.7 on a eu 55 points appariés sans aucune erreur d'appariement, d'une autre part on a remarqué que si on voulait garder un nombre élevé des points appariés il suffisait de prendre des points qui ont la même appartenance à la caractéristique échelle ou orientation en gardant le seuil de 0.85.



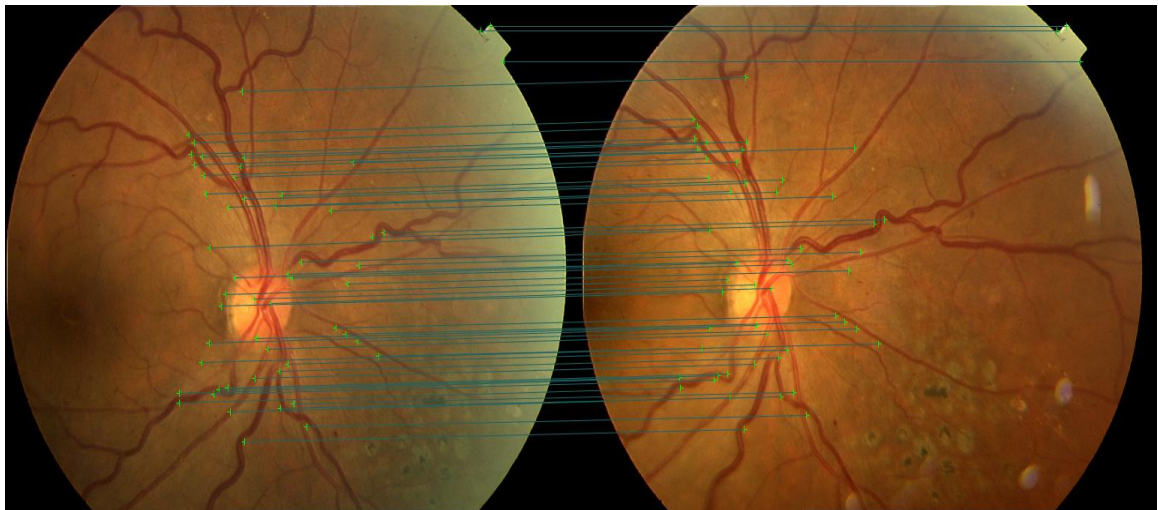
a) Image référence avec 1197 points détectés



b) Image à recaler avec 1301 points détectés



c) Appariement de 174 points d'intérêt avec un Seuil=0,85



d) Appariement de 55 points d'intérêt avec un Seuil=0,7

Figure 3.12 : La détection et l'appariement des points d'intérêt sur l'image référence et l'image à recaler avec le détecteur NSCT-SIFT pour deux seuils différents

3.7. Conclusion

Dans ce chapitre on a présenté en détail l'algorithme proposé NSCT-SIFT en montrant ses principes de fonctionnement et donnant des exemples d'extraction des points de contrôle sur des images tests et des images réelles. Cette méthode est basée sur la combinaison de deux techniques bien connues NSCT et SIFT. La NSCT est une transformation qui est capable de capturer les contours dans plusieurs directions et sur plusieurs sous bandes en gardant toujours la même taille de l'image (sans sous échantillonnage), cette transformée a servi pour le rehaussement de l'image originale, puis on a appliqué la transformée SIFT pour l'extraction des points clé.

Les résultats expérimentaux montrent que l'utilisation de SIFT avec la décomposition NSCT au troisième niveau et 4 directions donne les meilleurs résultats, concernant la localisation des points d'intérêt trouvés qui étaient plus précise et que les faux points détectés étaient aussi réduits ce qui a contribué à la diminution des erreurs d'appariement et optimiser le calcul de la transformée de recalage.

Le coefficient de corrélation et le RMSE ainsi que les angles de rotations trouvés par les transformées affine et projective illustrent une erreur minimale de recalage pour la transformée NSCT-SIFT qui est bien adaptée au recalage des images de la télédétection et les images médicales même en présence de bruit. L'application cible de ce nouveau détecteur est le mosaïquage des images agiographiques et cystoscopiques qui fera l'objet du chapitre suivant.

Chapitre 4
Mosaïquage et Blending

4.1. Introduction

L'analyse du diagnostic et le traitement du cancer de la vessie se base essentiellement sur les images fournies par l'endoscopie vésicale qui permet de localiser les lésions suspectes. L'angiographie rétinienne joue aussi un rôle important dans l'exploration et la recherche des pathologies ou anomalies anatomiques de la rétine.

Malheureusement ces deux examens nous offrent un champ de perception réduit, ce qui rend le contrôle visuel et la navigation dans l'image difficile. Le mosaïquage des images nous confère un champ de vision plus élargi résolvant ainsi la difficulté de repérage anatomique et réduisant le temps de la lecture et de la recherche des pathologies sur un cliché angiographique ou une séquence vidéo de l'endoscopie vésicale.

Après l'extraction des points d'intérêts qui ont pu donner une transformée de recalage satisfaisante on entame la procédure du mosaïquage. Cette dernière nécessite une phase de prétraitement.

4.2. Acquisition et prétraitement des images

Les images de l'angiographie rétinienne sont fournies par le service d'ophtalmologie du CHU de Sidi Bel Abbes avec une résolution de 591×873 pixels. Généralement neuf images [45] d'angiographie rétinienne avec certaines zones de chevauchement prises d'une façon ordonnées et précises sont suffisantes pour former une vue globale de la rétine.

Pour le mosaïquage des images de la vessie, les images cystoscopiques sont capturées à partir d'une séquence vidéo. Dans ce processus la paroi interne de la vessie est illuminée par une source de lumière telle que PDD (photodynamic diagnosis), les images extraites ont une résolution de 480×640 pixels.

Dans la première partie du prétraitement, on sous échantillonne les images constituantes de la séquence vidéo par un facteur de quatre pour réduire le temps de calcul et la résolution de l'image finale.

La sélection des images est faite d'une façon subjective. Des régions de 246×300 pixels centrés sur les images sont extraites permettant d'éliminer le fond noir qui ne contient aucune information.

4.3. Algorithme du mosaïquage

On utilise l'algorithme de recalage itératif ICP entre les deux nuages de points extraits par la méthode NSCT-SIFT, pour estimer les transformations géométriques qui doivent s'appliquer par la suite sur les deux images à recaler. On peut résumer l'ensemble des traitements insérés dans notre travail, comme le montre la figure 4.1 suivante:

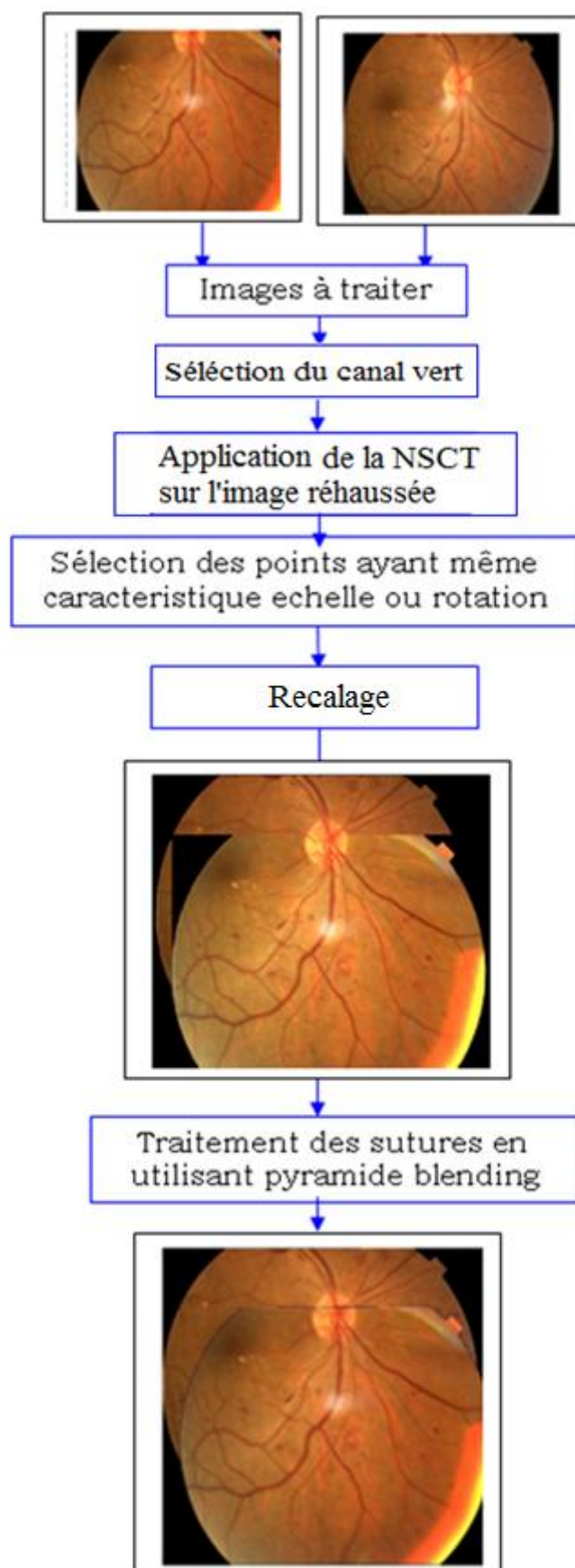


Figure 4.1 : Schéma bloc du mosaïquage

4.3.1. La sélection de l'image référence

Une fois que les correspondances entre les paires d'images sont établies l'image de référence I_{ref} est identifiée parmi les images disponibles, I_{ref} appartient à l'ensemble $\{I_1, \dots, I_N\}$, cette sélection joue un rôle important pour l'étape de mosaïquage car elle définit le système de coordonnées de base pour toutes les autres images.

L'idéal, c'est de choisir une image qui a des correspondances et des chevauchements avec l'ensemble des images à assembler mais cette image ne peut exister, donc l'idée est de calculer la surface de chevauchement de l'image I_{ref} et les autres images pour les classer et commencer le mosaïquage deux à deux. Dans le cas des images de l'angiographie rétinienne, on prend comme image de référence celle qui contient le disque optique et on cherche la surface d'intersection entre les autres images. La figure 4.2 montre la relation de l'image référence avec les autres images à mosaïquer.

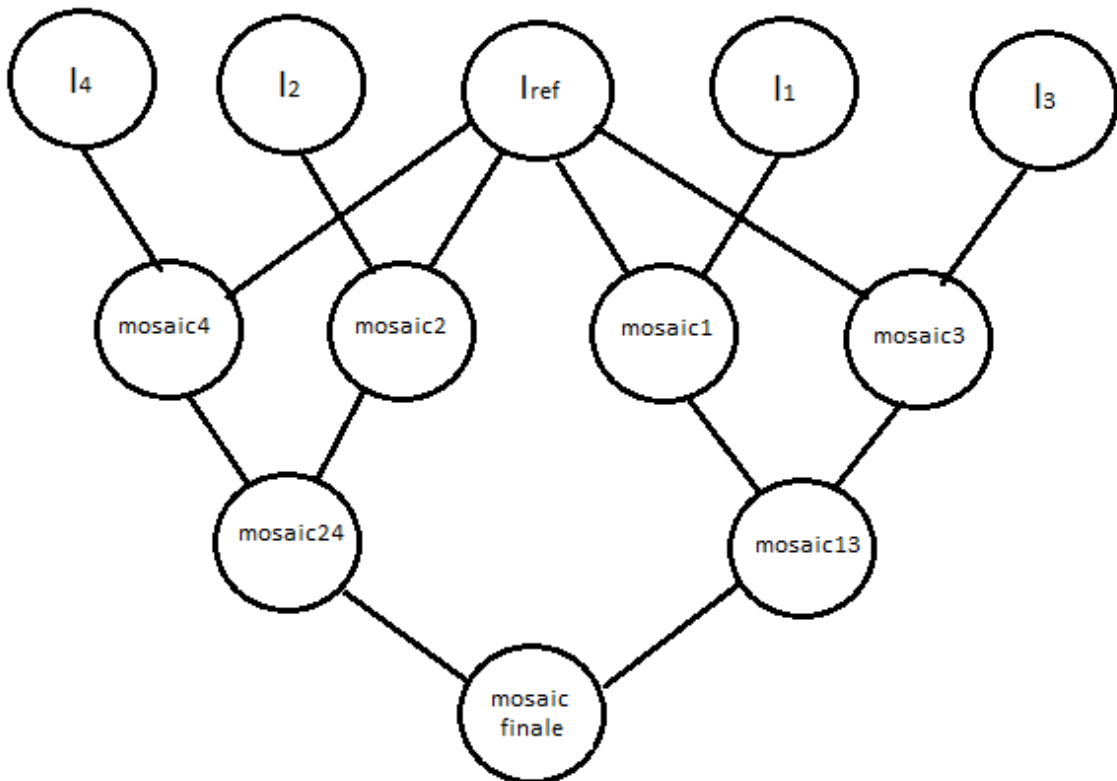


Figure 4.2 : Relation des images à mosaïquer selon la surface de chevauchement

4.3.2. Différentes méthodes du mélange (blending)

La principale préoccupation des méthodes de lissage de la transition est de produire une zone non perceptible entre deux images sur une région de chevauchement donnée. L'information de cette zone commune est fusionnée de telle sorte que les limites des images impliquées deviennent invisibles.

Dans le contexte de la photo-mosaïque aérienne à faible échelle (ordre des méga-pixels), la limitation de l'utilisation de deux images en chevauchement a été abordée en premier par Peleg [51], qui a introduit le concept de Seam-Eliminating Function SEF. Le SEF est basé sur une fonction de lissage de luminosité (c'est-à-dire une carte de pondération), obtenue à l'aide d'un algorithme itératif de relaxation coûteux, qui est utilisé pour lisser la transition d'un nombre arbitraire d'images chevauchées, en fixant les différences d'intensité le long des jointures à zéro. Le principal avantage de la méthode est que le changement progressif, lisse n'affecte pas ni le détail ni l'information utile de l'image auprès des jointures.

En 1983, Burt et Adelson [52] ont introduit le concept d'image "spline" pour obtenir une transition lisse entre plusieurs images. L'approche a eu un fort usage, étendant ses champs d'application à tout scénario d'imagerie, par opposition à Milgram [53] qui s'est concentré sur l'imagerie par satellite. C'était aussi la première approche pour la reconstruction d'image, c'est-à-dire le premier procédé capable de fusionner de façon transparente plusieurs images à partir de scènes différentes et sans rapport. Les images à fusionner sont décomposées en un ensemble de composantes passe-bandes d'images, et une spline séparée avec une largeur de transition appropriée est appliquée à chaque bande.

En 2003, Perez et al [54] ont proposé un mécanisme générique d'interpolation basé sur la résolution des équations de Poisson pour l'édition et le clonage sans jointure des régions de sélection.

Il s'agit de la première approche importante de la mosaïque d'images dans le domaine du gradient. La méthode est basée sur l'idée que, en mélangeant de manière appropriée le gradient d'une image donnée à celle d'un autre, il devient possible de fusionner de manière convaincante des régions d'image (à savoir des objets) avec un aspect transparent.

Agarwala et al.[55] ont proposé une technique en 2004 qui combine des méthodes appartenant aux deux principales classes d'algorithmes de mélange. Tout d'abord, l'optimisation de la coupe de graphe a été utilisée pour trouver l'endroit optimal pour la jointure dans la région de chevauchement. Deuxièmement, la fusion par domaine gradient a été appliquée pour réduire ou éliminer tout objet restant visible le long des jointures d'image. La méthode a plusieurs applications dans le champ du mosaïquage et réalise des résultats convaincants dans la matière.

En 2011, Szeliski et al [56] ont présenté une technique pour le mélange rapide de Poisson et la reconstitution dans le domaine de gradient qui associe, à chaque image d'entrée, une carte distincte de décalage à basse résolution, qui peut être représentée en utilisant une spline de faible dimension. Le système linéaire résultant est beaucoup plus réduit que le système original de Poisson et aussi que l'approximation de la quadtree d'une carte décalée. Puisque chacun des champs décalés est représenté à l'aide d'une spline de faible dimension, la représentation résultante est appelée multi-spline. Dans notre thèse on décrit brièvement le principe du poisson blending ainsi que le pyramid blending mais avant cette description on doit délimiter les sutures.

4.3.3. Sélection des sutures

Pour sélectionner les sutures de chaque image on crée un masque qui délimite les points à l'intérieur du cercle on utilisant la fonction d'un cercle avec la même taille de l'image originale.

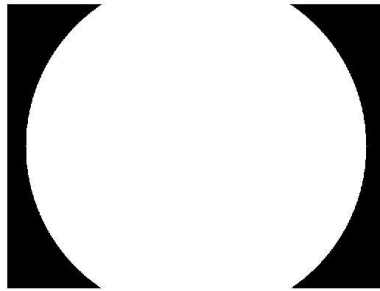


Figure 4.3 : Création du masque pour limiter les points à l'intérieur du cercle image

L'assemblage des deux masques permet de bien distinguer les frontières des sutures sur la figure 4.4 (a) qui illustre l'assemblage des masques de l'image 1 (qui est l'image de référence) et l'image 2 et la figure 4.4 (b) montre l'assemblage des masques de l'image 2 (image référence) avec l'image 1.



Figure 4.4 : Assemblage des masques des images à mosaïquer

4.4. Poisson blending

La technique du Poisson blending en traitement d'image a été introduite par P. Perez, M. Gangnet et A. Blake [54]. En résumé, elle permet de faciliter l'incrustation d'une partie d'une image A dans une image B, en calculant une solution qui détermine automatiquement les changements colorimétriques à appliquer aux pixels de A pour qu'elle s'incruste le mieux possible dans l'image B. Ceci s'obtient en résolvant l'équation de Poisson définie sur le domaine de l'image A.

Cette méthode consiste à copier des régions à partir d'une image source vers une image de la cible. Pour sélectionner des régions de l'image source qui sera utilisé, une troisième image représentant le masque troisième est nécessaire. Elle est de couleur blanche dans la région d'intérêt et noir ailleurs.

On a les deux images mosaïquées en prenant alternativement l'image 1 comme référence puis l'image 2:

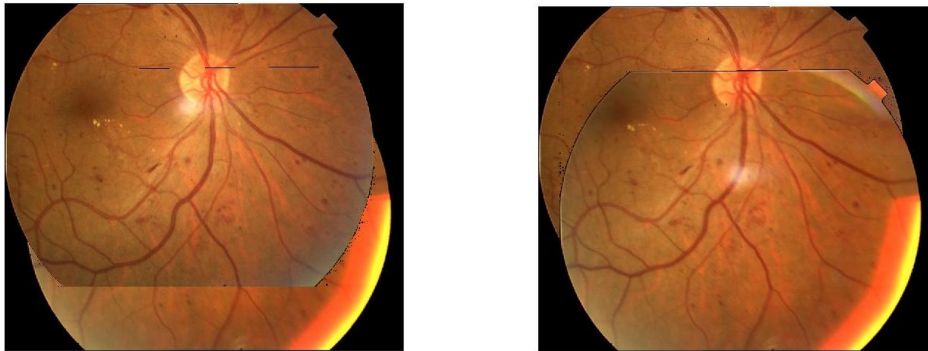


Figure 4.5 : Deux mosaïques de l'angiographie rétinienne avec sutures visibles

On calcule le masque de chauvechaument :



Figure 4.6 : Le masque de la région commune



Figure 4.7 : Résultat de poisson blending obtenu après mosaïquage

La figure 4.7 montre l'élimination presque totale des sutures qui étaient visibles sur la figure 4.5 donnant ainsi une image homogène avec un champ visuel plus élargi.

4.5. Correction des sutures par pyramid blending

Une fois que les coutures ont été mises en place et les objets non désirés ont été retirés, nous avons encore besoin de mélanger les images pour compenser les différences d'expositions et d'autres mauvais alignements.

Une solution intéressante à ce problème a été développée par Burt et Adelson [52]. Au lieu d'utiliser une simple moyenne des intensités sur la bande de recouvrement, la largeur de la bande est adaptée en fonction d'une pyramide passe-bande (pyramide Laplacienne) construite pour chaque image. Cette pyramide est une décomposition fréquentielle des intensités de l'image, réversible, et ne dégradant pas les données. Pour les basses fréquences, une large bande de la région de recouvrement est utilisée et permet de calculer le masque de poids α_i pour chaque pixel.

La bande de fusion utilisée est alors réduite pour chaque niveau, jusqu'à la plus haute fréquence (détails des images). Chaque niveau de la pyramide est alors fusionné en fonction du masque de poids calculé. La transformation inverse de la pyramide Laplacienne est appliquée et permet de reconstruire l'image fusionnée. Grâce à ce type de fusion, les basses fréquences sont moyennées sur tout le recouvrement, et les détails sur une zone plus faible.

Enfin, l'image composite est reconstruite par interpolation et sommation de tous les niveaux de la pyramide (images passe-bande). La figure 4.8 illustre le résultat obtenu à chaque niveau du pyramid blending pour la correction des sutures.

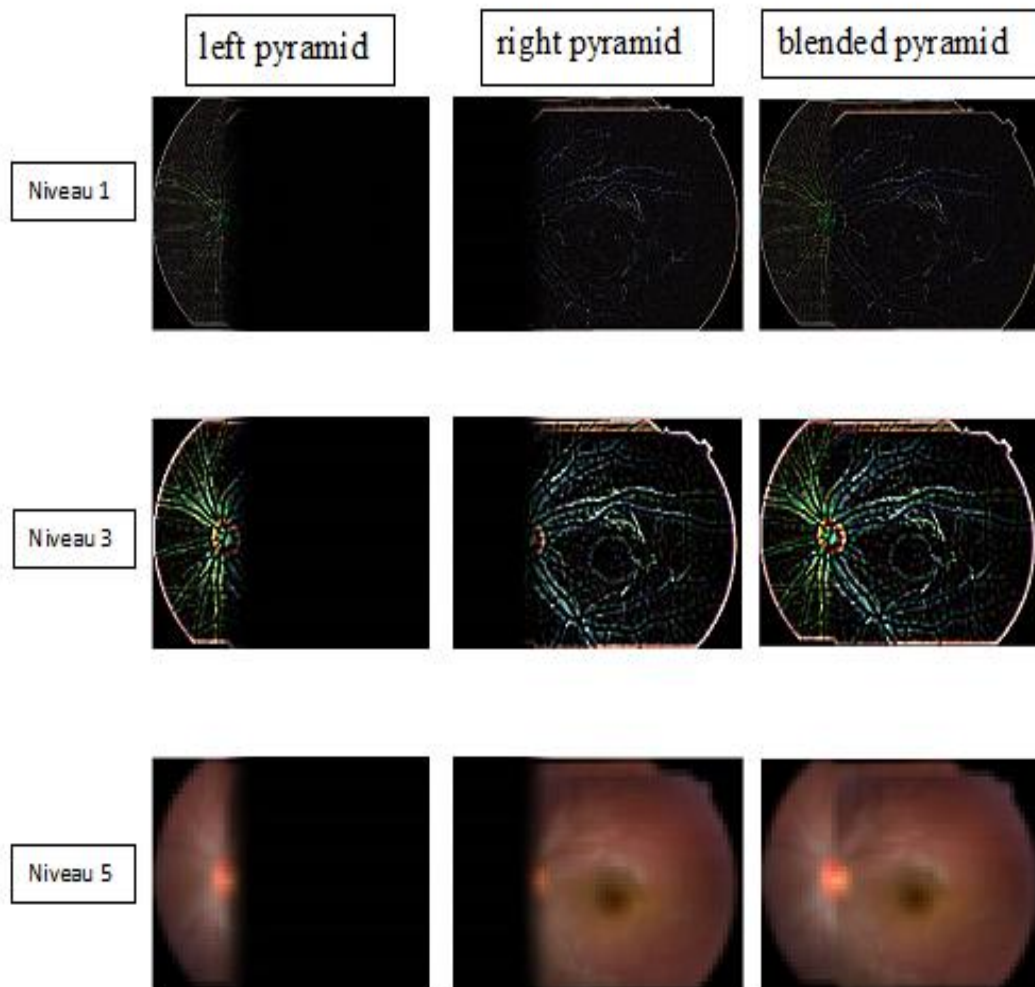


Figure 4.8 : Différents niveaux du 'pyramid blending'

4.6. Résultats du mosaïquage des images angiographiques de la rétine

Après application de l'algorithme du mosaïquage sur des images d'angiographie rétinienne fournies par le service d'ophtalmologie du CHU de Sidi Bel Abbas, on a obtenu des images panoramiques avec un champ visuel élargi. La lecture des panoramas obtenus par un ophtalmologiste a permis de diagnostiquer et interpréter les pathologies de chaque cliché en se référant au disque optique visible dans toute mosaïque, ce qui lui a permis de bien localiser les lésions de la rétine et les anomalies de l'arbre vasculaire.

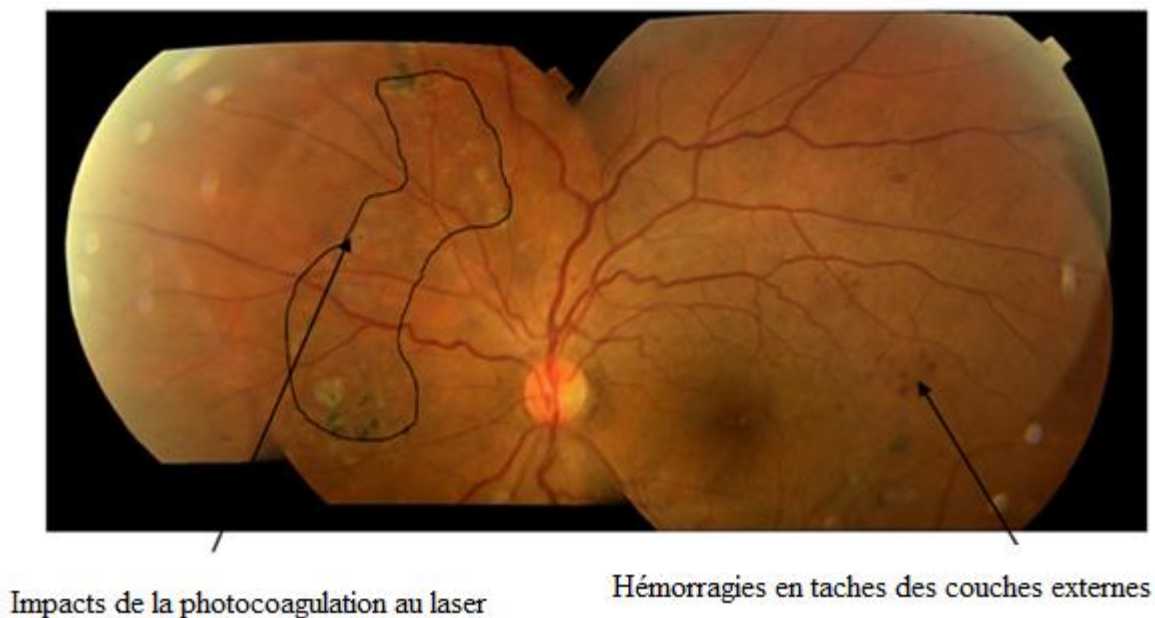


Figure 4.9 : Rétinopathie diabétique non proliférante sévère

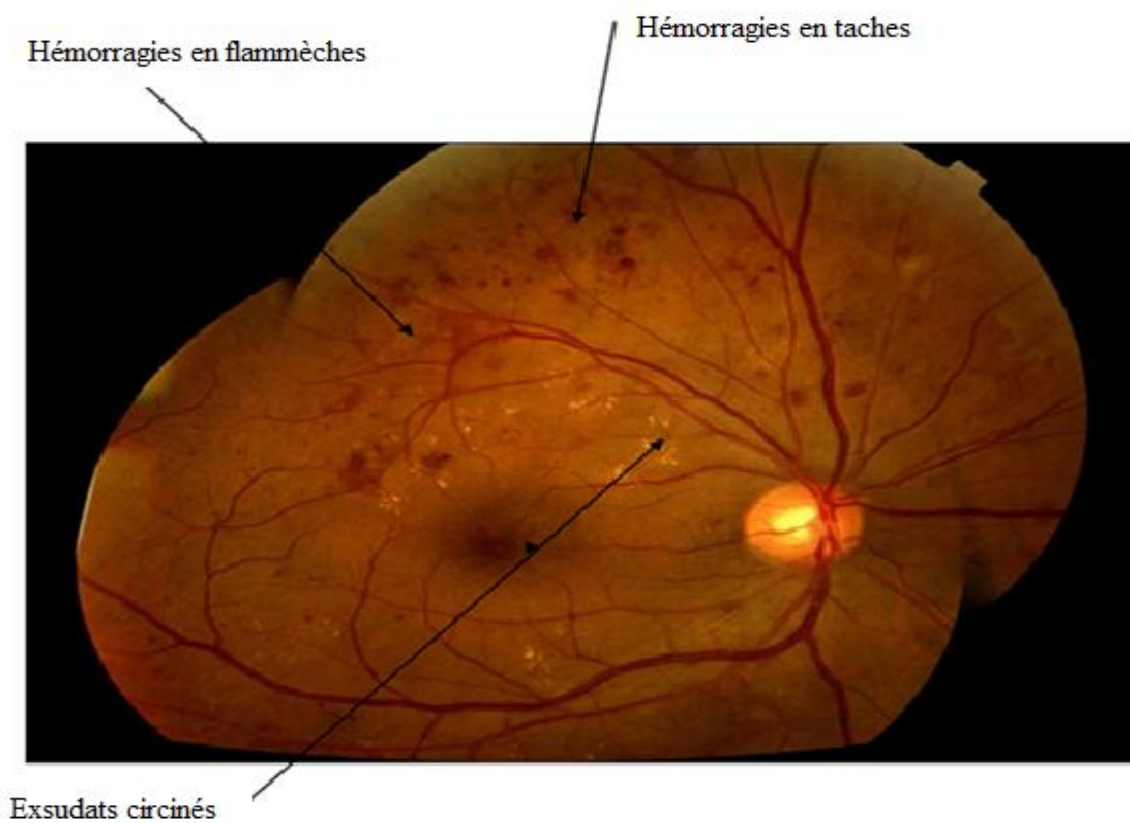


Figure 4.10 : Œdème maculaire tête flèche sur logette cystoïde



Figure 4.11 : Mosaïque formé par 8 images montrant une dilatation diffuse et tortuosité vasculaire veineuse

Anomalie vasculaire intrarétinienne avec formation de néovaisseaux

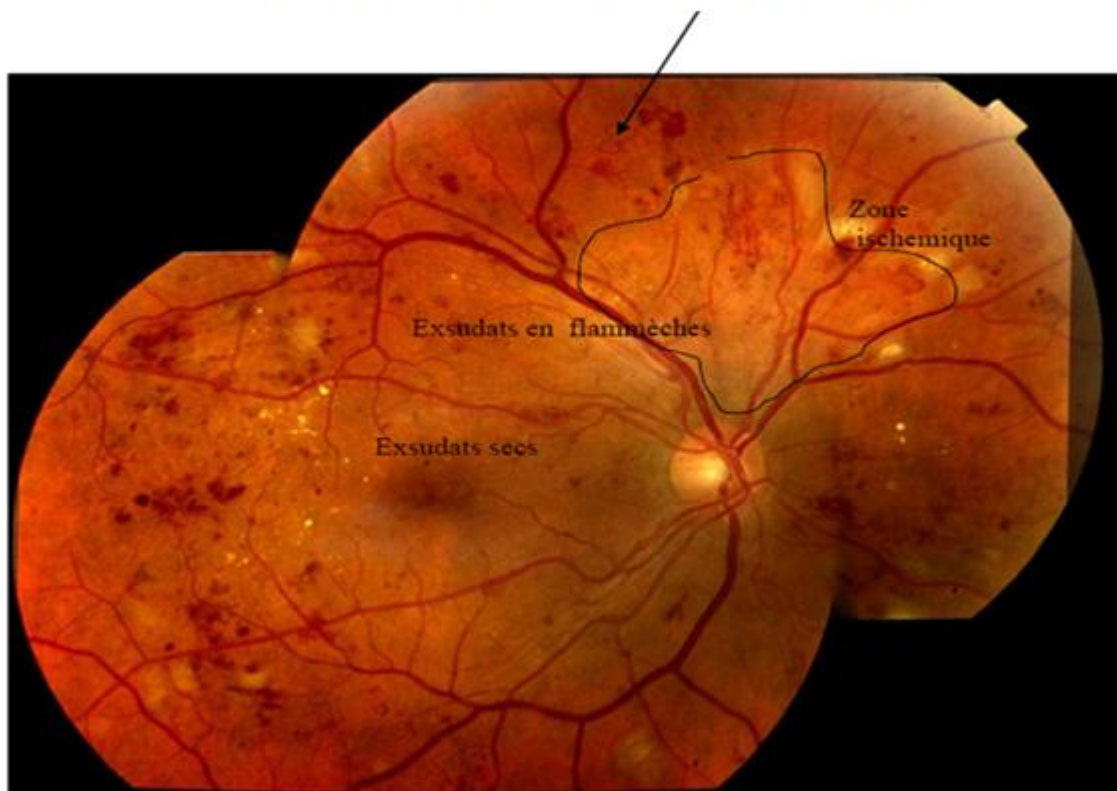


Figure 4.12 : Rétinopathie diabétique proliférante modérée

4.7. Résultats du mosaïquage des images cystoscopiques

La nécessité d'un support plus adapté pour localiser, évaluer et comparer des zones suspectes ainsi que pour préparer et guider des examens postérieurs a été exprimée par les urologues. Du point de vue du diagnostic, il est intéressant pour le clinicien de pouvoir étudier une lésion et la zone qui l'entoure dans son ensemble sans devoir chercher une image particulière ou passer d'une image à la suivante pour passer en revue la totalité de la zone.

Le fait de pouvoir observer de façon globale une zone d'intérêt doit permettre de mieux localiser des petites lésions satellites (multifocales) potentiellement à l'origine de récurrences. Une représentation globale issue d'une procédure de mosaïquage permet ce genre d'observation.

En cystoscopie classique (en lumière blanche), les tumeurs macroscopiques peuvent être détectées et donc traitées, c'est beaucoup moins facile pour les lésions dysplasiques et de type CIS.

C'est la raison pour laquelle des biopsies « aléatoires » de la muqueuse vésicale sont effectuées.

Les dysplasies « invisibles » sont majoritairement à l'origine de nombreuses récurrences tumorales qui auraient pu être évitées si elles avaient été diagnostiquées précocement et ainsi détruites. A la suite on va visualiser quelques panoramas obtenus et essayer de donner certaines observations.

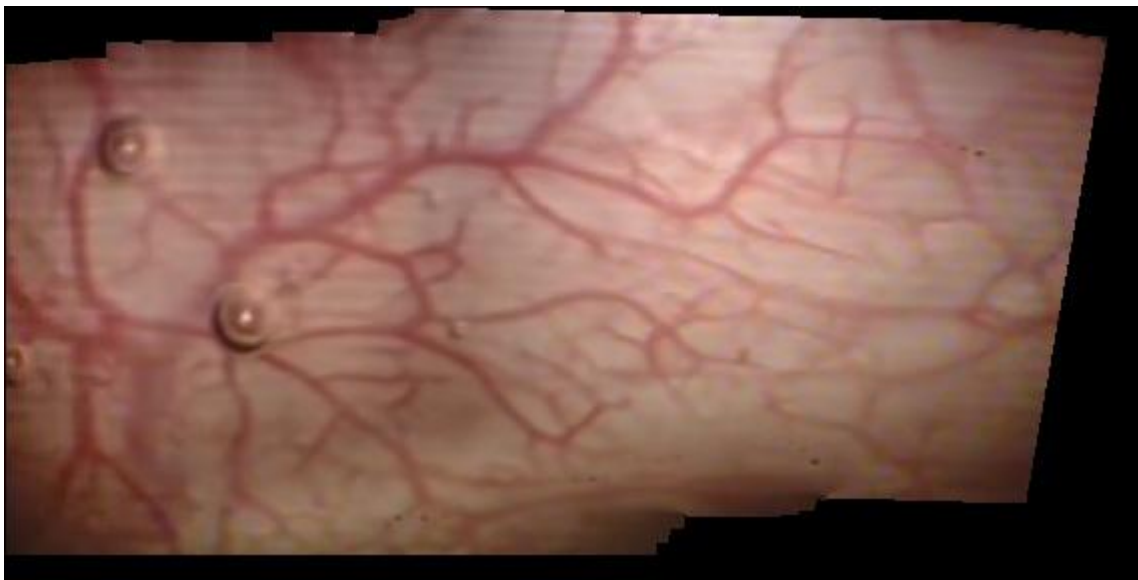


Figure 4.13 : 27 images cystoscopiques sont utilisées pour la construction de ce panorama avec une résolution de 288x567 en utilisant une transformation affine



Figure 4.14 : 27 images cystoscopiques sont utilisées pour la construction de ce panorama avec une résolution de 296x583 en utilisant une transformation projective

On remarque que les panoramas obtenus montrent très bien l'arbre vasculaire de la vessie, on peut percevoir deux bulles d'air dans la partie à gauche, la seule différence qu'on peut noter entre les deux recalages utilisés est la résolution des deux images.

Dans la suite de construction des panoramas on va se contenter d'utiliser la transformation affine pour le recalage,

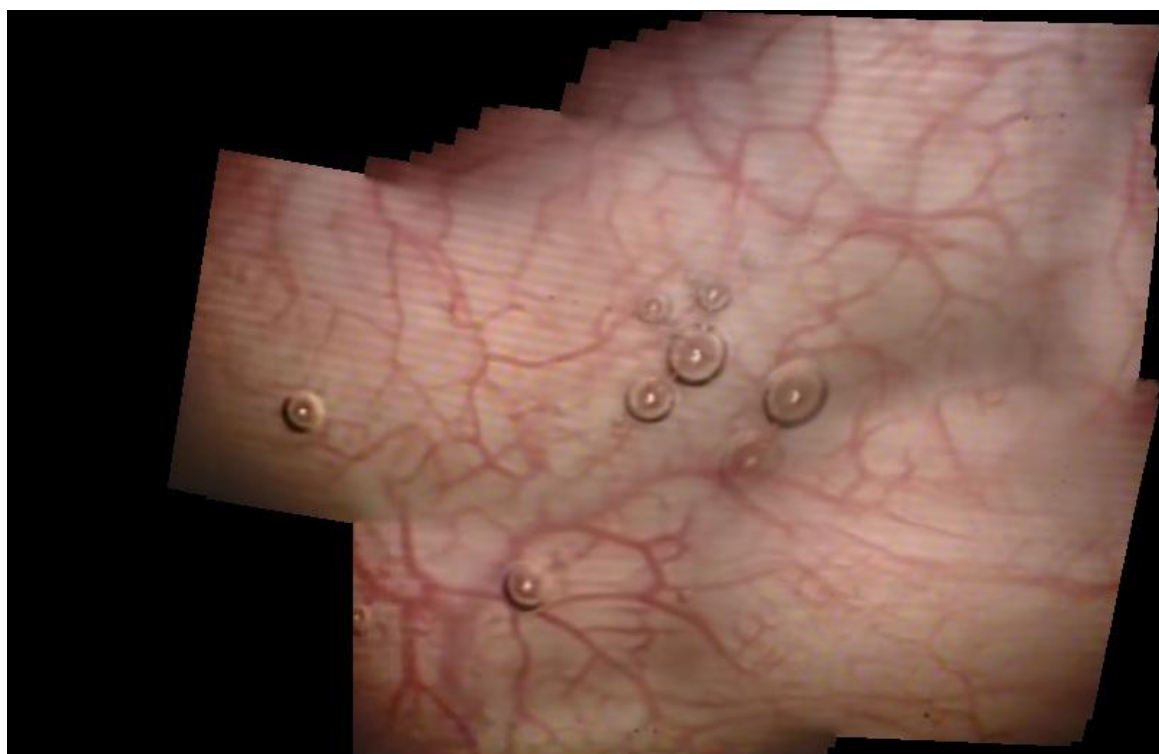


Figure 4.15 : 70 images forment ce panorama de la vessie avec une résolution de 802x517

On observe des bulles en centre et à gauche sans déceler des lésions suspectes.

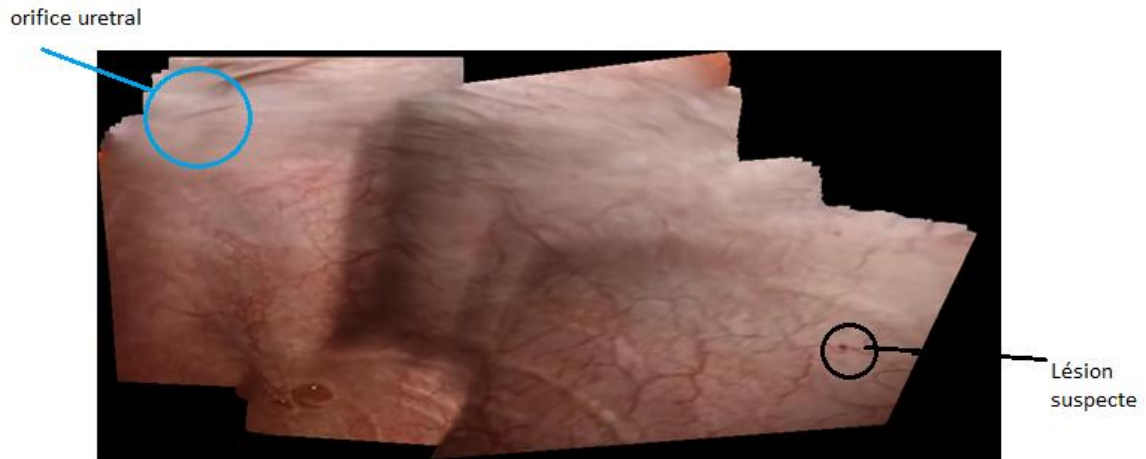


Figure 4.16 : Mosaïque formée par 300 images cystoscopiques avec une résolution de 947x428

En haut à gauche se trouve l'orifice de l'urètre, on peut remarquer en bas à droite une zone suspecte qui fera sujet d'une biopsie pour une éventuelle étude anatomopathologique.

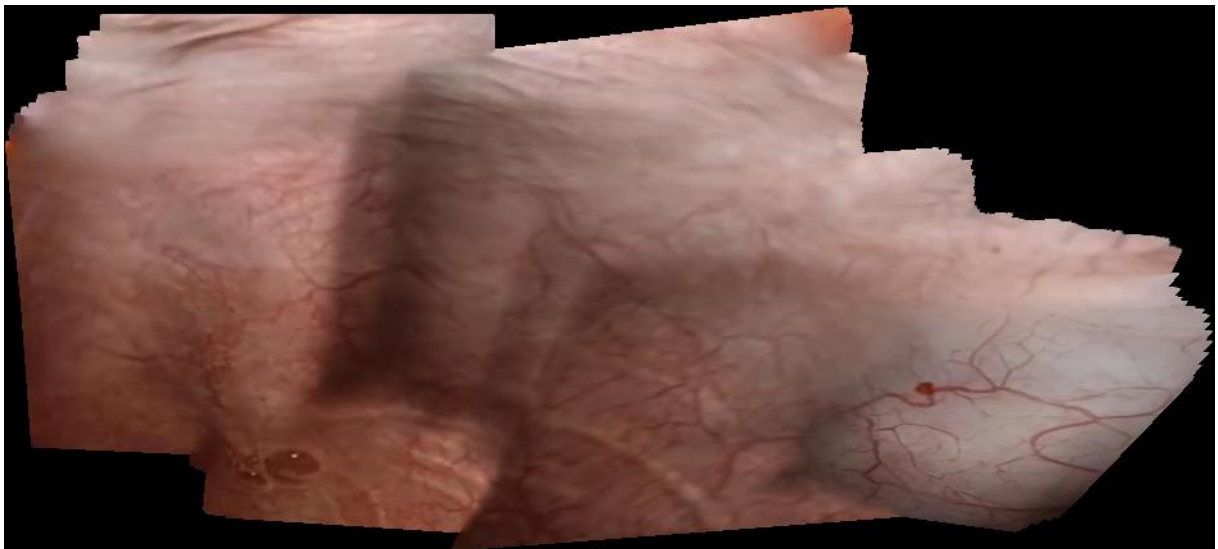


Figure 4.17 : Mosaïque formée par 353 images cystoscopiques avec la même résolution de 947x428

On a une image plus nette focalisée sur la lésion, l'urologue a rapproché l'endoscope près de la lésion ce qui a permis de la zoomer et bien l'observer.

4.8. La correction d'exposition

Le gradient du mélange pyramidal (pyramid blending) peut faire un bon travail de compensation des quantités modérées de différences d'exposition entre les images. Toutefois, lorsque les différences d'exposition deviennent grandes, d'autres approches peuvent être nécessaires.

Les méthodes d'égalisation d'histogramme adaptatif à contraste limité (contrast limited adaptive histogram equalization CLAHE [57]) séparent l'image en un certain nombre de carreaux, puis ajustent le contraste de sorte que l'histogramme du carreau ait la forme souhaitée, les carreaux sont ensuite rassemblés en utilisant l'interpolation bilinéaire . On a utilisé la fonction 'adapthisteq' de Matlab.

Afin d'optimiser l'algorithme pour une image spécifique, le nombre de carreaux, le nombre de bacs d'histogramme et le niveau d'écrtage sont ajustés. La répartition de l'histogramme peut également varier entre Poisson, exponentielle et uniforme.

Augmenter le nombre de carreaux revient à diminuer la taille de chaque carreau individuel et la taille des caractéristiques présentes dans chaque carreau. Pour les images couleurs on travaille sur chaque plan en utilisant 'adapthisteq' afin d'égaliser leurs histogrammes puis on concatène les résultats pour afficher l'image corrigée; la figure 4.18 montre les histogrammes avant et après traitement, on peut observer l'étalement de ces histogrammes vers la région lumineuse.

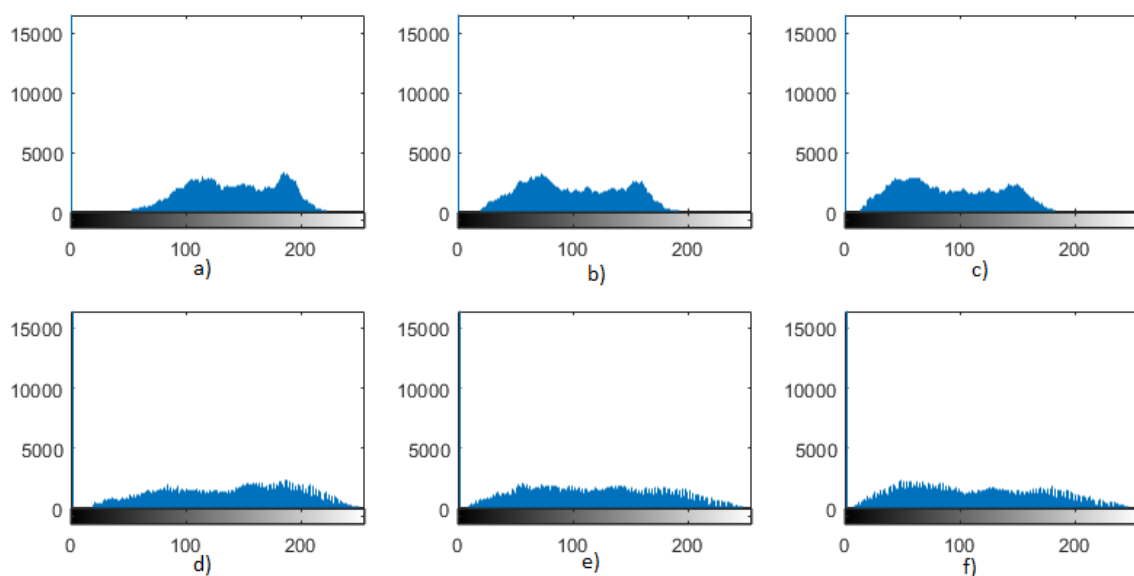


Figure 4. 18: Les histogrammes des trois canaux (a: canal rouge, b: canal vert; c: canal bleu) avant correction d'exposition lumineuse et après correction (d: canal rouge, e: canal vert; f: canal bleu)

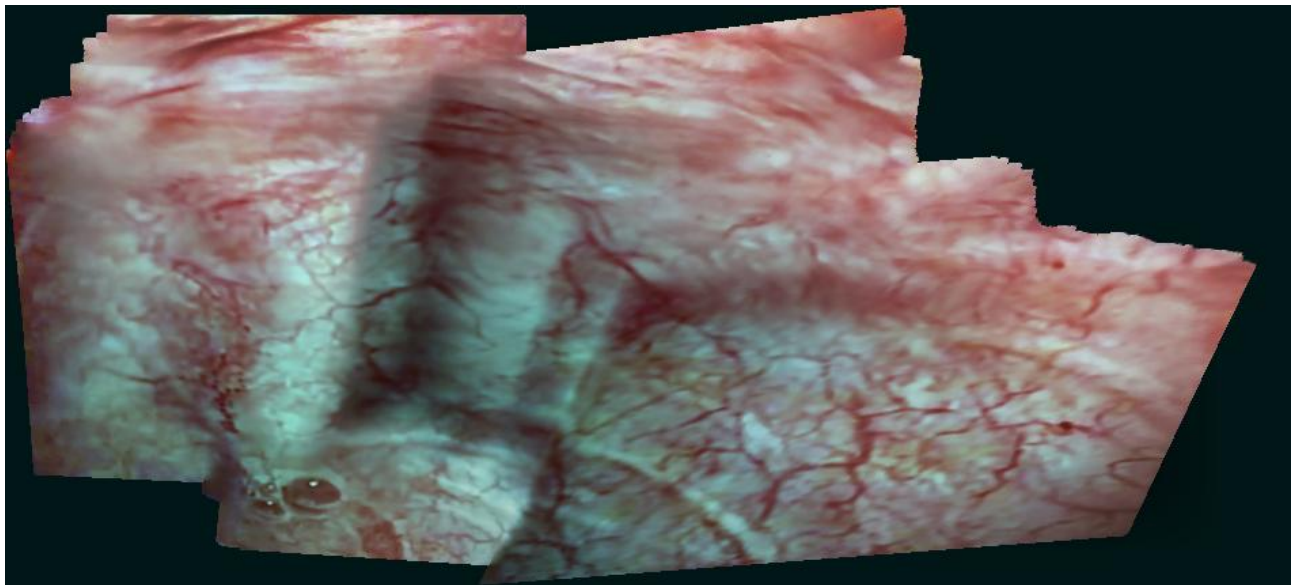


Figure 4.19 : Correction de l'exposition lumineuse en utilisant l'égalisation adaptative de l'histogramme

On remarque que l'arbre vasculaire est mieux visualisé ainsi que l'épithélium vésicale ce qui offre à l'urologue une bonne lecture facilitant le diagnostic des dysplasies ou des hémorragies.

4.9. Conclusion

En général l'application de la méthode proposée a permis d'obtenir des résultats satisfaisants; la précision du recalage qui est l'étape clé du mosaïquage a donné des vues panoramiques avec une bonne continuité de l'arbre vasculaire sans décalage visible soit pour les images angiographiques ou cystoscopiques.

Avec l'intervention de l'ophtalmologiste, l'interprétation des clichés était quasiment facile ce qui lui a permis de déceler certaines pathologies telles que la dilatation diffuse et tortuosité de l'arbre vasculaire diagnostiquée sur la figure 4.11 avec une vue plus élargie contenant 8 images angiographiques;

D'autres clichés figures 4.8 et 4.11 montrent l'impact de la photocoagulation au laser situé à droite et à gauche du disque optique, ce repérage permet à l'ophtalmologue le suivi des séances du bombardement au laser pour les régions non traitées.

D'autre part le mosaïquage des images cystoscopique a élargi le champ visuel du tissu de la vessie observée et a permis à l'urologue de situer la lésion suspecte par rapport à l'orifice urétral guidant ainsi la biopsie.

Cette procédure d'assemblage des images en un seul cliché a permis au clinicien de gagner énormément de temps et facilité le suivi du traitement et l'évolution de la pathologie.

Conclusion générale

Conclusion générale

L'angiographie rétinienne est un examen clinique ophtalmologique fondamental pour le diagnostic des pathologies rétiniennes. Il permet d'avoir une série d'images rétiniennes dont l'analyse nécessite fréquemment un mosaïquage d'images. Jusqu'ici, il existe plusieurs méthodes de mosaïquage subjective développées dans les travaux de recherche en ophtalmologie.

Dans notre travail, nous avons proposé une méthode de mosaïquage automatique basée sur l'insertion et l'exploitation d'un ensemble d'outils de traitement d'image à savoir la transformée SIFT et NSCT afin de remédier au problème de recalage des images rétiniennes et cystoscopiques. Le traitement effectué comprend: l'extraction des points clés après rehaussement des contours, la mise en correspondance des constellations de points et la création des mosaïques, cette méthode a permis d'éviter le problème des fausses correspondances des points de contrôle et d'optimiser le recalage.

On a appliqué aussi cet algorithme pour le mosaïquage des images de la cystoscopie qui est actuellement l'examen clinique de référence permettant de détecter et de diagnostiquer des lésions cancéreuses, le support photographique et le support vidéo sont communément utilisées. Sur une photographie, seule une petite région d'une zone d'intérêt restera enregistrée. Sur un support vidéo, une région plus grande sera enregistrée, mais avec une grande quantité de données redondantes. De plus, la recherche d'une zone particulière d'intérêt sur la vidéo (enregistrée à un instant donné non connu dans la séquence) peut prendre beaucoup de temps. Donc un support visuel contenant des zones assez larges de la vessie sans redondance de données permettrait au clinicien d'observer, en une seule fois, des zones de tissu plus étendues ce qui va faciliter le repérage de lésions satellites autour d'une lésion principale.

Après le mosaïquage il fallait trouver une méthode pour éliminer les sutures et corriger la différence d'exposition lumineuse; la méthode qu'on a adopté c'est la décomposition pyramidale en traitant chaque niveau à part par le blending puis rassembler l'ensemble, ce qui a donné des résultats satisfaisants avec une vue panoramique de la rétine et de la paroi interne de la vessie sans coutures entre les différents images du panorama.

Un des problèmes rencontré c'est que la rétine humaine n'est pas plane alors que les images semblent planes car l'ophtalmoscope ne capture qu'une petite partie de la rétine entière. L'une des solutions envisageable c'est d'utiliser les homographies pour s'approcher approximativement de la transformation de recalage d'une paire d'images donnée et corriger le décalage de l'arbre vasculaire qui peut en résulter.

Dans le cas du mosaïquage des images cystoscopiques, on 'a pas rencontré ce problème de décalage vu que le déplacement de l'endoscope se faisait en translation et la scène d'une image à l'autre ne différencier de quelque pixels c'est-à-dire que la zone de chevauchement est assez grande. On a constaté aussi que lorsque le nombre d'images à mosaïquer est important le temps de calcul augmente, cela est due principalement à l'étape de l'extraction des points d'intérêt qui utilise la décomposition NSCT combiné à SIFT (68,45 secondes pour la détection des points clé de deux images) et le nombre d'opérations accroit d'une façon exponentielle donné par $O(\sum_{i=0}^{\log_2 N} \frac{N}{2^i} = N \log_2 N)$ ce qui est proportionnel à N^2 , avec N le nombre d'images à mosaïquer.

Il reste à déterminer une méthode efficace pour le choix des images de la cystoscopie extraites de la séquence vidéo vue que l'urologue ne peut garder l'endoscope à distance et vitesse fixes de la paroi interne de la vessie durant tout l'examen ce qui donne des images avec des résolutions différentes et d'autres images avec du flou du au déplacement rapide de l'endoscope.

Références bibliographiques

- [1] S.E. Chen, QuickTime VR: An Image-based approach to Virtual Environment Navigation, SIGGRAPH '95, pp. 29-38, 1995.
- [2] R.Szeliski Image Mosaicing for Tele-Reality Applications, Proc. IEEE Workshop on Applications of Computer Vision, IEEE CS Press, Los Alamitos, Calif., pp. 44-53, 1994.
- [3-a] Maïtine Bergounioux Introduction au traitement mathématique des images méthodes déterministes Volume 76 of the series Mathématiques et Applications pp 13, 2015.
- [4] A.Behrens, T. Stehle, S. Gross, and T. Aach. "Local and global panoramic imaging for fluorescence bladder endoscopy". In Annual International Conference of the IEEE Engineering in Medicine and Biology Society., pp 6990, 2009.
- [3-b] Maïtine Bergounioux "Introduction au traitement mathématique des images - méthodes déterministes" Volume 76 of the series Mathématiques et Applications pp 8, 2015.
- [5] O. Berges a, K. Siahmed "Échographie de l'œil et de l'orbite" EMC-Ophtalmologie 2 pp 1-34, 2005.
- [6] [http:// www.has-sante.fr/portail/upload/docs/application/pdf/2012-10/fiche_dinformation_sur_langiographie_a_la_fluoresceine.pdf](http://www.has-sante.fr/portail/upload/docs/application/pdf/2012-10/fiche_dinformation_sur_langiographie_a_la_fluoresceine.pdf).
- [7] Eric Souied Mémoire "Nouvelle approche de l'Angiographie à la fluorescéine rétinienne adaptée à l'enseignement" pp 2, octobre 2004.
- [8-a] Nghiêm-Buffet et al " Rétinophotographies, clichés en autofluorescence et angiographies: technique et interprétation" pp 8 EMC Ophtalmologie Elsevier Masson , 2011.
- [8-b] S. Nghiêm-Buffet et al " Rétinophotographies, clichés en autofluorescence et angiographies: technique et interprétation" pp 9 EMC Ophtalmologie Elsevier Masson , 2011.
- [9-a] Marie-josé Bot-escluse "Abord clinique en ophtalmologie" pp 212 Springer-Verlag France, Paris 2013.
- [9-b] Marie-josé Bot-escluse "Abord clinique en ophtalmologie" pp 126 Springer-Verlag France, Paris.2013.
- [9-c] Marie-josé Bot-escluse "Abord clinique en ophtalmologie" pp 137 Springer-Verlag France, Paris 2013.
- [10] Ploeg M, Aben KK, Kiemeneij LA. "The present and future burden of urinary bladder cancer in the world". World J Urol vol 27 pp 289 , 2009.
- [11] A. Matau, C. Roy "Tumeurs de la vessie et de l'urètre " IRM du pelvis de l'homme et de la femme pp 94-111 , 2014.
- [12] C. Coulange " Cystoscopie" Progrès en urologie vol 20, pp 822-826 , 2010.

- [13] T.A.Pham, S.Barrat , M.Delalandre, J.-Y.Ramel " Approche robuste pour la pour la détection de point d'intérêt dans les images de documents techniques". Colloque International Francophone sur l'Ecrit et le Document, Bordeaux, France. pp.445-460, 2012.
- [14] K.Tuytelaars et al., "Local invariant feature detectors» A survey. Computer Graphics and Vision, 3 :177-280, 2007.
- [15] L.Rosenthaler, F. Heitger, O. Kubler, and R. von der Heydt, "Detection of general edges and keypoints," in G. Sandini, editor, Computer Vision – ECCV'92, pp. 78–86. Springer Verlag, Berlin, 1992.
- [16] H.P.Moravec, "Towards Automatic Visual Obstacle Avoidance". In Proceeding of Fifth International Joint Conference on Artificial Intelligence, Cambridge Massachusetts USA, pp. 584-587, 1977.
- [17] C.Harris, M.Stephens, "A Combined Corner and Edge Detector". In Proceeding of Forth Alvey Vision Conf, pp. 147-151, 1988.
- [18] S.Smith, J. Brady, "A New Approach to Low Level Image Processing". In International Journal of Computer Vision, pp. 45-78, 1997.
- [19] X. Wu, B. Bhanu, "Gabor Wavelets for 3-D Object Recognition". In proceeding of the 5 th International Conference on computer, Vol. 20, N° 23, pp. 537-542, 1995.
- [20] D. G. Lowe, "Object Recognition from Local Scale-Invariant Features". In Proceedings of the International Conference on Computer Vision (ICCV), Corfu, pp.1-8, 1999.
- [21] V.Gouet, "Mise en correspondance d'images en couleur application à la synthèse de vues intermédiaires". Thèse de Doctorat, Université de Montpellier II, 2000.
- [22] K.Rohr, "Über die Modellierung und Identifikation charakteristischer Grau-wertverläufe in Realwertbildern," en 12. DAGM-Symposium Mustererkennung, 1990.
- [23] K.Rohr , "Recognizing corners by fitting parametric," International Journal of Computer Vision, vol. 9, no. 3, pp. 213–230, 1992.
- [24] G.Medioni et Y. Yasumoto . Corner detection and curve representation using cubic B-spline. Computer Vision, Graphics, and Image Processing, 39(3) :267–278, 1987.
- [25] C.Schmid , R. Mohr et C. Bauckhage ."Evaluation of Interest Point Detectors". International Journal of Computer Vision, 37(2) :151–172, 2000.
- [26] C.Schmid, "Appariement d'images par Invariants Locaux de Niveaux de Gris Application à L'indexation d'une Base D'objets". Thèse de Doctorat, Institut National Polytechnique de Grenoble, 1996.
- [27] D.G.Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int'l J. Computer Vision, vol. 60, no. 2, pp. 91-110, Nov. 2004.

- [28] J.Zhou, A. L. Cunha, and M. N. Do, "The non subsampled contourlet transform: construction and application in enhancement," in Proc. IEEE Int. Conf. Image Processing, ICIP, vol. 1, pp. 469–472, 2005.
- [29] M. J.Shensa, "The discrete wavelet transform: Wedding the à trous and Mallat algorithms." IEEE Trans. Signal Proc., vol. 40, no. 10, pp. 2464–2482, October 1992.
- [30] Serief C., Barkat B., Bentoutou Y., and M. Benslama, "Robust feature points extraction for image registration based on the non subsampled contourlet transform". International Journal of Electronics and Communications (AEÜ), vol. 63, pp. 148–152, 2009.
- [31] B. S.Manjunath, C. Shekharm, and R. Chellappa, "A new approach to image feature detection with applications," Pattern Recognition, vol. 29, no. 4, pp. 627–640, 1996.
- [32] J.Sun, Y. Li, S. B. Kang et H.-Y. Shum. "Symmetric Stereo Matching for Occlusion Handling". IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 2, pp 399–406, San Diego, États-Unis, juin 2005.
- [33] R.Szeliski et D. Scharstein. "Symmetric Sub-Pixel Stereo Matching". European Conference on Computer Vision, volume 2, pp 525–540, Copenhagen, Danemark, mai 2002.
- [34] A. F. Bobick et S. S. Intille." Large Occlusion Stereo". International Journal of Computer Vision, 33(3) :181–200, septembre 1999.
- [35] Kratika Sharma and Ajay Goyal "Classification Based Survey of Image Registration Methods" Fourth International Conference on Computing, Communications and Networking 4th ICCCNT - pp 1-7 July 2013.
- [36] B. Zitova and J. Flusser, "Image registration methods: A survey," Image Vis. Comput., vol. 21, no. 11, pp. 977–1000, Apr. 2003.
- [37] A. Can, Hong Shen, James N. Turner, Howard L. Tanenbaum, and Badrinath Roysam " Rapid Automated Tracing and feature Extraction from Retinal Fundus Images Using Direct Exploratory Algoritms " IEEE Transactions onInformation Technology in Biomedecine vol 3, march 1999.
- [38] LI Jupeng, Chen Houjin, Yao Chang and Zhang Xinyuan. "A Robust Feature-Based Method for Mosaic of the Curved Human Color Retinal Images" International Conference on BioMedical Engineering and Informatics IEEE pp 845-849, 2008.
- [39] A.Can, C.V.Stewart, B. Roysam, and H. L. Tanenbaum. "A FeatureBased, robust, hierarchical algorithm for registering pairs of images of the curved human retina" Transactions on Pattern Analysis and Machine Intelligence ". IEEE, 24(3) :347, 2002.
- [40] P. J Besl and N. D McKay " A method for registration of 3-D shapes" Transactions on pattern analysis and machine intelligence IEEE, vol. 14, pp 239– 256, 1992.

- [41] S.Rusinkiewicz and M. Levoy, "Efficient Variant of the ICP Algorithm". Actes 3DIM, pp 145-152, Canada, Québec, juin 2001.
- [42] R. I. Hartley and A. Zisserman, "Multiple View Geometry," in Computer Vision, C. U. Press, Ed., 2004.
- [43] Thèse Construction rapide d'images panoramiques applicables `a l'exploration cystoscopique et `a l'endoscopie de fluorescence en cancérologie" de Yahir Hernandez mier soutenue le 22 octobre 2007.
- [44] Thèse "Extraction automatique de points d'intérêt à base de la transformée en contourlets non sous échantillonnée pour le recalage des images" de Serief chahira soutenue le 15 juillet 2009
- [45] Mohamed R. Metwall et al Estimation of the Optimal Set of Parameters for PAN-Sharpning of satellite images based on the Non-Sub-Sampled Contourlet Transform 29th National radio science conférence NRSC April 2012
- [46] Abdelkrim Ghaz, Kidiyo Kpalma and Abdennacer Bounoua, NSCT edge Enhancement for SIFT key points extraction, IOSR-JVSP, (2014), Vol. 4, No 2, 2014.
- [47] Jignesh Sarvaiya, Suprava Patnaik & Hemant Goklani "Image Registration using NSCT and Invariant Moment" International Journal of Image Processing (IJIP), Volume (4): Issue (2) , May 2010,p.119-130.
- [48] Feng Zhou ; Xiuli Ma ; Yang Li ; Xiaojun Zhou "Medical image enhancement based on NSCT" IET International Conference on, p 166 - 169, ICSSC 2013.
- [49] Meskine. F, Chikr EL Mezouar. M, Taleb. N. "A Rigid image registration based on the nonsubsamped contourlet transform and genetic algorithms". Sensors journal, vol.10(9), pp. 8553-8571, 2010
- [50] Philippe C. Cattin, Herbert Bay, Luc Van Gool, and Gábor Székely, "Retina Mosaicing Using Local Features" Med Image Comput Comput Assist Interv.;vol 9:pp 185 -192, 2006.
- [51] Peleg, S.:" Elimination of seams from photomosaics". Int. J. Comput. Vis (IJCV) 16, pp 90–94 1981.
- [52] Burt, P.J., Adelson, E.H.: "A multiresolution spline with application to image mosaics". ACM Trans. Graphics (TOG) 2(4), pp 217–236 (October 1983).
- [53] Milgram, D.L.:" Computer methods for creating photomosaics". IEEE Trans. Comput. 24(11), pp 1113–1119 ,1975.
- [54] Pérez, P., Gangnet, M., Blake, A.: "Poisson image editing" ACM Trans. Graphics (SIGGRAPH) 22(3), pp 313–318, 2003.
- [55] Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D.,Cohen M.:" Interactive digital photomontage". ACM Trans. Graphics (SIGGRAPH) ,2004.

- [56] Szeliski, R., Uyttendaele, M., Steedly D.: "Fast poisson blending using multi-splines". In: IEEE International Conference on Computational Photography (ICCP), pp. 1–8, April 2011.
- [57] S. M. Pizer, E. P. Amburn, J. D. Austin, et al.: Adaptive Histogram Equalization and Its Variations. *Computer Vision, Graphics, and Image Processing* 39 pp 355-368, 1987.

Résumé

L'utilisation des images panoramiques est largement répandue dans différents domaines tels que la robotique, la vision par ordinateur et la surveillance. Ces images sont également utilisées en médecine pour éliminer la redondance de l'information contenue dans les clichés des angiographies rétiniennes et de l'endoscopie vésicale. Ainsi elles apportent une aide énorme pour l'ophtalmologiste et l'urologue dans le diagnostic des pathologies de la rétine et de la vessie et permettre le suivi de l'évolution de ces pathologies et de leurs traitements.

L'objectif, affiché, de ce travail consiste à l'élargissement du champ visuelle et la création d'une vue globale de la rétine et de la vessie.

Pour parvenir à ce but; dans la première grande phase, on applique une décomposition multirésolution NSCT sur l'image originale pour rehausser les contours, dans la deuxième grande phase, on extrait les points clés à l'aide de la technique SIFT pour le calcul de la transformée servant au recalage et à l'assemblage des différentes parties de l'image panoramique.

L'extraction des points clés en rehaussant l'image originale par la NSCT et l'application de SIFT ont permis de donner des points d'intérêt plus stables et de diminuer les erreurs de correspondance entre les points des différentes portions d'images à mosaïquer même en présence de fort bruit. Aussi, on a pu corriger les sutures et la différence d'exposition lumineuse entre les images en utilisant le 'pyramid Blending' et améliorer la qualité visuelle.

Mots clé : SIFT, NSCT, pyramid Blending, rehaussement de contours, recalage.

ملخص

إن استعمال الصور البانورامية منتشرة في مختلف المجالات مثل الروبوتات، الرؤية عن طريق الكمبيوتر و المراقبة، كما تستخدم في المجال الطبي لإزالة المعلومات المتكررة في صور أوعية الشبكية و نسيج المثانة و ذلك لمساعدة طبيب العيون و أخصائي أمراض الكلى للكشف عن أمراض الشبكية و المثانة و مراقبة العلاج.

لبلوغ هدفنا المتمثل في الحصول على صورة إجمالية لشبكة العين و المثانة قمنا في أول الأمر بتحليل الصورة عن طريق تحويلات الصور متعددة-السلم و متعددة-الاتجاهات NSCT لتعزيز معالم الصور. في الخطوة الثانية تم استخراج النقاط الأساسية المثيرة للاهتمام عن طريق التحويل SIFT تستعمل هذه النقاط لحساب التحويل الملائم لجمع الصور و تركيبها في نظام احداثي موحد.

لقد تم استخراج النقاط الرئيسية عن طريق تعزيز محيطها بالتحويل NSCT ثم تطبيق التحويل SIFT الذي أعطى نقاطا أكثر استقرارا و صحح الخلل في التطابق بين أجزاء الصور التي نريد تركيبها وهذه الخوارزمية فعالة حتى بوجود تشويشات كما قمنا بتصحيح الحدود الفاصلة بين أجزاء الصور و الإختلاف في التعرض للضوء باستخدام الخلط الهرمي pyramid Blending للحصول على صورة عالية الجودة. بشكل عام إن النتائج التي تحصلنا عليها عن طريق التحويلات للحصول على صورة شاملة لشبكية العين و المثانة سمحت باختزال وقت الحسابات مما يتيح للأطباء استغلال هذه النتائج أثناء الفحص الكلينيكي للمريض و بالتالي يساعد في الكشف عن الأمراض و متابعة العلاج.