

Reconstruction 3D par vision active codée. Application à la vision endoscopique

mémoire présenté pour obtenir le grade de

Docteur de l'Université de Strasbourg
Discipline : Électronique, Électrotechnique, Automatique
(Spécialité : Vision et Robotique)

par

Chadi Albitar

Soutenance publique le 30 Mars 2009

Membres du jury

- Directeur de Thèse* : Pierre GRAEBLING, Professeur à l'Université de Strasbourg
Examineur : Christophe DOIGNON, Maître de Conférences HDR
à l'Université de Strasbourg
Rapporteur Externe : El Mustapha MOUADDIB, Professeur à l'Université de Picardie, Amiens
Rapporteur Externe : Eric MARCHAND, Chargé de recherche HDR à IRISA/INRIA, Rennes
Rapporteur Interne : Jean-Michel DISCHLER, Professeur à l'Université de Strasbourg
Examineur : Didier WOLF, Professeur à l'INPL, Vandœuvre-lès-Nancy

Remerciements

Je tiens tout d'abord à remercier les membres du jury et notamment les rapporteurs qui ont accepté de lire et d'évaluer mon travail de thèse :

- M. El Mustapha MOUADDIB, Professeur à l'Université de Picardie à Amiens,
- M. Eric MARCHAND, Chargé de recherche HDR à IRISA/INRIA à Rennes,
- M. Didier Wolf, professeur à l'INPL à Nancy,
- M. Jean-Michel DISCHLER, Professeur à l'Université de Strasbourg.

Je remercie également Michel de Mathelin, professeur à l'Université de Strasbourg et Directeur de l'équipe AVR pour m'avoir accueilli au sein de son groupe de recherche et avoir mis à ma disposition les moyens nécessaires pour mener cette thèse à son terme. Je remercie aussi l'IRCAD et particulièrement les membres de l'équipe de M. Luc Soler pour leur collaboration. Je remercie l'Institut Supérieur des Sciences Appliquées et de Technologies à Damas (Syrie) qui m'a accordé sa confiance en finançant la préparation de cette thèse.

Mes remerciements s'adressent à Pierre Graebing, professeur à l'Université de Strasbourg et directeur de cette thèse, et à mon encadrant, Christophe Doignon, Maître de Conférences HDR à l'Université de Strasbourg. Je vous remercie pour votre soutien tout au long de la thèse, vos conseils avisés, les relectures très constructives du manuscrit, mais aussi tout simplement pour la qualité de votre encadrement.

Je remercie vivement l'ensemble des permanents et doctorants de l'équipe AVR, pour les moments conviviaux que nous avons passés ensemble. Une dédicace spéciale aux anciens de l'eAVR que j'ai pu côtoyer : Benjamin, Estelle, Adlane, Yunjie, Ali et particulièrement à Cyrille, Ahmed et Wael pour leur amitié et pour tous les moments qu'on a pu partagés.

J'adresse mes remerciements particuliers à tous mes amis, ma troisième famille, pour leur sympathie et la bonne ambiance et leur soutien : Khaled et Rima, Alex et Yana, Alaa et Ruba, Ahmed et Yosra, Ali et Samia, Karima, Mohammed, Abdo, Juliette, Rabih, Benoît, Julien et Céline, Alexander et Natacha, Sébastien et Najat et Eva, Mourad et Aurélie et Emna. Je remercie également mes amis en Syrie : Wassim et Nassim, Eman, Shaka et Dima, Mihyar et Nisreen, Zaffer et Fatat. Un remerciement profond à Sarah-Fati qui a été plus qu'une sœur et qui a toujours répondu présente pour moi et pour ma femme. Je remercie également mes cousins Adnan et sa famille, Abdo et Lama et Sarah pour leur soutien et leur encouragement.

J'adresse ma reconnaissance et témoigne toute mon affection à mes parents, à ma belle famille et à mes frères Chafii, Hadi et Orwa qui m'ont soutenu tout au long de mes études.

Enfin, je dédie ce mémoire à mon épouse Waed qui m'a accompagné tout au long de ma thèse dans les moments joyeux et dans les moments difficiles. Je te remercie Waed pour ton amour, ton réconfort et ton dynamisme. J'ai compris avec toi ce que veut dire la bonne personne. Une dédicace spéciale à mon ange, ma fille Zoya, qui a rempli ma vie de bonheur.

Table des matières

Introduction	1
Chapitre 1 Chirurgie mini-invasive assistée par ordinateur : vision et robotique	5
1.1 Chirurgie mini-invasive (CMI)	5
1.2 Robotique et chirurgie mini-invasive	8
1.3 Vision et chirurgie mini-invasive	11
1.3.1 Vision 3D passive	12
1.3.2 Vision 3D active	15
1.3.3 Comparaison entre les deux principes	19
1.4 Conclusion	21
Chapitre 2 Codage de la lumière structurée : Conception d'un nouveau motif monochrome	23
2.1 Stratégies de codage de la lumière structurée	23
2.1.1 Multiplexage temporel	24
2.1.2 Codage direct	25
2.1.3 Voisinage spatial	26
2.2 Conception d'un nouveau motif	29
2.2.1 Contraintes imposées pour la conception du motif	29
2.2.2 Propriétés du motif	31
2.2.3 Génération du motif	32
2.2.4 Algorithmes de segmentation et de décodage associés	35
2.3 Résultats expérimentaux	37
2.4 Conclusion	39
Chapitre 3 Reconstruction 3D et étalonnage du dispositif de lumière structurée	45
3.1 Reconstruction 3D	45
3.2 Étalonnage du dispositif	47
3.2.1 Étalonnage de la caméra	47

3.2.2	Étalonnage du projecteur	50
3.2.3	Méthode d'étalonnage proposée	55
3.2.4	Évaluation de la méthode d'étalonnage proposée	57
3.3	Évaluation de la reconstruction métrique	63
3.4	Conclusion	64
Chapitre 4 Conception d'un outil endoscopique de vision 3D		67
4.1	Contraintes liées à l'application	67
4.2	Choix de la couleur de projection	68
4.3	Éléments d'optique diffractive	71
4.3.1	Principe	71
4.3.2	Réalisation des éléments d'optique diffractive	73
4.4	Conception d'un outil endoscopique de vision 3D	75
4.5	Conclusion	76
Chapitre 5 Résultats expérimentaux		79
5.1	Résultats expérimentaux : projection du motif proposé	79
5.2	Résultats expérimentaux pour le prototype d'endoscope 3D	85
5.2.1	Source laser et acquisition d'images	85
5.2.2	Étalonnage du système	86
5.2.3	Évaluation de la reconstruction métrique	90
5.2.4	Reconstruction 3D in vivo	93
5.3	Conclusion	94
Conclusion générale		99
Liste des publications		103
Bibliographie		

Introduction

La chirurgie mini-invasive est aujourd'hui une technique à part entière de la médecine moderne. On fait de plus en plus souvent appel à ce type de chirurgie parce qu'elle procure de nombreux avantages par rapport à la chirurgie, dite ouverte, plus classique. Son nom vient du latin : "minimus" qui signifie le plus petit et "invasion" qui désigne l'intervention. Elle est appelée ainsi car le principe est d'introduire des instruments chirurgicaux fins dans le corps d'un patient à travers des petites incisions et de guider les mouvements de ces instruments à l'aide d'un instrument optique appelé "endoscope" qui permet au praticien de voir la scène sur laquelle il intervient. Cette technique a été introduite et popularisée à la fin des années 1980 grâce notamment à l'amélioration des techniques de visualisation endoscopique et d'illumination. Les grands avantages que présente cette technique pour un patient sont la diminution des douleurs post opératoires et la réduction du temps d'hospitalisation. La chirurgie mini-invasive est aujourd'hui fréquemment utilisée dans plusieurs domaines comme la gynécologie, la chirurgie digestive, l'urologie et la chirurgie cardio-vasculaire.

Les systèmes robotisés sont de nos jours présents un peu partout. Mais ils n'ont fait leur apparition dans le domaine médical et chirurgical que très récemment. Ils ont été utilisés particulièrement pour réaliser la télé-chirurgie. En effet, l'utilisation de robots en chirurgie a été initialement proposée par l'armée des États-Unis afin de permettre à un chirurgien se trouvant à distance, d'opérer les soldats sur le champs de bataille. Dans ce cas, le robot n'a pas d'autonomie, son travail est de répéter, d'une manière identique, les gestes effectués à distance par le chirurgien. Grâce aux robots, le chirurgien bénéficie alors d'un grand confort en réalisant des tâches précises et efficaces. Ces apports (confort et sécurité) peuvent être transposés à la chirurgie mini-invasive. En effet, là non plus, le chirurgien n'est pas en contact direct avec la partie à opérer et, comme pour dans la télé-chirurgie, une assistance robotisée lui permettrait d'avoir un geste plus sûr et plus facile à réaliser. Depuis son apparition, la robotique médicale et chirurgicale n'a cessé de progresser et des médecins imaginent déjà le jour où ils pourront préparer une opération à l'avance, dans un environnement virtuel reproduisant exactement le corps du patient en question. Ils pourront alors simuler et enregistrer l'opération parfaite, qui sera ensuite reproduite par un robot sur le patient sous la supervision du médecin. Mais l'idée d'un robot effectuant d'une manière entièrement indépendante une opération pré-programmée n'est pas encore à l'ordre du jour. Quoiqu'il en soit, ces systèmes ont tous besoin d'être associés à des dispositifs de vision, ne serait ce que pour fournir une information visuelle au praticien.

En effet, l'imagerie est le seul moyen fournissant de l'information concernant la topologie des organes à traiter au chirurgien. Des technologies comme la tomographie par émission de positrons (TEP) et l'imagerie par résonance magnétique (IRM) offrent des informations détaillées

et riches du corps d'un patient, mais les médecins se basent encore sur leurs expériences cliniques afin d'atteindre les parties à traiter. On constate effectivement que des systèmes robotisés couplés à des dispositifs d'imagerie (scanner, IRM, échographie, vision 3D, etc) commencent à émerger dans le monde médical. À ce propos, on peut penser au système *Innomotion* proposé par la société allemande *Innomedic* qui fournit un positionnement précis et reproductible d'un instrument à l'intérieur d'un système d'IRM. Il est à l'heure actuelle le seul système robotique commercial compatible avec l'IRM.

La connaissance en trois dimensions de l'environnement dans lequel on évolue est un problème majeur pour de nombreuses applications, notamment médicales, et en particulier dans le cadre des Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO). 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. Par exemple, dans le cadre d'interventions mini-invasives, cela permettrait de restituer visuellement à un chirurgien, la profondeur de l'environnement dans lequel il opère, en s'appuyant sur la réalité augmentée par superposition des indices visuels de profondeur et ainsi lui offrir un plus grand confort et une plus grande précision dans son geste. 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.

C'est dans ce contexte que s'inscrit ce travail de thèse, l'objectif étant de développer un dispositif de reconstruction tridimensionnelle (3D) en temps réel de surfaces intra abdominales. Un tel dispositif pourrait alors avantageusement être mis à profit pour une assistance visuelle ou robotisée. Nous nous sommes alors intéressés à concevoir un outil endoscopique rigide capable de fournir l'information 3D basé sur l'utilisation d'une projection de lumière structurée. Nous avons conçu à cet effet un nouveau motif monochrome auquel un codage robuste a été associé afin de répondre aux contraintes imposées par l'application ainsi que par la réalisation matérielle du dispositif visé. L'objectif final de ce travail étant 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. Ce manuscrit relatant ce travail est composé de cinq chapitres :

Chirurgie mini-invasive assistée par ordinateur : dans ce chapitre, nous présentons l'état actuel de la chirurgie mini-invasive ainsi que des systèmes robotisés et les systèmes de vision qui ont fait leur apparition dans le monde chirurgical. Ensuite, nous nous focalisons sur la vision tridimensionnelle dans le domaine de la chirurgie et nous présentons une analyse de la vision passive et active en endoscopie. Nous finissons cette partie par présenter les objectifs visés par notre travail et tout l'intérêt qu'il peut représenter dans le monde de la chirurgie mini-invasive en particulier.

Codage de la lumière structurée : conception d'un nouveau motif monochrome : nous présentons dans ce chapitre un aperçu générale des différentes stratégies de codage de la lumière structurée afin de choisir la stratégie la plus appropriée pour notre application. Ensuite, nous détaillons les étapes de la conception du motif et nous décrivons également le motif proposé ainsi que les algorithmes de segmentation et de décodage associés. Finalement, les premiers résultats expérimentaux sont présentés dans le but de valider nos choix et de confirmer la robustesse de notre motif.

Reconstruction 3D et étalonnage du dispositif de lumière structurée : dans ce chapitre un aperçu général des méthodes d'étalonnage d'un système de lumière structurée est présenté. Nous présentons enfin le principe de reconstruction que nous avons retenu, puis nous décrivons la méthode d'étalonnage que nous avons mise en œuvre. Nous terminons ce chapitre en exposant des résultats expérimentaux qui permettent de valider la précision de la méthode proposée.

Conception d'un outil endoscopique de vision 3D : la conception de l'outil endoscopique tel que nous l'envisageons fait l'objet de ce chapitre. Dans un premier temps, nous présentons les contraintes très spécifiques liées à notre application : la reconstruction 3D des surfaces abdominales. Nous détaillons ensuite les considérations qui nous ont mené à choisir la longueur d'onde la plus appropriée pour la projection dans un environnement tel que la cavité abdominale. Nous présentons aussi la technologie des éléments d'optique diffractive qui a été retenue pour réaliser une miniaturisation de la projection de notre motif ainsi que les étapes de la réalisation de ces éléments. Nous terminons enfin par la présentation du prototype de l'outil et la réalisation matérielle actuelle.

Résultats expérimentaux : Dans ce chapitre, nous présentons des résultats expérimentaux obtenus avec un système de lumière structurée traditionnel utilisant la projection du motif proposé. Ensuite, le système d'expérimentation proposé dans le quatrième chapitre est étalonné afin d'évaluer la reconstruction 3D. Nous terminons par analyser les reconstructions 3D obtenues avec le prototype actuel, pour des projections sur des objets réels et finalement sur des surfaces d'organes de la cavité abdominale d'un porc.

En conclusion, nous dressons un bilan des travaux réalisés dans le cadre de cette thèse ainsi que les perspectives qu'ouvre ce travail pour des applications nécessitant la connaissance 3D de son environnement proche.

Chapitre 1

Chirurgie mini-invasive assistée par ordinateur : vision et robotique

Depuis quelques années, le monde de la chirurgie connaît de grands changements avec l'apparition de nouvelles techniques chirurgicales maintenant accessibles et utilisables grâce à la présence courante d'ordinateurs et de robots pour accomplir ou assister des gestes chirurgicaux. Parmi ces nouvelles techniques, il en est une dont on parle de plus en plus souvent : la chirurgie mini-invasive. Développée dans les années 80, elle consiste à limiter l'aspect intrusif des opérations classiques en ne nécessitant que de petites incisions sur le corps du patient pour accéder aux organes à traiter par le biais d'instruments adéquats. De nos jours, elle remplace de plus en plus souvent la chirurgie traditionnelle, essentiellement parce qu'elle est beaucoup moins traumatisante pour le patient. Elle présente également de nombreux autres avantages qui seront exposés par la suite. Cependant, sa mise en oeuvre implique des changements radicaux concernant les gestes chirurgicaux. En effet, ce mode opératoire nécessite obligatoirement l'utilisation de systèmes endoscopiques permettant aux chirurgiens de visualiser la zone qu'ils sont en train de traiter. D'autre part, il implique aussi l'utilisation d'instruments d'un type nouveau, parfois plus difficiles à manipuler que les instruments traditionnels, pour lesquels on peut envisager une manipulation assistée par un système robotisé.

Dans la suite de ce chapitre, nous présenterons l'état actuel de la chirurgie mini-invasive et nous précisons comment les systèmes robotisés et les systèmes de vision ont fait leur apparition dans le monde chirurgical avant de nous focaliser sur la vision tridimensionnelle dans le domaine de la chirurgie. Nous finirons cette partie en présentant les objectifs visés par notre travail et tout l'intérêt qu'il peut représenter dans le monde de la chirurgie mini-invasive en particulier, mais aussi dans d'autres domaines, comme la robotique notamment.

1.1 Chirurgie mini-invasive (CMI)

La chirurgie traditionnelle nécessite de pratiquer de larges incisions pour permettre aux chirurgiens d'accéder aisément aux organes et aux tissus à traiter. Cette manière d'opérer présente un certain nombre de désagréments qui ne sont d'ailleurs pas toujours en relation avec la nature même de l'intervention. Ils peuvent aussi avoir des conséquences médicales pour le patient. Parmi tous ces désagréments, on peut citer :

- les dommages causés aux tissus sains et à la paroi abdominale pour accéder à la cible,
- les risques élevés d'infection,
- une hospitalisation pouvant être longue,

- une récupération post-opératoire plus compliquée,
- la présence de cicatrices,
- etc.

Contrairement à ce qui se passe en chirurgie ouverte, la chirurgie mini-invasive permet au chirurgien d'atteindre sa cible à l'aide d'instruments longs et fins que l'on peut introduire dans le corps du patient par des trocarts eux-mêmes mis en place grâce à des petites incisions dont la taille peut aller de 0,3 à 1cm, comme cela est illustré par la Figure 1.1. C'est d'ailleurs la raison pour laquelle les instruments miniaturisés se développent de plus en plus, ils s'adaptent en effet aux avancées dans le domaine de la chirurgie mini-invasive (Figure 1.2).

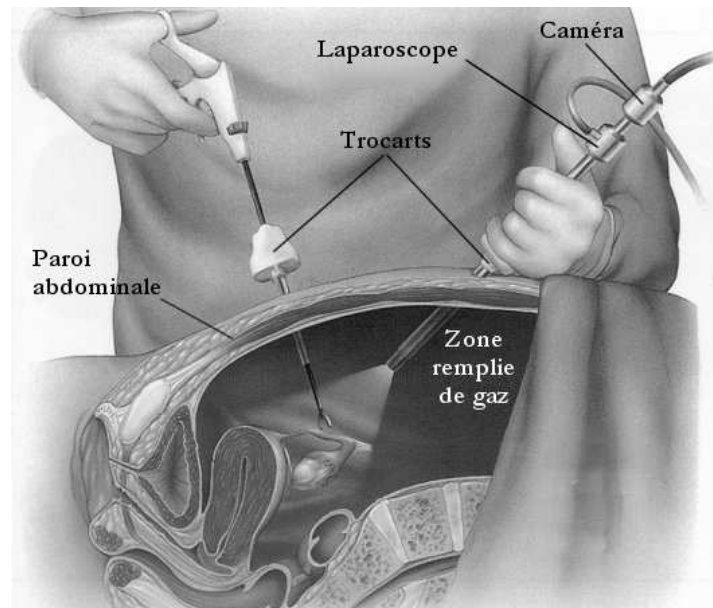


FIGURE 1.1 – Principe de la chirurgie mini-invasive (cas de la laparoscopie).

Il est clair que ce genre de technique limite les traumatismes opératoires, la douleur et les saignements per-opératoires. Moins de tissus sont touchés, les risques d'infection sont donc limités. De plus, comme on réduit les durées d'hospitalisation et de récupération pour le patient, les coûts post-opératoires deviennent aussi de moins en moins importants. Tout cela fait que la CMI est un mode opératoire qui se développe de plus en plus.

Selon l'endroit du corps où on pratique ce genre d'intervention, le nom utilisé varie. On parle d'arthroscopie lorsque l'intervention concerne une articulation, de laparoscopie ou coelioscopie, lorsque l'intervention se situe au niveau de la cavité abdominale, de thoracoscopie au niveau du thorax mais le principe reste le même.

La première conséquence de ce changement de mode opératoire est que le chirurgien ne voit plus directement la zone qu'il est en train d'opérer. Il est donc indispensable d'utiliser un système d'imagerie qui permet au chirurgien et à ses assistants d'assurer l'intervention en contrôlant les gestes chirurgicaux sur l'écran. Un tel système est basé sur l'utilisation d'un instrument particulier : un endoscope (aussi appelé laparoscope quand il s'agit de la cavité abdominale). Cet instrument est relié à une source de lumière froide et à une caméra, qui permet de visualiser l'intérieur de la cavité abdominale, par exemple, sur un écran de visualisation.

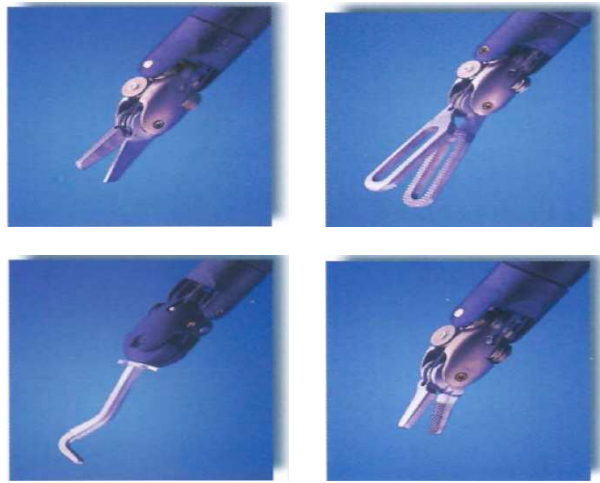


FIGURE 1.2 – Exemples d'instruments utilisés dans la chirurgie mini-invasive.



FIGURE 1.3 – Des endoscopes rigides (à gauche) et un fibroscope (à droite).

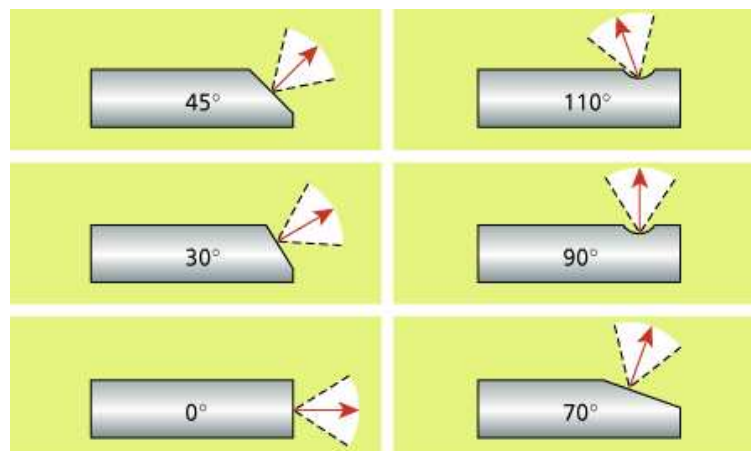


FIGURE 1.4 – Différents angles de vision pour des endoscopes standards.

Il existe des endoscopes rigides qui sont utilisés pour les applications dont l'accès à la zone d'inspection est rectiligne. On trouve également des endoscopes souples (fibroscopes) qui permettent d'exploiter des organes plus complexes à atteindre (Figure 1.3). Leur fonctionnement comporte une succession de lentilles placées à intervalles réguliers afin de transporter l'image jusqu'à l'oculaire. La lumière est quant à elle, transmise par des fibres optiques. L'angle de vision au niveau de l'objectif des endoscopes rigides standards peut varier de 0° à 110° selon les besoins et les différents modèles (Figure 1.4).

De part le fait de l'évolution des technologies : miniaturisation des instruments, évolution des systèmes vidéo, assistance par ordinateur et robotisation, et des nombreux avantages qu'elle présente, on constate que la chirurgie mini-invasive s'impose depuis une quinzaine d'années dans des domaines de plus en plus nombreux.

1.2 Robotique et chirurgie mini-invasive

La robotique est de plus en plus présente dans de nombreux secteurs d'activité. Dans l'industrie, les robots sont utilisés depuis nombreuses années. Ils effectuent des tâches répétitives avec rigueur et ils remplacent les ouvriers dans les tâches pénibles et dangereuses (peinture, soudage, emboutissage, etc.) sur les chaînes de production. Ils sont parfois aussi munis de systèmes de vision, ce qui leur procurent une souplesse d'exécution et des moyens de vérifier la qualité des produits fabriqués. Depuis une quinzaine d'années, les robots ont trouvé leur place dans la salle d'opération sous différentes formes. Les robots utilisés en chirurgie sont souvent classés en deux catégories [Long06] : les robots porte-endoscopes et les robots télé-manipulateurs. Cette classification reflète aussi la différence au niveau de leur complexité, de leur encombrement et de leur coût. La première catégorie concerne les systèmes robotisés qui servent à tenir une caméra et à aider le chirurgien à avoir une main libre durant l'opération. C'est la raison pour laquelle ils sont appelés systèmes porte-endoscopes.

Sur le marché, on peut trouver trois types de robots porte-endoscopes : AESOP (de *Computer Motion Inc.*), EndoAssist (de *ProSurgics*) et LapMan (de *MedSys*) (Figure 1.5). Les trois robots sont sous la forme d'un bras à 6 degrés de liberté pour AESOP (seulement 4 sont actifs), 4 pour EndoAssist et 3 pour LapMan. Le robot AESOP permet au chirurgien d'exécuter des commandes grâce à un logiciel de reconnaissance vocale intégré dans le système informatique de contrôle. Par contre, pour commander le robot EndoAssist, le chirurgien utilise un casque léger équipé d'un émetteur infra-rouge qui est reconnu par un récepteur placé au dessus de l'écran vidéo. Par conséquent, l'endoscope bouge dans la direction dans laquelle le chirurgien oriente son regard. Finalement, le robot LapMan est commandé par une télécommande placée entre les deux gants du chirurgien.

Les robots télé-manipulateurs ou manipulateurs esclaves sont, quant à eux, des robots porte-instruments complets. Jusqu'à présent, deux robots de ce type occupaient le devant de la scène : le robot "Zeus" et le robot "Da Vinci" (de *Intuitive Surgical*) (Figure 1.6). Ces deux robots ont prouvé leur efficacité dans la réalisation de nombreuses interventions difficiles. Le robot Da Vinci a été le premier robot capable de reproduire les mouvements du chirurgien par l'intermédiaire d'un bras et d'un poignet mécaniques. Il a été conçu de manière à incorporer tous les bras esclaves sur une même base qui est volumineuse (Figure 1.6). Le chirurgien suit ses actes grâce à un système d'imagerie 3D basé sur la stéréovision.



FIGURE 1.5 – Les robots (de gauche à droite) : AESOP (*Computer Motion Inc.*), LapMan (*MedSys*) et EndoAssist (*Prosurge*).

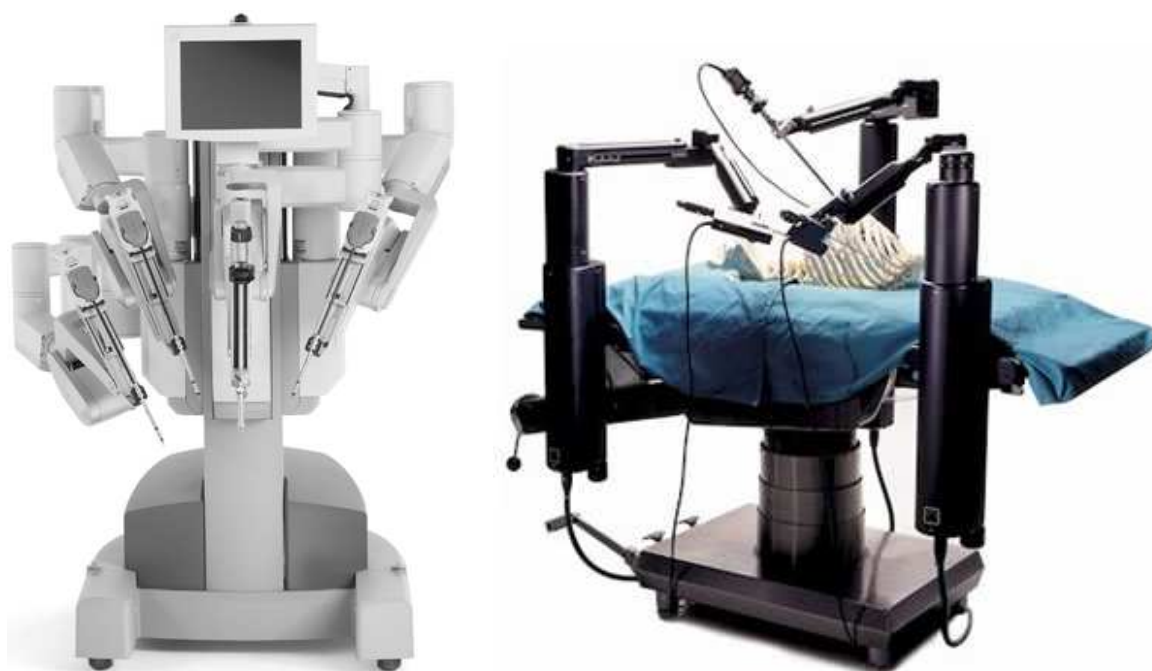


FIGURE 1.6 – Les robots Da Vinci (à gauche) et Zeus (à droite) de "*Intuitive Surgical*".

Ce système est composé d'un endoscope binoculaire couplé avec deux caméras à haute résolution (Figure 1.7). Le chirurgien, assis devant la console de contrôle, opère le patient d'une manière indirecte en commandant, par ses mains et ses doigts, des bras robotiques qui portent les instruments. Par conséquent, le système peut être commandé même à distance afin d'effectuer une téléchirurgie assistée par ordinateur [Marescaux01].



(a)



(b)

FIGURE 1.7 – Le système de commande du robot Da Vinci : l'endoscope stéréo binoculaire (à gauche) et la console de contrôle (à droite).

Le robot Zeus, quant à lui, se compose de trois bras robotisés reliés à une table d'opération. Le système est contrôlé par un ordinateur à l'aide d'une console ergonomique. Un des trois bras est utilisé pour le positionnement de l'endoscope afin de fournir la visualisation du site cible, tandis que les deux autres bras manipulent les instruments chirurgicaux selon les commandes du chirurgien. Le chirurgien a la possibilité de voir la scène en 2-D et 3-D grâce à une caméra binoculaire Storz¹. Le chirurgien peut aussi contrôler les mouvements de l'endoscope par des commandes vocales tandis que les mouvements des instruments sont contrôlés à partir de la console par des poignées qui ressemblent aux instruments chirurgicaux classiques. Ce système a été utilisé lors de la première intervention chirurgicale transatlantique réalisée en 2001 par le Pr. Marescaux qui a opéré de New-York une patiente située dans une salle d'opération à Strasbourg [Marescaux02]. Cette opération portait le nom de "Opération de Lindbergh".

Malgré les grands avantages de ces systèmes, ils restent inaccessibles à la majorité des centres hospitaliers en raison de leur coût qui est très important (de l'ordre de 1 million euros). Nous pouvons aussi remarquer que les systèmes robotiques s'appuient sur la vision en général, et sur la vision 3D en particulier, pour le contrôle et la coordination des instruments au cours de l'opération. Les systèmes de vision sont aussi utilisés afin de donner au chirurgien une impression tridimensionnelle de la scène.

1. Karl Storz Endoscopy, Tuttlingen, Allemagne

1.3 Vision et chirurgie mini-invasive

Nous avons présenté dans la section précédente un aperçu général concernant les avancées significatives dans le domaine de la chirurgie mini-invasive assistée par ordinateur dans lequel on peut constater que la robotique est de plus en plus impliquée pour aider les chirurgiens à effectuer des tâches simples et précises. Cependant, un inconvénient en chirurgie mini-invasive est que le chirurgien n'a pas d'accès visuel direct aux organes comme en chirurgie classique. En effet, l'endoscope est le seul moyen qui permet au chirurgien de voir les organes à opérer et de suivre ses gestes sur un écran en fournissant des images de la cible. Malgré la richesse d'informations qu'on peut trouver dans une telle image, le manque d'information tridimensionnelle limite les possibilités d'utilisation de la robotique dans ce domaine. L'analyse de l'influence des systèmes de visualisation 2D et 3D au cours des interventions laparoscopiques a montré que la dextérité et la précision de positionnement pourraient être améliorées en utilisant un système de vidéo endoscopique 3D au lieu d'un système de vidéo endoscopique standard [Pichler93]. Grâce à l'information 3D, le chirurgien peut se faire une estimation des dimensions de la scène et les instruments peuvent être localisés et repérés l'un par rapport à l'autre d'une part, et par rapport aussi aux surfaces des organes d'autre part.

Fournir une information tridimensionnelle de la surface d'un objet a suscité de nombreux travaux de recherche. La Figure 1.8 présente une classification habituellement retenue des solutions technologiques employées afin d'acquérir la forme 3D et les dimensions d'un objet d'intérêt.

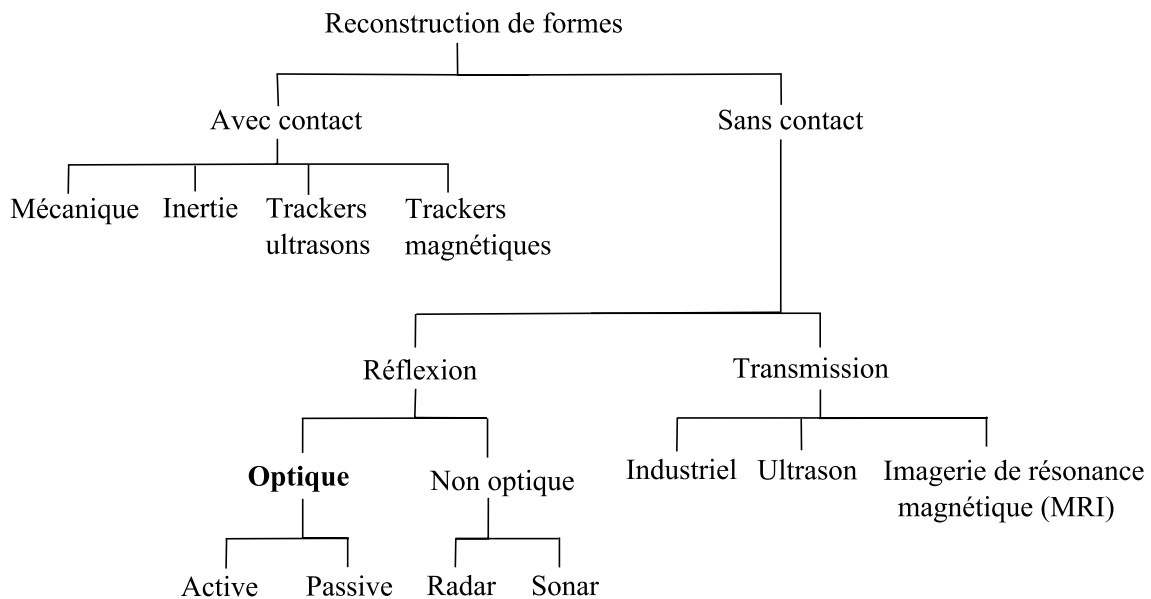


FIGURE 1.8 – Classification des différentes technologies concernant la reconstruction de formes.

Dans ce travail, nous nous intéressons à l'obtention de l'information tridimensionnelle par vision. Généralement, les méthodes qui se basent sur la perception optique sont classées en deux catégories : passive ou active. La vision passive se base sur l'acquisition des images de la scène à des points de vue différents et elle n'implique aucune structuration de la scène alors que la vision active consiste à combiner des caméras avec des sources de lumière structurée.

1.3.1 Vision 3D passive

Le principe de la reconstruction 3D en vision passive est d'imager une scène à reconstruire sous des points de vue différents et peut être avec des éclairages différents. L'adjectif *passive* vient du fait qu'on ne considère que la perception des radiations visibles dans la scène. Les méthodes proposées dans cette catégorie tendent de mettre à profit les indices visuels 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 [Seitz] : la stéréovision, la structure à partir du mouvement (*Structure from Motion (SFM)*), les ombres portées (*Shape from Shading (SFS)*) et la stéréo photométrie.

La stéréovision se base sur l'utilisation d'au moins deux caméras qui imagent la scène sous différents points de vue. Les mêmes primitives qui apparaissent dans toutes les images sont repérées afin de calculer leurs positions dans l'espace 3D par triangulation. De telles méthodes dépendent essentiellement de la calibration des caméras [Salvi02] et de la mise en correspondance d'indices visuels dans les images fournies par les caméras. Cependant, même si les contraintes épipolaires sont prises en compte, la recherche des correspondances reste un problème majeur pour ce type de technique [Armangue03]. Bien que des solutions ont été proposées dans la littérature, (par exemple [Hoff89] et [Stewart96]), ce problème rend la stéréovision limitée à des scènes simples ou bien structurées afin de restreindre le processus de mise en correspondance qui est généralement coûteux en temps de calculs.

La deuxième approche propose de calculer la structure 3D de la scène à l'aide d'une seule caméra en mouvement. Les correspondances sont alors calculées dans la séquence vidéo résultante. Généralement, la reconstruction 3D se base sur la méthode de factorisation [Tomasi92] qui n'a pas besoin d'estimation initiale de la solution. Par contre, une difficulté supplémentaire que soulève cette approche est que le chemin de la caméra doit être estimé car l'étalonnage précis d'une caméra qui se déplace librement est extrêmement difficile.

La technique utilisant les ombres portées est une approche simple pour récupérer la forme d'un objet qui suppose que les surfaces sont lisses. En supposant que la source de lumière directionnelle L est connue, la loi de Lambert (1.1) donne des solutions pour calculer la normale N sur la surface en chaque point en connaissant la réflectance K de la surface.

$$I = KNL \quad (1.1)$$

où I est l'intensité de l'image en chaque point.

Des contraintes supplémentaires sont nécessaires pour rendre le problème bien posé et il est généralement résolu par l'hypothèse de la similitude de réflectance de la surface et de l'orientation des points à proximité. Les difficultés peuvent être surmontées en acquérant plusieurs images avec différents éclairages. C'est d'ailleurs le principe de la dernière approche, la stéréo photométrie [Woodham80]. Dans ce cas, chaque image fournit une contrainte sur le calcul de la normale. Par conséquent, trois images sont suffisantes pour obtenir une solution unique pour l'estimation de la normale à la surface en chaque point.

1.3.1.1 Vision passive endoscopique 3D

Les techniques de la vision passive ont été utilisées dans le but de fournir de l'information 3D au chirurgien dans le cadre d'une opération endoscopique. Afin d'utiliser l'approche de la stéréovision, deux caméras regardant la scène de différents points de vue ont été utilisées pour former ce qu'on appelle un endoscope stéréoscopique. La figure 1.9 présente le schéma optique d'un endoscope binoculaire comme il a été proposé dans [Street00]. Dans cette figure, les éléments (3, 4) sont des systèmes optiques qui transfèrent les images vers les plans image des deux caméras CDD (5, 6). Le groupe des lentilles (7, 8 et 9) servent à ramener l'image du point C au centre de chaque caméra. Cet endoscope a fait l'objet de beaucoup de travaux afin de fournir d'une information stéréo au chirurgien durant l'opération.

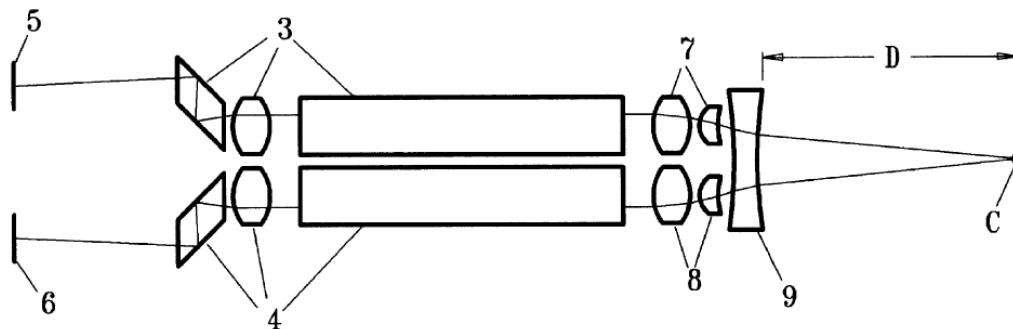


FIGURE 1.9 – Schéma optique de l'endoscope stéréoscopique binoculaire (source : [Street00]).

Dans [Devernay01] par exemple, les auteurs ont proposé une stratégie à cinq étapes pour la localisation des artères coronaires à l'aide de la réalité augmentée et de la stéréovision. La méthode proposée commence par préparer un modèle 3D du cœur à temps variant en utilisant la coronarographie et un CT-scanner ou IRM. Après l'étalonnage de l'endoscope stéréoscopique, la zone à opérer est reconstruite par triangulation et ensuite recalée avec le modèle 3D. Finalement, ces informations sont utilisées en cours d'opération avec les images endoscopiques et la réalité augmentée pour localiser les artères coronaires. En utilisant le même type d'approche, une autre méthode a été proposée dans [Stoyanov04] pour compenser la déformation des tissus durant une opération de chirurgie mini-invasive robotisée. La méthode nécessite l'alignement des images et impose des contraintes géométriques explicites sur le modèle de la déformation afin de fournir une estimation fiable de profondeur. Dans [Lau04], les auteurs ont proposé un algorithme basé image pour le suivi du mouvement du cœur. Cependant, l'inconvénient commun à ces méthodes est qu'elles considèrent un champ de vue fixe et par conséquent, elles ne peuvent pas fournir une information 3D toutes hors du champ de vue réduit.

D'autres méthodes ont proposé d'utiliser le "*Simultaneous localization and mapping (SLAM)*" afin de fournir une reconstruction 3D per-opératoire. Le *SLAM* est une technique habituellement utilisée surtout pour guider les robots et les véhicules autonomes pour mettre en place une cartographie d'un environnement inconnu tout en gardant simultanément la trace de leurs positions actuelles ([Burschka04a], [Choset01]). La reconstruction obtenue est à grande échelle et elle a besoin des indices visuels de la scène pour déterminer le facteur d'échelle.

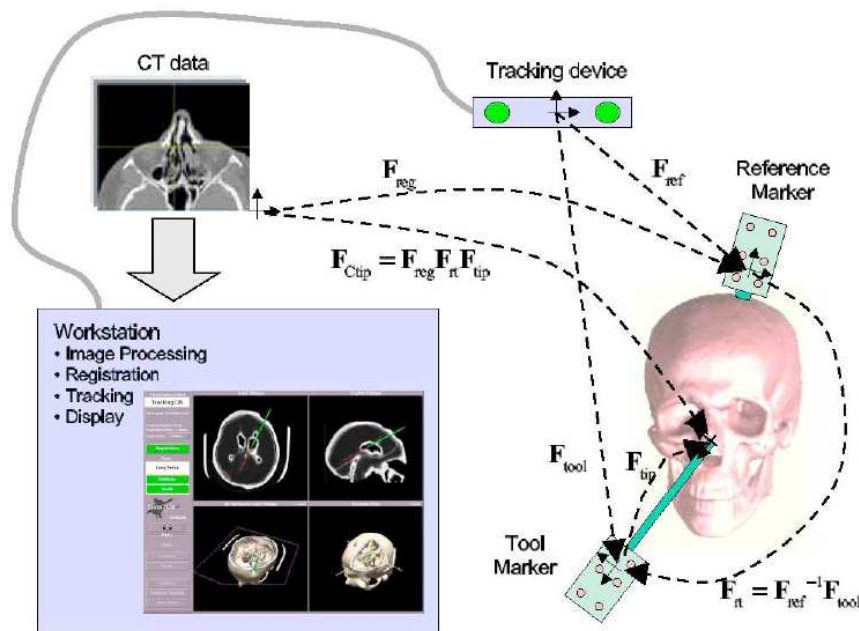


FIGURE 1.10 – Algorithme de reconstruction 3D basé sur la SLAM proposé dans [Burschka04b].

Cette technique a été utilisée dans [Burschka04b] avec un endoscope monoculaire pour obtenir une reconstruction à grande échelle. La reconstruction a été combinée avec les données de CT-scanner de la zone d'intérêt pour déterminer le facteur d'échelle. Les étapes de cet algorithme sont illustrées par la Figure 1.10. La même technique a été utilisée dans [Mountney06] avec un endoscope binoculaire afin d'estimer le mouvement de l'endoscope et construire une carte de la structure anatomique. L'inconvénient principal de ces méthodes est le besoin de la présence d'indices visuels qui se répètent dans les images acquises pour calculer les paramètres de la caméra. Cela n'est pas toujours possible pour les scènes qui sont en mouvement ainsi que pour les tissus lisses qui constituent la plupart des organes abdominaux.

La technique de *Structure from Motion (SFM)* a aussi été utilisée pour résoudre le problème de la reconstruction 3D des surfaces des organes internes. L'idée proposée dans [Hu07] est de reconstruire les surfaces en suivant des indices visuels sur les surfaces des organes au cours du mouvement de la caméra (Figure 1.11) en utilisant le système stéréoscopique du robot *Da Vinci*. Ensuite, les positions 3D de ces indices peuvent être calculées afin de fournir un champ de vue plus large. Les auteurs ont aussi mis à profit un algorithme basé sur les agents évolutionnaires pour récupérer les informations manquantes et assurer la robustesse de la reconstruction.

Il a été aussi proposé dans [Wu07] d'intégrer les contraintes géométriques de l'instrument chirurgical suivi dans la méthode de factorisation traditionnelle pour la reconstruction 3D des structures internes d'un patient à partir des images 2D. Deux types de contraintes ont été considérées : *in situ* et *free-position*. Les contraintes *in situ* comprennent des coordonnées 3D réelles détectées par l'instrument suivi et sont appliquées aux indices détectés. Les contraintes *free-position* sont des nouveaux indices visuels ajoutés par l'utilisateur en cours de l'acquisition afin d'améliorer la forme obtenue. Cependant, la procédure prend moins d'une seconde pour les contraintes *in situ* et moins de deux secondes pour les *free-position*.

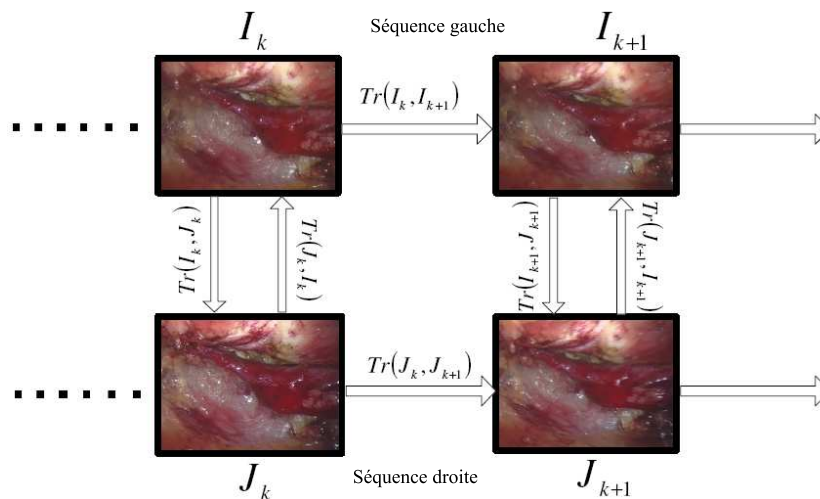


FIGURE 1.11 – Reconstruction 3D des tissus lisses basée sur la *SFM* avec un endoscope bino-culaire du robot *Da Vinci* [Hu07]. Les indices visuels sont repérés et suivis dans les séquences d’images gauches (I) et droites (J).

Malgré les améliorations apportées, le problème de la mise en correspondance pour de tels systèmes n’est pas résolu. Ce problème reste un problème majeur pour l’application de cette stratégie en chirurgie mini-invasive puisque les surfaces des organes abdominaux sont lisses et généralement mouillées, ce qui ne permet pas de repérer facilement les indices. Trouver les correspondances nécessite alors souvent la présence de marqueurs sur les surfaces à reconstruire.

1.3.2 Vision 3D active

Le principe de la vision active est de remplacer une des caméras d’un système stéréoscopique par une source de lumière afin de projeter des indices visuels dans la scène [Klette98]. Par conséquent, les problèmes soulevés en vision passive peuvent être résolus en se basant sur l’extraction des indices générés par les motifs projetés. Cette technique est appelée dans la littérature la projection de la lumière structurée (*Structured Light*). Cependant, le terme *vision active* est aussi utilisé pour désigner les méthodes qui proposent de contrôler le mouvement de la caméra, ainsi qu’éventuellement certains paramètres intrinsèques, afin de réaliser la tâche désirée en utilisant le concept qui a été élaboré dans [Swain93]. Mais nous considérons ici que ce terme correspond à la définition présentée auparavant. Le système à lumière structurée le plus simple consiste en une caméra et une source de lumière qui projette un faisceau laser sur la surface. Dans ce cas, il suffit de détecter la tâche de lumière dans l’image et de calculer sa profondeur par triangulation (Figure 1.12). Par conséquent, un balayage vertical et horizontal de la surface par le faisceau laser est nécessaire pour reconstruire toute la surface. Ceci prend généralement beaucoup de temps.

Pour accélérer le traitement, on remplace souvent le faisceau laser par une source qui projette un plan laser. Dans ce cas, un seul balayage est suffisant pour reconstruire la scène. En effet, cette technique est utilisée dans la plupart des scanners 3D. Cependant, le processus de balayage consomme beaucoup de temps et nécessite un système mécanique précis. Pour limiter ces déplacements, il a aussi été proposé de projeter un ou plusieurs motifs bidimensionnels contenant des points, des franges ou d’autres primitives géométriques (Figure 1.13).

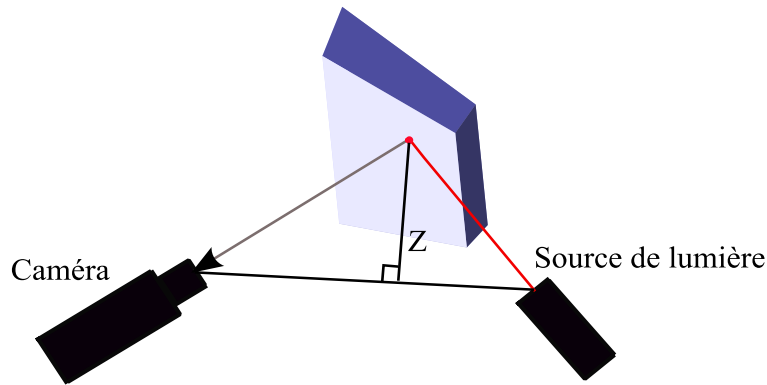


FIGURE 1.12 – Triangulation par vision active : la lumière projetée sur la scène forme temporairement un indice visuel qui est imagé par la caméra puis extrait.

Un codage est alors associé au motif projeté afin d'identifier ses composants dans l'image et d'établir facilement les correspondances. Plusieurs méthodes de codage ont été proposées dans la littérature qui se distinguent selon le nombre de motifs projetés, la nature des motifs et la stratégie de codage. Ces méthodes seront détaillées dans le chapitre 2.

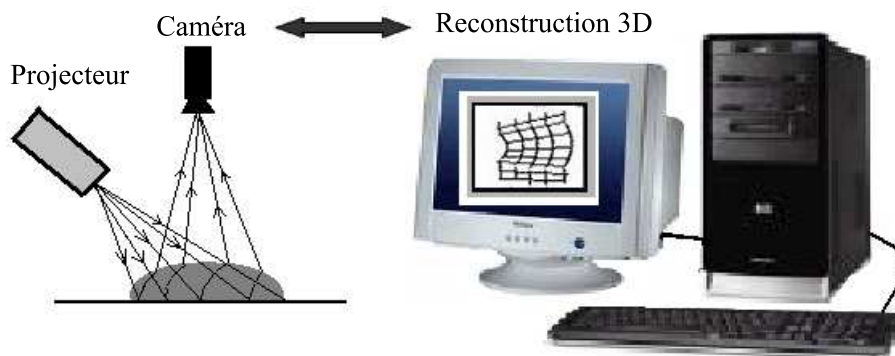


FIGURE 1.13 – Reconstruction 3D à l'aide de la projection d'un motif à lumière structurée.

1.3.2.1 Vision active endoscopique 3D

Différentes méthodes ont déjà été utilisées afin de mettre à profit la vision active pour calculer une information tridimensionnelle en chirurgie mini-invasive. Dans [Krupa03] par exemple, les auteurs ont proposé d'asservir visuellement la position 3D d'un instrument endoscopique par rapport à la surface de l'organe cible à l'aide de la projection d'une lumière structurée. L'instrument est alors repéré dans la scène à l'aide de trois marqueurs à diodes émettrices alignés et fixés à l'extrémité de l'instrument, et à l'aide de la projection de quatre faisceaux laser parallèles (Figure 4.5). L'objectif de l'asservissement 3D est de pouvoir déplacer l'instrument d'une position inconnue ou cachée vers une position désirée, d'une manière sécurisée.

Le premier prototype d'un endoscope 3D basé sur la lumière structurée a été proposé dans [Keller00]. Le principe utilisé consiste à projeter des motifs binaires sur les organes à l'aide d'un

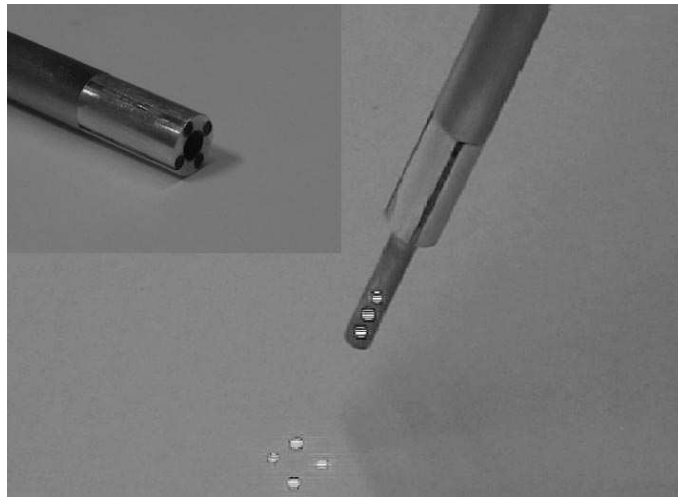


FIGURE 1.14 – Marquage du support de l'instrument endoscopique proposé dans [Krupa03].

projecteur miniaturisé et d'acquérir les images d'une manière très rapide afin d'arriver à reconstruire les surfaces en temps réel (Figure 1.15). Le mini-projecteur était composé d'un écran LCD ferro-reflet, rétro-éclairé avec une lumière froide à travers un tuyau optique flexible. L'acquisition d'images, accélérée de 30 images/seconde à 60 et 180 images/seconde, a été synchronisée avec la projection des motifs. Les auteurs ont considéré six motifs et par conséquent, six images ont été traitées simultanément afin d'extraire les informations 3D de la surface cible. L'objectif était de permettre au chirurgien d'exploiter les images des surfaces d'organes de différents points de vue en temps réel sans avoir besoin de déplacer l'endoscope. Cependant, le système n'est pas révélé fiable et les problèmes de synchronisation ne permettaient pas d'utiliser ce prototype sur des objets réels.

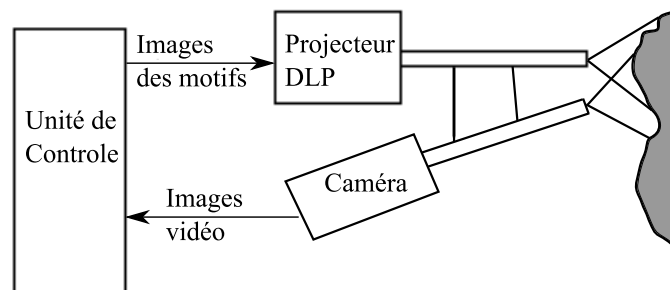


FIGURE 1.15 – Diagramme du fonctionnement du système proposé dans [Keller00].

La projection d'un motif monochrome générée par une source laser a aussi été adoptée par [Hasegawa01]. Afin de coder les pixels de l'image et obtenir une bonne résolution, le motif utilisé dans la technique développée subit un déphasage à chaque acquisition à l'aide d'un système similaire au celui d'un Cubicscope (Figure 1.16-a). Un mot de code de longueur égale au nombre de déphasages est alors associé à chaque pixel. Cependant, pour accélérer le traitement, il faut limiter le nombre de déphasages.

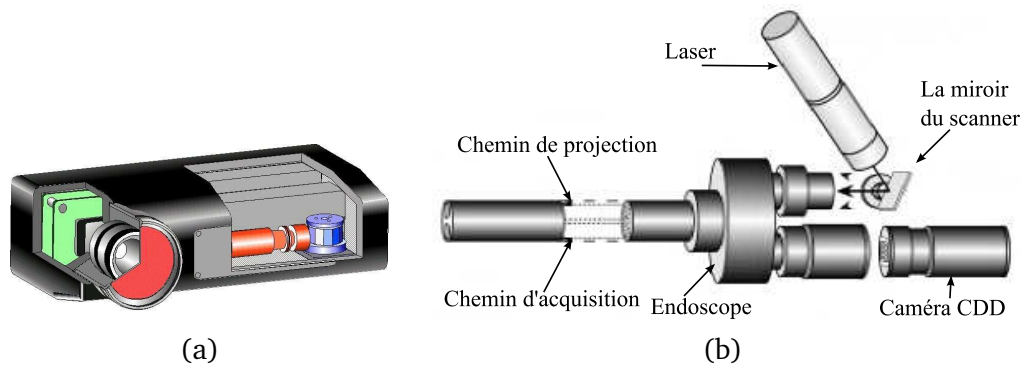


FIGURE 1.16 – (a) Le Cubicscope. (b) Configuration de l'endoscope de vision 3D basée sur la projection de bandes de laser [Hasegawa01].

Pour la réalisation du dispositif, la structure de l'endoscope stéréoscopique a été utilisée, une des deux caméras ayant été remplacée par le système de projection (Figure 1.16-b). Les auteurs ont considéré 8 déphasages pour des images de taille (512×242) pixels. Le système doit être bien synchronisé afin d'acquérir une image de la scène à chaque déphasage. D'après les auteurs, l'erreur de mesure de profondeur obtenue avec le système proposé est de 1% pour une plage de distances de 20 à 40 mm et le temps de traitement est de 0.6 seconde ce qui ne permettait pas de traitement en temps réel.

Le système proposé dans [Chan03] utilise aussi la structure d'un endoscope à deux canaux pour concevoir un système miniaturisé de vision 3D active. Un canal est attaché à une caméra CCD et l'autre est utilisé pour conduire la lumière structurée. Le système projette une grille de points lumineux générés avec une source laser et des optiques diffractives binaires (Figure 1.17). Un algorithme basé sur des contraintes géométriques épipolaires a été utilisé pour effectuer la mise en correspondance des points. L'erreur de la reconstruction 3D résultante était de 2.4% pour les mesures dans l'environnement du laboratoire. Ce système a fourni aussi de bons résultats lors de tests *in-vivo*. Cependant, son utilisation sur des surfaces complexes s'avère relativement problématique.

Un autre système endoscopique basé sur la projection laser afin de visualiser la forme de la région d'intérêt a été développé dans [Hayashibe01]. Ce système est composé d'une caméra rapide (955 fps) associée à un endoscope et à une source laser qui projette une forme lumineuse à l'intérieur du corps à l'aide d'un Galvano scanner et un autre endoscope (Figure 1.18). L'interface du système comprend un écran tactile où les coordonnées 3D du point touché à l'écran sont mesurées par le système afin de guider un robot. Pour une région correspondant au foie, d'une surface de 20 cm^2 , 4000 points ont été obtenus pour un temps d'échantillonnage de 1.2 ms. Le temps total de la mesure est de 5 secondes ce qui s'avère être trop long. Ce système a ensuite été amélioré dans [Hayashibe06] en utilisant la projection d'un plan laser. Le temps de traitement a été particulièrement diminué dans le mode brut où on ne considère que 20 lignes d'échantillonnage sur une surface de taille $8 \times 6.5 \text{ cm}$. Dans ce cas, le système peut traiter 5-6 images par seconde.

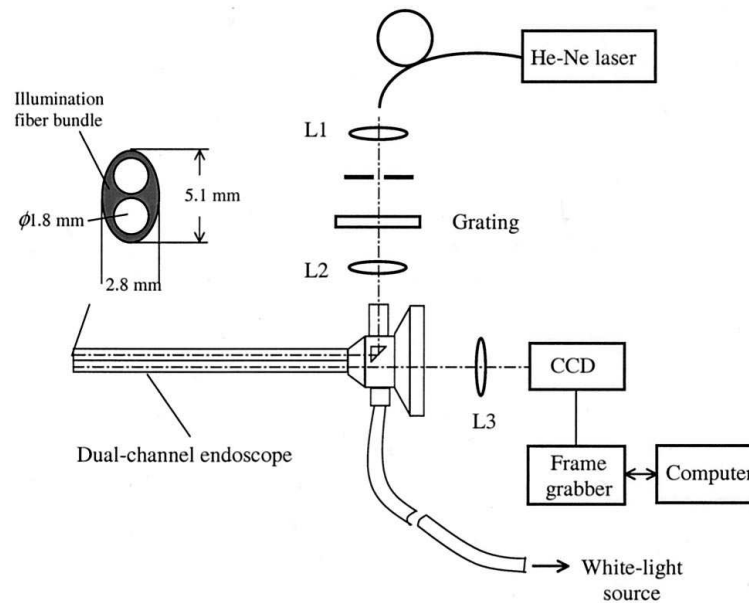


FIGURE 1.17 – La configuration du système miniaturisé proposé dans [Chan03].

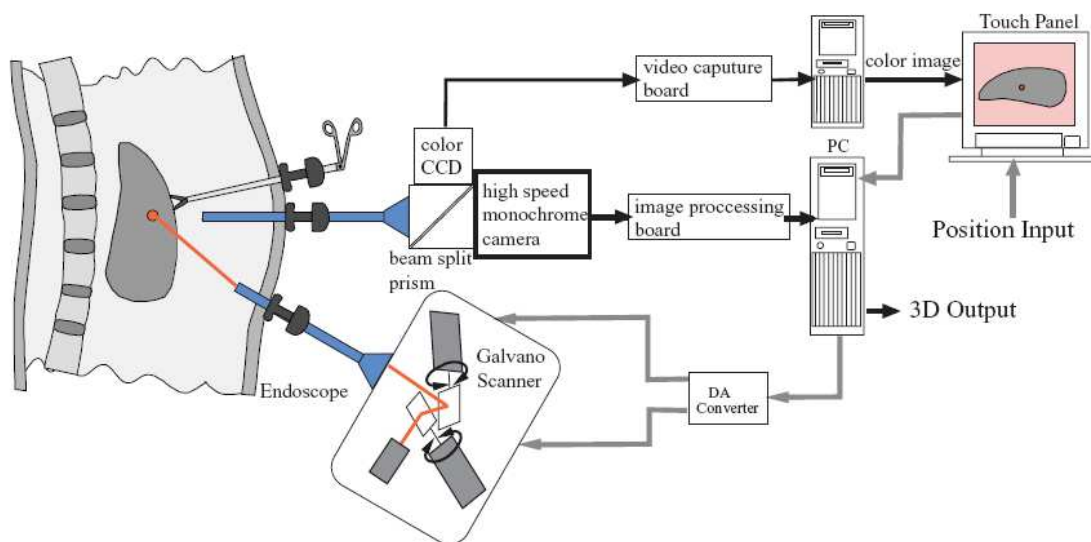


FIGURE 1.18 – Principe du système proposé dans [Hayashibe01].

1.3.3 Comparaison entre les deux principes

Aussi bien la vision passive que la vision active sont largement utilisées dans des applications nécessitant la connaissance de l'information tridimensionnelle de la scène imagée. Mis à part leurs principes, elles se différencient par la résolution de la reconstruction obtenue. La vision passive est connue pour la haute résolution de la reconstruction 3D obtenue puisque le calcul est effectué pour chaque pixel de l'image [El-Hakim95]. De plus, la précision atteinte est beaucoup plus élevée que celle obtenue par la vision active en particulier pour les objets bien définis et pour les contours [El-Hakim95].

La stéréovision est aussi considérée comme une technique fiable puisqu'elle ne contient pas de parties mécaniques mobiles. Malheureusement, l'inconvénient principal de cette technique est qu'elle requiert la présence d'indices visuels dans la scène afin d'établir rapidement des correspondances. Or, dans le cas particulier de la chirurgie laparoscopique, ces indices ne sont pas toujours présents, et quand ils le sont, ils sont souvent difficiles à détecter et à mettre en correspondance. Ce problème peut, dans certains cas, être résolu par l'ajout de marqueurs sur les zones à imager. Malheureusement, ceci rend l'utilisation des techniques basées sur ce principe assez difficile et même parfois impossible lorsque l'ajout de marqueurs est impossible.

Malgré sa précision moindre [El-Hakim95], la vision active peut représenter une alternative aux techniques relevant de la vision passive. En effet, de par sa nature, des indices connus et localisés sont introduits dans la scène imagée sans engendrer de contacts avec les organes. De plus, ces indices étant connus, il est plus facile de les détecter et de les mettre en correspondance. Les systèmes utilisant cette technique doivent reconstruire les surfaces qui traversent leurs champs de détection sans causer de modifications ou de danger aux environnements de travail. Malheureusement, il y a des applications où la projection des motifs lumineux n'est pas souhaitée à cause du bruit qu'elle peut provoquer lors de la perception de la surface cible. Dans ce cas, la solution possible est d'alterner la projection et l'acquisition afin de ne pas perturber l'affichage.

Le tableau 1.1 présente une comparaison rapide entre les deux techniques qui ont fait l'objet de beaucoup de travaux de recherche ([Keller00], [El-Hakim95] et [Fofi04]). Les techniques sont comparées selon plusieurs critères concernant leurs capacités et leurs limitations.

Critère	Vision passive	Vision active
Résolution	Haute résolution	Résolution moindre
Temps de traitement	Traitement en temps réel si les correspondances sont évidentes	Près de temps réel avec la projection d'un seul motif codé (<i>One shot pattern</i>)
Traitement de surfaces lisses	Oui	Oui
Traitement de surfaces courbes	Non	Oui
Traitement de surfaces non texturées	Non	Oui
Extraction d'indices aisée	Dépend de la scène	Automatique
Nécessité d'éclairage extérieur	Dépend de la scène	Elle a son propre éclairage
Robustesse	Dépend de la scène	Dépend du motif
Besoin de parties mécaniques mobiles	Non	Dépend de la méthode

TABLE 1.1 – Comparaison entre la vision active et passive.

Le choix d'une technique pour la reconstruction de formes par vision dépend donc essentiellement de la nature des surfaces se trouvant dans la scène et de la précision désirée. Dans le cas où on s'intéresse aux surfaces abdominales, il faut notamment être en mesure de surmonter le problème de la réflexion sur les surfaces mouillées et la nature lisse de ces surfaces. Souvent, les indices visuels géométriques facilement détectables dans une image sont inexistantes. Puisque nous nous intéressons à la reconstruction 3D des scènes non contrôlées telles que les images acquises à l'intérieur de l'abdomen, nous avons retenu une technique basée sur la vision active pour notre application.

1.4 Conclusion

La chirurgie mini-invasive prend progressivement de l'importance dans le monde chirurgical. Elle présente en effet de nombreux aspects attrayants. Elle réduit notamment de façon significative les traumatismes post-opératoires en limitant les incisions pratiquées sur le corps à des incisions de l'ordre du centimètre. Par contre, elle nécessite l'utilisation d'instruments adaptés et d'un système d'imagerie permettant au praticien de voir et d'atteindre la surface à opérer. Malheureusement, les instruments utilisés réduisent la dextérité du chirurgien et les imperfections des gestes sont amplifiées au bout des instruments par un effet de levier. De plus, les images fournies à travers un endoscope peuvent présenter des déformations géométriques et sont limitées en résolution, en contraste et en couleur. Ceci, ajouté au fait que le chirurgien perd la notion de profondeur, puisqu'il ne voit la scène sur laquelle il intervient qu'au travers d'un écran, rend le travail plus délicat et augmente le risque d'erreur. Ceci ne milite pas en faveur d'un plus grand confort, pour le chirurgien et le patient, ainsi que pour une sécurité accrue et une plus grande précision dans le geste chirurgical. Cependant, disposer en temps réel de l'information tridimensionnelle relative à l'environnement dans lequel opère un praticien permettrait d'envisager de nouvelles perspectives dans le domaine des Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO). Par exemple, dans le cadre d'interventions mini-invasives, cela permettrait de restituer visuellement à un chirurgien, la profondeur de l'environnement dans lequel il opère, en s'appuyant sur la réalité augmentée par superposition des indices visuels de profondeur. Le chirurgien bénéficierait alors d'un plus grand confort et gagnerait en précision d'intervention. De plus, l'utilisation de systèmes robotisés permettrait, d'une part, de contrôler la manipulation des instruments, et d'autre part, elle permettrait d'augmenter l'ergonomie des systèmes opératoires en permettant au chirurgien d'intervenir dans des postures confortables, éventuellement à distance. La connaissance en temps réel de la géométrie 3D de la cavité abdominale pourrait alors 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.

C'est pourquoi, dans le cadre de cette thèse, nous nous intéressons plus particulièrement à la conception de dispositifs capables de fournir des reconstructions 3D des structures internes de l'abdomen, en temps réel et avec une grande précision, dans le but d'intégrer ces données dans des systèmes de réalité augmentée ou des systèmes de robotique médicale. Actuellement, de tels dispositifs n'existent pas. Comme nous l'avons évoqué dans ce chapitre, seuls quelques systèmes permettant d'obtenir une information 3D très ponctuelle pour des applications bien précises commencent à émerger. Vue la comparaison présentée dans la section précédente entre les différentes techniques employées afin de fournir de l'information 3D par vision, et vue la particularité des surfaces abdominales, une solution basée sur la vision active codée avec un seul motif semble le plus appropriée. Bien qu'un peu plus difficile à concevoir, et nécessitant un décodage complexe, une telle technique permet de traiter des scènes dynamiques. La difficulté de la mise en œuvre d'une telle méthode réside alors dans le choix du motif lumineux à projeter, choix qui conditionne le code qui lui est associé ainsi que l'algorithme de décodage. De plus, le principe proposé doit être implémenté au sein d'un système endoscopique appelé à être utilisé dans un environnement particulier ; l'intérieur de la cavité abdominale. Il s'agit donc de prendre en compte l'interaction lumière-organe ainsi que la réalisation d'un système miniaturisé. L'objectif de cette thèse est donc de réaliser un prototype d'endoscope 3D mettant en œuvre une technique de vision 3D active adéquate.

Chapitre 2

Codage de la lumière structurée : Conception d'un nouveau motif monochrome

Dans le cadre de ce travail, nous nous intéressons à la reconstruction 3D de surfaces abstraites en temps-réel. Parmi toutes les techniques habituellement employées en vision pour reconstruire des surfaces, nous nous proposons d'utiliser une solution basée sur une projection de lumière structurée. Nous avons donc été amenés à mettre en œuvre une stratégie de reconstruction en fonction des contraintes imposées par notre application. Ces contraintes nous ont aussi amenés à concevoir un nouveau motif lumineux appelé à être projeté sur les surfaces à reconstruire. Le choix du type de stratégie à utiliser ainsi que les étapes de la conception du motif sont présentées dans ce chapitre. Nous décrivons également, plus en détails, le motif proposé ainsi que les algorithmes de segmentation et de décodage associés. Finalement, nous présentons les premiers résultats expérimentaux qui valident nos choix et qui confirment la robustesse de notre motif.

2.1 Stratégies de codage de la lumière structurée

Nous avons vu que dans le cadre d'une projection de lumière structurée, on remplace une des caméras d'un système stéréoscopique classique par un projecteur de motifs lumineux. Par conséquent, le problème classique de la mise en correspondance peut être résolu en matérialisant sur la surface à reconstruire des indices visuels qui permettent de la reconstruire. Pour ce faire, un code est associé à chaque élément du motif de manière à pouvoir l'identifier dans l'image. Ainsi, pour déterminer la géométrie 3D de la scène imagée, il suffit de décoder les éléments apparaissant dans l'image pour trouver leurs correspondants dans le motif et calculer la localisation 3D des points imagés par triangulation. Cependant, aussi bien la conception du motif que la stratégie de codage associée est guidée par l'obtention rapide et robuste des appariements.

Les méthodes utilisant la lumière structurée sont fonction de trois critères. Le premier critère est le nombre de motifs projetés. Dans ce cas, on peut considérer deux cas : les méthodes qui adoptent la projection de plusieurs motifs différents et celles qui ne considèrent qu'un motif unique (aussi appelé "*One shot*"). Les méthodes de la première catégorie sont généralement dédiées aux applications pour lesquelles on cherche à obtenir une haute résolution et une grande

précision dans les scènes statiques. Par contre, les méthodes de la deuxième catégorie s'intéressent plutôt au traitement de scènes en mouvement. Le deuxième critère est la nature du motif qui peut être binaire, en niveaux de gris ou en couleur. Généralement, le choix entre ces trois types d'encodage est essentiellement lié à la nature de la scène et à l'interaction surface-lumière. Le dernier critère est la stratégie de codage qui peut être périodique ou absolue. Ces critères décrivent la complexité de la méthode et ses capacités ainsi que ses limitations. Ils déterminent si la méthode peut traiter des scènes dynamiques ou si elle est applicable seulement aux scènes statiques. Enfin, la plupart des propriétés de chaque méthode est liée directement à la stratégie de codage du motif. C'est pourquoi nous nous focalisons dans la suite sur la classification et l'évaluation des stratégies de codage de la lumière structurée présentées dans [Salvi04] dans laquelle les méthodes sont classées en trois catégories : les méthodes mettant en œuvre un multiplexage temporel, celles utilisant un codage direct et celles adoptant le voisinage spatial.

2.1.1 Multiplexage temporel

Ces méthodes utilisent des projections successives de motifs différents. Chaque point de l'image est associé à un code absolu qui dépend de l'illumination du point pour chaque projection de motif. La première méthode basée sur ce principe a été proposée dans [Posdamer82] où n motifs binaires ont été utilisés afin d'encoder 2^m franges comme cela est illustré sur la Figure 2.1. La résolution dépend essentiellement du nombre de motifs considérés, cependant il n'est pas recommandé d'atteindre la résolution maximale du projecteur car la caméra n'est pas toujours en mesure de détecter les motifs lumineux qui peuvent être très fins. Pour déterminer le type des primitives à détecter et à décoder dans l'image, deux approches sont possibles : calculer le centre de chaque frange ou bien détecter leurs contours. La deuxième approche donne généralement de meilleurs résultats que la première [Posdamer82].

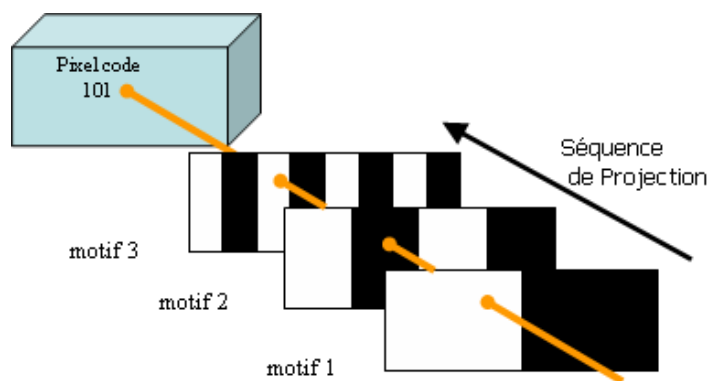


FIGURE 2.1 – Principe du multiplexage temporel.

Afin de proposer un codage robuste, il a aussi été proposé d'utiliser un code binaire accompagné d'un code correcteur basé sur la distance de *Hamming* [Minou81]. Dans cette méthode, 25 motifs ont été nécessaires pour réaliser un codage avec des mots de code de taille 5 et un code correcteur de taille 9. D'autres méthodes ont aussi proposé d'assurer la robustesse en augmentant le nombre de symboles du codage en considérant des motifs codés en niveaux de gris au lieu du codage binaire [Inokuchi84, Rocchini01].

D'autres méthodes permettent d'augmenter l'alphabet de codage en considérant un codage basé sur la couleur au lieu des niveaux de gris. Dans ce cas, chaque niveau de gris est associé à une couleur *RGB* parmi n couleurs considérées. Par conséquent, on peut encoder n^m franges avec m motifs au lieu d'encoder 2^m franges seulement avec le codage binaire [Caspi98]. Cependant, localiser précisément la projection des franges dans l'image a toujours été un problème pour la plupart de ces méthodes, particulièrement dans le cas où la surface est composée de régions ayant des réflectances différentes [Valkenburg98]. Récemment, une étude a proposé de résoudre ce problème en encodant les franges principales des motifs en bleu et rouge au lieu du noir et blanc classique, et à y ajouter des franges vertes de largeur d'un pixel entre ces franges [Rocchini01]. Par conséquent, les transitions entre les bandes principales sont détectées en localisant d'abord les franges vertes. L'inconvénient principal de ce type de méthodes est le grand nombre de motifs nécessaires pour atteindre une bonne résolution, car généralement, il est souhaitable de traiter le moins d'images possible afin d'accélérer le traitement.

Une autre possibilité est de mettre à profit le déphasage. En effet, le principe du déphasage est de projeter un motif d'intensité périodique et d'effectuer plusieurs acquisitions en déplaçant ce motif. Cette méthode exploite, quand elle est appliquée seule, une haute résolution spatiale mais malheureusement, la périodicité des motifs introduit une ambiguïté dans la détermination de la période du signal dans les images. C'est pourquoi on peut intégrer le codage multi-niveaux conjointement avec la méthode de déphasage pour résoudre le problème d'ambiguïté [Gühring01]. Finalement, des méthodes hybrides ont adopté la projection de plusieurs motifs en prenant en compte le voisinage spatial dans l'étape de décodage comme il a été proposé dans [Sato96] et [Hall Holt01]. Dans ce cas, le code de chaque pixel dépend à la fois de son éclairage et de son voisinage. L'aspect le plus intéressant de ces travaux est que cette méthode, contrairement aux autres méthodes de multiplexage temporel, peut traiter des scènes avec des mouvements lents. Cette capacité est due au suivi des contours des franges appliqué en parallèle avec la projection de motifs.

Les méthodes adoptant le multiplexage temporel sont généralement faciles à implémenter et permettent d'atteindre de très bonnes résolutions et de très bonnes précisions en termes de reconstruction. De plus, la détection des primitives projetées est facile, spécialement pour les méthodes qui utilisent un codage binaire. Cependant, elles n'autorisent pas le traitement des objets en mouvement puisqu'elles nécessitent la projection de plusieurs motifs et cela représente une limitation principale pour cette stratégie de codage. D'autre part, elles emploient forcément un dispositif mécanique pour réaliser les projections multiples et cela peut présenter un inconvénient majeur dans certains cas.

2.1.2 Codage direct

Les méthodes qui adoptent la stratégie de codage direct considèrent un motif unique pour lequel chaque élément contient lui-même le code qui lui est associé. On distingue deux types de codage : celui qui utilise des franges de niveaux de gris [Carrhill85] où chaque élément est codé par son intensité, et celui qui considère des franges colorées où l'élément est codé par sa couleur ([Tajima90] et [Sato99]). Théoriquement, on peut atteindre une bonne résolution avec de telles méthodes puisque chaque point possède un code bien distinct en considérant un large spectre de couleurs (Figure 2.2-a) ou de niveaux de gris (Figure 2.2-b).

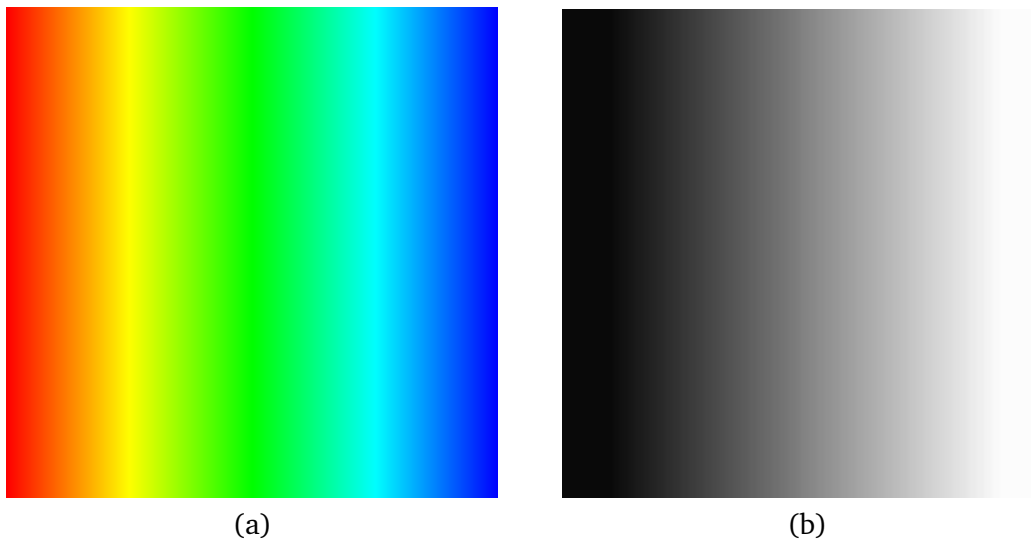


FIGURE 2.2 – (a) Codage direct avec des couleurs. (b) Codage direct avec des niveaux de gris.

Malheureusement, ces techniques s'avèrent être très délicates à utiliser. En effet, les variations de luminance peuvent causer des perturbations sur les mesures d'intensité. De plus, les scènes très colorées ainsi que la sensibilité des capteurs peuvent aussi produire du bruit sur les mesures de la saturation qui servent à distinguer entre les couleurs. Ces problèmes posent des grandes difficultés au niveau de décodage et limitent l'utilisation de ce type de codage.

Généralement, ces méthodes peuvent être appliquées à des scènes en mouvement puisqu'elles ne considèrent que la projection d'un motif unique, mais malheureusement, compte tenu de leur grande sensibilité, elles sont limitées à des applications bien spécifiques.

2.1.3 Voisinage spatial

Contrairement aux méthodes relevant du multiplexage temporel, les méthodes de cette catégorie s'intéressent, tout comme les méthodes basées sur le codage direct, à intégrer tout le code dans un motif unique (*One Shot Pattern*). Mais contrairement aux méthodes précédentes, le code de chaque élément du motif est ici fonction de sa valeur ainsi que de celle de ses voisins. L'avantage de cette stratégie est de n'avoir à faire qu'avec une seule projection. Par conséquent, ces méthodes peuvent être appliquées aux scènes en mouvement, mais compte tenu du fait que le codage est intégré dans les éléments qui composent le motif et non dans les images de ces éléments, leur résolution est limitée. Ces méthodes peuvent être classées en trois groupes. Le premier groupe de méthodes ne considère aucun fondement mathématique pour le codage donc cela peut donner lieu à des ambiguïtés dues à la présence des régions identiques dans le motif (voir par exemple [Graebing02], [Maruyama93], [Durdle98], [Ito95] et [Chen97]). De plus, les motifs proposés sont souvent assez complexes pour assurer une plus grande robustesse de traitement (Figure 2.3).

Les méthodes du deuxième groupe utilisent les séquences de *De Bruijn* pour définir des motifs optimaux (par exemple : [Hugli89], [Monks93], [Salvi98], [Petriu00] et [Zhang02]). Une séquence de *De Bruijn* d'ordre n avec un alphabet de k symboles est, par définition, un mot de code cyclique de longueur k^n dans lequel chaque mot de code de longueur n apparaît exactement une

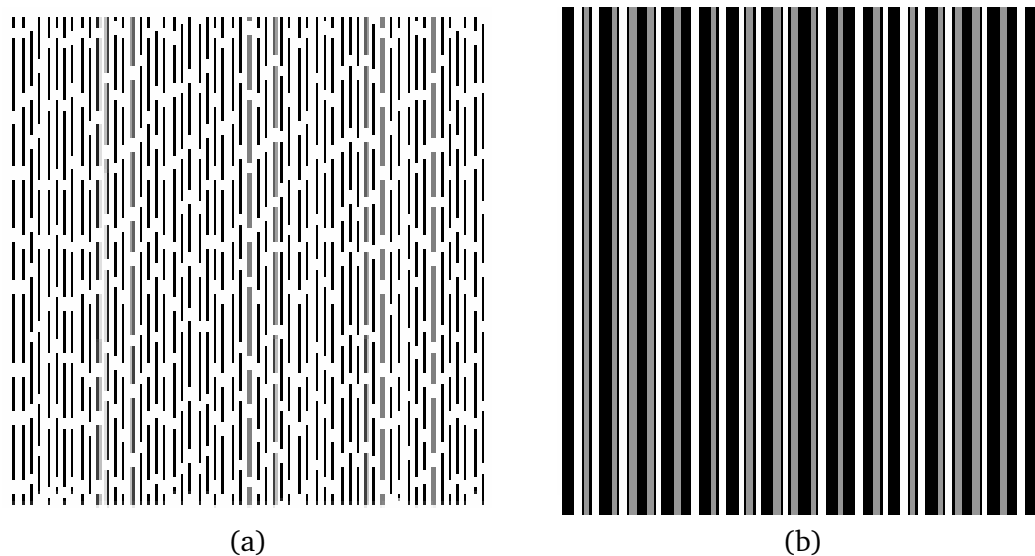


FIGURE 2.3 – Motifs proposés dans (a) [Maruyama93] et dans (b) [Durdle98].

seule fois (par exemple : en considérant un codage binaire, une séquence possible de *De Bruijn* d'ordre 3 est 00101110. Elle est de longueur 8 (2^3) et chaque mot de code de taille 3 est unique. Donc, la position de chaque élément dans la séquence est définie d'une manière absolue par le mot de code de taille n attribué à cet élément. Par conséquent, ce schéma de codage représente une solution pour lever l'ambiguïté évoquée précédemment.

Pour encoder les éléments d'un tel motif, on peut envisager deux solutions. La première considère un alphabet de couleurs où le motif est composé de franges colorées (Figure 2.4-a). La deuxième considère des primitives géométriques pour générer des motifs monochromes [Vuylsteke90]. Cependant, on constate que les motifs monochromes proposés peuvent être complexes (Figure 2.4-b) et que par conséquent, les algorithmes associés ne permettent pas de traitements rapides. Il n'en demeure pas moins que même si les ambiguïtés sont levées, la présence d'occultations, d'ombres ou de fortes variations de courbures des surfaces observées peut conduire à un manque de robustesse de ces techniques qui sont de ce fait adaptées essentiellement à des surfaces lisses.

Les méthodes du troisième groupe adoptent, quant à elles, la théorie des matrices parfaites (*Perfect Maps*). Une matrice parfaite $M(n, m, r, v, k)$ est une matrice de dimensions $n \times m$ dans laquelle chaque élément prend sa valeur dans un alphabet de k symboles et dans laquelle chaque fenêtre de dimensions $r \times v$ apparaît une et une seule fois dans M . Autrement dit si M a une propriété de fenêtre de $r \times v$. De plus, si M ne contient pas la fenêtre ne contenant que des valeurs nulles, M est appelée matrice pseudo-aléatoire ou *M-array* [MacWilliams76].

La robustesse des techniques utilisant les matrices parfaites vient de leur aptitude à décoder les parties visibles grâce aux propriétés des *M-arrays* qui permettent d'affirmer que la position de chaque élément dans la matrice est définie par la fenêtre à laquelle il appartient et d'une manière absolue. Par conséquent, les problèmes dus aux occultations et aux ombres que l'on rencontre avec les méthodes du deuxième groupe peuvent être surmontés.

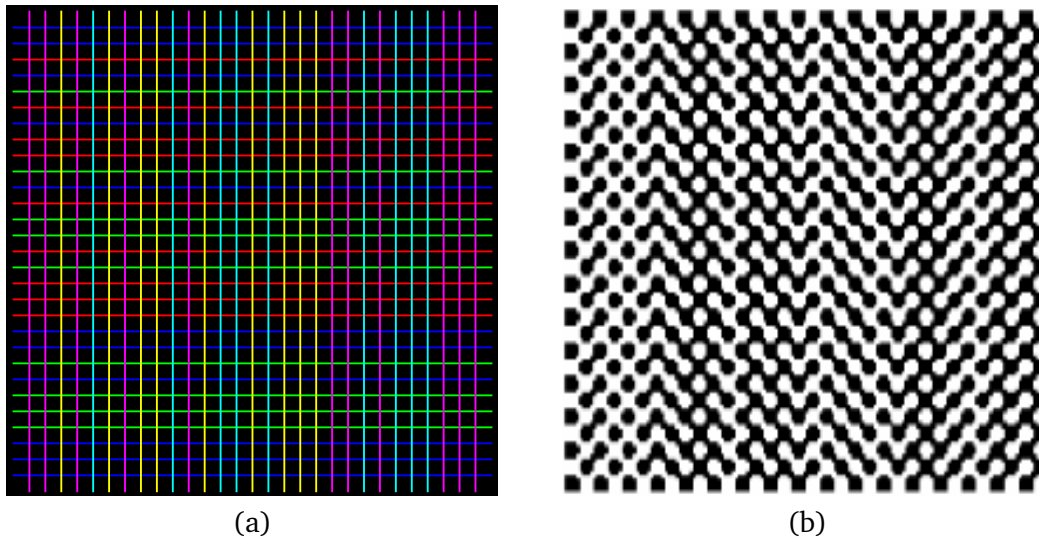


FIGURE 2.4 – Motifs codés selon les séquences de *De Bruijn* en utilisant (a) des couleurs [Salvi98] ou (b) des primitives géométriques [Vuylsteke90].

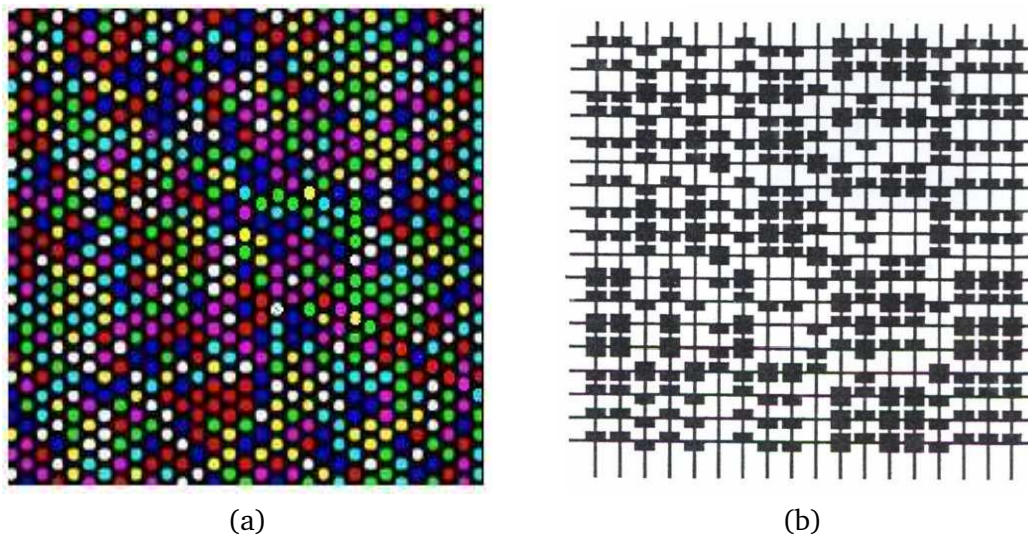


FIGURE 2.5 – Motifs codés sous la forme de *M-array* en utilisant (a) des couleurs [Adan04] ou (b) des primitives géométriques [Yee94].

Pour cette catégorie aussi, on peut envisager des techniques qui adoptent le codage des éléments composants le motif avec un alphabet de couleurs comme cela a été proposé dans [Adan04] (Figure 2.5-a), [Davies98], [Kiyasu95] ou encore [Griffin92]. Mais, on peut aussi songer à des méthodes qui proposent des motifs monochromes en utilisant des formes géométriques comme dans [Morita88], [Petriu92], [Spoelder00] et [Yee94] (Figure 2.5-b). Malgré le taux de décodage élevé de ces méthodes par rapport à celles basées purement sur la séquence de *De Bruijn* puisqu'elles adoptent un codage bidimensionnel robuste, les matrices pseudo-aléatoires restent difficiles à générer et la segmentation est délicate. Compte tenu de la robustesse que peut proposer une telle technique vis-à-vis les occultations et les ombres, c'est le principe que nous allons retenir pour définir le motif à utiliser pour notre application.

2.2 Conception d'un nouveau motif

L'application que nous visons traite des surfaces d'organes situées à l'intérieur de la cavité abdominale. Étant donnée la particularité de cet environnement et de la stratégie que nous avons choisie de suivre, nous avons été amenés à concevoir un nouveau motif auquel nous avons associé un codage robuste. Le motif que nous avons recherché doit répondre aux contraintes imposées par l'application, par l'environnement dans lequel il est appelé à être utilisé ainsi que par la réalisation d'un outil adéquat. Il doit en outre permettre des traitements rapides avec une bonne précision.

Dans la suite, nous détaillons les différentes contraintes qui ont dicté nos choix et celles que nous avons nous-même imposées à la génération du motif. Ensuite, nous présentons les approches et les algorithmes adoptés pour le traitement des images du motif. Finalement, les résultats expérimentaux sont présentés afin de valider nos choix et de confirmer l'efficacité et la robustesse du motif proposé.

2.2.1 Contraintes imposées pour la conception du motif

Nous nous intéressons à la reconstruction 3D et en temps réel des surfaces en mouvement à l'aide d'une projection de lumière structurée. Il est clair que parmi toutes les méthodes, celles qui font appel au multiplexage temporel ne conviennent pas pour des applications appelées à traiter des scènes dynamiques. Il faut donc choisir une méthode qui s'appuie sur une projection d'un unique motif, comme le font les méthodes reposant sur le voisinage spatial et le codage direct. Le codage direct est très sensible au bruit et aux variations de luminance. Par conséquent, une méthode qui met à profit le voisinage spatial est plus appropriée pour mieux répondre aux besoins de notre application. La difficulté de la mise en œuvre d'une telle méthode réside alors dans le choix du motif lumineux à projeter, choix qui est conditionné par le code qui lui est associé ainsi que par l'algorithme de décodage.

Dès que l'on s'intéresse à des scènes non contrôlées, comme cela est le cas pour les images acquises à l'intérieur de l'abdomen, nous nous trouvons confrontés aux problèmes que peut poser la présence d'ombres ou d'occultations. De telles situations peuvent produire des désordres ou des pertes d'information entre les composants du motif projeté et ainsi produire des erreurs lors de la phase de décodage. Pour résoudre au mieux ces problèmes, nous avons choisi de coder notre motif selon la formulation des "*M-arrays*". Bien qu'un peu plus difficile à concevoir, et nécessitant un décodage plus complexe, ce schéma de codage possède des propriétés mathématiques intéressantes et offre une bonne robustesse par rapport aux ombres et aux occultations.

De plus, afin d'augmenter encore la robustesse du motif, il est nécessaire d'utiliser un codage permettant de surmonter certaines erreurs qui peuvent se produire au cours de la segmentation ou du décodage. En d'autres termes, il faut définir un code autorisant une discrimination forte entre les mots de code. La distance de *Hamming* est un critère bien connu en théorie de l'information. Elle quantifie la différence entre deux séquences de symboles. Si cette distance est supérieure à 1, cela signifie que chaque paire de mots de code se différencie par au moins un symbole. Elle quantifie donc la capacité à identifier les mots de code incomplets en cas de perte ou de désordre. Ce critère est utilisé ici pour caractériser la différence entre deux mots de code correspondants à deux éléments distincts du motif. Il s'agit donc d'associer à chaque élément du motif un code judicieux. Par conséquent, il évalue la possibilité de surmonter les erreurs qui peuvent survenir dans les étapes de segmentation et de décodage.

Cependant, l'unicité des mots de code dans la plupart des méthodes correspond à une distance de *Hamming* de 1, ce qui n'est pas très efficace pour la correction d'erreurs. Il a été démontré qu'on peut obtenir de meilleurs résultats si la distance de *Hamming* est supérieure à 3 [Morano98]. Par contre, nous constatons que les méthodes proposées pour obtenir une telle distance utilisent souvent un alphabet avec un nombre de symboles assez important. C'est notamment le cas dans [Morano98] où plus de 8 symboles ont été utilisés avec une propriété de fenêtre 3×3 ou dans [Adan04] où 7 symboles ont été considérés pour atteindre une distance de *Hamming* moyenne de 3.15. Cependant, pour simplifier autant que possible la fabrication du motif ainsi que le décodage du motif projeté, il serait souhaitable de manipuler le moins de symboles possible.

Pour matérialiser ces symboles, il est possible d'utiliser un spectre de couleurs où chaque couleur serait associée à un symbole, ou bien de considérer des primitives de formes géométriques différentes telles qu'un symbole est attribué à une forme. Or, en pratique, on constate que les motifs monochromes, voire binaires, sont plus robustes que les motifs colorés. En effet, il est beaucoup plus facile de traiter une image dans laquelle on ne trouve que deux niveaux d'intensité. Il peut aussi s'avérer délicat de choisir des longueurs d'onde à utiliser pour le motif coloré dans le cas où les scènes sont très colorées. L'interaction lumière-matière peut alors jouer un rôle perturbateur avec plusieurs effets possibles comme la diffusion. À ce propos, nous avons mené des expérimentations sur des organes de porc pour voir ce qu'il en était. On peut constater qu'utiliser un spectre de couleurs n'est pas forcément judicieux (Figure 2.6). Il vaut donc mieux matérialiser les éléments du motif par des formes géométriques et ainsi n'utiliser qu'une lumière monochromatique.

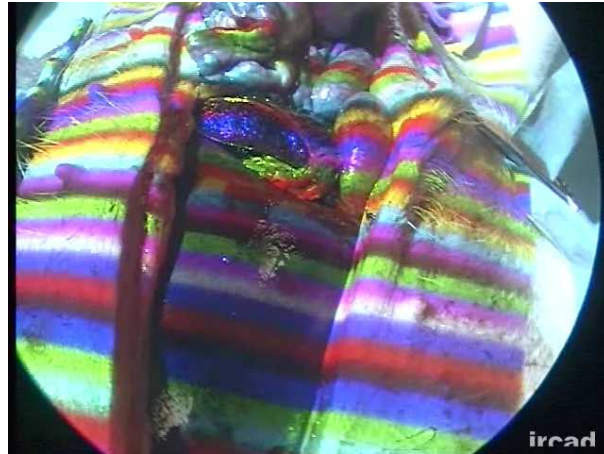


FIGURE 2.6 – Projection d'un spectre de couleurs sur les organes de porc.

Finalement, puisque le code d'un élément du motif dépend de son voisinage, il faut être en mesure de détecter correctement ce voisinage sans imposer de contraintes quant à la géométrie (position et orientation) du dispositif d'acquisition. Deux solutions ont été proposées pour résoudre ce problème. La première est de relier les primitives considérées entre elles pour faciliter l'identification des voisins [Yee94] (Figure 2.5-b). Malheureusement, cela peut induire de nouvelles difficultés de détection et d'interprétation lors de la segmentation des images. La deuxième solution définit les mots de code en tant que groupes de symboles sans prendre compte

de l'ordre de ces symboles. Dans ce cas, un nombre important de symboles est nécessaire pour construire un motif avec une bonne résolution comme c'est le cas dans [Adan04] où 7 couleurs étaient considérées (Figure 2.5-a) pour définir un motif de 986 points.

Pour résumer, l'idéal serait de disposer d'un motif codé à partir d'un alphabet construit autour d'un minimum de symboles de telle sorte que la distance de *Hamming* du code soit assez importante et qu'on soit capable de garantir, pour chaque composant du motif, la détection correcte du voisinage sans avoir besoin d'imposer de contraintes sur le dispositif d'acquisition.

2.2.2 Propriétés du motif

Afin de répondre aux contraintes que nous avons présentées dans la section précédente, nous avons cherché à concevoir un nouveau motif monochrome basé sur la théorie des *M-array*. Une propriété de fenêtre 3×3 a été considérée. Par conséquent, chaque élément du motif sera associé à un code de taille 9. Puisque nous visons à intégrer le système de vision à lumière structurée dans un outil endoscopique, des solutions technologiques présentant un encombrement aussi faible que possible doivent être envisagées pour mettre en œuvre la fabrication de l'outil. C'est ainsi que nous nous sommes tournés vers une technologie basée sur les optiques diffractives qui nous semble très appropriée pour réaliser un outil le plus compact possible. Contrairement à une optique standard (lentille ou miroir) qui est basée sur la réflexion et la réfraction de la lumière à l'interface de deux matériaux d'indices différents, les éléments d'optiques diffractives mettent à profit la diffraction de la lumière cohérente pour générer un motif prédéfini. Plus de détails concernant ces optiques seront présentés dans le chapitre 4. Or, il a été montré qu'on arrive à faire le meilleur usage de l'énergie lumineuse incidente avec ces éléments lorsque le motif présente une symétrie centrale. L'élément diffractif utilise seulement en moyenne 40% de l'énergie du laser incident pour générer un motif quelconque. Si par contre le motif présente une symétrie centrale, les éléments d'optiques diffractives peuvent exploiter jusqu'à 80% de l'énergie du rayonnement de laser. C'est la raison pour laquelle une contrainte de symétrie centrale a été rajoutée à la conception de motif.

Donc, le motif que nous souhaitons générer doit respecter les propriétés suivantes :

1. le code de chaque sous-matrice 3×3 est unique,
2. la distance de *Hamming* doit être élevée,
3. le motif présente une symétrie centrale,
4. la lumière émise est monochromatique,
5. le décodage doit être efficace et rapide.

La contrainte qui impose l'unicité du code de chaque sous-matrice 3×3 correspond à une distance de *Hamming* égale à 1. Pour rendre le motif encore plus robuste, nous avons souhaité que cette distance soit supérieure à 1. Pour ce faire, nous avons choisi d'imposer une contrainte supplémentaire telle que chaque sous-matrice 3×3 reste unique même si les symboles correspondant aux coins supérieurs de la fenêtre sont manquants. Puisque la matrice du motif que nous souhaitons élaborer est à symétrie centrale, chaque sous-matrice 3×3 restera unique même si les deux éléments des coins inférieurs sont également manquants. Par conséquent, une différence minimale est assurée entre les mots de code. Compte tenu du choix d'une source lumineuse monochromatique, les symboles de l'alphabet utilisé ne peuvent pas être codés par des couleurs

différentes et sont donc, de ce fait, représentés par des formes géométriques différentes. À cet effet, nous proposons alors de définir un alphabet réduit à trois symboles dont les primitives géométriques associées sont : un disque, un anneau et un tiret (Figure 2.7).



FIGURE 2.7 – Alphabet proposé (de gauche à droite) : un disque, un anneau et un tiret.

Le choix de l'anneau et du disque a été motivé par la volonté de simplifier aussi bien le traitement d'images associé à la détection de ces primitives que la reconnaissance des symboles. Le troisième symbole est associé à une primitive rectiligne, le tiret. Ce dernier symbole nous permettra, grâce à l'information directionnelle qu'il porte, de bien orienter les fenêtres d'observation lors de la recherche du voisinage de chaque élément et cela évite la complexité qui réside de la connexion entre les primitives ; le décodage s'en trouvera facilité.

2.2.3 Génération du motif

Compte tenu des propriétés et des contraintes que nous venons d'énumérer, il s'agit maintenant de générer un motif qui servira à la projection de lumière structurée. Dans [Morano98], un algorithme est proposé pour générer une matrice pseudo-aléatoire. L'algorithme proposé prend en compte la taille de l'alphabet, la propriété de fenêtre et la distance de *Hamming* pour générer une matrice de taille prédéfinie en utilisant l'approche dite "*bruteforce*". Par exemple, pour générer une matrice pseudo-aléatoire en utilisant trois symboles avec une propriété de fenêtre de 3×3 , on suit les étapes suivantes :

1. On choisit une sous-matrice de taille 3×3 d'une manière aléatoire au sein de l'alphabet considéré (Figure 2.8-a).
2. On génère aléatoirement une colonne de taille 3×1 et on l'ajoute à droite de cette sous-matrice (Figure 2.8-b).
3. On choisit une rangée de taille 1×3 de même manière et on l'ajoute sous la sous-matrice (Figure 2.8-c).
4. L'ajout est validé si les contraintes de la propriété de fenêtre et la distance de *Hamming* sont vérifiées. Si Les contraintes sont validées on procède, sinon on revient à la deuxième étape pour choisir une nouvelle colonne et une nouvelle ligne.
5. Les coordonnées de départ sont incrémentées de 1 et le processus horizontal et le processus vertical sont répétés (Figure 2.8-d).

On réitère ces opérations jusqu'à ce que la matrice soit remplie. Si la procédure arrive à un état où on ne peut plus ajouter d'éléments en respectant les contraintes imposées, l'algorithme recommence avec une autre sous matrice initiale 3×3 .

L'étude de la performance de cet algorithme [Morano98] montre que si on considère une matrice pseudo-aléatoire de taille 45×45 et une propriété de fenêtre 3×3 , un alphabet d'au moins 3 est symboles nécessaire pour atteindre une distance de *Hamming* supérieure à 1 et de plus que 8 symboles pour atteindre à une distance de *Hamming* supérieure à 3.

En s'inspirant du travail de [Morano98], nous avons implémenté un algorithme de génération de matrice respectant les contraintes citées dans la section précédente. Notre algorithme prend en compte la propriété de fenêtre, la taille de l'alphabet, la symétrie centrale et la contrainte supplémentaire que nous avons ajoutée. En vue de projeter le motif sur une surface de taille significative, nous avons cherché à maximiser les dimensions de la matrice. Ainsi, nous avons obtenu une matrice de dimensions (27×29) , soit 783 primitives, avec trois symboles (Figure 2.9). Puisque nous nous intéressons à la vision endoscopique pour laquelle les dimensions des surfaces imagées sont d'environ $(10 \times 10) \text{ cm}^2$ donc, la résolution que nous obtenons avec notre matrice est d'environ $(1 \text{ point} / 0,14 \text{ cm}^2)$. Cela semble être une résolution acceptable pour des applications dont la précision désirée est de l'ordre de 1mm. Le motif construit à partir des trois primitives géométriques est illustré sur la Figure 2.10.

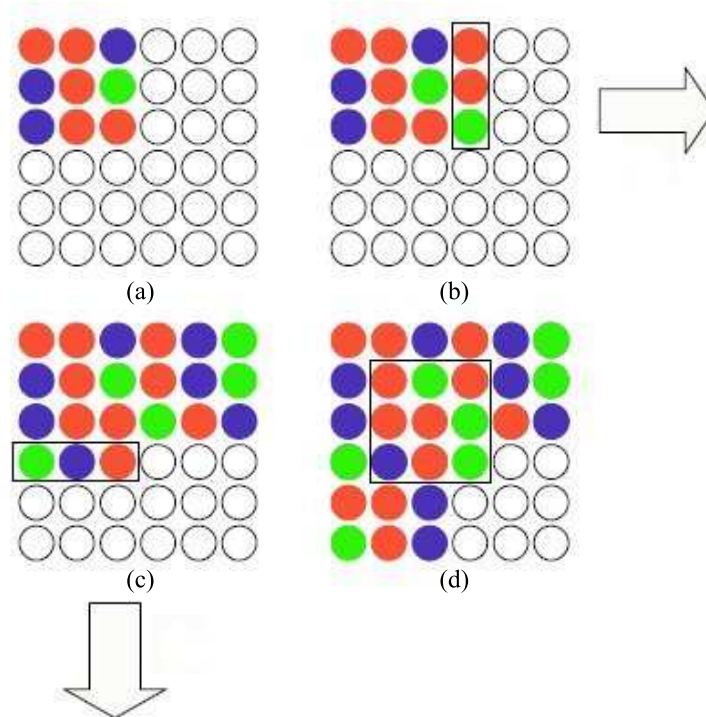


FIGURE 2.8 – Algorithme proposé dans [Morano98].

Pour vérifier la robustesse du motif que nous avons généré, nous avons choisi d'évaluer un critère qui est la distance de *Hamming* moyenne. Pour ce faire, nous avons calculé, pour chaque mot de code, la distance de *Hamming* avec tous les autres mots de code. Nous avons défini la fonction $H(x)$ comme étant le pourcentage de mots de code ayant une distance de *Hamming* égale à x . Ainsi, pour la matrice présentée dans la Figure 2.9, nous avons obtenu les résultats suivants : $H(1) = 0,05$, $H(2) = 0,71$, $H(3) = 3,45$, $H(4) = 10,40$, $H(5) = 20,24$, $H(6) = 27,07$, $H(7) = 23,55$, $H(8) = 11,83$ et $H(9) = 2,70$. Ces résultats numériques sont récapitulés sur la Figure 2.11. On notera que 99,95% des mots de code ont une distance de *Hamming* supérieure à 1, et que 95,79% des mots de code ont une distance de *Hamming* supérieure à 3. En d'autres termes, cela veut dire que 95,79% de mots de code se différencient par au moins 3 symboles et que plus de 65,15% ont au moins 6 symboles différents. Statistiquement, la distance de *Hamming* moyenne est $\bar{H} = 6,0084$. Cela diminue beaucoup la probabilité d'erreurs dans l'étape de décodage. Finalement, quand nous ne considérons, pour chaque mot de code d'un élément, que


```

2 1 1 0 2 1 2 0 1 1 2 1 0 0 2 0 0 2 1 0 0 0 1 2 2 2 1 1 2
1 0 0 2 2 2 0 2 2 1 0 2 1 1 2 0 2 0 0 1 1 0 0 1 0 2 0 0 2
0 2 0 0 2 0 0 0 1 1 1 0 1 1 2 0 0 2 0 1 1 2 1 1 2 1 1 2 1
2 1 2 1 1 1 2 1 0 0 0 1 2 0 1 2 0 1 0 1 1 2 2 0 2 0 0 1 1
2 2 1 2 1 2 0 1 0 2 2 1 2 0 1 0 0 0 2 2 2 1 1 2 0 0 0 2
0 2 1 1 0 2 0 2 0 2 1 0 0 2 0 0 1 0 2 2 1 0 0 1 0 2 2 1 2
0 1 2 2 0 1 1 2 1 1 2 0 2 2 0 1 0 1 0 2 0 1 1 0 2 2 2 0 2
2 0 0 1 2 0 2 2 2 1 0 2 0 2 2 2 0 1 0 0 0 2 1 0 1 0 1 1 1
1 2 0 0 2 1 0 1 0 0 1 2 0 1 2 1 2 2 2 0 1 1 2 0 2 2 1 2 0
1 1 1 2 2 2 1 2 0 1 1 1 2 1 0 0 1 1 1 0 2 0 1 2 0 1 2 0 2
2 1 1 0 0 2 2 1 1 0 1 2 1 0 2 2 0 1 0 1 1 2 0 1 1 1 0 1 2
1 1 0 2 1 2 1 2 1 1 2 2 2 1 2 2 1 1 0 0 1 1 0 0 0 2 0 2 0
0 0 2 0 2 1 0 0 0 1 2 2 0 1 1 1 0 2 1 2 2 1 2 0 2 1 1 0 1
1 2 0 2 2 1 2 2 0 2 0 1 2 0 0 0 2 1 0 2 0 2 2 1 2 2 0 2 1
1 0 1 1 2 0 2 1 2 2 1 2 0 1 1 1 0 2 2 1 0 0 0 1 2 0 2 0 0
0 2 0 2 0 0 0 1 1 0 0 1 1 2 2 1 2 2 1 1 2 1 2 1 2 0 1 1
2 1 0 1 1 1 0 2 1 1 0 1 0 2 2 0 1 2 1 0 1 1 2 2 0 0 1 1 2
2 0 2 1 0 2 1 0 2 0 1 1 1 0 0 1 2 1 1 1 0 2 1 2 2 2 1 1 1
0 2 1 2 2 0 2 1 1 0 2 2 2 1 2 1 0 2 1 0 0 1 0 1 2 0 0 2 1
1 1 1 0 1 0 1 2 0 0 0 1 0 2 2 2 0 2 0 1 2 2 2 0 2 1 0 0 2
2 0 2 2 2 0 1 1 0 2 0 1 0 1 0 2 2 0 2 1 1 2 1 1 0 2 2 1 0
2 1 2 2 0 1 0 0 1 2 2 0 1 0 0 2 0 0 1 2 0 2 0 2 0 1 1 2 0
2 0 0 0 2 1 1 1 2 2 2 0 0 0 1 0 2 1 2 2 0 1 0 2 1 2 1 2 2
1 1 0 0 2 0 2 2 1 1 0 1 0 2 1 0 2 1 0 0 0 1 2 1 1 1 2 1 2
1 2 1 1 2 1 1 2 1 1 0 2 0 0 2 1 1 0 1 1 1 0 0 0 2 0 0 2 0
2 0 0 2 0 1 0 0 1 1 0 0 2 0 2 1 1 2 0 1 2 2 0 2 2 2 0 0 1
2 1 1 2 2 2 1 0 0 0 1 2 0 0 2 0 0 1 2 1 1 0 2 1 2 0 1 1 2

```

FIGURE 2.9 – Matrice globale à symétrie centrale.

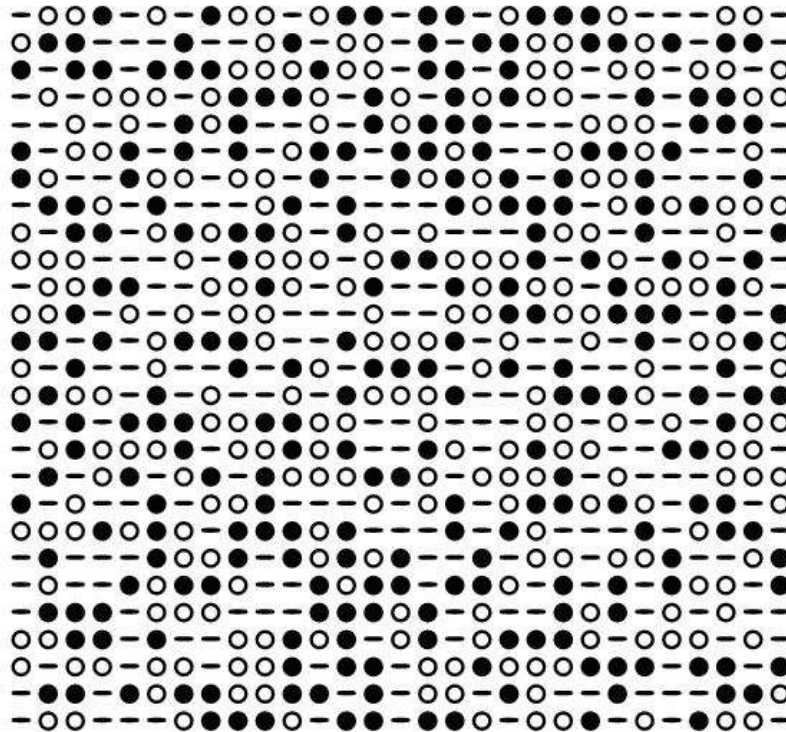


FIGURE 2.10 – Motif proposé : les disques, cercles et tirets correspondent respectivement aux 0, 1, 2 dans la matrice.

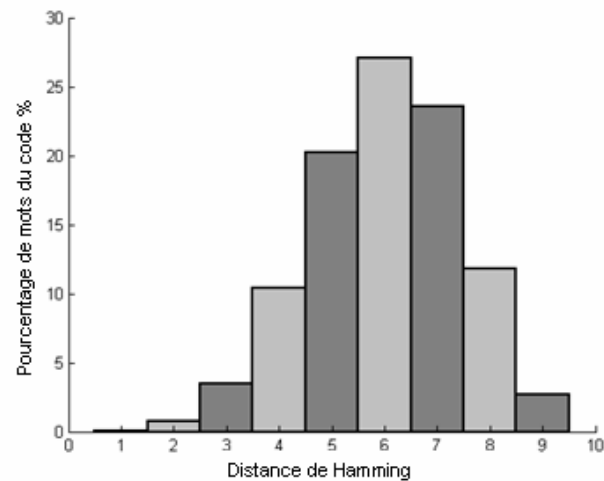


FIGURE 2.11 – Distribution de la distance de *Hamming* dans le motif proposé.

les mots de code de ses huit plus proches voisins, la distance de *Hamming* moyenne obtenue est alors de $\bar{H} = 6,173$. En effet, à notre connaissance, la distance moyenne la plus élevée obtenue à ce jour, pour d'autres motifs, est de 3,15 [Adan04].

Ce résultat permet d'avoir une évaluation de la robustesse de notre code associé au motif proposé. En utilisant notre motif, nous ne pouvons pas garantir une distance de *Hamming* constante pour tous les éléments du code. Cependant, à la vue des résultats (Figure 2.11), on constate que par rapport aux motifs proposés par ailleurs, nous avons diminué la probabilité d'avoir des erreurs suite à mauvaise détection ou classification de primitives. D'autre part, le pourcentage de mots de code différents de plus de 6 symboles est grand (65.15%). Par rapport aux autres motifs, celui que l'on propose permet de diminuer très sensiblement le taux d'erreurs lors de décodage. Ces résultats sont très encourageants et on peut cependant noter que des travaux en cours au laboratoire visent à encore les améliorer.

2.2.4 Algorithmes de segmentation et de décodage associés

Maintenant que nous avons défini un motif, celui-ci sera projeté sur les surfaces à reconstruire et une image de cette projection sera acquise. Il s'agit ensuite d'analyser cette image de manière à en extraire les primitives géométriques que nous avons considérées pour les décoder, en d'autres termes, pour identifier chaque primitive dans l'image.

Afin d'obtenir un décodage rapide, il faut associer au motif proposé des algorithmes de segmentation et de classification efficaces. Une approche basée sur les contours a été mise en œuvre pour la détection des symboles. En premier lieu, les contours ont été mis en évidence par le calcul du gradient de l'image en niveaux de gris à l'aide du filtre de Canny [Canny86]. Par la suite, les contours détectés qui correspondent aux primitives du motif sont isolés. Pour cela, seuls les contours ayant une longueur minimale et qui en même temps ne dépassent pas une longueur maximale sont qualifiés pour passer à l'étape suivante, la classification. Cette opération sert à diminuer l'influence produite par les contours des surfaces se trouvant dans la scène, par la texture des surfaces ou par l'algorithme de la détection. Les seuils sont choisis expérimentalement.

Dans l'étape de classification, nous devons aboutir à étiqueter chaque contour selon son appartenance à l'une des trois classes, une par symbole, sinon il est rejeté. La classification des primitives (anneau, disque ou tiret), est basée sur une statistique du premier ordre pour l'anneau, et du second ordre pour les deux autres primitives. En pratique, l'algorithme commence par estimer les centres des contours. Or, le résultat du filtre de Canny pour le symbole représenté par l'anneau correspond à deux cercles ou ellipses concentriques (voir la Figure 2.12). Par conséquent, pour reconnaître cette primitive, on choisit simplement de rechercher deux centres confondus ou bien très proches l'un de l'autre. La localisation de cette primitive dans l'image est réalisée en calculant la position moyenne des centres des deux contours.



FIGURE 2.12 – Détection de la primitive *anneau* : deux centres très proches.

Pour les contours restant, les distances entre les pixels de chaque contour et son centre sont calculées afin de déterminer le pixel le plus proche et le pixel le plus éloigné sur ce contour (distances euclidiennes minimales et maximales). Pour chaque contour, le rapport entre ces deux distances déterminera si la primitive correspond à un disque ou à un tiret (Figure 2.13). Si ce rapport est inférieur à un seuil déterminé expérimentalement, on considère que le contour correspond à un disque, sinon il correspond à un tiret. Une fois un tiret détecté, la direction de la ligne passant par son centre et par le pixel le plus éloigné sur son contour est calculée. Cette direction correspond à l'orientation locale du tiret.

Or, pour des surfaces localement lisses, cette orientation varie très peu dans un voisinage proche. Grâce à cette information, nous pouvons prédire les directions où se trouvent les voisins de chaque élément dans l'étape de décodage et cela simplifie considérablement la recherche des primitives voisines (Figure 2.14).

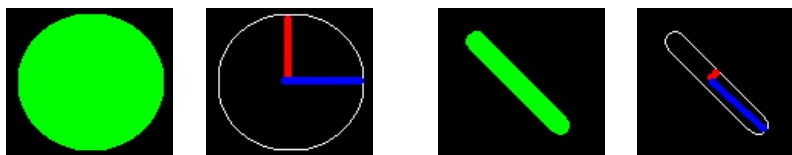


FIGURE 2.13 – Distinction entre le disque et le tiret : les distances entre le centre et le pixel le plus proche (en rouge) et le pixel le plus loin (en bleu).

Le code de chaque primitive est alors déterminé en fonction de ses voisins, selon l'ordre illustré sur la Figure 2.15. Dès que le mot de code de chaque primitive est déterminé, une matrice d'assignation établit la correspondance entre le code de chaque primitive et sa position (ligne-colonne) dans la matrice du motif. Pour trouver la position correspondant à chaque mot de code, lors de la phase d'analyse, il suffira alors d'effectuer une recherche exhaustive tenant compte des propriétés du motif.

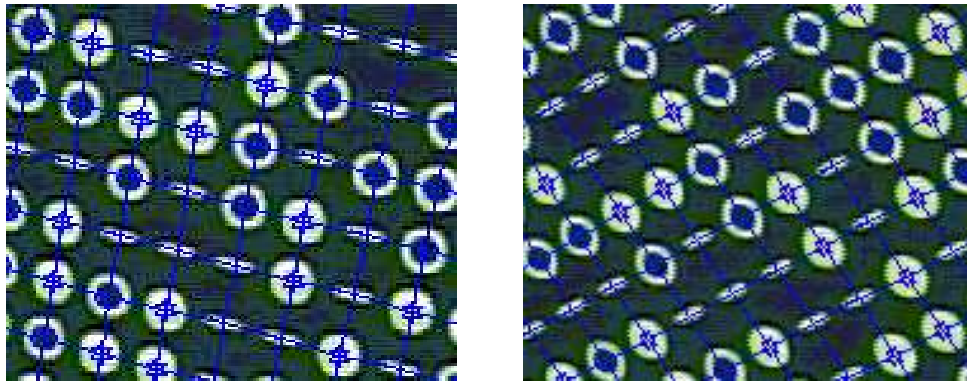


FIGURE 2.14 – Recherche des voisins à l'aide des orientations locales des tirets.

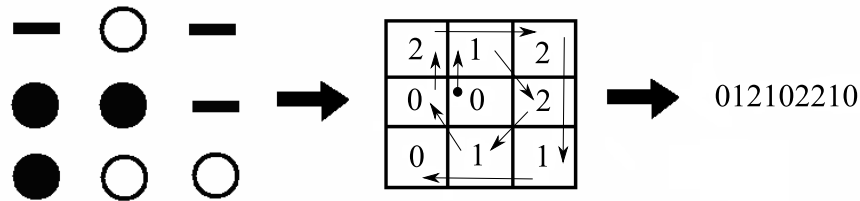


FIGURE 2.15 – Détermination du mot de code de chaque sous-matrice.

Afin de surmonter les erreurs qui peuvent survenir lors de la phase de décodage et d'améliorer encore la robustesse de notre technique, nous avons mis à profit la méthode de votes proposé dans [Morano98] pour déterminer correctement la position de chaque primitive dans la matrice. En effet, puisque nous considérons une propriété de fenêtre de 3×3 , chaque élément de la matrice pseudo-aléatoire est contenu dans une sous matrice 3×3 et il appartient également à chaque sous-matrice 3×3 de ses huit voisins. Comme chaque élément est associé à un mot de code défini par sa fenêtre, chaque fenêtre aura 9 mots de code correspondants à ses éléments composants. Par conséquent, un vote est effectué pour chaque fenêtre par ses éléments pour trouver la position de cette fenêtre dans la matrice correspondant au nombre maximum de voix. Dans la section suivante, nous constaterons que cela améliore notablement les résultats en fournissant une information plus fiable de la position des éléments dans la matrice.

2.3 Résultats expérimentaux

Les résultats que nous présentons dans cette section sont destinés à évaluer le choix du motif et à valider la méthode de codage proposée. Les expérimentations présentées ici ont été réalisées en laboratoire dans des conditions contrôlées. Le dispositif utilisé pour la validation est composé d'une caméra (AVT Pike F100C) qui fournit des images en couleurs (RGB) de taille 512×512 pixels, et un vidéo-projecteur (Sony VPL-CS6) pour projeter le motif dans la scène.

Après avoir projeté le motif sur une surface plane d'environ $(20 \times 20) \text{ cm}^2$, nous avons appliqué les algorithmes de détection et de décodage précédemment décrits sur les images de la projection. Nous avons ensuite fait subir des rotations à notre motif afin d'évaluer les capacités

de l'algorithme à détecter les voisins correctement en prenant en compte l'orientation du motif grâce à la primitive directionnelle du motif (le tiret). Dans la Figure 2.17, nous présentons les résultats de cette expérimentation. Les primitives matérialisées en bleu sont les primitives visibles et détectées, celles en rouge sont celles qui de surcroît ont été décodées correctement (attribution de la bonne adresse ligne-colonne du symbole dans le modèle du motif). Nous constatons que non seulement le décodage se fait correctement mais que la détection est pratiquement indépendante de la rotation appliquée au motif. Ceci peut être attribué logiquement à l'apport de la primitive directionnelle parmi les symboles de l'alphabet utilisé.

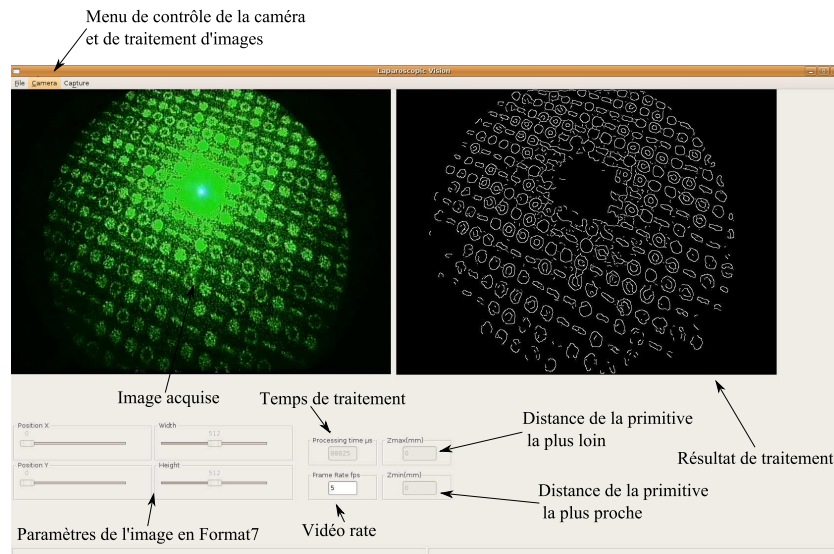


FIGURE 2.16 – Interface développée afin de permettre à l'utilisateur d'acquérir d'images et d'appliquer nos algorithmes de détection, de décodage et de reconstruction.

Nous avons également projeté le motif sur une surface cylindrique (Figure 2.18) et nous avons constaté que de nombreuses primitives, bien que détectées, ne sont pas décodées correctement à cause de détériorations, notamment près des discontinuités fortes de la surface, et des erreurs lors de la détection de voisins. Il en est de même lorsque nous avons à faire à des occultations et à des ombres projetées (Figure 2.19). Pour pallier ces problèmes, nous avons complété notre algorithme avec la méthode des votes présentée dans la section 2.2.4.

En appliquant maintenant notre algorithme, complété par la méthode des votes, nous observons immédiatement les améliorations (Figure 2.20, Figure 2.21 et Figure 2.22). Ces résultats sont synthétisés dans le tableau 2.1. Dans ce tableau, nous résumons pour chaque cas illustré par les figures (2.17-2.22) le nombre de primitives détectées et celles qui ont été décodées correctement de surcroît. Nous comparons ensuite ces chiffres à ceux obtenus sans utiliser la méthode de votes. Il est clair en observant ces chiffres que l'apport de la méthode est indéniable. Il est d'ailleurs d'autant plus grand, lorsque les surfaces deviennent complexes. Nous constatons aussi que le pourcentage des primitives visibles qui ont été détectées et décodées correctement est de plus de 96% pour la surface plane et de plus de 82% pour la surface cylindrique. Malgré les occultations et les ombres, plus de 90% des primitives détectées ont aussi été décodées correctement. Ces résultats sont encourageants et montrent l'efficacité du motif proposé.

Cependant, On constate que certaines primitives au bord de l'image ou celles se situant au niveau de ruptures de surfaces n'étaient pas bien décodées alors qu'elles étaient bien détectées. Cela est dû au fait que le décodage est fonction de la cohérence de sa position avec celles de ses voisins.

Scène	Primitives détectées	Primitives décodées	%	Primitives décodées sans les votes	%
Plan 1	684	672	98.2	581	85.1
Plan 2	586	563	96.1	396	67.4
Plan 3	611	588	96.2	496	81.2
Cylindre	508	421	82.9	290	57.1
Objets avec ombres	713	637	89.3	395	55.4

TABLE 2.1 – Résultats de la détection et du décodage.

De façon à respecter les objectifs que nous nous étions fixés, à savoir une reconstruction 3D en temps réel de surfaces intra abdominales, nous avons également estimé le temps nécessaire au traitement d'une image. On constate que lorsqu'on considère des images de taille (512×512) pixels, le temps de traitement est d'environ 50 ms. Bien qu'aucune optimisation n'est encore faite, cette durée, bien que ne correspondant pas à du vrai temps réel vidéo, s'en approche beaucoup et permet de visualiser correctement le suivi d'une séquence avec l'interface que nous avons développé (Figure 2.16).

2.4 Conclusion

Afin d'atteindre un des objectifs visés par nos travaux, la reconstruction 3D en temps réel de surfaces en mouvement, nous proposons d'utiliser une méthode basée sur la projection de lumière structurée. À cet effet, nous avons conçu un nouveau motif lumineux monochrome. Nous avons associé à ce motif un codage robuste mettant à profit la stratégie du voisinage spatial pour laquelle le code de chaque élément est une fonction de son symbole et de ceux de ses voisins. La mise en œuvre a été faite grâce à la théorie des matrices pseudo-aléatoires en utilisant un alphabet de trois symboles en considérant une propriété de fenêtre de 3×3 . Par conséquent, chaque élément sera défini par un mot de code de taille 9. Puisque nous souhaitons utiliser une lumière monochromatique pour l'application visée, nous avons représenté les symboles du motif par des primitives géométriques simples. Parmi les primitives considérées, nous en avons utilisé une qui contient une information directionnelle. Ceci simplifie la détection des voisins et réduit le temps nécessaire pour le décodage global. La robustesse du motif proposé et du code associé a été démontrée en évaluant la distribution de la distance de *Hamming* et la distance de *Hamming* moyenne. Les premiers résultats expérimentaux présentés dans ce chapitre valident le choix du motif et les traitements associés. Ainsi, ils confirment l'efficacité de la méthode de votes utilisée au niveau de décodage et la robustesse du code, même en présence d'occultations ou d'ombres. De plus, le traitement a été très rapide compte tenu de la taille des images considérées.

Il est maintenant possible de passer à la phase de la reconstruction 3D métrique. Le but de ce travail était d'utiliser ce motif pour la reconstruction 3D rapide afin de guider un robot dans un environnement complexe, comme la cavité abdominale, à l'aide d'un système de lumière structurée embarqué, mais il est clair qu'il peut également être utilisé dans d'autres applications nécessitant la connaissance 3D de son environnement.

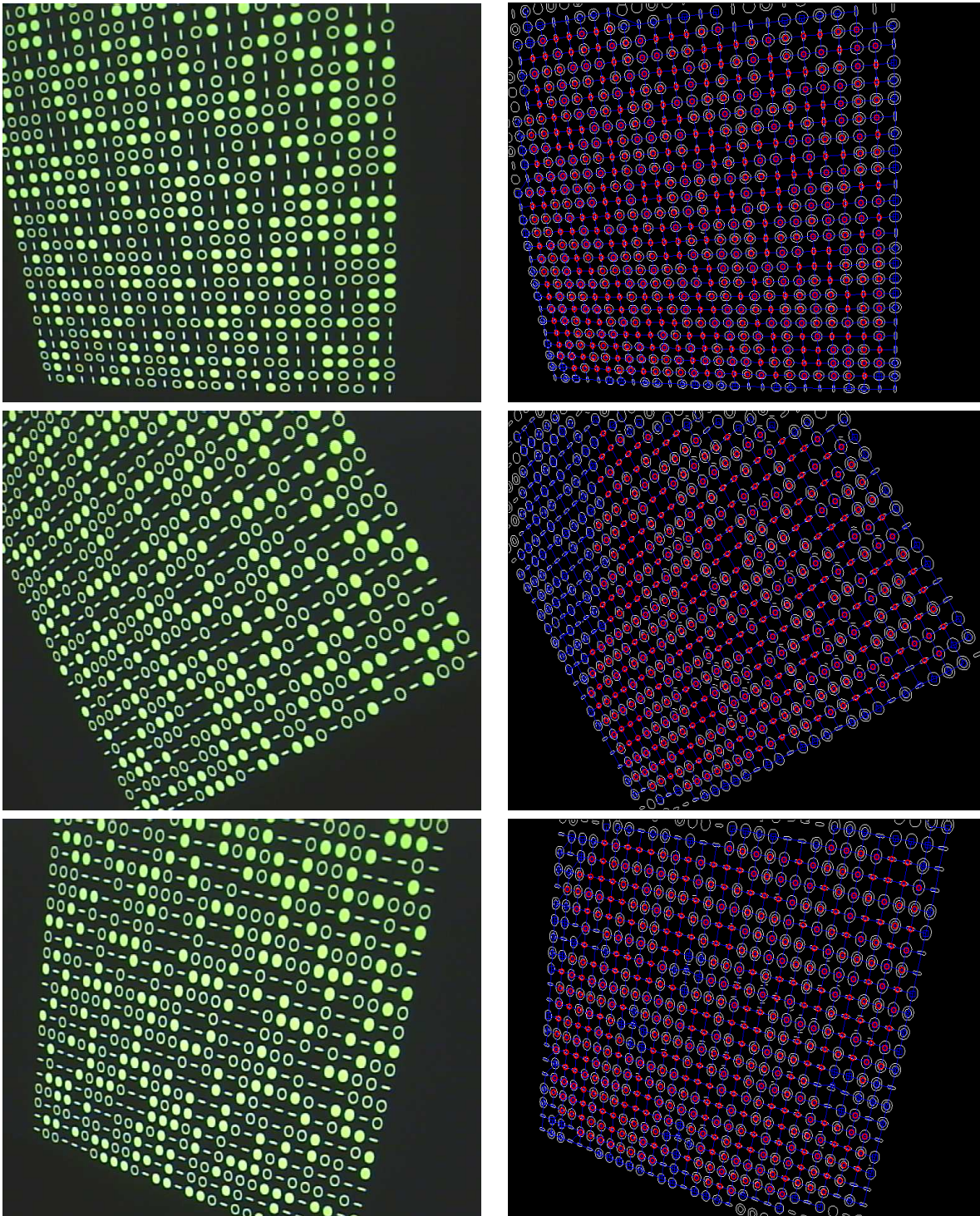


FIGURE 2.17 – Résultats de la détection et du décodage sur un plan pour différentes rotations du motif, sans prendre en compte la méthode des votes. Les primitives marquées en rouge représentent les primitives qui ont été bien détectées et décodées correctement et celles marquées en bleu sont les primitives qui ont été détectées mais qui n'ont pas été décodées correctement.

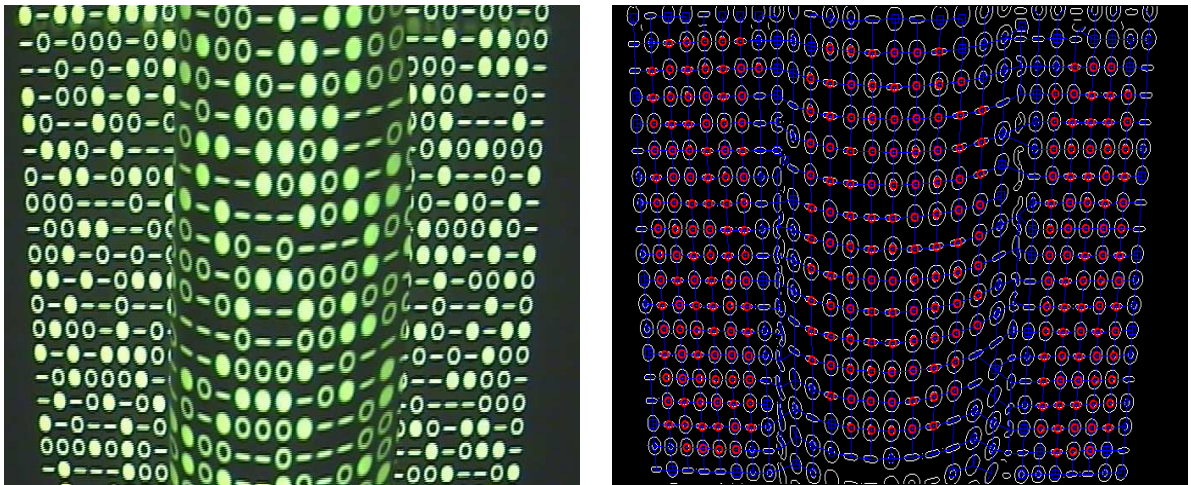


FIGURE 2.18 – Résultats de la détection et du décodage sur une surface cylindrique, sans prendre en compte la méthode des votes.

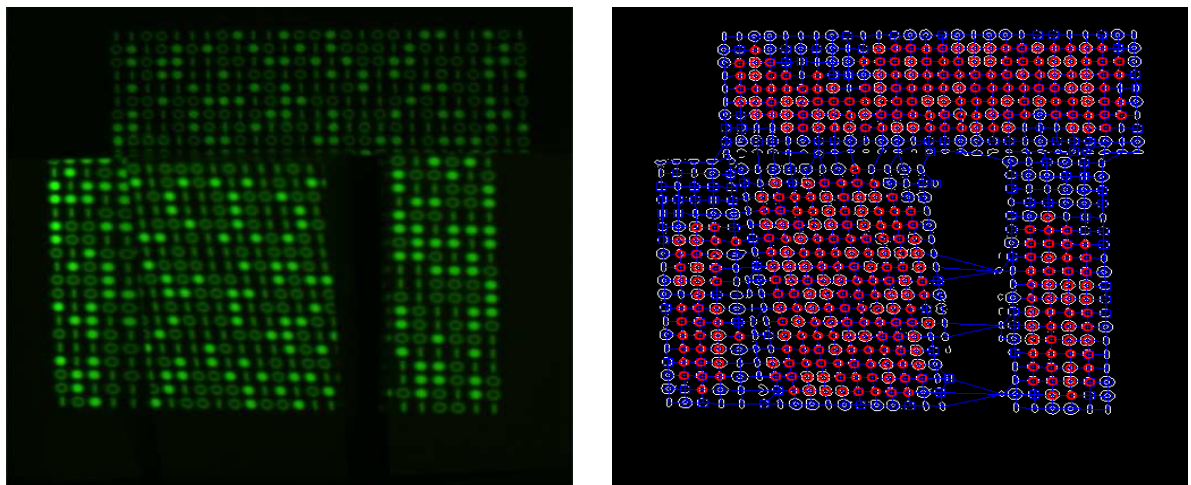


FIGURE 2.19 – Résultats de la détection et du décodage avec des occultations et des ombres, sans prendre en compte la méthode des votes sur trois plans.

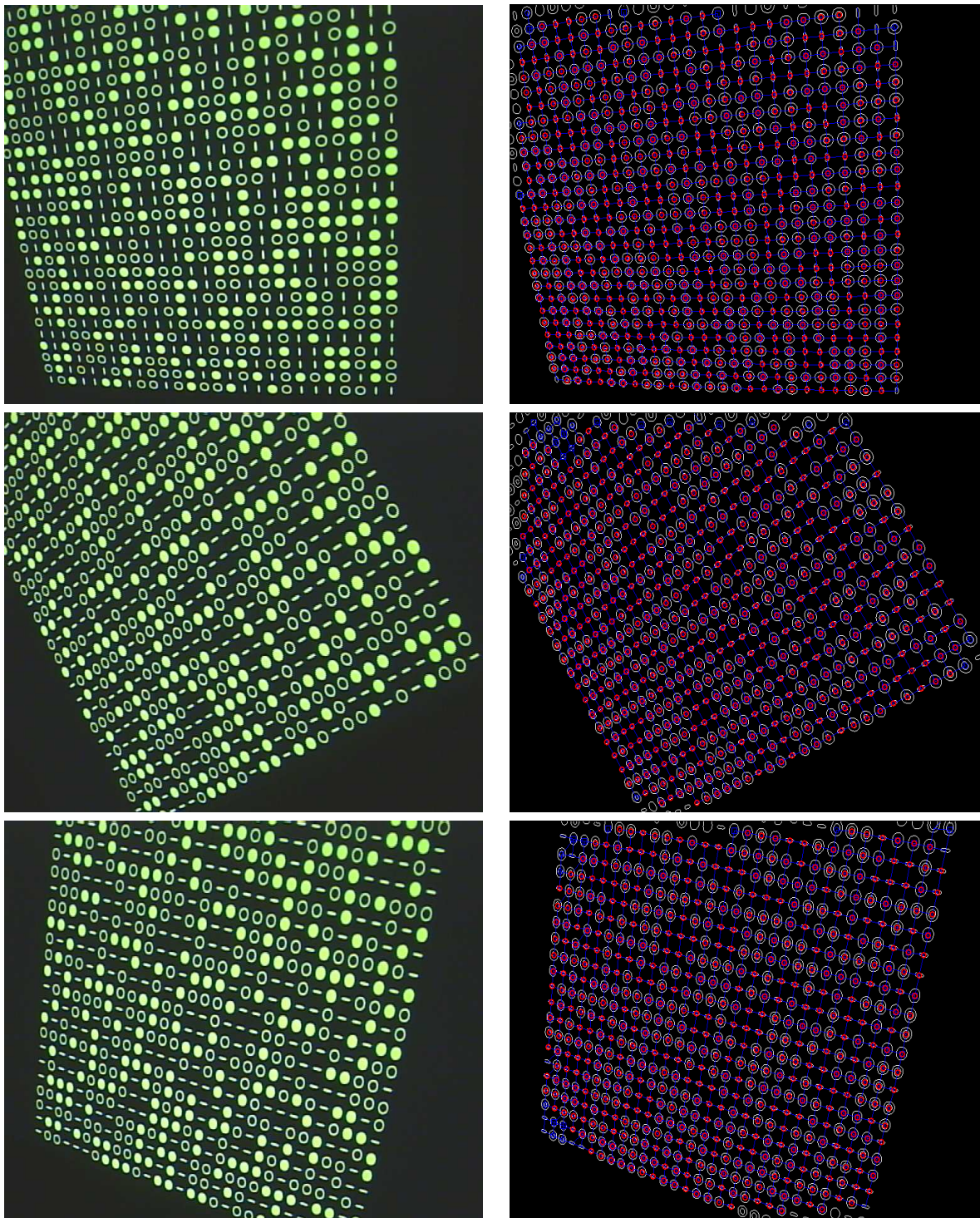


FIGURE 2.20 – Résultats de la détection et du décodage sur un plan pour différentes rotations du motif en prenant en compte la méthode des votes. Les primitives marquées en rouge représentent les primitives qui ont été bien détectées et décodées correctement et celles marquées en bleu sont les primitives qui ont été détectées mais qui n'ont pas été décodées correctement.

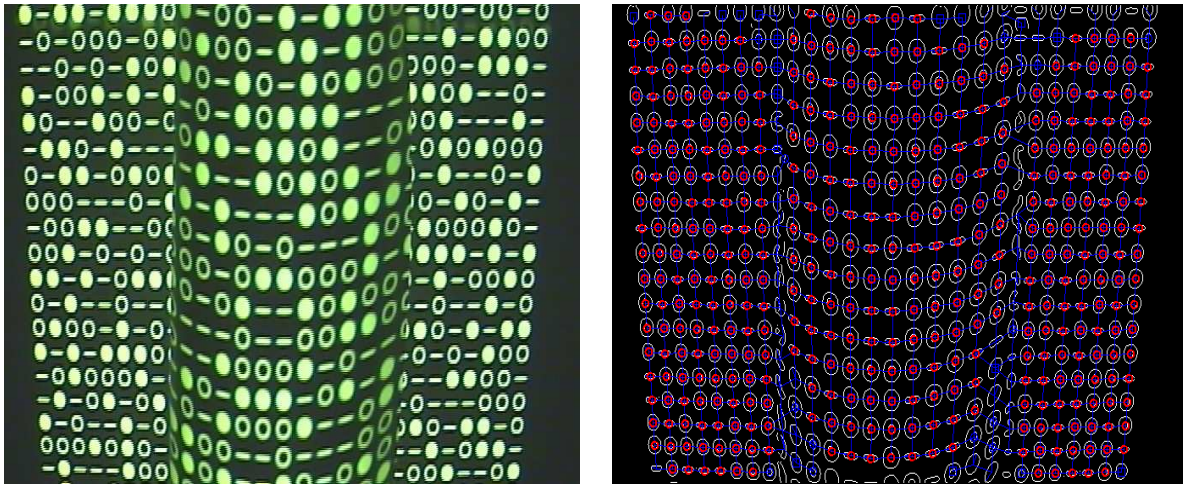


FIGURE 2.21 – Résultats de la détection et du décodage sur une surface cylindrique en prenant en compte la méthode des votes. La déformation des primitives à coté des discontinuités des surfaces ne permettait pas de bien détecter ces primitives et leurs voisins et par conséquent, elles n'étaient pas décodées correctement.

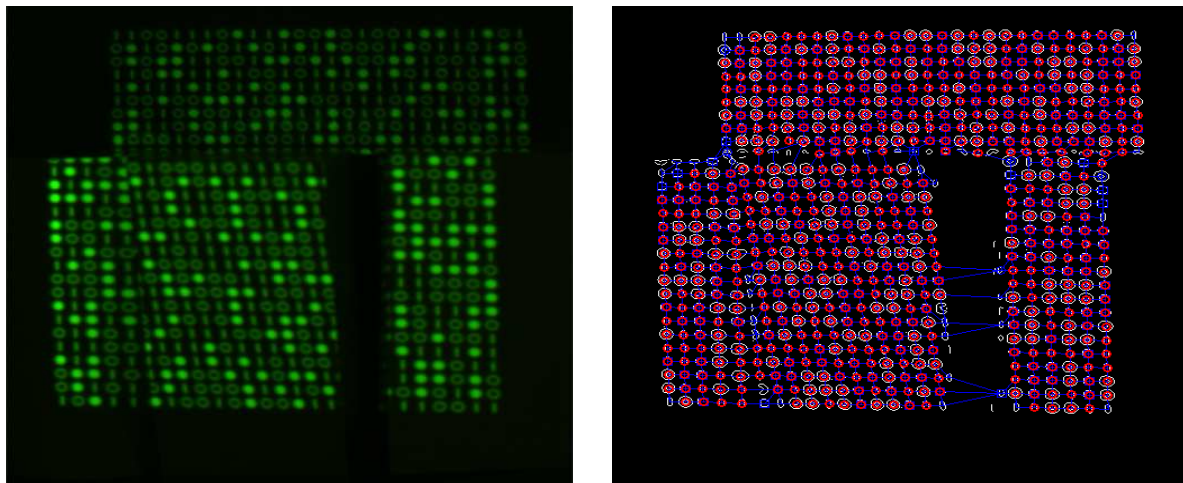


FIGURE 2.22 – Résultats de la détection et du décodage avec des occultations et des ombres, en prenant en compte la méthode des votes sur trois plans.

Chapitre 3

Reconstruction 3D et étalonnage du dispositif de lumière structurée

À partir du motif que nous venons de présenter, nous devons maintenant définir une méthode de reconstruction 3D pour des scènes dynamiques et en particulier pour des scènes intra abdominales. Comme nous avons adopté une solution basée sur une projection de lumière structurée, il s'agit de reconstruire dans l'espace les positions des points sur les surfaces imagées, matérialisés par la projection des primitives de notre nouveau motif. Dans ce chapitre, nous présentons d'abord le principe de reconstruction retenu. Cependant, la reconstruction tridimensionnelle métrique par vision, qu'elle soit passive ou active, nécessite toujours un étalonnage du système d'acquisition. Le but de cet étalonnage est de déterminer le comportement du motif projeté dans l'espace. En s'aidant de cette information, il est alors possible de déterminer la position d'une primitive, localisée dans l'image, dans un système de coordonnées 3D. Nous décrivons aussi la méthode d'étalonnage que nous avons mise en œuvre, après avoir fait un rapide tour d'horizon des techniques habituellement utilisées. Nous finirons par présenter quelques résultats expérimentaux qui permettront de valider et d'évaluer la précision de la méthode proposée.

3.1 Reconstruction 3D

Un dispositif de lumière structurée classique consiste en un projecteur de motifs et une caméra. Le principe de la reconstruction 3D, en utilisant un tel système, est alors relativement simple à exposer. Il consiste à acquérir une image de la surface à reconstruire, sur laquelle a été projeté le motif de lumière structurée. On identifie ensuite, dans l'image, les primitives projetées, puis on les associe, grâce au code qui leur a été attribué, à leurs correspondants sur le motif et on détermine leur position 3D dans l'espace à l'aide de données d'étalonnage. Ce principe est illustré sur la Figure 3.1. La caméra est représentée par le repère (O_c, X_c, Y_c, Z_c) lié au centre optique de son objectif. Le plan image, représenté par le repère (O, x, y) , est parallèle au plan (O_c, X_c, Y_c) et distant de f : la distance focale. Le point $P_i(x_i, y_i)$ est l'image du point $P_c(X_c, Y_c, Z_c)$ que l'on cherche à reconstruire. Donc, en identifiant et en localisant la primitive P_i dans l'image et en connaissant sa correspondance dans le repère du projecteur, on peut, grâce aux résultats de la calibration, déterminer sa position 3D dans le repère de la caméra. La position 3D peut être calculée en utilisant la triangulation classique comme on fait traditionnellement dans un système de stéréovision (par exemple : [Longuet-Higgins81], [Faugeras93], [Trucco98] et [Hartley00]).

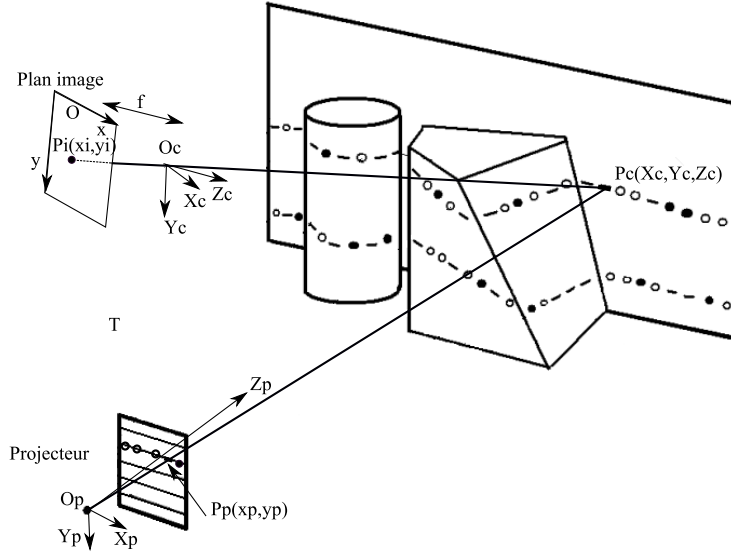


FIGURE 3.1 – Géométrie simple illustrant la triangulation avec la lumière structurée.

Dans ce cas, le projecteur et la caméra doivent être étalonnés séparément afin de déterminer leurs paramètres intrinsèques et extrinsèques, puis associés. Alors si $P_i(x_i, y_i)$ est l'image du point $P_c(X_c, Y_c, Z_c)$ matérialisé par la projection d'une primitive correspondant à la position $P_p(x_p, y_p)$ dans le repère du projecteur, les équations suivantes établissent les relations entre ces points :

$$\begin{pmatrix} P_i \\ 1 \end{pmatrix} = w_c \cdot \begin{pmatrix} P_c \\ 1 \end{pmatrix} \quad (3.1)$$

$$\begin{pmatrix} P_p \\ 1 \end{pmatrix} = w_p \cdot M \begin{pmatrix} P_c \\ 1 \end{pmatrix} \quad (3.2)$$

Dans ces équations, w_c est une matrice (3×4) représentant la projection perspective entre la scène et la caméra. Elle est défini en fonction des paramètres d'étalonnage de la caméra. La transformation w_p est une matrice (3×4) représentant la projection perspective entre la scène et le projecteur et qui contient les paramètres d'étalonnage du projecteur. La transformation entre le repère de la caméra et celui du projecteur est représentée par la matrice M (4×4) qui contient la matrice de rotation R et la translation T entre ces deux repères (Eq. 3.3).

$$M = \begin{pmatrix} R & T \\ 0^T & 1 \end{pmatrix} \quad (3.3)$$

Pour réaliser la mise en correspondance, la géométrie épipolaire peut être mis à profit. Cela consiste à déterminer le plan formé par le centre de la caméra, celui du projecteur et le point $P_c(X_c, Y_c, Z_c)$. Cela limite d'autant la recherche de correspondant. L'information sur la géométrie épipolaire est contenue dans la matrice (3×3) fondamentale F , qui permet de faire une reconstruction projective. Chaque point dans l'image et sa correspondance dans le plan du projecteur doivent vérifier la contrainte épipolaire liant entre ces deux points et la matrice fondamentale F (Eq. 3.4).

$$P_p \cdot F \cdot P_i = 0 \quad (3.4)$$

Par conséquent, en localisant une primitive dans l'image $P_i(x_i, y_i)$ et en déterminant sa correspondance dans le projecteur $P_p(x_p, y_p)$, la position tridimensionnelle du point de l'objet matérialisé par la projection de cette primitive peut être calculée. En effet, même si les contraintes épipolaires sont prises en compte, la recherche des correspondances reste un problème majeur pour ce type de technique de reconstruction [Armangue03] ainsi que quand le motif projeté contient plus qu'une seule primitive, la calibration du projecteur devient très difficile. Cependant, la manière dont nous avons conçu notre motif nous évite d'utiliser cette notion. Le code associé à chaque élément de notre motif ainsi que le fait d'avoir basé le codage sur l'utilisation de matrices parfaites, nous permet de repérer les correspondances entre les éléments extraits dans l'images et les éléments projetés sans ambiguïté.

D'autre part, en matérialisant la projection de chaque primitive du motif par une droite dans l'espace, la position 3D d'un point est calculée en déterminant l'intersection de cette droite dans l'espace avec celle qui matérialise le rayon de vue imageant ce point. Cette façon de procéder conduit inévitablement à des imprécisions. De plus, cela nécessite un étalonnage précis du dispositif de projection. Pour éviter ces problèmes, nous avons choisi de procéder de manière différente. Le principe est de caractériser le comportement de la structure lumineuse dans l'espace. Celui-ci est décrit en générant une sorte d'échantillonnage des positions de la structure lumineuse projetée dans l'espace. À l'aide des données de l'étalonnage, une correspondance est établie entre ces positions et leurs localisations dans l'image. Les coordonnées 3D d'un point sont alors obtenues en interpolant la position du point dans l'image avec les données d'étalonnage. Cette approche évite les calculs coûteux qui de plus peuvent introduire de nouvelles imprécisions.

Cependant, différentes techniques d'étalonnage sont envisageables en fonction de la méthode utilisée. Celles-ci vont être décrites dans la section suivante.

3.2 Étalonnage du dispositif

Comme nous venons de le voir, les données d'étalonnage permettent de procéder à une reconstruction 3D métrique. Celle-ci représente donc une étape primordiale dans le processus de reconstruction. C'est notamment de sa précision que dépendra la précision de la reconstruction finale. Généralement, la calibration d'un système à lumière structurée se fait en deux étapes ; la calibration de la caméra puis celle du projecteur. La première sert à calculer les paramètres de la caméra et la deuxième sert à déterminer la transformation entre le repère du projecteur et celui de la caméra.

3.2.1 Étalonnage de la caméra

La calibration des caméras dans les systèmes de vision par ordinateur est un sujet qui a été traité à des nombreuses reprises. Dans [Zhang04], on trouve un aperçu complet sur les différentes solutions proposées dans la littérature. Bien que certaines mesures de celles-ci peuvent être obtenues en utilisant des caméras non calibrées [Hartley94], l'étalonnage est essentiel pour obtenir des mesures métriques précises. En effet, le but de l'étalonnage est de modéliser la caméra par un modèle mathématique qui décrit son comportement et de calculer les paramètres de ce modèle afin de déterminer le comportement des rayons de vue dans l'espace. Par conséquent, la méthode de calibration dépend du modèle considéré. Le modèle de la caméra peut être

choisi selon la précision désirée [Ito91]. Cependant, dans la plupart des travaux, la caméra est représentée par un modèle classique appelé sténopé (*Pinhole model*) ([Duda73, Faugeras93]) qui approxime au mieux la géométrie réelle de nombreux dispositifs d'imagerie. Le modèle sténopé est illustré par la Figure 3.2 où C est le centre optique de la caméra, $\overrightarrow{CZ_C}$ est son axe optique, (O, u, v) est le système de coordonnées lié au plan image et f est la distance focale.

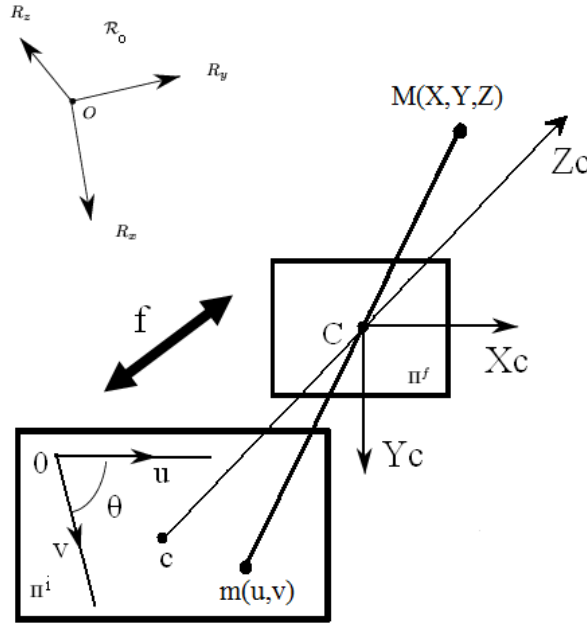


FIGURE 3.2 – Modèle sténopé de la caméra.

La relation entre un point $M(X, Y, Z)$ de l'espace, exprimé dans un repère global R_o , et sa projection $m(u, v)$ dans le plan image est donnée par l'équation 3.5.

$$s \begin{bmatrix} m(u, v) \\ 1 \end{bmatrix} = A_c P_0 \mathfrak{R}_c \begin{bmatrix} M(X, Y, Z) \\ 1 \end{bmatrix} \quad (3.5)$$

Dans cette équation, s représente un facteur d'échelle arbitraire et A_c la matrice 3×3 qui contient les paramètres intrinsèques de la caméra. P_0 une projection centrale standard (Eq. 3.6) et \mathfrak{R}_c est la matrice des paramètres extrinsèques qui décrit la position de la caméra dans l'espace et contient la transformation entre le repère global et le repère-caméra (Eq. 3.7) représentée par la matrice de rotation \mathfrak{R} et le vecteur de translation entre les deux repères t :

$$P_0 = (I_3 \quad | \quad 0) \quad (3.6)$$

$$\mathfrak{R}_c = \begin{pmatrix} \mathfrak{R} & t \\ 0 & 1 \end{pmatrix} \quad (3.7)$$

La matrice A_c est décrite par l'équation (Eq. 3.8). Les paramètres α et β sont les facteurs d'échelle des coordonnées dans le plan image selon les axes u et v et ils dépendent du capteur de la caméra et de la distance focale f . Les coordonnées (u_0, v_0) représentent l'image du centre optique

(le point principal) alors que γ dépend de l'angle θ entre les axes u et v (voir Figure 3.2). Ces paramètres sont les paramètres propres de la caméra et ils ne dépendent pas de sa position dans l'espace. Ils permettent de transformer les coordonnées d'un point qui sont connues dans le repère de la caméra dans le plan image.

$$A_c = \begin{pmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.8)$$

Généralement, les méthodes de calibration peuvent être distinguées selon que l'estimation des paramètres est linéaires ou non [Ito91]. Les méthodes linéaires décrivent le comportement de la caméra par des équations linéaires qui lient les coordonnées homogènes des points 3D de la scène et leurs projections 2D dans l'image (par exemple [Hall82] et [Faugeras86]). Cependant, ces modèles ne prennent pas en compte les distorsions des lentilles et par conséquent, les mesures sont moins précises. Par contre, les méthodes non linéaires peuvent prendre en compte ces distorsions (par exemple [Tsai87] et [Weng92]).

Généralement, l'étalonnage nécessite un objet étalon dont les dimensions sont connues avec une très grande précision. Cette référence peut être plane ou non. Les méthodes qui utilisent des objets étalons plans sont basées sur le fait que tous les points 3D connus appartiennent au même plan (un damier noir et blanc par exemple) (Figure 3.3). Par conséquent, l'erreur en utilisant ces méthodes est réduite puisqu'une des coordonnées est constante. Les méthodes non planaires cherchent à déterminer la relation entre les différents plans d'une scène et pour cela, elles utilisent généralement un dièdre ou un cube [Heikkila97]. D'autres méthodes se basent quant à elles sur les acquisitions d'un motif plan à des positions différentes dans l'espace (par exemple : [Zhang00] et [Doignon99]) sans avoir une relation entre les plans de la scène.

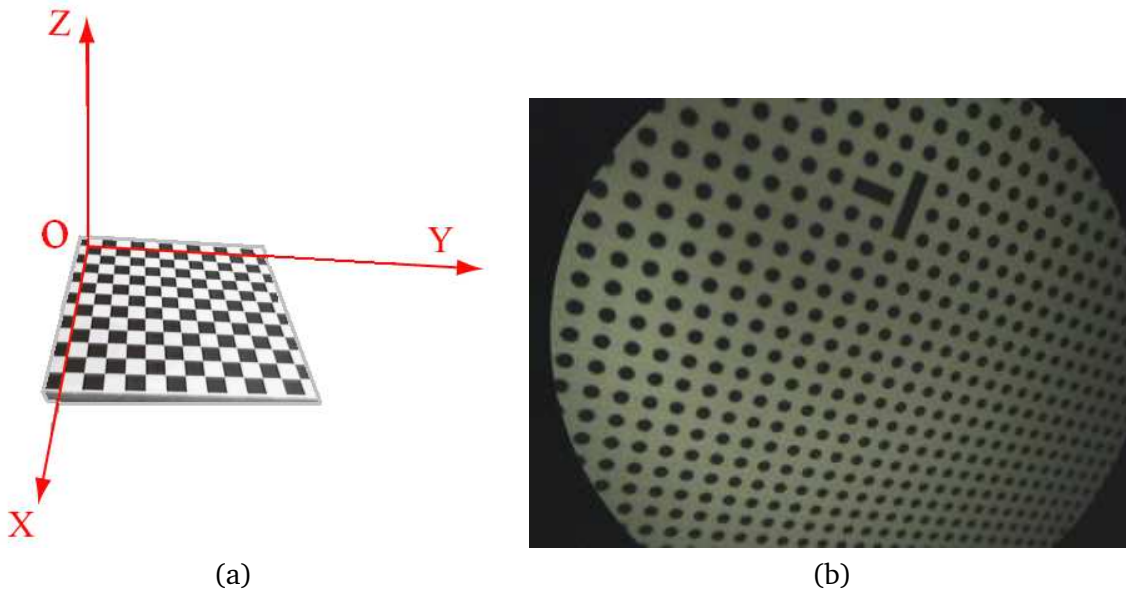


FIGURE 3.3 – Exemples de motifs plans utilisés pour la calibration de caméras (a) un damier noir et blanc [Zhang00]. (b) une matrice des points [Wengert06].

Généralement, la calibration peut être vue comme un processus qui sert à déterminer quatre transformations géométriques [Salvi02] : la matrice \mathfrak{R}_c de la transformation entre le repère de l'espace et le repère de la caméra, la projection du repère caméra sur le plan image, la modélisation la distorsion des lentilles et la dernière est la transformation qui convertit les résultats métriques en pixels dans le repère image. Ces quatre transformations sont illustrées par la Figure 3.4.

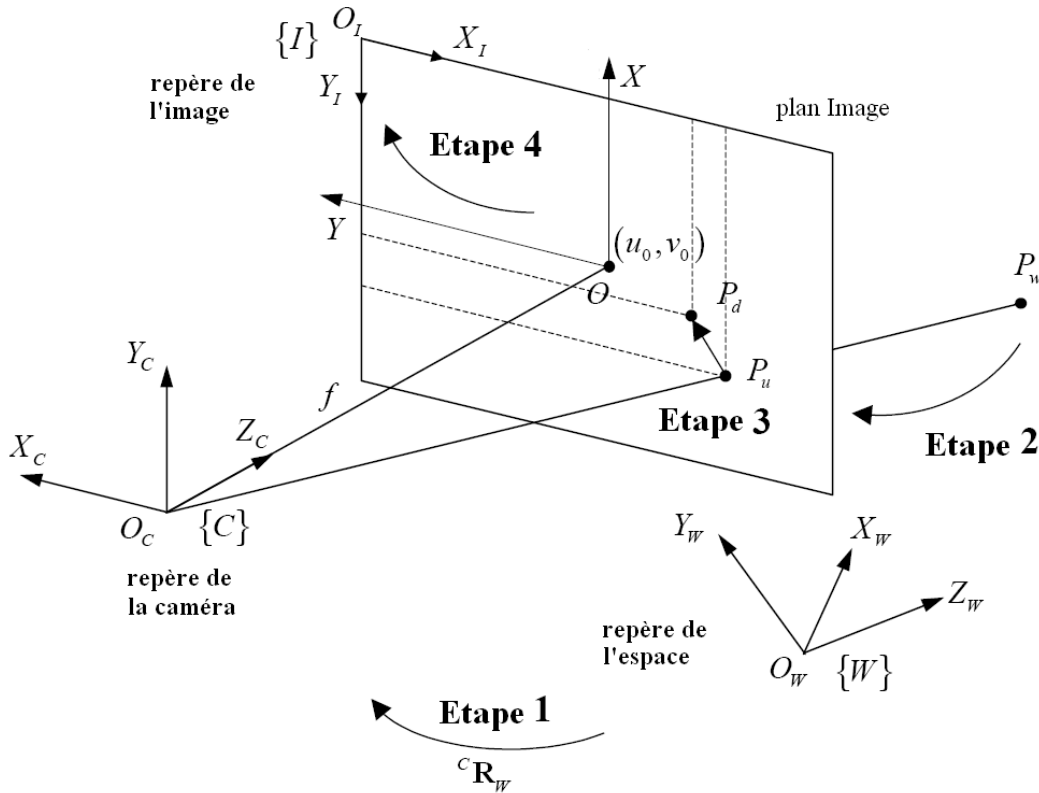


FIGURE 3.4 – Les quatre transformations opérées lors de l'étalonnage d'une caméra (source : [Salvi02]).

Dans [Salvi02], les auteurs détaillent les techniques les plus utilisées pour la calibration d'une caméra et présentent une étude comparative dans laquelle ces techniques sont testées et leurs précisions respectives sont évaluées. Comme on pouvait s'y attendre, les méthodes qui prennent en compte la distorsion des lentilles dans la modélisation de la caméra aboutissent à de meilleures précisions que celles qui ne rendent pas compte de la distorsion. C'est donc vers une de ces méthodes que nous avons orienté notre choix.

3.2.2 Étalonnage du projecteur

La deuxième étape de l'étalonnage d'un système à lumière structurée consiste dans le cas général, à calibrer le projecteur. Ce problème peut être traité de même manière que la calibration de la caméra présentée dans la section précédente. Les paramètres intrinsèques du projecteur représentent les caractéristiques du motif considéré alors que les paramètres extrinsèques du

projecteur peuvent être calculés à l'aide des résultats de la calibration de la caméra. On peut distinguer les projections d'une seule primitive (bande ou point lumineux) et les projections de motifs contenant plusieurs primitives [Ribo05]. Dans le premier cas, la transformation entre le repère du projecteur et le repère de la caméra doit être déterminée, et cela consiste à calculer la translation et les trois angles d'Euler entre les deux repères. Par contre, dans le deuxième cas, chaque primitive doit être étalonnée séparément de manière à déterminer pour chacune d'entre elles la position et l'orientation du plan de projection dans le repère de la caméra.

Dans [Chen87], la projection d'une bande de lumière dans le balayage de la surface avec une lumière structurée pour la vision 3D d'un robot est modélisée par un plan qui contient la bande et le centre optique du projecteur (Figure 3.5). L'équation de ce plan dans la repère caméra est donnée par l'équation suivante : (Eq. 3.9) :

$$ax_c + by_c + cz_c + d = 0 \quad (3.9)$$

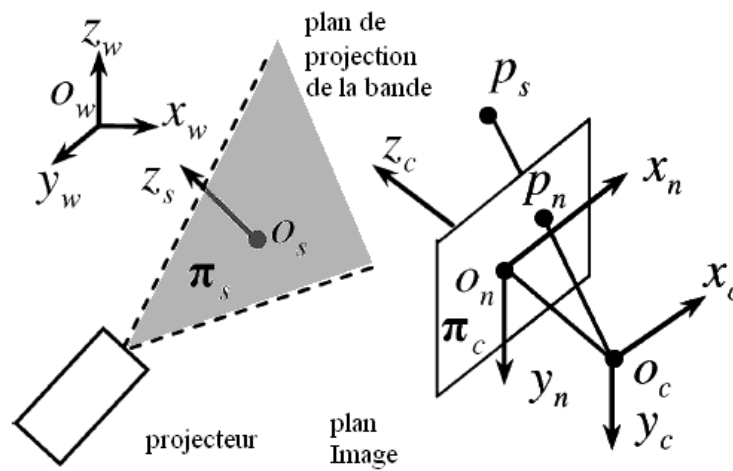


FIGURE 3.5 – Modèle du projecteur proposé dans [Chen87] et [Zhou04].

Dans ce cas, l'étalonnage consiste à calculer la matrice de transformation 4×3 qui permet de convertir les points dans le plan image appartenant à la projection de la bande dans le repère de la caméra. Cette matrice lie les coordonnées (x, y, z) d'un point dans le repère de la caméra avec sa projection (u, v) dans l'image avec un facteur d'échelle ρ par l'équation suivante (Eq. 3.10) :

$$\rho \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3.10)$$

Les auteurs ont proposé de calculer cette matrice à l'aide d'un ensemble de lignes non colinéaires matérialisées par les bords des objets fabriqués précisément que l'on dispose dans la scène. L'approche de correspondance entre des lignes dans l'espace et des points dans l'image a été étendue dans [Reid96] à une correspondance entre des plans dans l'espace et des points dans l'image. Cependant, la distorsion n'a pas été prise en compte et la précision diminue au fur et à mesure qu'on s'approche de l'objet à balayer. De plus, si en pratique il est difficile de disposer dans la

scène d'un grand nombre de points dont les coordonnées sont connues afin d'étalonner le projecteur d'une manière précise, il est également difficile d'assurer une précision suffisante de la localisation des points de l'espace qui appartiennent à la projection d'une bande.

Afin de résoudre ces problèmes, une nouvelle approche robuste a été présentée par [Huynh99]. Elle met à profit de l'invariance projective du birapport. Si p, q, r et m sont quatre points alignés sur une droite D , on appelle birapport de ces quatre alignés dans cet ordre, la quantité b (généralement désignée par (p, q, r, m)) définie comme suit (Eq. 3.11) :

$$b = \left(\frac{\theta_p - \theta_r}{\theta_q - \theta_r} \right) \setminus \left(\frac{\theta_p - \theta_m}{\theta_q - \theta_m} \right) \quad (3.11)$$

Il a été démontré [Semple52] que le birapport est conservé par une projection centrale. En d'autres termes, si p, q, r et m sont quatre points alignés sur une droite D et O est un point extérieur à D , et si D' est une droite qui coupe les droites Op, Oq, Or et Om respectivement en P, Q, R et M (voir Figure 3.6), alors $(p, q, r, m) = (P, Q, R, M)$. Par conséquent, la connaissance des projections de trois points et du birapport nous fournit précisément la projection du quatrième point (). Cependant, la configuration des points doit être bien choisie afin d'éviter les calculs imprécis du birapport ([Forsyth91], [Maybank92] et [Morin93]).

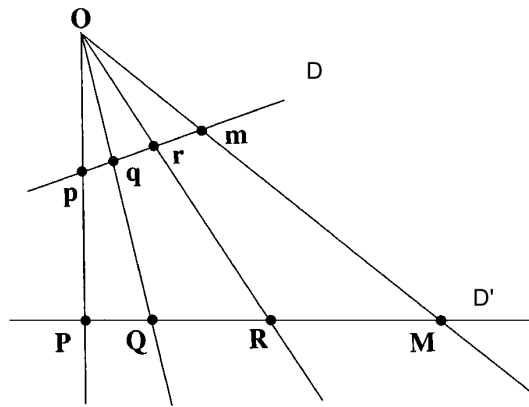


FIGURE 3.6 – Invariance projective du birapport, ici on a $(p, q, r, m) = (P, Q, R, M)$.

C'est cette propriété qui a été utilisée dans [Huynh99]. Les auteurs ont proposé d'utiliser 4 droites non coplanaires, chaque droite étant définie par 3 points colinéaires afin de calculer les coordonnées dans l'espace des points qui appartiennent aux bandes projetées, dans le but de déterminer la matrice de transformation (Eq.3.10) pour chaque bande lumineuse. Donc, en connaissant les points $\{P_i, Q_i, R_i\}, i = 1..4$ (Figure 3.7) colinéaires dans la scène, et en calculant les points $\{m_i\}$ qui correspondent à l'intersection des lignes formées par les trois points et la projection d'une bande dans le plan image, on peut calculer les points $\{M_i\}$ qui correspondent à la projection de la bande dans l'espace à l'aide de la méthode de birapport. En utilisant cette approche, on arrive à localiser de façon précise la projection de la bande. Pour évaluer cette méthode, les auteurs ont étalonné leur dispositif et se sont servis de ces résultats pour obtenir la reconstruction 3D d'objets étalons. Cependant, l'écart moyen et standard d'erreurs de reconstruction étaient, respectivement, 0.7240 mm et 0.4447 mm.

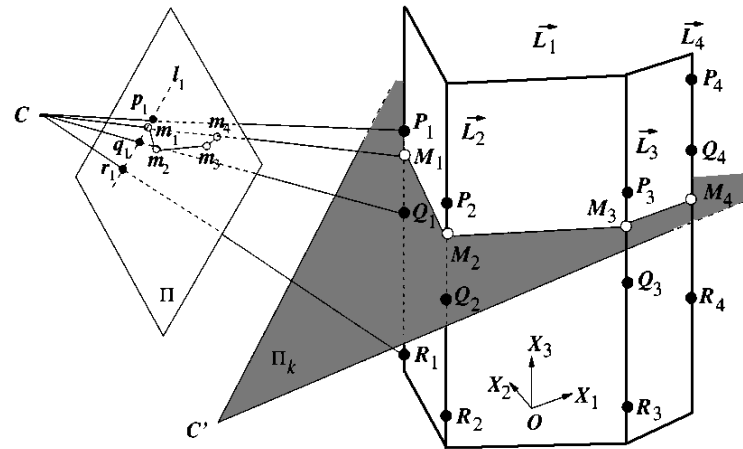


FIGURE 3.7 – Calibration à l'aide de 4 ensembles non coplanaires connus de 3 points colinéaires proposée dans [Huynh99].

Contrairement à cette méthode qui nécessite un objet 3D précis d'étalonnage, une méthode a été développée pour déterminer les paramètres du plan de projection d'une bande lumineuse sans avoir de connaissances particulières de la géométrie 3D de la scène [Zhou04]. La méthode exige que la caméra observe la projection de la bande sur une surface plane donnée et pour des orientations différentes (au moins deux). La nouvelle approche proposée est de générer suffisamment de points non-colinéaires de projection de la bande à l'aide d'un motif plan de calibration (Figure 3.8) afin de bien déterminer les paramètres (a, b, c, d) du plan de projection (Eq. 3.9). De plus, la méthode de birapport est utilisée pour obtenir des mesures précises. Les résultats expérimentaux présentés montrent que la méthode est robuste et présente l'avantage de pouvoir utiliser le même motif d'étalonnage pour calibrer la caméra et le projecteur. Les racines carrées des moyennes des erreurs de reconstruction selon X, Y et Z sont, respectivement, 0.059 mm, 0.103 mm et 0.255 mm [Zhou04]. Cette méthode est adaptée à la projection d'une bande lumineuse mais peut aussi être appliquée aux motifs lumineux contenant plusieurs bandes.

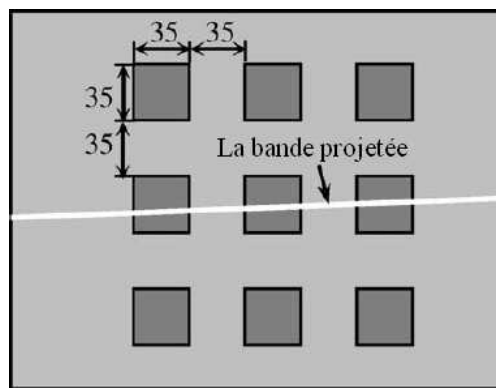


FIGURE 3.8 – Bande projetée sur le motif plan de calibration.

Dans [Chen02], les auteurs proposent une méthode d'auto-étalonnage avec une projection d'une grille de bandes encodée avec des niveaux de gris. La méthode suppose que la caméra ainsi que les paramètres intrinsèques du projecteur sont déjà calibrés. Par conséquent, on peut démontrer que les bandes peuvent être étalonnées avec une seule vue. De plus, l'étalonnage peut être refait automatiquement au cas où la position entre la caméra et le projecteur change. La méthode a été validée sur un système à deux degrés de liberté. L'avantage de cette méthode est qu'elle n'a besoin ni de dispositif d'étalonnage précis ni de connaissances préalables du mouvement de la caméra ou de la scène. De plus, la méthode a été validée dans [Chen03] et [Chen07] avec une projection d'un motif encodé par des couleurs (Figure 3.9) et la précision de la reconstruction 3D d'objets étalons a été d'environ 1%.



FIGURE 3.9 – Projection du motif encodé par des couleurs sur des objets étalons (source : [Chen03]).

Comme nous l'avons déjà mentionné, lorsque le motif projeté est plus complexe qu'un point ou une bande, une autre approche consiste à caractériser le comportement de projections des primitives dans l'espace. Le principe général consiste à projeter le motif sur une surface plane et à faire plusieurs acquisitions à des distances différentes et à déterminer les positions 3D des primitives dans chaque plan. On construit ainsi une sorte d'échantillonnage de l'espace matérialisé par différentes projections de primitives. Cela a été utilisé par [Graebbling02], [Marzani02] ou encore [Dipanda05] avec des systèmes assez similaires. Ce type de méthode donne de bons résultats, mais elle nécessite des déplacements précis qui font parfois appel à des dispositifs mécaniques. Dans [Marzani02], ce genre de méthodes a été employé pour étalonner un système de vision active qui projette une grille de points lumineux (361 points). L'espace de travail était compris entre 10 mm et 90 mm. La reconstruction 3D de plusieurs plans situés dans l'espace de travail a été évaluée. Une erreur moyenne de 0.115 mm a été mesurée pour la profondeur avec un écart moyen de 0.029 mm.

Dans cet ordre d'idées, les travaux présentés dans [Zhang06] sont parmi les travaux les plus intéressants. Ici, les auteurs proposent une nouvelle méthode d'étalonnage pour les systèmes de lumière structurée en utilisant des projecteurs DMD (Digital Micromirror Device). Ces pro-

jecteurs utilisent la technologie DLP (Digital Light Projective) qui se base sur la réflexion de la lumière sur des centaines de milliers de micro-miroirs mobiles (minimum 500 000) disposés en lignes et en colonnes sur une matrice DMD. Le concept clé dans cette méthode d'étalonnage est de traiter le projecteur comme une caméra et de calibrer la caméra et le projecteur indépendamment en utilisant la méthode traditionnelle d'étalonnage des caméras. Afin de traiter le projecteur comme une caméra, ils ont développé une nouvelle méthode qui permet au projecteur de "capturer" des images. Cela est fait en construisant une cartographie entre les pixels de la caméra CCD et ceux du projecteur DMD. En d'autres termes, l'image de la caméra CCD peut être transformée pixel par pixel pour former une image que l'on appelle l'image du projecteur DMD. Avec ce concept, l'étalonnage des systèmes à lumière structurée devient essentiellement le même que celui des systèmes traditionnels de stéréovision qui est déjà bien établi. Dans ce cas, il faut tenir compte de la contrainte épipolaire. L'erreur maximale mesurée de la reconstruction 3D est de 0,22 mm dans un espace de travail de volume de $342 \times 376 \times 658$ mm. En plus, cette méthode est rapide, robuste et précise. Une autre idée intéressante présentée dans ce travail est l'utilisation d'un damier rouge et bleu au lieu du damier classique noir et blanc (Figure 3.10-a). En effet, si le damier est éclairé par une lumière rouge, son image acquise avec la caméra à niveau de gris est la même que celle d'un damier classique noir et blanc (Figure 3.10-b) avec lequel on peut étalonner la caméra selon la méthode proposée dans [Zhang00]. Par contre, si l'éclairage est réalisé avec une lumière blanche, l'image du damier est un plan uniforme gris (Figure 3.10-c) car le capteur d'une caméra en niveau de gris a à peu près la même réponse pour le rouge et le bleu ce qui facilitera la détection du motif projeté sur le plan du damier au cours de la phase de l'étalonnage du projecteur.

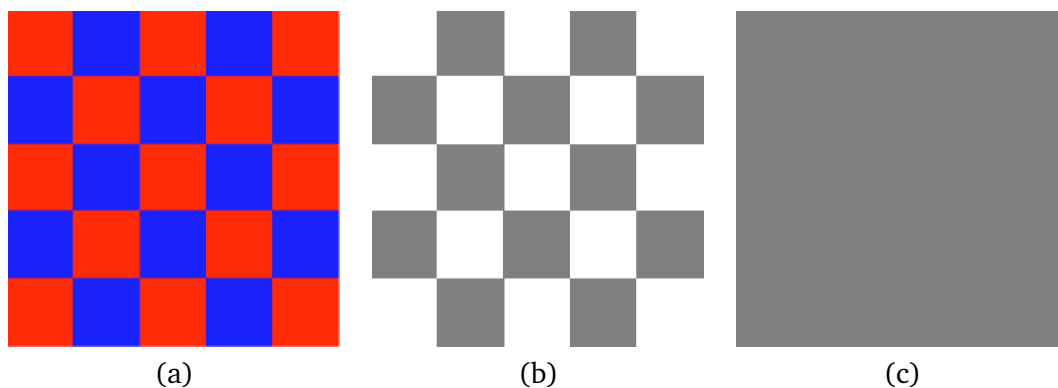


FIGURE 3.10 – (a) Damier utilisé pour l'étalonnage dans [Zhang06]. (b) Son image obtenue avec une caméra en niveau de gris avec un éclairage rouge (c) et avec un éclairage blanc.

Nous constatons que la méthode employée pour l'étalonnage du projecteur dépend essentiellement du type de motif projeté. De plus, si le motif contient plusieurs primitives, il semblerait que décrire le comportement de la projection de chaque primitive dans l'espace sous forme d'échantillonnage de l'espace de mesures à une description mathématique de ce comportement. C'est une méthode de ce genre que nous avons développée.

3.2.3 Méthode d'étalonnage proposée

Le motif que nous avons conçu est composé de 783 éléments organisés sous forme de matrice de dimensions 27×29 (Figure 2.10). L'objectif de notre étalonnage est d'obtenir la description du comportement 3D des projections de ces éléments dans l'espace. La méthode que nous proposons nécessite un dispositif d'étalonnage pour lequel les points 3D appartiennent à un même plan afin de réduire l'erreur. Nous nous sommes inspirés des travaux de [Zhang06] décrits dans le paragraphe précédent. Par contre, nous n'avons pas besoin de modifier l'éclairage durant notre procédure d'étalonnage. En effet, le motif que nous utilisons est monochrome et l'algorithme de segmentation mis en œuvre est capable de ne détecter que les éléments du motif projeté dans l'image même si le support de projection n'est pas uniforme. De plus, nous avons choisi un damier gris et blanc comme un dispositif d'étalonnage afin de faciliter la détection du motif quand il sera projeté sur le plan du damier.

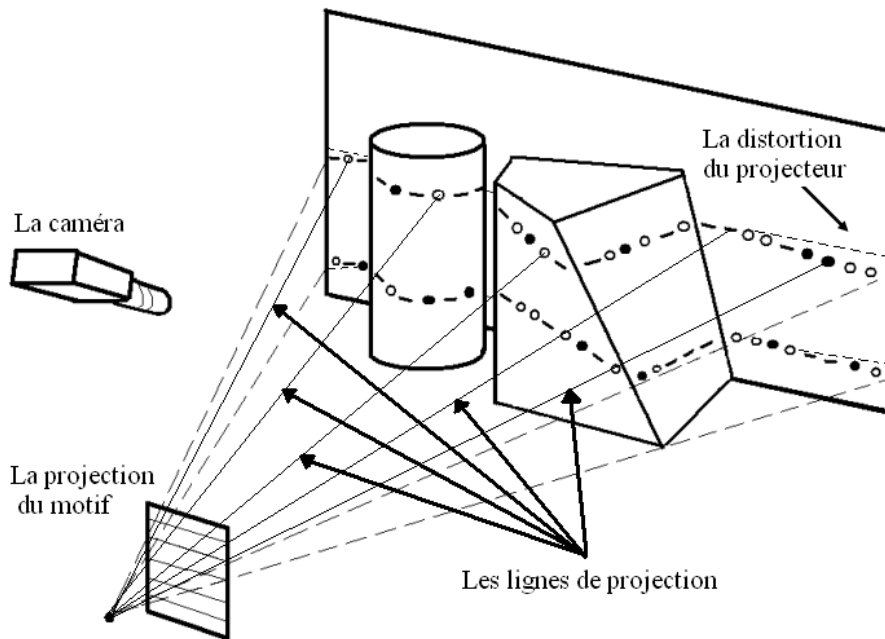


FIGURE 3.11 – Les lignes de projection des primitives sont considérées au lieu des plans.

La méthode que nous proposons comporte deux étapes principales. Dans un premier temps, la caméra est calibrée afin de calculer ses paramètres intrinsèques et extrinsèques. Ensuite, ces paramètres seront utilisés pour étalonner l'ensemble du dispositif. Pour calibrer la caméra, nous avons adopté la technique proposée dans [Zhang00] qui se base sur l'acquisition des images d'un damier à des positions différentes dans la scène. Par conséquent, le plan du damier pourra être reconstruit pour chaque position d'acquisition, dans le repère de la caméra. Le principe de notre méthode est de projeter notre motif sur le plan du damier pour chaque acquisition faite durant le processus d'étalonnage de la caméra. Par conséquent, les éléments du motif projeté appartiennent donc au plan du damier. On peut alors se servir de cette propriété pour décrire les comportements des projections des primitives en caractérisant chacune d'elles par un ensemble de points 3D. Dans notre cas, nous avons considéré des lignes de projection des primitives et par conséquent, le dispositif peut être étaloné d'une manière plus précise puisque nous évitons la distorsion causée par les lentilles du projecteur (Figure 3.11).

À chaque position du damier, deux images sont acquises ; la première ne contient que le damier (projecteur éteint) et la deuxième contient la projection du motif sur le plan du damier. Les images du damier sont utilisées pour calibrer la caméra selon la méthode présentée dans [Zhang00] afin de déterminer ses paramètres intrinsèques et extrinsèques ainsi que la distorsion de ses lentilles. À ce stade, nous disposons pour chaque position du damier, d'un ensemble de points (X_c, Y_c, Z_c) correspondant aux estimations de la position dans l'espace des points du plan du damier considérés dans la calibration (les coins des carrés). Puisque cette méthode prend en compte la distorsion des lentilles, on peut donc estimer les paramètres du plan du damier en fonction de (x, y, Z_c) où (x, y) sont les projections non déformées des points (X_c, Y_c, Z_c) dans le plan image au lieu de les estimer en fonction des coordonnées (X_c, Y_c, Z_c) . Cela consiste à estimer, pour chaque position du damier et à l'aide des n points reconstruits à cette position, les paramètres (a, b, c, d) du plan de damier vérifiant l'équation suivante (Eq. 3.12) :

$$\begin{bmatrix} x_1 & y_1 & Z_{c1} & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_n & y_n & Z_{cn} & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0 \quad (3.12)$$

La deuxième étape sert à traiter les images prises lors de la projection du motif en utilisant les algorithmes de traitement associés à ce motif (voir section 2.2.4). Par conséquent, une cartographie est construite entre chaque position du damier et les coordonnées (x_p, y_p) en pixels de chaque primitive dans le plan image. Cette cartographie servira, lors de la reconstruction, à calculer la profondeur de chaque primitive à chaque position du damier en utilisant les paramètres (a, b, c, d) estimés dans l'étape précédente et la contrainte de coplanarité entre le motif et le damier. Donc, nous obtenons pour chaque primitive, des coordonnées (x_p, y_p, Z_c) correspondant aux différentes positions du damier.

La dernière étape comporte l'estimation, et pour chaque primitive, la projection dans le plan (O, y, Z_c) de la ligne droite 3D qui passe par ses points $\{(y_p, Z_c)\}$, évitant ainsi l'emploi des équations définissant une droite 3D dans l'espace. Par conséquent, la profondeur Z_c et la coordonnée y dans l'image de la primitive d'ordre i dans la matrice seront liées par l'équation suivante (Eq. 3.13) :

$$Z_c = a_i \cdot y + b_i \quad (3.13)$$

L'ensemble des paramètres $\{a_i, b_i\}_{i=1..783}$ représente les données d'étalonnage du dispositif. Par conséquent, afin de reconstruire une surface sur laquelle notre motif est projeté, l'image de la projection sera traitée par les algorithmes de segmentation et de décodage afin de localiser les primitives dans l'image et de trouver leurs correspondances dans la matrice du motif. Chaque primitive identifiée et localisée en position (x_p, y_p) dans l'image est projetée sur la ligne de projection correspondant pour calculer sa profondeur Z_c . Finalement, les paramètres intrinsèques de la caméra sont introduits pour estimer la position 3D (X_c, Y_c, Z_c) de chaque primitive dans le repère de la caméra.

3.2.4 Évaluation de la méthode d'étalonnage proposée

Le système d'évaluation considéré consiste en une caméra digitale (AVT Pike F100C) et en un vidéo-projecteur de type Sony VPL-CS6 (Figure 3.12). La caméra a un capteur de taille (1024×1024) pixels et elle permet d'acquérir des images au *format7*. Avec ce format, on peut choisir le point d'origine de l'image, la taille de l'image ainsi que le mode d'encodage des couleurs. En pratique, nous avons choisi une région d'intérêt de taille (512×512) pixels située au milieu du capteur. Le damier utilisé pour la calibration contient (12×15) carrés gris et blancs de 10mm de coté. La surface du damier est d'environ (15×18) cm et elle représente la zone de projection du motif. Afin d'étalonner le système selon la méthode présentée dans la section précédente, nous avons considéré 14 positions différentes du damier dans la scène. La projection du motif est alternée afin d'avoir, à chaque position, deux images, une du damier et l'autre de la projection du motif codé sur la surface du damier (Figure 3.13).

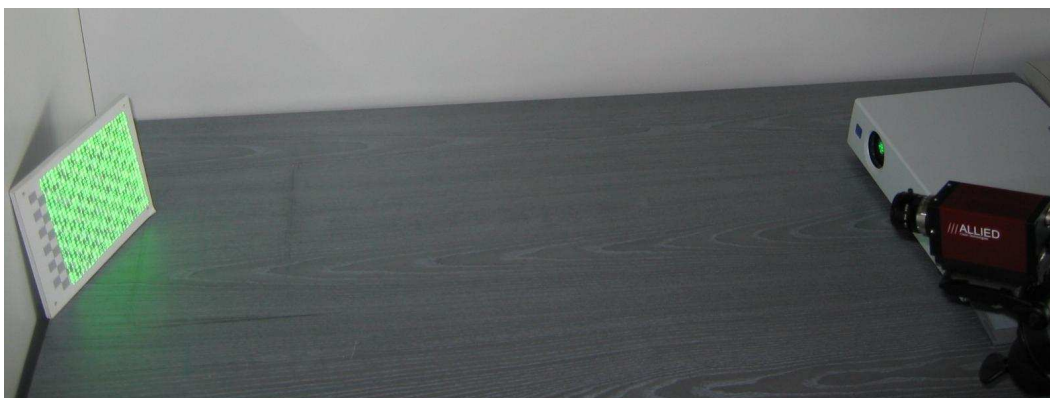


FIGURE 3.12 – Dispositif d'étalonnage. La caméra capture deux images pour chaque position du damier, une pour le damier (projecteur éteint) et l'autre de la projection du motif codé sur la surface du damier.

L'étalonnage de la caméra est obtenu en utilisant les images ne contenant que le damier et à l'aide de la librairie (*Camera Calibration Toolbox*²) implémentée dans l'environnement de *Mat-Lab*. Afin d'évaluer la précision de l'étalonnage, la re-projection des données sur le plan image est déterminée et la distance avec les données de départ est calculée. Les erreurs de la re-projection en pixels dans chaque image sont présentées dans la Figure 3.14. L'erreur moyenne mesurée selon (x, y) en pixels est de $(0.15875, 0.14521)$. La distorsion est illustrée sur la Figure 3.15. Nous remarquons qu'elle est faible dans la zone que nous avons considérée et que ses effets sont limités aux bords de l'image. Compte tenu de ces informations, il est possible de reconstruire le plan du damier dans le repère de la caméra à chaque position (Figure 3.16). Ces résultats sont utilisés pour estimer les équations de ces plans afin de les utiliser dans l'étape finale de l'étalonnage.

Les images des projections du motif sont traitées afin de détecter et de décoder les primitives dans chaque image (Figure 3.18). Ensuite, les coordonnées (x_p, y_p) des primitives sont projetées sur les plans estimés à la fin de l'étape de la calibration de la caméra. Cela nous permet de calculer la profondeur Z_c de chaque primitive à chaque acquisition.

2. http://www.vision.caltech.edu/bouguetj/calib_doc/

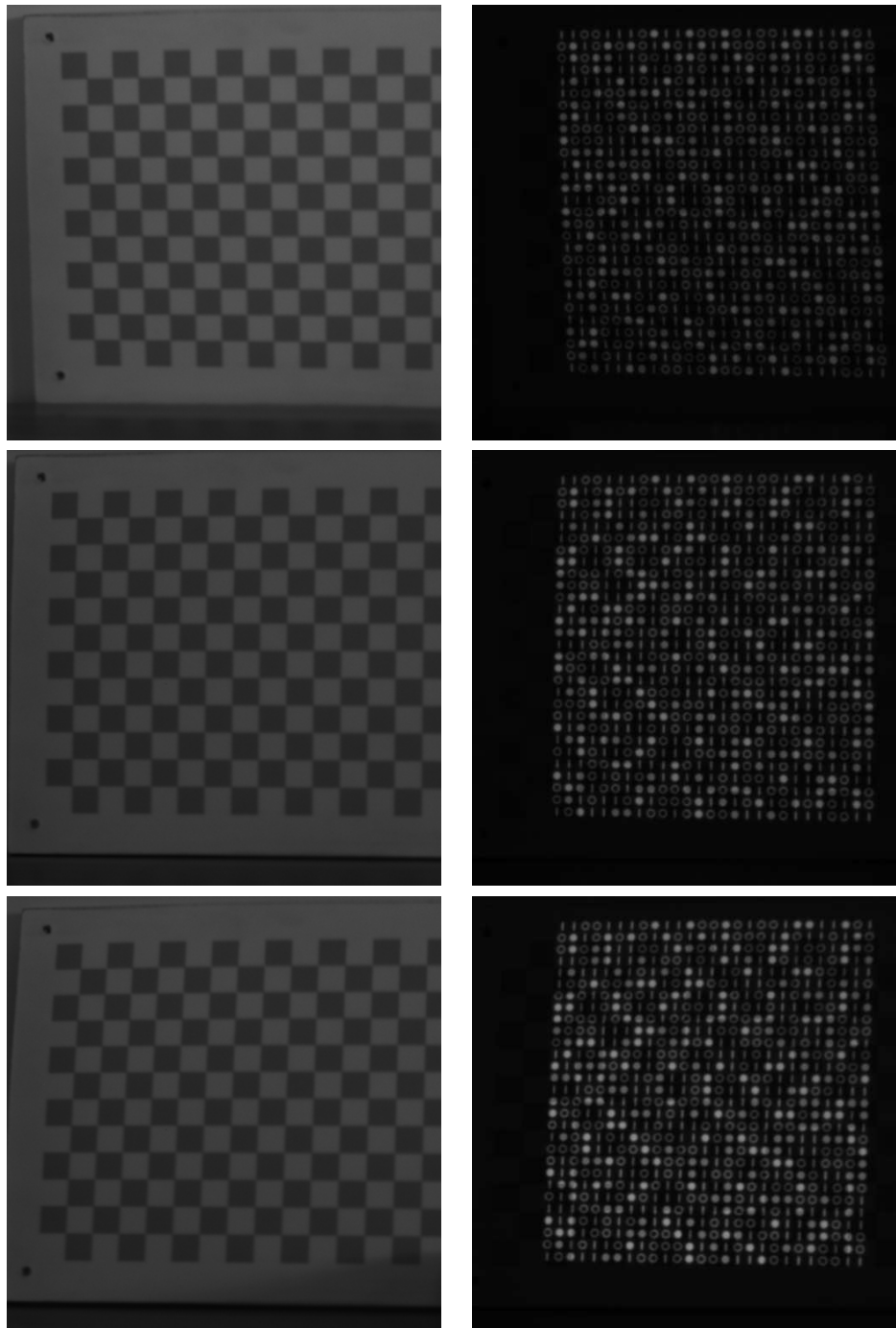


FIGURE 3.13 – Paires d’images acquises (à gauche le projecteur est éteint et à droite il est allumé).

Finalement, une estimation robuste est utilisée pour établir une relation entre la profondeur de chaque primitive et sa position dans le plan image. Dans la Figure 3.17, les cercles rouges correspondent aux points (y_p, Z_c) calculées pour chaque primitive et pour chaque acquisition et les lignes bleues sont l’estimation, pour chaque primitive, de la ligne qui passe par ces points. Ces lignes représentent la projection des lignes de projection des primitives dans le plan (y, Z_c) .

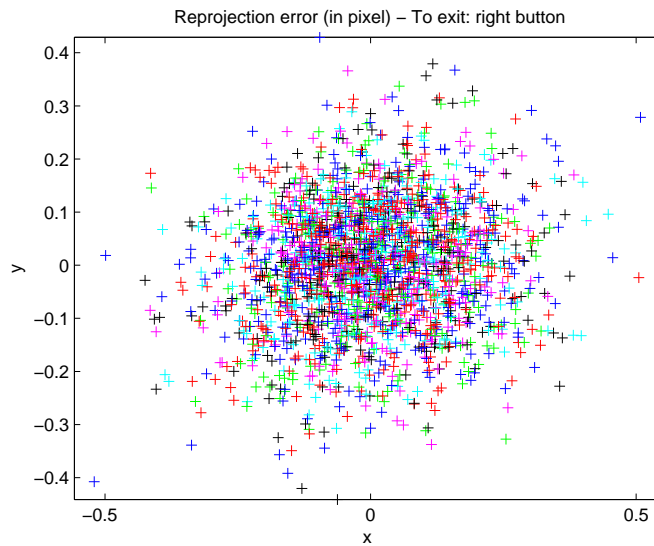


FIGURE 3.14 – Les erreurs de la re-projection des pixels pour les images acquises.

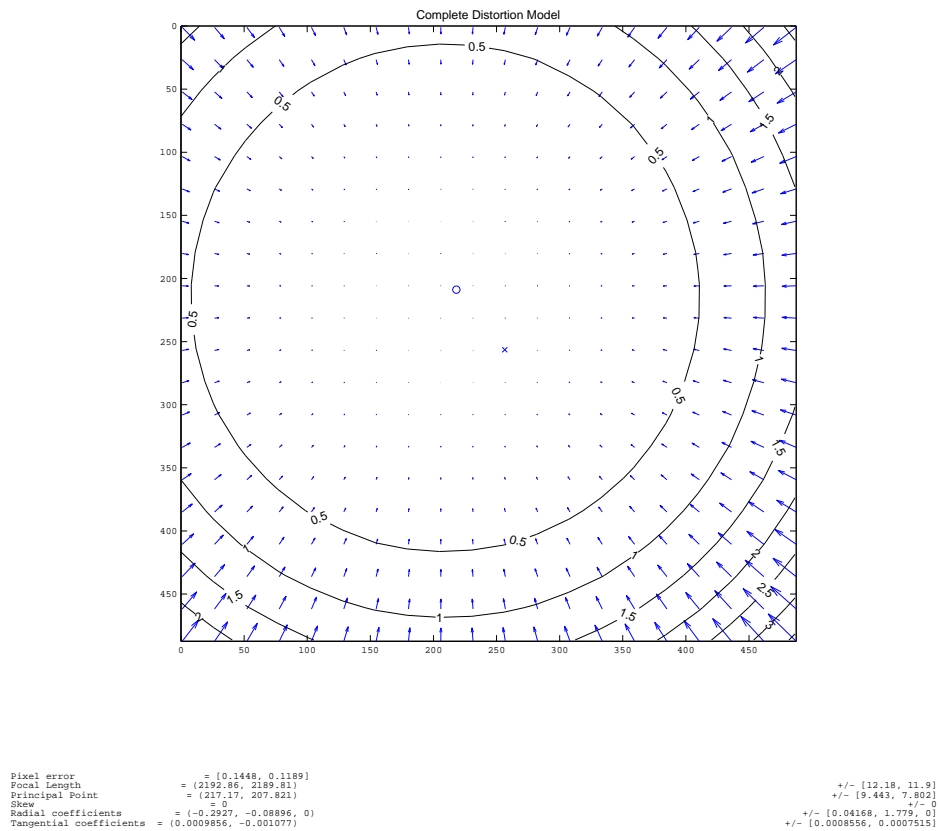


FIGURE 3.15 – Distorsion des lentilles de la caméra.

Ces résultats représentent, avec les paramètres intrinsèques de la caméra, les données de calibration du système qui seront utilisées plus tard pour obtenir la reconstruction 3D métrique.

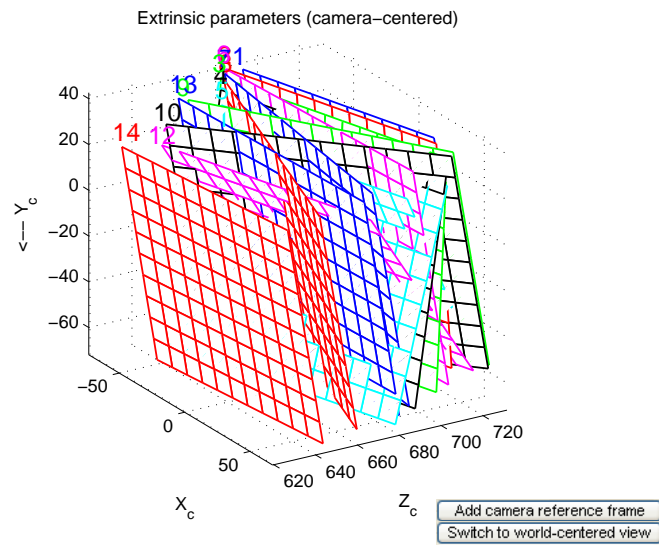


FIGURE 3.16 – Reconstruction du plan de damier à différentes positions.

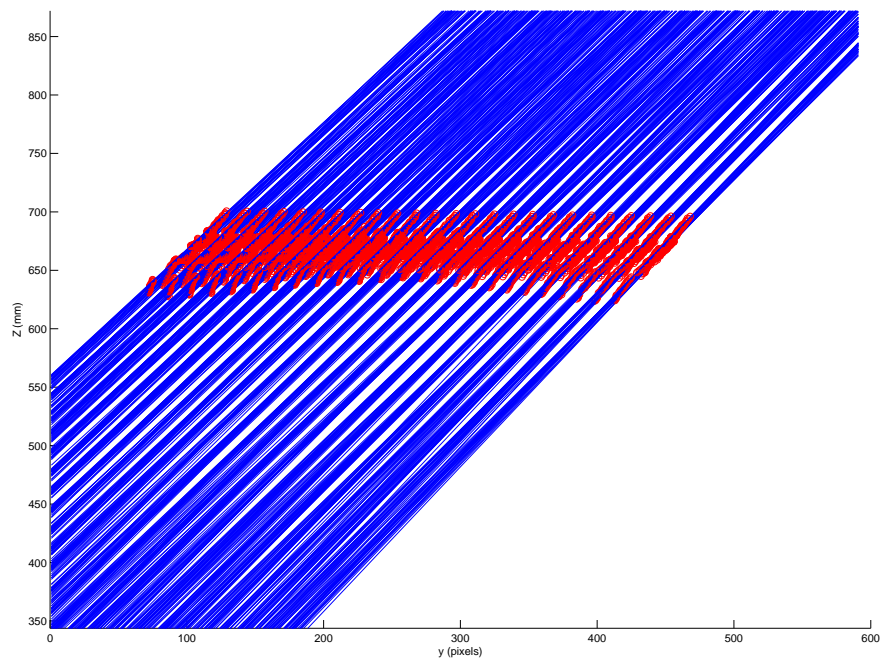


FIGURE 3.17 – Droites lient la profondeur Z_c et la coordonnée y de chaque primitive dans l'image.

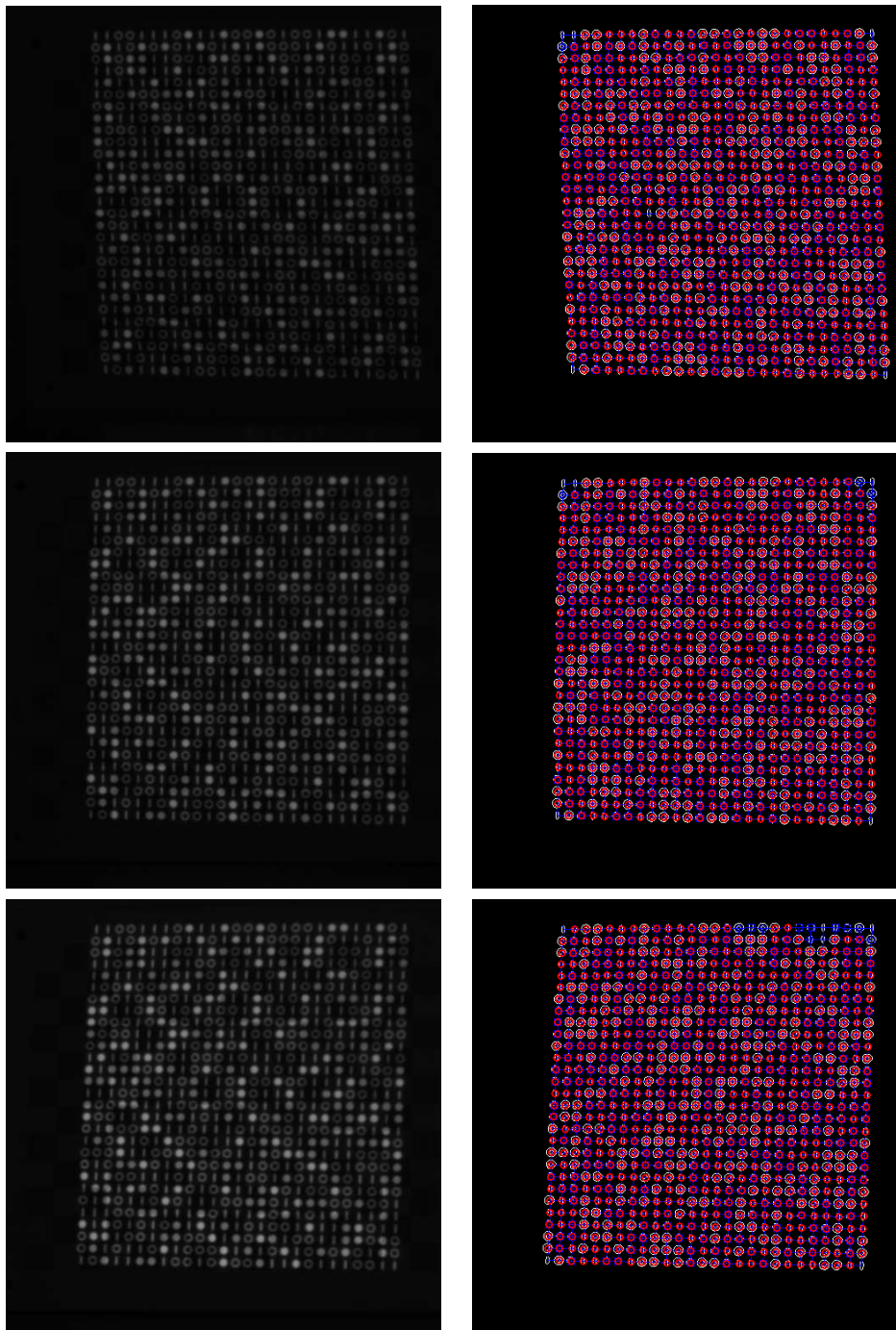


FIGURE 3.18 – Traitement des premières six images acquises de la projection du motif (les points rouges représentent les primitives décodées correctement).

3.3 Évaluation de la reconstruction métrique

Afin d'évaluer la précision de la reconstruction obtenue avec notre système, nous avons proposé de projeter le motif sur trois surfaces planes et parallèles. La distance réelle entre les surfaces a été mesurée d'une manière précise (Figure 5.14-a) et ils ont été placés dans l'espace de travail qui correspond à une zone de profondeurs comprise entre 600 et 800 mm. Ensuite, l'image de la projection a été traitée afin de décoder les primitives comme cela est illustré par la Figure 5.14-b dans laquelle les primitives qui ont été décodées correctement sont représentées par des points rouges. Finalement, les données de la calibration ont été utilisées pour reconstruire les plans (Figure 5.14-c). Plusieurs acquisitions ont été effectuées de cette manière pour des différents points de vue. Les plans sont estimés d'une manière robuste à l'aide de la commande *robustfit* sous *Matlab* (Figure 5.14-d). Les erreurs de reconstruction sont présentées sur la Figure 5.15 afin d'avoir une idée des erreurs moyennes et maximales. D'autre part, pour chaque acquisition, les distances et les orientations entre les plans ont été estimées.

Les résultats sont présentés dans le tableau 5.1 où la première colonne contient les couples de plans considérés, la deuxième colonne présente la distance réelle entre ces deux plans, la troisième colonne présente la distance calculée après la reconstruction et l'estimation des plans, la quatrième colonne présente l'erreur moyenne par rapport à cette estimation et à la distance réelle et la dernière colonne présente l'orientation estimée entre les plans.

Nous remarquons que l'erreur moyenne absolue ne dépasse pas 0.9 mm dans l'intervalle de profondeur considéré et que l'angle entre les plans ne dépasse pas 1° , sachant que les plans sont parallèles. Vu l'espace de travail considéré et en comparaison avec les précisions obtenues pour les autres motifs considérant la stratégie de codage de voisinage spatiale (voir section 3.2.2), ces valeurs de précision sont très intéressantes. La méthode d'étalonnage proposée ne nécessite pas la prise en compte de l'effet de la distorsion des lentilles sur les positions des primitives et elle ne dépend pas de type de projecteur pour calibrer l'ensemble de système. De plus, ces résultats nous paraissent suffisants pour appliquer la reconstruction 3D à l'intérieur de la cavité abdominale pour laquelle une précision inférieure à 1 mm est suffisante.

Surfaces considérées	Distance réelle (mm)	Distance moyenne (mm)	Erreur moyenne (mm)	Angle ($^\circ$)
1-2	16.50	16.27	0.48	0.15
1-2		16.00	0.84	0.53
1-2		16.75	0.75	0.87
1-2		15.98	0.88	0.77
2-3	16.70	16.71	0.21	0.11
2-3		16.58	0.52	0.09
2-3		16.98	0.75	0.74
2-3		16.78	0.45	0.91

TABLE 3.1 – Résultats de la reconstruction 3D obtenus pour trois plans parallèles.

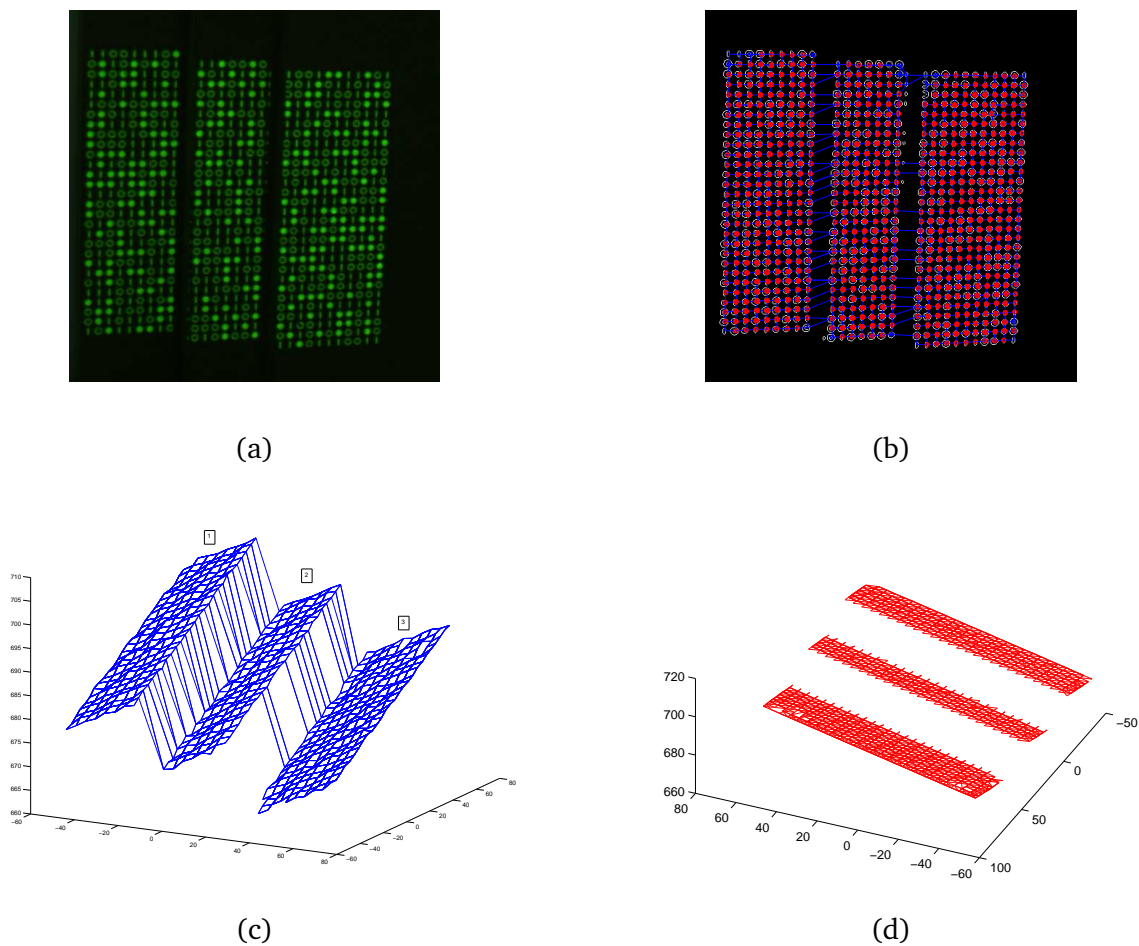
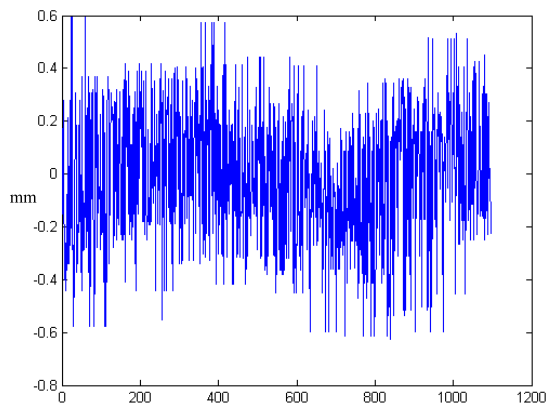


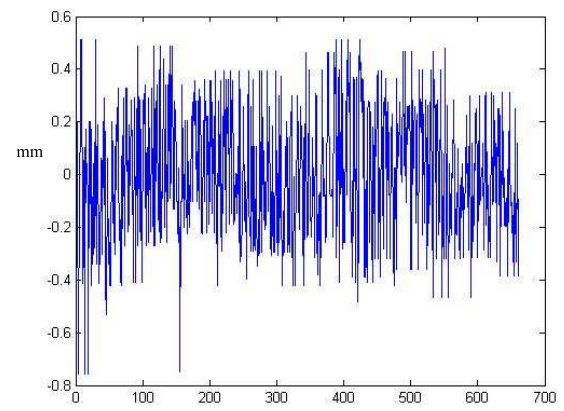
FIGURE 3.19 – (a) Projection sur trois surfaces planes et parallèles. (b) Le traitement de l'image : les points rouges représentent les primitives bien décodées. (c) La reconstruction 3D de la surface. (d) L'estimation des plans à l'aide de *MatLab*.

3.4 Conclusion

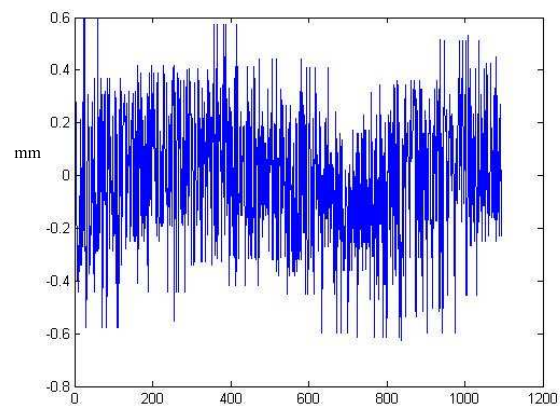
Dans le but d'obtenir la reconstruction 3D métrique, nous avons présenté dans ce chapitre la méthode d'étalonnage que nous proposons pour un système à lumière structurée. Cette méthode fait appel à un dispositif d'étalonnage plan représentant un damier gris et blanc sur lequel le motif codé est également projeté. La méthode de calibration que nous proposons n'est pas sensible aux éventuelles distorsions émanant du projecteur puisque nous considérons les lignes de projection issues de chaque primitive. Finalement, nous avons considéré un système composé d'une caméra et d'un vidéo projecteur afin d'évaluer la précision de la reconstruction 3D obtenue avec la projection de notre motif et la méthode d'étalonnage adoptée. Après avoir étalonné le système selon la méthode proposée, le motif a été projeté sur trois plans parallèles et les distances entre les surfaces ont été mesurées précisément. Après plusieurs acquisitions, l'erreur moyenne en profondeur a été inférieure à 1 mm et l'erreur en orientation a été inférieure à 1° . Ces précisions sont très intéressantes vu que nous avons considéré un intervalle de profondeur de 600 à 800 mm. De plus, elles nous semblent suffisantes pour la reconstruction 3D des surfaces d'organes de la cavité abdominale.



(1)



(2)



(3)

FIGURE 3.20 – Erreurs d'estimation des trois plans après la reconstruction.

Chapitre 4

Conception d'un outil endoscopique de vision 3D

À ce stade de l'étude, nous avons défini une technique de projection de lumière structurée dont l'objectif est de procéder à des reconstructions 3D de scènes intra-abdominales. La technique retenue repose sur le voisinage spatial. À cet effet, un nouveau motif a été défini ainsi que les algorithmes de segmentation et de décodage associés. Afin de mettre en œuvre ces méthodes, nous avons également développé une méthode d'étalonnage et utilisé un principe de reconstruction adopté. L'ensemble a été testé avec des dispositifs expérimentaux de laboratoire. Cependant, l'objectif de notre étude est d'intégrer les points précédents au sein d'un dispositif endoscopique capable de fournir en temps réel, et le plus précisément possible, la reconstruction tridimensionnelle des surfaces abdominales. Ce dispositif est appelé à restituer un nuage de points 3D, chaque point correspondant à la matérialisation de la projection d'un des éléments de notre motif sur la surface à analyser. Ce nuage de points pourra ensuite être utilisé, en association avec des techniques de réalité augmentée, pour restituer visuellement à un chirurgien, la profondeur de l'environnement dans lequel il opère. Il pourrait aussi permettre de piloter des systèmes robotisés dans le but d'assister le chirurgien dans ses actes chirurgicaux.

Dans ce chapitre, nous présentons l'outil endoscopique tel que nous l'envisageons. Il est évident que le cadre dans lequel nous souhaitons pouvoir utiliser cet outil, l'intérieur de la cavité abdominale, est assez particulier et impose de nombreuses contraintes. Dans un premier temps, nous présentons ces contraintes très spécifiques à notre application. Nous détaillons quelques considérations liées à l'interaction de la lumière avec les organes afin de choisir la longueur d'onde la plus appropriée pour la projection dans un environnement tel que la cavité abdominale. Ensuite, nous présentons la technologie des éléments d'optique diffractive qui a été retenue pour réaliser la projection de notre motif ainsi que les étapes de la réalisation de ces éléments. Nous terminons enfin par la présentation du prototype de l'outil et la réalisation matérielle actuelle.

4.1 Contraintes liées à l'application

Nous nous intéressons à la conception d'un outil endoscopique dans lequel un système de lumière structurée doit être incorporé. Généralement, le système de lumière structurée consiste en une caméra et en un moyen de projection de motif lumineux. Le but est de fournir la reconstruction 3D en temps réel de la surface sur laquelle le motif a été projeté. C'est pourquoi la projection d'un motif unique a été adopté. Pour encoder ce motif, il a été possible d'employer

différentes couleurs ou bien de définir des primitives géométriques afin de générer un motif monochromatique. Vu la nature très colorée de l'environnement de la cavité abdominale, la projection d'un motif contenant différentes couleurs peut poser des difficultés lors du choix de ces couleurs et au niveau de la segmentation. De plus, on ne dispose pas de moyen miniaturisé et flexible capable de projeter un motif contenant plusieurs couleurs. C'est pourquoi, durant la phase de conception du motif (voir section 2.2.2), nous avons imposé que le motif soit monochromatique. Cependant, le choix de la longueur d'onde de la lumière qui devra être émise doit prendre en compte la particularité de l'environnement dans lequel elle est projetée. En effet, la couleur de la projection doit être choisie parmi les couleurs présentant un bon comportement sur les surfaces des organes abdominaux afin de faciliter la détection des primitives projetées. L'endoscope que nous cherchons à concevoir est appelé à fournir l'information 3D des surfaces abdominales au cours d'une intervention mini-invasive, par conséquent, les dimensions de l'outil doivent être de l'ordre des dimensions des outils endoscopiques (voir section 1.1). Vu que notre motif est composé de primitives géométriques différentes, nous avons proposé une solution basée sur la technologie des optiques diffractives pour réaliser un moyen de projection de notre motif dont les dimensions sont compatibles avec les tailles envisagées. Ce sont donc essentiellement ces deux contraintes, auxquelles on pourrait rajouter une architecture compatible avec les dimensions d'un endoscope qui vont guider la conception.

4.2 Choix de la couleur de projection

Comme nous l'avons déjà mentionné, le motif que nous proposons est basé sur la projection d'une lumière monochromatique. Afin d'avoir une bonne perception des motifs projetés dans l'image, la couleur choisie doit correspondre à une longueur d'onde pour laquelle les surfaces des organes abdominaux présentent une bonne réflectance et le moins de diffusion possible. En effet, l'interaction de la lumière avec des tissus organiques n'est pas sans effet sur la perception que l'on peut avoir de la couleur réfléchi. Cette interaction a déjà fait l'objet d'un certain nombre d'études. Dans [Sambongi00] par exemple, l'influence de l'état des tissus vivants (tissus sains ou tuméfiés) sur la réflectance des organes a été étudiée à différentes longueurs d'onde. Deux organes ont fait l'objet de cette étude, le côlon et l'estomac. Cette étude est intéressante car les tissus étudiés correspondent à peu près à ceux que nous souhaitons observer. Les réflectances mesurées sur des surfaces d'organes atteints d'un cancer et de celles d'organes sains sont illustrées sur la Figure 4.1. Nous constatons que dans les deux cas et pour les deux organes, le comportement général des réflectances est similaire. On observe des maxima locaux pour des longueurs d'onde proches du rouge (680 nm) et du vert (540 nm).

D'autre part, le comportement de la réflectance de la peau à différentes longueurs d'ondes a également déjà été étudiée dans [Angelopoulou99]. Dans ce cas aussi, nous constatons un comportement similaire à la réflectance des organes étudiés dans [Sambongi00] et avec des réflectances maximales pour des longueurs d'onde proches du rouge et du vert (Figure 4.2).

Bien que l'interaction de la lumière avec la peau ne nous concerne pas directement pour notre étude, il semble qu'elle confirme néanmoins le fait que les couleurs proches du rouge et du vert (de longueur d'onde, respectivement 530nm et 650nm) sont appropriées pour l'application que nous visons. Le choix de ces couleurs est conforté par le fait que visuellement, ces couleurs sont facilement discernables à l'œil nu lorsqu'elles sont projetées sur des organes dans la cavité abdominale. Cela a pu être vérifié empiriquement lors de manipulations réalisées en collaboration

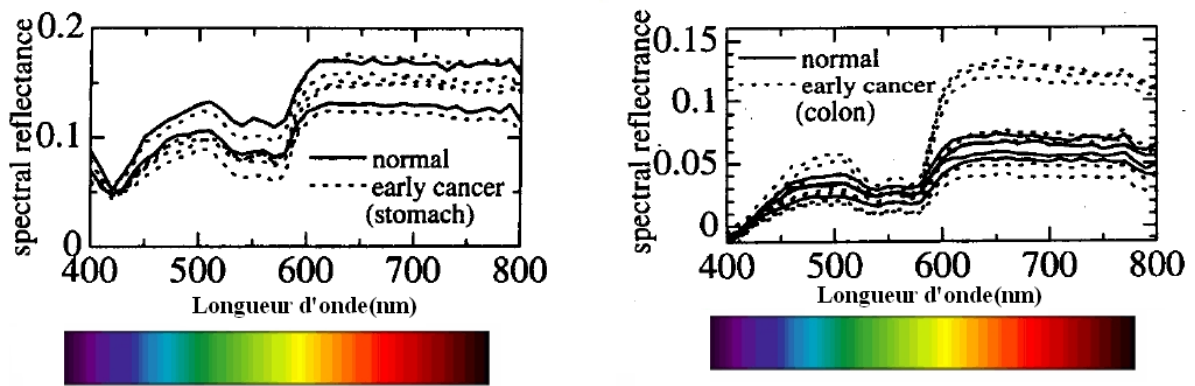


FIGURE 4.1 – Réflectance à différentes longueurs d'onde pour des tissus vivant provenant de l'estomac (à gauche) et du côlon (à droite) (source : [Sambongi00]).

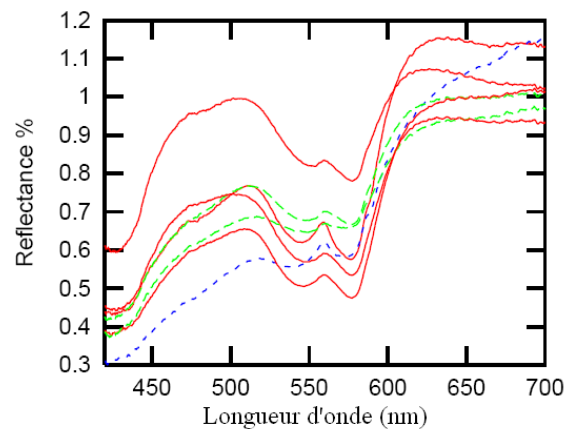


FIGURE 4.2 – Réflectance de la peau (source : [Angelopoulou99]).

avec les chirurgiens de l'IRCAD (Institut de Recherche contre les Cancers de l'Appareil Digestif). Pour ce faire, nous avons projeté des motifs colorés sur des structures internes de l'abdomen de porcs. De nombreuses combinaisons de couleurs ont été expérimentées et elles confirment toutes ces premières constatations. Des exemples sont présentés dans la Figure 4.3.

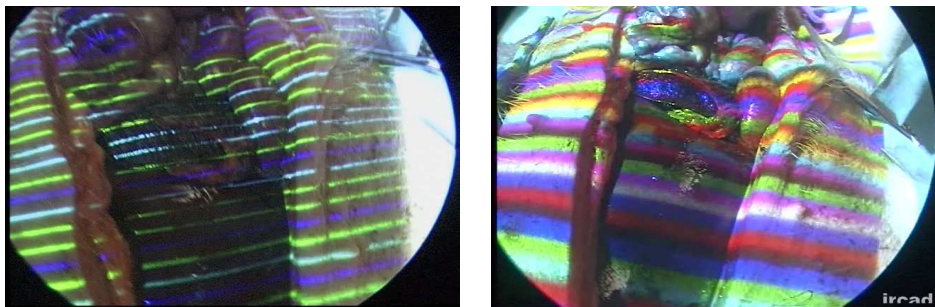


FIGURE 4.3 – Projection de motifs colorés sur des organes de l'abdomen d'un porc.

Afin de compléter ces expérimentations, nous avons projeté une matrice de points colorés sur les organes en utilisant un laser vert et un laser rouge puisque ces deux couleurs semblent convenir à l'utilisation que l'on souhaite faire de notre outil. Il est ressorti de cette étude que, contrairement au vert, la lumière rouge diffuse énormément lorsqu'elle est projetée sur les surfaces des organes abdominaux (Figure 4.4).

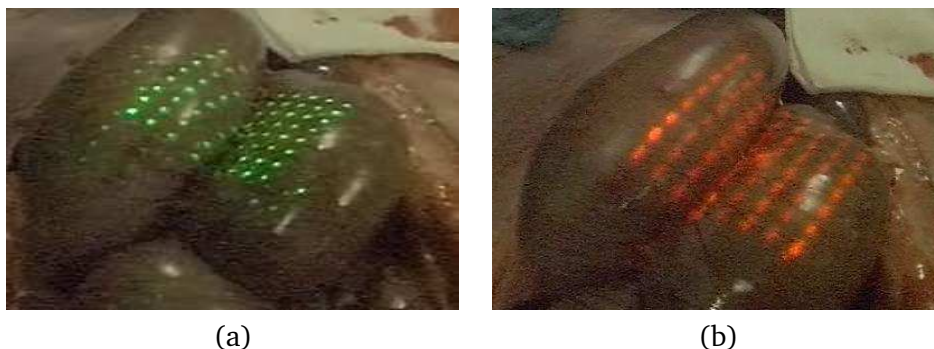


FIGURE 4.4 – Projection d'une matrice de points colorés sur des organes de l'abdomen d'un porc : (a) points verts (b) points rouges.

Au sein du laboratoire, un certain nombre de travaux ont déjà été menés pour lesquels des sources lumineuses sont utilisées à l'intérieur de la cavité abdominale. On peut notamment citer [Krupa03] dans lequel quatre faisceaux de laser parallèles et rouges ont été projetés sur les surfaces des organes abdominaux dans le but de détecter le centre des quatre taches correspondant à la projection de ces faisceaux. On a alors constaté une diffusion importante de la couleur rouge sur les surfaces abdominales (Figure 4.5).

Par conséquent, nous avons choisi d'utiliser une couleur verte (longueur d'onde de 530 nm) pour la projection de notre motif de façon plus précise.

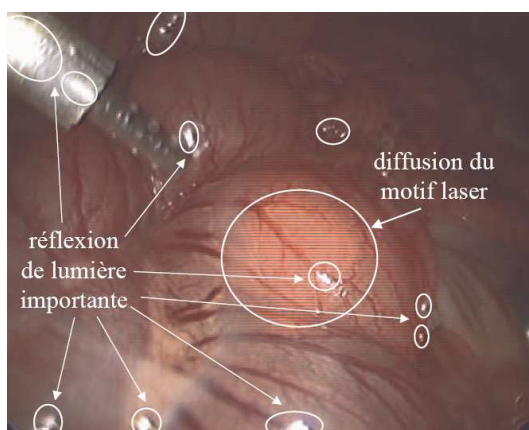


FIGURE 4.5 – Diffusion de la couleur rouge sur les surfaces abdominales (source : [Krupa03]).

4.3 Éléments d'optique diffractive

Les systèmes optiques font de plus en plus appel à la miniaturisation afin d'atteindre une haute résolution et la possibilité de concevoir des systèmes portables. La technologie de micro-fabrication offre des solutions permettant la réduction de la taille et la production en masse. De plus, cette technologie facilite l'assemblage des composants optiques. Récemment, des méthodes ont été proposées pour fabriquer des éléments optiques basés sur la diffraction de la lumière cohérente en utilisant la micro-fabrication (par exemple : [Swanson89, Tanigami89, Kress97, Kololuonma02] et [Wang02]).

4.3.1 Principe

La diffraction est un comportement ondulatoire de la lumière quand elle rencontre un obstacle qui n'est pas complètement transparent. Contrairement à une optique standard (lentilles et miroirs), les éléments d'optique diffractive (*Diffractive Optical Elements* (DOE)) sont des micro-structures qui font usage de la diffraction pour remodeler le faisceau de la lumière incident afin de créer un motif prédéfini. En d'autres termes, ces optiques sont conçues de telle sorte que le rayon laser qui les illumine se diffracte. Elles sont alors prévues pour que la composition de l'ensemble de diffractions forme un motif défini à l'avance. Il faut donc disposer d'une optique pour chaque motif que l'on souhaite projeter. Ces éléments sont devenus une des clés de la miniaturisation des systèmes optiques grâce à leur structure plane et mince et parce qu'ils peuvent être produits en utilisant des technologies de fabrication des semi-conducteurs [Dai02].

Les DOEs peuvent être de nature différente, ils peuvent être conçus soit en tant qu'éléments de Fresnel, soit en tant qu'éléments de Fourier. La différence entre ces éléments est liée essentiellement à la distance focale. Cette distance est plus courte pour les éléments de Fresnel que pour ceux de Fourier. Un élément diffractif de Fourier permet de séparer un faisceau incident en n faisceaux suivant n directions symétriques par rapport à l'axe optique du dispositif. De ce fait, les éléments de Fourier sont particulièrement bien adaptés à la génération des matrices d'éclairage. Par contre, les éléments de Fresnel sont généralement utilisés pour re-focaliser des faisceaux sur l'axe optique.

Le problème principal de ces optiques réside dans la difficulté de fabrication. En effet, le profil de phase gravé dans l'élément diffractif définit son fonctionnement. Or, pour que l'efficacité soit de 100%, il faudrait que le profil de phase soit continu. Malheureusement, on ne dispose pas de technologie capable de générer un profil continu de haute qualité avec une haute efficacité. C'est pourquoi, on utilise des approximations du profil continu [Goodman70] en le quantifiant en niveaux de phase discrets (Figure 4.6).

Différentes méthodes de fabrication des DOEs ont été proposées en se basant sur cette approche. La lithographie optique à multi-niveaux est parmi les méthodes les plus intéressantes [Swanson93], elle comporte la génération de photomasques binaires contenant l'information des niveaux de phase. Les photomasques sont configurés pour fournir 2^n niveaux de quantification de phase où n est le nombre de masques considérés. Ces photomasques peuvent être fabriqués par lithographie avec un faisceau d'électrons. Les photomasques sont appliqués successivement selon le nombre de niveaux considérés et la gravure du profil de phase est réalisée à l'aide de la technique de *reactive ion etching* (RIE) (Figure 4.7).

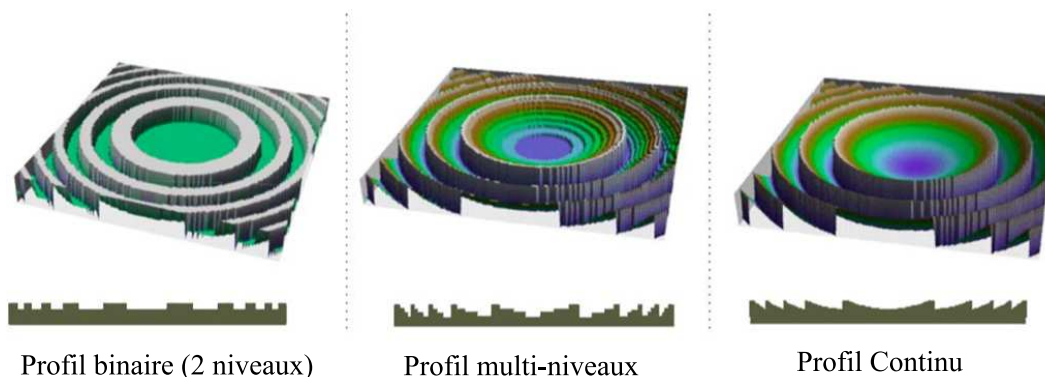


FIGURE 4.6 – Discretisation du profil de phase.

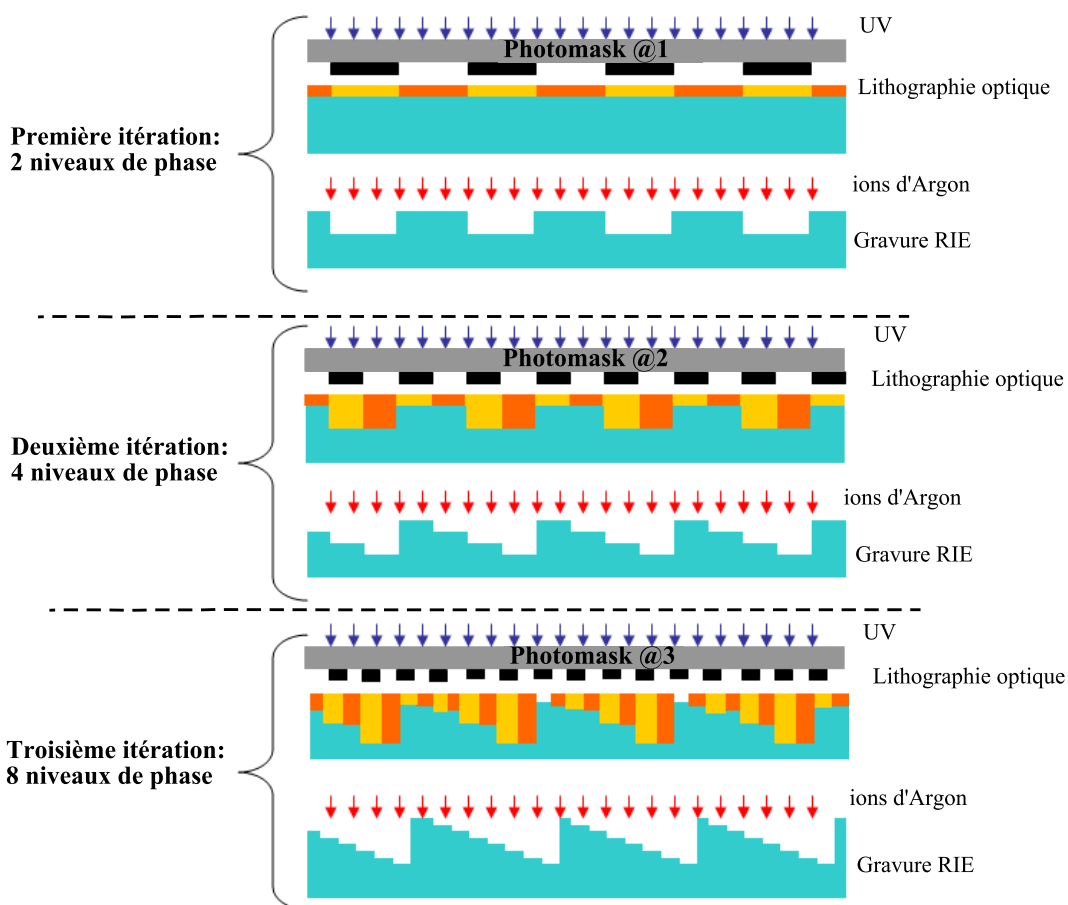


FIGURE 4.7 – Fabrication des DOEs avec la méthode de la lithographie optique à multi-niveaux.

Or, un élément diffractif réalisé avec un seul photomasque binaire (ce qui correspond à deux niveaux de phase) ne permet de restituer qu'une efficacité d'environ 40% de l'énergie du laser incident. Pour améliorer le rapport de l'efficacité énergétique, il y a lieu d'augmenter le nombre de niveaux de quantification. Afin d'atteindre 95% d'efficacité, on a besoin de considérer au moins 8 niveaux de quantification, ce qui revient à utiliser trois photomasques. Cependant, le

dans notre cas d'environ 20° pour un laser vert. Généralement, les simulations numériques sont réalisées à partir d'une image de taille 1024×1024 pixels qui contient le motif désiré et qui correspond à l'ouverture maximale. La taille réelle du motif dans cette image détermine aussi son angle propre de projection qui est inférieur à l'angle maximal (Figure 4.9).

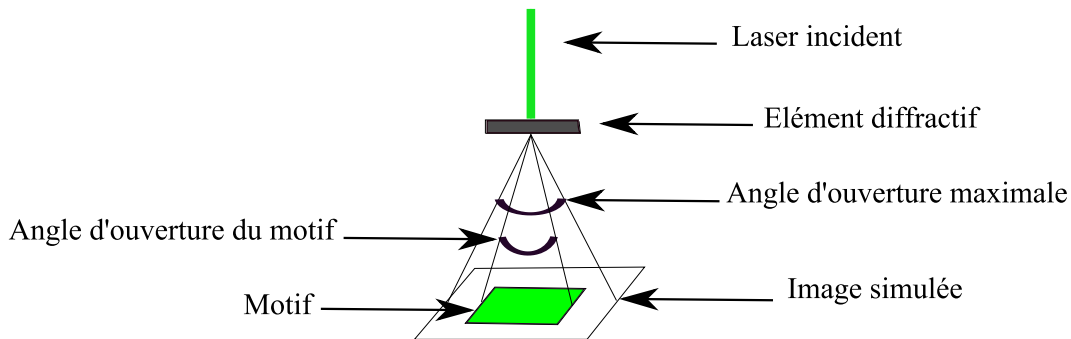


FIGURE 4.9 – Ouverture de projection.

Cependant, les éléments de la première génération ont été simulés avec un angle d'ouverture du motif de seulement 14° ce qui nécessite une distance importante de la surface pour qu'on puisse distinguer la projection du motif. Par la suite, plusieurs configurations de notre motif ont été simulées pour la deuxième génération des éléments avec une fabrication à quatre niveaux dans le but d'obtenir une image plus précise de la projection. La distribution de ces éléments dans chaque disque est sous la forme d'une matrice de taille 4×5 éléments (Figure 4.10-a). Les dimensions de chaque élément sont de 6×6 mm et le diamètre de la région qui contient les gravures dans chaque élément est de 1.4 mm. La réalisation des éléments est présentée sur la Figure 4.10-b.

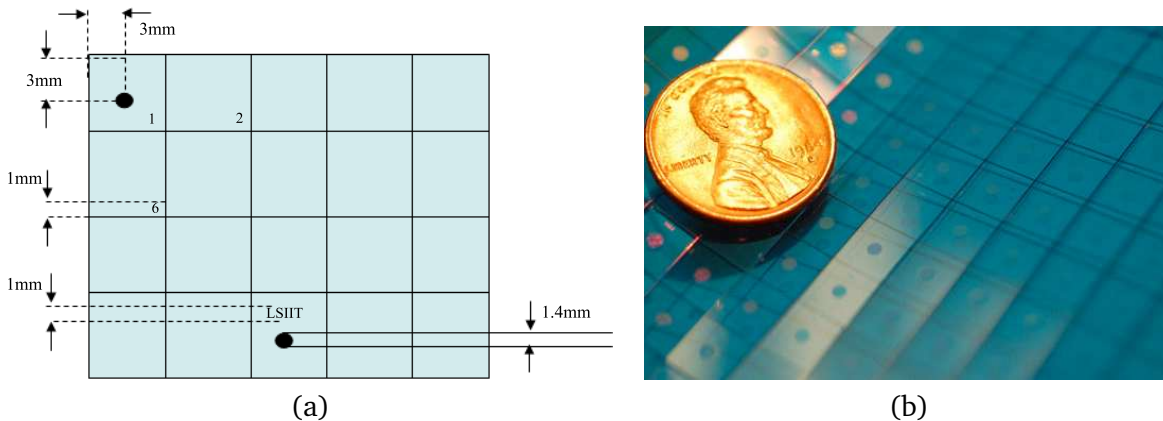


FIGURE 4.10 – (a) Distribution de nos éléments diffractifs sur chaque disque. (b) Réalisation de la deuxième génération des éléments sur un disque de quartz.

Les éléments ont été gravés sur des disques de quartz avec une résolution de gravure de 100 nm et une précision d'alignement de masques de 500 nm. La Figure 4.11 présente les images de la projection avec un laser vert et des éléments fabriqués à des angles d'ouverture différents.

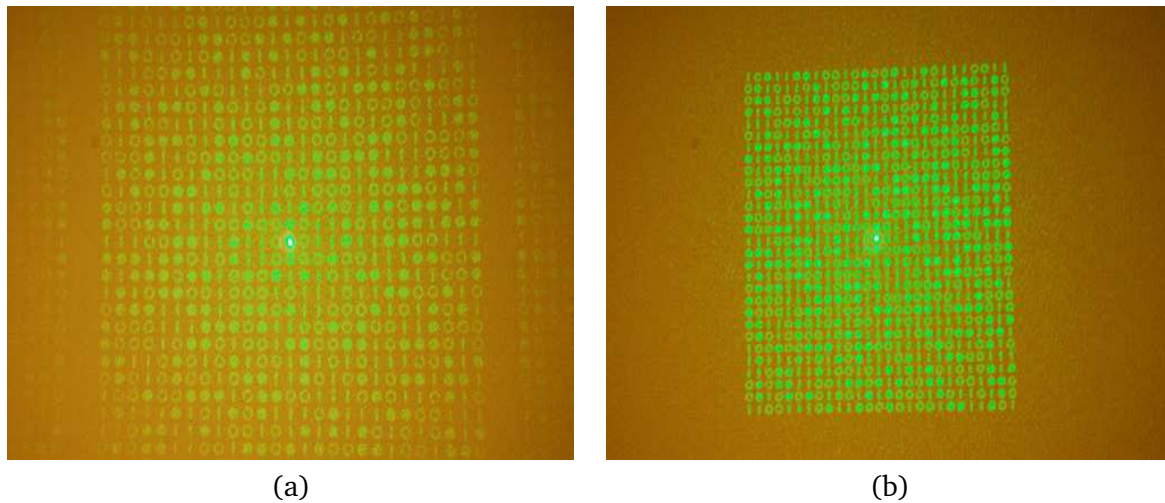


FIGURE 4.11 – Projections obtenues à l'aide d'un laser vert avec les éléments diffractifs fabriqués (angle d'ouverture de 20° (à gauche) et de 14° (à droite)).

On peut remarquer que la résolution des images de projection obtenues avec les éléments de la deuxième génération était meilleure que celle correspondant à la première génération car la quantification du profil de phase à quatre niveaux donne une approximation du profil de phase plus fine qu'une quantification à deux niveaux. De plus, la netteté des images obtenues avec les éléments conçus pour un angle d'ouverture de 20° est supérieure à celle des autres éléments (Figure 4.11-a). C'est pourquoi, nous avons retenu la fabrication d'un élément à quatre niveaux avec un angle d'ouverture de 20° comme correspondant à la configuration optimale pour notre application.

4.4 Conception d'un outil endoscopique de vision 3D

Après avoir conçu un motif monochromatique avec un codage robuste, celui-ci a été utilisé pour la générations des éléments d'optique diffractive afin de réaliser un moyen compact de projection. Dans le but de concevoir un outil endoscopique de vision tridimensionnelle basé sur cette technique et incorporant le moyen de projection proposé, nous envisageons d'employer l'endoscope stéréoscopique (Figure 1.9 Figure 1.7-a) en disposant l'élément d'optique diffractive éclairé par une source laser externe à la place d'une des caméras (Figure 4.12).

Ne disposant actuellement pas de cet endoscope stéréo, nous avons mis en œuvre un système composé de deux endoscopes de la société Storz faisant un angle entre eux (Figure 4.13). Ce système sera utilisé par la suite pour les expériences *in-vivo* dans le but d'évaluer la reconstruction 3D de ce système de lumière structurée sur des surfaces abdominales. La fibre optique amène le faisceau du laser jusqu'à l'élément diffractif et permet de rendre le système plus souple à l'utilisation. Un collimateur est utilisé pour assurer la bonne collimation du faisceau à la sortie de la fibre. Dans cette configuration, l'endoscope attaché à l'élément diffractif joue le rôle du chemin optique de la projection et l'autre endoscope joue le rôle de l'objectif de la caméra (de Storz). L'acquisition d'images se fait à l'aide d'un convertisseur qui convertit le signal vidéo analogique du contrôleur de la caméra en signal vidéo numérique selon la norme *FireWire 1394*.

La puissance du laser a été choisie de manière à obtenir une luminosité suffisante de la projec-

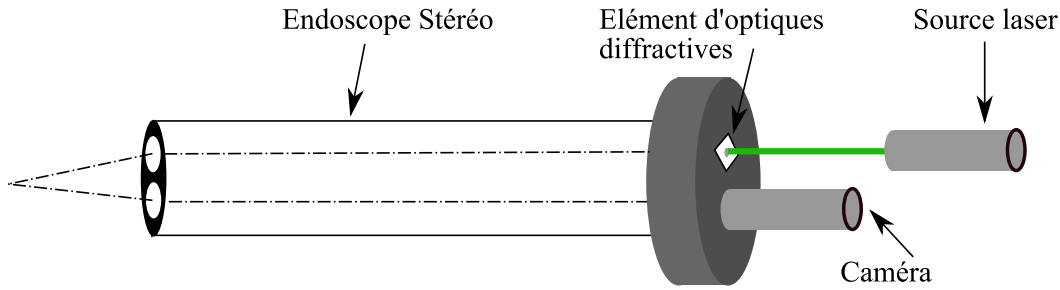


FIGURE 4.12 – Schéma de conception envisagée de l'outil endoscopique incorporant un système de vision 3D à lumière structurée.

tion afin que la caméra puisse détecter les primitives projetées. En effet, la luminosité dépend essentiellement de la puissance du laser, de l'efficacité de l'élément diffractif et du nombre de primitives dans le motif puisque la puissance du faisceau laser diffracté sera distribuée sur toutes les primitives. Comme nous l'avons mentionné, les éléments diffractifs ont été configurés afin d'avoir une efficacité d'environ 80%. Cependant, nous avons constaté expérimentalement que la puissance minimale du laser pour obtenir une bonne perception du motif projeté est d'environ 100 mW. Cette puissance est assez importante et elle est due au fait que le motif contient beaucoup de primitives (783 primitives). Actuellement, la puissance du laser utilisé dans notre système expérimental est de (200 mW).

Quand une partie du faisceau de laser est diffractée en passant par l'élément diffractif, l'autre partie traverse la surface de l'élément comme si elle était transparente. Cette partie de faisceau forme ce qu'on appelle la composante continue qui se superpose à la projection du point central qui devient très lumineux. Malheureusement, cela cause des saturations dans le capteur de la caméra correspondant à la région centrale du motif. Par conséquent, la détection des primitives dans cette région devient très difficile. Afin de surmonter ce problème dans le prototype final, nous envisageons d'ajouter un masque optique après le passage par l'élément diffractif afin de diminuer la puissance du motif dans la région centrale.

4.5 Conclusion

L'objectif de nos travaux est de concevoir un dispositif endoscopique incorporant un système de lumière structurée capable de fournir la reconstruction 3D en temps réel de surfaces abdominales. Pour cela, nous avons conçu un nouveau motif lumineux monochrome auquel nous avons associé un codage robuste. La première étape de la conception consiste à choisir la couleur de la projection en prenant en compte l'interaction de la lumière avec les surfaces des organes de l'abdomen. Suite à une étude sur cet aspect, nous avons retenu la couleur verte pour éclairer le motif. Nous avons fait appel à la technologie des optiques diffractives afin de réaliser un dispositif miniaturisé de projection. Deux générations d'éléments d'optique diffractive avec différentes configurations ont été dimensionnées et fabriqués dans le but de sélectionner la configuration optimale pour notre application. Grâce à sa bonne efficacité et à la bonne résolution des images de projection, notre choix s'est porté sur l'élément fabriqué à quatre niveaux de phase et avec un angle d'ouverture de 20° . La réalisation du dispositif complet est basée sur l'utilisation d'un en-

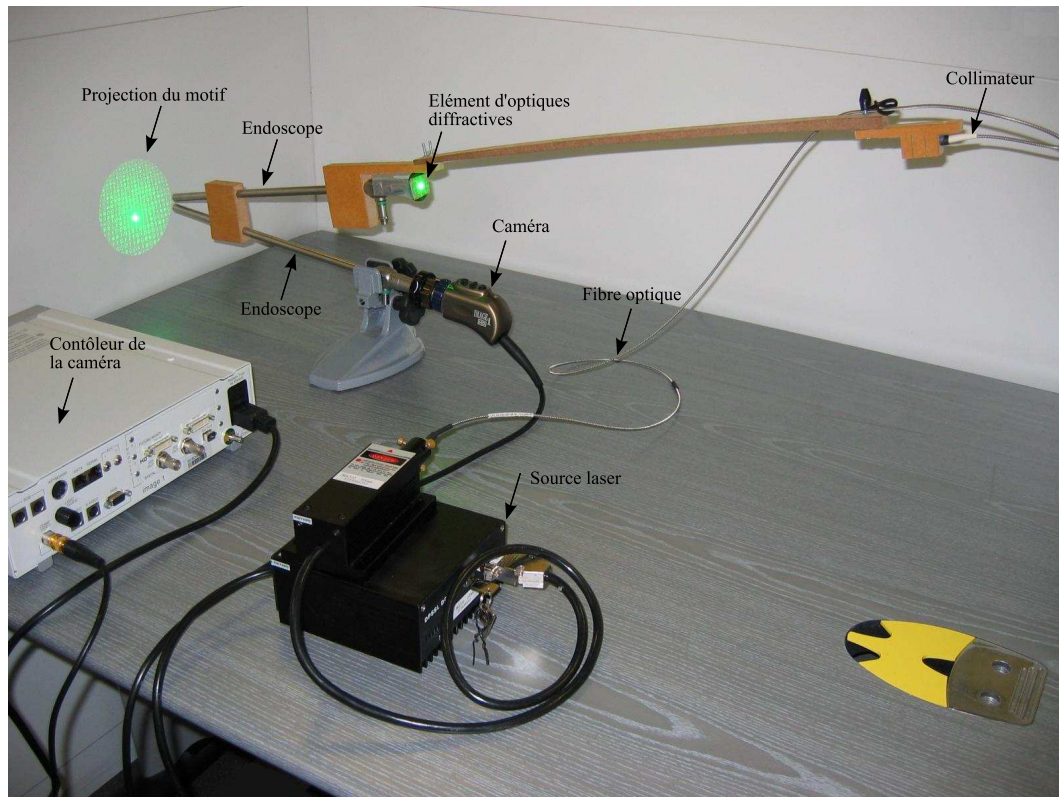


FIGURE 4.13 – Configuration du système d'expérimentation.

doscope stéréoscopique qui est actuellement simulé par deux endoscopes distincts. Nous avons mis en œuvre un système expérimental afin de valider notre conception et d'évaluer, par la suite, la reconstruction 3D *in-vivo*. La puissance minimale du laser a été déterminée mais le problème de la composante continue reste à résoudre dans la prochaine version du prototype.

Chapitre 5

Résultats expérimentaux

Nous avons conçu un nouveau motif pour la reconstruction 3D en temps réel de surfaces en milieu non contrôlés basée sur une technique de projection de lumière structurée. La conception du motif a été détaillée dans le deuxième chapitre de même que les algorithmes de segmentation et de décodage associés. Dans le troisième chapitre, nous avons présenté la méthode que nous proposons pour l'étalonnage du dispositif afin de déterminer le comportement du motif projeté dans l'espace et de calculer la reconstruction 3D métrique. Cette méthode peut également être utilisée dans divers types d'applications nécessitant la connaissance 3D de son environnement proche. Cependant, l'objectif de ces travaux plus précisément est de concevoir un outil endoscopique capable de fournir l'information 3D des surfaces abdominales en temps réel. La conception de cet outil a été détaillée dans le quatrième chapitre.

Dans ce chapitre, nous présentons dans un premier temps des résultats expérimentaux obtenus avec un système de lumière structurée traditionnel basé sur la projection notre motif. Ensuite, le système expérimental proposé dans le quatrième chapitre est étalonné afin d'évaluer la précision de la reconstruction 3D obtenue avec un tel système. Nous terminons par l'analyse des reconstructions 3D obtenues avec ce système pour des projections sur des objets réels et finalement sur des surfaces d'organes de la cavité abdominale d'un porc.

5.1 Résultats expérimentaux : projection du motif proposé

Comme nous l'avons déjà mentionné, le motif que nous avons conçu peut être utilisé dans des applications pour lesquelles la connaissance de l'information 3D est nécessaire. Compte tenu de la grande diversité des applications possibles, nous avons conçu ce motif de manière à ce que son codage soit robuste, que le traitement associé soit rapide, et que la reconstruction 3D obtenue soit précise. Afin de valider ces aspects, nous avons installé un système expérimental de lumière structurée composé d'une caméra (AVT Pike F100C) et d'un vidéo-projecteur de type Sony VPL-CS6 (Figure 5.1). Nous avons implémenté nos algorithmes de détection, de décodage et de reconstruction en C++ sous Linux et nous avons développé une interface (Figure 5.2) afin de gérer les communications avec la caméra à travers le port *FireWire* et d'exécuter nos algorithmes avec les images acquises. La caméra peut acquérir des images au *format7* du protocole vidéo *DCAM*, ce qui permet de spécifier la taille et la position des régions d'intérêt sur le capteur ainsi que le format d'encodage des pixels. Nous avons choisi d'acquérir des images en couleurs *RGB* de taille (512×512) pixels ce qui correspond à un bon compromis entre la taille d'une image et le temps d'exécution. Il est à remarquer que généralement les tailles des images considérées dans d'autres travaux sont de l'ordre de celles que nous considérons ici.



FIGURE 5.1 – Système de lumière structurée traditionnel projetant notre motif.

Le système a d'abord été étalonné selon la méthode proposée dans le troisième chapitre avec un damier composé de (12×15) carrés gris et blancs de 10 mm de côté afin d'obtenir une reconstruction 3D métrique. La précision de cet étalonnage a été évaluée. Les résultats ont été présentés dans la section 3.3 où nous avons trouvé que l'erreur moyenne est inférieure à 1 mm en profondeur et inférieure à 1° pour l'orientation sachant que nous avons considéré un espace de travail d'une profondeur comprise entre 600 et 800 mm. On peut rappeler que la précision obtenue avec notre méthode est largement suffisante pour le type d'applications que nous visons.

Pour estimer la reconstruction, nous avons projeté le motif sur des objets à géométrie simple afin de pouvoir tester les performances de nos algorithmes : la robustesse du codage associé, la précision de la reconstruction métrique et les performances en termes de temps d'exécution. Des résultats sont illustrés par les Figures 5.3 et 5.4. Les primitives correctement décodées (attribution correcte de position (ligne et colonne) dans la matrice du motif) sont marquées en rouge. La robustesse du codage proposé en présence d'ombres et de discontinuités de surfaces est évidente vu le grand pourcentage de primitives qu'elles ont été décodées correctement (Figure 5.3-b et Figure 5.4-b) : plus de 95% de primitives détectées ont été décodées correctement (voir aussi le tableau 2.1). On peut remarquer plus particulièrement sur la Figure 5.4-b l'efficacité de l'algorithme de détection de voisinage qui prend en compte les directions des tirets. Cependant, des primitives proches des contours des objets n'ont pas été décodées correctement car elles étaient déformées par ces contours. Par contre, il existe des primitives qui étaient déformées et qui ont été décodées correctement grâce à la méthode des votes malgré leurs déformations. Cependant, la déformation a causé des erreurs pour la détermination de leurs positions dans l'image ce qui a conduit à des erreurs de reconstruction. Les reconstructions métriques sont illustrées sur les Figures 5.3-c et 5.4-c où on peut apprécier visuellement la qualité des reconstructions par d'autres points de vues.

Les erreurs de reconstruction que nous venons d'évoquer, bien que ne remettant pas en cause le motif proposé, peuvent être gênantes. Ce sont des erreurs qui apparaissent, quelque soit le principe de lumière structurée utilisé. Grâce à l'utilisation du principe des matrices parfaites, leur nombre a déjà été grandement diminué. Ce genre d'inconvénient est inévitable. Cependant, pour en limiter un peu plus ces effets, il faudrait encore améliorer les algorithmes de segmentation afin de diminuer les erreurs de détection et de classification des primitives.

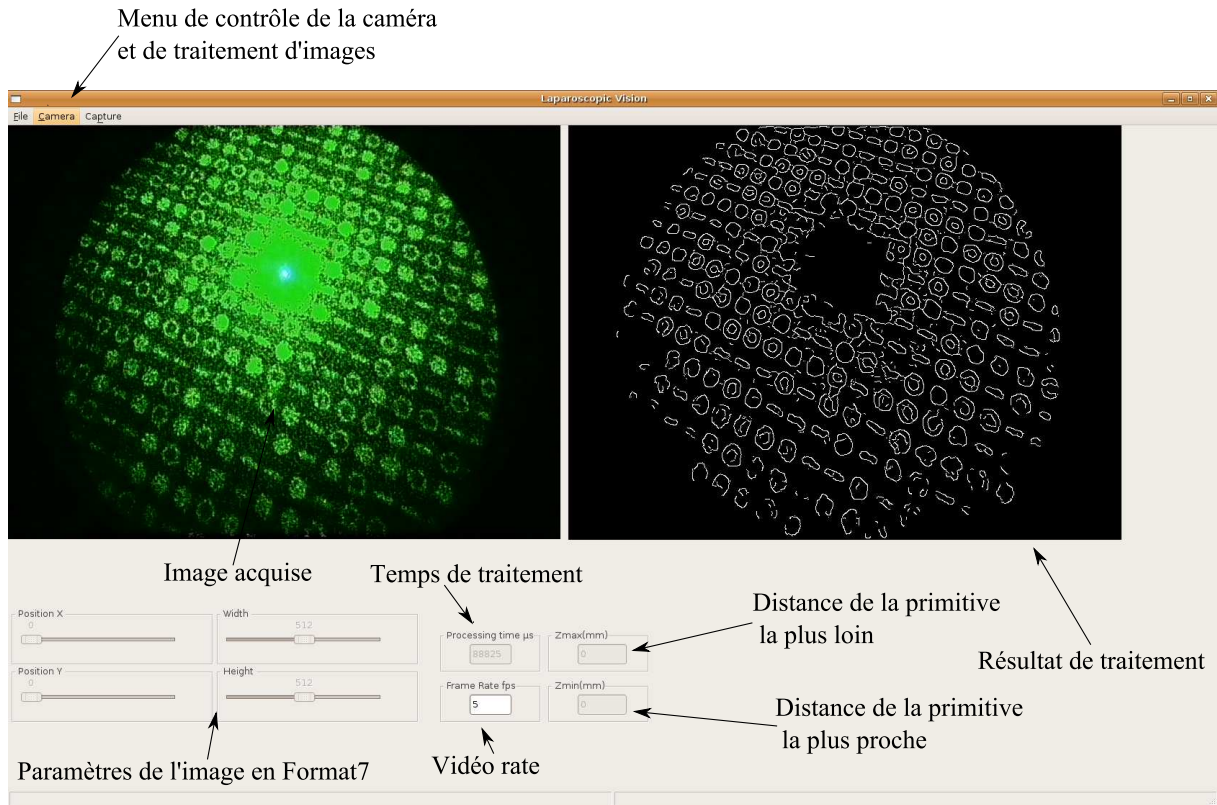
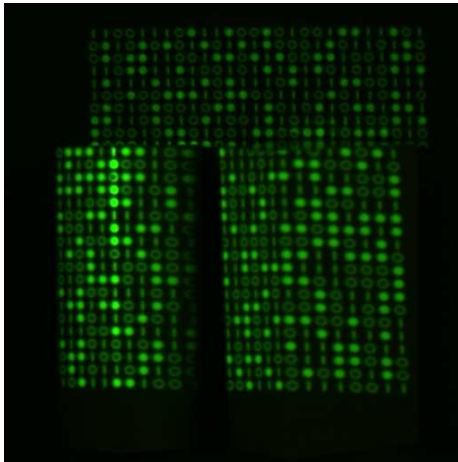


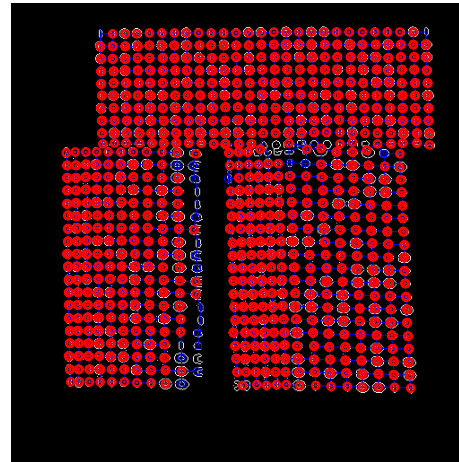
FIGURE 5.2 – Interface développée pour permettre de gérer la communication avec la caméra et d'appliquer nos algorithmes de détection, de décodage et de reconstruction aux images acquises.

Le système a aussi été testé sur des surfaces plus complexes. Nous présentons par exemple dans la Figure 5.5-a la projection de notre motif sur le visage d'une poupée. Les résultats de la segmentation et du décodage sont illustrés par la Figure 5.5-b et la reconstruction 3D est présentée par la Figure 5.5-c. Nous constatons que nos algorithmes ont réussi à décoder un grand pourcentage (plus de 85%) des primitives visibles et à reconstruire les surfaces matérialisées par la projection de ces primitives. On constate aussi dans ce cas la qualité et la précision de la description de surfaces définie par la reconstruction obtenue (Figure 5.5-d). Vu que la scène contient plusieurs surfaces (le visage de la poupée, son corps et l'arrière plan), ces surfaces ont été reconstruites grâce à la robustesse de notre méthode et malgré les discontinuités. Cependant, on remarque sur la Figure 5.5-d que les surfaces ont été séparées. Dans le but de mieux afficher les discontinuités entre ces surfaces, nous avons considéré que deux points appartiennent à une même surface si la distance entre eux ne dépasse pas un seuil prédéfini.

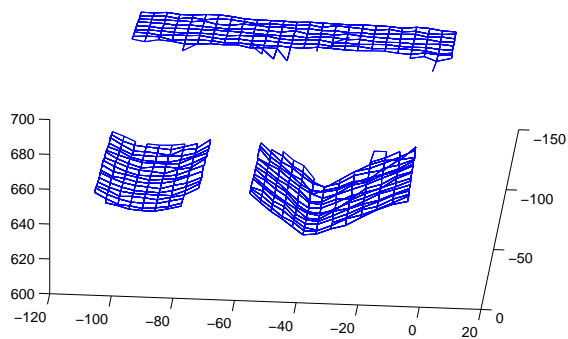
Finalement, on constate que les résultats obtenus ont été générés après des traitements pour lesquels le temps d'exécution nécessite en moyenne 65 ms pour un nombre de primitives traitées d'à peu près 700 primitives. Les champs de profondeur sont de 600 à 800 mm sur une région de surface d'environ 15×15 cm. Ces résultats sont encourageants et confirment l'efficacité de notre méthode en terme de temps d'exécution sachant qu'aucune optimisation n'a encore été faite. Cela montre que le motif proposé peut être utilisé pour plusieurs applications différentes nécessitant la connaissance de l'information tridimensionnelle de son environnement avec une précision de l'ordre de celle obtenue avec notre méthode.



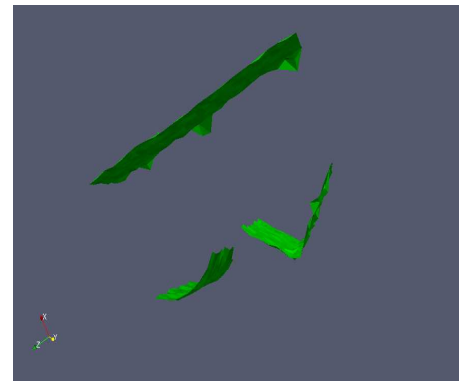
(a)



(b)



(c)



(d)

FIGURE 5.3 – (a) Projection sur des objets géométriques simples avec des ombres et des discontinuités. (b) Résultat de la détection et du décodage : les points rouges représentent les primitives bien décodées. (c) Reconstruction 3D des surfaces. (d) Rendu réaliste des surfaces reconstruites.

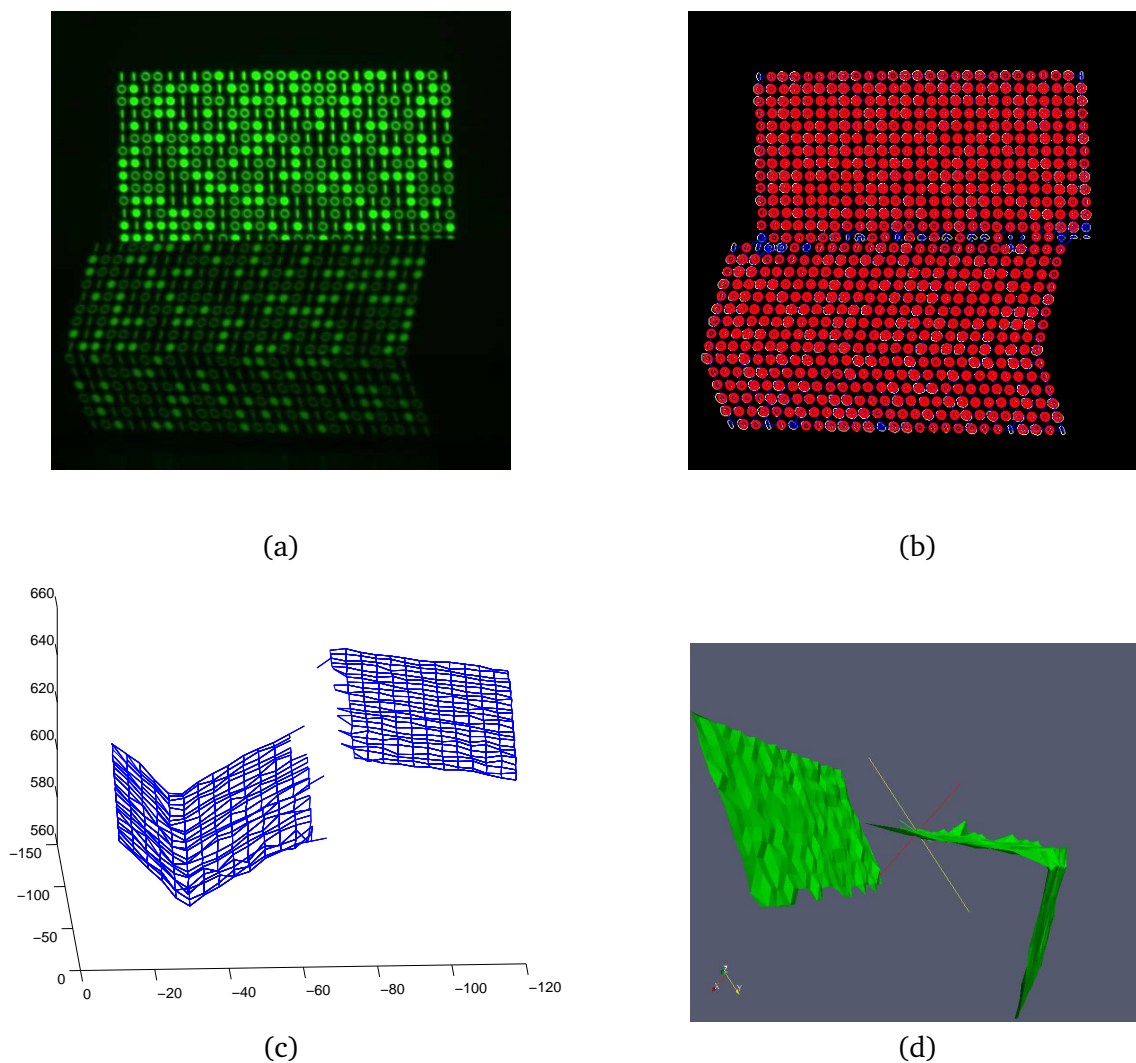
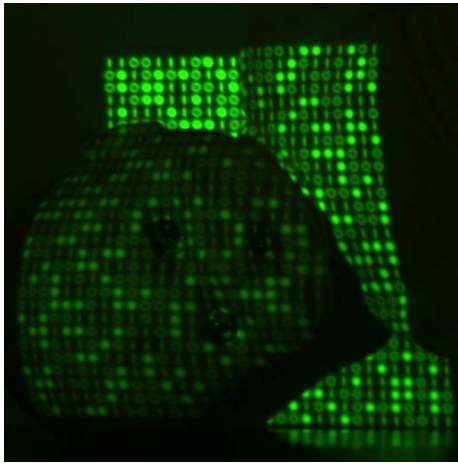
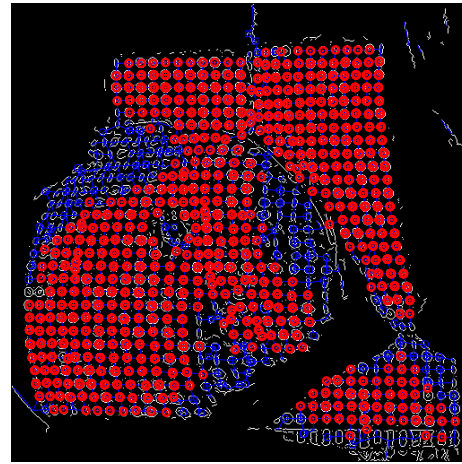


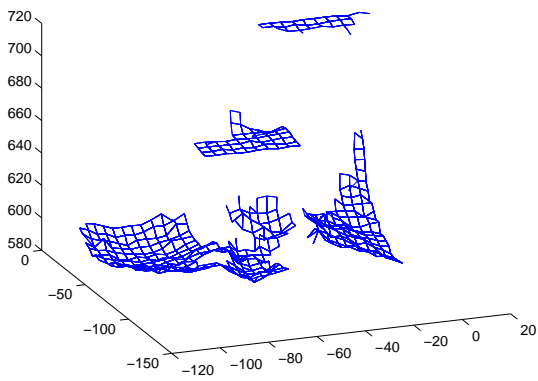
FIGURE 5.4 – (a) Projection sur des objets géométriques simples avec des inclinaisons différentes. (b) Résultat de la détection et du décodage : les points rouges représentent les primitives bien décodées. (c) Reconstruction 3D des surfaces. (d) Rendu réaliste des surfaces reconstruites.



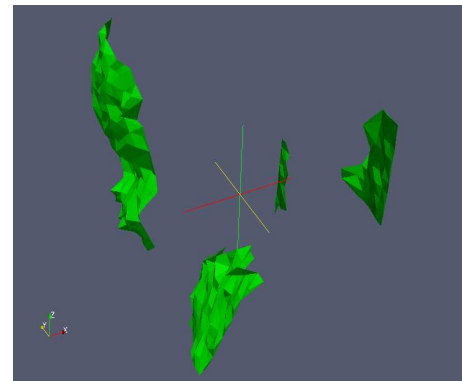
(a)



(b)



(c)



(d)

FIGURE 5.5 – (a) Projection sur le visage d'une poupée. (b) Résultat de la détection et du décodage : les points rouges représentent les primitives bien décodées. (c) Reconstruction 3D des surfaces. (d) Rendu réaliste des surfaces reconstruites.

5.2 Résultats expérimentaux pour le prototype d'endoscope 3D

Les résultats précédents ont démontré que notre méthode pouvait être utilisée pour des applications diverses. Mais pour atteindre l'objectif principal visé par ces travaux, il s'agissait de pouvoir reconstruire des surfaces à l'intérieur de la cavité abdominale. C'est la raison pour laquelle nous avons mis en œuvre nos techniques sur le prototype d'endoscope que nous avons proposé dans la section 4.4. Les expérimentations que nous avons menées ont permis d'évaluer son fonctionnement et d'apporter des critiques objectives quant aux techniques utilisées.

5.2.1 Source laser et acquisition d'images

De manière à respecter les contraintes citées dans le chapitre précédent, nous avons intégré l'élément d'optique diffractive dans le prototype d'endoscope et nous l'avons illuminé à l'aide d'une source laser émettant dans le vert. Le laser que nous avons employé est de type (*MGL – III*) d'une puissance de 200 mW et qui émet une lumière de longueur d'onde de 532 nm. À notre grande surprise, les images produites étaient d'une qualité médiocre (Figure 5.6). Elles sont sous cette forme inexploitable. En effet, deux phénomènes inattendus se sont produits : une forte sur-intensité centrale est observée causant une saturation dans une région importante dans l'image. De plus, on constate un bruit rendant l'image inexploitable. Ce bruit est dû à la source laser qui génère un bruit important. D'autre part, la caméra utilisée est une caméra à haute définition *Storz* connectée à un contrôleur qui traite et compresse les images avant de les envoyer vers une sortie *DVI*. Cette solution est incontournable puisque ce système est fourni avec l'endoscope. Par conséquent, l'accès direct aux images numériques acquises n'est pas disponible. Pour contourner cette difficulté et afin d'acquérir des images numériques non compressées avec cette caméra, nous avons utilisé la sortie vidéo composite du contrôleur où les images numériques sont transformées en un signal vidéo analogique. Les images ont été transformées par un convertisseur qui convertit le signal analogique en images numériques et qui envoie ces images sur une sortie *FireWire*. Une des conséquences de ces opérations est que les images obtenues sont non seulement très bruitées à cause du laser, mais aussi à cause des transformations successives qu'elles ont subies.

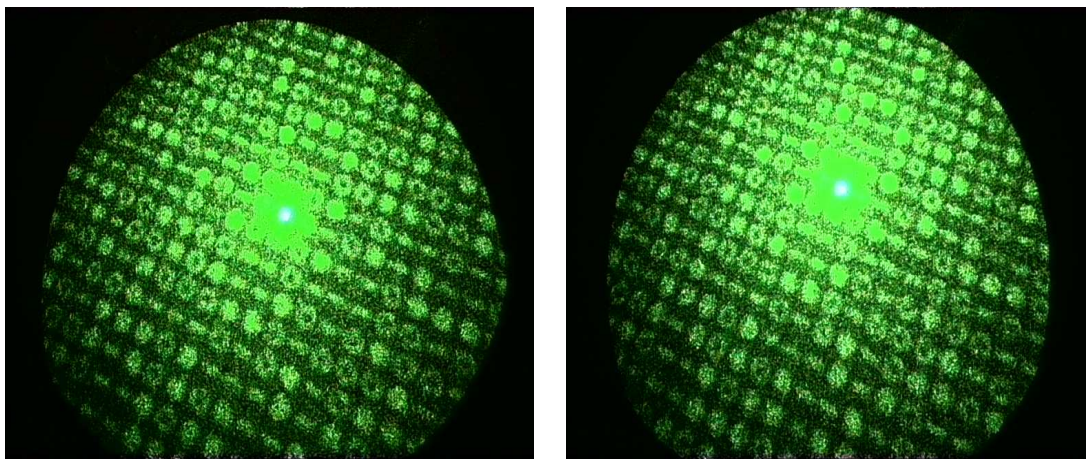


FIGURE 5.6 – Des images de la projection du motif avec le laser et l'optique diffractive. On remarque l'effet de saturation de la composante centrale et le bruit important résultant du laser.

Afin de diminuer l'effet du bruit provenant du laser, nous avons éloigné le laser de l'élément d'optique diffractive afin d'isoler le faisceau collimaté du bruit puisque ce dernier n'est pas collimaté et diverge. En faisant cela, nous avons amélioré les conditions de projection des images (Figure 5.7). Cependant, malgré que nous ayons considéré une étape de débruitage d'images avant d'appliquer nos algorithmes de traitement, la segmentation d'images a présente beaucoup d'imprécisions pour le dispositif expérimental que nous utilisons.

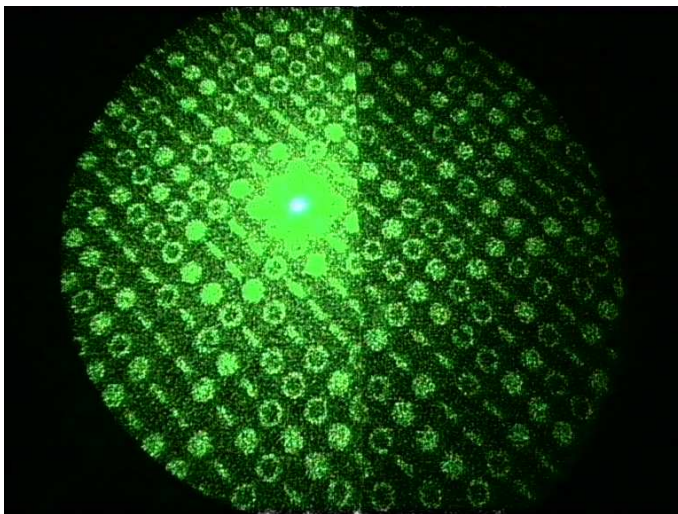


FIGURE 5.7 – Une image acquise avec notre système d'expérimentation après avoir éloigné la source laser de l'élément d'optique diffractive.

5.2.2 Étalonnage du système

Toute reconstruction 3D est généralement précédée d'un étalonnage du système. Pour notre système, nous avons appliqué la méthode d'étalonnage que nous proposons dans le troisième chapitre. Pour ce faire, nous avons utilisé un damier gris et blanc qui consiste en (12×15) carrés de 10 mm de coté. Parmi les configurations disponibles de la caméra utilisée, nous avons choisi d'acquérir des images de taille 576×768 pixels afin de mieux distinguer les primitives dans l'image.

Nous avons considéré 16 positions différentes du damier. Deux images ont été acquises à chaque position, une pour le damier (laser éteint) et l'autre de la projection du motif codé sur la surface du damier (Figure 5.10). Les images du damier sont utilisées pour calibrer la caméra de la même manière que le système expérimental. Afin d'évaluer la précision de l'étalonnage, la re-projection des données sur le plan image est déterminée et la distance avec les données de départ est calculée. Cependant, malgré le pré-traitement d'images, l'erreur moyenne de la re-projection était cette fois-ci de $(0.99548, 0.91755)$ pixel (Figure 5.8). Cette erreur est importante et nous l'attribuons à la mauvaise qualité des images acquises.

La distorsion des lentilles est présentée sur la Figure 5.12. On remarque que la distorsion est faible dans la région centrale de l'image mais elle est importante ailleurs. Cette distorsion importante est normale puisque nous utilisons ici une caméra endoscopique. Notre méthode d'éta-

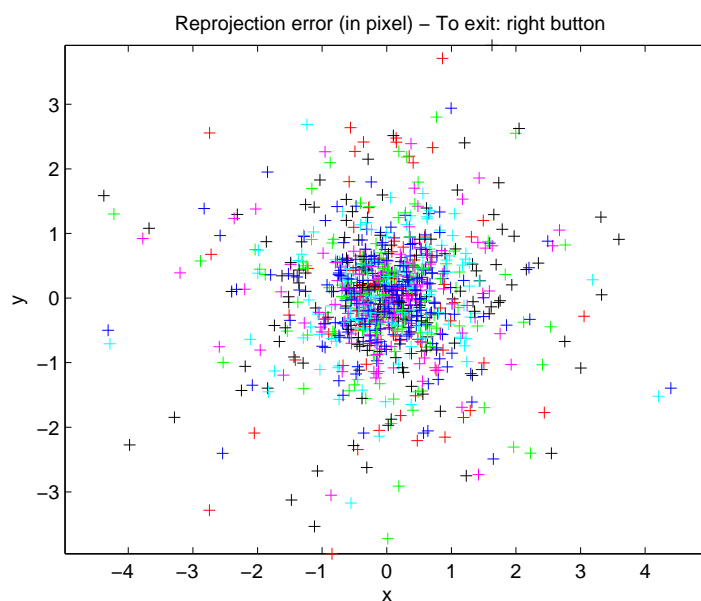


FIGURE 5.8 – Erreurs de re-projection dans les images lors de la calibration de la caméra.

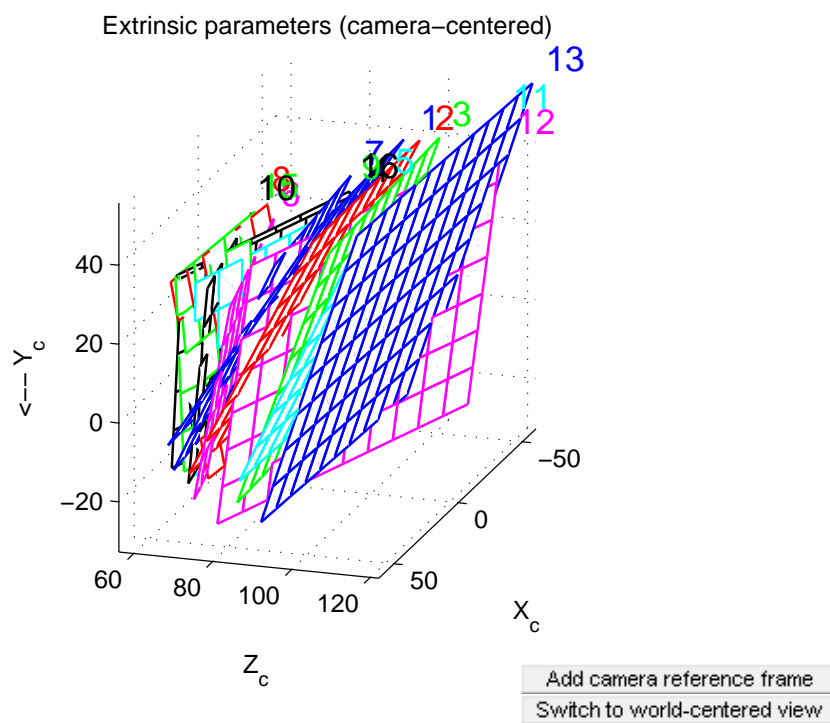


FIGURE 5.9 – Reconstruction du plan du damier pour les 16 positions considérées (Paramètres extrinsèques de la caméra).

lonnage ne prend pas en compte les distorsions afin d'accélérer le traitement mais cela n'est pas dérangeant dans la mesure où seule la partie centrale est prise en compte. Les paramètres extrinsèques sont illustrés par la reconstruction du plan du damier pour chaque position dans le repère de la caméra (Figure 5.9). Les paramètres de ces plans sont estimés afin de les utiliser dans l'étape finale de l'étalonnage.

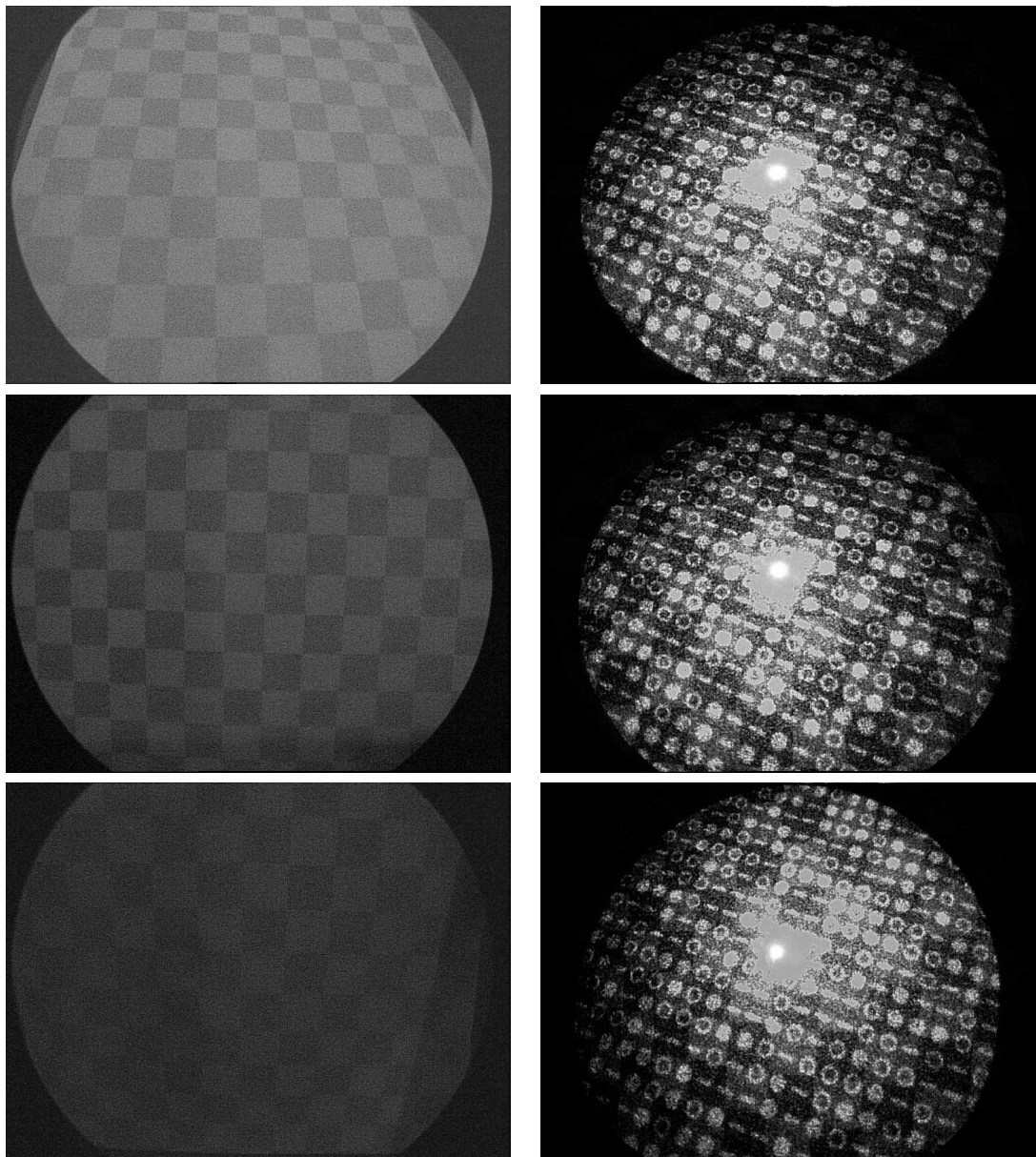


FIGURE 5.10 – Paires d'images acquises (à gauche laser éteint, à droite laser allumé).

Nous avons ensuite appliqué nos algorithmes de segmentation et de décodage aux images de la projection du motif sur le plan du damier afin de détecter et de décoder les primitives dans chaque image (Figure 5.11). Les coordonnées (x_p, y_p) des primitives sont projetées sur les plans estimés à la fin de l'étape de calibration de la caméra. Le but est de calculer la profondeur Z_c de chaque projection de primitive à chaque acquisition. Malheureusement, la mauvaise qualité des

images acquises a causé d'importantes erreurs sur l'identification et la localisation des primitives dans les images. Cela a un effet très important sur la précision de l'étalonnage.

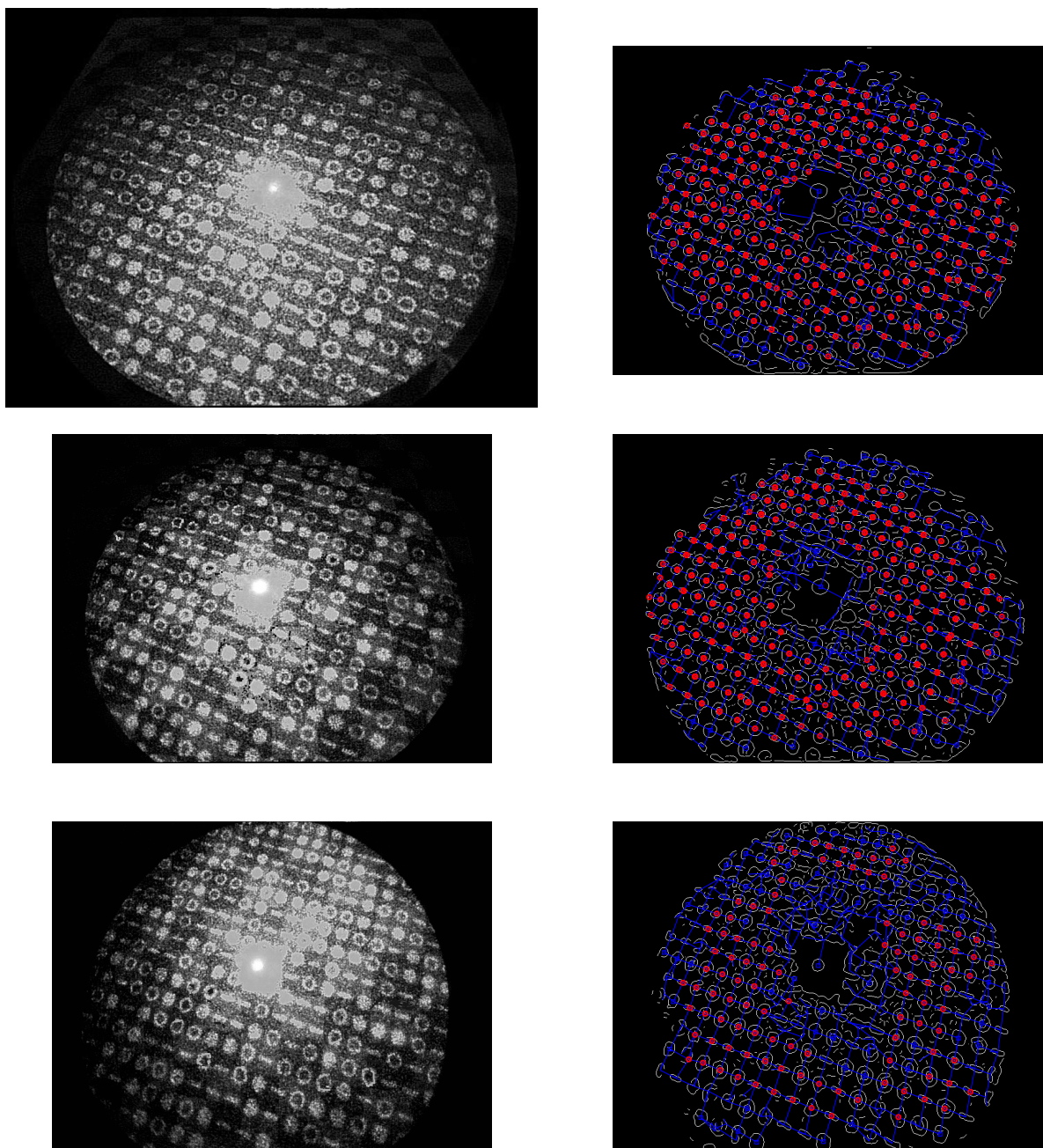


FIGURE 5.11 – Application de nos algorithmes de segmentation et de décodage aux images de la projection sur le plan du damier. On remarque la difficulté de détecter les primitives près de la composante centrale.

La relation entre la profondeur de chaque primitive et sa position dans le plan image est estimée. Dans la Figure 5.13, les cercles rouges correspondent aux points (y_p, Z_c) calculés pour

chaque primitive et pour chaque acquisition et les lignes bleues correspondent à l'estimation, pour chaque primitive, de la droite qui passe par ces points. Ces droites représentent la projection des rayons de projection des primitives dans le plan (y, Z_c) . Elles correspondent, avec les paramètres intrinsèques de la caméra, aux données de calibration du système qui seront utilisées dans la suite pour obtenir la reconstruction 3D métrique.

Malheureusement, le bruit important dans les images a causé beaucoup de mauvaises détections de primitives et par conséquent, étalonner toutes les primitives projetées a été impossible. Ceci a réduit considérablement l'efficacité de notre méthode à pouvoir reconstruire les primitives correctement décodées.

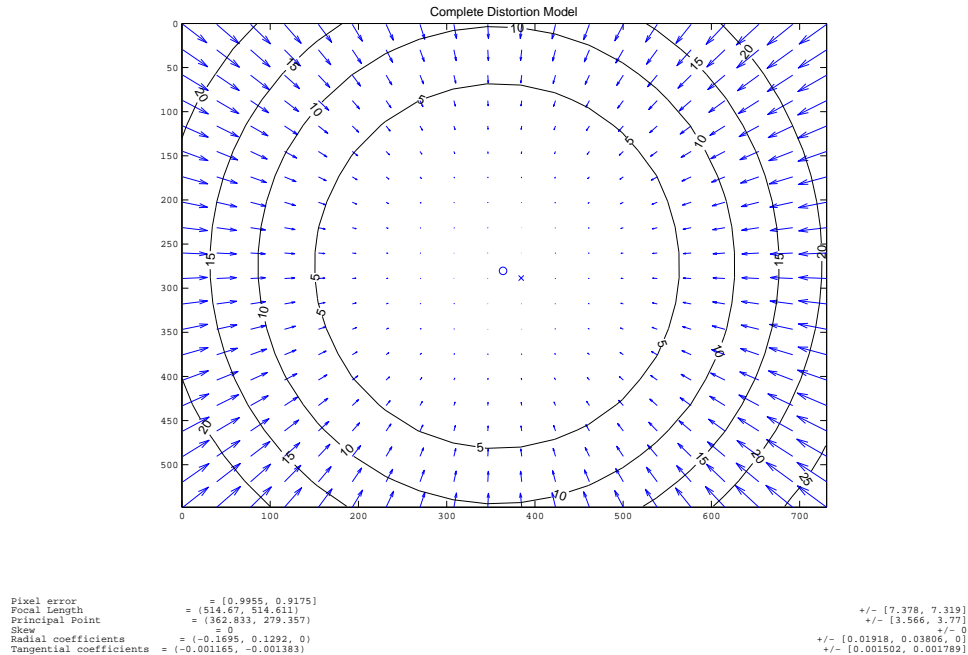


FIGURE 5.12 – Distorsion des lentilles de l'objectif de la caméra.

5.2.3 Évaluation de la reconstruction métrique

Avant de tester notre système sur des scènes abdominales réelles, nous avons d'abord évalué la précision de la reconstruction obtenue. De la même manière que nous avons opéré pour le dispositif expérimental, nous avons projeté le motif sur des surfaces planes et parallèles placées dans l'espace de travail (50 – 120 mm) (Figure 5.14-a). La distance réelle entre les surfaces est connue avec précision. Ensuite, l'image de la projection a été traitée afin de décoder les primitives comme cela est illustré par la Figure 5.14-b. Dans cette figure, les primitives correctement décodées sont représentées par des points rouges. Finalement, les données de calibration ont été utilisées pour reconstruire les plans (Figure 5.14-c). Dans cette figure, on peut constater très nettement l'influence de la sur-intensité centrale. Néanmoins, les primitives correctement décodées, autour de cette zone permettent de reconstruire correctement les points 3D correspondant. Cependant, la qualité médiocre des images a causé d'importantes erreurs lors de la localisation des primitives dans les images. Cela explique les irrégularités des surfaces obtenues. Plusieurs

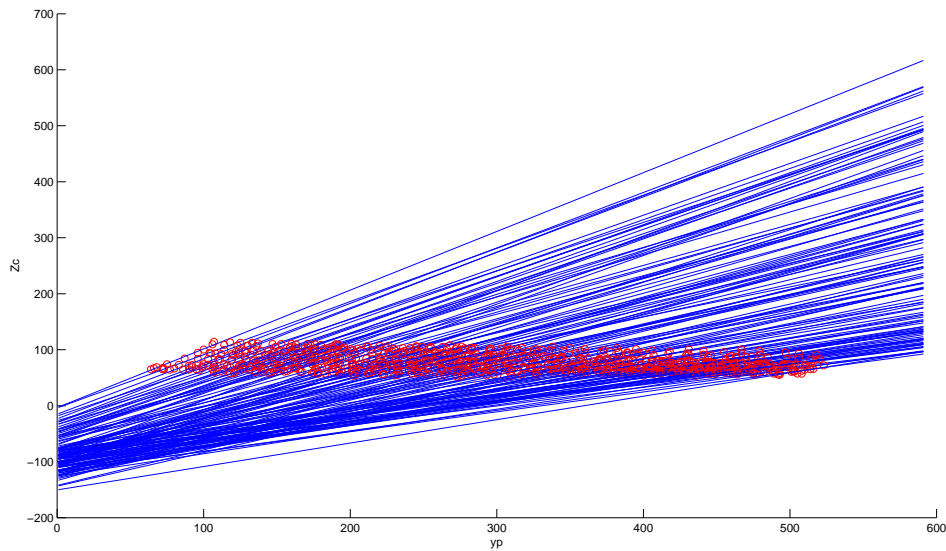


FIGURE 5.13 – Droites liant la profondeur Z_c et la coordonnée y de chaque primitive dans l'image.

acquisitions ont été effectuées de cette manière pour des différents points de vue. Les plans sont estimés d'une manière robuste à l'aide de la commande *robustfit* sous *Matlab* (Figure 5.14-d). Les erreurs de reconstruction sont illustrées par la Figure 5.15. La grandeur représentée sur cette figure est la distance entre chaque point de la surface reconstruite et la surface plane qui a été estimée. D'autre part, pour chaque acquisition, les distances et les orientations entre les plans ont été estimées.

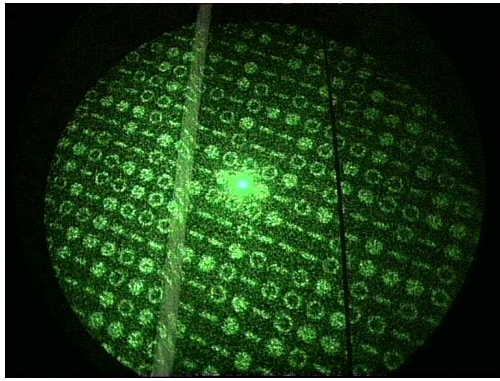
Les résultats de l'évaluation de la reconstruction sont résumés dans le tableau 5.1. La première colonne contient les couples de plans considérés, la deuxième colonne présente la distance réelle entre ces deux plans, la troisième colonne présente la distance calculée après la reconstruction et l'estimation des plans, la quatrième colonne présente l'erreur moyenne par rapport à cette estimation et à la distance réelle et la dernière colonne présente l'orientation estimée entre les plans.

Surfaces considérées	Distance réelle (mm)	Distance moyenne (mm)	Erreur moyenne (mm)	Angle (°)
1-2	16.50	15.77	1.66	1.76
1-2		13.57	3.84	1.08
1-2		14.85	2.42	1.74
2-3	16.15	13.14	3.81	2.11
2-3		14.58	2.52	2.89
2-3		14.98	2.75	2.74

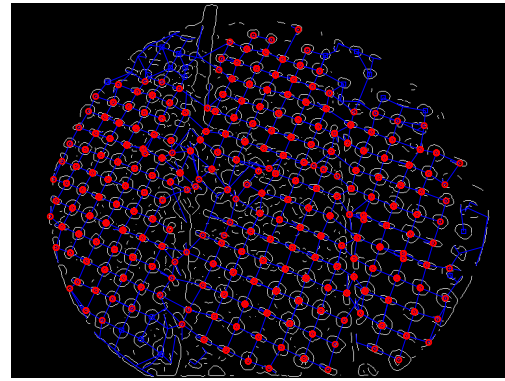
TABLE 5.1 – Résultats de la reconstruction 3D obtenues pour trois plans parallèles.

On constate malheureusement que contrairement à ce qui avait été le cas avec le système expérimental avec un vidéo projecteur, les erreurs mesurées pour notre prototype d'endoscope sont

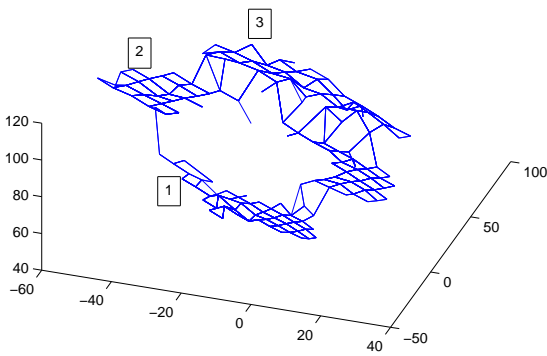
très grande comparées à l'objectif que nous nous étions fixé. L'erreur moyenne absolue est d'environ 3.5 mm dans l'intervalle de profondeur considéré et l'angle entre les plans ne dépasse pas 3° , sachant que les plans sont parallèles. Ces erreurs sont essentiellement dues aux erreurs d'étalement du dispositif. Cependant, vu les bons résultats que nous avons obtenus avec le système expérimental, nous espérons obtenir des résultats cohérents avec ceux-ci après avoir réglé les problèmes dû au bruit et dans le système d'acquisition du prototype actuel.



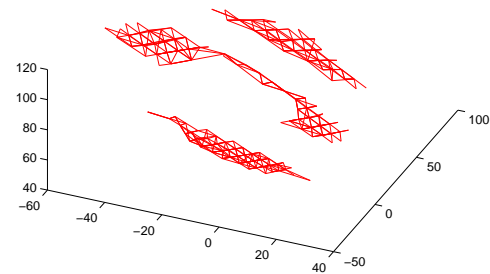
(a)



(b)



(c)



(d)

FIGURE 5.14 – (a) Projection sur trois surfaces planes et parallèles. (b) Traitement de l'image : les points rouges représentent les primitives bien décodées. (c) Reconstruction 3D de la surface. (d) Estimation des plans à l'aide de *MatLab*.

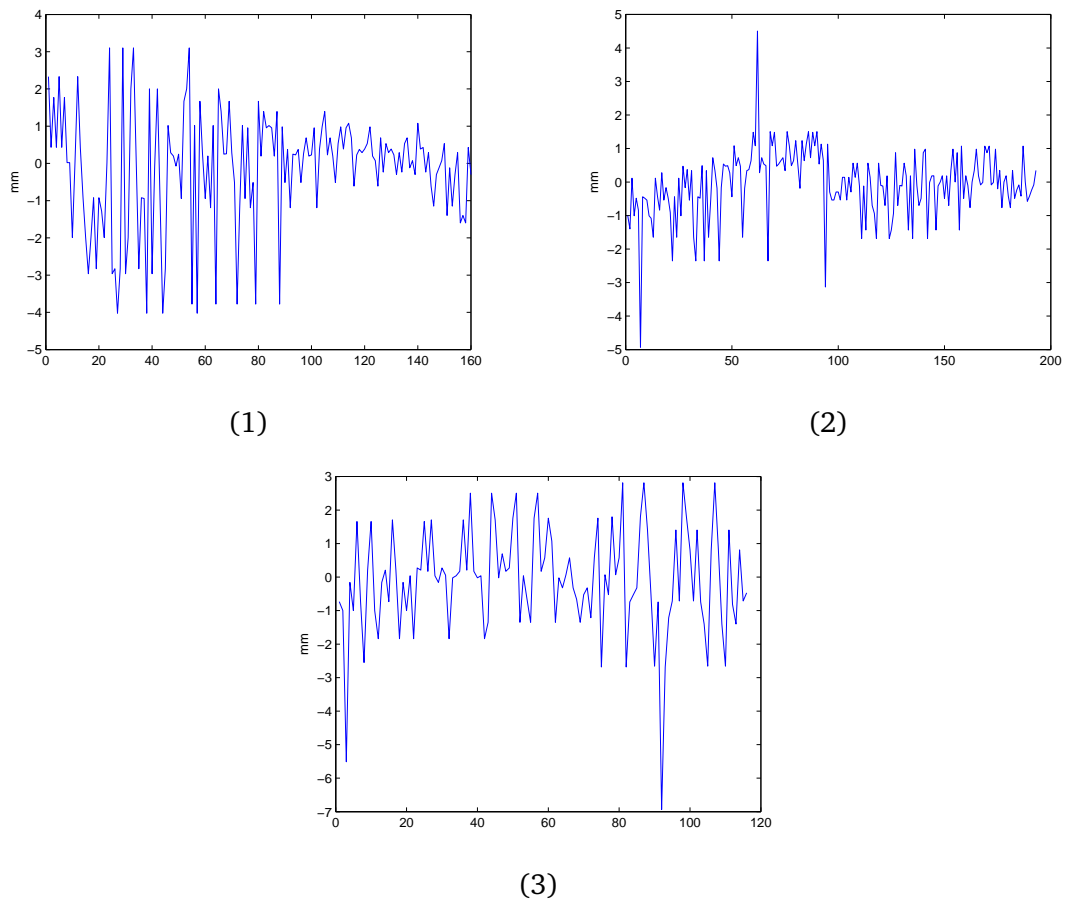


FIGURE 5.15 – Erreurs d'estimation des trois plans après la reconstruction.

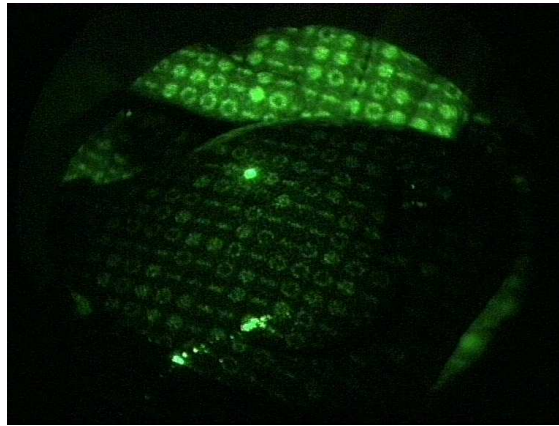
5.2.4 Reconstruction 3D in vivo

Malgré les résultats, jusque là peu satisfaisant, obtenus avec le prototype d'endoscope, nous avons souhaité valider la conception du dispositif et les aspects que nous avons traités lors des travaux dans le cadre de cette thèse. Nous avons étalonné le système comme nous l'avons détaillé précédemment. Dans le but de valider la méthode proposée ainsi que la conception envisagée en conditions réelles, à savoir sur des scènes abdominales réelles, nous avons pour cela testé le système avec des organes de l'abdomen d'un porc. Nous avons donc projeté notre motif sur les surfaces de ces organes et nous avons acquis les images de projection. Des exemples de ces acquisitions sont présentés respectivement sur les Figures 5.16-a, 5.17-a et 5.18-a. Les algorithmes de traitement ont été appliqués à ces images et les résultats sont illustrés sur les Figures 5.16-b, 5.17-b et 5.18-b. Les primitives qui ont été détectées et décodées correctement sont marquées par des taches rouges. Malgré les problèmes émanant du dispositif d'expérimentation actuel et que nous avons mentionnés, nous avons réussi à identifier plus de 60% des primitives visibles dans l'image. Les reconstructions 3D des points matérialisés par ces primitives sont illustrées sur les Figures 5.16-c, 5.17-c et 5.18-c. La qualité de ces reconstructions n'est pas aussi bonne que celle obtenue avec le système expérimental avec un vidéo projecteur. Cependant, bien qu'elle soit pas suffisante dans l'état actuel du prototype, la description obtenue des surfaces sur lesquelles le motif a été projeté est prometteuse.

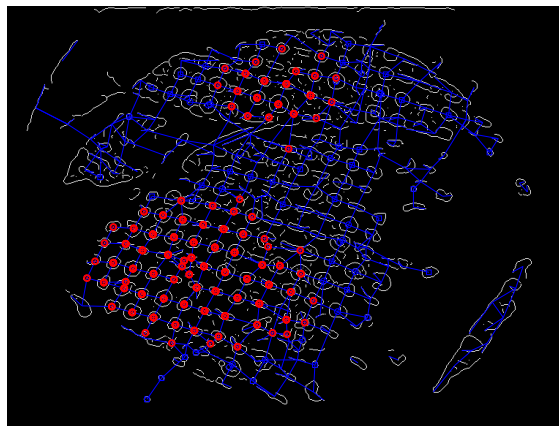
Compte tenu de la qualité médiocre des images, on peut regretter que pour les reconstructions 3D obtenues, le nombre de points pris en compte pour pouvoir exploiter correctement ces surfaces est trop faible. Ceci est essentiellement dû à la prédominance de la sur-intensité de la composante centrale. Ce problème peut être résolu en ajoutant un masque après l'élément diffractif qui servira à diminuer la puissance du laser dans cette région. Le bruit est une autre origine de la mauvaise qualité des images. Comme nous l'avons déjà mentionné, le laser que nous disposons actuellement présente un niveau de bruit important. C'est pourquoi nous l'avons éloigné de l'élément diffractif (environ 1 m) dans le but d'isoler le faisceau collimaté. Bien que cette solution ne convient pas à la conception de l'endoscope 3D envisagée, une solution possible est de remplacer le laser actuel par un autre laser pour lequel le bruit généré est plus faible.

5.3 Conclusion

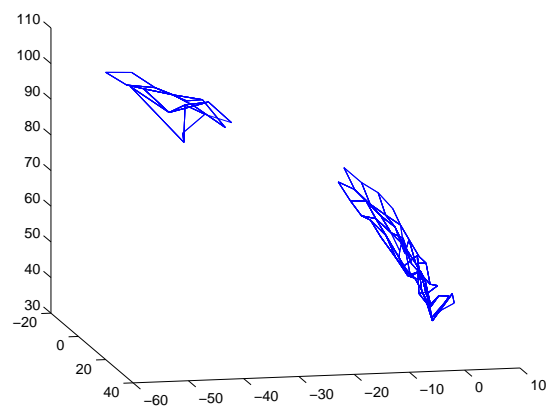
Nous avons présenté dans ce chapitre des résultats expérimentaux dans le but de valider notre méthode ainsi que la conception envisagée pour le dispositif qui met à profit cette méthode. Afin de montrer la qualité et la précision de notre méthode, nous avons commencé par exposer des résultats de reconstruction 3D obtenue avec un système expérimental à lumière structurée qui a été étalonné selon le principe décrit. Ces résultats montrent aussi l'efficacité du motif proposé ainsi que la robustesse du codage associé. Ensuite, nous avons essayé d'étalonner le prototype de l'endoscope 3D proposé afin de valider la conception envisagée du dispositif final. Malheureusement, le système actuel présente plusieurs problèmes comme le niveau de bruit dans les images acquises qui est essentiellement dû au laser et au système d'acquisition d'images. Ces problèmes ont causé d'importantes erreurs lors de l'étalonnage de la caméra et du motif. Malgré ces problèmes pratiques, nous avons tout de même évalué la précision de la reconstruction 3D obtenue dans l'état actuel du système. Comme on pouvait s'y attendre, les erreurs ont été importantes, d'environ 3.5 mm sur l'erreur moyenne absolue de profondeur et une erreur d'orientation de 3° ont été constatées. Ces erreurs sont dues à la mauvaise qualité des images acquises, à l'origine de beaucoup d'erreurs lors de la calibration de la caméra et de la segmentation des images. Finalement, nous avons testé ce système sur des scènes abdominales réelles. Pour ce faire, nous avons projeté le motif sur les surfaces d'organes de l'abdomen d'un porc et nous avons appliqué nos algorithmes de traitement. Les résultats que nous avons obtenus sont prometteurs.



(a)

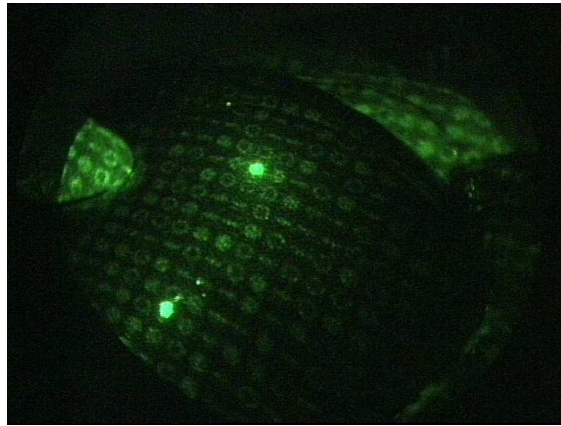


(b)

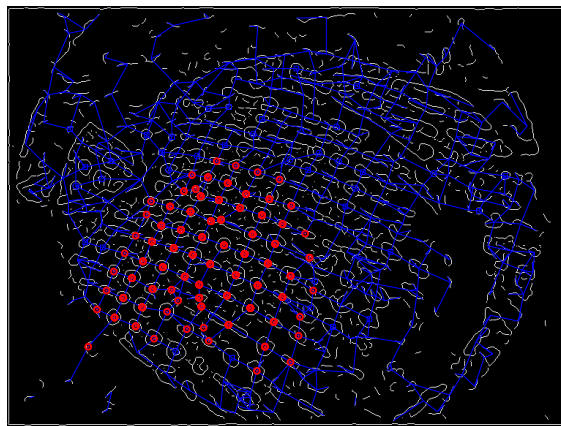


(c)

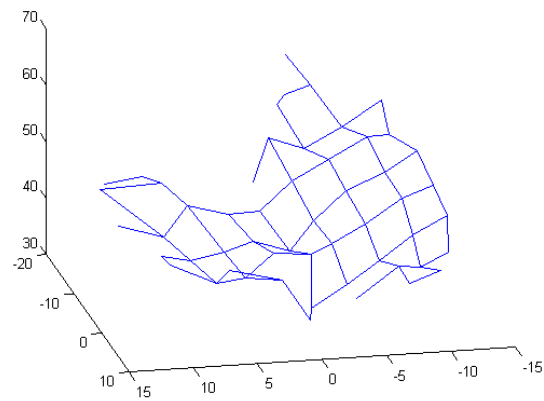
FIGURE 5.16 – (a) Projection du motif sur les organes de l'abdomen d'un porc. (b) Résultat de la segmentation et du décodage. (c) Reconstruction 3D.



(a)

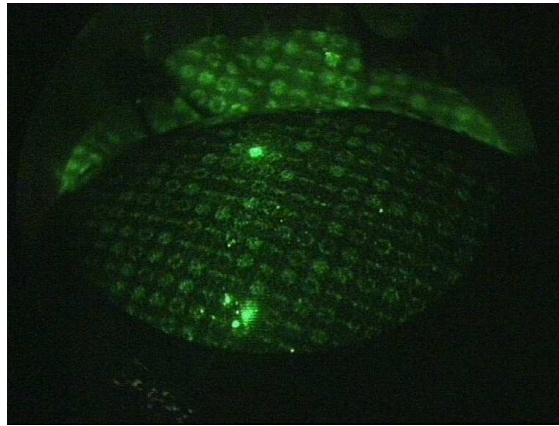


(b)

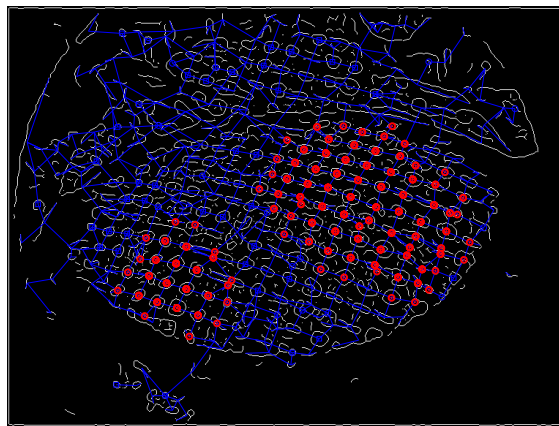


(c)

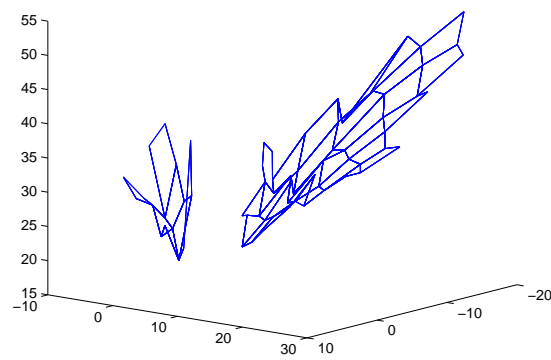
FIGURE 5.17 – Projection du motif sur les organes de l'abdomen d'un porc. (b) Résultat de la segmentation et du décodage. (c) Reconstruction 3D.



(a)



(b)



(c)

FIGURE 5.18 – Projection du motif sur les organes de l'abdomen d'un porc. (b) Résultat de la segmentation et du décodage. (c) Reconstruction 3D.

Conclusion générale

Dans le cadre des travaux menés au cours de cette thèse, nous nous sommes intéressés plus particulièrement à la conception de dispositifs endoscopiques capables de fournir des reconstructions 3D précises des structures internes de l'abdomen en temps réel. Le but poursuivi est d'utiliser ces données dans le cadre d'une assistance au geste chirurgical lors d'une intervention par voie laparoscopique. Cette assistance pourra être de nature visuelle ou robotisée. Actuellement, de tels dispositifs n'existent pas, seuls quelques systèmes permettant d'obtenir une information 3D très ponctuelle commencent seulement à apparaître.

Problématique

Parmi toutes les techniques habituellement utilisées pour reconstruire des surfaces dans l'espace, nous avons retenu, pour concevoir notre dispositif, une solution basée sur une projection de lumière structurée. Un tel système est similaire à un système stéréoscopique traditionnel pour lequel on remplace une des caméras par un projecteur de motifs lumineux. Ainsi, pour déterminer la géométrie 3D de la scène observée, on projette un motif de lumière sur la surface à reconstruire. Un certain nombre de points de la surface sont alors matérialisés par cette projection. Une image de cette projection est acquise et à l'aide des données d'étalonnage, les coordonnées 3D de ces points sont déterminées par triangulation. Cependant, ceci n'est possible qu'après avoir réalisé une mise en correspondance entre les éléments qui composent le motif projeté et leurs positions respectives dans l'image. Il est donc nécessaire d'associer un code aux primitives projetées, de manière à pouvoir les identifier dans l'image. Ce problème, lié au choix du motif utilisé et au codage qui lui est associé, est un problème récurrent, quelle que soit la technique de projection de lumière structurée employée.

Néanmoins, dans notre cas précis, ce motif doit être mis en œuvre dans un endoscope. Cela impose de nouvelles contraintes dont nous devons tenir compte. En effet, l'endoscope est appelé à être utilisé en conditions cliniques, il devra donc être d'un encombrement compatible, d'une utilisation facile pour un chirurgien et devra restituer des résultats précis dans un laps de temps très court.

Solutions et contributions

Afin d'atteindre un de nos objectifs, nous avons choisi une solution basée sur la vision active après une analyse de la vision passive et de la vision active en endoscopie. Cela nous a amené à choisir un motif lumineux à projeter ainsi qu'un codage associé. Le motif a été aussi choisi pour faciliter l'intégration du système de vision au sein d'un dispositif endoscopique. La première étape dans ce travail a été consacrée à surmonter cette difficulté par la conception d'un nouveau motif lumineux. Le motif proposé est monochrome. Cela permet de traiter des images dans lesquelles on ne trouve que deux niveaux d'intensité ce qui est généralement plus simple

que de traiter des images en couleurs. La stratégie de codage qui a été retenue est basée sur le voisinage spatial, par conséquent le code de chaque élément est une fonction de son symbole associé et de ceux de ses voisins. Le principe du codage est basé sur la théorie des matrices parfaites dans lesquelles la position de chaque élément dans la matrice est définie d'une manière non-équivoque par la fenêtre à laquelle appartient cet élément. Puisque nous avons choisi de générer un motif monochrome, l'alphabet du codage est représenté par des primitives géométriques. Afin d'accélérer le traitement sur ordinateur, nous avons considéré trois primitives géométriques simples : un cercle, un disque et un tiret. Ce dernier supporte une information directionnelle. En effet, nous nous basons sur la stratégie de voisinage spatial et par conséquent, la détection du voisinage est indispensable pour identifier les éléments. Grâce à l'information directionnelle, nous pouvons prédire les directions où se trouvent les voisins de chaque élément.

Le but étant d'incorporer le système dans un dispositif endoscopique, nous nous sommes tournés vers la technologie des optiques diffractives afin de réaliser un moyen miniaturisé de projection. Pour ce genre d'optiques, il a été montré que l'on arrive à faire un meilleur usage de l'énergie lumineuse incidente lorsque le motif présente une symétrie centrale, ce que nous avons tenu compte lors de la conception.

Enfin, vu que nous souhaitons employer ce motif avec des scènes non contrôlées telles que les scènes intra-abdominales, il nous faut définir un code garantissant une bonne robustesse. Pour ce faire, nous avons cherché à respecter une distance de *Hamming* élevée. En effet, la distance de *Hamming* est un critère bien connu en théorie de l'information qui quantifie la différence entre deux séquences de symboles.

Pour ce faire, nous avons implémenté un algorithme pour générer un motif répondant aux contraintes imposées. Finalement, nous avons obtenu une matrice de 27×29 éléments dans laquelle, chaque élément a un code de longueur 9 correspondant à une fenêtre 3×3 . Nous disposons donc à l'heure actuelle d'un motif dont la robustesse du code associé se compare favorablement avec celle des autres motifs proposés par ailleurs. En effet, le codage du motif vérifie une distance de *Hamming* moyenne supérieure à 6. À notre connaissance, la distance moyenne la plus élevée obtenue à ce jour, pour d'autres motifs, est de 3.15. Nous avons également associé des algorithmes de segmentation de décodage à ce motif. Ces algorithmes ont été testés et validés puisqu'ils ont fourni de très bons résultats même en présence d'ombres ou d'occultations. De plus, notre motif respecte les contraintes imposées pour une implémentation efficace avec des optiques diffractives, ce qui permet d'envisager son intégration dans un outil endoscopique.

Afin d'obtenir la reconstruction 3D métrique, l'étalonnage du dispositif est indispensable. C'est pourquoi nous avons développé une méthode d'étalonnage afin de décrire le comportement de la structure lumineuse projetée dans l'espace. Cette méthode utilise un objet étalon plan qui sert à la fois à calibrer la caméra et le motif projeté. Le principe consiste à projeter le motif sur une surface plane, à acquérir des images à des positions différentes et à déterminer les positions 3D des primitives sur chaque plan grâce aux paramètres d'étalonnage de la caméra. On obtient les coordonnées 3D d'un point en interpolant sa position dans l'image avec les données d'étalonnage. Cette méthode évite les distorsions de l'objectif du projecteur car ici nous considérons individuellement les rayons de projections des primitives. Le dispositif expérimental que nous avons réalisé et qui met en œuvre les techniques que nous avons adoptées a démontré que les résultats sont encourageants. En effet, l'erreur moyenne obtenue pour la position a été inférieure à 1 mm et celle de l'orientation a été inférieure à 1° . Ces précisions sont très intéressantes vu

que nous avons considéré un intervalle de profondeur compris entre 600 et 800 mm. De plus, elles nous semblent suffisantes pour exploiter la reconstruction 3D des surfaces d'organes de la cavité abdominale. Par ailleurs, on constate que pour des images de taille égale à (512×512) pixels, le temps de traitement est d'environ 65 ms. Cette durée s'approche beaucoup du temps réel vidéo et permet de visualiser correctement le suivi d'une séquence avec l'interface que nous avons développée ce qui est rarement le cas avec des systèmes existants.

Afin de mettre à profit les résultats que nous avons obtenus, nous avons conçu un outil endoscopique incorporant les aspects que nous avons validés auparavant en respectant les contraintes très spécifiques imposées par notre application. Dans un premier temps, nous avons mené une étude sur l'interaction de la lumière avec les organes afin de choisir la longueur d'onde la plus appropriée pour la projection dans un environnement tel que la cavité abdominale. Suite à cette étude, nous avons retenu la couleur verte pour illuminer le motif. Afin de réaliser un dispositif miniaturisé de projection, nous avons fait appel à la technologie des éléments d'optique diffractive. Deux générations d'éléments d'optique diffractive avec différentes configurations ont été fabriquées dans le but de sélectionner la configuration optimale pour notre application. Grâce à sa bonne efficacité et à la bonne résolution des images de projection, notre choix s'est porté sur l'élément fabriqué à quatre niveaux de phase et avec un angle d'ouverture de 20° . En effet, notre conception envisagée est basée sur l'utilisation d'un endoscope stéréoscopique. Ne disposant actuellement pas de cet endoscope, nous avons proposé un système d'expérimentation contenant deux endoscopes distincts simulant un endoscope stéréo.

Résultats et perspectives

Le motif que nous avons conçu peut être utilisé pour des applications diverses ayant besoin d'information 3D vu que son codage est robuste, que le traitement associé est rapide et que la reconstruction 3D obtenue est précise. Les résultats expérimentaux de la reconstruction 3D avec un système à lumière structurée traditionnel projetant notre motif sur des surfaces à géométrie simple et sur des surfaces plus complexes permettent d'apprécier l'efficacité de nos algorithmes employés dans le traitement des images ainsi que la qualité de la reconstruction 3D. Malheureusement, le bruit généré par le laser et le système d'acquisition a produit des erreurs importantes lors la phase d'étalonnage du système. Cela a des répercussions importantes sur la reconstruction 3D métrique obtenue avec un tel système. Cette reconstruction a présenté une erreur moyenne absolue d'environ 3.5 mm en position et une erreur d'environ 3° en orientation. Ces erreurs sont évidemment très importantes. Cependant, le fait qu'elles soient grandes en utilisant la même méthodologie qui a donné de bons résultats comme ceux qui ont été présentés plus haut avec le système traditionnel, démontre bien que les problèmes sont dus à une source de laser mal adaptée et à des optiques diffractives qui ne sont pas encore tout à fait optimisées. Ceci a été renforcé par le fait que les tests in-vivo en utilisant notre prototype sur des surfaces abdominales d'un porc a, malgré l'état actuel du prototype, donné des résultats prometteurs. Malgré ces conditions de travail, nous avons obtenu une bonne description des surfaces qui se situaient dans un espace d'une profondeur comprise entre 20 et 120 mm.

Les perspectives de ce travail sont très larges. À court terme, il faut notamment résoudre les problèmes du système d'expérimentation en choisissant un autre laser qui présente moins de bruit que celui que nous disposons actuellement, et en trouvant une autre solution pour acquérir les images avec la caméra endoscopique afin d'améliorer la qualité des images acquises. De plus, il faut résoudre le problème de la composante centrale qui a causé des saturations dans l'image.

Une solution possible est d'ajouter un masque après l'élément d'optique diffractive adapté à la distribution de la puissance du laser. Une fois ces problèmes pratiques réglés, nous souhaitons réévaluer la reconstruction 3D fournie par un système qui aura été dimensionné et réalisé de telle sorte que l'on puisse complètement le valider dans des conditions réelles.

Mais d'autres pistes sont encore à exploiter. Notamment, celle concernant la conception du motif. En effet, bien que très performant, notre motif est d'une part unique, et d'autre part, présente une robustesse qui peut encore être améliorée. Des travaux dans cette direction ont déjà été entamés. L'utilisation des données 3D fournies par cet instrument reste aussi à exploiter, que ce soit en utilisant la réalité augmentée ou en faisant appel à des dispositifs robotisés pour l'assistance au geste chirurgical. Quoi qu'il en soit, le travail réalisé au cours de cette thèse aura permis de contribuer à l'amélioration du geste chirurgical et permettrait, je l'espère, de contribuer aussi à un plus grand confort et une meilleure sécurité pour toutes les personnes appelées à être opérées ainsi qu'aux chirurgiens.

Liste des publications

– Conférences Internationales

1. Fast 3D vision with robust structured light coding .
C. Albitar, P. Graebling, C. Doignon. *Visualization, Image-guided Procedures and Modeling Conference, SPIE Medical Imaging 2009* , Orlando, USA, Février 2009.
2. Position-Based Visual Servoing Using a Coded Structured Light Sensor.
T. Heitzmann, C. Doignon, **C. Albitar** and P. Graebling. *IEEE International Workshop on Robotic and Sensors Environments*, Ottawa, Canada, Octobre 2008.
3. Robust structured light coding for 3D reconstruction.
C. Albitar, P. Graebling, C. Doignon. *IEEE International Conference on Computer Vision, ICCV'07* , Rio de Janeiro, Brazil, Octobre 2007.
4. Design of a monochromatic pattern for a robust structured light coding.
C. Albitar, P. Graebling, C. Doignon. *IEEE International Conference on Image Processing, ICIP'07* , San antonio, USA, Septembre 2007.

– Conférences Nationales

1. Reconstruction en temps réel de surfaces par vision active codée.
C. Albitar, P. Graebling, C. Doignon. *16ème congrès francophone de Reconnaissance des Formes et Intelligence Artificielle, RFIA'08* , Amiens, France, Janvier 2008.
2. Décodage en temps réel de la lumière structurée codée pour les environnements complexes.
P. Graebling, **C. Albitar**, C. Doignon. *Métrie 3D par vision, GdR ISIS*, Lyon, France, Mars 2008.
3. Reconstruction 3D de surfaces corporelles pour le guidage d'instruments en chirurgie robotisée. (Conférence invité)
P. Graebling, C. Doignon, **C. Albitar** and A. Ayadi. *4th International Conference on Computer Integrated Manufacturing, CIP'07*, Sétif, Algérie, Novembre 2007.
4. Décodage en temps réel de la lumière structurée codée pour des applications robotiques dans les environnements complexes.
C. Albitar, P. Graebling, C. Doignon. *Journées Nationales de Recherche en Robotique, JNRR'07*, Obernai, France, Octobre 2007.
5. Conception d'un motif monochrome pour le codage robuste de la lumière structurée.
C. Albitar, P. Graebling, C. Doignon. *Onzième congrès francophone des jeunes chercheurs en vision par ordinateur, ORASIS'07* , Obernai, France, Juin 2007.

Bibliographie

- [Adan04] A. Adan, F. Molina, et L. Morena. «Disordered patterns projection for 3D motion recovering». In *Proceedings of the 2nd Int. Symp. on 3D Data Processing, Visualization and Transmission, 3DPVT'04*, Vol. 33, n° 6, pages 262–269, 2004.
- [Angelopoulou99] E. Angelopoulou. «The reflectance spectrum of human skin». Rapport technique No. MS-CIS-99-29, University of Pennsylvania, Department of Computer and Information Science, 1999.
- [Armangue03] X. Armangue et J. Salvi. «Overall view regarding fundamental matrix estimation». *Image and Vision Computing, IVC*, Vol. 21, n° 2, pages 205–220, 2003.
- [Burschka04a] D. Burschka et G. Hager. «Vision-based inertial system for mobile robots». *IEEE International Conference on Robotics and Automation "ICRA"*, Vol. 1, pages 409–415, 2004.
- [Burschka04b] D. Burschka, M. Li, M. Ishii, R. H. Taylor, et G. D. Hager. «Scale-Invariant Registration of Monocular Endoscopic Images to CT-Scans for Sinus Surgery». pages 413–421, 2004.
- [Canny86] F. J. Canny. «A computational approach to edge detection». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, n° 6, pages 679–698, 1986.
- [Carrhill85] B. Carrhill et R. A. Hummel. «Experiments with the Intensity Ratio Data Sensor». *Computer Vision Graphics and Image Processing*, Vol. 32, n° 3, pages 337–358, December 1985.
- [Caspi98] D. Caspi, N. Kiryati, et J. Shamir. «Range Imaging with Adaptive Color Structured Light». *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 20, n° 5, pages 470–480, 1998.
- [Chan03] M. Chan, W. Lin, C. Zhou, et J. Y. Qu. «Miniaturized Three-Dimensional Endoscopic Imaging System Based on Active Stereovision». *Applied Optics*, Vol. 42, n° 10, pages 1888–1898, 2003.
- [Chen87] C. Chen et A. Kak. «Modeling and calibration of a structured light scanner for 3-D robot vision». *IEEE International Conference on Robotics and Automation, ICRA*, 1987.
- [Chen97] C. S. Chen, Y. P. Hung, C. C. Chiang, et J. L. Wu. «Range Data-Acquisition Using Color Structured Lighting and Stereo Vision». *Image and Vision Computing*, Vol. 15, n° 6, pages 445–456, June 1997.

- [Chen02] S. Chen et Y. Li. «Self recalibration of a structured light vision system from a single view». *IEEE International Conference on Robotics and Automation, ICRA'02*, Vol. 3, pages 2539–2544, 2002.
- [Chen03] S. Y. Chen et Y. F. Li. «Self-recalibration of a colour-encoded light system for automated 3-D measurements». *Measurement Science and Technology*, Vol. 14, n° 1, 2003.
- [Chen07] S. Chen, Y. Li, et J. Zhang. «Realtime Structured Light Vision with the Principle of Unique Color Codes». *IEEE International Conference on Robotics and Automation, ICRA'07*, pages 429–434, 2007.
- [Choset01] H. Choset et K. Nagatani. «Topological simultaneous localization and mapping (slam) : Toward exact localization without explicit localization». *IEEE Transactions on Robotics and Automation*, Vol. 17, pages 125–137, 2001.
- [Dai02] C.-L. Dai, H.-L. Chen, C.-Y. Lee, et P.-Z. Chang. «Fabrication of diffractive optical elements using the CMOS process». *Journal of Micromechanics and Microengineering*, Vol. 12, n° 1, pages 21–25, 2002.
- [Davies98] C. J. Davies et M. S. Nixon. «A Hough transform for detecting the location and orientation of 3-dimensional surfaces via color encoded spots». *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 28, n° 1, pages 90–95, 1998.
- [Devernay01] F. Devernay, F. Mourgues, et E. Coste-Maniere. «Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery». *Proceedings of the International Workshop on Medical Imaging and Augmented Reality*, pages 16–20, 2001.
- [Dipanda05] A. Dipanda et S. Woo. «Towards a real-time 3D shape reconstruction using a structured light system». *Pattern Recognition*, Vol. 38, n° 10, pages 1632–1650, 2005.
- [Doignon99] C. Doignon et G. Abba. «A practical multi-plane method for a low-cost camera calibration technique». *In Proceedings of the European Control Conference, ECC'99*, 1999.
- [Duda73] R. O. Duda et P. E. Hart. *Pattern Classification and Scene Analysis*. Wiley interscience publication, New-York, 1973.
- [Durdle98] N. Durdle, J. Thayyoor, et V. Raso. «An improved structured light technique for surface reconstruction of the human trunk». *IEEE Canadian Conference on Electrical and Computer Engineering*, Vol. 2, pages 874–877, 1998.
- [El-Hakim95] S. F. El-Hakim, J. A. Beraldin, et F. Blais. «A Comparative Evaluation of the Performance of Passive and Active 3-D Vision Systems». *SPIE Proceedings, Conference on Digital Photogrammetry*, Vol. 2646, pages 14–25, 1995.
- [Faugeras86] O. D. Faugeras et G. Toscani. «The Calibration Problem for Stereo». *IEEE Computer Vision and Pattern Recognition*, pages 15–20, 1986.
- [Faugeras93] O. Faugeras. *Three-dimensional computer vision : a geometric viewpoint*. MIT Press, 1993.
- [Fofi04] D. Fofi, T. Sliwa, et Y. Voisin. «A Comparative Survey on Invisible Structured Light». *proceedings of SPIE on Machine vision applications in industrial inspection XII*, Vol. 5303, pages 90–98, 2004.

-
- [Forsyth91] D. Forsyth, J. Mundy, A. Zisserman, C. Coelho, A. Heller, et C. Rothwell. «Invariant Descriptors for 3-D object Recognition and Pose». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, n° 10, pages 971–991, 1991.
- [Goodman70] J. W. Goodman et A. M. Silvestri. «Some effects of Fourier domain phase quantization». *IBM Journal of Research and Development*, Vol. 14, pages 478–484, 1970.
- [Graebling02] P. Graebling, A. Lallement, D. Y. Zhou, et E. Hirsch. «Optical high-precision three-dimensional vision-based quality control of manufactured parts by use of synthetic images and knowledge for image-data evaluation and interpretation». *Applied Optics*, Vol. 41, n° 14, pages 2627–2643, 2002.
- [Griffin92] P. Griffin, L. Narasimhan, et S. Yee. «Generation of uniquely encoded light patterns for range data acquisition». *Pattern Recognition*, Vol. 25, n° 6, pages 609–616, 1992.
- [Gühring01] J. Gühring. «Dense 3D surface acquisition by structured light using off-the-shelf components». *Proceedings of Videometrics and Optical Methods for 3D Shape Measurement, SPIE*, Vol. 4309, pages 220–231, 2001.
- [Hall Holt01] O. Hall Holt et S. Rusinkiewicz. «Stripe Boundary Codes for Real-Time Structured-Light Range Scanning of Moving Objects». *International Conference on Computer Vision, ICCV'01*, Vol. II, pages 359–366, 2001.
- [Hall82] E. L. Hall, J. B. K. Tio, C. A. McPherson, et F. A. Sadjadi. «Measuring Curved Surfaces for Robot Vision». *Computer Journal*, Vol. 15, n° 12, pages 42–54, 1982.
- [Hartley94] R. I. Hartley. «Euclidean Reconstruction from Uncalibrated Views». pages 237–256, 1994.
- [Hartley00] R. Hartley et A. Zisserman. *Multiple view geometry in computer vision*. Cambridge Univ. Press, 2000.
- [Hasegawa01] K. Hasegawa et Y. Sato. «Endoscope system for high-speed 3D measurement». *Systems and Computers in Japan*, Vol. 32, n° 8, pages 30–39, 2001.
- [Hayashibe01] M. Hayashibe et Y. Nakamura. «Laser-pointing endoscope system for intraoperative 3D geometric registration». *Proceedings of IEEE International Conference on Robotics and Automation, ICRA*, Vol. 2, pages 1543–1548, 2001.
- [Hayashibe06] M. Hayashibe, N. Suzuki, et Y. Nakamura. «Laser-scan endoscope system for intraoperative geometry acquisition and surgical robot safety management». *Medical Image Analysis, Special Issue on Functional Imaging and Modelling of the Heart*, Vol. 10, n° 4, pages 509–519, 2006.
- [Heikkila97] J. Heikkila et O. Silven. «A Four-Step Camera Calibration Procedure with Implicit Image Correction». pages 1106–1112, 1997.
- [Hoff89] W. Hoff et N. Ahuja. «Surfaces from stereo : integrating feature matching, disparity estimation, and contour detection». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, n° 2, pages 121–136, 1989.

- [Hu07] M. Hu, G. Penney, P. Edwards, et D. Hawkes. «3D Reconstruction of internal organ surfaces for minimal invasive surgery». *MICCAI*, pages 68–77, 2007.
- [Hugli89] H. Hugli et G. Maitre. «Generation and use of color pseudorandom sequences for coding structured light in active ranging». *Proceeding of Industrial Inspection, SPIE*, Vol. 1010, 1989.
- [Huynh99] D. Q. Huynh, R. A. Owens, et P. E. Hartmann. «Calibrating a Structured Light Stripe System : A Novel Approach». *International Journal of Computer Vision*, Vol. 33, n° 1, pages 73–86, September 1999.
- [Inokuchi84] S. Inokuchi, K. Sato, et F. Matsuda. «Range Imaging System for 3-D Object Recognition». *International Conference on Pattern Recognition*, pages 806–808, 1984.
- [Ito91] M. Ito. «Robot vision modelling—camera modelling and camera calibration». *Advanced Robotics*, Vol. 5, n° 3, pages 321–335, 1991.
- [Ito95] M. Ito et A. Ishii. «A three-level checkerboard pattern projection method for curved surface measurement». *Pattern Recognition*, Vol. 28, n° 1, pages 27–40, 1995.
- [Keller00] K. Keller et J. Ackerman. «Real-time Structured Light Depth Extraction». *proceedings of SPIE on Three Dimensional Image Capture and Applications III*, Vol. 3958, pages 11–18, 2000.
- [Kiyasu95] S. Kiyasu, H. Hoshino, K. Yano, et S. Fujimura. «Measurement of the 3-D shape of specular polyhedrons using an m-array coded light source». *IEEE Transaction on Instrumentation and Measurement*, Vol. 44, n° 3, pages 775–778, 1995.
- [Klette98] R. Klette, A. Koschan, et K. Schluns. *Computer Vision : Three-Dimensional Data from Images*. Springer-Verlag, 1998.
- [Kololuonma02] T. Kololuonma, K. Kataja, S. Juuso, et J. Aikio. «Fabrication and characterization of hybrid-glass-based axicons». *Optical Engineering*, Vol. 41, n° 12, pages 3136–3140, 2002.
- [Kress97] B. Kress. *Optimisation du calcul et de la fabrication par microlithographie d'optiques diffractives numériques pour leur intégration dans des micro-systèmes. Une approche de CAO/CFAO*. Thèse de doctorat, Université de Strasbourg I (Louis Pasteur), France, 1997.
- [Krupa03] A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel, J. Leroy, L. Soler, et J. Marescaux. «Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing». *IEEE Transactions on Robotics and Automation*, Vol. 19, n° 5, pages 842–853, 2003.
- [Lau04] W. Lau, N. Ramey, J. Corso, N. Thakor, et D. Hager. «Stereo-Based Endoscopic Tracking of Cardiac Surface Deformation». *MICCAI*, Vol. 3217, pages 494–501, 2004.
- [Long06] J. A. Long, J. L. Descotes, O. Skowron, J. Troccaz, P. Cinquin, B. Boillot, N. Terrier, et J. J. Rambeaud. «Utilisation de la robotique en chirurgie laparoscopique urologique : état de l'art». *Progrès en urologie*, Vol. 16, n° 1, pages 3–11, 2006.

-
- [Longuet-Higgins81] C. Longuet-Higgins. «A computer algorithm for reconstructing a scene from two projections». *Nature*, Vol. 293, pages 61–62, 1981.
- [MacWilliams76] F. J. MacWilliams et N. J. A. Sloane. «Pseudo-random sequences and arrays». *Proceedings of the IEEE*, Vol. 64, n° 12, pages 1715–1729, 1976.
- [Marescaux01] J. Marescaux, J. Leroy, M. Gagner, F. Rubino, D. Mutter, M. Vix, S. E. Butner, et M. K. Smith. «Transatlantic robot-assisted telesurgery». *Nature*, Vol. 413, pages 379–380, 2001.
- [Marescaux02] J. Marescaux, J. Leroy, F. Rubino, M. K. Smith, M. Vix, M. Simone, et D. Mutter. «Transcontinental robot-assisted remote telesurgery : feasibility and potential applications». *Annals of Surgery*, Vol. 235, n° 4, pages 487–492, 2002.
- [Maruyama93] M. Maruyama et S. Abe. «Range Sensing by Projecting Multiple Slits with Random Cuts». *Pattern Analysis and Machine Intelligence*, Vol. 15, n° 6, pages 647–651, June 1993.
- [Marzani02] F. Marzani, Y. Voisin, L. L. Y. Voon, et A. Diou. «Calibration of a three-dimensional reconstruction system using a structured light source». *Optical Engineering*, Vol. 41, n° 2, pages 484–492, 2002.
- [Maybank92] S.-J. Maybank. «The cross-ratio and the j-invariant». *Geometric Invariance in Computer Vision*, pages 107–109, 1992.
- [Minou81] M. Minou, T. Kanade, et T. Sakai. «A Method of Time-Coded Parallel Planes of Light for Depth Measurement». *Trans. Institute of Electronics and Communication Engineers of Japan*, Vol. E64, n° 8, pages 521–528, 1981.
- [Monks93] T. P. Monks et J. N. Carter. «Improved stripe matching for colour encoded structured light». *Fifth International Conference on Computer Analysis of Images and Patterns*, pages 476–485, 1993.
- [Morano98] R. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, et J. Nissanov. «Structured Light Using Pseudo-random Codes». *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 20, n° 3, pages 322–327, 1998.
- [Morin93] L. Morin. *Quelques contributions des invariants projectifs à la vision par ordinateur*. Thèse de doctorat, Institut National Polytechnique de Grenoble (LIFIA), 1993.
- [Morita88] H. Morita, K. Yajima, et S. Sakata. «Reconstruction of Surfaces of 3-D Objects by M-array Pattern Projection Method». *International Conference on Computer Vision, ICCV88*, pages 468–473, 1988.
- [Mountney06] P. Mountney, D. Stoyanov, A. J. Davison, et G. Z. Yang. «Simultaneous Stereoscope Localization and Soft-Tissue Mapping for Minimal Invasive Surgery.» *International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI'06*, Vol. 4190, pages 347–354, 2006.
- [Petriu92] E. M. Petriu, T. Bieseman, N. Trif, W. S. McMath, et S. K. Yeung. «Visual object recognition using pseudo-random grid encoding». *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1617–1624, 1992.
- [Petriu00] E. M. Petriu, Z. Sakr, H. Spoelder, et A. Moica. «Object recognition using pseudo-random color encoded structured light». *In Proceedings of the*

- 17th IEEE Instrumentation and Measurement technology Conference, Vol. 3, pages 1237–1241, 2000.
- [Pichler93] C. V. Pichler, K. Radenacher, V. Grablowitz, W. Boekmann, G. Rau, G. Jakse, et V. Schumpelick. «An ergonomic analysis of stereo-video-endoseopy». *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1408–1409, 1993.
- [Posdamer82] J. L. Posdamer et M. D. Altschuler. «Surface Measurement by Space-Encoded Projected Beam System». *Computer Graphics and Image Processing*, Vol. 18, n° 1, pages 1–17, 1982.
- [Reid96] I. D. Reid. «Projective Calibration Of A Laser-Stripe Range Finder». *Image and Vision Computing*, Vol. 14, pages 659–666, 1996.
- [Ribo05] M. Ribo et M. Brandner. «State of the art on vision-based structured light systems for 3D measurements». *International Workshop on Robotic and Sensor Environments*, pages 2–6, 2005.
- [Rocchini01] C. Rocchini, P. Cignoni, C. Montani, P. Pingi, et R. Scopigno. «A low cost 3D scanner based on structured light». *EUROGRAPHICS*, Vol. 20, n° 3, pages 299–308, 2001.
- [Salvi98] J. Salvi, J. Batlle, et E. Mouaddib. «A robust coded pattern projection for dynamic 3d scene measurement». *Int. Journal of Pattern Recognition Letters*, Vol. 19, pages 1055–1065, 1998.
- [Salvi02] J. Salvi, X. Armangu, et J. Batlle. «A comparative review of camera calibrating methods with accuracy evaluation». *Pattern Recognition*, Vol. 35, n° 7, pages 1617–1635, 2002.
- [Salvi04] J. Salvi, J. Pagès, et J. Batlle. «Pattern codification strategies in structured light systems». *Pattern Recognition*, Vol. 37, n° 4, pages 827–849, 2004.
- [Sambongi00] M. Sambongi, M. Igarashi, T. Obi, M. Yamaguchi, N. Oyama, M. Kobayashi, Y. Sano, S. Yoshida, et K. Gono. «Analysis of spectral reflectance of mucous membrane for endoscopic diagnosis». *In Proceedings of the 22nd Annual Engineering in Medicine and Biology Society International Conference*, pages 1026–1029, 2000.
- [Sato96] K. Sato. «Range Imaging Based on Moving Pattern Light and Spatio-Temporal Matched Filter». *International Conference on Image Processing, ICIP'96*, Vol. I, pages 33–36, 1996.
- [Sato99] T. Sato. «Multispectral Pattern Projection Range Finder». *Three-Dimensional Image Capture and Applications*, pages 28–37, 1999.
- [Seitz] S. M. Seitz. «An Overview of Passive Vision Techniques». *Course Notes, The Robotics Institute, Carnegie Mellon University*, pages 1–3.
- [Semple52] J. G. Semple et G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1952.
- [Spoelder00] H. J. W. Spoelder, F. M. Vos, E. M. Petriu, et F. C. A. Groen. «Some aspects of pseudo random binary array based surface characterization». *IEEE Transaction on Instrumentation and Measurement*, Vol. 49, n° 6, pages 1331–1336, 2000.

-
- [Stewart96] C. V. Stewart, R. Y. Flatland, et K. Bubna. «Geometric constraints and stereo disparity computation». *International Journal on Computer Vision*, Vol. 20, n° 3, pages 143–168, 1996.
- [Stoyanov04] D. Stoyanov, A. Darzi, et G. Z. Yang. «Dense 3D Depth Recovery for Soft Tissue Deformation During Robotically Assisted Laparoscopic Surgery». *International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI'04*, pages 41–48, 2004.
- [Street00] G. S. B. Street. «Stereo-Endoscope». *United States Patent No. 6104426*, issued 08/15/2000.
- [Swain93] M. J. Swain et M. A. Stricker. «Promising directions in active vision». *International Journal on Computer Vision*, Vol. 11, n° 2, pages 109–126, 1993.
- [Swanson89] G. J. Swanson et W. B. Veldkamp. «Diffractive optical elements for use in infrared systems». *Optical Engineering*, Vol. 28, n° 6, pages 605–608, 1989.
- [Swanson93] G. J. Swanson et W. B. Veldkamp. «High-efficiency, multilevel, diffractive optical elements». *United States Patent NO. 5218471*, 1993.
- [Tajima90] J. Tajima et M. Iwakawa. «3-D data acquisition by Rainbow Range Finder». *Proceedings of the 10th International Conference on Pattern Recognition*, Vol. I, pages 309–313, 1990.
- [Tanigami89] M. Tanigami, S. Ogata, S. Aoyama, T. Yamashita, et K. Imanaka. «Low-wavefront aberration and high-temperature stability molded micro Fresnel lens». *IEEE Photonics Technology Letters*, Vol. 1, n° 11, pages 384–385, 1989.
- [Tomasi92] C. Tomasi et T. Kanade. «Shape and Motion from Image Streams under Orthography : a Factorization Method». *International Journal of Computer Vision*, Vol. 9, n° 2, pages 137–154, 1992.
- [Trucco98] E. Trucco et A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [Tsai87] R. Y. Tsai. «A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses». *Robotics and Automation*, Vol. 3, n° 4, pages 323–344, 1987.
- [Valkenburg98] R. J. Valkenburg et A. M. McIvor. «Accurate 3d Measurement Using A Structured Light System». *Image and Vision Computing*, Vol. 16, n° 2, pages 99–110, February 1998.
- [Vuylsteke90] P. Vuylsteke et A. Oosterlinck. «Range Image Acquisition with a Single Binary-Encoded Light Pattern». *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, n° 2, pages 148–164, 1990.
- [Wang02] C. Wang, Y. C. Chan, et Y. L. Lam. «Fabrication of diffractive optical elements with arbitrary surface-relief profile by direct laser writing». *Optical Engineering*, Vol. 41, n° 6, pages 1240–1245, 2002.
- [Weng92] J. Weng, P. Cohen, et M. Herniou. «Camera Calibration with Distortion Models and Accuracy Evaluation». *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, n° 10, pages 965–980, October 1992.

- [Wengert06] C. Wengert, M. Reeff, P. C. Cattin, et G. Szkely. «Fully Automatic Endoscope Calibration for Intraoperative Use». in *Bildverarbeitung für die Medizin*, pages 19 – 23, 2006.
- [Woodham80] R. J. Woodham. «Photometric method for determining surface orientation from multiple images». *Journal of Optical Engineering*, Vol. 19, n° 1, pages 138–144, 1980.
- [Wu07] C. H. Wu, Y. N. Sun, et C. C. Chang. «Three-Dimensional Modeling From Endoscopic Video Using Geometric Constraints Via Feature Positioning». *IEEE Transactions on Biomedical Engineering*, Vol. 54, n° 7, pages 1199–1211, 2007.
- [Yee94] S. Yee et P. Griffin. «Three-dimensional Imaging System». *Optical Engineering*, Vol. 33, n° 6, pages 2070–2075, 1994.
- [Zhang00] Z. Y. Zhang. «A Flexible New Technique for Camera Calibration». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, n° 11, 2000.
- [Zhang02] L. Zhang, B. Curless, et S. M. Seitz. «Rapid shape acquisition using color structured light and multi-pass dynamic programming». pages 24–36, 2002.
- [Zhang04] Z. Zhang. *Emerging Topics in Computer Vision*, chapitre 2. Prentice Hall Professional Technical Reference, 2004.
- [Zhang06] S. Zhang et P. S. Huang. «Novel method for structured light system calibration». *Optical Engineering*, Vol. 45, n° 8, pages 083601–, 2006.
- [Zhou04] F. Q. A. Zhou et G. J. Zhang. «Complete calibration of a structured light stripe vision sensor through planar target of unknown orientations». *Image and Vision Computing*, Vol. 23, n° 1, pages 59–67, January 2004.

Résumé

Reconstruction 3D par vision active codée. Application à la vision endoscopique

L'objectif de ce travail est de développer un outil endoscopique rigide capable de fournir en temps réel l'information 3D basée sur l'utilisation de la lumière structurée. Nous avons conçu à cet effet un nouveau motif monochrome auquel un codage basé sur la stratégie du voisinage spatial et la formulation des M-array's en utilisant 3 symboles a été associé. De plus, une primitive directionnelle a été considérée. L'originalité de ce motif se situe également dans la robustesse de son codage caractérisée par une distance de Hamming moyenne supérieure à 6. Les algorithmes de traitement proposés ont été validés. L'erreur moyenne obtenue a été inférieure à 1 mm pour la position et à 1° pour l'orientation dans un espace de travail de profondeur de 600 à 800 mm. Les optiques diffractives ont été employées pour la réalisation de l'outil. La couleur verte a été retenue pour illuminer le motif. Les tests in-vivo sur des surfaces abdominales d'un porc ont donné des résultats prometteurs.

Mots-clés : Reconstruction 3D, lumière structurée, encodage de motifs, temps réel, étalonnage, matrices parfaites, chirurgie mini-invasive.

Abstract

3D Reconstruction using codified active vision. Application to endoscopic vision.

The aim of this work is to develop a rigid endoscopic tool which can provide real-time 3D reconstruction using a structured light projection. To this end, we designed a new monochromatic pattern to which a code based on the spatial neighborhood strategy and the theory of M-array's using 3 symbols was associated. In addition, a directional primitive was considered. The originality of this pattern is also in the robustness of its code characterized by a mean Hamming distance greater than 6. The proposed processing algorithms were validated. The average error was less than 1 mm for position and 1° for the orientation in a workspace of depth from 600 to 800 mm. Diffractive optical elements were used for the implementation of the tool. The green color was chosen to illuminate the pattern. The tests in-vivo on the internal surfaces of a pig abdomen have shown promising results.

Keywords : 3D reconstruction, structured light, pattern codification, real-time, calibration, perfect maps, mini-invasive surgery.