

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

المديرية العامة للبحث العلمي و التطوير التكنولوجي

Direction Générale de la Recherche Scientifique et du Développement Technologique



Projet national de recherche : rapport final

TRAITEMENT DES IMAGES MÉDICALES APPLIQUÉ À LA CHIRURGIE COELIOSCOPIQUE

Chef de projet : BESSAID Abdelhafid

Affiliation: Laboratoire de génie biomedical

Organisme de domiciliation : Université Abou Bekr BELKAID Tlemcen

Organisme pilote : CERIST

Programme national de Recherche : 2011-2013



تقرير عام لمشروع البحث
Rapport général du projet PNR

وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المديرية العامة للبحث العلمي و التطوير التكنولوجي
Direction Générale de la Recherche Scientifique et du Développement Technologique

I-Identification du projet:

PNR

Technologies de l'information et de la communication.

1-التعريف بالمشروع

Organisme pilote

CERIST

Domiciliation du projet :

Laboratoire de génie biomédical

Intitulé du projet

عنوان المشروع

Traitement des images médicales appliqué à la chirurgie coelioscopique

Chercheurs impliqués dans le projet

أعضاء المشروع و المؤسسة المستخدمة

Nom et prénom الاسم و اللقب	Grade الرتبة	Etablissement employeur المؤسسة المستخدمة	Observation
BESSAID Abdelhafid	Prof	Université de Tlemcen	Membre actif
MESSADI Mahammed	MCB	Université de Tlemcen	Membre actif
KHEMIS Kamila	MCB	Université de Tlemcen	Membre actif
LOUDJEDI MOUEDDEN Ismet Salim	MCB	Université de Tlemcen	Membre actif
BELGHERBI Aicha	Doctorante	Université de Tlemcen	Membre actif
SELKA Faiçal	Doctorant	Université de Tlemcen	Membre actif

Déroulement du projet :

Rappeler brièvement les objectifs du projet et les taches prévues

تذكير مختصر بأهداف المشروع و المهام المسطرة :

Objectifs :

Le projet de recherche que nous présentons a pour objectif de mettre en œuvre des opérations de traitement et d'analyse des images de l'abdomen permettant, de visualiser sur l'écran des images de bonne qualité permettant de localiser et extraire les différentes organes anatomiques comme le foie, la vésicule biliaire, l'identification d'éventuelles pathologies, et de faire le suivi des instruments chirurgicaux à partir des images de coelioscopie.

Taches prévues pour la première année :

- Etat de l'art sur la problématique, Recherche bibliographique
- Procédure d'acquisition des images coelioscopiques
- Elaboration de techniques d'amélioration et de filtrage de ces images.
- Implémentation d'algorithmes de segmentation, de caractérisation et de reconnaissance d'objets.

**RAPPORT D'ACTIVITÉ (Veuillez expliquer les activités que
vous avez mené dans le cadre du projet et les résultats auxquels vous êtes parvenu)
40 pages minimum**

حصيلة النشاط (يرجى شرح و توضيح النشاطات و الأعمال التي قمتم بها) 40 صفحة على الأقل

Rapport scientifique détaillé du projet (rédaction libre selon la nature du projet tout en respectant le plan suivant) :

الحصيلة العلمية المفصلة للمشروع (يتم تدوين الأعمال و النشاطات التي قمتم بها بشكل حر حسب طبيعة مشروعكم مع مراعاة المخطط الآتي) :

- Page de garde - الواجهة
- Table de matières avec les titres des chapitres et des sections ainsi que les numéros de page - الفهرس
- Introduction. - المدخل (يراعى فيه التذكير بالإشكالية و أهداف المشروع و المهام المسطرة...)
- Contenu du travail (théorie et Expérimentation) - محتوى انجاز المشروع (الجانب النظري و التطبيقي)
- Conclusions - الخاتمة و خلاصة النتائج
- Bibliographie - المراجع
- Annexes - الملاحق

- Information financière - معلومات مالية

Veuillez inscrire vos revenus et vos dépenses dans le tableau De la nomenclature des dépenses.

يرجى القيام بتدوين صرف النفقات حسب جدول مدونة النفقات

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

Université Abou BEKR BELKAID

DEPARTEMENT DE GENIE ELECTRONIQUE ET ELECTRONIQUE



**LABORATOIRE DE GENIE
BIOMEDICAL DE TLEMCEEN (GBM)**

Rapport Général du Projet PNR

Thème

Traitement des images médicales appliqué à la chirurgie coelioscopique

Présenté par :

BESSAID Abdelhafid	Prof	Université de Tlemcen
MESSADI Mahamed	MCB	Université de Tlemcen
KHEMIS Kamila	MCB	Université de Tlemcen
LOUDJEDI MOUEDDEN Ismet Salim	MCB	Université de Tlemcen
BELGHERBI Aicha	Doctorante	Université de Tlemcen
SELKA Faïçal	Doctorant	Université de Tlemcen

Année Universitaire 2013-2014

Sommaire

Résumé.....	6
1. Introduction :	6
2. La chirurgie endoscopique.....	7
2. Objectif :	9
3. La réalité augmentée :	9
4. Définition d'un point d'intérêt :	14
5. Détection des points d'intérêts:.....	15
5.1. Détecteur de Moravec :	16
5.2. Détecteur de Harris :	16
- Résumé Algorithmique :	19
- Résultats du détecteur de Harris:	20
- Influence du prétraitement sur le détecteur de Harris:	21
<i>Idées proposées:</i>	25
5.3. Le Détecteur de Shi-Tomasi.....	26
- Résumé Algorithmique :	28
- Résultats du détecteur de Shi –Tomasi:	28
6. Appariement des points:	30
6.1. Algorithme à une dimension:	31
6.2. Généralisation à plusieurs dimensions:	32
7. Reconstruction de points 3D:	33
7.1. Structure From Motion.....	34
8. Conclusion:	34
Références.....	35
Description de l'application du suivi des points	37
Introduction:	37
Description:	37
Schéma:	39
Perspectives:	40
Quelques résultats que nous avons trouvés et logiciel proposé	40
Information financière.....	46

Résumé

Le principe fondamental de ce projet est de limiter les accès à l'intérieur du corps humain en utilisant de petites incisions (céolioscopie, laparoscopie, etc.) comme canaux pour diagnostiquer et opérer, au lieu des grandes incisions utilisées en chirurgie ouverte. Les bénéfices pour les patients sont importants : durée d'hospitalisation réduite, traumatisme postopératoire minimisé, complications et infections moins nombreuses. Par conséquent, les coûts d'hospitalisation sont également réduits. Pour rendre l'opération possible, le chirurgien doit pouvoir disposer d'un retour visuel lui permettant de voir à l'intérieur du corps ainsi que d'instruments adaptés au passage par les canaux choisis.

Dans ce domaine, plusieurs méthodes sont évaluées pour l'extraction des points d'intérêts dans des images céolioscopique. Nous sommes intéressés dans ce travail à développer une méthode de planification de chemin. La mise en œuvre pratique de cette méthode de planification nécessite l'estimation et la mesure de paramètres géométriques tels que la position de l'aiguille dans le porte-aiguille. La 1ère approche que nous présentons consiste à appliquer une opération de filtrage et d'augmentation de contraste. On élimine, ensuite les éléments principaux et on garde uniquement les points d'intérêts dans le réseau sanguins. Cette méthode donne des meilleurs résultats et permet d'avoir une idée sur des mouvements des objets (pistolet).

Mot clés- céolioscopie, segmentation, morphologie mathématique, points d'intérêts, suivie et guidage

1. Introduction :

L'évolution de l'imagerie médicale a eu des répercussions directes sur la chirurgie. En permettant de mieux visualiser l'intérieur du corps, elle a permis l'avènement d'une chirurgie mini-invasive, qui ne nécessite pas de grandes ouvertures du corps. La chirurgie mini-invasive repose sur l'utilisation d'un endoscope composé d'une optique cylindrique muni d'un système d'éclairage, introduit dans l'organisme par une toute petite incision, voire même glissé par les voies naturelles. Ce type de chirurgie procure de nombreux avantages (morbidity post-opératoire et le temps d'hospitalisation sont réduits, et il résulte de l'opération des cicatrices mineures), par rapport à la chirurgie classique. Cependant, cette technique présente quelques inconvénients pour le praticien tels que, la limitation du champ visuel, la perte de la vision 3D directe, la mobilité réduite des instruments de chirurgie, le manque de rétroaction tactile. Ceci peut donc causer la désorientation visuelle-spatiale lors de l'exploration des structures in-vivo complexes, ces limitations ont encouragé plusieurs axes de recherche dans le domaine de la vision par ordinateur, donnant ainsi une véritable impulsion à la chirurgie laparoscopique et aussi, dans le cadre des Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO), qui sont des systèmes de navigation chirurgicaux, où ils assistent les chirurgiens, durant une intervention chirurgicale sur des images endoscopiques. **L'objectif de ces recherches est d'apporter des informations en temps réel au chirurgien afin de réduire les limitations préalablement décrites. En effet, disposer de l'information tridimensionnelle, et en temps réel, relative à l'environnement dans lequel opère un praticien permettrait d'envisager de nouvelles perspectives dans ce domaine, tel que la restitution visuelle de la profondeur pour offrir un plus grand confort et une plus grande précision dans le geste chirurgical, la reconstruction de surface, qui permet d'élargir le champ visuel du praticien et les applications en réalité augmentée qui apporte une aide au praticien sous une information virtuelle insérée dans l'environnement réel. D'autre part, la connaissance 3D en temps réel de la géométrie de la cavité abdominale pourrait aussi, par exemple, permettre de piloter des systèmes robotisés dans le but d'assister le chirurgien dans ses actes chirurgicaux pour aboutir à une plus grande sécurité pour le**

patient. Parmi ces travaux de recherches, l'application de la réalité augmentée, avec un recalage d'un modèle 3D sur des environnements rigides. Dans le cas de la chirurgie laparoscopique, cela n'est pas satisfait à cause de la déformation des tissus où aucune application n'est en routine clinique.



Fig. 1. Mise en place des trocarts dans la salle d'opération pour une intervention laparoscopique in vivo

2. La chirurgie endoscopique

Parmi les innombrables progrès qu'a connus récemment la chirurgie, le développement de la chirurgie minimalement invasive, ou chirurgie endoscopique, tient une place importante. L'endoscopie est une technique de diagnostic qui consiste à examiner l'intérieur des organes ou des cavités internes du corps, à l'aide d'un appareil appelé endoscope. L'intérêt de ces méthodes est que, comparées à des opérations ouvertes du même type, elles réduisent considérablement le traumatisme du patient. Cette amélioration est la conséquence directe de la forte diminution de la taille des incisions pratiquées, ou même de leur suppression, lorsqu'une voie d'accès naturelle existe. On observe ainsi d'une part une diminution des séquelles esthétiques liées à la taille des cicatrices, mais aussi d'autre part une chute spectaculaire de la durée totale d'hospitalisation. Les champs d'application de l'endoscopie sont très étendus : on les retrouve en orthopédie, en urologie, en gynécologie, en neurochirurgie, mais aussi en chirurgie vasculaire ou en chirurgie cardiaque. On emploie suivant les cas les termes d'arthroscopie, de bronchoscopie, de coloscopie, de gastroscopie, de cystoscopie ou encore d'hystérocopie. Dans le cas de la chirurgie digestive qui nous intéresse, on parle indifféremment de coelioscopie ou de laparoscopie. L'histoire de l'endoscopie moderne débute en 1806 avec l'invention par Philip Bozzini, chirurgien italien installé à Vienne, d'un appareil appelé *lichtleiter*. Cet instrument comporte un système de miroirs qui permet l'envoi, à l'intérieur de la cavité d'un animal vivant, des rayons lumineux générés par une bougie, ainsi que leur retour jusqu'à l'œil du praticien. Réprimandé par la faculté de médecine pour sa "curiosité", Bozzini n'a pas pu tester son invention, véritable ancêtre de l'endoscope, sur un humain. En 1853, le français Antoine Jean Desormeaux améliore le *lichtleiter* de Bozzini et choisit d'utiliser une lampe à mèche comme source de lumière. Il augmente la visibilité mais cause, en contrepartie, des risques de brûlures. Desormeaux, qui utilisa son invention sur de nombreux patients, principalement dans le domaine de l'urologie, est généralement considéré comme le père de l'endoscopie. L'éclairage de la cavité est longtemps resté un problème crucial. En 1865, l'irlandais Cruise améliore l'intensité lumineuse en remplaçant l'essence et la térébenthine par un mélange de camphre et de pétrole. En 1873, le viennois Maximilien Nitze s'inspire de l'invention d'Edison et utilise un fil de platine chauffé à blanc par le passage d'un courant électrique, ce qui améliore encore l'illumination mais surtout diminue les risques de brûlures. C'est finalement l'arrivée de la lampe à incandescence à la fin du XIX^{ème} siècle, et surtout sa miniaturisation, qui permet l'apparition des premiers endoscopes à ampoule. Les progrès

modernes et l'apparition des fibres optiques permettent de nos jours l'éclairage efficace et sans danger pour les cavités. En 1869, Pantaleoni réalisa la première hystéroscopie, en 1879 Leistner pratique la première cystoscopie, en 1881 Miculicz la première gastroscopie et en 1918 Takagi la première arthroscopie du genou. Si les premières opérations se cantonnaient au diagnostic ou à la réalisation de biopsies, la tentation d'aller vers la thérapie et la chirurgie fut présente dès le début avec des actes de plus en plus complexes.



Fig. 2. Exemple de cours de chirurgie laparoscopique à l'IRCAD (sur des cochons). En haut à gauche, l'ensemble des outils utilisés. En bas à gauche on réalise le pneumopéritoine après avoir mis en place l'insufflateur. On peut observer au premier plan les deux trocars qui vont accueillir les instruments et au fond celui qui sera utilisé par la caméra. L'image de droite montre la disposition du praticien et des aides lors de l'opération.

La laparoscopie a été plus longue à franchir le pas et se cantonna longtemps dans le domaine du diagnostic. C'est en 1901 que George Kelling observe pour la première fois la cavité abdominale d'une chienne préalablement remplie d'air et c'est en 1911 que le suédois Hans Christian Jacobaeus utilise le premier cette méthode sur un être humain pour observer le thorax et l'abdomen. Dans les années 50, le gynécologue Raoul Palmer découvre l'importance du maintien d'une pression constante dans la cavité abdominale. Le premier insufflateur automatique fut développé par l'allemand Kurt Semm en 1960. En 1982, la miniaturisation des caméras vidéos rend possible les premières opérations de vidéo-laparoscopie et ce n'est qu'en 1987 que le chirurgien lyonnais Philippe Mouret réalise la première cholécystectomie (ablation de la vésicule biliaire) par coelio-chirurgie. Les techniques de chirurgie laparoscopique sont en effet spécifiques et assez délicates à maîtriser. Par exemple, le chirurgien

n'opère plus directement avec ses mains mais à l'aide de longs instruments dont l'axe passe par un point fixe, ce qui a pour conséquence d'inverser tous les mouvements. Le chirurgien ne regarde plus ses mains mais un écran et les informations de relief ne lui sont plus accessibles directement. Il lui est de plus difficile de s'orienter car la caméra est manipulée par un assistant à qui il doit demander d'effectuer les mouvements désirés.

2. Objectif :

L'objectif de ce projet est de faire de la réalité augmentée par reconstruction de surface, en d'autres termes d'établir un modèle 3D par reconstruction de surface et de l'insérer dans l'environnement réel qui correspond au champ visuel du praticien produit par une caméra endoscopique. Ainsi, à partir de ces surfaces, il sera possible d'augmenter le champ de vision en navigation endoscopique, de proposer un rendu en relief, ou bien encore d'insérer des objets synthétiques ou naturels issus d'autres sources (imageurs TDM, IRM,...), en d'autres termes de concevoir un outil compatible avec la chirurgie endoscopique qui devra être capable de fournir l'information tridimensionnelle de la surface imagée en temps réel pour une assistance visuelle ou robotisée. Dans ce rapport, on décrit un état de l'art des systèmes de réalité augmentée en chirurgie mini-invasive, qui mène à éclairer les étapes à franchir pour faire de la reconstruction de surface avec une caméra monoscopique :

- Un prétraitement de l'image.
- La détection de point d'intérêts dans les images et ou l'extraction de zones caractéristiques.
- Le suivi de ces points dans plusieurs images cela permet d'estimer leur profondeur relative et donc d'estimer la position de la surface de l'organe.
- Construction de modèle 3D à partir de ces points qui peut ensuite être utilisée pour diverses applications (recalage non rigide, réalité augmentée...).

La détection et le suivi devront donc être précis et rapide, afin que modèles 3D reconstruits insérés apparaissent de façon fluide et cohérente dans les images.

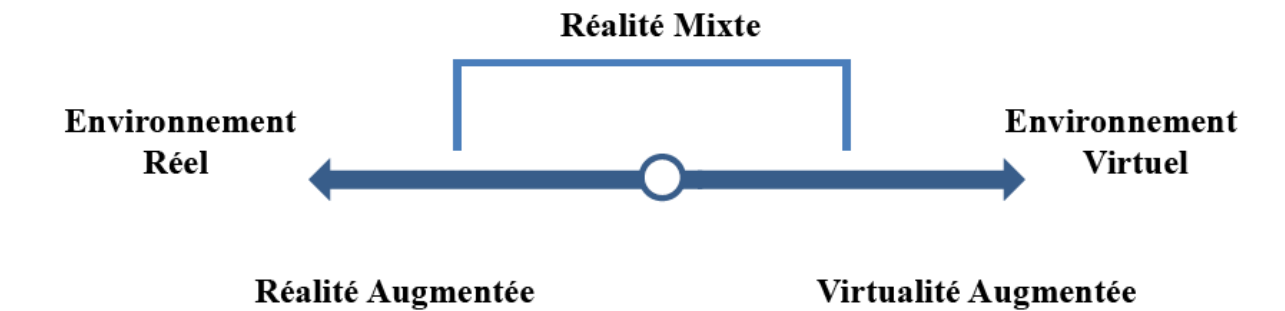
Par la suite On décrit la première partie de notre approche pour faire de la réalité augmentée, qui est la détection des points d'intérêts ainsi que leurs suivis. Néanmoins il est difficile de localiser une région d'intérêt dans une image et trouver sa région correspondante dans une autre dans le cas des images en chirurgie laparoscopique. Ceci est dû aux images qui peuvent être mal contrastées, mal illuminées ou bruitées, et aussi au fait de l'apparition des tissus, qui varie énormément d'une manière homogène, à une manière fortement texturée et aussi, à une apparition de régions qui contiennent des réflexions spéculaires dépendantes du point de vue de l'endoscope. Il est aussi nécessaire de prendre en compte l'occlusion selon les instruments chirurgicaux, et des effets dynamiques comme la fumée de cautérisation et le saignement.

3. La réalité augmentée :

On appelle réalité augmentée l'insertion d'éléments virtuels dans une séquence vidéo « réelle ». A la différence de la réalité virtuelle, où la totalité de la scène est générée artificiellement, une séquence en réalité augmentée contient à la fois des éléments réels et virtuels. R. Azuma [1] définit la réalité augmentée comme ayant les propriétés suivantes :

- elle combine les éléments réels et virtuels dans un environnement réel.
- elle est interactive, et par conséquent fonctionne en temps réel.
- elle cherche à ajuster des objets réels et virtuels les uns par rapport aux autres.

La réalité augmentée consiste le plus souvent à ajouter des objets virtuels dans un environnement réel, et donc à « augmenter » notre vision, mais il ne faut pas la limiter à cela : elle peut au contraire consister à retirer des objets réels dans des images, pour pouvoir insérer d'autres éléments. De plus, potentiellement, cela ne concerne pas uniquement la vue, mais aussi l'ouïe, l'odorat, le toucher. Il s'agit donc, plus que d'ajouter des éléments à la réalité, de modifier cette réalité. D'après Milgram [2] il existe une continuité entre l'environnement virtuel et réel que l'on peut modéliser avec le schéma ci-dessous.



Entre autre il n'y a pas de rupture entre la virtualité augmentée, qui consiste à ajouter des éléments réels dans un environnement virtuel, et la réalité augmentée. Toutes deux appartiennent au domaine de la « réalité mixte », mélange d'éléments réels et virtuels.

Pour le cas des systèmes mis en place durant la navigation en chirurgie laparoscopique, on est confronté aux problèmes de la compensation des mouvements physiologiques (déformation des tissus) or la recherche courante dans le secteur, suppose généralement que l'environnement est statique, il est difficile de satisfaire la plupart des procédures chirurgicales (abdominale, vasculaire), le fait de la mobilité de l'endoscope, des contraintes du mouvement diaphragmatique responsable de la respiration du patient et du déplacement des organes, causés par les instruments introduits dans le corps du patient.

Dans un premier temps, les systèmes de navigation en chirurgie coelioscopique étaient basés sur une mise en correspondance rigide (Recalage) de l'information, qui provenait des systèmes d'acquisitions tels qu'IRM, TDM. En pratique, ils recalaient des modèles 3D rigides extraits des images préopératoires IRM/TDM sur la vue endoscopique (Réalité Augmentée). Ils recalaient dans un même repère les instruments et les modèles 3D et les affichaient dans un environnement virtuel (Virtualité Augmentée). L'erreur de recalage de ces systèmes satisfaisait les exigences en neurochirurgie [3], [4] et en orthopédie [5], [6], car les structures considérées dans ces applications sont majoritairement rigides.

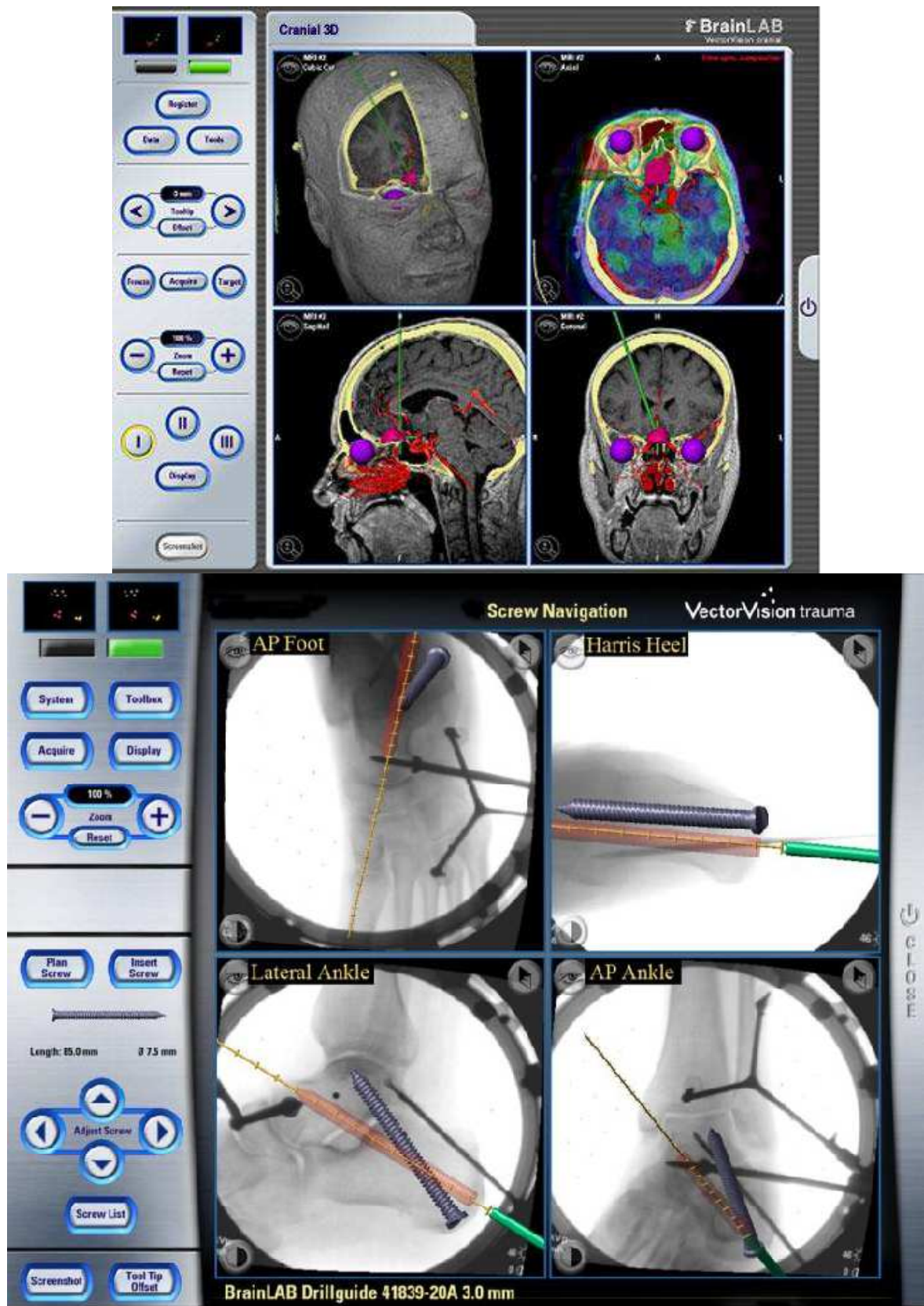


Fig.3. Le système VectorVision de la société BrainLab.

Par la suite il fut établi, que c'était finalement rarement vrai [3], [4], car au moment de l'ouverture de la calotte crânienne, il peut y avoir un affaissement du cerveau, ou alors une pression qui le fait bouger [7] causant ainsi une perte en précision du système.

Pour ce qui est des systèmes de guidage en chirurgie abdominale, des travaux ont adopté une méthode de recalage ponctuelle 3D/2D basée sur des marqueurs radio-opaques [8] collés sur la peau du patient,

qui est ensuite passé au scanner à rayon X; pour superposer le modèle 3D dans les images vidéo, il faut positionner les marqueurs localisés sur la reconstruction 3D sur ceux visibles dans les images vidéo. Cette technique a présenté un système de guidage par réalité augmentée, qui permet de réaliser des ponctions de cibles dans l'abdomen (niveau du Foie), l'aiguille doit être repérée en temps réel afin de lui superposer son modèle virtuel dès qu'elle sera introduite sous la peau.

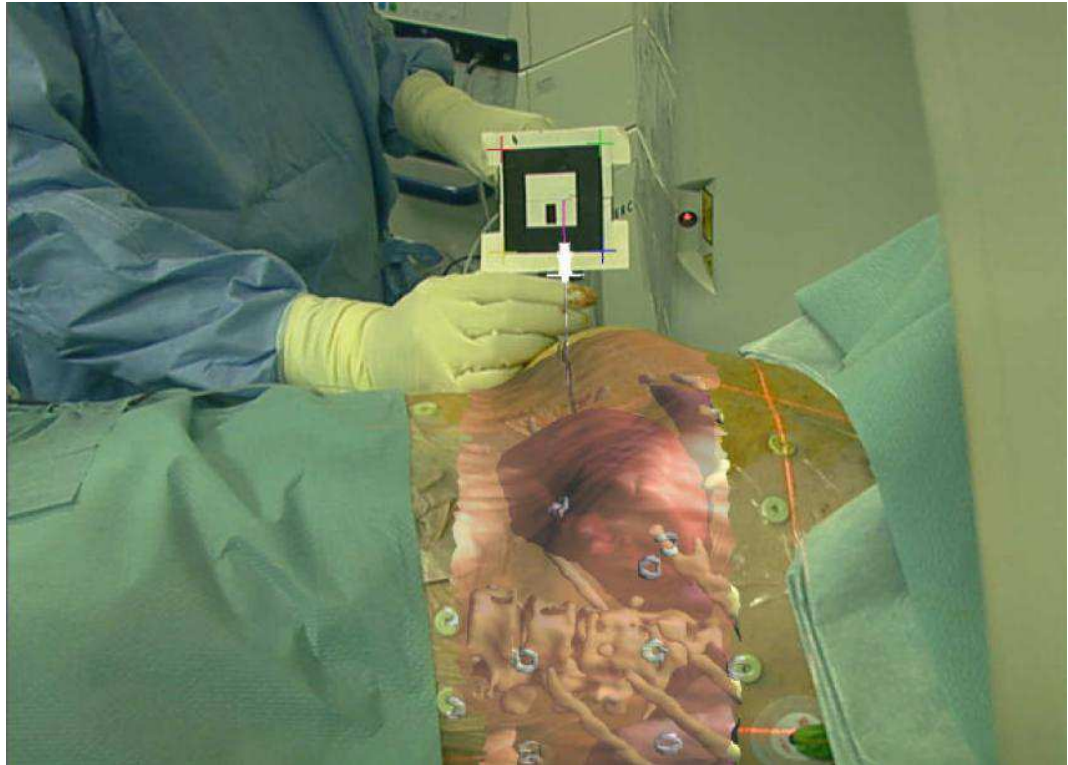


Fig.4. Système de vision utilisant une méthode de recalage 3D/2D.

Les approches récentes pour la visualisation d'un modèle 3D de l'environnement facilitant ainsi la navigation in-vivo [16] s'articulent sur la méthode SLAM (Simultaneous Localization And Mapping). La méthode proposée augmente le champ visuel efficace en créant in-crémentalement un modèle 3D du tissu pendant que le laparoscope est dirigé autour de la scène chirurgicale. Le modèle 3D est employé pour augmenter la vue courante du laparoscope avec l'information de l'extérieur du champ visuel courant. Le modèle est projeté sur une caméra virtuelle située dans le même endroit que le laparoscope, de ce fait créant un plus grand champ visuel que le laparoscope physique lui-même. La méthode établit un modèle 3D du tissu *in situ* et récupère la position du laparoscope relativement au tissu. Avec la disponibilité d'un endoscope stéréo, le SLAM visuel a été avec succès utilisé comme rapporté dans [17], où c'est particulièrement utile pour le laparoscope avec la vision restreinte parce qu'une carte globale du champ visuel d'opération peut être intégrée avec la vision stéréo mobile permettant une meilleure localisation de la sonde endoscopique.

La méthode du SLAM peut faire face non seulement au mouvement de l'endoscope, mais aussi à la compensation périodique des mouvements d'organes (MC-SLAM) [18], par un modèle à niveau élevé d'apprentissage. Le modèle de mouvement appris est explicitement incorporé au cadre statistique du SLAM, permettant au mouvement dynamique du tissu d'être estimé même lorsqu'il est en dehors du champ visuel courant de l'endoscope, qui à notre connaissance, est le premier travail pour l'évaluation en ligne simultanée du mouvement d'endoscope et de la structure dynamique. Dans MC-SLAM, un modèle de mouvement périodique est présenté pour compenser la déformation dynamique de tissu, de

ce fait permettant la cartographie dynamique et la localisation précise de l'endoscope. Conceptuellement, ceci présente trois étapes originales :

La première est d'apprendre une première évaluation du modèle périodique de respiration. La deuxième et troisième étapes sont la prévision du modèle de mouvement de respiration et la prévision du mouvement de tissu dans la carte.

Dans une deuxième approche pour l'extraction des modèles 3D à partir de scènes réelles filmées à l'aide d'une simple caméra monoculaire, les techniques de structure de mouvement qui sont très proches de la conception du SLAM, sont liées à l'effet de perception cinétique en profondeur, et à l'estimation de la structure 3D d'un objet rigide à partir d'images 2D de la structure prise par une caméra en mouvement. Les techniques de structure de mouvement sont basées sur la possibilité d'utiliser des points d'image correspondants dans des vues multiples. Étant donné une image avec deux ou plusieurs points de vue, un point 3D peut être reconstruit par triangulation. Un pré requis important est la détermination de la pose et du calibrage de la caméra, qui peut être exprimé par une matrice de projection. La théorie géométrique de structure du mouvement permet aux matrices de projection et des points 3D, d'être calculée utilisant simultanément des points correspondants dans chaque vue.

Les techniques de structure de mouvement sont employées dans une étendue d'applications comme par exemple, la construction automatique des modèles virtuelles d'objets réels à partir de séquence vidéo [19] la segmentation et la reconnaissance de structure basée sur la projection de nuage de points durant une séquence vidéo permet ainsi d'établir un model 3D de la scène [20]. Elle a aussi été utilisée pour résoudre le problème de la reconstruction 3D des surfaces des organes internes [21] par le suivi d'indices visuels surfaciques des organes au cours du mouvement de la caméra en utilisant le système stéréoscopique du robot *Da Vinci*, par la suite, les positions 3D de ces indices ont été calculées afin de fournir un champ de vue plus large.



Fig.5. Nouvelle version du robot Da Vinci, cette fois en 3Det en haute définition

Une nouvelle approche sur la reconstruction de surface 3D déformable à partir d'une séquence d'images 2D [22], est un problème également connu sous le nom de structure de mouvement souple (déformable). Cette déformation peut être représentée comme une combinaison linéaire des formes rigides, ou des modes de déformation, avec des coefficients variables dans le temps [23]. Cette approche utilise un modèle de déformation quadratique qui est en mesure de représenter des mouvements non rigides, non-linéaires tels que l'étirement, le cisaillement, la flexion et la torsion. Ceci convient particulièrement aux applications de navigation en chirurgie laparoscopique abdominale et vasculaire où une grande déformation du tissu se présente, qui est due au mouvement physiologique du patient (respiration, battement cardiaque).

Notre stratégie quant à la reconstruction de surface avec une caméra monoscopique, repose sur une identification et un suivi de points le plus robuste possible avec un traitement en temps réel, ceci car les méthodes décrites ci-dessus utilise soit des systèmes stéréoscopiques et ou repose sur un traitement qui est lent et qui n'apporte pas d'information en temps réel au praticien.

4. Définition d'un point d'intérêt :

Dans de nombreuses applications (mise en correspondance d'images stéréoscopiques, détection ou reconnaissance d'objets, suivi...), la comparaison systématique de tous les pixels entre deux images n'est généralement pas envisageable car elle nécessite un temps de calcul très important. Le processus peut être simplifié en réduisant le domaine de recherche aux zones de l'image localement intéressantes.

Les points référençant ces zones sont dénommés **points d'intérêt**. Informellement, un point d'intérêt correspond à une discontinuité des niveaux de gris (voire des couleurs), de la texture, de la géométrie,... de l'image. Il est souvent assimilé à un coin.

Les images de la figure 1 donnent une panoplie de types de coins qui se situent à des jonctions de type « L », « V », « T », « Y », « X » et « damier ».

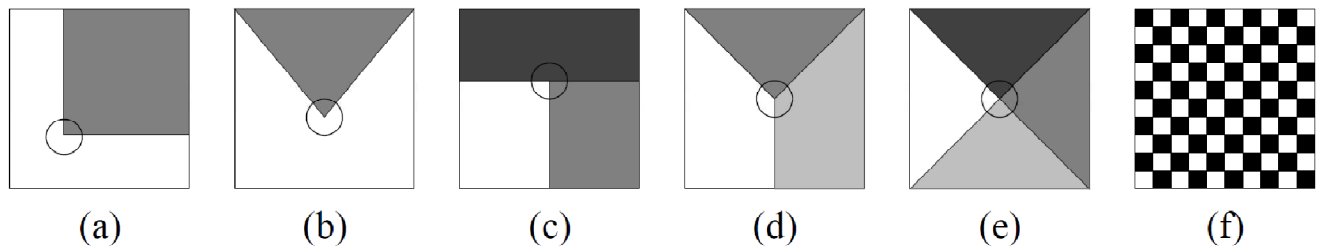


Fig.6. Différents types de coins : (a) jonction en « L », (b) jonction en « V », (c) jonction en « T », (d) jonction en « Y », (e) jonction en « X » et (f) jonction en « damier ».

Il est possible d'illustrer les différents types de coins par les images de la figure 5. Il est constaté clairement d'avoir une discontinuité parfaite à cause des bruits du capteur, de la floutée et de la compression des images.

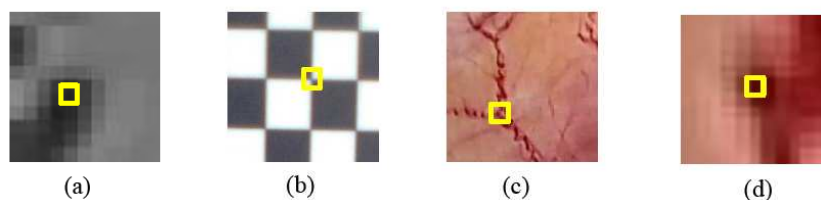


Fig.7. Différents types de coins dans le cas des images réels

La sélection de ces points d'intérêts est une étape importante parce qu'elle a une influence directe sur la performance du suivi. Plusieurs stratégies ont été mises en place pour la détection de ces points d'intérêt.

5. Détection des points d'intérêts:

De nombreuses méthodes ont été proposées pour détecter des points d'intérêts. Elles peuvent être classées suivant trois catégories :

- **Approches contours** : l'idée dans un premier temps est de détecter les contours dans une image. Les points de courbures maximales ainsi que les intersections des contours, correspondent à des points d'intérêts.
- **Approches intensité** : l'idée est de calculer la moyenne de la fonction d'intensité, dans plusieurs directions pour en extraire directement les points de discontinuités sur l'image.
- **Approches à base de modèles** : les points d'intérêts sont identifiés dans l'image par mise en correspondance de la fonction d'intensité avec un modèle théorique (préalablement connu) de cette fonction des points d'intérêts considérés.

La partie qui va suivre portera sur les premiers principaux détecteurs qui sont basés sur le calcul du gradient de l'image, où on détermine une fonction d'énergie suivant la variation des intensités de ces gradients, dans diverses directions. Les détecteurs présentent une amélioration successive du détecteur de Moravec.

5.1. Détecteur de Moravec :

L'idée du détecteur de Moravec est de considérer le voisinage d'un pixel (une fenêtre) et de déterminer les changements moyens de l'intensité dans le voisinage considéré lorsque la fenêtre se déplace dans diverses directions $\{(1,0), (1,1), (0,1), (-1,1)\}$ Plus précisément. On considère la fonction de changement d'intensité :

$$E(x, y) = \sum_{u,v} W(u, v) [I(x + u, y + u) - I(u, v)]^2 \quad 5.1$$

Où :

- 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 conséquence, le principe du détecteur de Moravec est donc de rechercher les maxima locaux de la valeur minimale de E en chaque pixel (au-dessus d'un certain seuil) pour les déplacements $\{(1,0), (1,1), (0,1), (-1,1)\}$.

Le détecteur de Moravec fonctionne dans un contexte limité. Il souffre en effet de nombreuses limitations :

- Réponse anisotrope en raison du caractère discret des directions de changement (rotation par pas de 45°).
- Réponse bruitée en raison qu'il ne considère aucun filtrage de l'image.
- Le détecteur de Moravec répond de manière trop forte aux bords en raison du fait que seul le minimum de E est pris en compte en chaque pixel.

Harris et Stephen ont identifié certaines limitations et, en les corrigeant, en ont déduit un détecteur de coins très populaire : le *détecteur de Harris*.

5.2. Détecteur de Harris :

Pour définir la notion de coins dans des images, Harris[10], [11]regarde la moyenne des changements d'intensité directionnelle d'une petite fenêtre autour d'un point d'intérêt **(5.1)**, ceci donc nous renvoie que, le détecteur de Harris est une version améliorée qui corrige les limitations préalablement décrite du détecteur de Moravec.

- Réponse anisotrope:

Pour améliorer cet aspect, Harris propose d'effectuer le développement de Taylor de la fonction d'intensité au voisinage du pixel (u, v) .

$$I(x + u, y + v) = I(u, v) + x \frac{\partial I}{\partial x} + y \frac{\partial I}{\partial y} + O(x^2, y^2) \quad 5.2$$

D'où

$$E(x, y) = \sum_{u,v} W(u, v) [x \frac{\partial I}{\partial x} + y \frac{\partial I}{\partial y} + O(x^2, y^2)]^2 \quad 5.3$$

En négligeant le terme $O(x^2, y^2)$ (valable pour les petits déplacements), on obtient la forme quadratique de l'énergie:

$$E(x, y) = Ax^2 + 2Cxy + By^2 \quad 5.4$$

$$\text{Avec : } A = \sum_{u,v} \left(\frac{\partial I}{\partial x} \right)^2 * W \quad B = \sum_{u,v} \left(\frac{\partial I}{\partial y} \right)^2 * W \quad \text{et}$$

$$C = \sum_{u,v} \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right) * W$$

- Réponse bruitée:

Pour améliorer cela, Harris et Stephen propose de déterminer les dérivées par convolution avec une fenêtre gaussienne. Mathématiquement, ceci équivaut à filtrer l'image avec une gaussienne avant de calculer le gradient.

$$W(u, v) = e^{-(u^2+v^2)/2\sigma^2} \quad 5.5$$

La gaussienne W , avec laquelle sont convoluées les composantes du gradient, permet de tenir compte du gradient dans une fenêtre circulaire autour de chaque pixel, sans coupure brusque, ainsi, les calculs des gradients sont moins approximatifs.

Forte réponse aux bords:

Pour prendre en compte le comportement général de la fonction E localement, on écrit :

$$E(x, y) = (x, y) \cdot M(x, y)^T \quad 5.6$$

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

Cette matrice est une matrice symétrique 2x2 qui caractérise la moyenne des changements d'intensités dans toutes les directions. Les paramètres A, B et C de la matrice M impliquent les gradients de l'image qui sont souvent calculés en utilisant l'opérateur Sobel décrit ci-dessous.

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

On peut montrer que les valeurs propres de la matrice M nous renseignent sur le changement d'intensité moyen maximal et le changement d'intensité moyen pour la direction orthogonale.

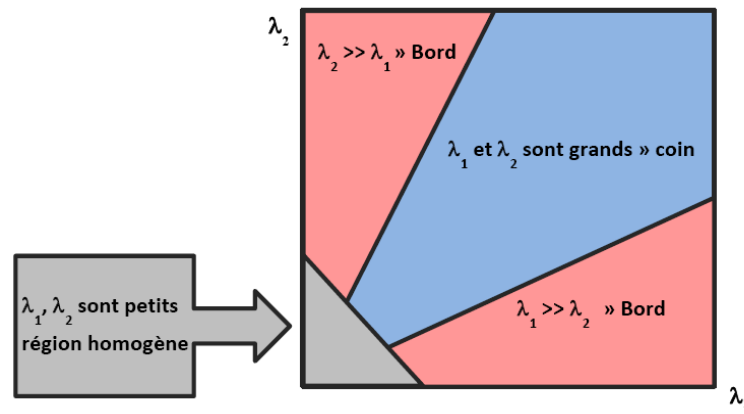


Fig.8. Représentation de la réponse des valeurs propres pour la détermination des régions.

La matrice M caractérise le comportement local de la fonction E, les valeurs propres $[\lambda_1, \lambda_2]$ de la matrice M renseignent sur les caractéristiques de la fenêtre considérée.

- Une valeur propre élevée : contour
- Deux valeurs propres élevées : coin

Dans un but d'éviter le calcul fastidieux des valeurs propres, Harris et Stephen proposent l'opérateur suivant pour définir un coin dans une image :

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

$$\text{Avec } \text{Trace}(M) = \lambda_1 + \lambda_2 = A + B \quad \text{Det}(M) = \lambda_1 \cdot \lambda_2 = AB - C^2 \quad 5.7$$

et : k compris entre $0,02 \cong 0,06$

L'intérêt majeur de ce critère est d'éviter le calcul explicite des valeurs propres. De manière simplifiée c'est d'interpréter géométriquement les courbures principales de la fonction d'auto corrélation du signal. Une courbure importante dans les deux directions principales indique la présence d'un coin qui est retenu comme point d'intérêt.

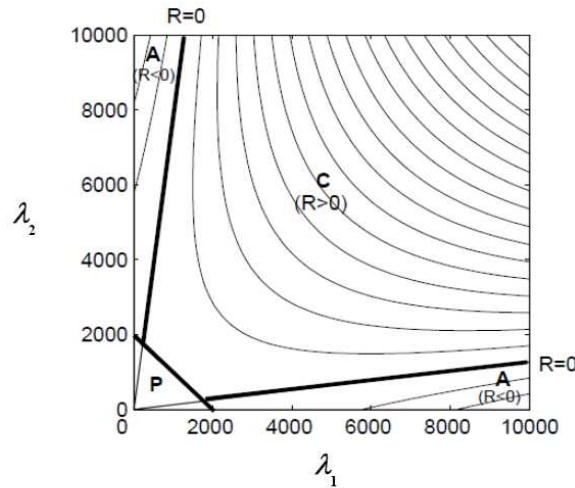


Fig.9. Courbes de niveau de la fonction de réponse du détecteur de Harris.

Les types de régions sont délimités par le signe de la réponse et le seuil: coins (C), contour (A) et plat (P).

Le paramètre « k » est utilisé pour le calcul des réponses du détecteur de Harris. À aucun moment, Harris et Stephens ne mentionnent une quelconque valeur stratégique pour k. Certaines implémentations utilisent $k = 0.04$. Cette valeur a été confirmée par Harris mais il ne sait pas comment elle a été diffusée. Elle est partie intégrante de l'implémentation « temps-réel » exploitant le jeu d'instructions MMX d'INTEL [24]. Son rôle est essentiellement de faire la distinction entre les coins et les contours. Plus « k » est faible, plus la probabilité de confondre un point de bord pour un point de coin est élevé.

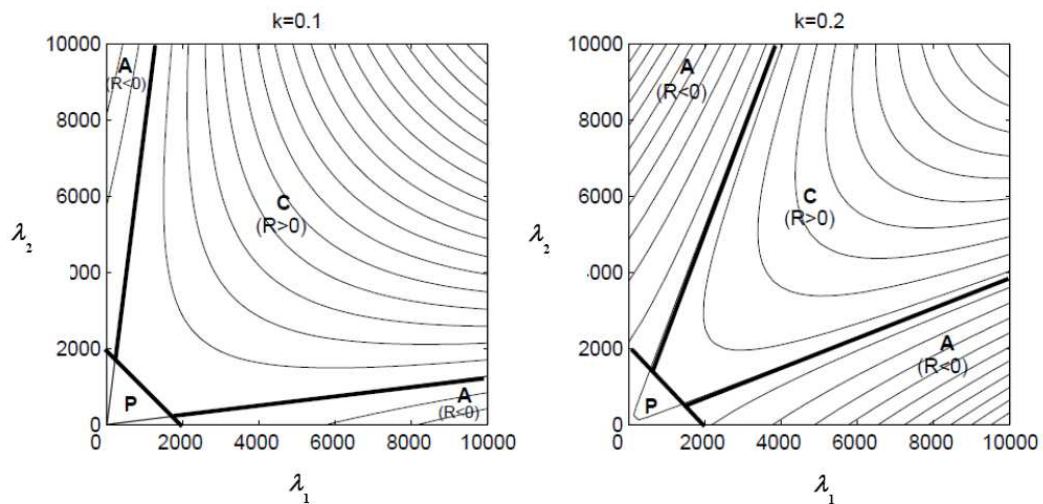


Fig.10. Effet du paramètre k sur la fonction de réponse du détecteur de Harris.

- Résumé Algorithmique :

- Calculer les images des gradients X et Y.

- Filtrer l'image par un filtre Gaussien.
- Pour chaque pixel de l'image :
 1. Calculer suivantes :
 - ❖ $\text{Trace}(M) = \lambda_1 + \lambda_2 = M_{1,1} + M_{2,2}$.
 - ❖ $\text{Déterminant}(M) = \lambda_1 \cdot \lambda_2 = M_{1,1} \cdot M_{2,2} - M_{1,2} \cdot M_{2,1}$.
 2. Et on calcule la réponse suivante :
 - ❖ $R = \text{Déterminant}(M) - k (\text{Trace}(M))^2$.
 - ❖ Les coins correspondent aux maximums locaux de R où:
 - $R > 0$
 - $\text{Trace}(M) > t$
- Constante $k = 0.04$ (typiquement).
- Constante t qui définit le seuil des valeurs propres pris en compte ($t = \lambda_1 + \lambda_2$).

- Résultats du détecteur de Harris:

Taille de la fenêtre 3×3

$k = 0.04$

Seuil de la constante $t = 0.01$

Seuil pour dessiner les points: $\min\{\text{Valeur}\{R\} + (\max\{\text{Valeur}\{R\} - \min\{\text{Valeur}\{R\}\})$

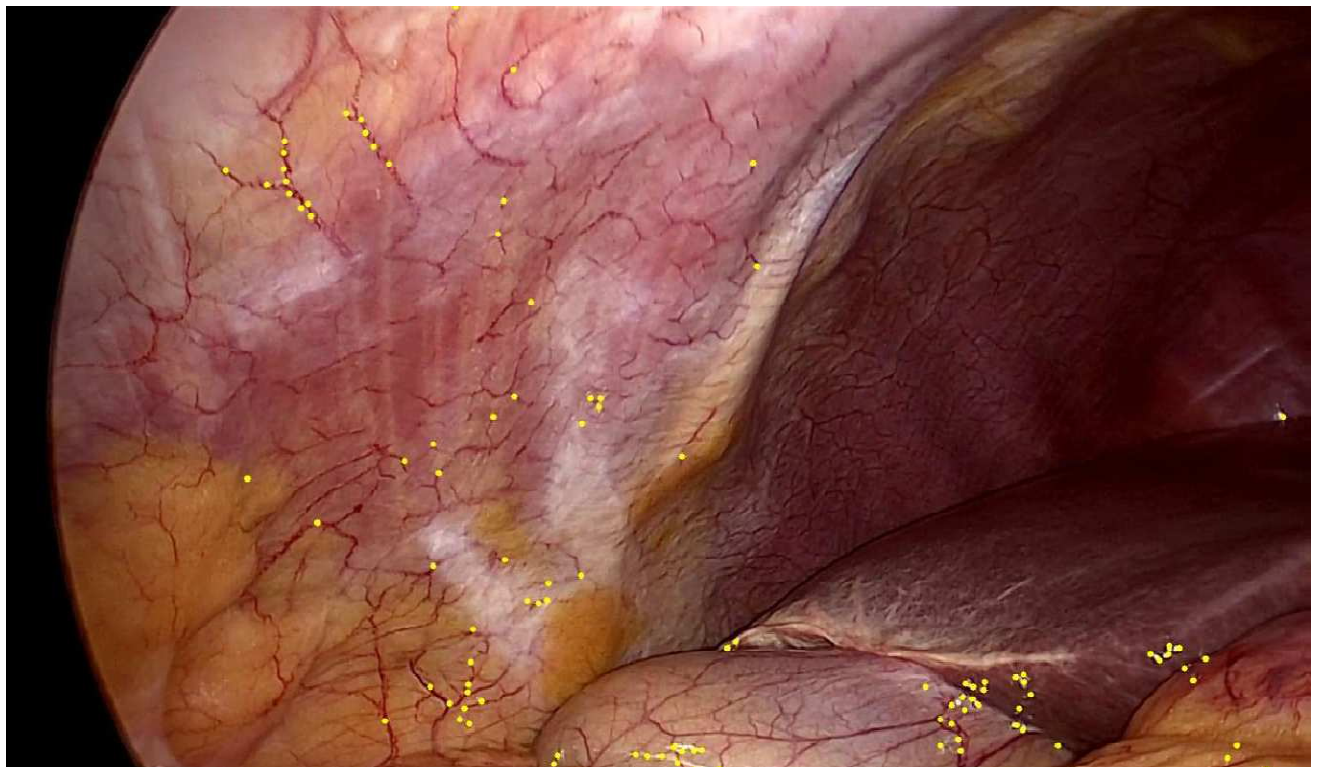


Fig.11. 116 Points détectés pour une image à l'instant (t).



Fig.12. 205 Points détectés pour une image à l'instant ($t+1$).

Nous avons appliqué le détecteur de Harris sur une séquence d'images, les deux figures 11 et 12 sont séparées par 30 millisecondes entre les deux.

- Influence du prétraitement sur le détecteur de Harris:

Par la suite nous avons appliqué une égalisation d'histogramme sur l'image en niveau de gris, avant d'appliquer le calcul des gradients, dans le but d'uniformiser le contraste sur toute l'image et de faire ressortir d'autres points saillants.

Rappelons que l'égalisation d'histogramme est une méthode d'ajustement du contraste d'une image. Elle consiste à appliquer une transformation sur chaque pixel de l'image (Transformation ponctuelle). Cette transformation est construite à partir de l'histogramme cumulé de l'image de départ. L'égalisation d'histogramme permet de mieux répartir les intensités sur l'ensemble de la plage de valeurs possibles, en «étaillant» l'histogramme. L'égalisation est intéressante pour les images dont la totalité, ou seulement une partie, est de faible contraste (l'ensemble des pixels sont d'intensité proches). La méthode est rapide, facile d'implémentation, et complètement automatique.

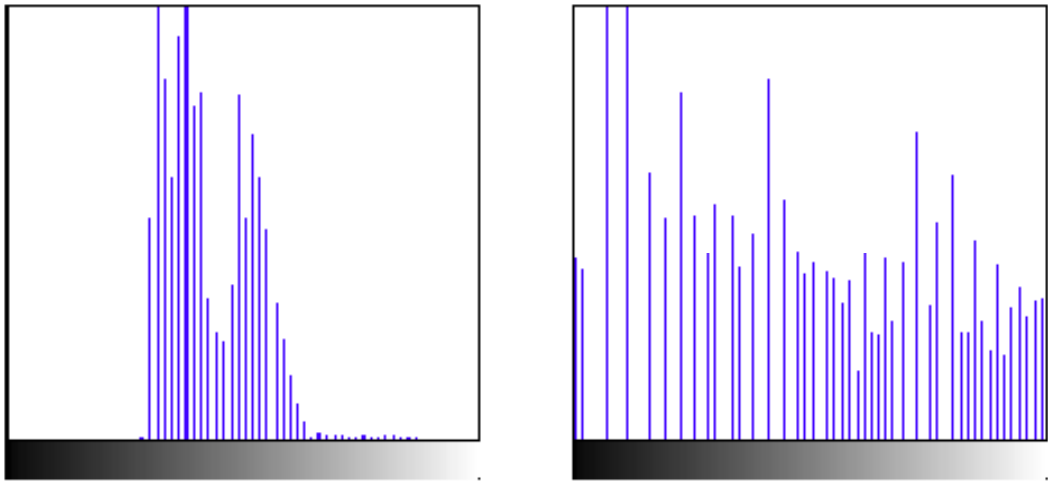


Fig.13. Egalisation d'histogramme.

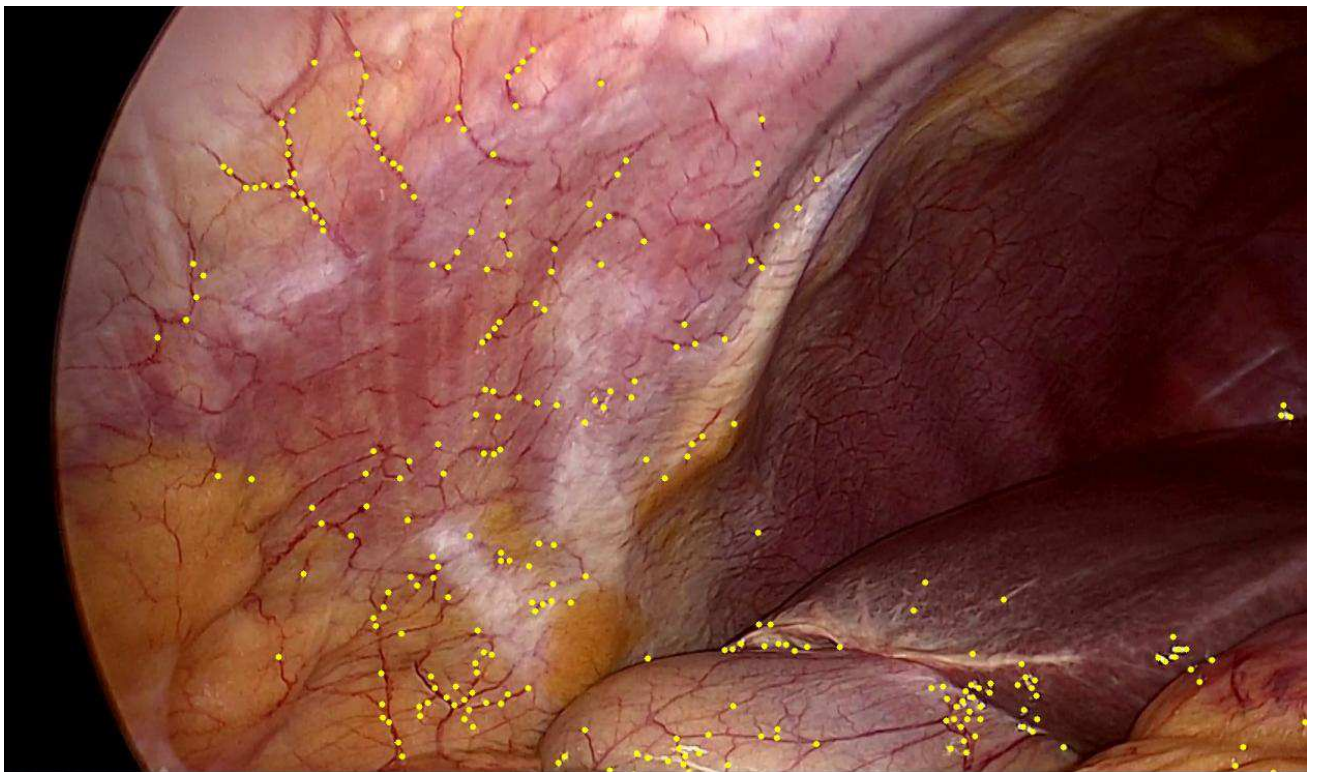


Fig.14. 309 Points détectés sur l'image à l'instant (t).

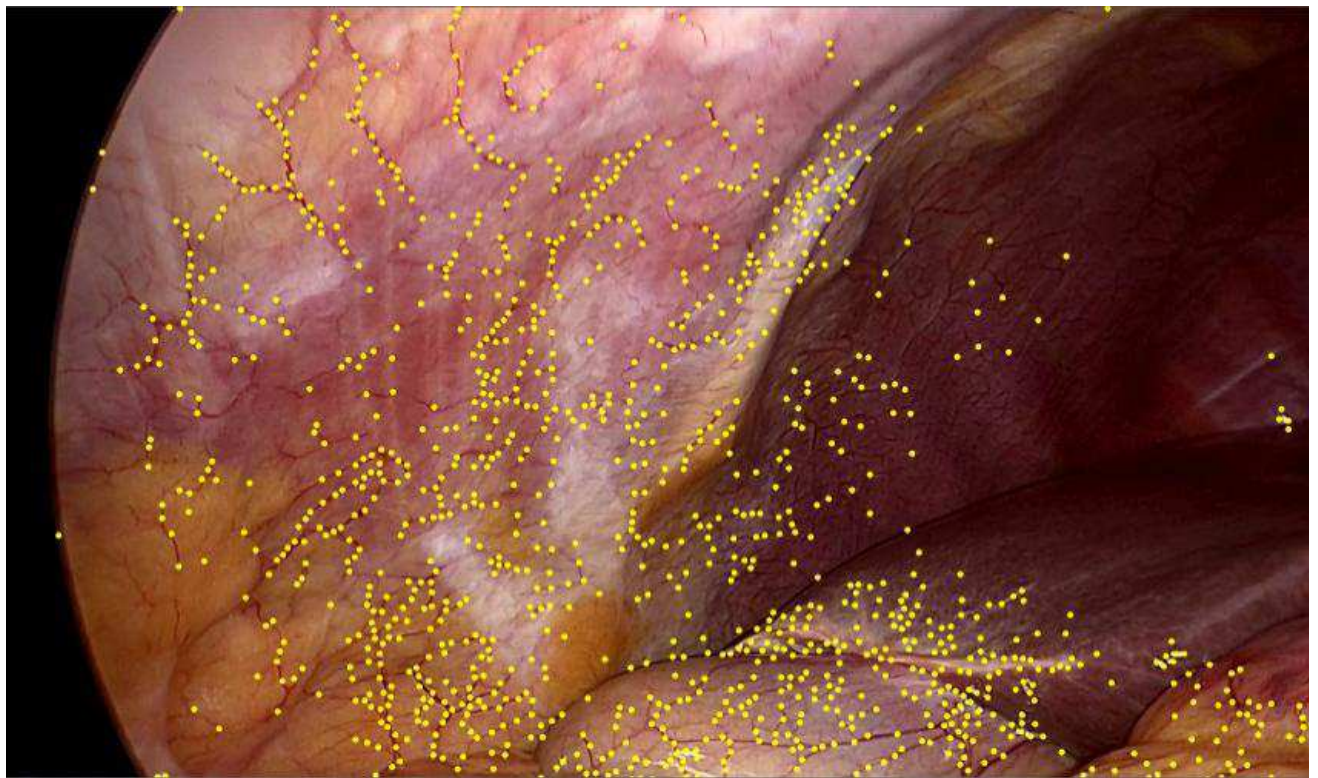


Fig.15. 484 Points détectés sur l'image à l'instant (t+1).

Sur la suite des tests nous avons fait varier le seuil de la constante t à 0.001 pour prendre plus en considération des coins avec de faible gradient dans les deux directions. Les mêmes tests avec et sans égalisation d'histogramme sur les mêmes séquences.

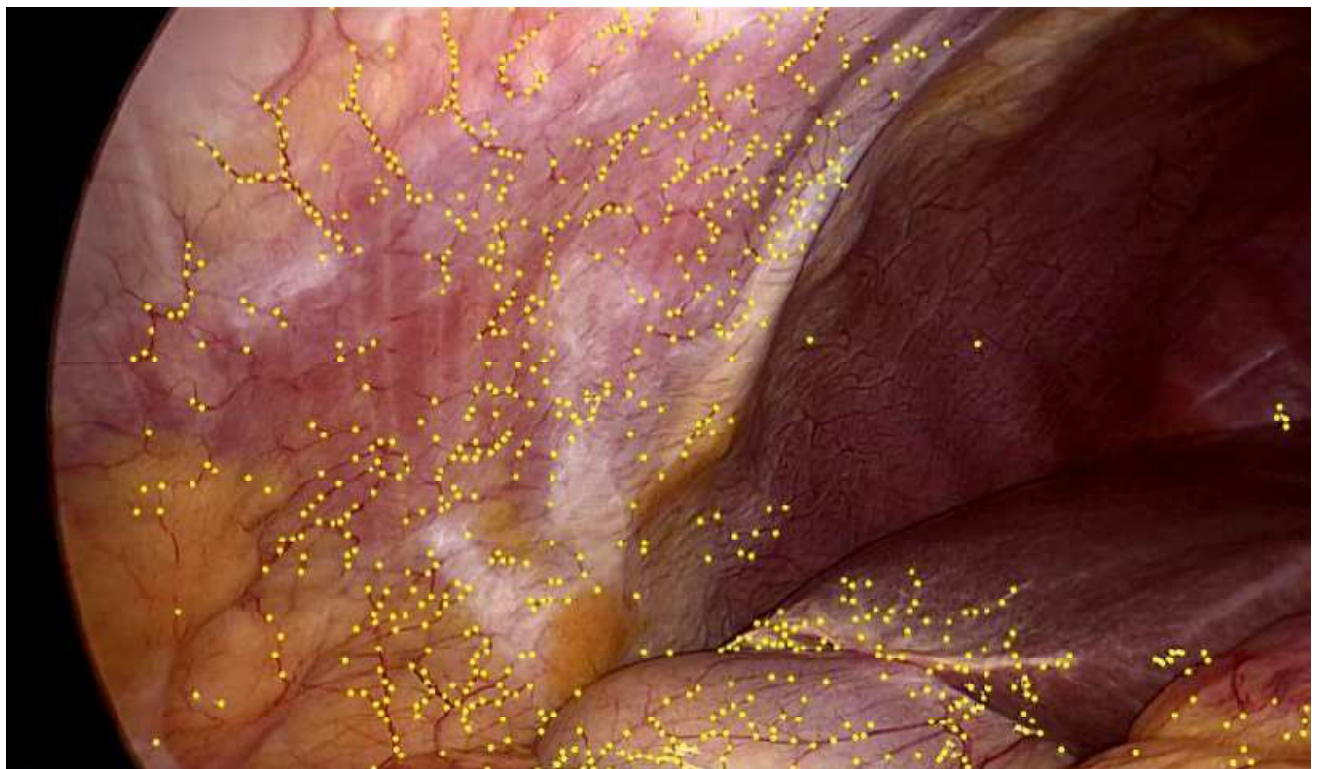


(A) Sans égalisation 658 points détectés.

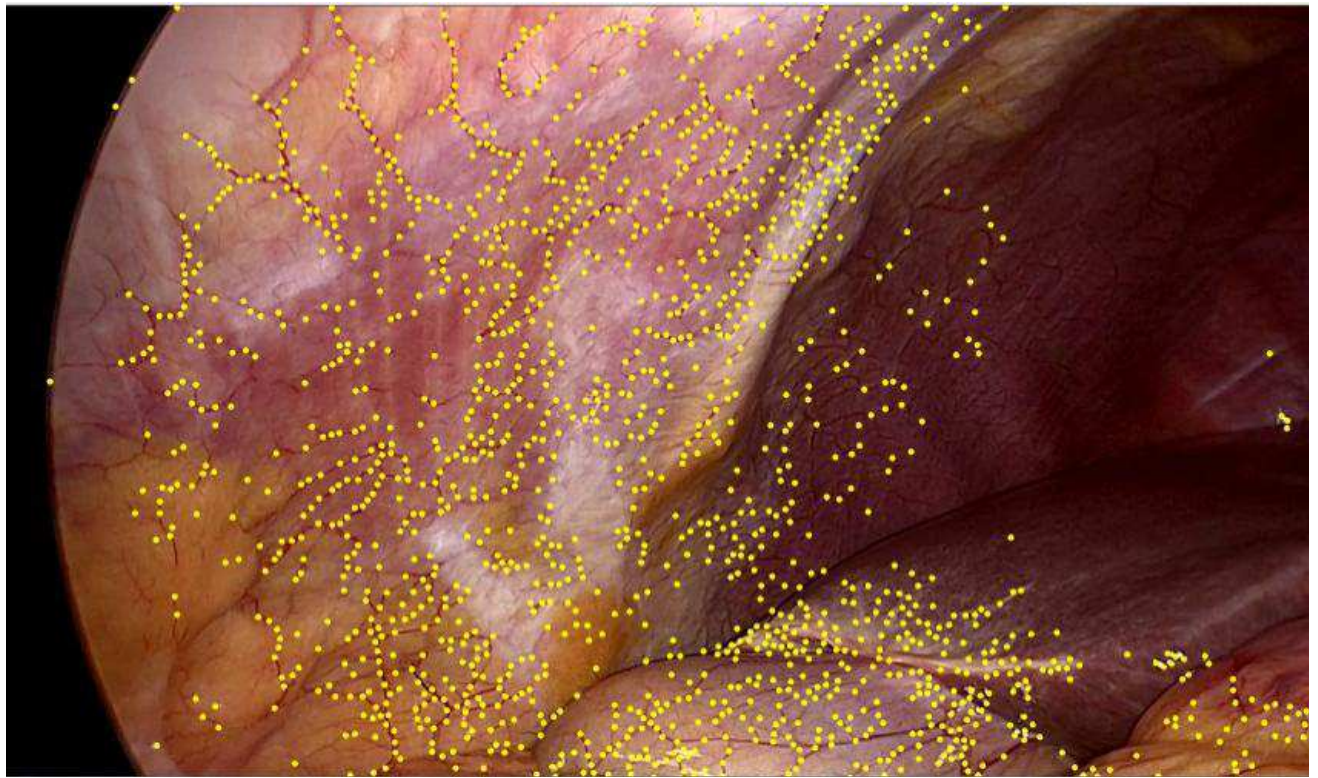


(B) Avec égalisation 1314 point détectés.

Fig.16. Points détectés sur les images à l'instant (t).



(A) Sans égalisation 901 points détectés



(B) Avec égalisation 1725 points détectés

Fig.17. Points détectés sur les images à l'instant $(t+1)$.

Il est à noter clairement qu'on arrive à détecter des points sur les régions sombres de l'image ceci est dû, que le détecteur de Harris repose sur le calcul des gradients de l'image qui est proportionnel au contraste de l'image.

Idées proposées:

Dans un but de ne garder qu'un coin détecté présentant les plus grandes valeurs propres sur une région hautement texturé, Nous imposons une distance minimale entre deux coins détectés, ceci permettra de réduire le nombre de points à suivre, lors de l'étape d'appariement des points et donc d'alléger le temps de calcul.

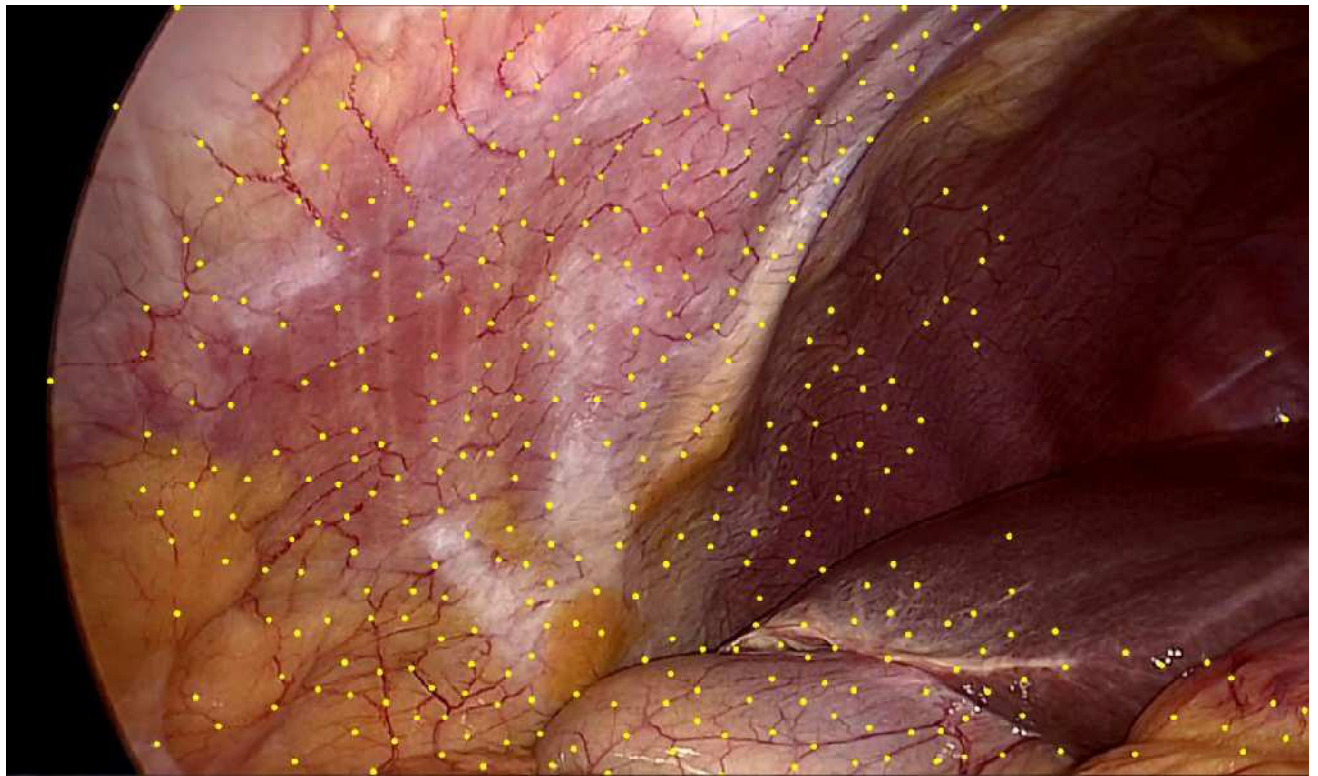


Fig.18. Points détectés avec respect d'une distance.

5.3. Le Détecteur de Shi-Tomasi

Le détecteur de coin de Shi-Tomasi [13], [14] est basé sur le détecteur de coin de Harris. Cependant, une variation dans les critères de sélection a fait que ce détecteur soit beaucoup mieux que l'original, et qu'il soit plus utilisé pour la détection des coins [25].

Le détecteur de coin de Harris opère avec un critère de sélection de coin. En effet un score est calculé pour chaque pixel et si le score est au-dessus d'une certaine valeur, le pixel est marqué comme un coin. Par ailleurs, le score est calculé en utilisant deux valeurs propres et un paramètre k qui lui, influe considérablement sur la considération d'un coin ou non. Il est attribué deux valeurs propres avec un paramètre à une fonction, qui les manipule et rend un score.

Shi et Tomasi ont suggéré que la fonction doit être reconsidérée. Seules les valeurs propres devraient être utilisées pour vérifier si le pixel était un coin ou non. Pour Shi-Tomasi, il est calculé comme ceci :

Dans un premier temps, ils recherchent une valeur seuil ϵ permettant de caractériser des points d'intérêt à partir des valeurs propres λ_1 et λ_2 de telle sorte que si $\min(\lambda_1, \lambda_2) > \epsilon$ alors le point sera considéré comme un coin.

Pour ϵ , ils considèrent une valeur moyenne qui définit la mesure d'une valeur propre pour une région qui présente un contraste uniforme et une d'une région fortement texturées la valeur de ϵ , est choisie à l'intermédiaire.

Pour calculer le minimum des valeurs propres d'une matrice carrée 2×2 , la méthode classique est de calculer le polynôme caractéristique, et de chercher ses racines.

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \text{ avec : } A = \left(\frac{\partial l}{\partial x}\right)^2 \quad B = \left(\frac{\partial l}{\partial y}\right)^2 \quad A = \left(\frac{\partial l}{\partial x}, \frac{\partial l}{\partial y}\right)$$

$$M - \lambda \cdot I_d = \begin{bmatrix} A - \lambda & C \\ C & B - \lambda \end{bmatrix} \text{ avec : } I_d \text{ la matrice d'identité}$$

Le calcul du déterminant de cette matrice donne l'équation ci-dessous:
Il suffit de déterminer les solutions de cette équation.

$$\lambda^2 - \lambda(A + B) + A \cdot B - C^2 = 0 \quad 5.8$$

C'est une équation du second ordre avec deux solutions réelles.

Dans le cas présent la matrice M, est une matrice symétrique dont les valeurs diagonales sont positives, il suffit de déterminer une seule solution qui n'est que la plus petite valeur pour résoudre le système de ce cas.

La plus petite solution de l'équation **(5.8)** est:

$$x = +B/2 - \sqrt{(A/2 + B/2)^2 - (A \cdot B - C^2)}$$

donc elle représente la valeur minimale de (λ_1, λ_2)

Dans leur papier, Shi et Tomasi ont démontré expérimentalement que ce critère de score était bien meilleur. Si ϵ est plus grand qu'une certaine valeur prédéfinie, il peut être marqué comme un coin.

Ainsi, la région d'effet pour un point d'être un coin est particulièrement simplifiée en contraignant une intersection entre deux régions pour obtenir la figure ci-dessous:

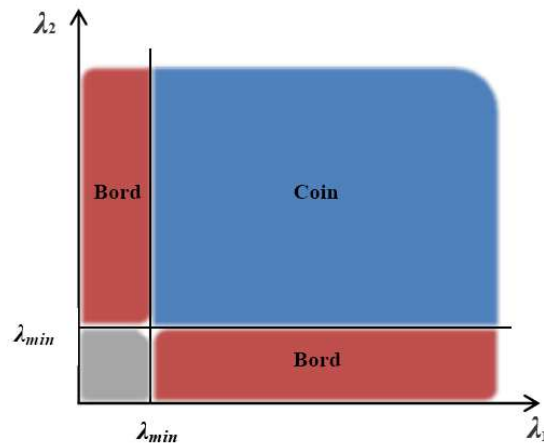


Fig.19. Réponse de Shi-Thomasi suivant les valeurs propres et le seuil λ_{min}

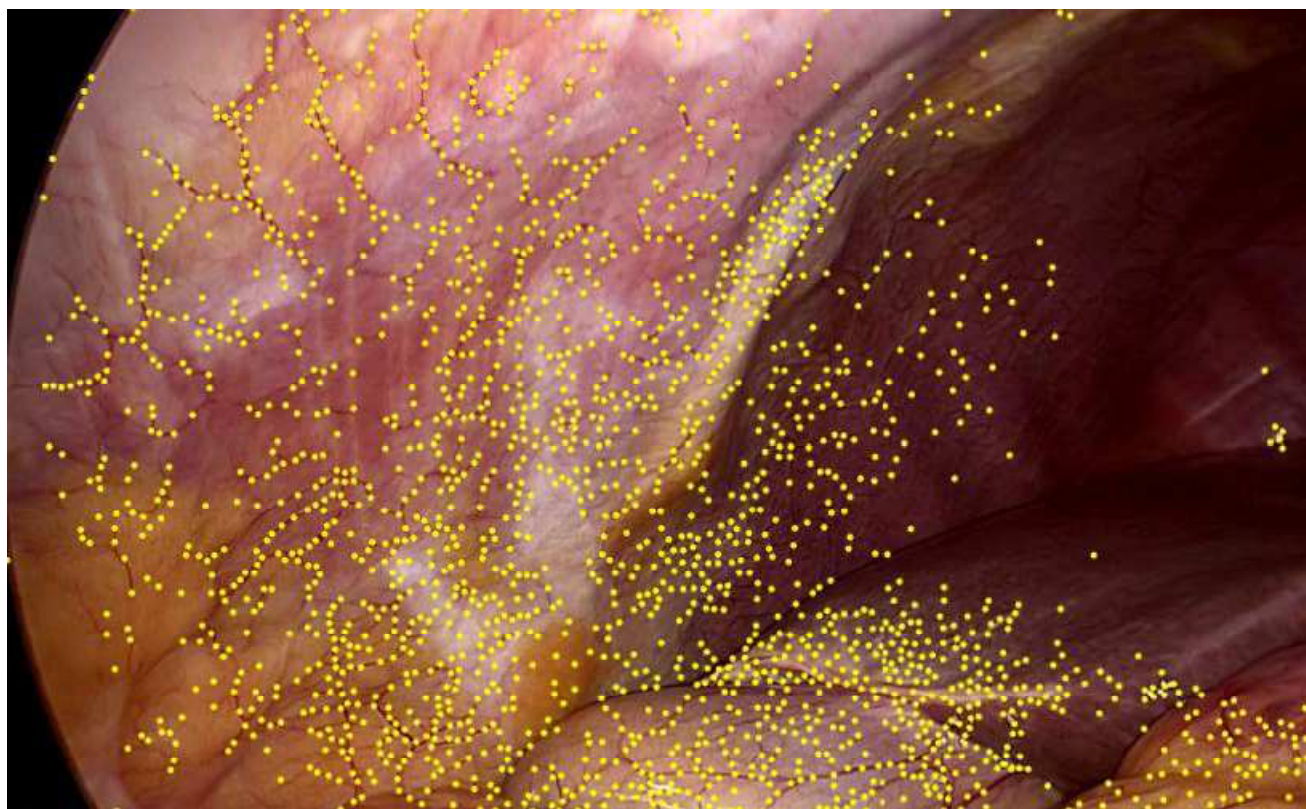
- Résumé Algorithmique :

- Déterminer ϵ
- Calculer les images des gradients X et Y
- Filtrer l'image par un filtre Gaussien
- Pour chaque pixel de l'image :
 - 1- Calculer la matrice M sur une fenêtre de taille définie
 - 2- On détermine le minimum de (λ_1, λ_2)
 - Si $\epsilon < \min(\lambda_1, \lambda_2)$:
 - 3- On retient Les coins correspondent aux maximums locaux de $\min(\lambda_1, \lambda_2)$

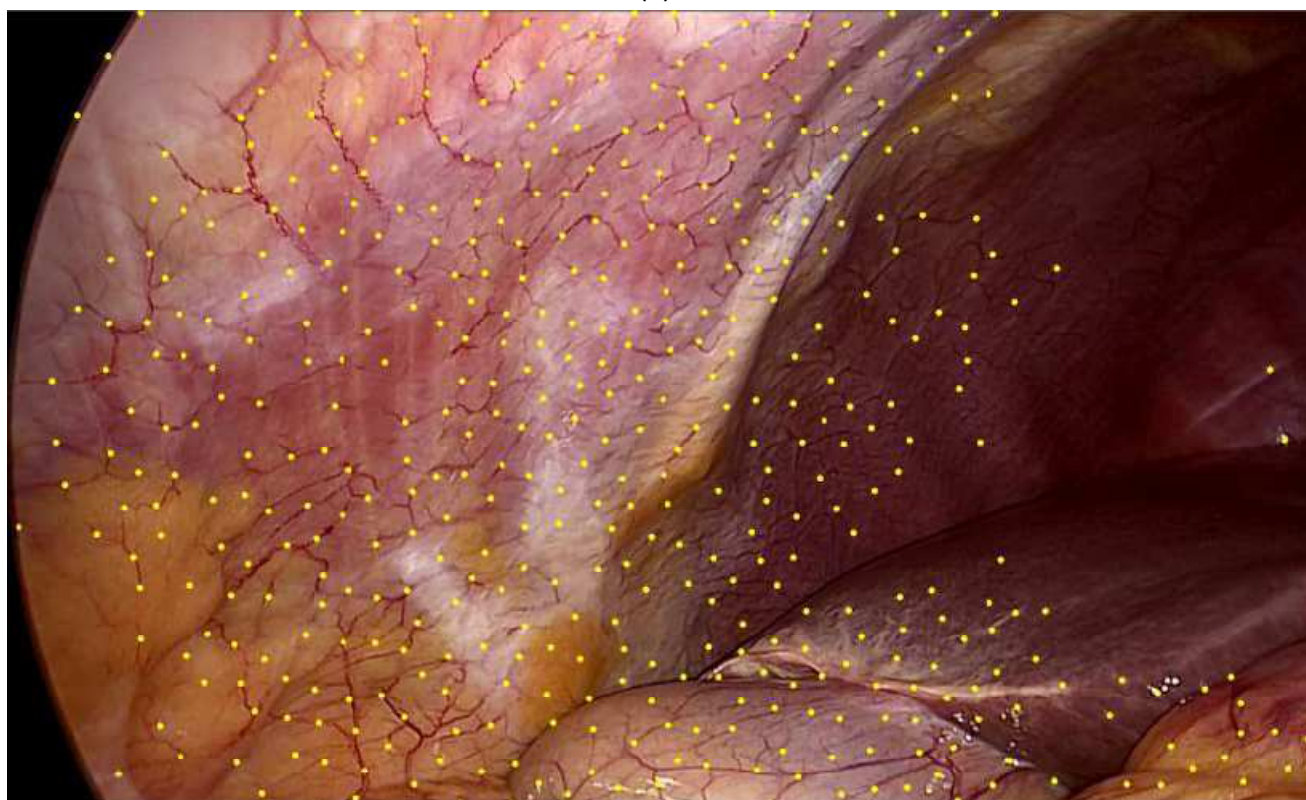
- Résultats du détecteur de Shi –Tomasi:

Taille de la fenêtre 3×3

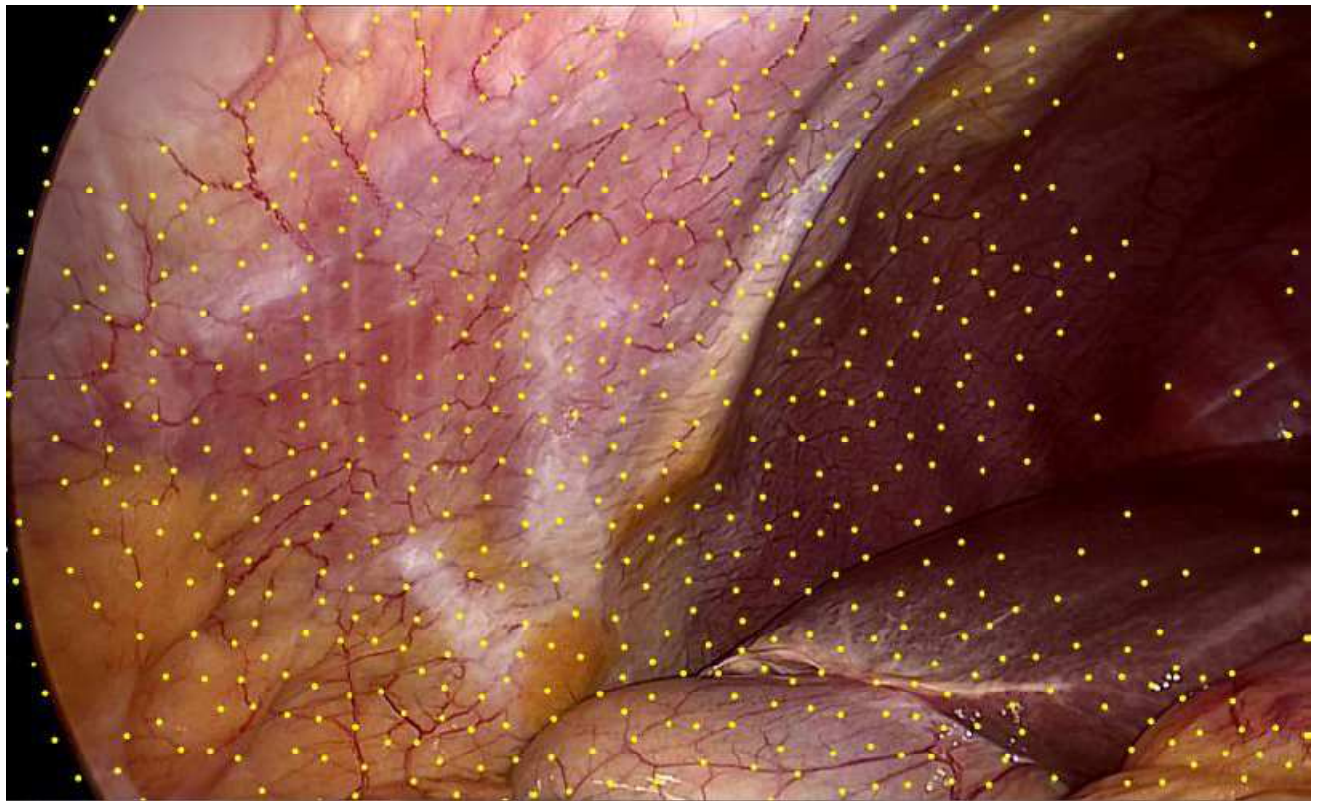
$\epsilon = 0.01$



(a)



(b)



(c)

Fig.20. Résultat détecteur de Shi-Tomasi : (a) distance minimum 10, (b) distance minimum 35, (c) avec une égalisation d'histogramme et distance minimum 35.

Avec le détecteur de Shi-Tomasi, on arrive à détecter des points d'intérêts sur des régions où le gradient est pourtant faible dans les deux directions, ces régions peuvent être prises comme des régions homogènes avec le détecteur de Harris si on n'adapte pas le seuil t .

6. Appariement des points:

Une fois que les points sont extraits dans une image, plusieurs stratégies sont possibles pour les appairer. On s'intéresse à l'algorithme **de Lucas Kanade**[14] qui essaye itérativement de réduire au minimum la différence entre deux fenêtres de référence sur une zone de recherche de l'image. La technique peut être utilisée pour l'alignement d'image, le dépistage, l'analyse de flux optique et l'évaluation de mouvement.

Soit I et J deux images en niveau de gris. Les quantités $I(\mathbf{x}) = I(x, y)$ et $J(\mathbf{x}) = J(x, y)$ représentent la valeur en niveau de gris du point $\mathbf{x} = [x, y]^T$ où x et y sont les coordonnées en pixel de l'image. La coordonnée du coin gauche supérieure correspond donc à $[0, 0]^T$. L'image I sera référencée comme première image et J la suivante. Soit n_x et n_y la largeur et la hauteur de l'image, donc la coordonnée du coin droit inférieur correspond à $[n_x-1, n_y-1]^T$.

On considère un point $\mathbf{u} = [u_x, u_y]^T$ dans l'image I . Le but du suivi des points d'intérêts est de trouver la coordonnée du point $\mathbf{v} = \mathbf{u} + \mathbf{d} = [u_x + dx, u_y + dy]^T$ tel que $I(\mathbf{u})$ et $J(\mathbf{v})$ sont similaires. Le vecteur $\mathbf{d} = [dx, dy]^T$ correspond à la vitesse de l'image au point $\mathbf{x}$, aussi connu sous flux optique au point $\mathbf{x}$. À cause du problème d'ouverture il est essentiel de définir la notion de similarité dans le voisinage 2D. Soit w_x et w_y deux entiers qui déterminent la taille de la fenêtre, on définit la vitesse d'image $\mathbf{d}$ comme un vecteur minimisant la fonction résiduel ci-dessous:

$$F(d) = \sum_{x=ux-wx}^{ux+wx} \sum_{y=uy-wy}^{uy+wy} (I(x, y) - I(x + dx, y + dy))^2 \quad 6.1$$

Cette fonction aussi appelée la somme des différences au carrée (SSD) de deux blocs comme illustre la figure ci-dessous:

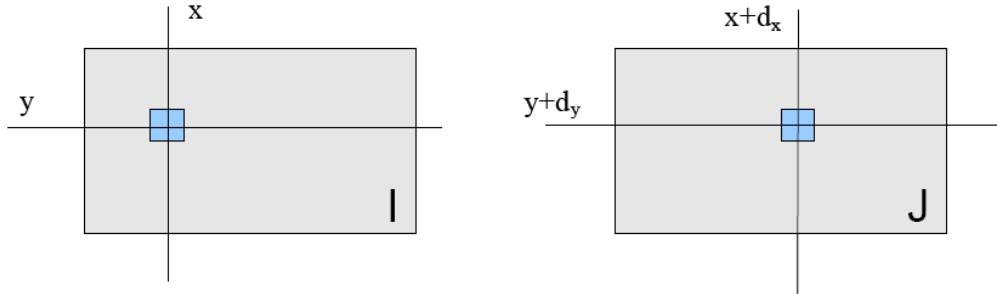


Fig. 22. Calcul de la SSD sur une taille de fenêtre prédéfinie pour un point au coordonnées (x, y)

6.1. Algorithme à une dimension:

La problématique d'appariement des points d'intérêts [12][15] peut être décrite comme suit :

On considère des fonctions $F(x)$ et $G(x)$ qui donne les valeurs de pixel respectives à chaque emplacement x dans deux images à des instant $(t, t+1)$, où x est un vecteur.

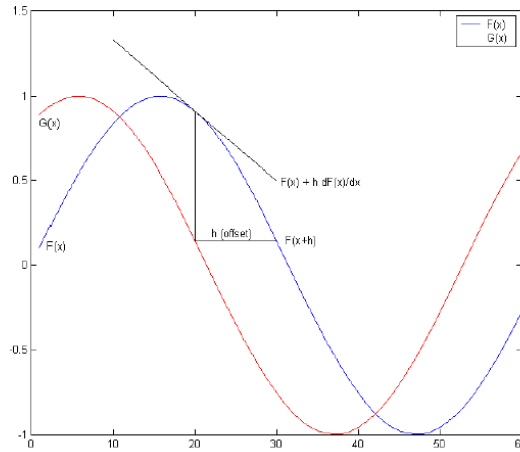


Fig.23. Appariement dans le cas d'une seule dimension

Il suffit de trouver le vecteur de disparité $h = \frac{G(x) - F(x)}{F'(x)} = 0$ **(6.2)** qui réduit au minimum une certaine mesure (E) de la différence entre $F(x + h)$ et $G(x)$, pour x dans une certaine région d'intérêt R.

$$E = \sum_x (F(x + h) - G(x))^2 \quad 6.3$$

Pour minimiser l'erreur de h on pose $\frac{dE}{dh} = 0$, et aussi une approximation linéaire de la fonction. $F(x + h) = F(x) + hF'(x)$ qui permet non seulement de généraliser l'aspect à deux dimensions mais aussi de parer le cas où la fonction **(6.2)** serait non défini pour $F'(x)=0$.

$$h = \frac{\sum_x F'(x)[G(x)-F(x)]}{\sum_x F'(x)^2} \quad 6.4$$

L'algorithme commence par une estimation du vecteur h , où le gradient d'intensité spatial à chaque point de l'image, va permettre de modifier l'évaluation actuelle de h pour obtenir un h qui apporte un meilleur matching. C'est un processus itératif qui fait appel à la méthode de Newton-Raphson.

$$h_{k+1} = h_k + \frac{\sum_x W(x)F'(x+h_k)[G(x)-F(x+h_k)]}{\sum_x W(x).F'(x+h_k)^2} \quad 6.5$$

$$\text{Où } W(x) = \frac{1}{|G'(x)-F'(x)|} \quad 6.6$$

Pour élargir le domaine de convergence de l'algorithme on élimine les fortes transitions (fréquences hautes) dans l'image, ceci peut être accompli par un lissage de l'image, c'est à dire en remplaçant chaque pixel de l'image par une moyenne pondérée des pixels voisins. Le compromis de ce lissage est qu'il supprime les petits détails et rend ainsi le matching moins précis.

Il faut savoir aussi que si la fenêtre considérée pour le lissage est beaucoup plus grande que la taille de l'objet auquel nous essayons de correspondre, l'objet peut être supprimé entièrement et donc aucun suivi ne sera possible.

Par ailleurs en considérant la fonction **(6.6)** qui n'est que la proportion inverse à une évaluation de $F''(x)$, Ils améliorent la fonction **(6.5)** en tenant compte que l'approximation linéaire dans **(6.2)** est bonne où $F'(x)$ est presque linéaire et est au contraire plus mauvaise où $F''(x)$ est grande donc le calcul du paramètre h devient plus précis ce qui permet d'avoir une convergence plus rapide.

6.2. Généralisation à plusieurs dimensions:

$$h = [\sum_x (\frac{\partial F}{\partial x})^T [G(x) - F(x)]][\sum_x (\frac{\partial F}{\partial x})^T (\frac{\partial F}{\partial x})]^{-1} \quad (6.7)$$

$$\text{où } \frac{\partial}{\partial} x = [\frac{\partial}{\partial} x_1 \frac{\partial}{\partial} x_2 \dots \frac{\partial}{\partial} x_n]^T$$

$G(x) = F(xA + h)$, où A exprime la matrice de transformation linéaire entre $F(x)$ et $G(x)$. En utilisant l'aspect quadratique lors de l'approximation linéaire ça nous permet non seulement de gérer l'aspect de translation mais aussi celui de la rotation le changement d'échelle et l'étirement d'image.

$$E = \sum_x [F(xA + h) - G(x)]^2 \quad 6.8$$

Pour déterminer la quantité ΔA et Δh ils utiliseront l'approximation linéaire

$$F(x(A + \Delta A) + (h + \Delta h)) = F(xA + h) + (x\Delta A + \Delta h) \frac{\partial F(x)}{\partial x} \quad 6.9$$

7. Reconstruction de points 3D:

La reconstruction 3D à partir d'images, désigne la technique qui permet d'avoir une représentation tridimensionnelle d'un objet ou d'une scène à partir d'un ensemble d'images prises sous différents points de vue de l'objet ou de la scène.

D'une manière plus générale, Le principe de la reconstruction 3D est le suivant: on dispose d'une ou plusieurs représentations en 2D d'un objet et on souhaite déterminer les coordonnées des éléments visibles sur ces représentations dans un repère de l'espace réel 3D. Les méthodes proposées dans cette catégorie tendent à suivre des indices visuels (points d'intérêt) qu'on peut trouver dans une image comme la texture, les ombres ou le mouvement, pour calculer la structure 3D de la scène. Quatre techniques sont généralement utilisées pour résoudre ce type de problème [26] : la stéréovision, la structure à partir du mouvement (*Structure from Motion (SFM)*), les ombres portées (*Shape from Shading (SFS)*) et la stéréo photométrie.

Dans le cadre de notre travail où les représentations 2D sont des images représentant le tissu abdominal, il s'agit de trouver la profondeur des points visibles sur l'image, car la notion de profondeur, est perdue durant le processus de projection qui produit l'image.

Les points visibles sur les images sont les projections des points réels qu'on peut alors situer sur des droites (droite épipolaire). Si deux ou plusieurs vues de l'objet sont prises, la position dans l'espace des points réels peut alors être obtenue par intersection de ces droites: c'est le principe de triangulation fondamental de toute reconstruction 3D à partir d'images. La figure reprend

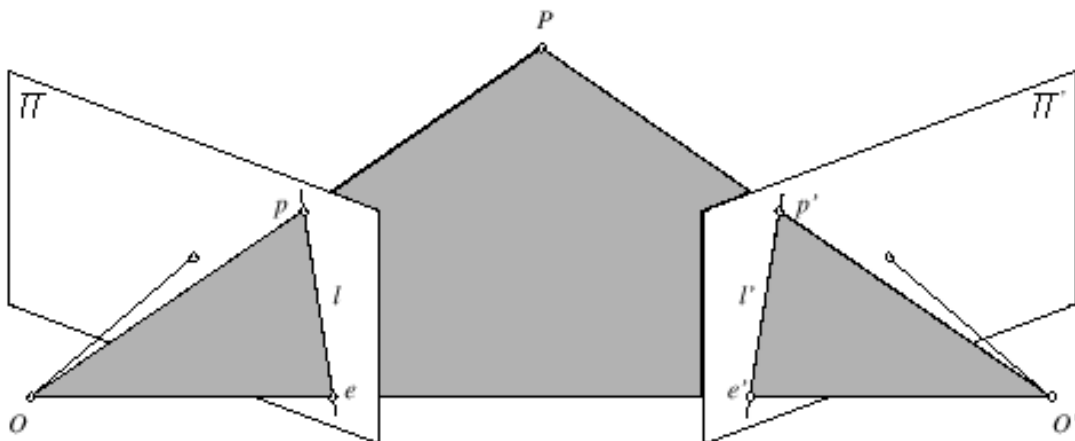


Fig. 24. Triangulation d'un point

- 5 points dans le même plan: P, p, p', O et O'
- Les droites l et l' sont les droites épipolaires (conjuguées)
- Les points e et e' sont respectivement les épipôles gauche et droit
- e est en fait O' vu dans l'image gauche
- e' est O vu dans l'image droite

7.1. Structure From Motion

Cette technique a pour but de faire une reconstruction 3D (*Structure*) de la scène, à l'aide d'une seule caméra en mouvement et d'estimer son déplacement entre les prises de vue (*Motion*). Les correspondances sont alors calculées dans la séquence vidéo résultante.

Le principe de la méthode est le suivant:

Tout d'abord, deux images sont sélectionnées parmi les différentes vues de l'objet et un repère de la scène initial est défini. Ensuite, la position et l'orientation des caméras à l'origine des autres images sont déterminés dans ce repère grâce aux associations de points. Des nouveaux points 3D sont alors ajoutés à la reconstruction. À chaque étape, la reconstruction est affinée et enrichie. Une fois la reconstruction effectuée et la position relative des caméras retrouvée, des ajustements peuvent être réalisés.

8. Conclusion:

Dans cette première partie d'étude, nous avons présenté les premières étapes pour établir un modèle 3D surfacique, qui suit la déformation des tissus, ceci par une détection de points ensuite par un suivi de ces points d'image en image.

Par la suite nous validerons la robustesse du suivi de ces points à l'aide des techniques de géométrie épipolaire, et nous aborderons l'étape de reconstruction des surfaces à l'aide des techniques de structure pour la motion à partir du suivi des points détectés.

Néanmoins, il est à noter que la robustesse de la vision par ordinateur en chirurgie laparoscopique, est affectée par un certain nombre de facteurs incluant le manque de caractéristiques, les régions spéculaires, le mouvement de la caméra, la déformation tissulaire, la fumée chirurgicale et l'occlusion, ce qui donc nécessite l'élaboration d'algorithmes robustes donnant des résultats précis, ce qui va nous amener à revoir constamment la validité de ces algorithmes.

Références

- [1] Azuma R., Baillyot., Behringer R. *et al.* Recent Advances in Augmented Reality. cga, November 2001.
- [2] Milgram P. et Kishino F. A Taxinomy of Mixed Reality Visual Displays, *IEICE Trans. Information Systems*, vol. E77-D, no. 12, p. 1321–1329, 1994.
- [3] D.T. Gering, A. Nabavi, R. Kikinis, W. Eric L. Grimson, N. Hata, P. Everett, F. Jolesz. An Inte-grated Visualization System for Surgical Planning and Guidance using Image Fusion and Interven-tional Imaging. In second International Conference on Medical Image computing And Computer-Assisted Intervention, pages 808-819,1999
- [4] Gumprecht HK, Widenka DC, Lumenta CB BrainLab VectorVision Neuronavigation System: technology and clinical experiences.Technology and clinical experiences in 131 cases (1999)
- [5] B.Peuchot A. Tanguy, M. Eude. Virtual Reality as an Operative Tool During Scoliosis Surgery .in first International Conference on computer Vision,Virtual reality and robotics in Medicine (CVRMed'95), pages 550-554, Nice, France 1995.
- [6] P.Merloz, F. Lavalley, J. Tonneti , L. Pittet . Image-guided spinal surgery:Technology, operative technique, and clinical practice. Operative Technique in Orthopaedics, 2000.
- [7] Warfield, S. K., Haker, S. J., Talos, I.-F., Kemper, C. A., Weisenfeld, N., Mewes, A. U. J., et al. (2005). Capturing intraoperative deformations: research experience at Brigham and Women's Hospital. *Medical image analysis*, 9(2), pages 145-62.
- [8] Nicolau, S., Garcia, A., Pennec, X., Soler, L., Ayache, N., & Fasquel, J.-b. Guidage de ponctions percutanées à l'aide d 'un système de vision utilisant une méthode de recalage 3D / 2D. *Medical Image Analysis*, 3-5, 2003.
- [9] Mountney, P., Stoyanov, D., & Yang, G.-zhong. (2010). Introducing techniques based on laparo-scopic or endoscopic images]. *IEEE Signal Processing Magazine*, (July), 14-24.
- [10] Harris, C., & Stephens, M. (1988). A combiend Corner AndEdge Detector, 147-152.
- [11] Derpanis, K. G. (2004). The Harris Corner Detector. *International Journal*, 2-3.
- [12] Lucas, B. D. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. *Imaging*, 130, 121-129.
- [13] Tomasi, C. (1991). Detection and Tracking of Point Features Technical Report CMU-CS-91-132. *Image (Rochester, N.Y.)*, (April).
- [14] Jianbo Shi & Carlo Tomasi. Good Feature To Track, IEEE CCVPR 1994.
- [15] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 56, no. 1, pp. 221 –255, March 2004.
- [16] Mountney, P., & Yang, G.-Z. (2009). Dynamic view expansion for minimally invasive surgery using simultaneous localization and mapping.Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, 2009.
- [17] Mountney, P., Stoyanov, D., Davison, A., & Yang, G.-Z. (2006). Simultaneous stereoscope localization and soft-tissue mapping for minimal invasive surgery: MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 347-54.

- [18] Mountney, P., & Yang, G.-Z. (2010). Motion compensated SLAM for image guided surgery. Medical image computing and computer-assisted intervention: MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention, 13(Pt 2), pages 496-504.
- [19] Zisserman, Fitzgibbon, & Cross, G. VHS to VRML: 3D graphical models from video sequences. Proceedings IEEE International Conference on Multimedia Computing and Systems, pages 51-57. IEEE Comput, 1999.
- [20] Brostow, G. J., Shotton, J., Fauqueur, J., & Cipolla, R. Segmentation and Recognition using Structure from Motion Point Clouds, pages 1-14, 2008.
- [21] Hu, M., Penney, G., Edwards, P., Figl, M., & Hawkes, D. (2007). 3D reconstruction of internal organ surfaces for minimal invasive surgery. MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 68-77.
- [22] Bue, A. D., Science, C., Mary, Q., Agapito, L., & Aguiar, P. M. Q. Non-rigid structure from motion using quadratic deformation models, 2009.
- [23] Bregler, C., Hertzmann, a, & Biermann, H. Recovering non-rigid 3D shape from image streams. Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662), pages 690-696. IEEE Comput.
- [24] Pascaline PARISOT, Suivi d'objets dans des séquences d'images de scènes déformables –De l'importance des points d'intérêt et du maillage 2D, Thèse soutenu le 23 janvier 2009.
- [25] <http://www.aishack.in/2010/04/interesting-windows-in-the-harris-corner-detector/>
- [26] S. M. Seitz. «An Overview of Passive Vision Techniques». *Course Notes, The Robotics Institute, Carnegie Mellon University*, pages 1–3.

Description de l'application du suivi des points

Introduction:

Dans le but d'établir une reconstruction 3D déformable du tissu abdominal, en utilisant les techniques de Structure From Motion à partir d'une séquence d'images 2D, une première étape est la détection des points d'intérêts et le suivi dans le temps de ces points. Pour ce fait nous avons élaboré une application de détection et de suivi des points d'intérêts. Pour ce fait l'application permet la lecture d'un fichier vidéo sur laquelle on effectue notre traitement, à la fin du processus du suivi on génère un fichier pour chaque image qui contient les coordonnées des points pour l'étape de reconstruction. On rappelle que tout le processus de détection et du suivi des points d'intérêts suit notre approche sur l'analyse du signal RGB.

Description:

Les différents processus de l'application sont sous forme d'états, cela permet d'avoir différents modes de fonctionnement, décomposés sous forme de branches au sein de la même application afin de faciliter par la suite la lecture et ou la modification du code. Le choix et la sélection des modes s'effectuent par une assignation de touche au clavier. Dans un premier temps on charge une vidéo sur laquelle on veut effectuer notre traitement. Au lancement de l'application on passe directement en pause dans le mode Edit. La lecture et ou le processus du traitement peut être fait en mode lecture continue ou en mode séquencé image par image. On peut interagir avec ces deux modes par des touches du clavier, par exemple la touche `p` pour le mode en lecture continue et `n` pour le mode image par image.

L'initialisation et la modification des points ce fait dans le mode Edit de l'application. Elle peut se faire soit à l'aide de la souris en ajoutant ou en enlevant des points dans le mode manuel, ou bien à l'aide des détecteurs de Harris ou de Shi-Tomasi dans le mode automatique, ou bien encore une combinaison des deux modes (automatique et manuel). L'état Edit met la lecture et le traitement de la vidéo en pause. La détection automatique se fait sur une ROI de l'image voir Fig.1. L'étape du suivi se fait aussi sur cette même ROI, le point n'est plus suivi en dehors de cette zone et il est retenu comme étant un point perdu. On met ses coordonnées à (-1, -1), on peut aussi désigner un point pour qu'il ne soit plus suivi en mettant ses coordonnées à (-1, -1) ou bien en le supprimant. Le fait de mettre les coordonnées à (-1, -1) va nous renseigner dans le processus de reconstruction que le point a était perdu dans le frame correspondant.

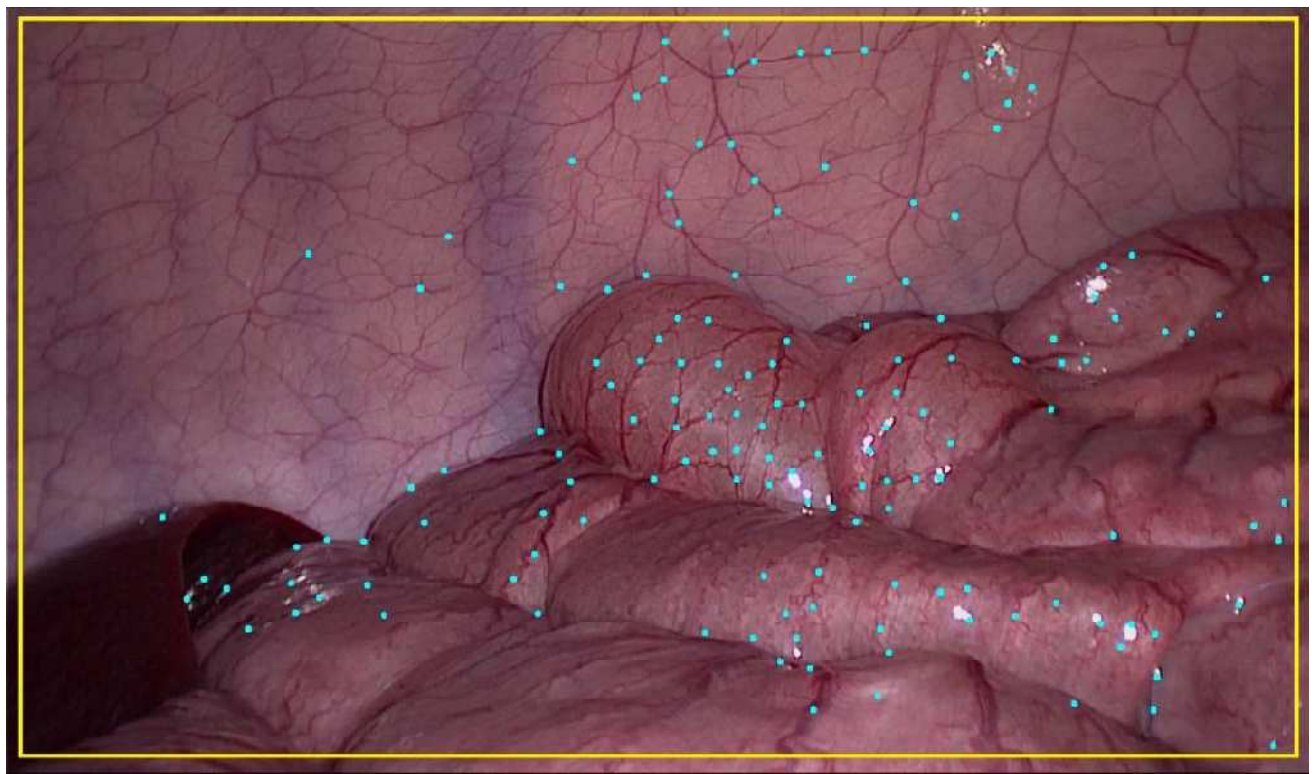


Fig.1. Détection des points dans une ROI.

Dans la partie de détection automatique des points on a ajouté la fonctionnalité d'enlever les points spéculaires du traitement de la détection des points d'intérêts, ceci à l'aide d'une image masque, des zones hautement lumineuse (forte intensité) voir Fig.2. On rappelle que les points spéculaires sont à l'origine des points perdus en suivi (déplacement de la lumière) et des artéfacts dans l'étape de reconstruction.

La particularité du mode automatique dans le suivi des points, est qu'on réintègre de nouveaux points à partir d'un certain seuil de points perdu dans le suivi, ceci sans repasser par l'état Edit et en gardant toujours les anciens points qui sont toujours suivi dans la séquence, tout étant en mode lecture séquencé ou en continue.

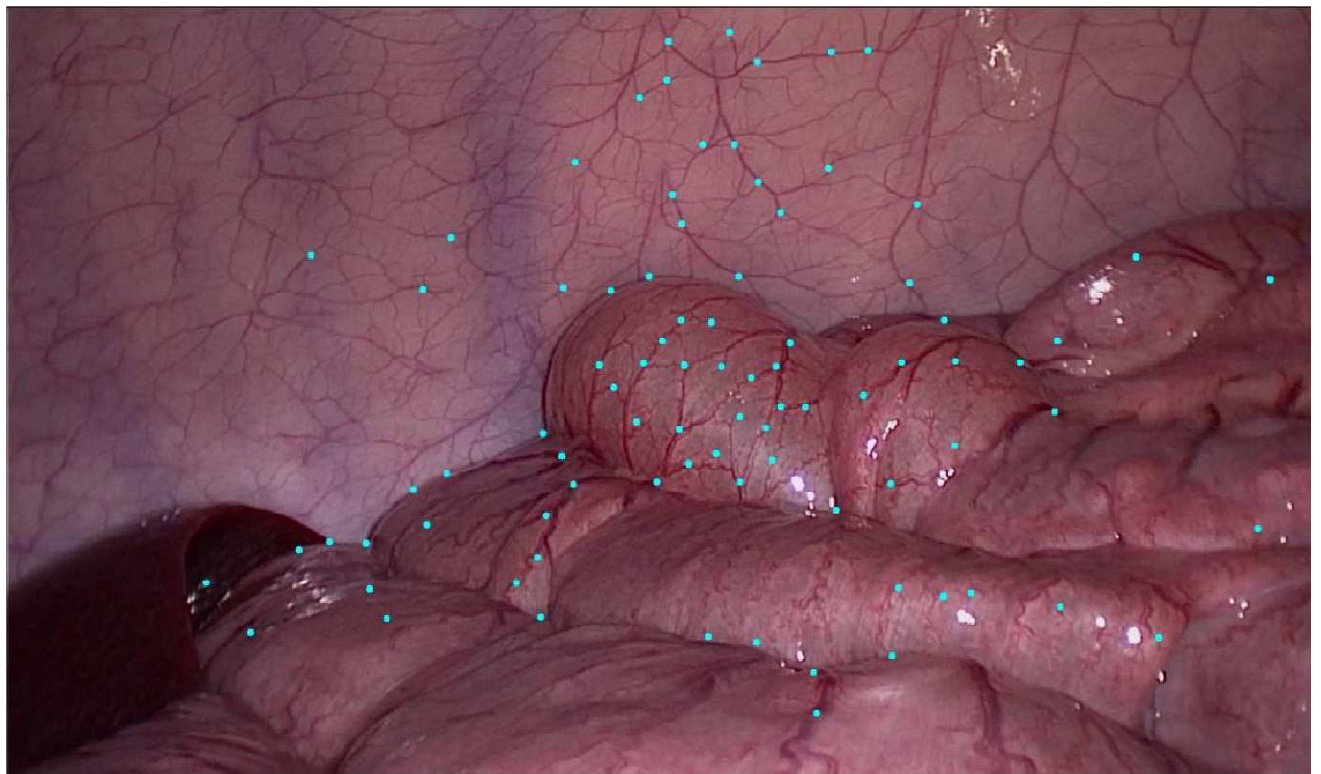
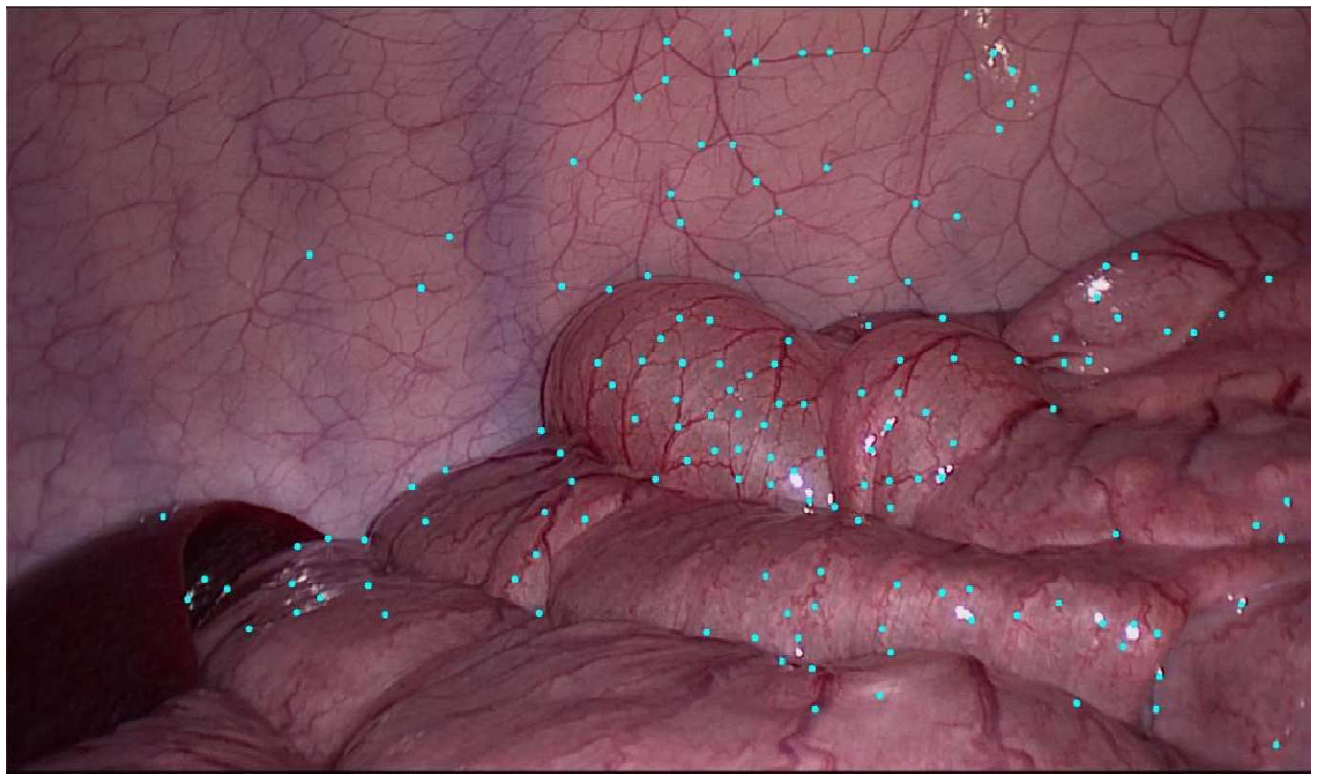


Fig.2. En haut détection normal, en bas détection sans les points spéculaires.

Schéma:

Le schéma ci-dessous Fig.3, reprend l'aspect général de la structure avec les différents modes de l'application.

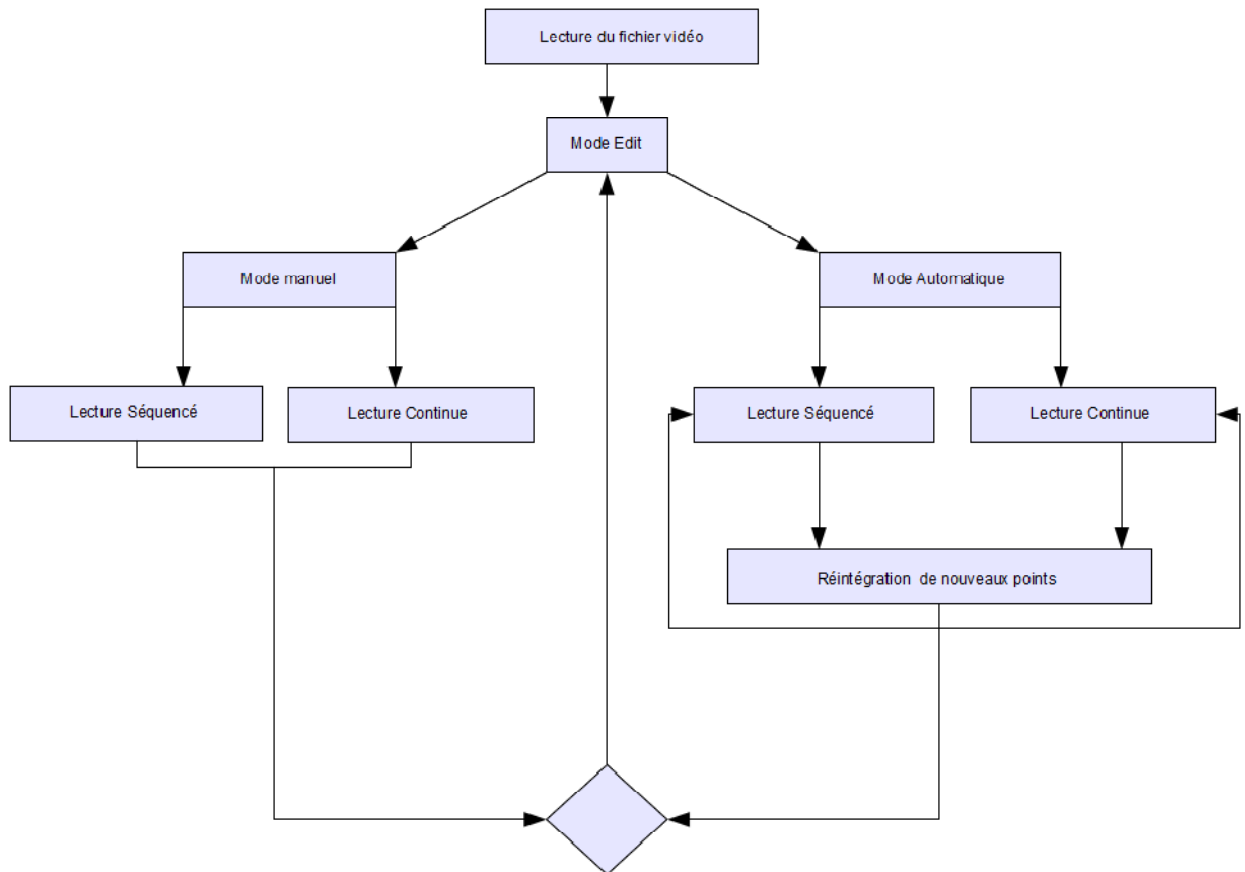


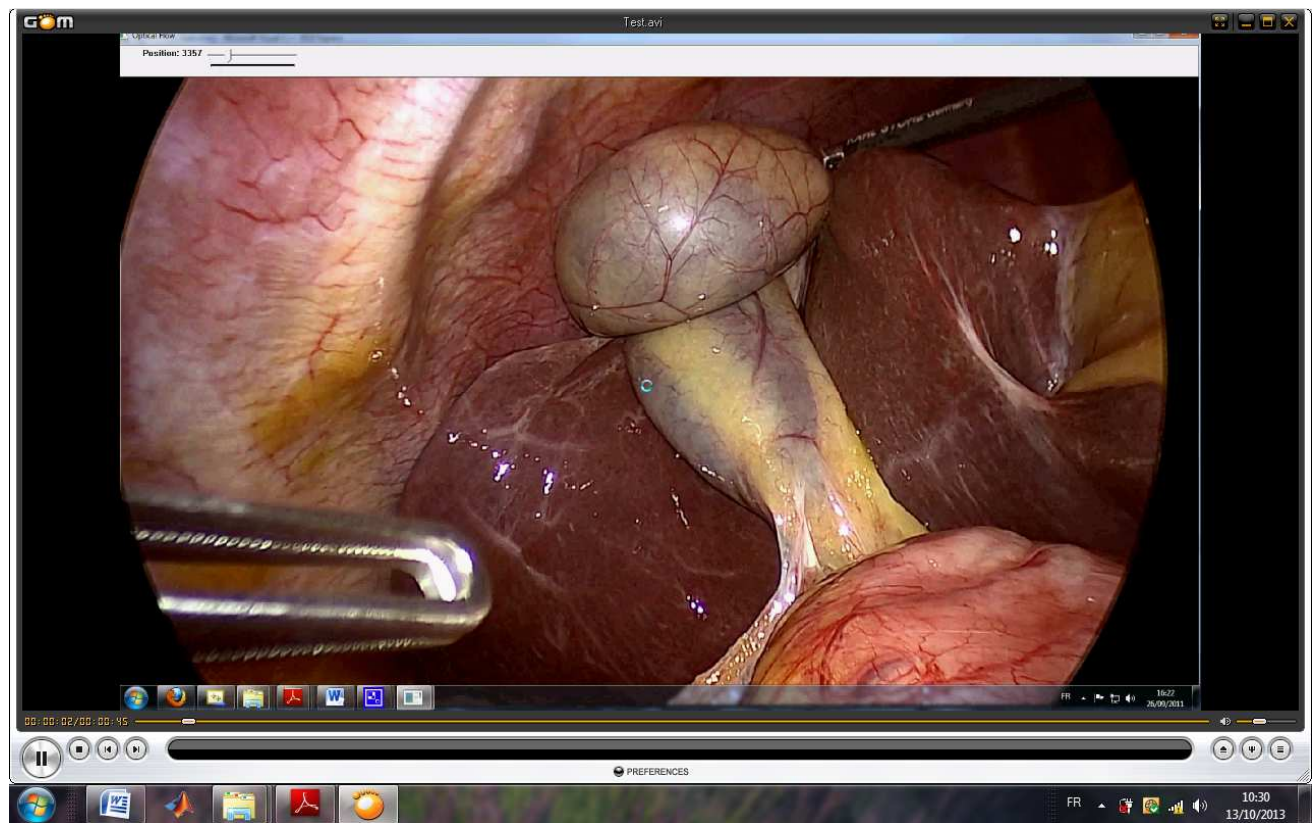
Fig.3. Structure général de l'application.

Perspectives:

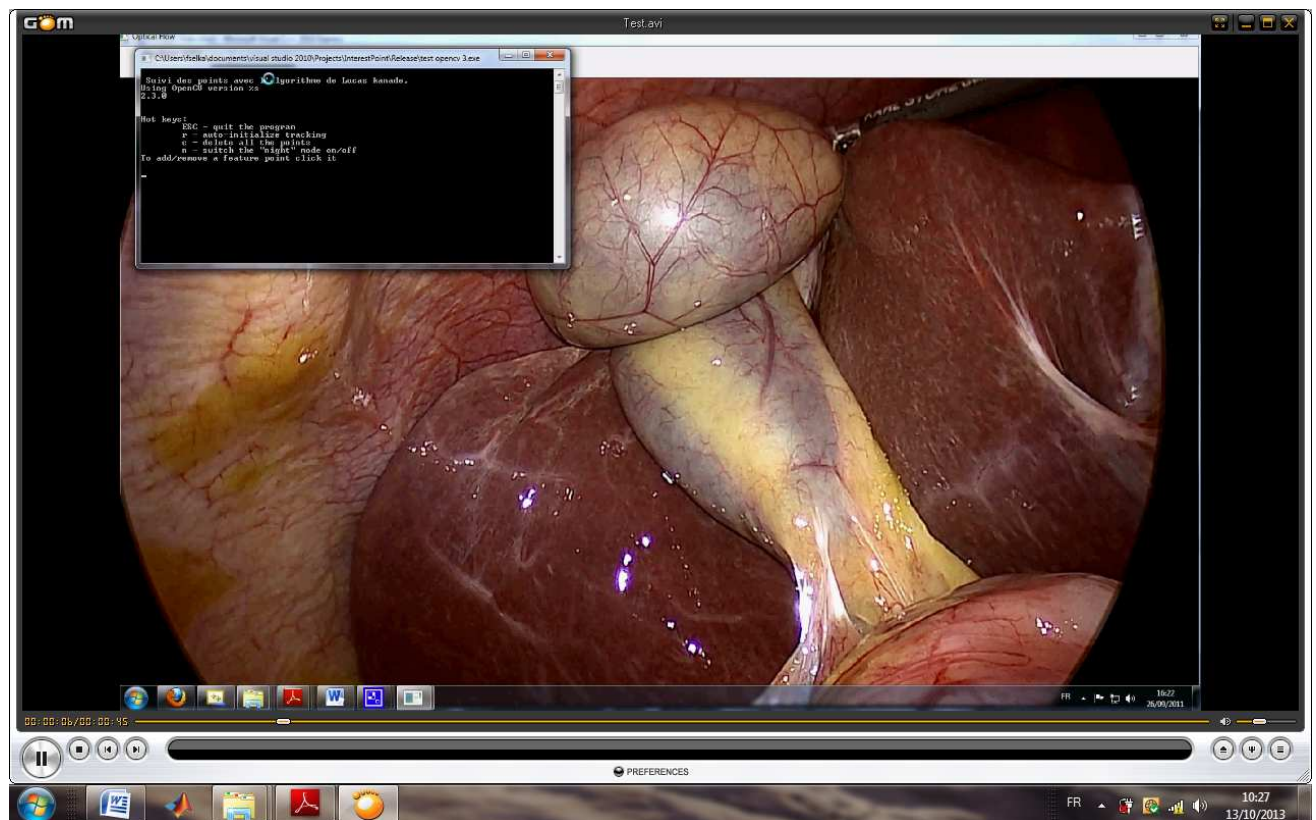
- Externaliser le choix de lecture des fichiers vidéo.
- Externaliser le choix du canal sur lequel effectuer le traitement ainsi que les combinaisons des canaux.
- Externaliser le choix de la ROI.
- Création de boutons pour les paramètres prédéfinis des fonctions d'initialisation et de suivi des points d'intérêts.
- Effectuer un bloc de prétraitement (amélioration du contraste, rehaussement de l'information) dans la chaîne de traitement.

Quelques résultats que nous avons trouvés et logiciel proposé

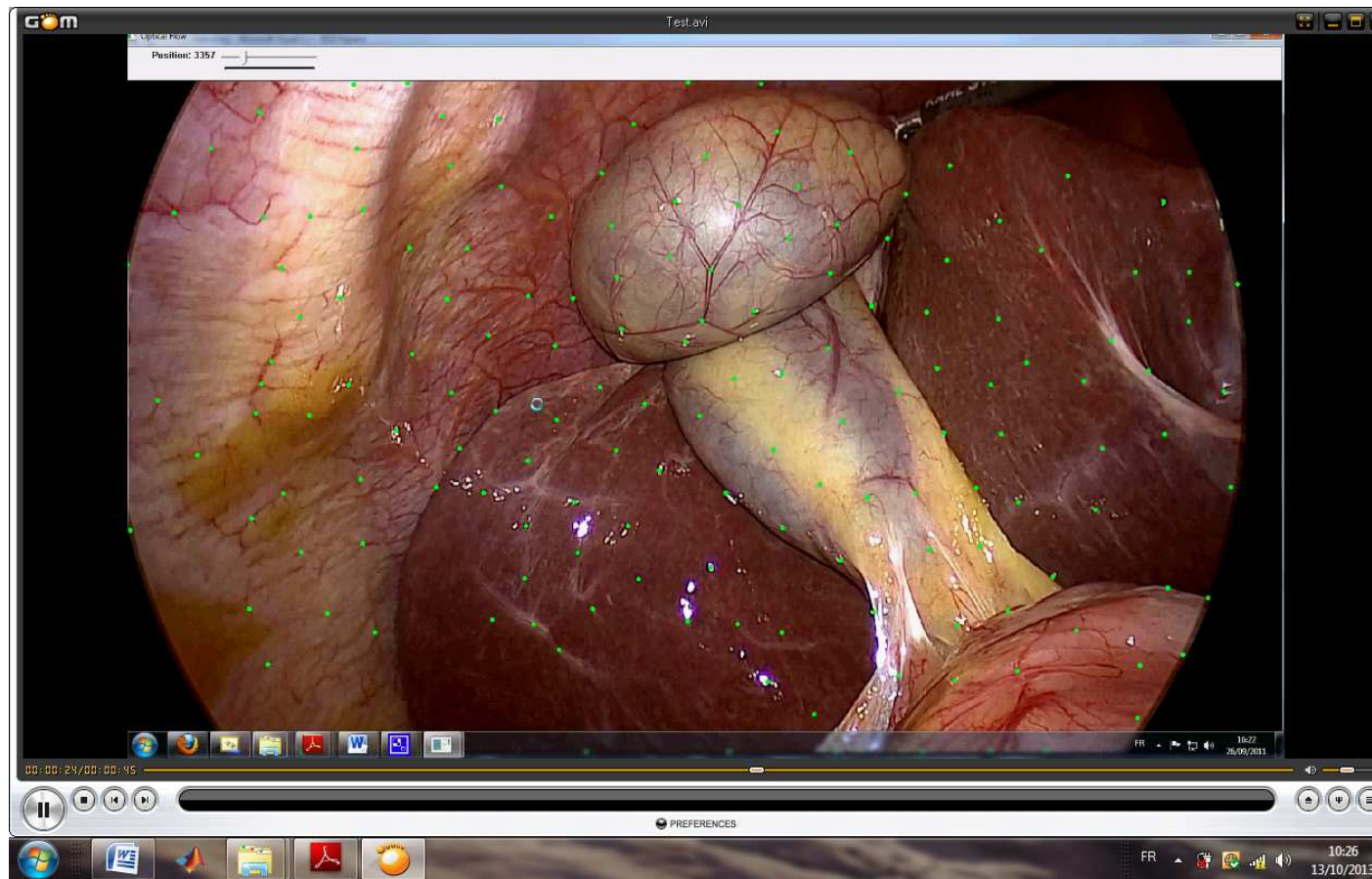
Le logiciel commence à lire des séquences d'images.

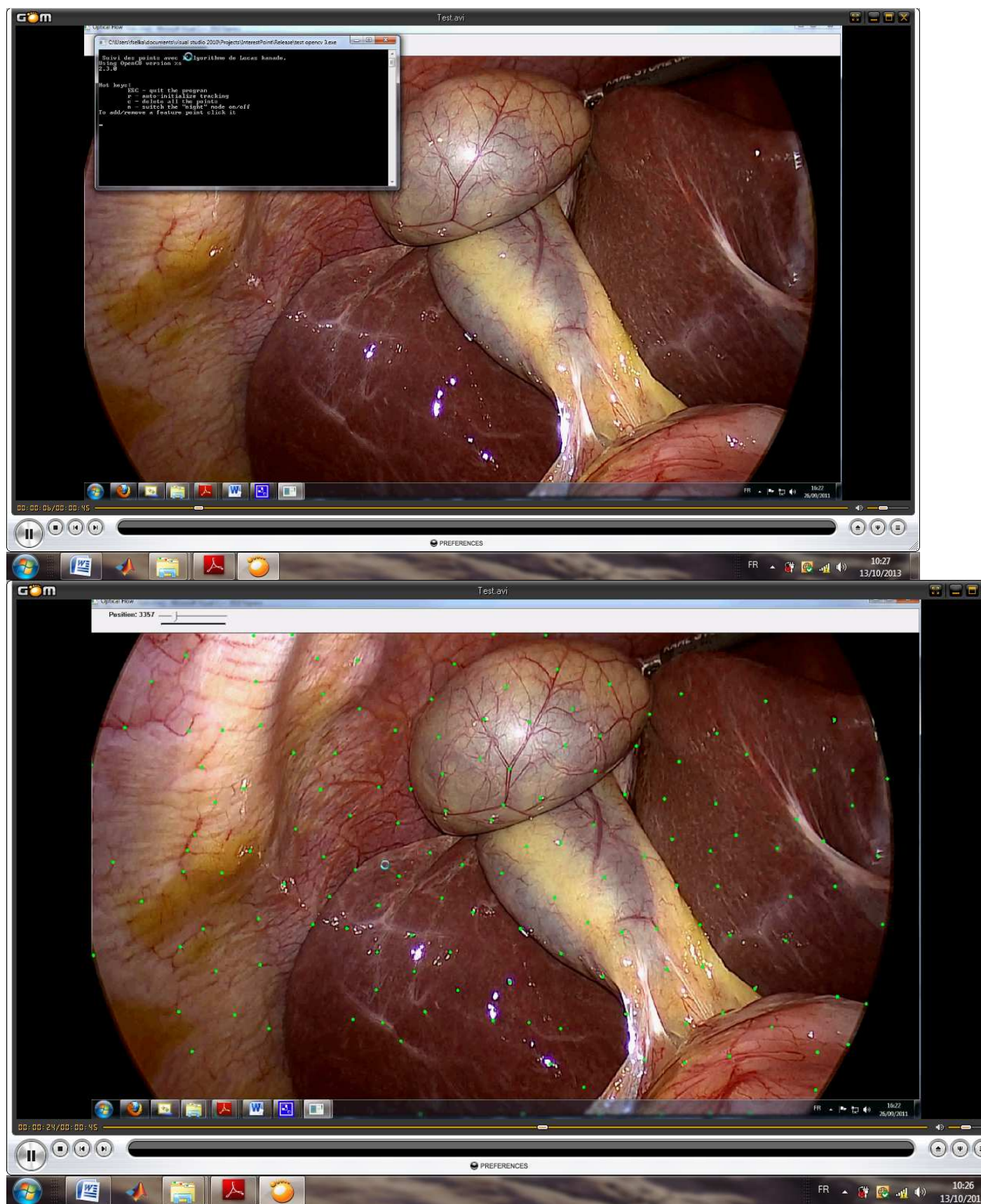


Séquence d'images initiale

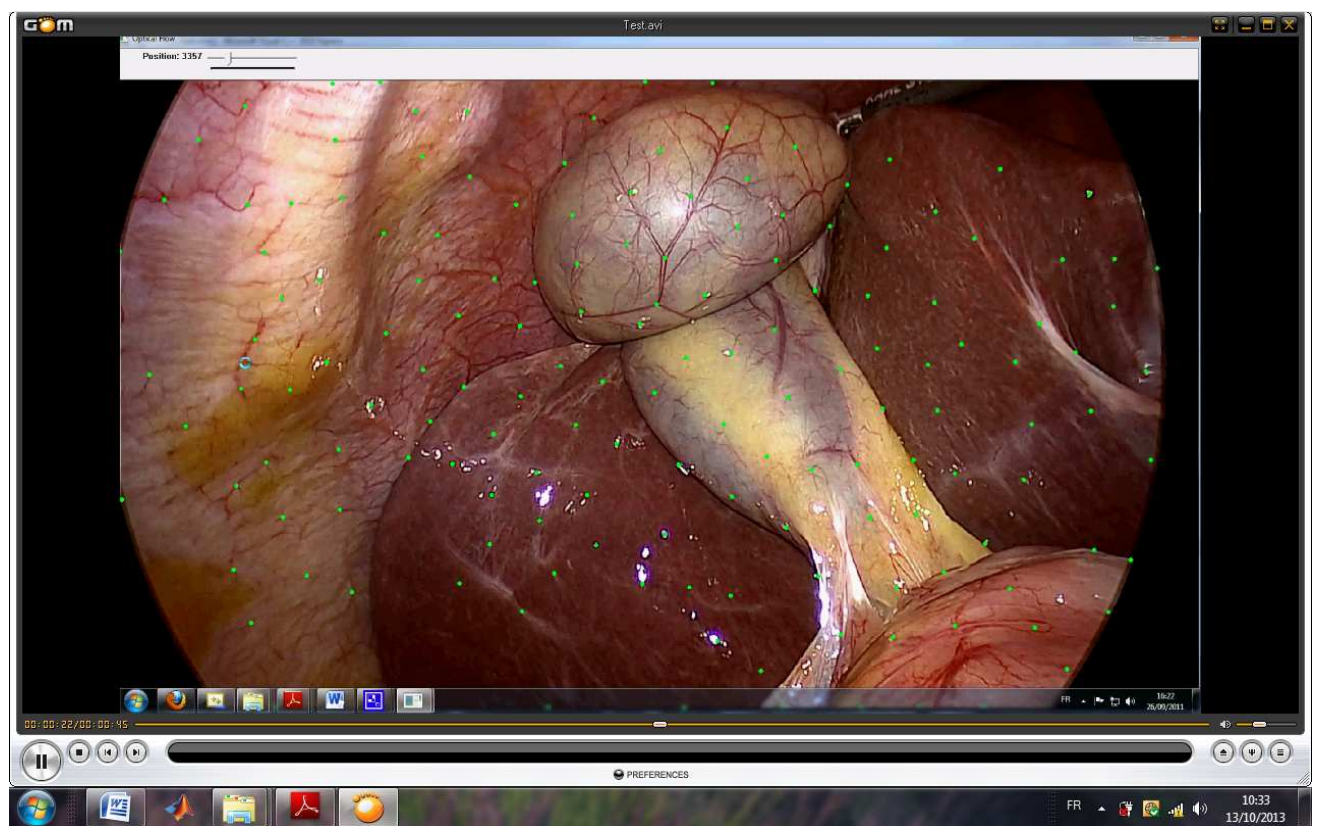


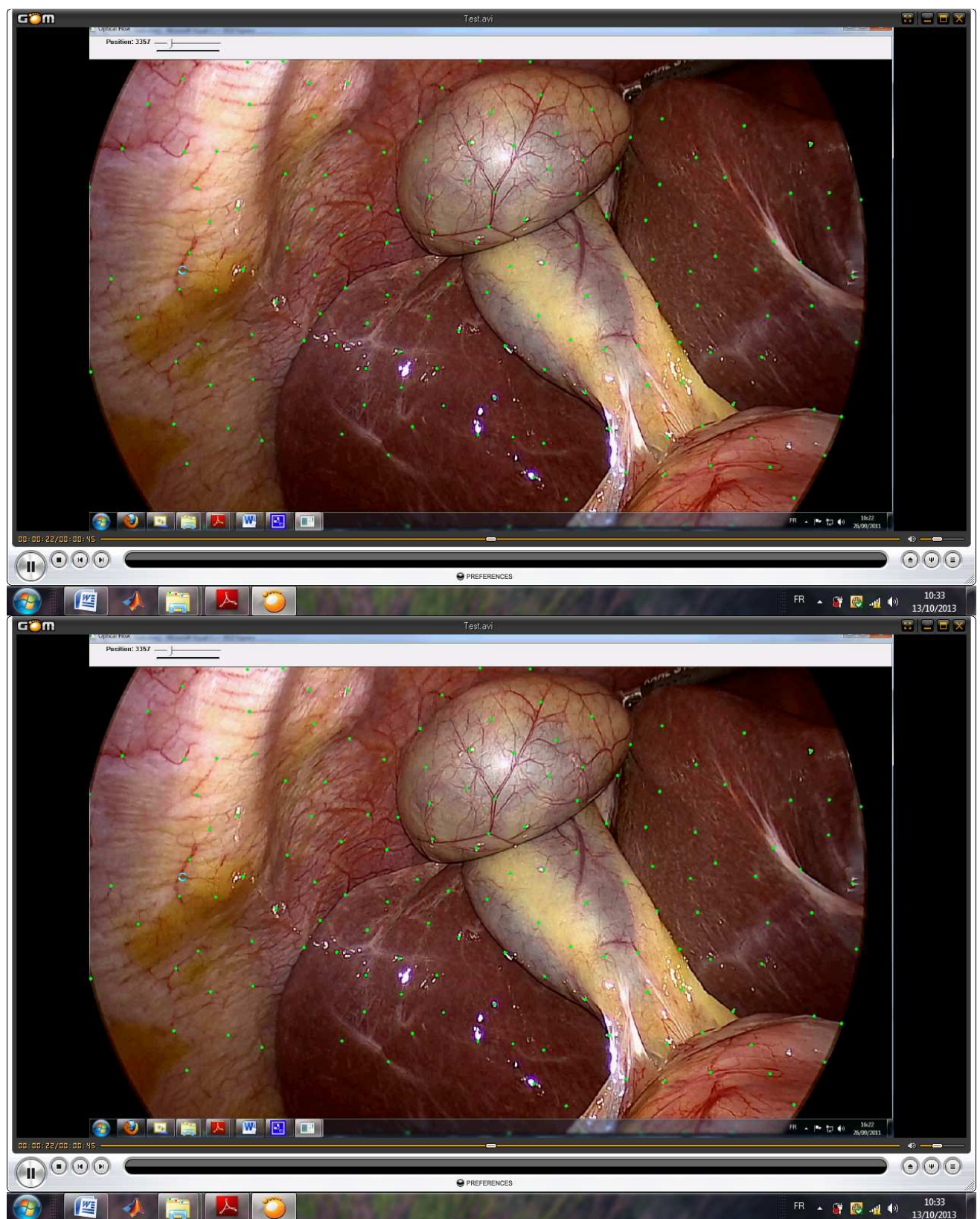
Ensuite le logiciel que nous avons implémenté commence à détecter des points.

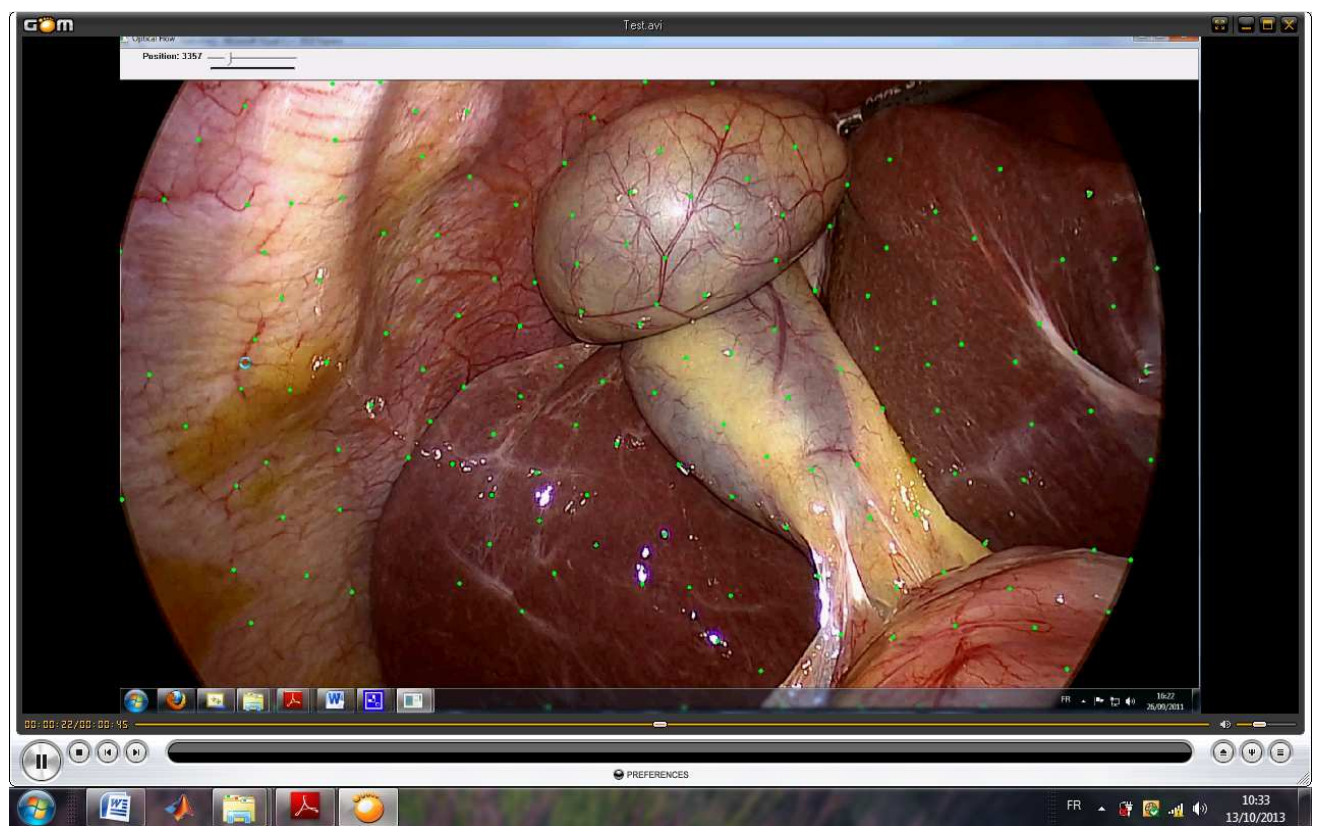




Détecter des points







Détecter des points

Information financière

Aucun financement n'a été fait pas pour réaliser ce travail