



HAL
open science

Contribution à la cartographie 3D des parois internes de la vessie par cystoscopie à vision active

Achraf Ben Hamadou

► To cite this version:

Achraf Ben Hamadou. Contribution à la cartographie 3D des parois internes de la vessie par cystoscopie à vision active. Médecine humaine et pathologie. Institut National Polytechnique de Lorraine, 2011. Français. NNT : 2011INPL051N . tel-01749440

HAL Id: tel-01749440

<https://hal.univ-lorraine.fr/tel-01749440>

Submitted on 29 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Contribution à la cartographie 3D des parois internes de la vessie par cystoscopie à vision active

THÈSE

présentée et soutenue publiquement le 19 septembre 2011

pour l'obtention du

Doctorat de l'Institut National Polytechnique de Lorraine

Spécialité Automatique, Traitement du Signal et des Images, Génie Informatique

par

Achraf Ben Hamadou

Composition du jury

<i>Président :</i>	Walter C. P. M. Blondel	PU à Nancy université, CRAN
<i>Rapporteurs :</i>	Su Ruan	PU à l'université de Rouen, LITIS
	David Fofi	PU à l'université de Bourgogne, Le2i Le creusot
<i>Examineurs :</i>	Sylvie Treuillet	MdC HDR à la Polytech d'Orléans, PRISME
	Christian Daul	PU à Nancy université, CRAN
	Charles Soussen	MdC à Nancy université, CRAN
<i>Invités :</i>	Didier Wolf	PU à Nancy université, CRAN
	François Guillemin	PUPH au Centre Alexis Vautrin Nancy, CRAN



Centre de Recherche en Automatique de Nancy
UMR 7039 - CNRS - Nancy Université

2 avenue de la Forêt de Haye 54516 Vandœuvre-Lès-Nancy

Mis en page avec la classe thloria.

Remerciements

Ce travail de thèse a été réalisé au sein du Centre de Recherche en Automatique de Nancy (CRAN UMR 7039 CNRS/Nancy Université), dans les murs de l'Institut National Polytechnique de Lorraine (ENSEM/INPL).

Je souhaite tout d'abord exprimer mes remerciements à M. le professeur Didier Wolf pour m'avoir accueilli au sein du groupe Ingénierie pour la Santé (IPS) du CRAN pour réaliser cette thèse.

Je remercie mon directeur de thèse, M. Christian Daul, professeur des universités à l'INPL, pour m'avoir encadré avec compétence et disponibilité tout au long de ma thèse. Ses connaissances étendues en traitement d'images ont été un support et une inspiration pour mon travail.

Je tiens à remercier mon co-directeur de thèse, M. Charles Soussen, Maître de conférences à l'université de Henry Poincaré. Qu'il soit assuré de ma profonde reconnaissance pour m'avoir encadré tout au long de cette thèse. Je le remercie aussi pour ces conseils et toutes les corrections qu'il a apportées à mon rapport de thèse.

Je remercie particulièrement Madame Su Ruan et M. David Fofi, respectivement professeurs de l'université de Rouen et de l'IUT du Creusot. Je leur sais gré d'avoir accepté d'être rapporteurs de cette thèse. Je remercie aussi Madame Sylvie Treuillet, de l'université d'Orléans, qui a accepté d'examiner cette thèse.

Je remercie aussi M. François Guillemin, qui a apporté toute son expertise médicale à ce travail et qui a initié ce sujet de recherche. Je ne saurais oublier M. Walter Blondel pour sa participation à ce travail.

Je pense aussi à tous les membres de l'équipe IPS du laboratoire CRAN, qui m'ont accueilli parmi eux en toute hospitalité.

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

*À mon père abdelmajid et à ma mère Karama
qui m'ont soutenue et encouragé tout au long de mes études.
je vous aime*

À mon cher frère Amine.

À ma chère sœur Dorra.

*À ma chère épouse Sana et mon adorable fils Mahdi
qui illuminent chaque jour de ma vie
Je vous adore.*

*À toute ma famille
A tous mes amis
Je vous aime*

Table des matières

Introduction générale	1
Chapitre 1 Contexte médical de la thèse	7
1.1 Le cancer	7
1.1.1 Le cancer des tissus épithéliaux	8
1.1.2 Le cancer de la vessie	9
1.1.2.1 Caractéristiques des tissus urothéliaux	9
1.1.2.2 Types et classification des cancers de la vessie	10
1.1.2.3 Techniques actuelles de dépistage du cancer de la vessie	12
1.1.3 Exploration visuelle de la paroi interne de la vessie	14
1.2 Motivation et besoins cliniques	14
1.2.1 Limites des supports visuels disponibles en cystoscopie conventionnelle.	14
1.2.2 Construction de nouveaux supports visuels panoramiques	15
1.3 Objectifs des travaux de thèse	16
Chapitre 2 Cartographie 2D et 3D : application aux données endoscopiques	19
2.1 Cartographie 2D des parois internes de la vessie	19
2.1.1 Modèle de formation d’images	20
2.1.2 Mosaïquage d’images	24
2.1.2.1 Imagerie panoramique	24
2.1.2.2 Recalage d’images	26
2.1.2.3 Transformation perspective	27
2.1.3 Mosaïquage d’images cystoscopiques	28
2.1.3.1 A_{DQ} : Algorithme basé sur la distance quadratique	29
2.1.3.2 A_{IM} : Algorithme basé sur l’information mutuelle	29
2.1.4 Étude comparative	30

2.1.4.1	Robustesse	30
2.1.4.2	Précision	31
2.1.4.3	Vitesse de calcul	34
2.1.5	Bilan	35
2.2	Endoscopie 3D	36
2.2.1	Vision passive	36
2.2.2	Vision active	38
2.2.3	Comparaison des méthodes de vision actives et passives	42
2.3	Objectifs et verrous scientifiques de la thèse	44
2.4	Conclusion	45

Chapitre 3 Prototype expérimental : calibrage et reconstruction de points 3D 47

3.1	Prototype expérimental	48
3.2	Calibrage des systèmes de vision en lumière structurée	50
3.2.1	Calibrage de la caméra	50
3.2.2	Calibrage des projecteurs de lumière structurée	54
3.2.3	Bilan sur les méthodes de calibrage de projecteurs de lumière structurée	57
3.3	Méthode proposée pour le calibrage du projecteur de points laser	57
3.3.1	Principe général de la méthode	57
3.3.2	Étapes de la méthode de calibrage proposée	59
3.4	Reconstruction de points 3D	69
3.5	Évaluation de la reconstruction 3D	70
3.5.1	Expérimentations	70
3.5.2	Résultats	71
3.5.2.1	Précision de la reconstruction en fonction du nombre d’images de calibrage	72
3.5.2.2	Précision en fonction de la distance de la pièce test au capteur	73
3.5.2.3	Précision en fonction de l’inclinaison de la pièce test . . .	75
3.5.3	Commentaires et bilan général sur les résultats	76
3.6	Conclusion	77

Chapitre 4 Construction de surfaces étendues de la paroi interne de la vessie 79

4.1	Données disponibles et transformations géométriques	80
4.1.1	Données disponibles	80

4.1.2	Transformation projective	80
4.1.3	Mouvement 3D entre deux acquisitions consécutives	81
4.1.4	Mouvement 2D entre deux acquisitions consécutives	82
4.2	Stratégie de la construction de surfaces 3D étendues	83
4.2.1	Formulation du problème	83
4.2.2	Techniques d'estimation de mouvement 3D	84
4.2.2.1	Estimation de mouvement 3D à partir de points 3D	85
4.2.2.2	Estimation du mouvement 3D à partir d'images 2D	85
4.2.3	Discussion et choix effectué	86
4.3	Estimation du mouvement 3D guidée par recalage 2D/2D d'images	87
4.3.1	Principe général de la méthode	87
4.3.2	Fondement mathématique de la méthode	88
4.3.2.1	Relation mathématique entre les transformations $\mathbf{T}_{3D}^{k-1,k}$ et $\mathbf{T}_{2D}^{k-1,k}$	88
4.3.2.2	Calcul de la transformation perspective \mathbf{T}_{2D} associée à une transformation \mathbf{T}_3	
4.3.3	Algorithme développé	91
4.3.3.1	Choix de la fonction coût \mathcal{F}	92
4.3.3.2	Minimisation de la fonction coût \mathcal{F}	94
4.4	Évaluation de la méthode proposée	96
4.4.1	Expérimentations	96
4.4.1.1	Simulation des données 3D/2D.	97
4.4.1.2	Application de l'algorithme sur données simulées	98
4.4.1.3	Critères quantitatifs de l'évaluation	98
4.4.2	Étude de la précision inhérente à la méthode	99
4.4.3	Étude de la robustesse inhérente à la méthode	101
4.5	Conclusion	107

Chapitre 5 Évaluation de l'algorithme de cartographie 3D 109

5.1	Description des expérimentations	110
5.1.1	Objectif des expérimentations	110
5.1.2	Fantômes utilisés	110
5.1.2.1	Fantôme plan (\mathcal{X}_1)	111
5.1.2.2	Fantôme cylindrique (\mathcal{X}_2)	112
5.1.2.3	Fantôme en forme de vagues (\mathcal{X}_3)	112
5.1.3	Scénarios testés	112
5.1.4	Critères d'évaluation	114

Introduction générale

Cette thèse a été effectuée au CRAN (Centre de Recherche en Automatique de Nancy, UMR 7039) dans le groupe thématique IPS (Ingénierie pour la Santé). Un des axes de recherche du groupe IPS réside dans l'amélioration du diagnostic du cancer de la vessie à l'aide de l'examen cystoscopique. C'est également l'objectif de cette thèse.

La cystoscopie est actuellement l'examen clinique de référence qui permet l'exploration visuelle des parois internes épithéliales de la vessie. L'instrument utilisé pour cet examen est le cystoscope (endoscope dédié aux voies urinaires). Il est constitué d'une fine tige qui est introduite dans la vessie à travers l'urètre et qui permet l'observation des parois internes de la vessie. Le clinicien effectue l'examen cystoscopique en balayant les surfaces internes de la vessie tout en prenant comme référence des repères anatomiques tels que les uretères ou la bulle d'air située à la surface du liquide qui remplit l'organe. Ce remplissage déplisse les parois de la vessie. La vessie est convenablement remplie (pleine) est alors stable dans sa morphologie au moment de l'examen. L'observation de chaque région d'intérêt à un instant donné est réalisée à travers une « fenêtre » (champ de vue) réduite. En effet, chaque image de la séquence vidéo ne visualise qu'au mieux qu'une petite surface de quelques centimètres carrés de la paroi interne.

Les données classiquement archivées pour des observations post-examens ou post-opératoires sont une photographie (image à champ de vue réduit de la zone d'intérêt), un croquis effectué à la main et qui situe la photographie dans la vessie et éventuellement la séquence vidéo. Les lésions vésicales sont en général multifocales. Ces lésions s'étalent souvent sur des surfaces plus grandes que celles visibles dans les images de la séquence vidéo et les photographies archivées. Que ce soit durant l'examen ou avec les données archivées, le clinicien ne peut donc pas visualiser la lésion dans sa globalité, ce qui ne facilite pas le diagnostic.

Par ailleurs, sur une même image, il est souvent impossible de voir à la fois la lésion et un repère anatomique. Pour cette raison, seul le clinicien qui a effectué l'examen sera capable d'exploiter les données archivées. Le suivi de l'évolution d'une lésion par un deuxième clinicien sera impossible. La séquence vidéo ne résout pas non plus ces pro-

blèmes, même si l'ensemble des régions d'intérêt a été scanné par le cystoscope. En effet, la lésion et les repères anatomiques sont dispersés dans plusieurs images. Parcourir la séquence vidéo pour reconstituer mentalement une seule image qui contiendrait toute l'information d'intérêt est très fastidieux et long, voire même souvent impossible. Par ailleurs, la séquence vidéo est rarement conservée (stockée) à cause de la grande quantité de données. Ces données vidéo sont d'ailleurs fortement redondantes, les images consécutives d'une séquence visualisent de grandes zones communes (homologues).

Un support visuel contenant des zones étendues de la vessie et sans redondance de données permettrait au clinicien d'observer, en une seule fois, des zones de tissus contenant à la fois la lésion entière et un repère anatomique. Voir des zones élargies permettrait donc de faciliter l'exploration visuelle des parois internes de la vessie au moment de l'examen. Ces zones élargies pourraient aussi être utilisées pour réaliser des observations post-examen ou post-opératoires, pour faciliter le suivi du patient, le diagnostic ou pour le stockage des données informatives sur un minimum d'espace mémoire. La construction d'un tel support est généralement effectuée à l'aide de méthodes de cartographie 2D (appelées aussi méthodes de mosaïquage). La cartographie 2D est une technique de traitement d'images qui permet la construction d'une mosaïque (images panoramiques) regroupant, en une seule image (dans un repère global), une scène représentée dans une séquence d'images 2D (avec des repères locaux). L'élément central de ces méthodes est le processus de recalage d'images qui consiste à trouver le lien entre les repères des images d'une séquence. Cette information géométrique peut ensuite être utilisée pour ramener les images de la séquence dans un système de coordonnées commun (repère global de l'image panoramique).

Au laboratoire CRAN, deux travaux de thèse ont été effectués pour le mosaïquage des parois internes de la vessie. Les résultats obtenus lors de ces deux thèses montrent le potentiel de la construction d'images panoramiques des parois internes de la vessie pour pallier les limites des supports visuels actuels en cystoscopie (séquence d'images vidéo). Malgré ce potentiel, les images panoramiques 2D ne permettent pas la visualisation de la surface complète de la vessie. La visualisation 2D s'avère donc inadaptée. Par ailleurs, les urologues se représentent mentalement (et de façon naturelle) les images cystoscopiques en trois dimensions, les organes creux étant des entités tridimensionnelles (3D). Une représentation 3D des surfaces internes de la vessie a pour but de faciliter l'interprétation des données acquises lors d'un examen cystoscopique grâce à l'augmentation du réalisme des données disponibles. En effet, les urologues auront une appréciation visuelle des vraies courbures et des formes des surfaces internes de la vessie. De plus, reconstruire la géométrie des parois peut conduire à de nouveaux critères de diagnostic comme des indices

sur la distribution spatiale des lésions qui est très utile pour cerner l'évolution des lésions d'un examen à un autre.

La construction de carte 3D pour les parois internes de la vessie est un objectif qui nécessite des développements à la fois sur le plan instrumental et en terme de traitement d'images. En effet, les systèmes cystoscopiques actuels ne fournissent pas d'informations anatomiques tridimensionnelles comme c'est requis pour la cartographie 3D. De nos jours, la construction des formes anatomiques tridimensionnelles en endoscopie est un sujet de recherche en vogue. Les travaux réalisés dans ce domaine tournent essentiellement autour de deux axes, à savoir.

- Développements instrumentaux. Il s'agit de travaux pour la fabrication de nouveaux systèmes endoscopiques permettant de fournir des informations 3D sur les surfaces d'organes imagées. La majorité de ces développements s'oriente vers l'utilisation de la vision active pour la reconstruction de points 3D localisés sur la surface imagée par le système endoscopique.
- Développements en terme de méthodes de traitement d'images. Ces travaux s'orientent plutôt vers le développement et l'adaptation de méthodes de traitement d'images pour la reconstruction 3D des surfaces imagées tout en utilisant des systèmes endoscopiques existants (acquisition d'image 2D). Ces méthodes n'impliquent donc pas de modifications fondamentales des instruments. Ces développements se basent sur la détection et le suivi de primitives images (*e.g.*, coins, segments de droite, *etc.*). Il s'agit par exemple des techniques de « *structure from motion* ».

Dans le cas de la cystoscopie, il n'y a aucun travail ayant pour objectif la reconstruction 3D des parois internes de la vessie. Les travaux réalisés par le passé en « endoscopie tridimensionnelle » ne sont pas transposable en cystoscopie. En effet, une étude menée au sein du CRAN montre que la nature des images cystoscopiques (textures et condition d'éclairage à forte variabilité intra- et inter-patient, *etc.*) ne favorise pas l'utilisation des techniques de détection de primitives images (par exemple les détecteurs de Harris et SIFT) pour accomplir une cartographie des parois internes de la vessie. D'ailleurs, les systèmes endoscopiques actuels basés sur la vision active ne permettent pas d'acquérir simultanément des informations de couleurs/textures et anatomiques des surfaces.

La cartographie 3D des parois internes de la vessie est à l'origine de défis et de verrous spécifiques à l'application. Ainsi, il faut développer un algorithme de reconstruction 3D qui permet, pour chaque image d'une séquence vidéo, de reconstruire des points 3D localisés sur la portion de la paroi imagée. Les points 3D reconstruits pour chaque image sont exprimés dans un repère « local » propre à l'image. Le second défi à relever réside dans le développement d'une méthode permettant de placer l'ensemble des points 3D reconstruits

dans un repère (référentiel) « global » pour reconstruire la globalité de la surface imagée d'une séquence vidéo cystoscopique. La surface reconstruite jouera le rôle du support 3D sur lequel seront collées les images 2D de la séquence.

L'objectif de cette thèse est de développer de nouvelles méthodes de traitement d'images nécessaires à la reconstruction de surfaces 3D étendues de la paroi interne de la vessie. Le but est d'obtenir une surface 3D correspondant à une surface étendue de la paroi de la vessie. La texture et les couleurs vues dans les images 2D classiquement acquises avec les cystoscopes doivent pouvoir être plaquées de façon cohérente (sans discontinuités) sur les surfaces reconstruites.

Le travail qui est en cours au CRAN (à savoir la réalisation d'un banc optique pour valider le principe de la vision active appliqué à un cystoscope rigide) ne fait pas partie de ce travail de thèse. Dans le cadre de ce travail, nous avons construit un système à vision active qui simule le fonctionnement d'un cystoscope 3D. Nous utilisons ce système (prototype expérimental) ainsi que les données fournies par ce dispositif pour concevoir et tester les algorithmes nécessaires à la cartographie 3D (méthode d'étalonnage du prototype, de reconstruction de point 3D, de recalage 2D et de mosaïquage 3D). La thèse est structurée de la façon suivante.

Le chapitre 1 présente le contexte applicatif, la motivation clinique, les objectifs de nos travaux et la problématique liée à ces derniers. Dans la dernière partie de ce chapitre, nous exposons les contraintes scientifiques liées à l'application.

Le chapitre 2 donne un état de l'art sur les techniques de cartographie 2D et 3D. Nous passons en revue les différentes techniques de cartographie 2D appliquées à la cystoscopie tout en discutant leurs performances. Cette partie de l'état de l'art se termine par une étude comparative de ces techniques. La deuxième partie du chapitre est consacrée à la cartographie 3D en endoscopie en général. Nous y présentons les différentes approches, que ce soit en terme d'instrumentation ou en terme de traitement d'images et de reconstruction de données tridimensionnelles à partir de systèmes endoscopiques. Cette étude nous a permis d'orienter les solutions que nous proposons dans le cadre de cette thèse pour aboutir à la reconstruction des surfaces de la paroi interne de la vessie.

Le chapitre 3 décrit le système à vision active (prototype non miniaturisé) que nous avons construit et utilisé pour la validation des approches proposées dans cette thèse. Nous présentons également les outils et les méthodes de traitement d'images que nous proposons pour le calibrage de notre système ainsi que pour la reconstruction de données 3D pour chaque acquisition d'une séquence.

Le chapitre 4 présente l'approche retenue pour la cartographie 3D des parois internes de la vessie en exploitant les données 2D et 3D fournies par le prototype pour chaque

image d'une séquence. Le chapitre inclut également une évaluation de la performance en terme de précision et de robustesse inhérente à la solution proposée. Cette évaluation est réalisée exclusivement avec des données simulées à partir de fantômes numériques.

Le chapitre 5 complète l'évaluation proposée dans le chapitre 4. Il s'agit d'évaluer la capacité de notre méthode de cartographie 3D complète (incluant la reconstruction de données 3D et le mosaïquage 3D) à retrouver précisément les formes 3D des surfaces imagées. Les évaluations entreprises dans ce chapitre sont réalisées avec des données acquises par notre système à vision active pour des fantômes réalistes de vessies. En effet, n'ayant pas encore de cystoscope 3D utilisable en condition clinique, nos algorithmes sont éprouvés sur des fantômes réalistes.

Nous terminons ce document par une conclusion qui met en avant les contributions et les perspectives de nos travaux.

1

Contexte médical de la thèse

Le début de ce chapitre décrit le cancer de la vessie ainsi que les caractéristiques de ces lésions cancéreuses. Nous proposons ensuite un bref aperçu des techniques de diagnostic du cancer de la vessie en portant un intérêt particulier sur l'examen cystoscopique qui est l'examen standard pour l'exploration visuelle des parois internes de la vessie. Nous montrons également les limites du support visuel (séquence d'images cystoscopiques) associé à cet examen. Ces limites sont à l'origine des motivations et les objectifs de cette thèse.

1.1 Le cancer

Le terme « *cancer* » fait référence aux maladies pour lesquelles les cellules se reproduisent et disséminent sans contrôle dans le corps humain. Actuellement le cancer est considéré comme la maladie la plus mortelle dans le monde. Selon les dernières statistiques, le cancer provoque chaque année quelque 7,6 millions de décès dans le monde soit 60% des 12,7 millions de cas de cancer [WHO, 2009]. Ces chiffres devraient augmenter dans les prochaines années, en partie, du fait de l'allongement de la durée de vie. Dans le monde, on estime que 26,4 millions de personnes seront diagnostiquées pour un cancer chaque année d'ici 2030 avec 17 millions de décès [WHO, 2009].

Le développement du cancer se caractérise par la multiplication anormale de cellules malignes dans un tissu sain. Les cellules cancéreuses ont subi des modifications de leur génome qui induisent :

- une perte de l'apoptose (mort cellulaire programmée indispensable à l'équilibre du corps pour tenir compte du renouvellement des tissus),
- une perte du contrôle de la multiplication cellulaire et

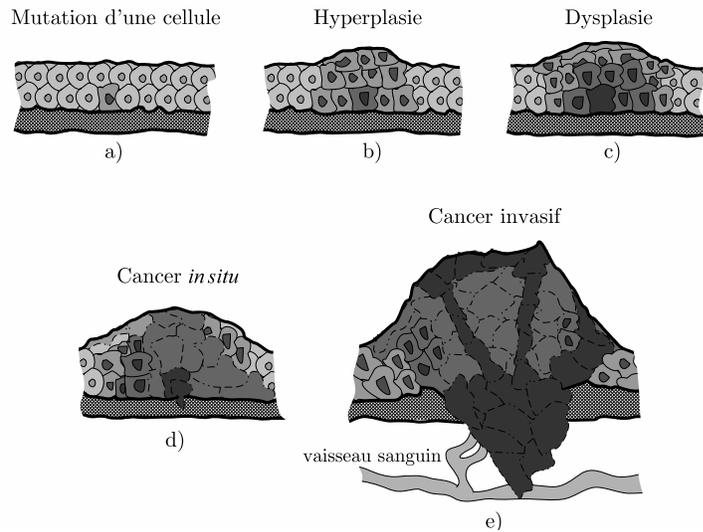


FIGURE. 1.1 – Représentation simplifiée des différentes étapes de l'évolution d'un tissu épithélial cancéreux (illustration extraite de [Hernandez-Mier, 2007]). a) Mutation d'une cellule. b) Augmentation anormale du nombre de cellules et de la masse du tissu (hyperplasie). c) Développement anormal du tissu (dysplasie). d) Apparition d'une lésion maligne confinée à l'épithélium. e) Propagation de cellules malignes qui peuvent envahir d'autres tissus.

– un pouvoir d'invasion des structures voisines.

Les cellules cancéreuses sont des clones d'une première cellule dont certaines caractéristiques ont changé. Une tumeur cancéreuse peut être définie comme une population de cellules génétiquement instables qui sont faiblement liées entre elles [Abelev and Kopnin, 2000].

85% de tous les cancers prennent naissance dans les couches superficielles (épithélium) de la peau et des organes creux, dont 50% au sein des de la vessie, du colon, de l'œsophage, de l'estomac et du col de l'utérus.

1.1.1 Le cancer des tissus épithéliaux

L'épithélium est un tissu dont les cellules sont jointives et solidaires les unes des autres grâce à des jonctions intercellulaires. Il n'est pas vascularisé et repose sur une membrane basale au travers de laquelle l'apport des nutriments et l'évacuation des déchets s'effectue grâce au tissu conjonctif sous-jacent.

Dans le cas des tissus épithéliaux, on distingue différents stades d'évolution [Alberts *et al.*, 2005]. Schématiquement, un cancer se développe à partir d'une cellule mutante (voir figure 1.1(a)), dont les modifications internes, en particulier l'incapacité de mourir par apoptose, facilitent la prolifération et la survie. L'hyperplasie (voir figure 1.1(b)) correspond à une augmentation du nombre de cellules du tissu. Ces cellules sont soit normales, soit porteuses d'anomalies initiant une dégénérescence maligne. La *dysplasie* correspond à une

nouvelle étape de mutation où le nombre de cellules et leur développement augmentent de façon significativement anormale (voir figure 1.1(c)). Si l'état des cellules affectées est très anormal, tant au niveau du développement que de l'apparence et que la tumeur produite par ces cellules est confinée en deçà de la muqueuse, elle est appelée carcinome *in situ* (CIS, voir figure 1.1(d)) ou carcinome intra-muqueux. Un CIS peut rester confiné pendant une période prolongée. Si certaines cellules subissent des mutations supplémentaires, la tumeur peut envahir le tissu sous-jacent et des cellules affectées peuvent atteindre les systèmes de circulation sanguine ou lymphatique. La tumeur devient donc un vrai cancer invasif (voir figure 1.1(e)). Elles peuvent se propager dans d'autres organes et former des *métastases*.

Une lésion à un stade pré-cancéreux est une lésion témoignant d'un processus de cancérisation à un stade précoce non invasif. Elle comporte un risque de progression tumorale, mais peut aussi régresser ou persister sans s'aggraver.

1.1.2 Le cancer de la vessie

Parmi les cancers des organes creux, celui de la vessie est le 7^{ème} cancer le plus répandu dans le monde, soit 3,7% de l'ensemble des cancers [Tyczynski and Parkin, 2003]. Selon le CIRC (Centre Internationale de Recherche sur le Cancer), les tumeurs de la vessie ont représenté la 5^{ème} cause de cancer et la 3^{ème} cause de décès en France. 80% des tumeurs vésicales surviennent entre 50 et 80 ans.

1.1.2.1 Caractéristiques des tissus urothéliaux

La vessie est un organe qui fait partie du système urinaire (voir figure 1.2(a)). Elle a la forme d'une poche qui reçoit l'urine produite par les reins. Elle est constituée d'un tissu musculaire qui, en se contractant, permet à l'urine d'être expulsée. Lorsque la vessie est vide, l'épaisseur moyenne de sa paroi est d'environ 5 mm. En état de distension moyenne, son épaisseur ne dépasse pas 2 mm. La paroi de la vessie est constituée de plusieurs couches (voir figure 1.2(b)). Une couche de *cellules urothéliales* recouvre l'intérieur de la vessie ainsi que l'intérieur des reins, des uretères et de l'urètre. Cette couche, aussi appelée *épithélium transitionnel* repose sur une membrane basale. Sous l'urothélium se trouve une fine couche de tissu conjonctif appelée *lamina propria* aussi appelée *sous-muqueuse* ou *chorion*. Cette couche mesure en moyenne 1,4 mm d'épaisseur. La couche suivante, plus profonde, est une zone de tissu musculaire appelée *muscularis propria*. Enfin, une couche de tissu adipeux sépare la vessie des autres organes.

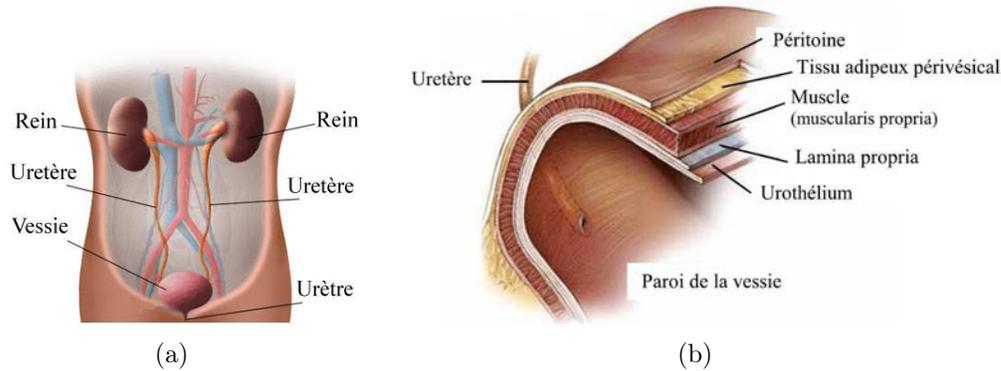


FIGURE. 1.2 – (a) Schéma montrant les principaux composants du système urinaire. (b) Illustration des différentes couches formant la paroi de la vessie [© 2003 American Society of Clinical Oncology].

1.1.2.2 Types et classification des cancers de la vessie

En fonction du pronostic (prévision de l'évolution ou de l'issue de la maladie en fonction des facteurs pouvant l'influencer), les tumeurs de la vessie peuvent être divisées en trois groupes :

- Les tumeurs bénignes,
- les tumeurs malignes qui infiltrent le chorion,
- les tumeurs à potentiel malin indéterminé présentant un caractère récidivant et extensif et qui constituent le groupe le plus important.

Du point de vue histologique, deux types de tumeurs peuvent être distingués selon leur grade (proportionnel à l'importance des anomalies des tissus tumoraux) : les *tumeurs papillaires* (infiltrantes ou non infiltrantes) et les *lésions tumorales planes* (carcinomes *in situ*).

- **Tumeurs papillaires :**

Non infiltrantes. La plupart des tumeurs de vessie sont superficielles. Selon leur aspect, elles sont appelées polype superficiel, papillome ou tumeur superficielle. Ces tumeurs n'infiltrent pas la *lamina propria*.

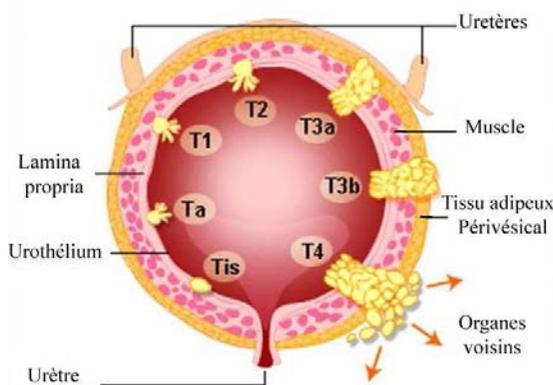
Infiltrantes. Dans la plupart des cas, leur racine s'enfonce profondément dans la paroi de la vessie, jusqu'à la couche *musculaire* (qui contient les fibres musculaires de la vessie) ou au delà.

- **Les lésions tumorales « planes » :**

Carcinome in-situ (CIS). Les lésions urothéliales planes désignent un spectre d'anomalies histologiques intra-épithéliales de l'urothélium allant de l'hyperplasie plane au CIS. Le CIS est une tumeur plate, non papillaire, située dans l'épaisseur même de la muqueuse. Le CIS est un cancer épithélial résistant et agressif avec un poten-

tiel d'infiltration. Même si le CIS peut se présenter comme une lésion isolée, il peut être *multifocal* et *diffus*. Le CIS est fréquemment associé aux tumeurs papillaires. La présence du CIS proche des tumeurs papillaires est souvent associée à un risque plus important de progression de la tumeur et de récurrence. A cause de ces caractéristiques, le CIS peut passer inaperçu à l'examen cystoscopique en lumière blanche conventionnelle.

La classification *TNM*, *Tumor Node Metastasis* en anglais, classe les tumeurs par rapport à leur degré de pénétration (stade) dans la paroi de la vessie (voir figure 1.3).



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

Ta Tumeur papillaire non-invasive (confinée à l'urothélium).

T1 La tumeur envahit le tissu mou (*lamina propria* de la vessie).

T2 La tumeur envahit le muscle de la vessie.

T3 La tumeur envahit le tissu adipeux périvésical (T3a invasion microscopique, T3b invasion macroscopique).

T4 La tumeur envahit un organe voisin (prostate, utérus, vagin, *etc.*).

FIGURE. 1.3 – Différents degrés d'invasion (stades) d'une tumeur de vessie.

Les tumeurs superficielles peuvent poser deux problèmes évolutifs essentiels : soit un risque de récurrence sans progression où le *grade* et le *stade* resteront identiques à ceux de la tumeur initiale, soit un risque de récurrence avec progression (ce dernier cas est de loin le plus fréquent). La progression peut être définie comme l'aggravation du stade et/ou du grade par rapport à celui de la tumeur initiale. D'une manière générale, plus le stade et le grade sont élevés plus le risque de récurrence et de progression est grand.

1.1.2.3 Techniques actuelles de dépistage du cancer de la vessie

Le dépistage en cancérologie vise, à l'aide d'examens cliniques, à détecter la présence d'un cancer précoce ou d'une lésion pré-cancéreuse avant que des symptômes cliniques apparaissent. Un examen clinique doit pouvoir indiquer de façon fiable la présence ou l'absence d'une maladie chez un individu. La qualité de l'examen peut être mesurée par l'appréciation de sa *validité intrinsèque*. Cette validité intrinsèque est composée de deux paramètres : la *sensibilité* et la *spécificité*.

La sensibilité et la spécificité peuvent être définies en utilisant les résultats d'un examen clinique : le *vrai positif* (résultat positif d'un test sur un organe ou tissu malade), le *faux positif* (résultat positif d'un test sur un organe ou tissu sain), le *faux négatif* (résultat négatif d'un test sur un organe ou tissu malade) et le *vrai négatif* (résultat négatif d'un test sur un organe ou tissu non malade). La *sensibilité* (Se) est la probabilité d'avoir un test positif lorsque la maladie est présente :

$$Se = \frac{\text{Nombre de vrais positifs}}{\text{Nombre de sujets malades}} \quad (1.1)$$

La spécificité (Sp) est la probabilité d'avoir un test négatif lorsque la maladie est absente :

$$Sp = \frac{\text{Nombre de vrais négatifs}}{\text{Nombre de sujets non malades}} \quad (1.2)$$

Il existe plusieurs examens cliniques permettant de diagnostiquer une lésion cancéreuse dans la vessie. Nous présentons dans la suite uniquement les techniques les plus courantes.

La cytologie urinaire. Dans cet examen, les cellules urothéliales desquamées à partir d'urines mictionnelles ou d'un lavage vésical sont étudiées. C'est l'examen biologique courant le plus performant, sa spécificité étant de 93%. En revanche, son point faible est la sensibilité qui dépend étroitement du grade tumoral : 27% pour les tumeurs de bas grade et 77% pour les tumeurs de haut grade. La cytologie urinaire est donc intéressante pour le diagnostic de tumeurs agressives de haut grade. Il existe jusqu'à 12% de faux-positifs (inflammation vésicale, lithiase urinaire, *etc.*). Pour éviter ces inconvénients, elle peut être associée à la cytométrie de flux, permettant un taux de diagnostic de 95 %.

L'échographie vésicale sus-pubienne. Elle utilise des ultrasons pour chercher des polypes dans la vessie. Cet examen présente une sensibilité faible de 74% mais une bonne spécificité (95%). Compte tenu de la résolution, son exploitabilité est limitée, surtout quand il s'agit de tumeurs planes, situées sur le dôme vésical, de moins de 2 mm de

diamètre. Cet examen ne peut donc être utilisé pour le diagnostic ou la surveillance. Elle s'effectue par voie sus-pubienne et cherche une image caractéristique de végétation endo-luminale, à base d'implantation pédiculée ou sessile. Son intérêt reste limité.

Uroscanner. C'est un examen scanner avec ou sans injection de produit de contraste à base d'iode. Il visualise les parois de la vessie et les tumeurs qui s'y développent. Dans le cas d'une tumeur invasive, il permet de détecter l'infiltration des méats urétéraux avec la répercussion sur le haut appareil urinaire (*i.e.*, obstacle à l'évacuation des urines aboutissant à une dilatation des cavités rénales et de l'uretère).

La cystoscopie en lumière blanche. À l'aide d'un cystoscope (endoscope pour les voies urinaires) souple ou rigide, elle permet de voir la tumeur vésicale, de la localiser, d'en préciser ses caractères morphologiques et d'explorer le reste de la vessie à la recherche d'éventuelles lésions. Plus de détails sur cet examen seront présentés dans la section 1.1.3. Lorsqu'un tissu anormal est observé, il est nécessaire de prélever un échantillon de tissu pour effectuer une analyse microscopique. Les échantillons utilisés pour la **biopsie** sont obtenus par ablation chirurgicale par voie endoscopique. Une analyse histopathologique permet de préciser le diagnostic de cancer de la vessie. Il est recommandé de faire des biopsies à différents endroits de la vessie (plus ou moins aléatoirement) de façon à déterminer avec précision l'étendue du cancer ou de la dysplasie. La biopsie, suivie de l'analyse histopathologique de l'échantillon prélevé, est le seul moyen de pouvoir donner un diagnostic définitif sur la présence ou non de cancer. Cette façon de procéder est donc le « gold standard ».

La cystoscopie de fluorescence. Cette technique représente un des progrès essentiels de ces dix dernières années pour le diagnostic du cancer de la vessie. Son principe consiste à cibler les cellules tumorales en utilisant un agent fluorophore avec forte affinité pour les lésions tumorales. Un de ces agent est par exemple l'acide aminolévulinique (ALA). Il permet de renforcer le contraste visuel entre les cellules tumorales et les cellules saines bénignes. Introduit dans la vessie et éclairé par une lumière bleue, l'agent photosensible permet de visualiser les tumeurs en les rendant qui deviennent fluorescentes.

Parmi ces techniques de dépistage, la cystoscopie en lumière blanche reste aujourd'hui l'examen de référence qui permet l'exploration visuelle de la paroi interne de la vessie en vue de rechercher ou de suivre des lésions superficielles. Notre étude se focalise sur le traitement d'images vidéo générées par un cystoscope qui éclaire la scène avec une lumière blanche.

1.1.3 Exploration visuelle de la paroi interne de la vessie

Lors d'un examen cystoscopique, l'urologue introduit un cystoscope (rigide ou souple, voir figure 1.4) par l'urètre après une désinfection du gland. Un gel anesthésique est instillé dans l'urètre et étalé sur le tige du cystoscope. Puis il commence son exploration en progressant lentement dans le conduit urinaire puis dans la vessie. L'urologue visualise à l'aide d'un moniteur une séquence vidéo relative à la région de la paroi interne de la vessie illuminée par le cystoscope. L'examen cystoscopique, permet (avec l'équipement adéquat) d'enregistrer la séquence vidéo et de prendre des photographies de « petites » zones d'intérêt. Le support photographique permet de conserver une représentation précise de la zone, d'effectuer des comparaisons dans le temps en cas de récurrence et de disposer d'un document objectif médico-légal [Gattengo and Chopin, 2001]. Actuellement, les supports visuels disponibles en postopératoire sont la séquence vidéo cystoscopique et les impressions papier de certaines images considérées par l'urologue comme présentant un intérêt.

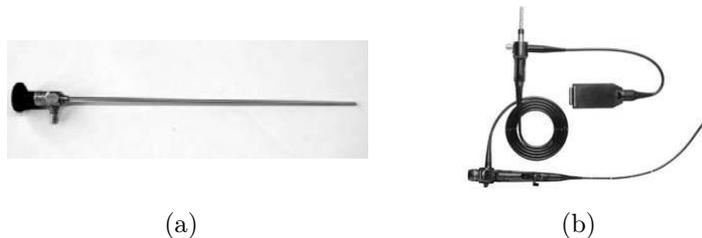


FIGURE. 1.4 – Deux types de cystoscope utilisés en clinique. (a) Cystoscope rigide Karl Storz 27005BA. (b) Cystoscope souple Olympus EndoEYE.

1.2 Motivation et besoins cliniques

1.2.1 Limites des supports visuels disponibles en cystoscopie conventionnelle.

L'urologue note à la main sur un croquis de la vessie les sites où il a observé des anomalies, pris des photos ou réalisé des biopsies en vue de les analyser. En complément, un enregistrement vidéo peut être utilisé. Cependant, le fait de rechercher des images d'intérêts dans une longue séquence d'images peut s'avérer contraignant, surtout si l'urologue est amené à rechercher des images de sites colocalisés d'un examen à un autre. De plus, le *champ de vision* limité du cystoscope (surface de tissu visualisée de l'ordre du cm^2 par image) ne permet d'estimer de façon précise ni l'extension d'une lésion cancéreuse, ni la distribution spatiale des lésions qui sont généralement multifocales (voir figure 1.5). La

localisation des lésions sur un croquis réalisé par le clinicien est hautement dépendante de son expérience. Ces contraintes limitent l'exploitation des supports actuels pour la détection visuelle et le suivi des lésions. D'où l'intérêt de concevoir de nouveaux supports permettant la visualisation de surfaces étendues de la vessie.

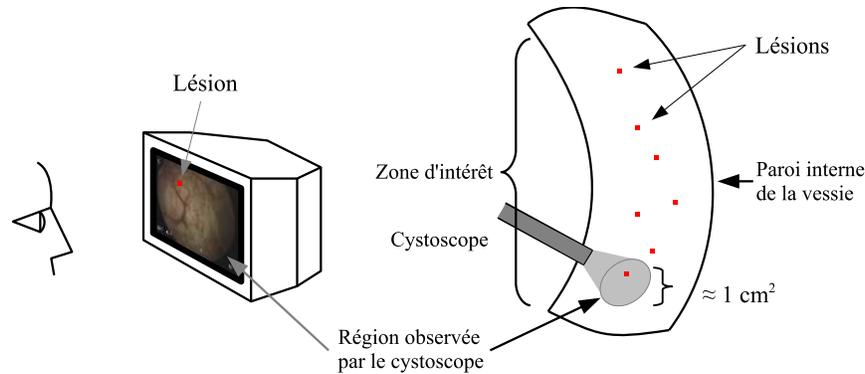


FIGURE. 1.5 – Illustration de la petite région ($\approx 1 \text{ cm}^2$) pouvant être relevée par une image cystoscopique. Quelques lésions multifocales sont éparpillées sur la zone d'intérêt.

1.2.2 Construction de nouveaux supports visuels panoramiques

Visualisation panoramique 2D De récentes études réalisées notamment au sein du CRAN [Miranda-Luna, 2005; Hernandez-Mier, 2007], ont le potentiel d'un support visuel panoramique pour pallier les limites des supports visuels actuels en cystoscopie (séquence d'images vidéo). Ce support est une image panoramique (unique) construite moyennant des méthodes de recalage et de mosaïquage 2D qui utilisent des séquences d'images cystoscopiques. Les pixels de toutes les images de la séquence sont placés dans un repère (système de coordonnées) global (celui de l'image panoramique). L'image panoramique permet la visualisation en une seule fois des surfaces étendues de la paroi interne de la vessie. [Miranda-Luna, 2005] a proposé une méthode de mosaïquage d'images basée sur l'utilisation de l'information mutuelle maximisée par une technique de gradient descendant stochastique. La méthode de mosaïquage d'images cystoscopiques proposée par [Hernandez-Mier, 2007] est basée sur l'estimation du flot optique entre des images successives. Un bilan quantitatif de la performance de ces deux travaux est présenté dans le chapitre 2. La figure 1.6 montre un exemple d'image panoramique construite à partir d'une séquence de 450 images cystoscopiques cliniques. Il est possible de distinguer sur cette image deux polypes assez distants qui ne pouvaient donc être visualisés en une seule image de la séquence.

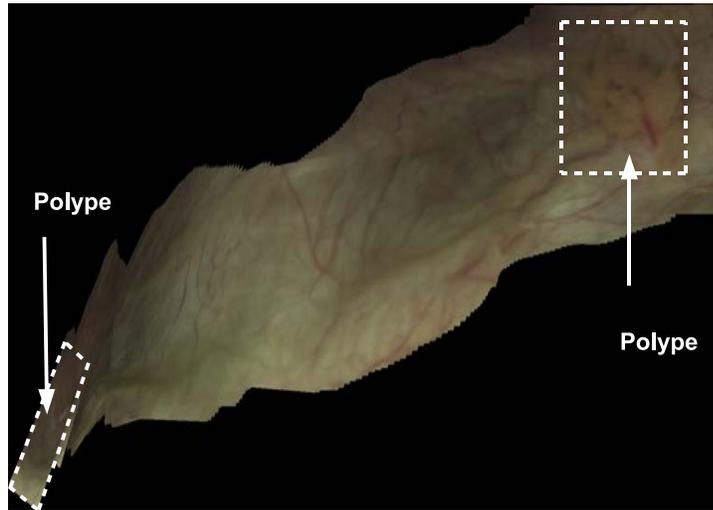


FIGURE. 1.6 – Exemple d’image panoramique (1479×1049 pixels) construite à partir d’une séquence de 450 images cystoscopiques cliniques [Hernandez-Mier, 2007].

Vers la visualisation panoramique 3D Malgré le potentiel que présentent les images panoramiques 2D pour faciliter l’exploration visuelle de la vessie, elles ne permettent pas la visualisation de la surface complète de la vessie. La visualisation 2D s’avère donc inadaptée. Par ailleurs, les urologues se représentent mentalement (et de façon naturelle) les images cystoscopiques en trois dimensions, les organes creux étant des entités tridimensionnelles (3D). Une représentation 3D des surfaces internes de la vessie a pour but d’améliorer l’interprétation des données acquises lors d’un examen cystoscopique grâce à l’augmentation du réalisme du support cystoscopique. En effet, les urologues auront une appréciation visuelle des vraies courbures et des formes des surfaces internes de la vessie. De plus, reconstruire la géométrie des parois peut conduire à de nouveaux critères de diagnostic comme des indices sur la distribution spatiale des lésions qui est très utile pour cerner l’évolution des lésions d’un examen à un autre.

1.3 Objectifs des travaux de thèse

Cette thèse vise, principalement, à développer une méthode de cartographie 3D des parois internes de la vessie. Cette cartographie 3D réside dans la reconstruction et la visualisation de surfaces internes étendues de la vessie.

Le choix de la modalité d’imagerie est très important, car l’instrument doit permettre l’acquisition d’informations relatives non seulement à la forme 3D, mais aussi à la couleur et la texture des surfaces imagées. Par exemple, les dispositifs d’imagerie 3D tels que l’imagerie à résonance magnétique (IRM) ou la tomographie par ordinateur permettent d’acquérir

des données anatomiques 3D avec une haute précision. Cependant, ils ne restituent pas des informations de couleurs et textures de la surface des tissus. L'endoscopie reste aujourd'hui la seule modalité qui permet de restituer au clinicien des images de l'intérieur du corps humain, avec des couleurs et textures naturelles tout en étant au plus proche des organes. Actuellement, aucun système endoscopique commercialisé n'est capable de fournir des données 3D. La section 2.2 du chapitre 2 sera consacrée à une revue de différents travaux de recherche visant à développer des systèmes endoscopiques et des algorithmes de reconstruction 3D pour fournir des données endoscopiques 3D.

Les travaux de cette thèse comportent donc des développements sur le plan instrumental (développement d'un prototype expérimental simulant un cystoscope 3D, calibrage, *etc.*) et en terme de traitement des images (reconstruction 3D, recalage de données 2D et 3D, *etc.*). Les problèmes scientifiques liés à ces développements sont abordés dans les chapitres qui suivent dans ce mémoire.

2

Cartographie 2D et 3D : application aux données endoscopiques

Ce chapitre présente les méthodes existantes de cartographie à partir de données endoscopiques. Le terme cartographie (ou mosaïquage) désigne l'élaboration d'une représentation spatiale (carte ou mosaïque ou image panoramique) à large champ de vue à partir de plusieurs champs de vue restreints. La carte à générer peut être bi-dimensionnelle (mosaïque 2D) ou tri-dimensionnelle (relief ou topographie 3D). Les applications de la cartographie en imagerie médicale restent limitées à quelques modalités et organes. En particulier, dans le cas de la vessie, très peu de solutions émergent de la littérature. Ce chapitre présente, en premier lieu, une étude comparative des principales méthodes connues pour la cartographie 2D de la paroi interne de la vessie à partir d'images cystoscopiques. Nous abordons ensuite la cartographie 3D en endoscopie. Pour ce faire, nous recensons les techniques de reconstruction de données 3D endoscopiques basées sur la vision active ou passive en détaillant leurs avantages et inconvénients dans le cas de la vessie. Cet état des lieux sert de base pour orienter les choix scientifiques et techniques à effectuer dans cette thèse pour la cartographie 3D des parois internes de la vessie.

2.1 Cartographie 2D des parois internes de la vessie

Cette section présente les différentes notions théoriques de base liées à la cartographie 2D. Nous commençons par détailler le modèle de formation d'images endoscopiques qui relie la scène imagée et les images acquises par l'intermédiaire des paramètres optiques de l'instrument. Ensuite, nous détaillons le principe des méthodes de cartographie 2D et

présentons les principaux travaux publiés sur la cartographie 2D appliquée aux images cystoscopiques.

2.1.1 Modèle de formation d'images

Le modèle caméra le plus courant dans la littérature est le *modèle sténopé*. Son principe peut être schématisé par une « boîte » ne laissant pénétrer la lumière que par un petit trou percé sur une des faces de la boîte (trou d'épingle, d'où « *pinhole camera* » en anglais). L'image inversée de la scène se forme sur la face parallèle à celle avec le trou. L'image est formée par les rayons de lumière issus de la scène et passant par le trou. Si le trou est réduit à un point, un photon qui est issu de la scène et qui tombe sur le plan image en passant par le sténopé ne peut suivre qu'une trajectoire rectiligne unique. La projection perspective crée des images inversées selon l'illustration de la figure 2.1. Pour une camera CCD (*charge coupled device*), le sténopé représente le centre de l'optique qui focalise la lumière sur la matrice de capteurs absorbant les photons. L'image est également inversée sur ce plan image. Cette image est inversée une deuxième fois lors de la lecture « ligne par ligne » des capteurs CCD.

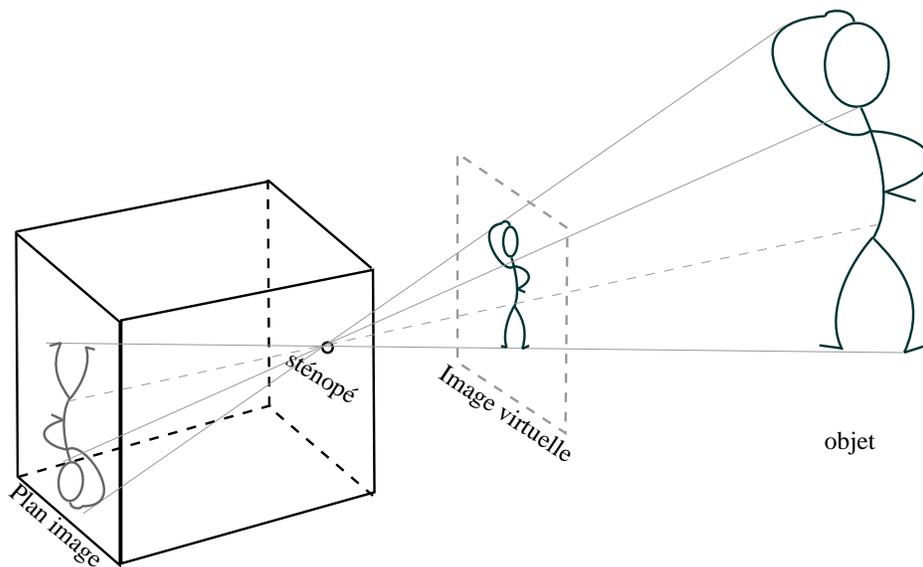


FIGURE. 2.1 – Représentation du modèle sténopé de formation d'images.

Le modèle sténopé permet de modéliser mathématiquement le fonctionnement d'une caméra idéale [Forsyth and Ponce, 2003]. Néanmoins, l'optique associée aux caméras est souvent à l'origine de distorsions dans les images acquises. Ainsi illustré par la figure 2.2, le processus de formation des images peut être modélisé mathématiquement par quatre transformations géométriques.

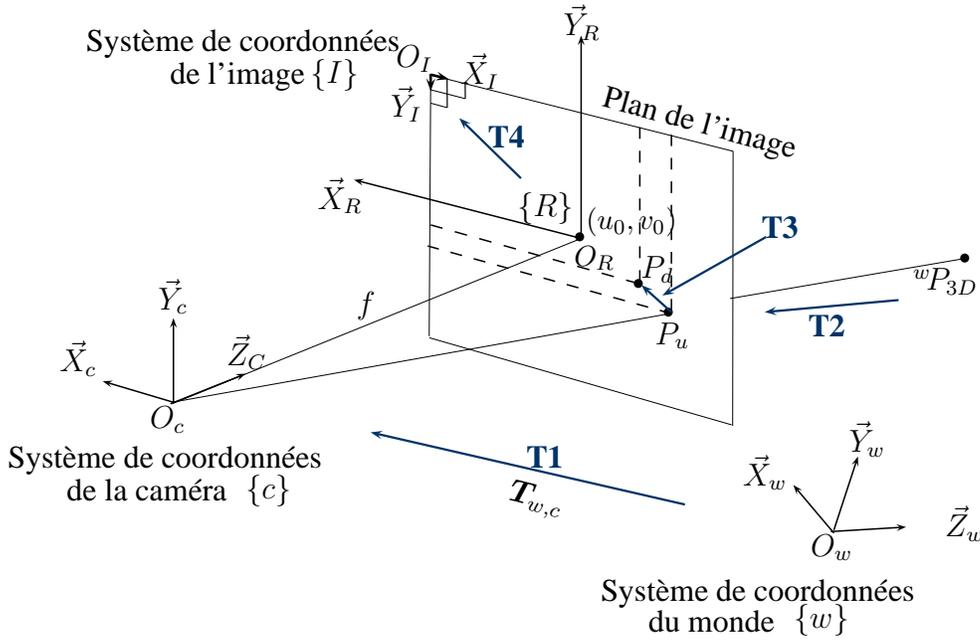


FIGURE. 2.2 – Illustration des quatre transformations nécessaires pour la projection d'un point dans l'espace 3D sur le plan image 2D. La transformation 1 est une transformation rigide 3D entre le repère monde et le repère local à la caméra. La transformation 2 est une projection 3D/2D du repère local à la caméra vers le repère image $\{R\}$. La transformation 3 corrige les distorsions radiales dans le plan image. La transformation 4 (rotation et homothétie) exprime le point projeté dans le système de coordonnées situé sur le coin supérieur gauche de l'image.

Transformation T1 : Cette transformation définit le lien géométrique entre le repère monde $\{w\}$ (défini en général à l'aide d'un plateau d'étalonnage) et le repère caméra $\{c\}$ dont l'origine O_c correspond au centre optique de la caméra (le sténopé dans la figure 2.1). Ainsi illustré par l'équation (2.1) en coordonnées homogènes des points 3D, ce lien correspond à une transformation rigide composée d'une rotation et d'une translation 3D. Un point de coordonnées ${}^w P_{3D} = (x_w, y_w, z_w)^T$ dans le repère monde devient le point de coordonnées ${}^c P_{3D}$ défini par $(x_c, y_c, z_c)^T$ dans le repère 3D de la caméra.

$$\begin{bmatrix} {}^c P_{3D} \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{R} & \vec{\mathbf{t}} \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{T}_{w,c}} \begin{bmatrix} {}^w P_{3D} \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^w P_{3D} \\ 1 \end{bmatrix} \quad (2.1)$$

La matrice 4×4 $\mathbf{T}_{w,c}$ est la matrice des paramètres extrinsèques. Elle regroupe le vecteur translation $\vec{\mathbf{t}} = [t_x, t_y, t_z]^T$ et la matrice de rotation $\mathbf{R} = \mathbf{R}_x(\theta)\mathbf{R}_y(\phi)\mathbf{R}_z(\psi)$ qui se décom-

pose en trois rotations respectivement autour des axes \vec{X}_w , \vec{Y}_w et \vec{Z}_w du repère monde. Ces trois matrices, à partir desquelles les paramètres r_{ij} de l'équation (2.1) peuvent être calculés, sont définies comme suit :

$$\mathbf{R}_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}; \quad \mathbf{R}_y(\phi) = \begin{bmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{bmatrix};$$

$$\mathbf{R}_z(\psi) = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.2)$$

avec θ , ϕ et ψ les angles de rotation autour des axes \vec{X}_w , \vec{Y}_w et \vec{Z}_w .

Transformation T2 : Elle définit la projection du point de coordonnées ${}^c\mathcal{P}_{3D}$ dans le plan image de la caméra. Cette projection est obtenue en utilisant le théorème de Thalès. ${}^c\mathcal{P}_{3D}$ est projeté dans le repère rétinien de l'image $\{R\}$ dont l'origine est la projection du centre optique dans le plan image et dont les axes \vec{X}_R , \vec{Y}_R sont respectivement parallèles à \vec{X}_c et \vec{Y}_c . Les coordonnées (x_u, y_u) du point projeté P_u s'écrivent

$$x_u = f \frac{x_c}{z_c} \text{ et } y_u = f \frac{y_c}{z_c}, \quad (2.3)$$

où f est la distance focale de la caméra. Sur la figure 2.2, f est défini comme la distance entre le centre optique O_c et sa projection orthogonale O_R sur le plan de l'image. Par analogie, f est aussi représenté dans la figure 2.1 par la distance entre le sténopé et le plan image. Nous pouvons reformuler l'équation (2.3) sous la forme d'une équation matricielle homogène :

$$z_c \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}. \quad (2.4)$$

Il est à noter que le point P_u ne tient pas compte de la distorsion induite par l'optique. La transformation 3 a pour but de modéliser cette distorsion comme une transformation 2D dans le plan image.

Transformation T3 Dans le cas général, les distorsions dans les images ont une composante radiale et une composante tangentielle. La première est due à la symétrie des lentilles sphériques, la deuxième aux défauts d’alignement des lentilles. La distorsion tangentielle est négligeable pour la majorité des caméras actuelles. La distorsion radiale qui est la plus significative apparaît pour des distances focales allant de 2,8 à 12 mm [Svoboda, 2005]. Les rayons de lumière qui passent par les bords de la lentille entraînent une distorsion plus importante que celle causée par les rayons passant au centre de la lentille. La figure 2.3 représente schématiquement ce phénomène.

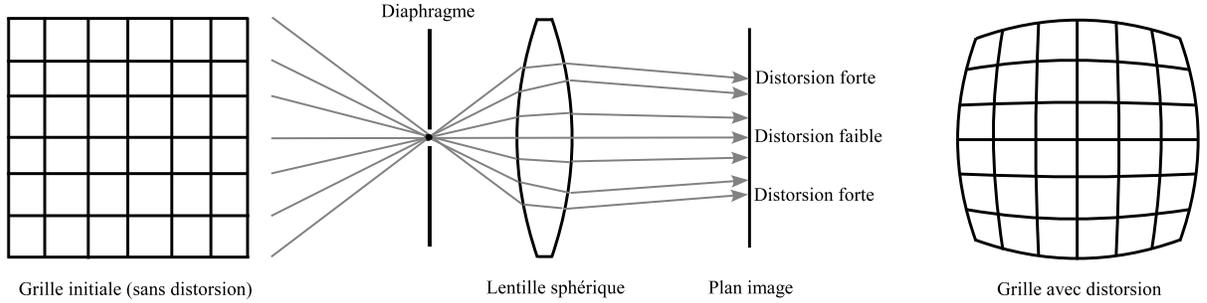


FIGURE. 2.3 – Illustration de l’effet de la distorsion radiale sur un quadrillage initial rectiligne. Les rayons qui ne passent pas par le centre de la lentille sphérique sont déviés des positions idéales dans le plan image. Les zones de la lentille qui sont plus éloignées du centre produisent des déviations plus importantes que les zones centrales.

La distorsion radiale peut être modélisée par un polynôme en r^2 d’ordre n (2.5) qui déplace un point $P_u = (x_u, y_u)^T$ de l’image en un point $P_d = (x_d, y_d)^T$ avec distorsion radiale.

$$\begin{cases} x_d = x_u(1 + k_1r^2 + k_2r^4 + \dots + k_nr^{2n}), \\ y_d = y_u(1 + k_1r^2 + k_2r^4 + \dots + k_nr^{2n}), \\ r = \sqrt{(x_u)^2 + (y_u)^2} \end{cases} \quad (2.5)$$

Les k_i sont les coefficients qui génèrent la distorsion radiale alors que r est la distance entre P_u et le centre optique de l’image (O_R dans la figure 2.2). Il a été montré dans [Miranda-Luna *et al.*, 2004] qu’un polynôme d’ordre 2 ($n = 2$) est suffisant pour modéliser précisément des distorsions importantes comme celles produites par un système endoscopique.

Transformation T4 : Un dernier changement de repère permet de passer du point P_d exprimé dans le système de coordonnées du repère rétinien de l’image $\{R\}$ au point P_{im} de coordonnées $(x_{im}, y_{im})^T$ dans le repère qui prend pour origine le coin supérieur gauche de l’image. Les paramètres de cette transformation (voir Fig. 2.2) sont les dimensions S_x

et S_y (exprimées en micromètres) des capteurs (CCD) de la caméra et les coordonnées (u_0, v_0) du centre optique O_R dans le repère $\{I\}$.

$$\begin{bmatrix} x_{im} \\ y_{im} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{S_x} & 0 & u_0 \\ 0 & \frac{1}{S_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} \quad (2.6)$$

D'une manière pratique et dans le cas où les capteurs CCD sont carrés ($S_x = S_y$), les paramètres intrinsèques de la caméra peuvent être regroupés d'une manière plus compacte dans une matrice 3×3 que nous appelons \mathcal{K} . Dans ce cas, la modélisation de la distorsion radiale est alors appliquée en tenant compte du centre optique (u_0, v_0) après points images de coordonnées $({}^u x_{im}, {}^u y_{im})$ obtenues par la projection des points 3D du repère $\{C\}$ dans le repère $\{I\}$ en utilisant la matrice \mathcal{K} :

$$z_{3D}^{i,k} \begin{bmatrix} {}^u x_{im} \\ {}^u y_{im} \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{f}{S_x} & 0 & u_0 \\ 0 & \frac{f}{S_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathcal{K}} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (2.7)$$

Dans l'équation (2.7), l'exposant u indique que les points de coordonnées $({}^u x_{im}, {}^u y_{im})$ sont sans distorsion et pas visible dans les images. Cependant, si les distorsions radiales sont préalablement corrigées dans les images, la projection devient linéaire et $({}^u x_{im}, {}^u y_{im})$ deviennent (x_{im}, y_{im}) .

Ensemble de paramètres projectifs. L'ensemble de ces paramètres (intrinsèques et extrinsèques) est estimé par des méthodes de calibrage (détermination de la fonction de transfert d'une caméra). Un grand nombre de méthodes de calibrage de caméras sont recensées dans la littérature (voir [Salvi *et al.*, 2002] pour une étude comparative).

2.1.2 Mosaïquage d'images

2.1.2.1 Imagerie panoramique

Le mosaïquage d'images (« *image mosaicing* » en anglais) est une méthode qui permet de placer l'ensemble des images consécutives d'une séquence vidéo dans un repère unique pour obtenir une représentation panoramique de la scène observée. Cette représentation, appelée mosaïque ou image panoramique, est construite en trouvant les relations géométriques entre les images de la séquence. Ces transformations géométriques sont ensuite

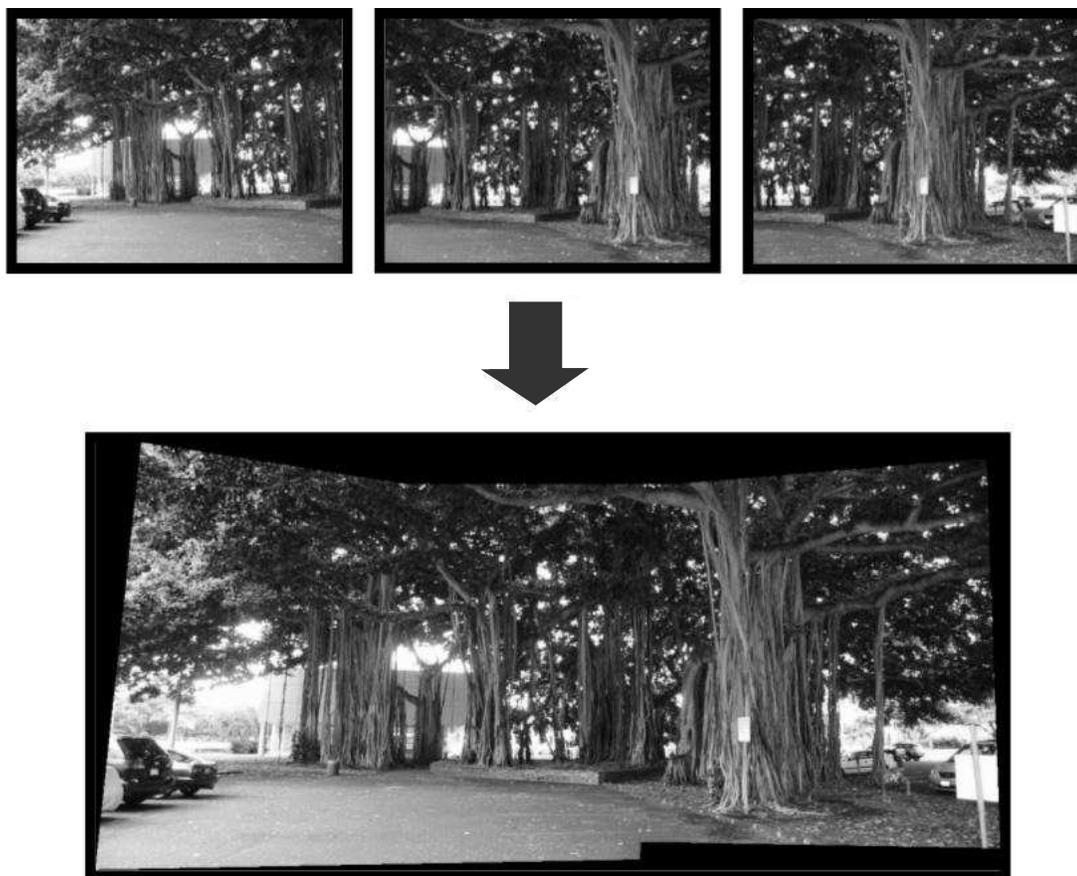


FIGURE. 2.4 – Exemple d'image panoramique construite à partir de trois images.

utilisées pour placer l'ensemble des images dans un référentiel (un système de coordonnées) commun. Les liens géométriques entre les images sont obtenus par des algorithmes de *recalage d'images*. Ces derniers peuvent être mis en œuvre efficacement pour des images qui présentent un recouvrement minimal (taille minimale des régions homologues). Le mosaïquage d'images est un moyen efficace pour augmenter le champ de vision effectif d'une scène et pour supprimer la redondance des informations. La figure 2.4 illustre ce procédé par un exemple de mosaïquage issu du domaine de la photographie grand public.

2.1.2.2 Recalage d'images

Le recalage d'images (de l'anglais « *image registration* ») est dit *monomodal* lorsque les images sont toutes acquises avec le même type de capteur. C'est le cas dans notre étude.

Soit deux images \mathbf{I}_{sc} (image source à transformer) et \mathbf{I}_{ci} (image cible). Recaler ces deux images revient à trouver, les paramètres d'une transformation géométrique $\tilde{\mathbf{T}}_{2D}$ au moyen d'une méthode d'optimisation $\arg \text{opt}$, qui maximise la mesure de ressemblance (notée sim) entre des structures homologues détectées par un algorithme de segmentation g dans les deux images. Ce principe est formulé dans l'équation (2.8). Dans le cas où les images ne présentent pas de primitives géométriques détectables d'une manière suffisamment robuste, aucune fonction g n'est appliquée aux images et la mesure de similarité peut être calculée directement à partir des niveaux de gris (ou des couleurs) des pixels des images. Ω désigne la méthode d'interpolation appliquée à l'image \mathbf{I}_{sc} transformée par \mathbf{T}_{2D} pour recalculer les niveaux de gris ou les couleurs de l'image transformée.

$$\tilde{\mathbf{T}}_{2D} = \arg \text{opt}_{\mathbf{T}_{2D}} \text{sim} (g(\mathbf{I}_{ci}), \Omega(g(\mathbf{T}_{2D}(\mathbf{I}_{sc})))) . \quad (2.8)$$

Les structures homologues, la mesure de similarité, le type de la transformation géométrique et la méthode d'optimisation sont à choisir selon l'application et la nature des images, entre autres. Dans le cas des examens cystoscopiques de la vessie, les images ne contiennent pas de primitives géométriques (contours, points saillants) détectables d'une manière suffisamment rapide, robuste et systématique dans toute la séquence d'images. Cela est dû à la grande variabilité inter- et intra-patients en terme de texture de la paroi interne de la vessie. Les valeurs des pixels (niveaux de gris) des deux images constituent la seule information disponible pour effectuer le recalage [Miranda-Luna *et al.*, 2008; Hernández-Mier *et al.*, 2010].

La vessie est un organe mou qui est sujet à des déformations élastiques, notamment lorsqu'il est en contact avec d'autres organes. Cependant, lors d'un examen cystoscopique

clinique standard, la vessie est remplie avec du sérum physiologique afin de tendre sa paroi et faciliter l'examen visuel, ce qui par ailleurs rigidifie l'organe et minimise les mouvements de la paroi. Ainsi, compte tenu de la vitesse d'acquisition des images en endoscopie (au moins 25 images par seconde), la paroi interne de la vessie peut être considérée comme rigide entre deux acquisitions séparées de 40 ms (ou moins). De plus, du point de vue anatomique, la géométrie de la vessie est approximativement ellipsoïdale et le cystoscope ne permet donc de visualiser que des petites surfaces quasi-planaires de la surface de la paroi [Hernández-Mier *et al.*, 2010; Miranda-Luna *et al.*, 2008]. Si l'on suppose que les seules non-linéarités de l'instrument (distorsions radiales liées aux courtes distances focales) peuvent être corrigées, alors il est possible de modéliser précisément les transformations non linéaires entre les images cystoscopiques par une **transformation perspective**. C'est la transformation géométrique linéaire la plus générale.

2.1.2.3 Transformation perspective

Le déplacement d'un point de coordonnées (x, y) dans l'image source \mathbf{I}_{sc} aux coordonnées (x', y') du point homologue dans le repère de l'image cible \mathbf{I}_{ci} est formulé par :

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{w} \begin{bmatrix} u \\ v \end{bmatrix}, \text{ avec } \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \underbrace{\begin{bmatrix} \underbrace{S_f \cos(\phi)}_{a_{11}} & \underbrace{-S_x \sin(\phi)}_{a_{12}} & \underbrace{t_x}_{a_{13}} \\ \underbrace{S_y \sin(\phi)}_{a_{21}} & \underbrace{S_f \cos(\phi)}_{a_{22}} & \underbrace{t_y}_{a_{23}} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}_{\mathbf{T}_{2D}} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.9)$$

où (t_x, t_y) désigne les translations dans le plan image, S_f est un facteur d'échelle, ϕ est l'angle de rotation dans le plan, (S_x, S_y) sont les paramètres de cisaillement, (a_{31}, a_{32}) sont les paramètres de perspective et $a_{33} = 1$.

Les coordonnées (x', y') sont, la plupart du temps, non entières dans l'image transformée. Une méthode d'interpolation est donc nécessaire pour estimer le niveaux de gris de l'image transformée pour la position (x', y') . Les méthodes les plus simples sont les interpolations par *le plus proche voisin*, *bilinéaire* et *bicubique*.

Le recalage d'images a trouvé de nombreuses applications dans le domaine de la médecine. Des synthèses bibliographiques très complètes sont disponibles dans [Brown, 1992; Zitová and Flusser, 2003; Maintz and Viergever, 1998; Pluim *et al.*, 2003]. Des exemples d'applications existent en mammographie [Jalink *et al.*, 1996], en angiographie X [Chou *et al.*, 1997], en ophtalmologie [Can *et al.*, 2002] ou en microscopie [Vercauteren *et al.*, 2005].

2.1.3 Mosaïquage d'images cystoscopiques

Le mosaïquage de la paroi interne de la vessie pose des problèmes scientifiques spécifiques qui sont autant de verrous à lever. Premièrement, les images ne contiennent pas de primitives détectables et présentent une grande variabilité inter- et intra-patient en termes de textures. Deuxièmement, il est impossible d'avoir des informations sur la position ou le mouvement du cystoscope dans la mesure où le clinicien le déplace librement. Finalement, une séquence classique d'images cystoscopiques (plusieurs minutes) comporte un nombre très important d'images (plusieurs milliers). Dans les applications citées précédemment (mammographie, angiographie, *etc.*), le problème du mosaïquage est plus simple car le nombre d'images est plus faible et de plus, il est possible d'utiliser des connaissances *a priori* sur les mouvements entre points de vue.

Dans la littérature on trouve très peu de travaux sur le mosaïquage de la paroi interne de la vessie. Récemment [Behrens, 2008; Behrens *et al.*, 2009; Behrens *et al.*, 2010] se sont intéressés à la construction d'images panoramiques de fluorescence de la paroi interne de la vessie. Ces images diffèrent de celles acquises en lumière blanche en cystoscopie conventionnelle (modalité la plus répandue). La modalité de fluorescence induite améliore la détection des lésions cancéreuses dans les tissus de la vessie. La lumière fluorescente améliore aussi le contraste des structures vasculaires, ce qui facilite la détection de primitives dans les images cystoscopiques (voir figure 2.5). L'algorithme SIFT (*Scale Invariant Feature Transform*) [Lowe, 2004] est utilisé dans [Behrens, 2008] pour la détection de points de référence. Les points détectés dans deux images consécutives sont appariés en utilisant l'algorithme RANSAC (RANDOM SAMPLE CONSENSUS) [Fischler and Bolles, 1981] pour guider le recalage des paires d'images consécutives.

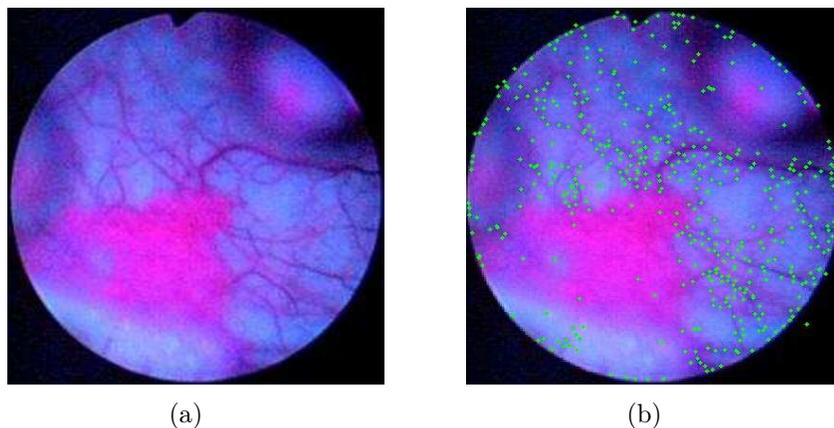


FIGURE. 2.5 – (a) Exemple d'image cystoscopique de fluorescence [Behrens, 2008]. Cette image présente des structures vasculaires bien contrastées et détectables d'une manière robuste. (b) Points de référence détectés par la méthode SIFT.

Si cette approche est rapide et efficace pour le recalage d'images cystoscopiques de fluorescence (peu répandue), elle n'est pas applicable pour des images cystoscopiques acquises en lumière blanche (examen le plus fréquent). En effet, pour cette dernière modalité, aucune primitive image n'est détectable d'une manière systématique et robuste. Les niveaux de gris des pixels des images demeurent alors la seule information disponible pour le recalage d'images cystoscopiques.

Au sein du CRAN deux thèses ont été consacrées à la cartographie 2D. [Miranda-Luna, 2005] a proposé un premier algorithme de mosaïquage (A_{IM}) basé sur la maximisation de l'information mutuelle entre les images à recaler. Les travaux de [Hernandez-Mier, 2007] ont permis de proposer une nouvelle approche avec un algorithme (A_{DQ}) plus rapide basé sur la minimisation de la distance quadratique entre les niveaux de gris des images à recaler. Nous avons effectué une étude comparative de ces deux algorithmes de construction d'images panoramiques de la paroi interne de la vessie en termes de robustesse, de précision et de vitesse de calcul. Cette étude a été menée avec des données réelles extraites d'examens cystoscopiques cliniques et sur des données acquises pour un fantôme de vessie en contrôlant le déplacement 3D de l'endoscope. Avant de détailler les résultats de cette étude, nous décrivons brièvement le principe des deux traitements pour le recalage de deux images consécutives.

2.1.3.1 A_{DQ} : Algorithme basé sur la distance quadratique

Soit \mathbf{I}_k la $k^{\text{ème}}$ image d'une séquence de K images cystoscopiques. L'algorithme A_{DQ} utilise la mesure de dissimilarité \mathcal{S}_{DQ} définie par la somme des distances quadratiques entre les niveaux de gris des pixels de l'image cible \mathbf{I}_k et ceux des pixels homologues de l'image source $\mathbf{T}_{2D}(\mathbf{I}_{k+1})$ transformée. L'équation (4.20) formule cette mesure de dissimilarité.

$$\mathcal{S}_{DQ}(\mathbf{I}_k, \mathbf{T}_{2D}(\mathbf{I}_{k+1})) = \sum_{\mathbf{I}_k \cap \mathbf{T}_{2D}(\mathbf{I}_{k+1})} (\mathbf{I}_k - \mathbf{T}_{2D}(\mathbf{I}_{k+1}))^2. \quad (2.10)$$

Lorsque les régions homologues de $\mathbf{T}_{2D}(\mathbf{I}_{k+1})$ et \mathbf{I}_k se superposent, \mathcal{S}_{DQ} tend vers 0. La minimisation de \mathcal{S}_{DQ} est réalisée par une méthode itérative basée sur l'algorithme de Baker et Matthews [Baker and Matthews, 2004] qui estime le flot optique entre les images à recaler.

2.1.3.2 A_{IM} : Algorithme basé sur l'information mutuelle

L'algorithme A_{IM} est basé sur l'approche EMMA (*EMpirical entropy Manipulation and Analysis*) de Viola et Wells [Viola and Wells III, 1997]. A_{IM} aligne les images \mathbf{I}_k et

\mathbf{I}_{k+1} en maximisant l'information mutuelle entre \mathbf{I}_k et $\mathbf{T}_{2D}(\mathbf{I}_{k+1})$. L'information mutuelle est une mesure statistique qui combine les entropies $H(\mathbf{I}_k)$ et $H(\mathbf{T}_{2D}(\mathbf{I}_{k+1}))$ des parties communes des images \mathbf{I}_k et $\mathbf{T}_{2D}(\mathbf{I}_{k+1})$ et l'entropie conjointe $H(\mathbf{I}_k, \mathbf{T}_{2D}(\mathbf{I}_{k+1}))$. Cette mesure de similarité s'exprime par :

$$\mathcal{S}_{IM}(\mathbf{I}_k, \mathbf{T}_{2D}(\mathbf{I}_{k+1})) = H(\mathbf{I}_k) + H(\mathbf{T}_{2D}(\mathbf{I}_{k+1})) - H(\mathbf{I}_k, \mathbf{T}_{2D}(\mathbf{I}_{k+1})) \quad (2.11)$$

La maximisation de \mathcal{S}_{IM} est réalisée à l'aide de l'algorithme du gradient où le critère \mathcal{S}_{IM} et son gradient sont évalués de façon stochastique (les pixels dont les niveaux de gris sont utilisés pour estimer l'information mutuelle sont tirés au sort à chaque itération).

2.1.4 Étude comparative

Pour effectuer la comparaison des performances (robustesse, précision et vitesse) des deux algorithmes A_{DQ} et A_{IM} , nous commençons par définir le critère d'arrêt utilisé : nous considérons que ces algorithmes itératifs ont convergé lorsque les transformations \mathbf{T}_{2D} obtenues pour deux itérations consécutives conduisent à un déplacement moyen des pixels de l'image source qui est inférieur à la longueur d'un pixel. En effet, il a été constaté que pour ce seuil de déplacement, la position de l'image source sur l'image cible ne varie plus de façon significative et que la cohérence visuelle est assurée par cette superposition (pas de discontinuité de textures entre les images [Miranda-Luna *et al.*, 2008]).

2.1.4.1 Robustesse

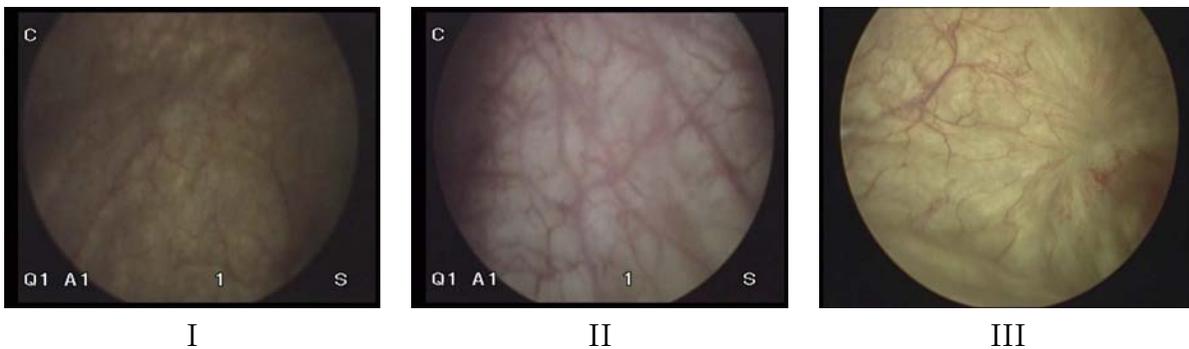


FIGURE. 2.6 – I, II et III : Images extraites de trois séquences cystoscopiques cliniques. Les images choisies présentent une forte variabilité de texture, de contraste et d'intensité lumineuse. Cette variabilité est due à la fois à la scène (variabilité inter-patients) et à l'instrument (variabilité engendrée par les cystoscopes).

Pour évaluer quantitativement la robustesse du recalage aux différents types de trans-

formations géométriques (translation, rotation, changement de perspective), nous avons utilisé trois images extraites de trois séquences cystoscopiques cliniques (voir figure 2.6). Les images choisies présentent des niveaux d’illumination et des textures très variables. Pour nos tests, ces données font office d’images cibles notées \mathbf{I}_0 . Les images sources \mathbf{I}_1 sont obtenues en simulant un mouvement 3D du cystoscope. Plus précisément, les images \mathbf{I}_1 sont calculées en appliquant aux images références \mathbf{I}_0 une transformation $\mathbf{T}(= \mathbf{T}_{2D}^{-1})$ déterminée à partir du déplacement 3D simulé. Conformément au modèle (2.9), les transformations \mathbf{T}_{2D} à déterminer incluent trois translations (deux translations t_x et t_y et une troisième translation t_z correspondant au facteur d’échelle S_f) et trois rotations (une rotation ϕ dans le plan image et deux rotations hors plan ψ et α liées aux paramètres a_{31} et a_{32} dans l’équation (2.9)). La simulation des couples d’images $(\mathbf{I}_0, \mathbf{I}_1)$ pour différentes transformations nous a permis de déterminer les bornes des intervalles de variation des valeurs des paramètres conduisant à un recalage réussi quand le déplacement moyen des quatre coins de l’image (voir tableau 2.1).

Même si pour les deux méthodes, les limites des translations sont approximativement les mêmes (avec un léger avantage en faveur de l’information mutuelle), la méthode basée sur l’information mutuelle témoigne de la plus grande robustesse vis-à-vis des variations du facteur d’échelle et des variations des rotations dans et hors plan. Cette robustesse vis-à-vis des changements de perspective et d’échelle est importante lorsque le clinicien décide de changer radicalement la trajectoire du cystoscope pour aller d’une région de la vessie à une autre ou d’un point de repère (urètre par exemple) à un autre.

2.1.4.2 Précision

Pour évaluer quantitativement la précision du recalage, nous avons construit un fantôme à partir d’une vessie de porc. Pour les urologues, les textures des parois de vessies de porcs sont visuellement très proches, voire indissociables, de celle des vessies humaines (voir figure 2.7 (a)). La vessie est incisée, ouverte, mise à plat et photographiée dans sa globalité avec un appareil photographique haute définition. Ensuite, l’acquisition d’une séquence d’images est simulée numériquement à partir de cette photographie. La surface couverte dans cette simulation est de l’ordre de 16 cm^2 . Comme illustré sur la figure 2.7 (a), la première image de la séquence se situe dans le coin supérieur gauche de la photographie. La séquence des images est simulée selon un scénario de transformations perspectives comportant, notamment, des translations horizontales de 10 pixels (10 images sur la trajectoire horizontale en haut de l’image), des combinaisons de translations verticales de 10 pixels et de rotations dans le plan de 2° (10 premières images sur la verticale à partir du coin supérieur droit), des combinaisons de translations verticales de 10 pixels et de

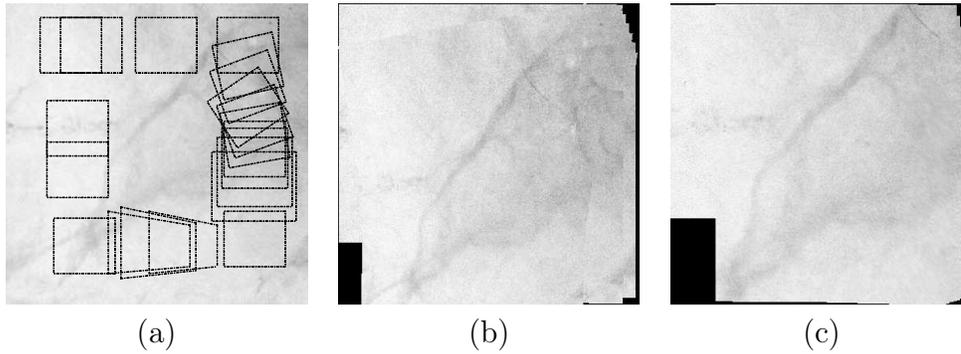


FIGURE. 2.7 – (a) Photographie d'une vessie de porc utilisée comme fantôme : les carrés indiquent la séquence d'images simulées. Pour des raisons de lisibilité, chaque carré a été réduit mais en réalité, l'ensemble des carrés couvrent la totalité de l'image (y compris la zone centrale) ce qui explique que l'on est capable de reconstruire une image panoramique qui recouvre l'ensemble de la photographie. Images panoramiques générées à partir de la séquence d'images simulées par l'algorithme A_{IM} (b) et A_{DQ} (c).

changements de facteur d'échelle de 5% (10 dernières images sur la verticale à droite de l'image) et des translations horizontales de 10 pixels (10 images sur l'horizontale en bas de l'image)

Chaque paire d'images ($\mathbf{I}_k, \mathbf{I}_{k+1}$) de la séquence est recalé en utilisant chacun des deux algorithmes. La mesure d'erreur $\bar{\epsilon}_{k,k+1}$ utilisée pour quantifier la précision du recalage correspond à la moyenne des distances euclidiennes entre les positions des pixels homologues des images cibles \mathbf{I}_k et recalées $\mathbf{T}_{2D}(\mathbf{I}_{k+1})$. Idéalement, $\bar{\epsilon}_{k,k+1} = 0$ (en d'autres termes, \mathbf{I}_{k+1} et \mathbf{I}_k sont parfaitement superposées). Ainsi illustré sur le graphique de la figure 2.8, dans le cas de translations pures, puis de la combinaison des translations avec les rotations hors plan puis dans le plan, les deux algorithmes donnent des résultats relativement comparables pour les images [1-15] et [30-50], avec des erreurs moyennes de recalage $\bar{\epsilon}_{k,k+1}$ valant respectivement 0.2, 0.6 et 3.5 pixels pour ces combinaisons.

TABLEAU. 2.1 – Robustesse définie par les limites obtenues pour chaque paramètre pour les algorithmes A_{DQ} et A_{IM} . La dernière colonne fournit les intervalles dans lesquels s'inscrivent 90% des transformations d'un examen cystoscopique standard [Hernandez-Mier, 2007].

Paramètres de la transformation	Limites des intervalles		
	A_{DQ}	A_{IM}	Intervalle type d'un examen
Translation (t_x et t_y)	± 25 pixels	± 30 pixels	± 5 pixels
Facteur d'échelle (S_f)	$\pm 15\%$	$\pm 25\%$	$\pm 2\%$
Rotation dans le plan (ϕ)	$\pm 10^\circ$	$\pm 20^\circ$	$\pm 1^\circ$
Rotations hors plan (ψ and α)	$\pm 10^\circ$	$\pm 20^\circ$	$\pm 1^\circ$

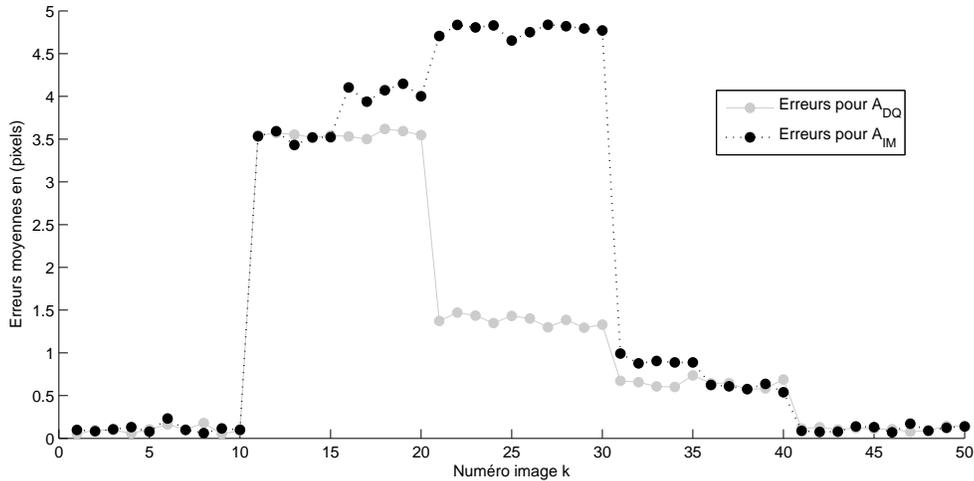


FIGURE. 2.8 – Erreurs moyennes de recalage $\bar{\epsilon}_{k,k+1}$ pour les deux algorithmes A_{DQ} et A_{IM} . Pour l'image numéro k en abscisse, l'ordonnée donne la valeur de $\bar{\epsilon}_{k,k+1}$. Les valeurs de $\bar{\epsilon}_{k,k+1}$ sont quasiment équivalentes pour les deux algorithmes pour les images [1-15] et [30-50] dans la plus part de la séquence. Dans la partie de changement de facteur d'échelle (images [20-30]), A_{DQ} donne de meilleurs résultats.

Dans le cas des translations combinées à un changement du facteur d'échelle (images [12-30]), l'algorithme A_{DQ} donne de bien meilleurs résultats ($\bar{\epsilon}_{k,k+1} \approx 1.5$ pixels) que l'algorithme A_{IM} ($\bar{\epsilon}_{k,k+1} \approx 4.5$ pixels). Ces erreurs de recalage restent, cependant, suffisamment petites pour ne pas affecter la cohérence visuelle de la carte générée (comme visible sur la figure 2.7 (b), pour l'algorithme basé sur l'information mutuelle, la carte reste cohérente en terme de textures et similaire à l'image originale dans la figure 2.7 (a)).

Il est important de souligner que les valeurs des transformations choisies pour la simulation des images tests sont bien plus grandes que les valeurs des transformations typiquement rencontrées au cours de l'examen clinique. En effet, lors d'un examen, les images sont acquises avec une fréquence de 25 images/seconde et la vitesse de déplacement du cystoscope est au maximum de quelques millimètres/seconde. Dans ces conditions, les variations des valeurs des paramètres de transformation reliant deux images successives sont relativement petites : moins de 5 pixels en translations avec des rotations ne dépassant pas 1° et un facteur d'échelle ne dépassant pas 5%. Les tests sur ce fantôme indiquent que dans le cas d'un examen cystoscopique réel, les deux méthodes donnent des erreurs de recalage inférieures à un pixel. Les figures 2.9 (a) et 2.9 (b) montrent deux images panoramiques construites à partir d'images cystoscopiques cliniques.

La figure 2.9 (a) présente une image panoramique constituée de 650×182 pixels et



FIGURE. 2.9 – Exemple de deux images panoramiques construites à partir de deux séquences d’images cystoscopiques en utilisant l’algorithme A_{IM} (a) et A_{DQ} (b).

construite à partir de 500 images en utilisant l’algorithme A_{IM} . Sur cette image panoramique, il n’y a pas de discontinuités visibles de la texture, ce qui témoigne d’une bonne cohérence visuelle. L’image panoramique de la figure 2.9 (b) a une taille de 1479×1049 pixels. Elle est construite à partir de 450 images en utilisant l’algorithme A_{DQ} . Sur cette image panoramique, deux polypes sont visibles simultanément et facilement localisables l’un par rapport à l’autre.

2.1.4.3 Vitesse de calcul

Les algorithmes A_{IM} et A_{DQ} ont été implémentés en C et en utilisant la bibliothèque libre de vision par ordinateur OpenCV. L’évaluation de ces deux algorithmes est faite dans les mêmes conditions d’expérimentation, c’est-à-dire avec un ordinateur de 2 Go de RAM muni d’un processeur Intel Dual core(TM) à 2.40 GHz. La méthode d’optimisation de l’algorithme A_{IM} nécessite en moyenne 250 itérations et entre 50 à 60 secondes pour le recalage de deux images consécutives. Les résultats de ce test sont obtenus pour une solution initiale correspondant à la transformation unitaire. Ce choix de solution initiale a été fait ici, car il s’agit de tests du comportement des algorithmes, mais en pratique, lorsqu’on traite une séquence d’images, nous préconisons de choisir pour initialisation la transformation estimée au temps précédent, ce qui permet de réduire significativement le

nombre d'itérations de l'algorithme¹

La construction de l'image panoramique de la figure 2.9 (a) nécessite aussi 8 heures 27 minutes de temps de calcul. Dans les mêmes conditions, l'algorithme A_{DQ} est 100 fois plus rapide que A_{IM} . En moyenne, A_{DQ} nécessite 12 itérations pour recalculer deux images consécutives et le temps de recalage varie entre 0.3 à 0.6 seconde par paire d'images. L'image panoramique (b) de la figure 2.9 est construite en 3 minutes 20 secondes. L'algorithme A_{DQ} permet ainsi la construction d'images panoramiques partielles de la surface interne de la vessie au cours de l'examen cystoscopique.

2.1.5 Bilan

L'étude comparative des deux algorithmes A_{DQ} et A_{IM} a permis de cerner leurs limites et avantages. En terme de précision de recalage, les deux algorithmes donnent des résultats comparables avec un avantage en faveur de l'algorithme A_{DQ} lors des changements de facteur d'échelle. Cependant, l'algorithme A_{IM} est globalement plus robuste que A_{DQ} surtout lorsque plusieurs paramètres subissent simultanément des changements significatifs mais pas pour des valeurs réalistes des paramètres. L'algorithme A_{DQ} est environ 100 fois plus rapide (quelques dixièmes de secondes pour recalculer deux images).

Les résultats obtenus avec les deux algorithmes de cartographie 2D conçus au CRAN sont jugés comme étant convaincants par les urologues. La cohérence des résultats a été confirmée par des urologues et des chirurgiens *via* une inspection visuelle des cartes reconstruites pour diverses données cliniques. Toutefois, la vessie constitue un espace volumique que les urologues et les chirurgiens se représentent mentalement en 3D.

Ainsi, une représentation 3D des surfaces internes de la vessie a pour but d'améliorer l'interprétation des données acquises lors d'un examen cystoscopique grâce à l'augmentation du réalisme du support cystoscopique. De plus, reconstruire la géométrie des parois rendrait possible le calcul de nouveaux critères de diagnostic comme des indices liés à la distribution spatiale des lésions d'intérêt dans le suivi de l'évolution des tissus d'un examen à un autre. C'est donc pour toutes ces raisons que mon travail s'est orienté vers la cartographie 3D.

1. Le critère de temps est beaucoup moins important que les critères de robustesse et de précision, car en pratique, le clinicien n'examine les images panoramiques qu'un jour après l'examen cystoscopique.

2.2 Endoscopie 3D

La reconstruction d'informations 3D à partir d'une séquence vidéo est un sujet de recherche très en vogue dans le domaine de la vision par ordinateur. On trouve diverses approches élaborées selon la nature des applications et des objets à numériser (statique ou en mouvement, déformable ou rigide, *etc.*) [Butime *et al.*, 2010]. Dans le cadre de l'endoscopie, les méthodes dites optiques sont les mieux adaptées pour la reconstruction d'informations 3D. Ces méthodes sont regroupées en deux familles : les méthodes passives et les méthodes actives. La vision passive se base sur l'acquisition d'images de la scène à partir de points de vue différents alors que la vision active consiste à associer des caméras avec des sources de lumière structurée. Le concept de lumière structurée sera défini en détail dans les paragraphes ci-dessous. Dans ce qui suit, nous présentons les principaux travaux en endoscopie 3D utilisant ces deux approches de vision.

2.2.1 Vision passive

La vision passive fait référence aux systèmes standards d'acquisition d'images qui ne nécessitent pas un éclairage structuré supplémentaire de la scène. Les seules données disponibles sont acquises à partir d'une ou plusieurs caméras. En endoscopie, on distingue deux types d'instruments commercialisés basés sur la vision passive.

- Les endoscopes monoculaires. Outre le canal d'entrée pour l'éclairage de la scène, ils sont équipés d'un seul canal optique d'acquisition d'images qui se termine par une caméra. Ce sont les endoscopes les plus anciens et les plus utilisés, notamment pour les examens cystoscopiques.
- Les endoscopes binoculaires. Ils sont équipés de deux canaux optiques d'acquisition d'images (voir figure 2.10) se terminant chacun par une caméra. Un des premiers endoscopes binoculaires est celui conçu par la société Olympus. Il est utilisé dans les systèmes de chirurgie robotisée (par exemple, le système Da Vinci de la société Intuitive Surgical Inc.). Ces endoscopes binoculaires autorisent une perception tridimensionnelle de la zone opératoire. Cependant, pour cette perception 3D, ces instruments nécessitent l'utilisation de dispositifs particuliers comme des casques, des lunettes polarisées *etc.* D'autre part, une telle configuration matérielle ne permet qu'une perception visuelle de la profondeur et non pas une reconstruction 3D des surfaces.

Les méthodes de reconstruction d'informations 3D à partir de systèmes de vision passive sont nombreuses (voir [Hartley and Zisserman, 2003; Horn and Brooks, 1989; Zhang *et al.*, 1999]). Cependant, seules quelques unes sont exploitables en endoscopie. Ces méthodes sont dé-



FIGURE. 2.10 – Tête d'un endoscope stéréoscopique. L'endoscope est équipé de trois canaux optiques : deux pour l'acquisition d'images avec deux caméras et le troisième (le plus petit) pour l'illumination de la scène en lumière blanche.

taillées dans la suite.

Reconstruction 3D à partir du mouvement. Ces méthodes sont appelées « *structure from motion* » (SFM) en anglais. Les images utilisées par ces méthodes sont acquises par une caméra en mouvement balayant une scène à reconstruire en 3D. Le principe de ces méthodes est basé sur la détection de points de références dans la séquence d'images en utilisant des filtres particuliers comme le détecteur de Harris [Harris and Stephens, 1988] et le SIFT [Lowe, 2004]. Comme les images se recouvrent (c'est-à-dire, présentent des régions homologues), les points de référence détectés doivent être appariés. Il est ainsi possible de reconstruire le mouvement 3D de la caméra et donc la structure de la scène imagée en se basant sur la géométrie épipolaire. [Hartley and Zisserman, 2003] présente un état de l'art complet sur ces techniques. Dans [Deguchi *et al.*, 1996; Thormahlen *et al.*, 2002], ces techniques sont utilisées en endoscopie monoculaire pour la reconstruction des formes anatomiques 3D (à une échelle près) d'organes comme l'estomac et le côlon. Les images utilisées dans ces travaux présentent des primitives assez facile à détecter.

Reconstruction 3D à partir de la stéréo vision. Les images utilisées dans ces techniques sont acquises avec un système de vision stéréoscopique passive. Ce dernier est formé par deux caméras dont la géométrie (paramètres intrinsèques et extrinsèques) est préalablement connue (calibrée). Les champs de vision des deux caméras doivent se recouvrir de telle manière que des régions homologues sont acquises par les deux caméras. Cette géométrie multi-images permet, par l'intermédiaire d'une méthode de triangulation (intersection de trajectoires rectilignes issues de points homologues des deux images) de retrouver la troisième dimension. Dans le même ordre d'idées que les méthodes SFM,

ces techniques nécessitent de résoudre des problèmes de segmentation et de mise en correspondance de primitives images. [Faugeras, 1993; Hartley and Zisserman, 2003] donnent un état de l'art sur les méthodes de reconstruction 3D reposant sur la vision stéréoscopique passive. En endoscopie binoculaire, on trouve quelques travaux de reconstruction 3D à partir de la stéréo vision [Liu *et al.*, 2006; Mourgues *et al.*, 2001; Stoyanov *et al.*, 2005; Devernay *et al.*, 2001]. Ces travaux reposent sur l'hypothèse que les surfaces imagées contiennent des informations suffisamment riches pour extraire des primitives images. En cystoscopie, les images ne présentent pas de telles primitives détectables d'une manière robuste et précise. Cela compromet l'utilisation des approches basées sur l'utilisation de la vision passive pour la reconstruction 3D des parois internes de la vessie.

Reconstruction 3D à partir des ombrages. Appelée « *shape from shading* » (SFS) en anglais, c'est une technique de reconstruction 3D dont le principe a été établi par Horn au début des années 1970 [Horn, 1989]. Elle consiste à retrouver le relief (« *shape* ») d'une surface imagée à partir d'une seule photographie de cette surface, en exploitant les ombrages (« *shading* »). L'ombrage désigne l'ensemble des phénomènes faisant apparaître un même élément de surface plus ou moins clair sur une photographie.

Cette technique de reconstruction 3D a été exploitée en endoscopie dans plusieurs travaux [Okatani and Deguchi, 1997; Yeung *et al.*, 1999; Forster and Tozzi, 2000; Wang *et al.*, 2009; Wu *et al.*, 2007; Wu *et al.*, 2010]. Ces travaux reposent sur des hypothèses sur la réflectance des surfaces imagées. En orthopédie par exemple, les textures de la surface des os sont homogènes et réfléchissent fortement la lumière blanche, le contraste qui en résulte explique la robustesse des méthodes SFS [Wu *et al.*, 2010]. Dans la vessie, les textures d'images endoscopiques de la paroi présentent une grande variabilité intra- et inter-patients, ce qui rend l'utilisation des techniques de SFS inadaptée dans le cadre de cette étude. De plus, pour ces méthodes appliquées à l'endoscopie, il est difficile d'avoir un réglage de l'intensité de la source de lumière blanche qui puisse autoriser une reconstruction juste et robuste.

2.2.2 Vision active

La deuxième famille de méthodes de reconstruction 3D fait appel à l'utilisation de la vision active. Un système de vision active est composé d'une ou plusieurs caméras et d'une source projetant une lumière structurée (« *structured light* » en anglais). Ces systèmes de vision en lumière structurée trouvent leurs applications dans de nombreux domaines comme l'inspection de pièces manufacturées [Lin and Lin, 1999], la vision robotique [Tamas *et al.*, 2009], *etc.*

Classiquement, un projecteur de lumière LASER (lumière monochromatique) est utilisé pour générer la lumière structurée. La lumière passe à travers une lentille (ou une mire) pour projeter un motif 2D sur la scène à reconstruire. Dans la littérature, plusieurs formes géométriques sont considérées selon le besoin des applications. On peut citer les motifs formés par des droites parallèles [Li *et al.*, 2009], une matrice de points [Dipanda *et al.*, 2003], une grille [Guisser *et al.*, 2000], un ensemble de formes géométriques codées [Salvi *et al.*, 2004].

Physiquement, un point du motif de la mire est projeté sur la surface de l'organe en suivant une trajectoire rectiligne. L'intersection entre la droite 3D et la surface donne un point 3D localisé sur la surface (noté P_{3D}). Lors de l'acquisition d'image, ce point est projeté sur le plan image de la caméra (projection rectiligne et déplacement du point projeté dans le plan image : voir section 2.1). Comme l'illustre la figure 2.11, pour une acquisition donnée, retrouver les coordonnées d'un point P_{3D} correspondant au point P_{2D} dans l'image acquise en supposant qu'elle est sans distorsion revient à trouver l'intersection des deux droites 3D formées par les centres optiques des deux systèmes projectifs et par les projections de P_{3D} sur la surface de la scène et dans le plan image, respectivement.

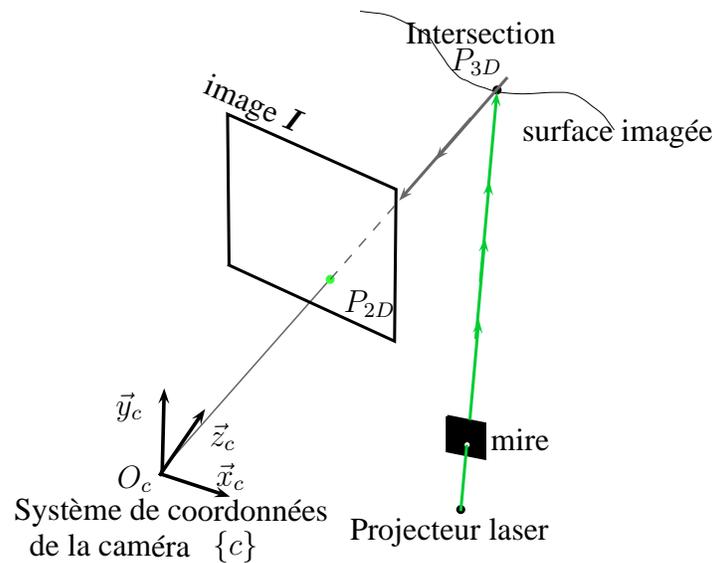


FIGURE. 2.11 – Principe de reconstruction des coordonnées 3D d'un point d'une mire projeté par le projecteur de lumière (point projeté P_{3D} sur la surface de l'organe) et acquis par la caméra à un instant donné.

Calibration. Le système de vision doit être préalablement calibré afin de pouvoir reconstruire les données 3D. Cette étape indispensable inclut à la fois le calibrage de la caméra et l'étalonnage du projecteur de lumière structurée. Pour la caméra, le calibrage

visé à trouver les paramètres relatifs à la projection de chaque élément du motif (paramètres intrinsèques et extrinsèques définis dans la section 2.1.1). Pour le projecteur, le calibrage permet de trouver les paramètres de la trajectoire relative à chaque élément du motif dans l'espace. Par exemple, si le motif projeté est formé de points, le calibrage permet de trouver l'ensemble des droites 3D passant par ces points et toutes issues du centre optique du projecteur.

Défis technologiques. En endoscopie, quelques travaux de recherche ont été consacrés à l'intégration matérielle du principe de vision active dans des instruments dont la finalité réside dans la mesure d'informations 3D relatives à l'anatomie d'organes humains. Ces développements concernent essentiellement l'aspect instrumental de l'endoscopie 3D. Les solutions instrumentales, que ce soit pour le diagnostic dans des organes creux ou la chirurgie mini-invasive, sont fortement conditionnées par la géométrie des endoscopes, à savoir par la longueur et surtout le diamètre des tiges de l'endoscope. Par exemple, en cystoscopie, le diamètre de la tige ne dépasse pas 9 millimètres. Par contre, en chirurgie mini-invasive, il est possible d'utiliser des endoscopes avec un diamètre plus grand, car il y a moins de contraintes anatomiques. En chirurgie mini-invasive trans-abdominale (laparoscopie)², il est possible d'utiliser simultanément deux endoscopes (ou plus) en même temps moyennant plusieurs incisions de la paroi abdominale. Dans le cadre de la cystoscopie, ces contraintes dimensionnelles sont à l'origine du défi à relever pour intégrer la projection de lumière structurée à un système endoscopique. Dans ce cadre, plusieurs solutions ont été proposées. Les travaux de [Iyengar *et al.*, 2001; Hayashibe *et al.*, 2002; Hayashibe *et al.*, 2006] proposent de faire passer la lumière structurée en dehors de la tige du système endoscopique. On peut citer aussi les travaux présentés dans [Fuchs *et al.*, 1998; Keller and Ackerman, 2000] qui traitent du développement d'un laparoscope capable de fournir des informations à la fois de couleur/texture et anatomiques (surfaces 3D). Pour arriver à reconstruire simultanément la couleur et la forme 3D des tissus visualisés, le laparoscope est équipé d'une caméra à très haute vitesse et d'un projecteur numérique de motifs de type DLP (« *Digital Light Processive* », voir figure 2.12). En effet, la caméra peut acquérir des images à une vitesse allant de 30 à 180 images par secondes de manière synchronisée avec la projection de lumière structurée.

Ce développement restreint le champ d'application du système à la laparoscopie. En effet, lors d'un examen cystoscopique, l'endoscope passe uniquement par la voie naturelle urinaire et il faut donc que les canaux d'acquisition et de projection soient tous les deux inclus d'une manière compacte dans la tige du système endoscopique. C'est le cas de l'un des

2. Endoscopie utilisée pour visualiser l'intérieur de la cavité abdominale.

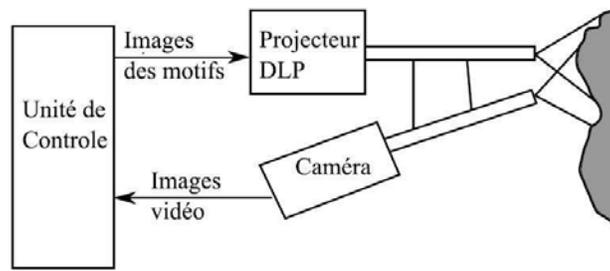


FIGURE. 2.12 – Illustration du mode de fonctionnement du système projection/reconstruction 3D développé pour la laparoscopie [Keller and Ackerman, 2000].

premiers systèmes endoscopiques à vision active, proposé dans [Schubert and Müller, 1998]. L'instrument développé est composé d'un endoscope muni d'un projecteur laser qui projette un cercle formant un cône. Ce système est destiné à la mesure de la taille des crevasses en bronchoscopie. Ce système fournit des informations anatomiques trop « ponctuelles » dans le cadre de la cartographie 3D d'organes.

Le système endoscopique développé à l'Université de Hong Kong [Chan *et al.*, 2003] est à la fois miniaturisé et compact. Comme l'illustre la figure 2.13, il est composé de trois canaux optiques regroupés dans une seule tige de 5 mm de diamètre. Ces canaux servent respectivement à éclairer la scène en lumière blanche, à projeter la lumière laser structurée et à acquérir des images. Dans le même souci de miniaturisation, les travaux

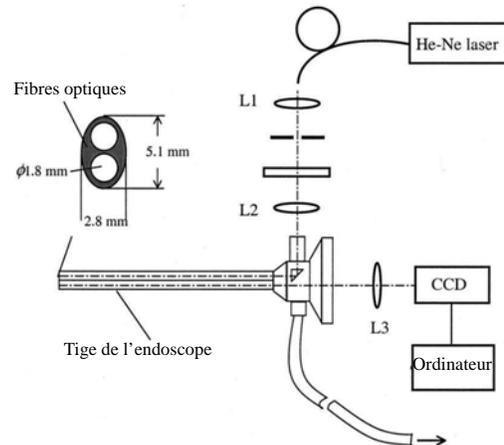


FIGURE. 2.13 – Configuration de l'endoscope 3D miniaturisé présenté dans [Chan *et al.*, 2003].

de [Hasegawa and Sato, 2001; Hasegawa *et al.*, 2002] utilisent un endoscope binoculaire conventionnel (avec deux canaux de sortie se terminant chacun par une caméra) ce qui permet d'éviter la construction d'un nouvel endoscope à trois canaux. L'idée est d'utiliser un des deux canaux d'acquisition d'images pour la projection sur la surface des organes de la lumière laser structurée (voir figure 2.14). L'avantage de cette configuration est

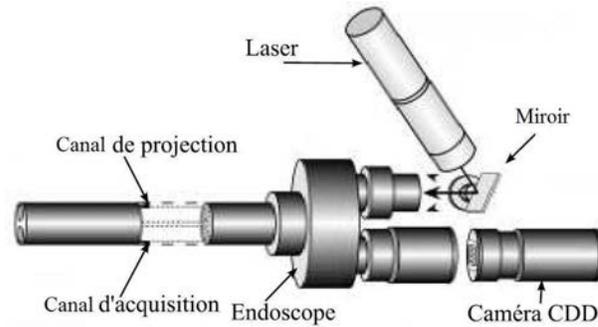


FIGURE. 2.14 – Configuration de l'endoscope 3D basé sur la projection de bandes laser. Un scanner comportant un miroir est utilisé pour effectuer le déphasage et acheminer les bandes laser dans l'un des deux canaux optiques de l'endoscope binoculaire [Hasegawa and Sato, 2001].

que le système endoscopique final a pu être miniaturisé sans avoir recours à un développement instrumental coûteux supplémentaire. Le motif projeté (généralisé par une source laser), composé d'un ensemble de bandes, subit un déphasage par acquisition. Les auteurs ont considéré 8 déphasages. Le système doit donc être bien synchronisé pour réaliser une acquisition par déphasage. D'après les auteurs, les erreurs de mesure de profondeur obtenues à l'aide de ce système sont de 1% pour un espace de travail allant de 20 à 40 mm.

2.2.3 Comparaison des méthodes de vision actives et passives

Les deux approches de vision (passive et active) sont largement utilisées pour la reconstruction d'informations 3D. Cela dit, le choix de l'une ou l'autre dépend des besoins de l'application (coût de l'instrumentation, résolution de la reconstruction, *etc.*) et de la nature des surfaces à reconstruire (avec ou sans textures et/ou couleurs, surfaces à fortes courbures ou non, existence d'occlusions ou non, *etc.*). Le tableau 2.2 (qui sera commenté ci-dessous) regroupe les caractéristiques (ou critères) principales des deux approches qui ont orienté nos choix instrumentaux pour la cartographie des parois internes de la vessie.

Des tests effectués dans le cadre de la thèse de [Hernandez-Mier, 2007] ont montré que les images cystoscopiques acquises en lumière blanche n'incluent pas d'indices visuels détectables d'une manière robuste dans une séquence entière. Même s'il est possible d'en extraire quelques-uns dans une image, ils ne figurent pas forcément dans les images qui suivent (voir figure 2.15). De plus, il existe une grande variabilité inter- et intra-patient en terme de texture des parois internes de la vessie. Compte tenu de ces observations, les méthodes de reconstruction 3D qui reposent sur la détection et la mise en correspondance d'indices visuels ne sont pas suffisamment robustes dans le cadre de notre application à la vessie. D'autre part, en endoscopie, les techniques basées sur l'ombrage souffrent

TABLEAU. 2.2 – Comparaison des caractéristiques des deux approches de vision passive et active pour l’endoscopie 3D.

Caractéristiques	Vision passive	Vision active
Surfaces sans informations riches (couleurs et/ou textures)	méthode non applicable	méthode applicable
Détection d’indices visuels	Dépend de la scène	Facile
Robustesse	Dépend de la scène	Dépend du motif projeté
Besoin d’une instrumentation particulière	Non	Oui
Résolution des données reconstruites	Haute résolution	Résolution basse

notamment du fait que les images sont acquises avec l’instrument qui est placé très près de la surface. Cette façon d’acquérir les données implique une très faible variabilité des intensités lumineuses dans les images rendant les techniques SFS peu adaptées en urologie.

Par contre, les techniques de reconstruction 3D basées sur la vision active présentent la meilleure robustesse, notamment vis-à-vis de la grande variabilité des textures que l’on peut trouver dans des images cystoscopiques. L’avantage de la vision active est donc lié au fait que les techniques de reconstruction 3D ne dépendent pas de la nature des surfaces imagées. Dans le cas de la vessie, les images acquises correspondent à des morceaux de surfaces à faibles rayons de courbure. Il n’est donc pas nécessaire de reconstruire de denses ensembles de données 3D des surfaces.

Les travaux instrumentaux présentés dans ce chapitre montrent qu’il est techniquement possible de concevoir des endoscopes 3D. Dans le cas de la vessie, il est nécessaire de reconstruire des surfaces 3D tout en préservant l’information de texture qui est à plaquer sur ces surfaces. Deux solutions instrumentales basées sur la vision active semblent autoriser un tel objectif. La première solution est de construire un nouvel endoscope avec un canal supplémentaire pour la projection d’un motif (approche validée dans [Chan *et al.*, 2003]). La deuxième possibilité consiste à utiliser un endoscope binoculaire conventionnel dont l’un des deux canaux d’acquisition est réservé à la projection d’un motif (approche validée dans [Hasegawa and Sato, 2001; Hasegawa *et al.*, 2002]).

En considérant globalement les remarques de ce paragraphe, la vision active semble être l’approche la plus adaptée pour construire des informations 3D en cystoscopie. Le système cystoscopique à développer dans cette étude sera donc basé sur la lumière structurée.

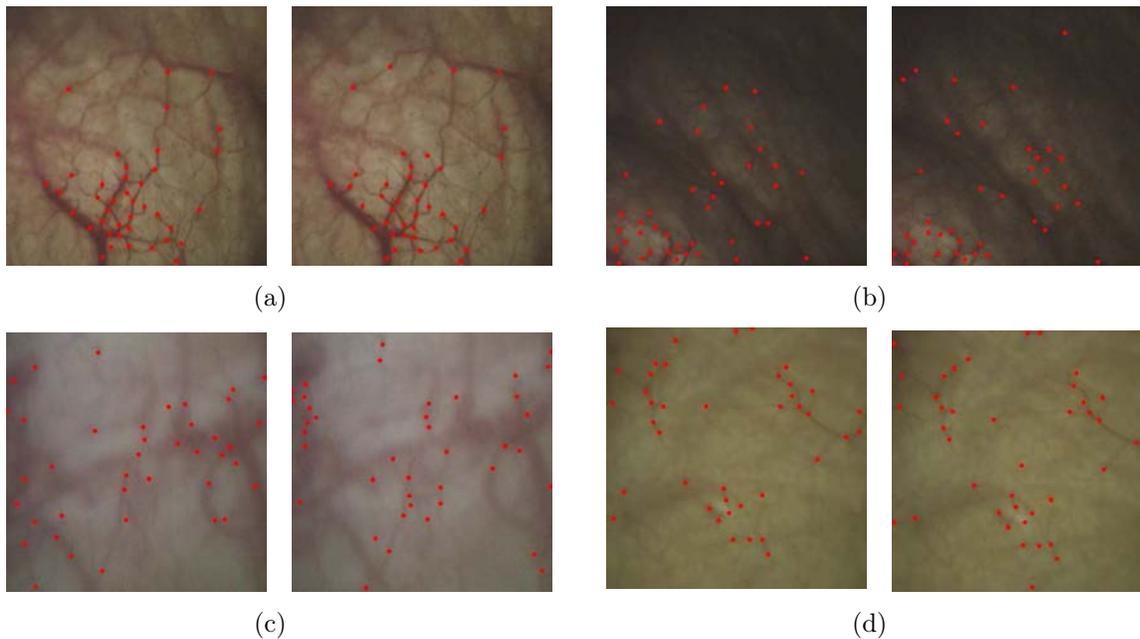


FIGURE. 2.15 – Exemples de points détectés par l’algorithme de Harris sur des paires d’images consécutives et issues de 4 séquences cystoscopiques différentes [Hernandez-Mier, 2007]. Les coins détectés sont marqués par des points rouges. (a) Images bien illuminées avec des structures bien marquées (contrastées). Les points détectés se concentrent sur une petite région de l’image. (b) Images avec des régions de fortes différences d’éclairage. Il n’y a pas de points détectés dans les régions obscures et les points détectés sur la première image ont des points homologues sur la deuxième image seulement dans la partie plus claire. (c) Images bien illuminées et avec des structures visibles mais légèrement floues. Les points détectés sur la première image ont des points homologues sur la deuxième seulement sur la partie gauche des images. (d) Images avec une bonne illumination mais avec peu de primitives. Les mêmes points sont détectés sur les deux images, mais ils se concentrent seulement sur les structures fortement contrastées.

2.3 Objectifs et verrous scientifiques de la thèse

Rappelons que la finalité clinique de cette étude est de fournir au clinicien une méthode intégrée et automatique pour construire une carte 3D avec couleurs/textures relatives aux surfaces internes de la vessie. Aussi, le principe du système cystoscopique 3D pour lequel nous avons opté est basé sur la lumière structurée. Nos développements techniques concernent donc les problèmes métrologiques et techniques liés à l’instrumentation et les problèmes de traitement d’images liés à la reconstruction d’informations 3D et au recalage mosaïquage 3D.

En effet, même si les travaux présentés précédemment [Hasegawa *et al.*, 2002; Chan *et al.*, 2003] valident la possibilité de reconstruire des informations 3D en projetant un motif lumineux à travers la tige d’un endoscope, le « verrou instrumental » n’est pas pour autant levé. En effet, le type de lumière à projeter, la forme du motif et les composants optiques du

projecteur notamment doivent être choisis en fonction de la scène (*i.e.*, de l'application).

Du point de vue traitement d'images, le vrai défi scientifique majeur est de pouvoir reconstruire en 3D la surface de l'organe alors qu'on ne connaît rien du mouvement du cystoscope. Il s'agit d'arriver à placer toutes les images locales acquises par le cystoscope dans un référentiel commun pour obtenir une surface. Pour cela, il s'agira de positionner les informations 3D reconstruites pour l'ensemble des acquisitions d'une séquence dans un référentiel commun. Cet objectif fait appel à des méthodes de recalage de données tridimensionnelles.

Enfin, les solutions à développer doivent permettre l'acquisition simultanée d'informations 3D et de textures relatives aux surfaces imagées. Les solutions proposées dans la littérature permettent la reconstruction d'une information 3D au détriment des couleurs et textures. Dans [Chan *et al.*, 2003], le motif laser projeté est formé d'une matrice de points, avec une grande densité spatiale, les couleurs et textures de la scène se trouvant ainsi complètement occultées. Le défi à relever ici est donc de choisir un motif permettant à la fois une reconstruction de surfaces 3D et une conservation maximale des informations de texture et de couleur. Pour conserver un maximum d'informations dans les images et respecter les habitudes des urologues, les points projetés ne devront pas apparaître dans les cartes reconstruites.

2.4 Conclusion

Dans ce chapitre, nous avons présenté en premier lieu les techniques de recalage et de mosaïquage 2D appliquées aux images cystoscopiques pour la construction d'images panoramiques. Ces méthodes, qui sont à ma connaissance les seules qui ont été publiées pour la vessie, facilitent effectivement le diagnostic et la détection de lésions dans la paroi interne. Une étude comparative de deux de ces méthodes a été effectuée pour cerner leurs limites et faire un bilan sur l'apport du mosaïquage d'images en cystoscopie. Cependant, une représentation tridimensionnelle des parois internes de la vessie s'avère nécessaire pour faciliter davantage le diagnostic clinique et obtenir de nouveaux critères de discrimination tissulaire.

La deuxième partie de ce chapitre a été consacrée à l'endoscopie 3D. Nous avons présenté un aperçu des travaux de reconstruction 3D en endoscopie. Ces travaux reposent sur deux types de vision : passive et active. Nous avons montré que la vision active est la mieux adaptée pour la reconstruction 3D en cystoscopie. Enfin, nous avons présenté les objectifs scientifiques de notre thèse ainsi que les différents verrous à lever. Le chapitre 3 détaille d'abord l'aspect instrumental développé. Ensuite, nous présentons une nouvelle

méthode de calibrage du système cystoscopique. C'est une méthode flexible facile à utiliser pas uniquement dans le milieu médical, mais aussi pour tout autre système à base de vision active similaire à celui développé dans cette thèse.

Contributions

- Établir un état de l'art sur les méthodes de cartographie 2D et 3D applicables aux surfaces internes de la vessie.
- Établir une étude comparative de deux méthodes de mosaïquage des surfaces internes de la vessie.

Publications

- [Ben-Hamadou *et al.*, 2009b], *XXIIe colloque GRETSI (traitement du signal et des images) (GRETSI)*.
- [Ben-Hamadou *et al.*, 2009a], *SPIE European Conferences on Biomedical Optics (ECBO)*.

3

Prototype expérimental : calibrage et reconstruction de points 3D

Dans ce chapitre, nous présentons notre prototype expérimental de vision en lumière structurée. Ce prototype est un système non miniaturisé de vision active que nous avons construit afin de montrer la faisabilité de la reconstruction surfacique à partir de données cystoscopiques enrichies (images cystoscopiques enrichies où figurent, en plus des couleurs liées à la texture de l'organe, des motifs géométriques projetés par un projecteur laser sur la surface de l'organe). Le système cystoscopique réel que nous envisageons sera donc une version miniaturisée (intégrée sur un endoscope) de ce prototype expérimental. Le prototype développé est composé d'une caméra et d'un projecteur d'un petit nombre d'objets ponctuels situés sur une mire. La caméra et le projecteur sont liés physiquement de manière rigide. Pour envisager la reconstruction d'informations 3D à partir d'une séquence d'images cystoscopiques, le système de vision doit être entièrement calibré. Cela concerne à la fois la caméra et le projecteur. Nous cherchons à rendre ce processus le plus flexible possible, c'est-à-dire, à ne pas utiliser d'équipements additionnels de positionnement qui s'avèrent chers et qui compliquent une utilisation en condition clinique. Si pour le calibrage de caméras plusieurs techniques classiques sont utilisables, il n'en va pas de même du calibrage de projecteurs de lumière structurée. Nous présenterons une méthode originale et simple qui vise à reconstruire la position 3D des projections des centres des points de la mire sur la surface de l'organe dans le système de coordonnées de la caméra. Par abus de langage, nous désignerons par « points laser 3D » ces projections. Nous évaluerons finalement la précision de la méthode en testant la reconstruction 3D d'une pièce usinée (étalon) avec notre prototype expérimental.

3.1 Prototype expérimental

Principe de l’endoscopie en lumière structurée. Les travaux de [Chan *et al.*, 2003] démontrent la possibilité d’intégrer la projection d’un motif lumineux dans un système endoscopique miniaturisé (le diamètre de la tige de l’endoscope est de 5 mm). Actuellement, ce type d’instrument n’est pas encore commercialisé pour l’exploration de la vessie. Nous avons développé un prototype expérimental en nous basant sur l’idée de [Chan *et al.*, 2003], à savoir la reconstruction de points 3D laser à partir de leur projection dans des images endoscopiques. Notre objectif *in fine* est la reconstruction 3D de surfaces étendues de l’intérieur de la vessie à partir de ces points. Il est important de noter que le système endoscopique présenté dans [Chan *et al.*, 2003] ne fournit pas simultanément des informations 3D et des textures relatives aux surfaces imagées. En effet, le motif projeté est une grille dense de points dont la projection occulte les textures de la surface imagée. Par conséquent, ce dispositif permet seulement d’acquérir des informations 3D relatives aux coordonnées des points lasers. Dans ce travail, nous optons pour un motif moins dense (très faible nombre d’objets ponctuels) pour permettre à la fois la reconstruction d’informations 3D (points situés sur la surface interne de la vessie) et l’examen visuel de la texture des surfaces par les urologues. Ainsi, les images acquises seront très peu dégradées par la présence du motif projeté. Les cliniciens du centre Alexis Vautrin de Nancy³ avec lesquels nous collaborons confirment que la projection d’un motif constitué de quelques points situés à la périphérie du champ de vision est acceptable, car le clinicien se focalise essentiellement sur le centre des images⁴.

Choix du motif. Nous avons choisi un motif composé d’un ensemble de huit points laser⁵ situés sur un cercle (ce qui donne lieu à 8 rayons laser dans l’espace : voir figure 3.1). La position des points projetés dans les images acquises par la caméra dépend de la profondeur de la surface imagée (distance entre la caméra et la surface). Comme le volume de la vessie est limité, les positions des spots laser changent, mais restent tout de même sur les bords du champ de vue des images cystoscopiques. Le choix de la nature et de la longueur d’onde de la lumière à projeter est un facteur important. Il est guidé par le besoin de localiser aisément les points laser dans les images cystoscopiques acquises par une procédure de segmentation de points dans une image. Par exemple, une lumière laser rouge complique cette localisation du motif car il y a très peu de contraste entre le laser

3. Une urologue (Marie-Ange D’Hallewin) et un chirurgien (François Guillemin)

4. En pratique, il sera possible de projeter les points uniquement sur une image sur deux. En effet, afficher seulement les images sans points spots laser conduit à un flux vidéo exploitable par le clinicien

5. Ainsi détaillé dans la thèse, il est aussi possible de projeter des motifs avec des formes variables et avec un nombre différent de points.

rouge et les textures rougeâtres de la paroi interne de la vessie. Par contre, le laser vert fournit un contraste suffisamment important pour simplifier la segmentation des spots laser dans les images. Le processus de segmentation sera abordé dans la section 3.3.2 du chapitre. En résumé, le prototype expérimental développé et utilisé dans ce travail est

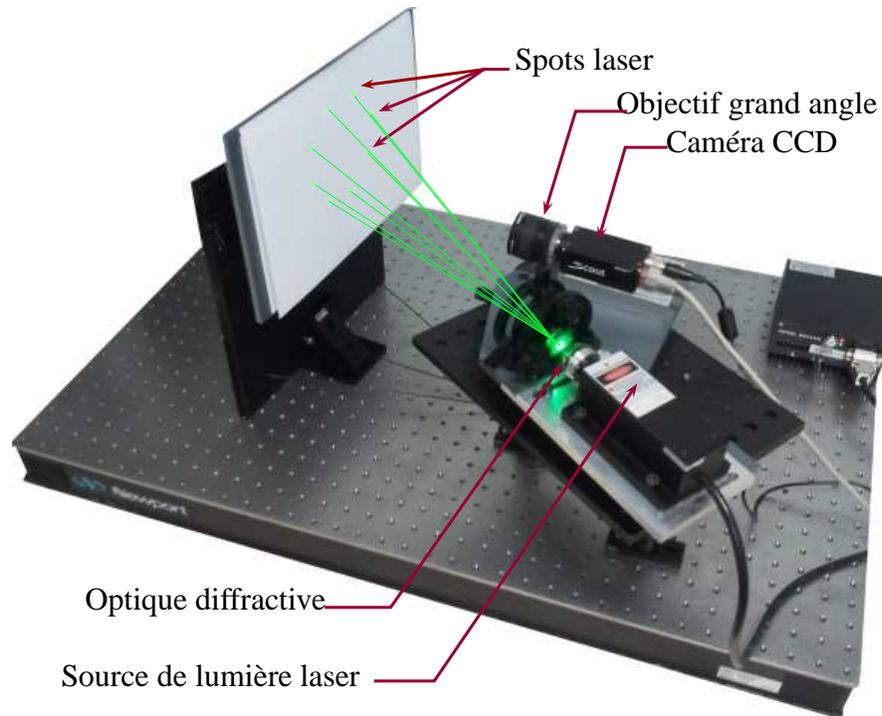


FIGURE. 3.1 – Configuration du prototype expérimental. La mire contenant le motif géométrique (8 points situés sur un cercle) est située à l'intérieur du boîtier correspondant au projecteur de lumière structurée. La caméra CCD et le projecteur (source laser et mire diffractive) sont liés de manière rigide. Les faisceaux issus du projecteur (spots lasers verts) sont ajoutés dans cette figure pour schématiser la projection linéique des huit points du motif sur la surface, mais en réalité, elles ne sont pas visibles à l'œil nu. L'optique de la caméra a une distance focale de 6 mm.

composé des éléments suivants (voir figure 3.1) :

- une caméra de type Basler équipée d'un objectif grand angle pour reproduire la qualité des images cystoscopiques (qualité en terme de distorsions par exemple). Les images sont acquises à une fréquence de 25 acquisitions par seconde et ont une taille de 780×580 pixels.
- une source générant un faisceau collimaté de lumière laser.
- une optique diffractive pour générer le motif lumineux composé de huit points placés sur un cercle. Le faisceau de lumière laser passe à travers l'optique pour se diffracter en huit faisceaux formant un cône.

Dans nos expériences, l'éclairage (source de lumière blanche d'un cystoscope) est remplacé par la lumière de la pièce dans laquelle les expérimentations sont faites.

3.2 Calibrage des systèmes de vision en lumière structurée

Le calibrage (« *calibration* » en anglais) d'un système de vision en lumière structurée est une étape essentielle à la reconstruction d'informations 3D relatives à la scène imagée. Cette étape concerne à la fois la caméra et le projecteur de lumière structurée. Généralement, on commence par le calibrage de la caméra, indépendamment du projecteur [Chen and Li, 2003; Chen *et al.*, 2007; GAO *et al.*, 2008]. Lorsque la caméra a été calibrée, le projecteur est calibré à son tour. Ce traitement en deux temps est lié au fait que contrairement à la caméra, le projecteur n'est pas un capteur, ce qui implique la nécessité d'utiliser la caméra du système de vision pour faire le calibrage. Dans la suite de la section, nous passons brièvement en revue les approches les plus connues pour le calibrage des caméras puis le calibrage des projecteurs de lumière structurée, et nous établissons des choix de méthodes adaptées à notre problème.

3.2.1 Calibrage de la caméra

Le calibrage de la caméra consiste à déterminer les paramètres intrinsèques et extrinsèques du modèle « *pinhole* » décrivant le processus de formation d'images. Les paramètres de ce modèle ont été détaillés dans la section 2.1.1 du chapitre 2. Cependant, par soucis de clarté, nous rappelons les transformations géométriques et les notations retenues pour le modèle caméra.

Paramètres extrinsèques. Ils regroupent les paramètres de la transformation $T_{w_k,c}$ qui définissent la pose 3D de la caméra dans la scène (c'est-à-dire, relation géométrique entre le repère caméra $\{c\}$ prenant le centre optique comme origine et le repère monde $\{w_k\}$ défini par une pièce de calibrage) pour une acquisition k donnée :

$$T_{w_k,c} = \begin{bmatrix} \mathbf{R}_k & \vec{\mathbf{t}}_k \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

où \mathbf{R}_k est une matrice de rotation 3D et $\vec{\mathbf{t}}_k$ est un vecteur de translation 3D.

Paramètres intrinsèques. Ils regroupent les paramètres optiques propres à la caméra. Ils modélisent la projection d'un point de la scène dans le plan image *via* la matrice caméra

\mathcal{K} et les coefficients k_1 et k_2 définis comme suit.

- La matrice caméra \mathcal{K} décrit la projection d'un point donné de la scène exprimé dans le repère caméra sur la matrice CCD sans tenir compte de la distorsion radiale.

$$\mathcal{K} = \begin{bmatrix} \frac{f}{S_x} & 0 & u_0 \\ 0 & \frac{f}{S_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

où f est la distance focale de la caméra, S_x et S_y sont les dimensions des capteurs CCD de la caméra et (u_0, v_0) sont les coordonnées de la projection du centre optique dans l'image.

- k_1 et k_2 sont les coefficients d'un polynôme qui modélise la distorsion radiale due aux optiques de la caméra (voir section 2.1.1 du chapitre 2).

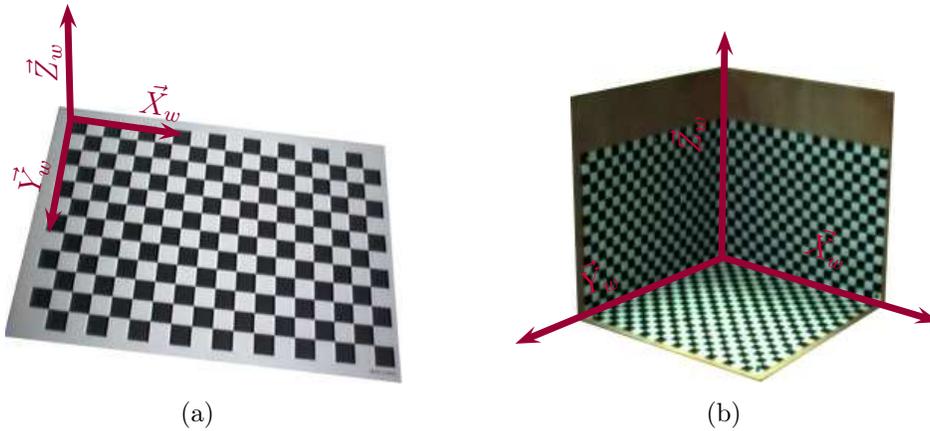


FIGURE. 3.2 – Deux exemples de pièces de calibrage. (a) Le damier est une mire de calibrage souvent utilisé avec la méthode de [Zhang, 2000]. (b) L'objet de calibrage formé de trois plans orthogonaux est une pièce souvent utilisée pour le calibrage de caméra [Guerchouche and Coldefy, 2008]. Il offre une plus grande précision par rapport à l'objet planaire au prix d'une plus grande complexité de la méthode de calibrage.

Méthodes d'estimation des paramètres du modèle caméra. Classiquement, l'estimation de l'ensemble de ces paramètres nécessite l'utilisation d'une mire de calibrage dont la géométrie est connue avec précision. La plus simple des mires de calibrage est formée d'un plan contenant un ensemble de points de référence définis par exemple par les coins d'un damier ou par une grille de centres de disques (voir l'exemple de la figure 3.2.(a) pour le damier). Les méthodes qui utilisent ce type de mire de calibrage nécessitent au moins deux acquisitions de la même scène sous différents points de vue. Elles se basent sur le fait que toutes les positions des points de référence 3D sont connues dans un

repère défini par la mire de calibrage et appartiennent au même plan. D'autres méthodes de calibrage cherchent à déterminer la relation entre différents plans d'une scène. Dans ce cas de figure, un dièdre ou un cube est généralement utilisé (voir l'exemple de la figure 3.2.(b)). Le calibrage de caméras revient à estimer les meilleurs paramètres (intrinsèques et extrinsèques) qui relient les points 3D définis dans le repère monde à leurs correspondants extraits des images. Les principales méthodes d'estimation de ces paramètres sont détaillées dans la suite du paragraphe.

Les premières méthodes de calibrage de caméra (par exemple [Hall *et al.*, 1982; Faugeras and Toscani, 1986]) ne prennent pas en compte la modélisation de la distorsion radiale dans le modèle caméra. Ce dernier devient donc linéaire et les paramètres du modèle se résument aux éléments des matrices $\mathbf{T}_{w_k,c}$ et \mathcal{K} . Ces méthodes sont les moins précises (sauf pour les capteurs à grandes distances focales, donc sans distorsion).

Dans le cas de faibles distances focales ($f \leq 12 \text{ mm}$), les techniques les plus précises sont celles qui intègrent la distorsion radiale dans le modèle caméra. Ce modèle devient alors non linéaire. L'estimation des paramètres du modèle est faite moyennant des techniques d'optimisation non linéaire qui minimisent l'erreur de projection dans les images des points 3D de la mire de calibrage. En d'autres termes, les points 3D de la mire, projetés dans les images avec les paramètres de calibrage estimés par l'optimisation, doivent être le plus proche possible en moyenne des points correspondants vus (et segmentés) dans les images.

Dans le processus d'optimisation, les valeurs initiales des paramètres du modèle doivent être suffisamment proches de la solution pour assurer leur convergence. Une solution à ce problème consiste à effectuer une première estimation des paramètres du modèle caméra en excluant les coefficients des distorsions radiales, et en utilisant une technique d'estimation linéaire. Dans un deuxième temps, une estimation non linéaire est effectuée pour déterminer tous les paramètres du modèle en prenant le résultat de la première étape d'estimation comme valeur initiale. Les valeurs initiales des coefficients de distorsions sont en général nulles. Parmi les méthodes basées sur une estimation non linéaire, nous pouvons citer les travaux de [Tsai, 1987], [Salvi *et al.*, 1998] (version modifiée de [Faugeras and Toscani, 1986]), [Heikkila and Silven, 1997] et [Zhang, 2000].

Choix effectué. Nous avons utilisé la méthode de Zhang [Zhang, 2000] pour calibrer la caméra de notre prototype expérimental⁶. Cette méthode figure parmi les méthodes les plus utilisées dans la littérature. L'avantage de cette méthode réside dans sa flexibilité.

6. Une implémentation de cette technique de calibrage est fournie dans la librairie libre de vision par ordinateur OpenCV développée par Intel Corporation : <http://opencv.willowgarage.com/wiki/>

En effet, le seul équipement nécessaire pour effectuer le calibrage est un plan avec des points de référence. La caméra acquiert des images de ce plan sous différentes positions et orientations inconnues. Notre plateau de calibrage est formé d'un plateau en PVC « *polyvinyl chloride* » sur lequel nous avons collé une feuille contenant une grille de disques dont la dimension et les positions sont connues dans le repère monde $\{w\}$. La figure 3.3 montre un exemple d'image acquise pour notre plateau de calibrage. La segmentation des disques dans les images est effectuée par la technique de [Ouellet and Hébert, 2008] qui détecte et calcule l'équation des ellipses dans une image. Nous avons choisi une mire avec des disques plutôt qu'une mire avec un damier car le calcul des centres des ellipses est plus précis que la détection de coins.

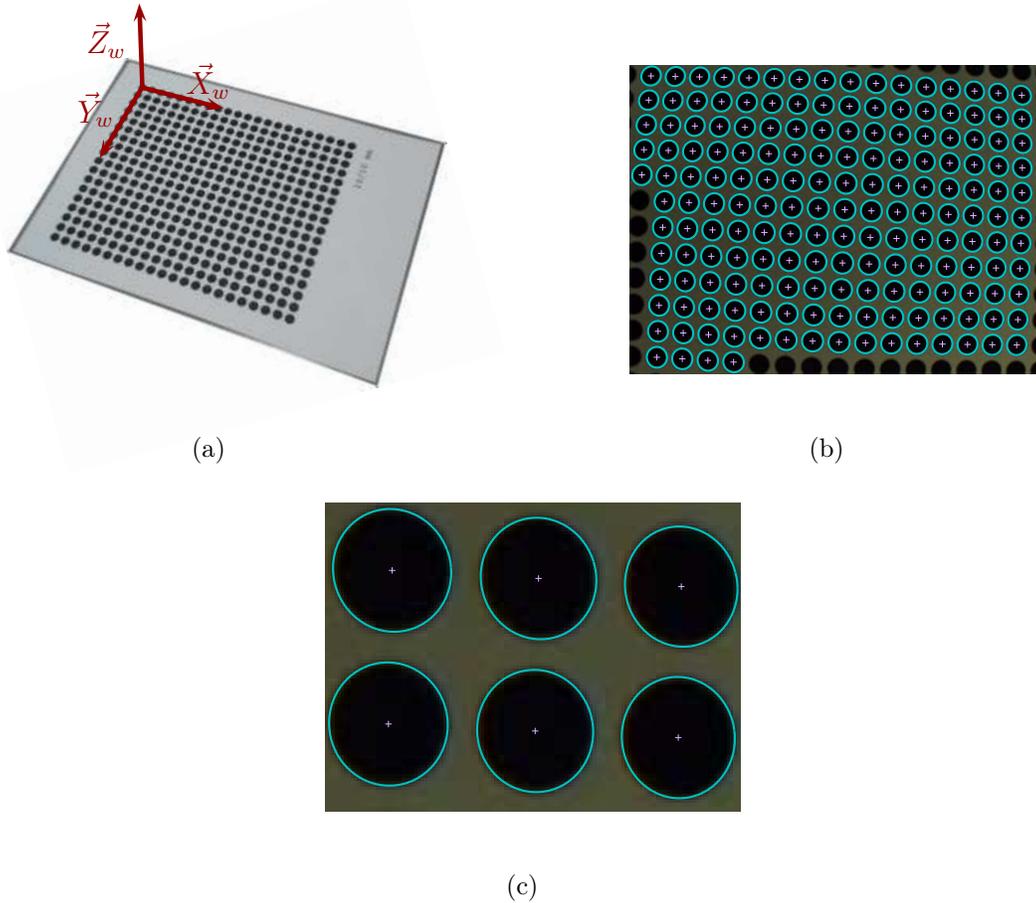


FIGURE. 3.3 – (a) Mire utilisée pour le calibrage de la caméra. (b) Exemple d'image acquise pour la mire de calibrage dont les ellipses ont été segmentées. La segmentation a été effectuée en utilisant la technique présentée dans [Ouellet and Hébert, 2008]. (c) Zoom sur l'image qui met en lumière la précision de la détection des ellipses.

Nous avons utilisé un ensemble de 20 images (poses) du plateau de calibrage. Dans le cadre de cette thèse, nous avons besoin uniquement des paramètres intrinsèques (la

matrice \mathcal{K}) et des coefficients de distorsion radiale (k_1 et k_2). En effet, les points 3D doivent être reconstruits dans le repère de la caméra et non dans un repère monde (non disponible dans une vessie). Un exemple de valeurs obtenues en calibrant notre caméra est donné ci-après.

$$\begin{aligned}\mathcal{K} &= \begin{bmatrix} 752.39 & 0 & 405.18 \\ 0 & 755.21 & 274.28 \\ 0 & 0 & 1 \end{bmatrix} \\ k_1 &= -0.21 \\ k_2 &= 0.17\end{aligned}$$

Nous avons vérifié qu'en modélisant la distorsion radiale par un polynôme d'ordre quatre, les coefficients k_3 et k_4 des monômes de degrés 3 et 4 sont quasi nuls. Les deux coefficients k_1 et k_2 permettent de prendre en compte précisément la distorsion [Miranda-Luna *et al.*, 2004]. Les paramètres du modèle caméra trouvés sont évalués en comparant les positions de la projection des points 3D du plateau de calibrage dans les images (positions calculées) aux positions 2D de ces mêmes points vus et segmentés dans les images (positions de référence). L'erreur moyenne est de l'ordre de 0.2 pixel. Cette erreur est comparable à celles qu'on trouve dans la littérature.

3.2.2 Calibrage des projecteurs de lumière structurée

Le calibrage du projecteur de lumière structurée consiste à modéliser mathématiquement la projection d'un motif lumineux dans l'espace. Pour calibrer le projecteur, on utilise les images acquises par la caméra du système. Il existe essentiellement deux approches pour modéliser le processus de projection.

La première approche repose sur le modèle sténopé. On remarque que la caméra projette un ensemble de points 3D (situés sur la scène imagée) en un ensemble de points 2D dans son plan focal. En revanche, le principe du projecteur est de projeter un ensemble de points 2D (situés dans le plan de la mire du projecteur) sur la scène. On peut donc considérer le projecteur comme un système projectif de type « caméra inversée » [Mosnier *et al.*, 2009; Draréni *et al.*, 2009].

La deuxième approche consiste à modéliser mathématiquement la trajectoire de chaque élément du motif dans l'espace. Dans notre cas, le motif projeté contient un ensemble de points, leur projection est donc décrite par un ensemble de droites 3D (droites à exprimer dans le repère de la caméra) [Huynh *et al.*, 1999; Zhou and Zhang, 2005; Yamauchi *et al.*, 2008]. Dans les prochains paragraphes, nous nous focalisons sur l'estimation des coordonnées 3D

des éléments du motif projetés sur la mire de calibrage du projecteur. L'avantage de ce dernier modèle de projection réside dans son adaptabilité à différents types de projecteurs. En effet, ce modèle est plus général que le modèle caméra inversée. Ce dernier repose sur l'hypothèse que la projection est centrale (c'est-à-dire, possède un centre optique). Cette hypothèse n'est pas vérifiée pour tous les projecteurs [Yamauchi *et al.*, 2008].

Quel que soit le modèle de projection du motif lumineux, il est indispensable de retrouver les coordonnées 3D du motif projeté sur la mire de calibrage du projecteur. Dans les deux prochains paragraphes, nous présentons les approches utilisées pour résoudre ce problème.

Cas où le motif est composé de droites. Ce cas est très couramment rencontré dans la littérature, c'est pourquoi nous lui consacrons un paragraphe. La projection d'un motif composé de droites donne lieu à un ensemble de plans dans l'espace, le calibrage consiste à déterminer l'équation de plans formés par la projection des droites dans l'espace. Les équations des plans sont souvent déterminées à l'aide d'une propriété géométrique appelée *invariance du birapport* (« *cross ratio* » en anglais) en géométrie projective [Jean-Denis, 2009]. De nombreux auteurs [Chen and Kak, 1987; Huynh *et al.*, 1999; Zhou and Zhang, 2005] ont exploité cette propriété pour calibrer ce type de projecteur.

Choisir un motif composé de segments de droites implique, après projection, un ensemble de lignes situées sur la surface à reconstruire, donc un motif projeté très informatif sur la courbure de la surface. Cependant, l'inconvénient de ce choix de motif est que les lignes projetées occultent des zones importantes dans l'image, ce qui perturbe le diagnostic médical. Pour l'endoscopie de la vessie, il s'agit de préserver autant que possible les informations de couleur et de texture. C'est pour cette raison que nous optons pour un motif composé d'un ensemble des points.

Cas où le motif est composé de points. Dans la littérature il existe très peu de travaux qui concernent le calibrage de projection de motifs composés de points. Le système de vision en lumière structurée proposé dans [Marzani *et al.*, 2002] est formé d'une caméra et d'un projecteur de lumière structurée de type laser. Le motif projeté est formé d'une grille de 361 (19×19) points. L'angle de projection de chaque rayon laser par rapport à l'axe principal de la projection (c'est-à-dire, la direction orthogonale à la mire contenant le motif à projeter) est fixe et connu. La méthode proposée par [Marzani *et al.*, 2002] pour calibrer ce système nécessite un plateau de calibrage parfaitement orthogonal à l'axe principal de projection du projecteur de points laser. Pour cela, *un plan motorisé permettant la translation du plateau de calibrage selon l'axe principal de projection entre*

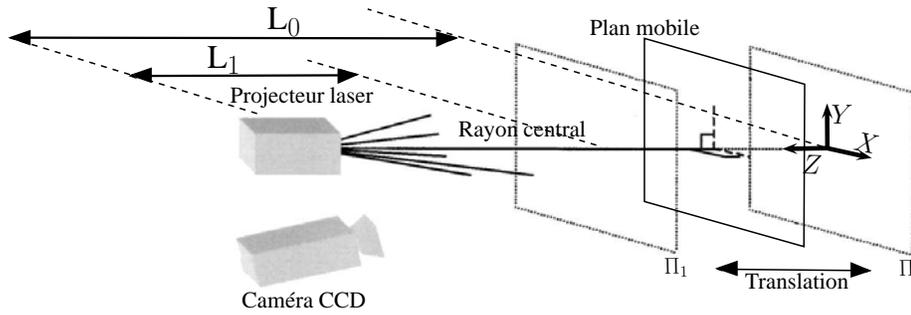


FIGURE. 3.4 – Configuration du système de vision active proposé dans [Marzani et al., 2002]

deux positions limites Π_I et Π_0 (voir figure 3.4) est requis. Les limites Π_I et Π_0 définissent l'espace de travail. Elles doivent être choisies avec précaution car l'espace de travail doit couvrir la profondeur de l'objet à reconstruire. L'intersection entre un rayon laser avec le plan dans les positions Π_I et Π_0 donne deux points laser dans l'image. Ces deux points sont utilisés pour calculer le segment de droite épipolaire correspondant au rayon en question. Les positions sur l'axe Z (voir figure 3.4) correspondant aux extrémités du segment épipolaire sont définies par les profondeurs de Π_I et Π_0 . Après le calibrage du système, le calcul de la profondeur de chaque point laser détecté dans l'image est effectué à l'aide de la position du point sur le segment de droite épipolaire le plus proche. Sans distorsion radiale dans les images, il existe une relation linéaire connue entre la position d'un point sur la ligne épipolaire et la profondeur du point sur la droite 3D correspondante. L'avantage de cette modélisation est qu'elle ne nécessite pas le calibrage de la caméra. Cependant, son inconvénient majeur est l'utilisation d'équipements de positionnement précis pour le placement du plateau de calibrage parfaitement perpendiculaire à l'axe de projection. Par ailleurs, la méthode dépend de la précision du matériel de translation utilisé. La reconstruction de points 3D de plusieurs plans situés dans l'espace de travail a conduit à une erreur moyenne de 0.115 mm avec un écart moyen de 0.029 mm pour un espace de travail allant d'une profondeur de 10 à 90 mm.

[Chan et al., 2003] utilisent des équipements similaires à ceux de [Marzani et al., 2002] pour le calibrage d'un système endoscopique qui intègre un canal optique additionnel pour la projection d'une grille de points laser. La projection est décrite par un modèle sténopé. Un plateau de calibrage est positionné orthogonalement à l'axe principal de projection de l'endoscope. Ensuite, le plateau de calibrage est déplacé avec précision à différentes distances par rapport à la tête de l'endoscope. L'angle de projection de chaque point de la grille étant connu exactement par rapport à l'axe principal de l'endoscope, il est possible de calculer les coordonnées des points laser sur le plateau de calibrage de la même manière

que dans [Marzani *et al.*, 2002]. Les positions 3D des points laser sont ensuite utilisées avec leurs correspondants dans *l'image projetée* (c'est-à-dire, image du motif construite numériquement) pour estimer les paramètres du modèle sténopé de projection. Notons que quelle que soit la profondeur du plateau de calibrage, l'image projetée ne change pas. L'erreur de la reconstruction 3D résultante est de 2.4% de la profondeur réelle des pièces testées.

3.2.3 Bilan sur les méthodes de calibrage de projecteurs de lumière structurée

Parmi les méthodes de calibrage de projecteurs présentées, nous retenons uniquement celles qui adressent le calibrage de projecteurs d'une lumière structurée formée de points. Les méthodes de [Marzani *et al.*, 2002; Chan *et al.*, 2003] nécessitent l'utilisation d'équipements de positionnement très précis. En plus de leur coût élevé, la performance de ces méthodes dépend fortement de la précision des équipements employés. En vision industrielle (par exemple pour faire des mesures dimensionnelles de pièces manufacturées), on accepte ce type de méthodes, car il est possible de disposer d'équipements précis pour garantir une bonne performance du calibrage. Par contre, leur utilisation dans un contexte médical est proscrite de par leur manque de flexibilité. Cela motive le développement d'une nouvelle méthode de calibrage adaptée à l'examen endoscopique des organes creux.

3.3 Méthode proposée pour le calibrage du projecteur de points laser

3.3.1 Principe général de la méthode

Dans un souci de flexibilité⁷, nous avons développé une nouvelle méthode de calibrage de projecteurs de lumière structurée en points qui ne nécessite aucun équipement de positionnement. Notre méthode requiert simplement un plateau de calibrage planaire composé d'une plaque en PVC et une feuille blanche sur laquelle est imprimé un ensemble d'au moins quatre disques non alignés. Sur la figure 3.5, les quatre disques sont situés dans les coins du plateau de calibrage et le plateau est déplacé en deux positions distinctes (plans \mathcal{P}_1 et \mathcal{P}_2 correspondants à $k = 1$ et $k = 2$).

7. L'idéal serait de tenir manuellement le plateau dans le champ de vue de la caméra au moment du calibrage sans pour autant connaître la position du plateau.

Notre méthode de calibrage nécessite l'acquisition d'au moins deux images I_k de la mire de calibrage placées en différents points de vue ($k \in \{1, \dots, K\}$ et $K > 1$). La figure 3.6 montre un exemple de paire d'images de calibrage.

Contrairement aux méthodes de [Marzani *et al.*, 2002] et [Chan *et al.*, 2003], le positionnement du plateau peut être fait manuellement sans avoir recours à un équipement de positionnement particulier. Il suffit de déplacer approximativement le plateau selon l'axe \vec{z}_c de la caméra.

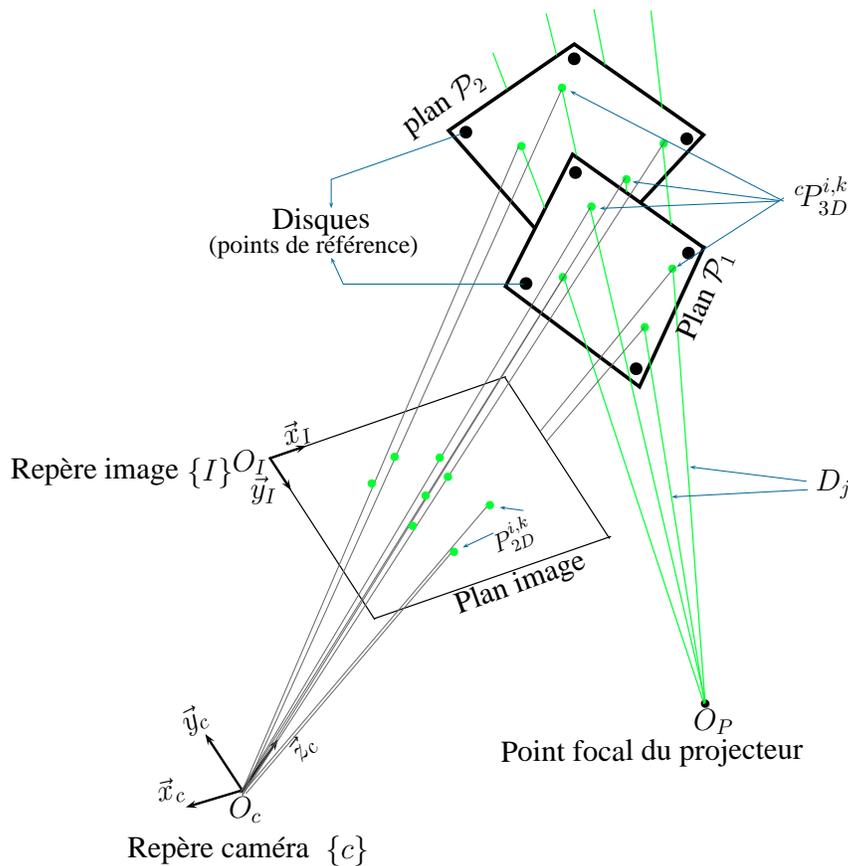


FIGURE. 3.5 – Principe du calibrage du projecteur de points laser. Le plateau de calibrage comprend une feuille blanche et quatre disques noirs situés dans les coins du plateau. Les centres des quatre disques sont utilisés pour calculer les équations des plans \mathcal{P}_k .

Dans la figure 3.5, le faisceau laser principal se diffracte pour générer N rayons (droites 3D) lasers D_j avec $j \in \{1, \dots, N\}$ ($N = 8$ rayons pour notre prototype expérimental présenté sur la figure 3.1, en supposant que chaque objet projeté est ponctuel). Ces rayons passent tous par le centre optique O_P du projecteur. L'intersection entre les droites D_j et un plan \mathcal{P}_k de calibrage (dans une position k donnée) donne un ensemble de points 3D notés $cP_{3D}^{i,k}$. $cP_{3D}^{i,k}$ est donc le $i^{\text{ème}}$ point laser 3D sur le $k^{\text{ème}}$ plan dont les coordonnées sont exprimées dans le repère caméra. Les points image qui correspondent à ces points 3D sont

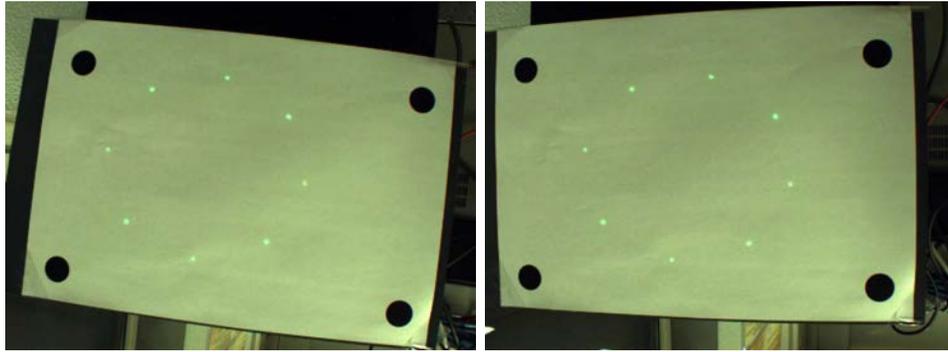


FIGURE. 3.6 – Exemple d’images acquises pour la mire de calibrage du projecteur dans deux différentes positions.

désignés par $P_{2D}^{i,k}$. Physiquement, on peut résumer l’observation d’un point du motif par la caméra en deux projections successives :

- le point du motif est projeté sur le plan de la mire de calibrage.
- le point laser obtenu est à son tour projeté sur le plan focal de la caméra lors de l’acquisition des images cystoscopiques.

En pratique, les paramètres du projecteur (coordonnées du point O_p et les équations des N rayons laser D_j , tous exprimés dans le repère caméra) doivent être retrouvés à partir des images I_k en se basant sur les paramètres de la caméra préalablement calibrée. Notre méthode de calibrage de projecteur est composée des quatre étapes suivantes.

1. Segmentation des spots laser et des quatre disques dans les images. Cette étape conduit au calcul des points $P_{2D}^{i,k}$ correspondant aux centroïdes des spots laser et au calcul des centres des disques projetés.
2. Calcul de l’équation des plans \mathcal{P}_k dans le repère de la caméra, puis estimation des coordonnées des points laser 3D ${}^cP_{3D}^{i,k}$ ($\forall i \in \{1, \dots, N\}$ et $k \in \{1, \dots, K\}$) à partir des points $P_{2D}^{i,k}$.
3. Mise en correspondance, pour tout i et k , des points 3D ${}^cP_{3D}^{i,k}$ et des rayons laser D_j . À l’issue de cette étape, l’algorithme a déterminé l’ensemble des points ${}^cP_{3D}^{i,k}$ associés à chaque rayon D_j .
4. Détermination des paramètres des N rayons laser D_j sous la contrainte qu’ils s’intersectent (en O_p , valeur définitive et très précise des coordonnées du centre optique du projecteur). Les équations de ces rayons sont toujours exprimées dans le repère de la caméra.

3.3.2 Étapes de la méthode de calibrage proposée

Étape 1 : Segmentation d'image. Avant tout traitement, les images I_k sont traitées pour corriger la distorsion radiale due à l'optique de la caméra (voir la section 2.1.1 du chapitre 2).

On décrit successivement les choix effectués pour l'algorithme de segmentation des points laser et l'algorithme de segmentation des disques du plateau de calibrage.

La segmentation des spots laser est basée sur leur couleur et sur le fait que le fond de l'image est blanc, ou plus précisément gris clair dans les images acquises par la caméra. La distinction des pixels appartenant aux spots laser des pixels du fond de l'image repose sur la différence nette entre la teinte des spots laser et celle des pixels blancs.

Les images sont transformées de l'espace colorimétrique RVB vers l'espace TSV défini par trois composantes : teinte, saturation et valeur (« *hue, saturation, value* » en anglais). La forme de l'espace couleur TSV est illustrée dans la figure 3.7. La composante T (teinte) indique la perception de la couleur. Elle correspond à un angle de rotation en degré autour de l'axe de la composante V. À titre d'exemple, T est égale à 0° pour la teinte rouge et vaut 120° pour la teinte verte. La composante S (saturation) décrit la pureté de la couleur, c'est-à-dire, son caractère vif (saturé) ou terne. La composante V (valeur) indique la quantité de lumière de la couleur, c'est-à-dire son aspect clair ou sombre. La valeur des deux composantes S et V est comprise entre 0 et 1. Pour plus de détails sur la conversion d'images de l'espace RVB vers l'espace TSV on peut se référer au troisième chapitre du livre « *Digital color image processing* » [Koschan and Abidi, 2008].

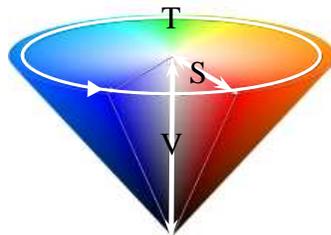


FIGURE. 3.7 – Représentation schématique de l'espace couleur TSV. Pour $S = 0$, V donne des intensités de pixels d'une image en niveaux de gris.

Après la conversion de l'image (RGB vers TSV), les pixels appartenant aux spots laser sont segmentés moyennant un seuillage sur les composantes T, S et V de l'image. La composante T permet d'identifier les pixels qui sont de couleur verte. Le seuillage sur la composante S permet d'éviter la segmentation des zones blanches de la mire de calibrage. En effet, ces zones sont représentées par un faible S dans l'espace TSV (partie centrale du haut du cône de la figure 3.7). Enfin, le seuillage sur la composante V permet d'éviter la segmentation des disques noir. Les zones sombre de l'images sont caractérisées par un

faible V .

Les seuils utilisés dans cette thèse sont fixés en se basant sur une image test regroupant une palette de couleurs (voir figure 3.8). Les seuils pour les valeurs de T , S et V définissent trois intervalles valant respectivement $[72^\circ, 144^\circ]$, $[0.3, 1]$ et $[0.9, 1]$.

L'intervalle $[72^\circ, 144^\circ]$ est centré sur 120° (teinte verte) et permet de détecter la teinte des spots laser. L'intervalle $[0.9, 1]$ permet de détecter les pixels ayant un minimum de quantité de lumière (pixels clairs). L'intervalle $[0.3, 1]$ de détecter les pixels de couleurs vives et surtout d'éliminer les pixels blancs. La figure 3.12 montre le résultat de l'étape de segmentation des spots laser dans une image acquise pour la mire de calibrage.

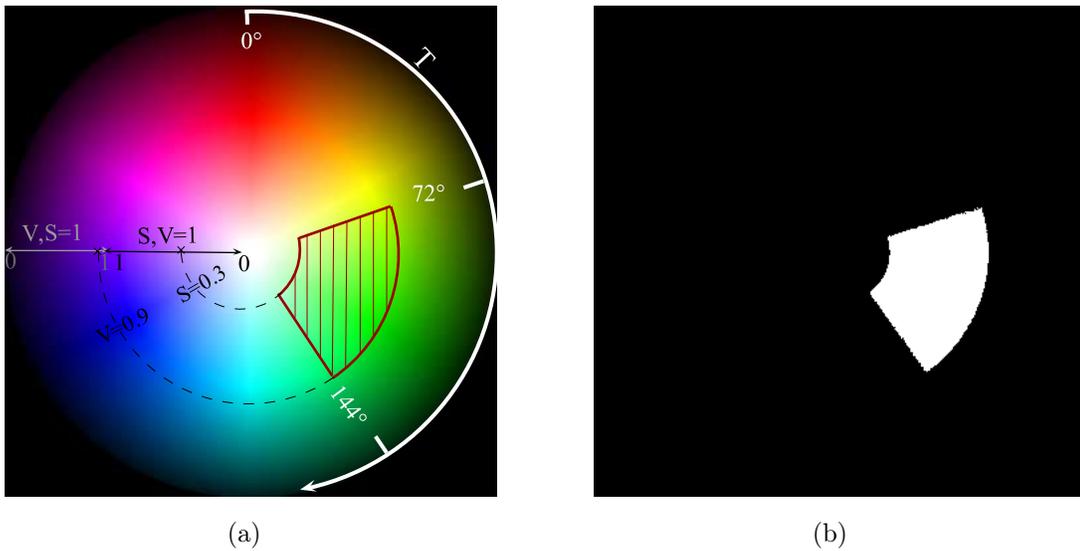


FIGURE. 3.8 – (a) Image test utilisée pour retrouver les seuils sur les composantes T , S et V permettant d'identifier les spots verts dans les images cystoscopiques. La zone hachurée représente les couleurs que peuvent prendre les pixels d'un spot laser dans les images cystoscopiques. (b) Résultat de la segmentation par un seuillage sur les composantes T , S et V . Les pixels blancs sont les pixels segmentés. Les valeurs T , S et V de ces pixels appartiennent respectivement aux intervalles $[72^\circ, 144^\circ]$, $[0.3, 1]$ et $[0.9, 1]$.

Lorsque les pixels appartenant aux spots laser ont été segmentés pour les différentes acquisitions, nous calculons pour chaque ensemble connexe de pixels segmentés un point $P_{2D}^{i,k}$ qui est le centroïde du $i^{\text{ème}}$ ensemble connexe de pixels dans la $k^{\text{ème}}$ image.

La deuxième étape de segmentation d'image concerne les disques situés dans les coins de la mire de calibrage dans les images \mathbf{I}_k . Cette étape de segmentation repose sur le fait que les disques projetés par la caméra ont des formes elliptiques. Nous utilisons la méthode de [Ouellet and Hébert, 2008] qui consiste à estimer les paramètres des ellipses à partir des informations de gradient de l'image \mathbf{I}_k . Finalement, nous exploiterons dans l'étape 2 ci-dessous la position des centres des 4 ellipses.

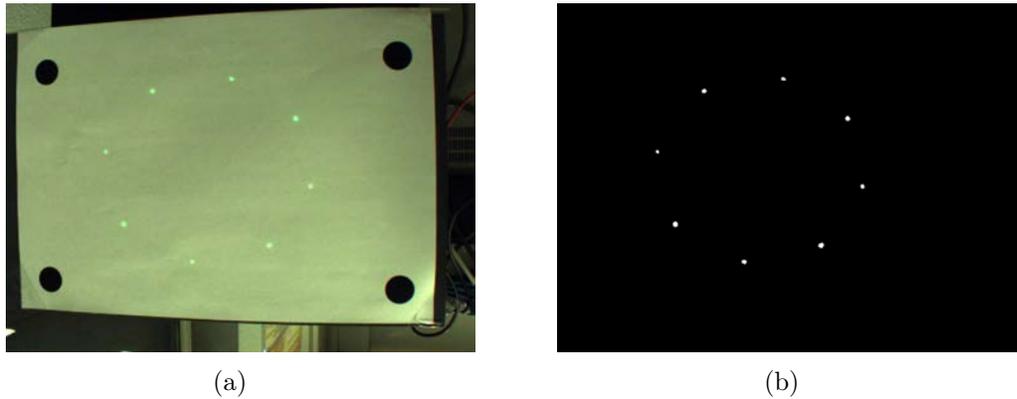


FIGURE. 3.9 – (a). Image acquise pour la mire de calibrage du projecteur. (b) Résultat de la segmentation des spots laser.

Étape 2 : Calcul des plans \mathcal{P}_k et calcul des points $cP_{3D}^{i,k}$

Pour calculer les coordonnées 3D des points laser, nous devons d'abord déterminer l'équation cartésienne des plans \mathcal{P}_k qui supportent le plateau de calibrage dans différentes positions ($k \in \{1, \dots, K\}$) dans le repère de la caméra. La caméra étant préalablement calibrée, nous pourrions, à partir des points segmentés dans les images, définir autant de droites 3D qu'il y a de points laser segmentés. Les coordonnées 3D des points laser seront alors trouvées en calculant l'intersection entre les droites 3D et l'équation de leurs plans respectifs (voir figure 3.10).

Dans la figure 3.10, les points ${}^w m_{3D}$ présentent les centres des quatre disques noirs de la mire de calibrage. Pour ne pas alourdir les notations, nous n'introduisons pas d'indice supplémentaire correspondant au numéro du disque. Les quatre points ${}^w m_{3D}$ (un par disque noir) sont exprimés dans $\{w_k\}$ le repère local au plan de calibrage (sur lequel sont collés les disques noirs). Donc quand le plateau de calibrage est déplacé, les quatre centres ${}^w m_{3D}$ gardent les mêmes coordonnées. Les coordonnées des points ${}^w m_{3D}$ selon \vec{Z}_w valent toutes 0 dans le repère $\{w_k\}$ quelque soit la position k du plateau de calibrage. Nous désignons par $m_{2D,k}$ les correspondants des points ${}^w m_{3D}$ dans le repère d'une image \mathbf{I}_k . Pour chaque acquisition k , l'ensemble de ces points de référence $m_{2D,k}$ et les coordonnées de leurs correspondants ${}^w m_{3D}$ est utilisé pour calculer la transformation $T_{w_k,c}$ qui relie le système de coordonnées monde $\{w_k\}$ au repère de la caméra $\{c\}$ (voir section 2.1.1 du chapitre 2). Il est à noter que les coordonnées des points ${}^w m_{3D}$ dans le repère monde ne changent pas. Par contre, du point de vue de la caméra, elles changent selon la position k de la mire de calibrage dans l'espace défini par le repère caméra.

La transformation $T_{w_k,c}$ est composée d'une rotation \mathbf{R}_k et d'un vecteur de translation \vec{t}_k . Pour faire le lien avec les notations du chapitre 2, l'indice k est ajouté pour indiquer le

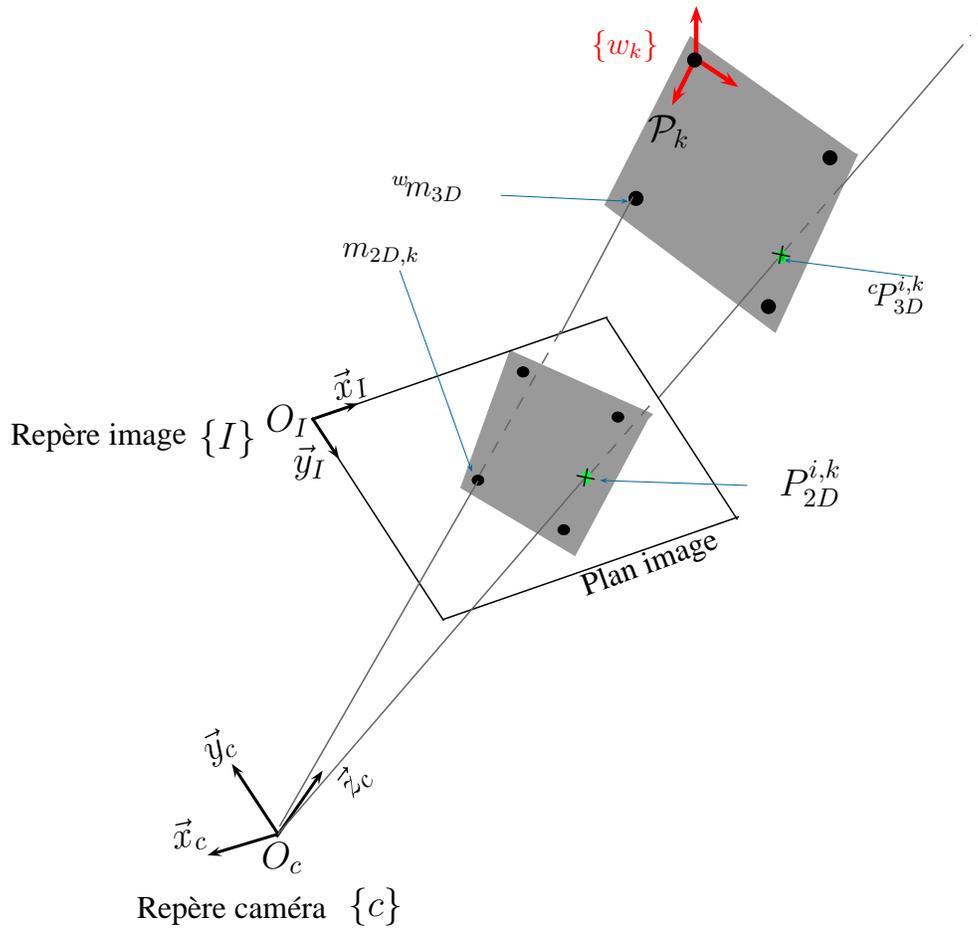


FIGURE. 3.10 – Illustration du calcul des coordonnées de points laser ${}^cP_{3D}^{i,k}$ dans le repère caméra. Ces points sont situés à l'intersection des plans \mathcal{P}_k avec les droites de la projection des points laser 2D $P_{2D}^{i,k}$ dans l'espace.

numéro de l'acquisition en question. Pour une acquisition donnée, la relation entre le plan de calibrage (repère monde) et le plan de l'image peut être représentée par une matrice homogène (3×3) \mathcal{H}_k que l'on appelle homographie. Par conséquent, l'expression de \mathcal{H}_k qui relie les points ${}^w m_{3D}$ à leurs correspondants dans l'image selon le modèle caméra (c'est-à-dire, détaillé dans la section 2.1.1 du chapitre 2) peut être écrite comme suit :

$$\begin{aligned}
 \alpha \begin{bmatrix} m_{2D,k} \\ 1 \end{bmatrix} &= \begin{bmatrix} \mathcal{K} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{R}_k & \vec{\mathbf{t}}_k \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{T_{w_k,c}} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} \mathcal{K} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{r}_{.1,k} & \mathbf{r}_{.2,k} & \mathbf{r}_{.3,k} & \vec{\mathbf{t}}_k \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} \mathcal{K} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{r}_{.1,k} & \mathbf{r}_{.2,k} & \vec{\mathbf{t}}_k \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \\
 &= \mathcal{H}_k \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{3.1}
 \end{aligned}$$

où X , Y et Z sont les coordonnées 3D d'un point de référence dans le repère $\{w_k\}$ et $\mathbf{r}_{.1,k}$, $\mathbf{r}_{.2,k}$ et $\mathbf{r}_{.3,k}$ sont les trois colonnes de la matrice \mathbf{R}_k et α est un facteur d'échelle à l'aide de (3.1). Disposant d'un minimum de quatre points de référence non colinéaires sur la mire de calibrage, il est possible de déterminer \mathcal{H}_k en résolvant un système surdéterminé construit à partir de l'équation (3.1) et de l'ensemble de points référence ${}^w m_{3D}$ et m_{2D} . Les matrices \mathcal{K} et \mathcal{H}_k étant connues, il est possible de calculer $\mathbf{r}_{.1,k}$, $\mathbf{r}_{.2,k}$ et $\vec{\mathbf{t}}_k$. La troisième colonne $\mathbf{r}_{.3,k}$ peut être retrouvée en calculant le produit vectoriel de $\mathbf{r}_{.1,k}$ et $\mathbf{r}_{.2,k}$, puisque les trois colonnes de \mathbf{R}_k doivent être orthonormales.

À l'aide des matrices $T_{w_k,c}$ ainsi estimées, les coordonnées des points de référence ${}^w m_{3D}$ sont ramenées dans le repère de la caméra $\{c\}$ pour retrouver l'équation des plans \mathcal{P}_k de calibrage dans le repère de la caméra.

Les points $P_{2D}^{i,k}$ (centres des spots laser segmentés) sont projetés dans l'espace en utilisant l'inverse de la matrice caméra \mathcal{K} (voir equation de la section 2.1.1 du chapitre 2). Celle-ci génère pour chaque point $P_{2D}^{i,k}$ une droite 3D notée $r_{i,k}$. Finalement, les coordonnées de chaque point ${}^c P_{3D}^{i,k}$ sont obtenues en calculant l'intersection de la droite $r_{i,k}$ avec le plan estimé pour l'acquisition k .

Cette deuxième étape a permis de calculer les coordonnées des points laser ${}^c P_{3D}^{i,k}$ dans

le repère caméra. Cependant, ces points ne sont pas encore mis en correspondance avec leurs rayons laser respectifs. C'est l'objectif de l'étape suivante.

Étape 3 : mise en correspondance automatique entre les points ${}^cP_{3D}^{i,k}$ et les rayons lasers D_j

L'objectif de cette étape est donc de regrouper les points laser en fonction des rayons desquels ils émanent. En d'autres termes, pour chaque point laser, ${}^cP_{3D}^{i,k}$ doit être attribué au rayon laser D_j qui lui correspond. Pour effectuer cette mise en correspondance, nous proposons de faire une première estimation des droites D_j . L'ensemble des droites D_j ainsi estimées autorise la mise en correspondance en attribuant chaque point ${}^cP_{3D}^{i,k}$ à la plus proche droite D_j .

L'idée générale de l'étape 3 est d'optimiser le point ${}^1\hat{O}_P$ (première estimation de la position du centre optique du projecteur) de sorte que tous les points ${}^cP_{3D}^{i,k}$ soient tous simultanément le plus proche possible d'une des droites D_j . Les équations des N droites D_j sont calculées à partir du centre optique du projecteur ${}^1\hat{O}_P$ et des points laser 3D calculés pour la première acquisition (plan \mathcal{P}_1 de la figure 3.10). Ainsi, à chaque étape du processus d'optimisation, les nouvelles droites D_j sont définies par les points ${}^cP_{3D}^{i,1}$ et les coordonnées actualisées du point ${}^1\hat{O}_P$.

La position ${}^1\hat{O}_P$ du centre optique du projecteur est estimée en minimisant la distance de Hausdorff dirigée (DHD) calculée entre les droites D_j et les points 3D ${}^cP_{3D}^{i,k}$:

$${}^1\hat{O}_P = \arg \min_{O_P} \text{DHD} \left(\mathcal{D}, {}^cP_{3D}^{i,k} \right), \quad (3.2)$$

tels que \mathcal{D} représente l'ensemble des rayons D_j passant par O_P et la distance DHD est définie par l'équation (3.3).

$$\text{DHD} \left(\mathcal{D}, {}^cP_{3D}^{i,k} \right) = \max_{j \in \{1, \dots, N\}} \left(\min_{i \in \{1, \dots, N\}; k > 1} d({}^cP_{3D}^{i,k}, D_j) \right) \quad (3.3)$$

d désigne une distance euclidienne. En d'autres termes, on cherche la position ${}^1\hat{O}_P$ qui fait passer chaque rayons laser au plus près de l'un des points 3D laser. Le choix de la distance de Hausdorff s'argumente par le fait qu'elle peut indiquer le degré de superposition de deux structures géométriques (ici un nuage de points et des droites 3D) sans connaître le lien (correspondance) entre ces structures.

Puisque la distance de Hausdorff n'est pas analytiquement dérivable, la minimisation de $\text{DHD} \left(\mathcal{D}, {}^cP_{3D}^{i,k} \right)$ doit être faite par une méthode d'optimisation ne nécessitant pas le calcul du gradient de la fonction coût. Nous avons implémenté la méthode du sim-

plexe [Nelder and Mead, 1965] qui est souvent utilisée pour les problèmes de superposition des structures géométriques (par exemple, de surfaces 3D : voir [Posada *et al.*, 2007; Markelj *et al.*, 2010]). La solution de départ dans notre approche consiste en un simplexe composé d'un ensemble de quatre sommets⁸. Chaque sommet représente une solution de la position du point O_P . Le simplexe se déforme progressivement en enlevant le « pire » des sommets (celui qui correspond à la valeur la plus élevée de $\text{DHD}(\mathcal{D}, {}^cP_{3D}^{i,k})$) et en le remplaçant par un nouveau sommet (donnant une valeur plus faible de $\text{DHD}(\mathcal{D}, {}^cP_{3D}^{i,k})$) calculé à partir des trois meilleurs sommets. Itérativement, les quatre sommets du simplexe se rapprochent et le volume du simplexe se rétrécit pour converger vers la position finale ${}^1\hat{O}_P$. Pour plus de détails sur la méthode du simplexe, nous renvoyons le lecteur à l'annexe A.

8. O_P étant défini de 3 coordonnées, le simplexe doit être composé de 4 sommets. Cette solution initiale est discutée plus loin.

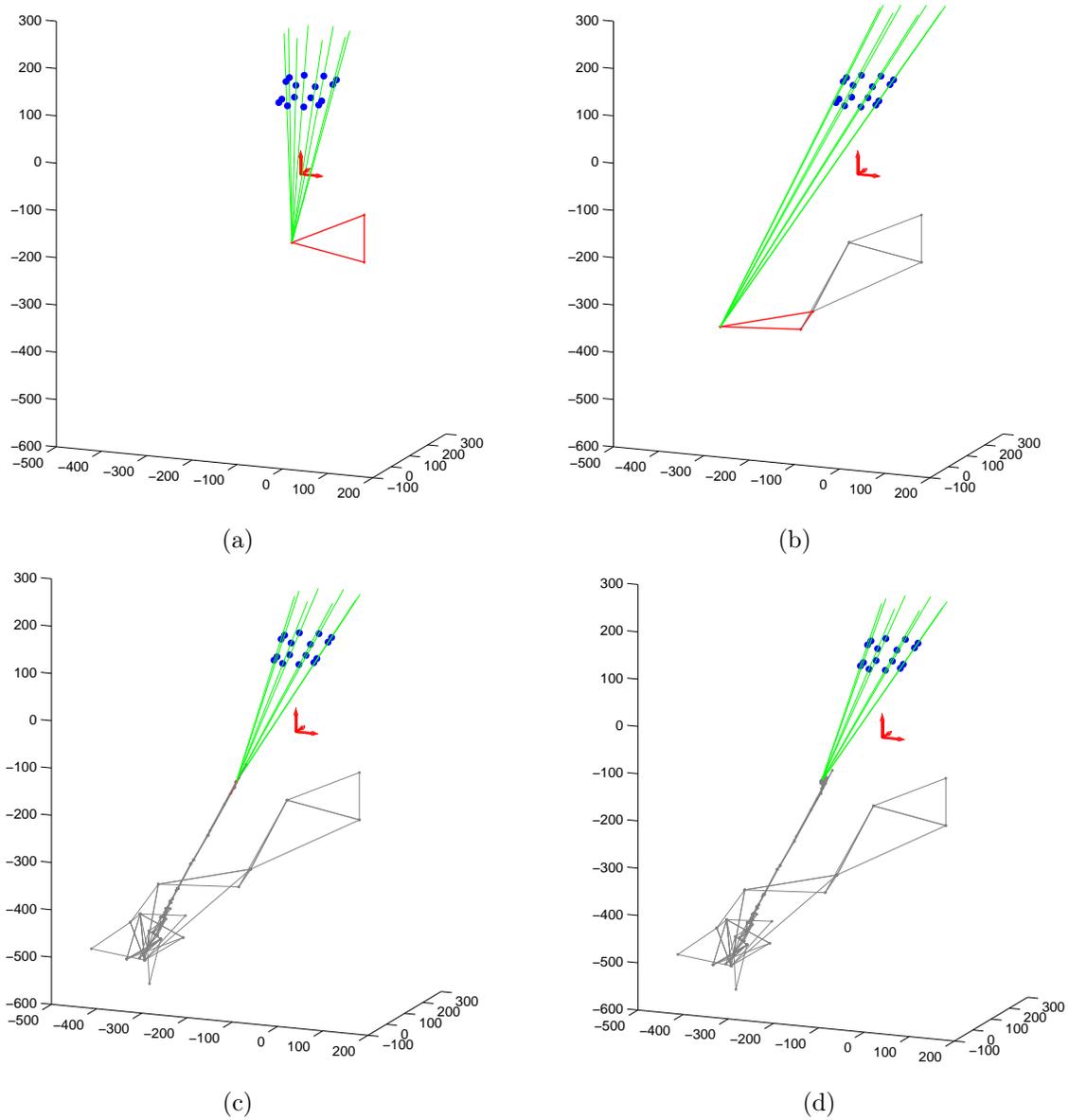


FIGURE. 3.11 – Évolution du simplexe dans le processus d'estimation de $\mathcal{V}\hat{\mathcal{O}}_P$ (position du centre optique du projecteur) en utilisant seulement deux acquisitions du plateau de calibrage. Les sommets du triangle rouge représentent les trois meilleurs sommets du simplexe de l'itération en cours. Les figures (a), (b), (c) et (d) correspondent respectivement aux itérations 1, 4, 50 et 118 (itération finale). Dans (c) et (d), le triangle rouge est très petit. Il est situé au niveau de l'intersection des droites vertes. Dans cet exemple nous avons choisi un simplexe de départ loin de la solution finale pour montrer la robustesse de la méthode.

La figure 3.11 illustre l'évolution de la position des trois premiers sommets du simplexe pour les itérations 1, 4, 50 et 118 du processus d'optimisation. Dans cet exemple, nous avons utilisé uniquement deux acquisitions de la pièce de calibrage ($K = 2$). Le simplexe initial est formé des quatre sommets d'un tétraèdre construit par des règles de symétrie à partir du centre optique de la caméra (le point de coordonnées $[0, 0, 0]^T$). Nous avons vérifié que ce choix assure la convergence du simplexe vers une solution optimale. En effet, dans les systèmes de vision en lumière structurée, le projecteur est généralement placé à côté de la caméra. Les itérations sont arrêtées quand le volume du simplexe (tétraèdre) devient très petit (inférieur à 0.1 mm^3). Les points bleus représentent les points ${}^cP_{3D}^{i,k}$. Les droites vertes sont les droites laser qui passent par le meilleur sommet du simplexe de l'itération en cours (O_P à chaque itération) et par les points ${}^cP_{3D}^{i,1}$ (du plan \mathcal{P}_1). Finalement, le repère en rouge définit le repère de la caméra. Pour la première itération (voir figure 3.11.(a)), on peut remarquer que les droites laser ne passent pas par les points 3D. Au fur et à mesure des itérations (voir figure 3.11.(b), (c) et (d)), le simplexe se déforme pour que ces droites laser passent au plus près des points 3D. Même si la position ainsi estimée pour le centre optique du projecteur est approximative, elle est suffisamment précise pour faire la mise en correspondance entre les points ${}^cP_{3D}^{i,k}$ et les rayons laser D_j . En effet, chaque ${}^cP_{3D}^{i,k}$ peut être affecté à sa droite respective D_j comme suit :

$$j = \arg \min_j d({}^cP_{3D}^{i,k}, D_j). \quad (3.4)$$

L'équation (3.4) est écrite pour chaque point ${}^cP_{3D}^{i,k}$. Ce processus d'optimisation renvoie pour chaque ${}^cP_{3D}^{i,k}$ l'indice de la droite D_j qui lui est le plus proche. Dans l'étape 4, afin de simplifier la lecture, la notation du point ${}^cP_{3D}^{i,k}$ est changée en ${}^cP_{3D}^{j,k}$ pour signifier que le $i^{\text{ème}}$ point du plan k est aussi celui qui appartient à D_j .

Étape 4 : Détermination finale des équations des rayons laser D_j

La dernière étape de notre méthode de calibrage du projecteur permet d'ajuster avec précision les équations des droites D_j . Cet ajustement est effectué en estimant plus finement la position du centre optique O_P du projecteur sous l'hypothèse que chaque droite D_j passe par O_P et par le barycentre ${}^c\bar{P}_{3D}^j$ de tous les points 3D ${}^cP_{3D}^{j,k}$ d'une droite (moyennage sur k). En termes de régression linéaire par moindres carrés, on peut facilement montrer que si on approxime une droite passant au plus près d'un ensemble de points, elle doit passer nécessairement par le barycentre des points.

Puisque les ${}^c\bar{P}_{3D}^j$ sont connus et fixes, les rayons laser D_j ne dépendent que de la position O_P du centre optique du projecteur. Les coordonnées d' O_P sont estimées en minimisant

simultanément l'ensemble des distances entre les points $cP_{3D}^{j,k}$ et les rayons laser D_j :

$$O_P = \arg \min_{O_P} \sum_{k=1}^K \sum_{j=1}^N d(cP_{3D}^{j,k}, D_j(O_P, c\bar{P}_{3D}^j)) \quad (3.5)$$

où N est le nombre de points laser par image et K est le nombre d'images. Nous utilisons à nouveau l'algorithme du simplexe pour la minimisation de la fonction coût de l'équation (3.5). Comme le point ${}^1\hat{O}_P$ estimé dans l'étape précédente est proche du centre optique du projecteur, le simplexe de départ pour cette minimisation est construit à partir des coordonnées de ce point.

3.4 Reconstruction de points 3D

Les paramètres nécessaires pour pouvoir effectuer la reconstruction de données 3D à partir d'images acquises avec le prototype sont la matrice de la caméra \mathcal{K} , les coefficients k_1 et k_2 de la distorsion radiale et l'ensemble des droites D_j modélisant la trajectoire de chaque point laser projeté dans l'espace.

La reconstruction des points 3D (schématisée dans la figure 3.13) est effectuée comme suit pour chaque acquisition.

1. Corriger la distorsion radiale des images acquises en utilisant la matrice \mathcal{K} et les coefficients k_1 et k_2 .
2. Détecter les centres P_{2D}^i des points laser dans les images par segmentation couleur (voir section 3.3.2).
3. Utiliser l'inverse de \mathcal{K} la matrice de la caméra pour calculer l'ensemble des droites r_i issues du centre optique de la caméra et passant par les points P_{2D}^i .
4. Trouver la correspondance entre les rayons caméra r_i et les rayons laser D_j passant par le même point P_{3D}^i . La trajectoire du rayon laser D_j projetée dans le plan de la caméra donne une droite épipolaire [Hartley and Zisserman, 2003]. La projection du point P_{3D}^i (à reconstruire) dans le plan de l'image doit être très proche, si ce n'est exactement, sur la droite épipolaire qui correspond à D_j . Donc, P_{2D}^i est assigné à la plus proche droite épipolaire. À ce stade de la reconstruction, la mise en correspondance entre r_i et D_j est donc effectuée. En pratique, pour des raisons numériques, les r_i et D_j ne s'intersectent pas exactement. Le point P_{3D}^i est donc défini par la position médiane sur le plus court segment liant les deux droites r_i et D_j (voir figure 3.13).

Il est à noter que la méthode de segmentation des spots laser introduite dans la section 3.3.2 reste aussi valable dans le cas de reconstruction 3D à l'intérieur de la vessie. En effet, cette méthode de segmentation permet d'une manière générale de séparer des pixels verts du reste des pixels dans une image. Dans le cas des images cystoscopiques, la teinte dominante est rougeâtre ce qui assure une grande performance de la segmentation des spots laser (voir l'exemple de la figure 3.12).

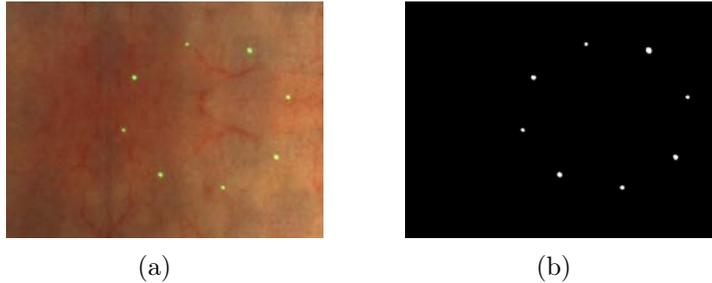


FIGURE. 3.12 – (a). Image acquise avec notre prototype pour un fantôme de vessie construit à partir d'une texture de vessie réaliste (photographie de vessie de porc). (b) Résultat de la segmentation des spots laser.

3.5 Évaluation de la reconstruction 3D

Nous étudions la qualité de la reconstruction de points 3D à partir de notre prototype expérimental. Nous utilisons un objet test dont la géométrie est connue avec précision. Nous vérifions si les points 3D reconstruits décrivent précisément la géométrie de l'objet test. En évaluant la précision de la reconstruction de données 3D, nous validons la qualité du calibrage de notre prototype expérimental.

3.5.1 Expérimentations

La pièce test utilisée a la forme d'un échelon composé de deux demi-plans horizontaux parallèles (voir figure 3.14). La distance (connue au centième de millimètre) entre ces deux plans est de 21.95 mm . Le choix de cette pièce test est inspiré des expérimentations effectuées dans les travaux de [Chan *et al.*, 2003].

Notre système de vision doit être capable de reconstruire précisément la géométrie de la pièce sous différents points de vue (différentes profondeurs et orientations). Pour chaque acquisition, un ensemble de huit points 3D est reconstruit en utilisant la démarche détaillée dans la section 3.4. Nous faisons en sorte que lors des acquisitions, les spots laser soient répartis sur les deux demi-plans parallèles de la pièce (quatre par demi-plan). À partir

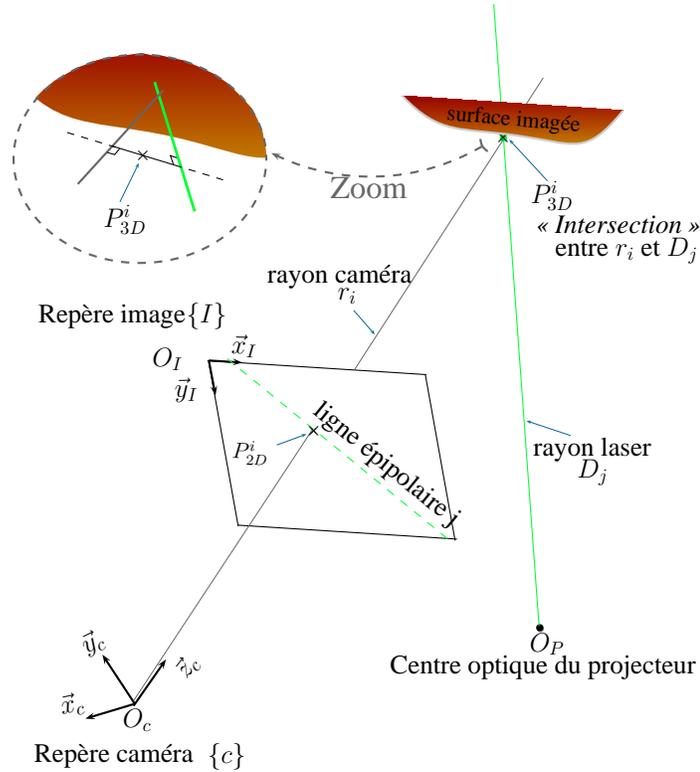


FIGURE. 3.13 – Illustration du calcul des coordonnées d'un point laser 3D. Le point P_{2D}^i segmenté dans l'image est attribué à la plus proche ligne épipolaire (droite verte en pointillés). Le point P_{3D}^i est situé à mi-distance sur le plus court segment entre le rayon caméra r_i et le rayon laser D_j .

des deux plans estimés, nous calculons la hauteur h et l'angle θ entre les demi-plans (θ est censé être nul car les demi-plans sont parallèles).

En plus de ces indicateurs géométriques (h et θ), nous considérons deux mesures d'erreur supplémentaires pour quantifier la précision de la reconstruction 3D. La première, notée ϵ_{3D}^i , correspond à la plus courte distance euclidienne entre le $i^{\text{ème}}$ point 3D reconstruit (P_{3D}^i) et le plan calculé qui lui correspond. La seconde, notée $\epsilon_{\%}^i$, représente le rapport entre l'erreur ϵ_{3D}^i et la profondeur *réelle* du point reconstruit par rapport à la caméra (troisième coordonnée du point localisé sur le plan et le plus proche de P_{3D}^i). L'intérêt de cette dernière valeur normalisée est de ramener l'erreur à la distance de mesure (la précision de la reconstruction 3D dépend autant de l'erreur que de la distance de mesure). Les valeurs moyennes de tous les ϵ_{3D}^i et $\epsilon_{\%}^i$ sont respectivement notées $\bar{\epsilon}_{3D}$ et $\bar{\epsilon}_{\%}$.

3.5.2 Résultats

L'évaluation de la précision et de la robustesse de la reconstruction avec notre méthode de calibrage est donc réalisée avec les données et les objectifs suivants.

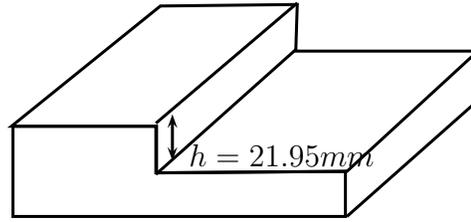


FIGURE. 3.14 – Géométrie de la pièce échelon utilisée.

- Pour étudier la robustesse (sensibilité de la reconstruction vis-à-vis du nombre d’images de calibration), nous avons testé la reconstruction avec des paramètres de calibration calculés respectivement avec 2, 10 et 20 images de calibration. Pour cela, 120 acquisitions de la pièce placée à différentes positions (distances du capteur et inclinaisons de la pièce) sont utilisées pour évaluer l’impact du nombre d’images de calibration sur la reconstruction.
- Pour étudier la précision de la reconstruction, nous avons utilisé 60 images de la pièce test placée à diverses distances de la caméra, mais avec une inclinaison constante par rapport au capteur. Les distances entre la caméra et la pièce appartiennent à l’intervalle $[150, 300]$ mm . Les paramètres de calibration utilisés sont ceux obtenus avec 2 images de calibration.
- Pour vérifier la précision de la reconstruction en fonction de l’inclinaison de la pièce par rapport à la caméra, nous avons encore acquis 60 images. La pièce est placée à 250 mm de la caméra et l’angle d’inclinaison de la pièce varie entre -60° et 60° (l’angle 0° correspond à l’axe principal de la caméra qui est perpendiculaire à la pièce). Pour ces tests nous avons à nouveau utilisé les paramètres de calibration calculés avec 2 images.

3.5.2.1 Précision de la reconstruction en fonction du nombre d’images de calibration

Nous disposons pour ces tests de trois jeux de paramètres de calibration calculés pour 2, 10 et 20 images. À chaque acquisition de la pièce (et à l’aide des paramètres de calibration), un ensemble de 8 points 3D est reconstruit en utilisant la méthode décrite dans la section 3.4. Ces points sont ensuite utilisés pour trouver l’équation des deux plans parallèles (voir l’exemple de la figure 3.15). Enfin, les indices géométriques θ et h sont calculés et comparés à leurs valeurs réelles. Le tableau 3.1 résume les différentes mesures 3D et les erreurs de reconstruction obtenues pour les 120 acquisitions. Il est à noter que les erreurs de reconstruction 3D sont petites ($\bar{\epsilon}_{3D} \approx 0.15$ mm et $\bar{\epsilon}_\% \approx 0.07$ $\%$). La valeur de h est très proche de sa valeur connue et l’angle θ tend vers 0. La position et la forme de la pièce sont

donc estimées avec précision en moyenne. Tout en étant plus flexible, notre méthode de calibrage a donné des résultats comparables que ceux obtenus dans [Marzani *et al.*, 2002; Chan *et al.*, 2003].

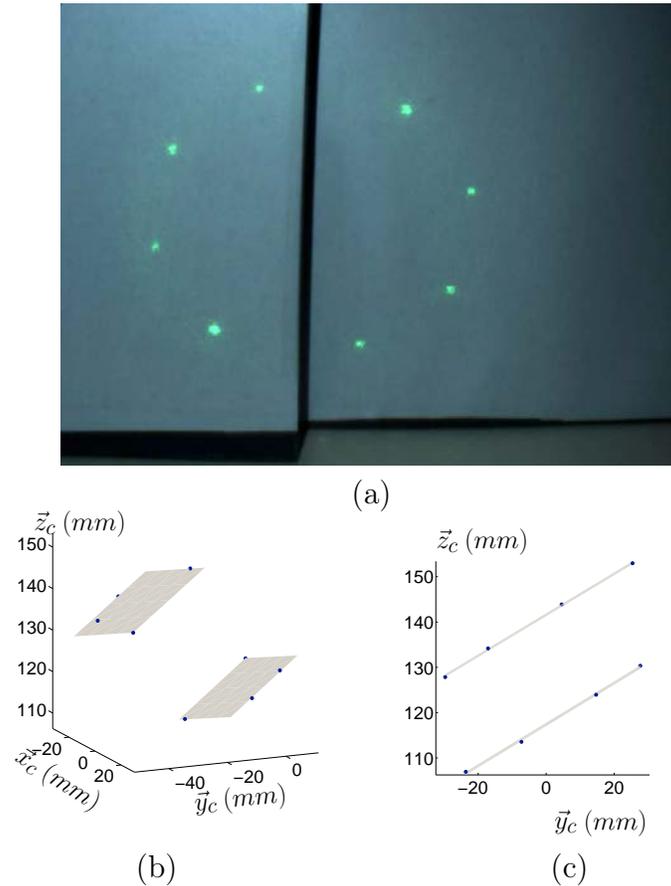


FIGURE. 3.15 – Exemple de reconstruction 3D de la pièce test. (a) Exemple d'image acquise pour la pièce échelon. (b) et (c) Deux différents points de vue des plans parallèles estimés à partir des huit points 3D reconstruits à partir de l'image (a). Le projecteur est calibré avec seulement deux images.

Ces tests expérimentaux montrent aussi que les résultats de la reconstruction ne dépendent quasiment pas du nombre d'acquisitions utilisées pour le calibrage. En effet, les erreurs calculées avec différentes configurations (en utilisant 2, 10 ou 20 images pour le calibrage) sont approximativement les mêmes. En conclusion, on peut admettre que deux acquisitions sont suffisantes pour calibrer le projecteur avec une bonne précision.

3.5.2.2 Précision en fonction de la distance de la pièce test au capteur

Les valeurs de la hauteur h obtenues pour différentes distances de mesure sont présentées dans la figure 3.16(a). Nous remarquons que plus la distance de la pièce à la

TABLEAU. 3.1 – Mesures 3D et erreurs obtenues pour la reconstruction de la pièce échelon. Les valeurs données entre parenthèses pour h représentent l'écart avec la valeur étalon $h = 21.95$ mm.

	Nombre d'images de calibrage		
	2	10	20
Moyenne d'angle θ (0°)	0.3°	0.3°	0.3°
Moyenne de la hauteur h (21.95 mm)	21.62 (-0.33 mm)	21.7 (-0.25 mm)	21.67 (-0.28 mm)
$\bar{\epsilon}_{3D}$ (mm)	0.150	0.142	0.146
$\bar{\epsilon}_\%$ (%)	0.075	0.071	0.073

caméra augmente, plus les erreurs de mesures sur h augmentent. La distribution des erreurs est donnée dans la figure 3.16(b). Cette distribution est caractérisée par une valeur moyenne de 0.22 mm et par un écart type de 0.29 mm. L'erreur de reconstruction est sub-millimétrique pour un intervalle de distance relativement grand [150, 300] mm.

La figure 3.17(a) montre l'évolution de la mesure de l'angle θ en fonction de la profondeur de la pièce. Les valeurs de θ ont tendance à s'écartier de 0 lorsque la profondeur de la pièce augmente. Cependant, la valeur moyenne et l'écart type de θ valent respectivement 0.26° et 1.06° (voir figure 3.17(b)). L'erreur de l'estimation de l'angle θ est donc relativement faible. Cela signifie qu'indépendamment de la profondeur, les deux demi-plans tendent à rester parallèles.

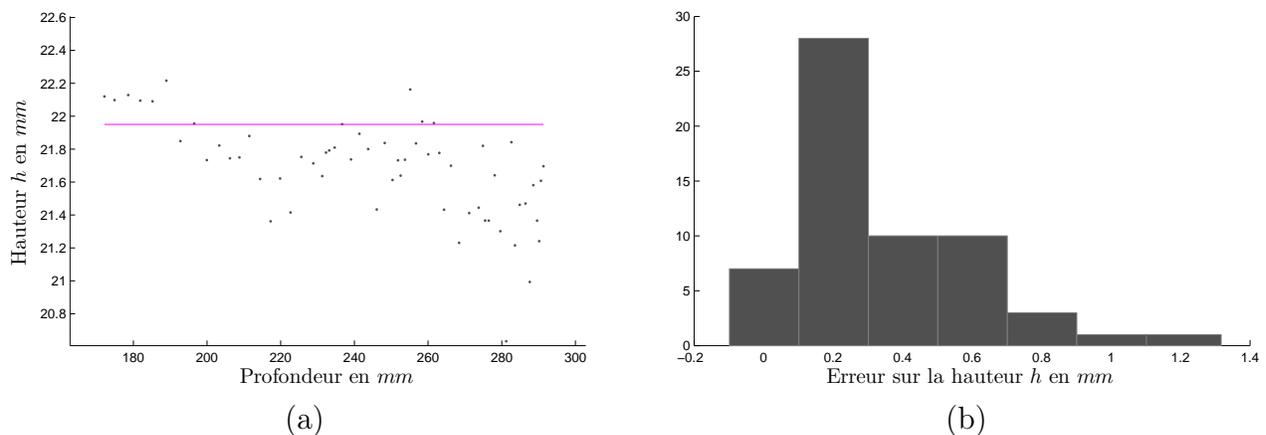


FIGURE. 3.16 – (a) Évolution de la hauteur h en fonction de la distance (profondeur) de la pièce test par rapport à la caméra. La barre horizontale indique la valeur théorique de h (21.95 mm). (b) Distribution des erreurs effectuées pour h .

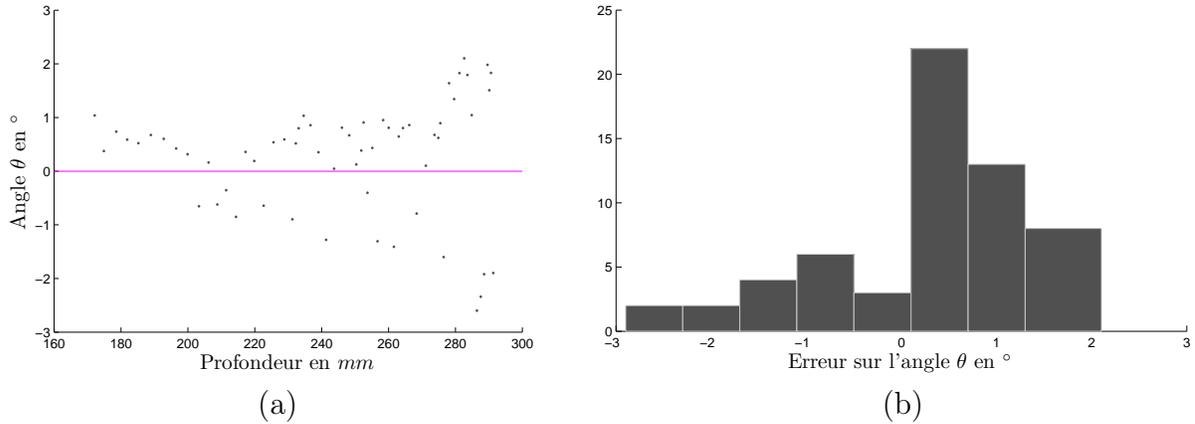


FIGURE. 3.17 – (a) Évolution de l’angle θ en fonction de la distance (profondeur) entre la pièce test et la caméra. La barre horizontale indique la valeur théorique de θ (0°). (b) Distribution des erreurs sur θ .

3.5.2.3 Précision en fonction de l’inclinaison de la pièce test

La figure 3.18(a) montre l’évolution de la hauteur h en fonction de l’angle d’inclinaison de la pièce par rapport à l’axe principal de projection des points laser. La hauteur h tend vers sa valeur réelle (21.95 mm), lorsque l’inclinaison tend vers 0° . La moyenne et la l’écart type de l’erreur sur l’hauteur h sont respectivement, 0.47 mm et 0.37 mm (voir figure 3.18(b) pour la distribution des valeurs). Ces résultats montrent que la dimension (distance entre les deux plans) de la pièce test a été conservée avec précision pour de différentes orientations de la pièce.

Le figure 3.19(a) présente les différentes valeurs de l’angle θ obtenues pour les différents angles d’inclinaison de la pièce variant de -60° à 60° . On remarque que plus la pièce est orthogonale à l’axe principal de projection des points laser, plus l’angle θ tend vers 0° . Les erreurs sur θ augmentent lorsque l’angle de l’inclinaison de la pièce augmente en valeur absolue. La distribution des valeurs de θ est caractérisée par une moyenne de 0.33° et un écart type de 0.81° (voir figure 3.19(b)). Ces valeurs obtenues montrent que la structure de la pièce (les deux plans doivent rester parallèles) est conservée même lorsque la pièce est fortement inclinée. Il est à noter que ces tests sont relativement sévères (inclinaison allant jusqu’à $\pm 60^\circ$) puisque dans le cas des examens cystoscopiques les cliniciens essayent intuitivement de placer la tige du cystoscope orthogonale à la surface imagée.

La dégradation de la qualité de reconstruction 3D observée dans le cas d’une forte inclinaison de la pièce test peut être expliquée en partie par le fait qu’il y a une grande réflexion de lumière laser sur la surface de la pièce. Plus l’inclinaison est importante, plus le spot laser se déforme sur les surfaces de la pièce test ce qui rend la segmentation des

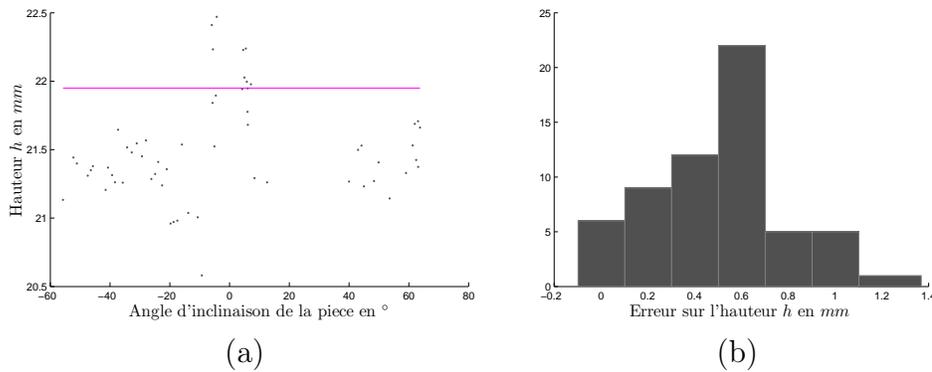


FIGURE. 3.18 – (a) Évolution de la hauteur h en fonction de l’inclinaison de la pièce test par rapport à l’axe principal de projection de points laser (c’est-à-dire, angle entre le vecteur directeur de l’axe de projection et le vecteur normal aux plans de la pièce). La barre horizontale indique la valeur théorique de h (21.95 mm). (b) Distribution des erreurs effectuées pour h .

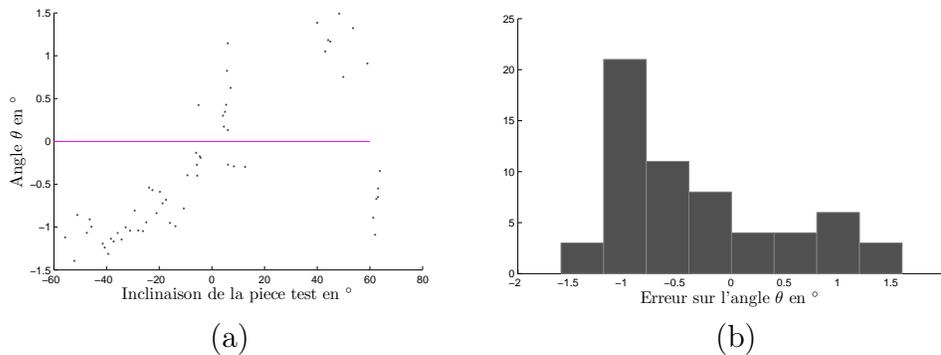


FIGURE. 3.19 – (a) Évolution de l’angle θ en fonction de l’inclinaison de la pièce test par rapport à l’axe principal de projection de points laser (c’est-à-dire, angle entre le vecteur directeur de l’axe de projection et le vecteur normal aux plans de la pièce). La barre horizontale indique la valeur théorique de θ (0°). (b) Distribution des erreurs sur θ .

points laser moins précise. Ce phénomène est réduit d’une manière importante lors de l’application clinique, car les tissus biologiques humains absorbent plus la lumière laser et par conséquent les spots laser sont plus ponctuels [Chan *et al.*, 2003].

3.5.3 Commentaires et bilan général sur les résultats

Flexibilité de la méthode. Notre méthode de calibrage de projecteur de lumière structurée est bien adaptée pour les systèmes portables de vision active. En effet, ces systèmes sont généralement maintenus à la main pour scanner et reconstruire une scène en 3D. Il est avantageux de pouvoir calibrer ce type de système sans utiliser d’équipements de positionnement chers et précis. Notre méthode ne nécessite aucun de ces équipements. Il suffit de maintenir l’instrument à la main en face d’une mire de calibrage simple (feuille

blanche avec quatre disques placés dans les coins) et le déplacer orthogonalement à la mire. Nous avons montré que deux acquisitions de la mire suffisent pour conduire à une reconstruction 3D précise et pour un espace de travail assez large.

Précision de la méthode. Les résultats obtenus témoignent de la grande précision de la reconstruction 3D pour diverses positions et inclinaisons de l'axe de l'instrument par rapport aux surfaces. Or durant un examen cystoscope, les cliniciens ont tendance à maintenir naturellement l'axe de l'instrument perpendiculairement à la surface. C'est dans cette situation que la reconstruction est la plus précise. Il est à noter également que la précision obtenue permet de faire de la mesure tridimensionnelle de pièces manufacturées. Cependant, nous n'avons pas besoin d'une telle précision dans le cas de la cystoscopie où il faut plutôt reconstruire une forme réaliste.

Robustesse de la méthode. Outre les résultats précités, nous avons testé la robustesse de la méthode du simplexe *en fonction* de la position *initiale* du centre optique du projecteur (*valeur initiale du simplexe*). Nous avons pu vérifier expérimentalement que la méthode converge systématiquement vers une solution optimale en initialisant le simplexe avec un tétraèdre régulier centré sur le centre optique de la caméra et dont la longueur des arêtes est assez grande pour couvrir approximativement la distance entre l'optique de la caméra et celle du projecteur.

3.6 Conclusion

Nous avons présenté dans ce chapitre le prototype expérimental utilisé dans la suite de cette thèse. Nous avons aussi présenté une nouvelle méthode flexible de calibrage de projecteurs de points lumineux qui ne nécessite qu'un simple plateau de calibrage constitué d'une mire imprimée sur une feuille fixée sur un plateau en PVC. La méthode proposée ne nécessite aucun système de positionnement. Nous avons montré que notre méthode permet une reconstruction de points 3D avec une grande précision (erreur 3D relative moyenne d'environ 0.07%). Un autre avantage de notre méthode de calibrage est qu'elle peut être utilisée pour le calibrage de projecteurs de points formant différents motifs. Dans notre cas, les points du motif lumineux forment un cercle, mais la méthode fonctionne aussi pour un ensemble de points non régulièrement espacés voire situés aléatoirement dans le plan.

Contributions

- Conception et réalisation d'un prototype expérimental à vision active.
- Développement d'une nouvelle méthode flexible et précise pour le calibrage de projecteur de points lumineux.

Publication

- [Ben-Hamadou *et al.*, 2010b] présenté à l'occasion de la conférence internationale *IEEE International Conference on Image Processing (ICIP)*.

4

Construction de surfaces étendues de la paroi interne de la vessie

Ce chapitre présente une méthode de reconstruction de cartes 3D étendues de la paroi interne de la vessie à partir d'une séquence d'«images cystoscopiques enrichies». Ces données (8 points 3D reconstruits et une image 2D enrichie des projections des 8 points) sont connues dans le repère lié au cystoscope et propre à chaque acquisition (voir chapitre 3). Il s'agit maintenant de placer l'ensemble des données de la séquence dans un repère global commun afin de générer :

1. un nuage de points 3D situés sur la surface de la paroi interne de la vessie.
2. une image panoramique 2D représentant la texture de l'ensemble de la surface imagée.

La reconstruction du nuage de points 3D permettra, par interpolation, de générer une surface 3D sur laquelle l'image panoramique 2D pourra être projetée. Dans le chapitre 2, nous avons vu que la construction d'une image 2D panoramique nécessite le recalage de chaque paire d'images consécutives de la séquence (estimation d'une transformation 2D perspective). De la même manière, la reconstruction du nuage de points 3D revient implicitement à estimer le mouvement 3D du cystoscope (transformation 3D rigide) entre chaque paire d'acquisitions consécutives.

Dans ce chapitre, nous proposons une nouvelle méthode pour l'estimation du mouvement 3D, guidée par un recalage d'images 2D enrichies. Notre méthode se base sur la relation mathématique qui lie les transformations rigides 3D et les transformations perspectives 2D qui mettent en relation deux images cystoscopiques consécutives pour ramener le problème d'estimation de mouvement à un problème de recalage d'images. Nous

évaluerons finalement la précision et la robustesse inhérentes à la méthode proposée sur des données simulées à partir d'un fantôme numérique simulé) de vessie. Cette évaluation nous permettra de valider le principe de notre méthode. Une évaluation plus poussée de l'algorithme est proposée dans le chapitre 5 à partir de fantômes physiques (réels).

4.1 Données disponibles et transformations géométriques

Dans cette section, nous décrivons les liens géométriques qui relient les données d'une même séquence d'acquisition. Nous commençons par définir les notations et les relations géométriques entre les données d'une même acquisition. Puis, nous établissons les relations entre les données issues de deux acquisitions consécutives. Nous distinguons deux types de relations géométriques :

- la transformation 3D rigide qui décrit le mouvement du cystoscope
- la transformation 2D perspective qui modélise le mouvement 2D entre les régions homologues de deux images consécutives.

4.1.1 Données disponibles

Notre prototype expérimental (voir section 3.1 du chapitre 3) permet d'acquérir des images couleur de la surface interne de la vessie contenant en plus un ensemble de spots laser. La méthode de reconstruction 3D (voir chapitre 3) permet de trouver les coordonnées 3D du centre de chaque spot laser localisé sur la surface imagée (voir figure 4.1). Ainsi, pour une acquisition k donnée, les données disponibles sont formées des éléments suivants :

- une image couleur \mathbf{I}_k ,
- un ensemble de N paires de points 3D/2D ($P_{3D}^{i,k}$ et $P_{2D}^{i,k}$). L'indice i fait référence au $i^{\text{ème}}$ point et appartient à $\{1, \dots, N\}$, avec $N = 8$ pour notre prototype expérimental. Chaque point $P_{3D}^{i,k}$ est situé sur la surface interne de la vessie, supposée fixe quand k varie, et a pour projection $P_{2D}^{i,k}$ dans l'image \mathbf{I}_k .

4.1.2 Transformation projective

Les coordonnées points $P_{3D}^{i,k}$ sont définies dans le repère caméra $\{c_k\}$ dont l'origine est notée O_{c_k} . Les coordonnées des points $P_{3D}^{i,k}$ sont définies comme suit :

$$P_{3D}^{i,k} = \begin{bmatrix} x_{3D}^{i,k} \\ y_{3D}^{i,k} \\ z_{3D}^{i,k} \end{bmatrix}. \quad (4.1)$$

Les coordonnées des points $P_{2D}^{i,k}$ sont définies dans le repère image $\{\mathbf{I}_k\}$ comme suit :

$$P_{2D}^{i,k} = \begin{bmatrix} x_{2D}^{i,k} \\ y_{2D}^{i,k} \end{bmatrix}. \quad (4.2)$$

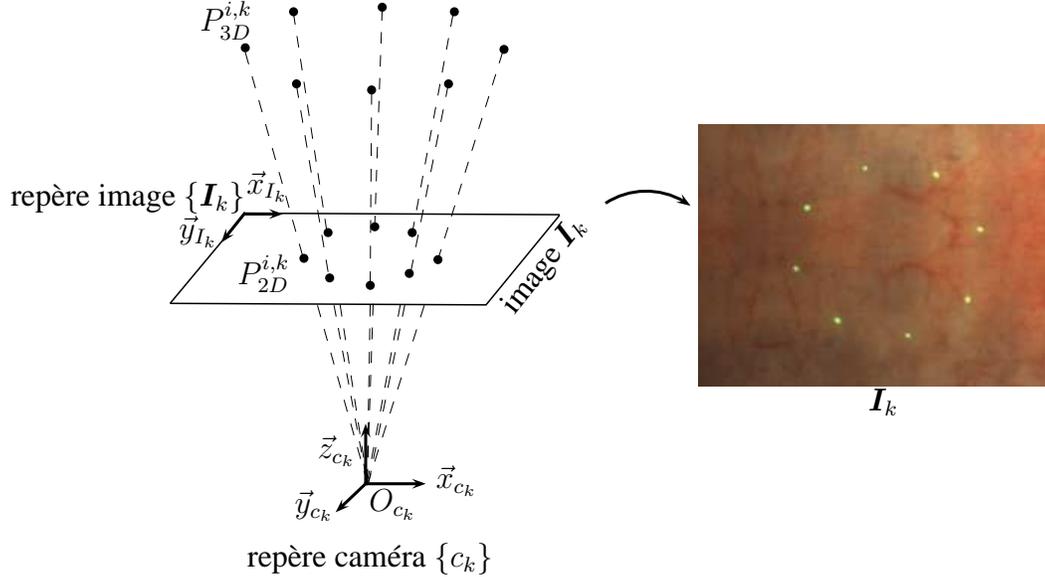


FIGURE. 4.1 – Représentation des repères caméra et image, et des données disponibles pour une acquisition k donnée.

Il est à noter que la distorsion radiale est corrigée dans les images au cours du processus de reconstruction 3D des points laser (voir section 3.4 du chapitre 3). Dans ce chapitre, nous supposons donc que les images \mathbf{I}_k sont dénuées de distorsions.

La relation géométrique entre les points $P_{3D}^{i,k}$ et leurs correspondants $P_{2D}^{i,k}$ dans l'image est formulée dans l'équation (4.3) où \mathcal{K} est la matrice caméra préalablement calculée dans le processus de calibrage du prototype. La matrice \mathcal{K} , de taille 3×4 , est entièrement définie par l'équation 2.7 du chapitre 2. Cette matrice est constante et indépendante du point de vue k .

$$z_{3D}^{i,k} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} = \mathcal{K} P_{3D}^{i,k} \quad (4.3)$$

4.1.3 Mouvement 3D entre deux acquisitions consécutives

Le déplacement d'une caméra (ou dans notre cas d'un cystoscope) correspond à une transformation rigide [Hartley and Zisserman, 2003]. On notera $\mathbf{T}_{3D}^{k-1,k}$ cette transformation 3D rigide. $\mathbf{T}_{3D}^{k-1,k}$ est une matrice de taille 4×4 englobant six degrés de liberté : trois angles de rotation (θ , ϕ , ψ) et trois paramètres de translation (t_x , t_y , t_z). Ces rotations

et translations sont effectuées respectivement autour et selon les axes \vec{x}_{c_k} , \vec{y}_{c_k} et \vec{z}_{c_k} du repère caméra. Les paramètres de $\mathbf{T}_{3D}^{k-1,k}$ sont définis dans l'équation (4.4).

$$\mathbf{T}_{3D}^{k-1,k} = \begin{bmatrix} \cos \theta \cos \psi & -\cos \phi \sin \psi + \sin \phi \sin \theta \sin \psi & \sin \phi \sin \psi + \cos \phi \sin \theta \cos \psi & t_x \\ \cos \theta \sin \psi & \cos \phi \cos \psi + \sin \phi \sin \theta \sin \psi & -\sin \phi \cos \psi + \cos \phi \sin \theta \sin \psi & t_y \\ -\sin \theta & \cos \phi \sin \theta & \cos \phi \cos \theta & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.4)$$

$\mathbf{T}_{3D}^{k-1,k}$ permet d'exprimer les coordonnées des points $P_{3D}^{i,k}$ dans le repère $\{c_{k-1}\}$.

$$\begin{bmatrix} \widehat{P}_{3D}^{i,k-1} \\ 1 \end{bmatrix} = \begin{bmatrix} \widehat{x}_{3D}^{i,k-1} \\ \widehat{y}_{3D}^{i,k-1} \\ \widehat{z}_{3D}^{i,k-1} \\ 1 \end{bmatrix} = \mathbf{T}_{3D}^{k-1,k} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \quad (4.5)$$

Il est important de faire la différence entre les points $\widehat{P}_{3D}^{i,k-1}$ (**transformation des points à l'instant k dans le repère $k-1$**) et $P_{3D}^{i,k-1}$ (**points reconstruits à l'instant $k-1$**), car ces points ne représentent pas les mêmes positions sur la surface imagée. En effet, quand le clinicien déplace le cystoscope dans l'organe, les points lasers projetés changent de localisation sur la surface de l'organe.

4.1.4 Mouvement 2D entre deux acquisitions consécutives

Le mouvement 2D entre deux images consécutives \mathbf{I}_{k-1} et \mathbf{I}_k est modélisé par une transformation perspective $\mathbf{T}_{2D}^{k-1,k}$ (les paramètres $a_{mn}^{k-1,k}$ de cette transformation perspective sont décrits dans la section 2.1.2.2 du chapitre 2). Par analogie au mouvement 3D, $\mathbf{T}_{2D}^{k-1,k}$ déplace les régions communes aux deux images \mathbf{I}_k sur l'image \mathbf{I}_{k-1} dans le repère de l'image \mathbf{I}_{k-1} . La transformation d'un point $P_{2D}^{i,k}$ du repère $\{\mathbf{I}_k\}$ est formulée par :

$$\beta^{i,k-1} \begin{bmatrix} \widehat{P}_{2D}^{i,k-1} \\ 1 \end{bmatrix} = \beta^{i,k-1} \begin{bmatrix} \widehat{x}_{2D}^{i,k-1} \\ \widehat{y}_{2D}^{i,k-1} \\ 1 \end{bmatrix} = \mathbf{T}_{2D}^{k-1,k} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11}^{k-1,k} & a_{12}^{k-1,k} & a_{13}^{k-1,k} \\ a_{21}^{k-1,k} & a_{22}^{k-1,k} & a_{23}^{k-1,k} \\ a_{31}^{k-1,k} & a_{32}^{k-1,k} & 1 \end{bmatrix}}_{\mathbf{T}_{2D}^{k-1,k}} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} \quad (4.6)$$

Le scalaire $\beta^{i,k-1}$ est le produit scalaire de la 3-ème ligne de $\mathbf{T}_{2D}^{k-1,k}$ et de $[P_{2D}^{i,k}, 1]^T$.

4.2 Stratégie de la construction de surfaces 3D étendues

Nous commençons par présenter le principe du positionnement des points 3D dans un repère global et commun. Cela nous conduira à estimer le mouvement 3D du cystoscope entre deux acquisitions $k - 1$ et k . Nous passerons brièvement en revue les techniques existantes pour estimer le mouvement tout en discutant leur adaptabilité à notre contexte. Finalement, nous présenterons l'approche choisie.

4.2.1 Formulation du problème

Sans perte de généralité, nous choisissons les repères $\{c_1\}$ et $\{\mathbf{I}_1\}$ liés à la première acquisition comme repères globaux pour construire respectivement la surface représentée par le nuage de points 3D et l'image panoramique (voir l'illustration schématique de la figure 4.2).

Les points 3D $P_{3D}^{i,k}$ de chaque acquisition k sont exprimés dans le repère commun $\{c_1\}$ à l'aide de la transformation globale $\mathbf{T}_{3D}^{1,k}$. Par opposition aux transformations « locales » $\mathbf{T}_{3D}^{k-1,k}$ qui font le lien entre deux acquisitions consécutives, les transformations « globales » $\mathbf{T}_{3D}^{1,k}$ permettent de passer du repère $\{c_k\}$ au repère global $\{c_1\}$. Les transformations $\mathbf{T}_{3D}^{1,k}$ sont définies par :

$$\mathbf{T}_{3D}^{1,k} = \mathbf{T}_{3D}^{1,k-1} \times \mathbf{T}_{3D}^{k-1,k} = \prod_{j=2}^k \mathbf{T}_{3D}^{j-1,j} \quad (4.7)$$

À l'instar du placement des points 3D dans un repère global, la construction de l'image panoramique est effectuée en plaçant les pixels de chaque image \mathbf{I}_k dans le repère global $\{\mathbf{I}_1\}$ à l'aide d'une transformation perspective globale $\mathbf{T}_{2D}^{1,k}$.

$$\mathbf{T}_{2D}^{1,k} = \mathbf{T}_{2D}^{1,k-1} \times \mathbf{T}_{2D}^{k-1,k} = \prod_{j=2}^k \mathbf{T}_{2D}^{j-1,j} \quad (4.8)$$

Les deux supports panoramiques (*i.e.*, le nuage de points 3D représentant la surface de la vessie et l'image panoramique 2D) sont liés car le point $P_{2D}^{i,k}$ ramené dans le repère $\{\mathbf{I}_1\}$ correspond physiquement au point $P_{3D}^{i,k}$ ramené dans le repère $\{c_1\}$. Ce lien permettra de coller (projeter) l'image panoramique sur la surface 3D définie à partir du nuage de points 3D.

Le défi à relever pour la cartographie 3D est donc de **trouver les différentes transformations $\mathbf{T}_{3D}^{k-1,k}$ et $\mathbf{T}_{2D}^{k-1,k}$ modélisant les mouvements 3D et 2D entre chaque paire d'acquisitions consécutives $k - 1$ et k de la séquence.**

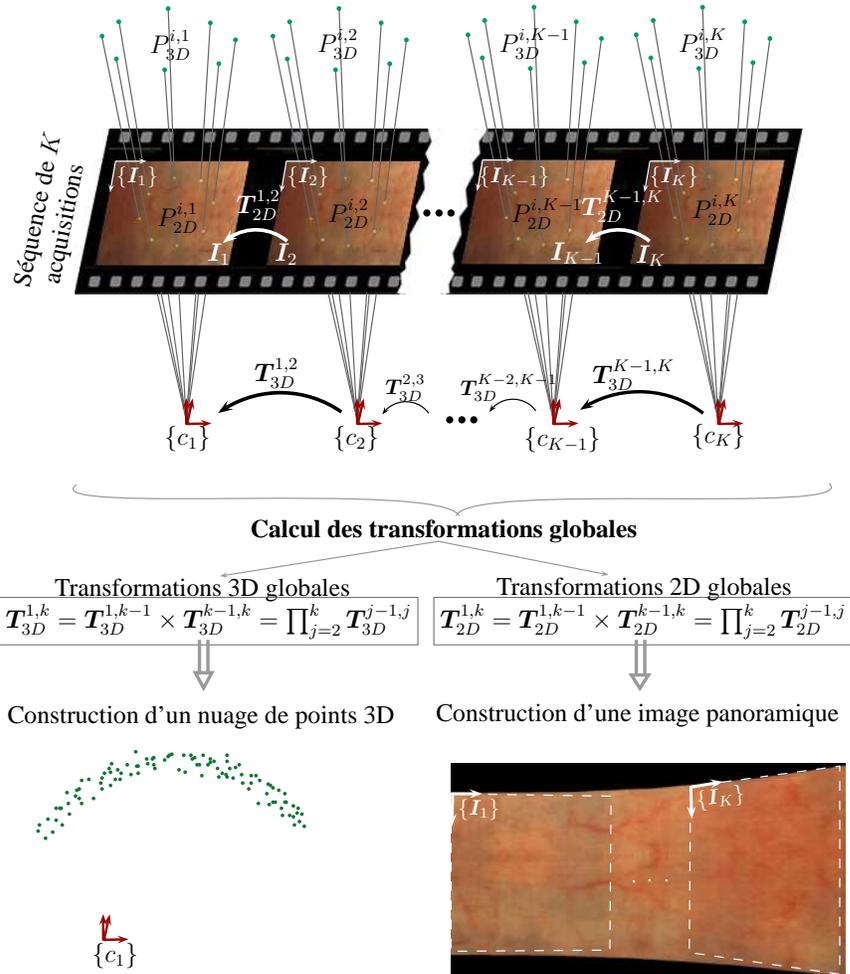


FIGURE. 4.2 – Construction d'un nuage de points 3D dans un repère global et d'une image panoramique à partir d'une séquence de K acquisitions cystoscopiques. Les carrés blancs dans l'image panoramique indiquent les positions de la première et la dernière image de la séquence.

Au chapitre 2, nous avons abordé avec détail les techniques de recalage d'images cystoscopiques (*i.e.*, l'estimation des transformations perspectives $\mathbf{T}_{2D}^{k-1,k}$). Dans la section qui suit, nous passons brièvement en revue les principales méthodes d'estimation de mouvement 3D.

4.2.2 Techniques d'estimation de mouvement 3D

Dans la littérature, on distingue deux familles de méthodes. Les premières utilisent des données 3D [Salvi *et al.*, 2007] alors que les secondes sont uniquement basées sur une séquence d'images 2D [Hartley and Zisserman, 2003].

4.2.2.1 Estimation de mouvement 3D à partir de points 3D

L'estimation de mouvement 3D (matrice $\mathbf{T}_{3D}^{k-1,k}$) peut être réalisée à l'aide de méthodes de recalage qui superposent des points acquis aux instants $k - 1$ et k . Ces méthodes reposent sur l'existence de points homologues contenus dans les acquisitions $k - 1$ et k , c'est-à-dire représentant des points identiques sur la surface observée à partir de deux points de vue différents. Souvent, ces méthodes sont utilisées pour la modélisation 3D d'objets ou de scènes. Dans [Sansoni *et al.*, 2009], on trouve une synthèse des applications qui se basent sur le recalage de données 3D dans divers domaines comme la médecine légale, l'industrie ou l'archéologie. Pour ces applications, un système de reconstruction 3D fournit un ensemble de données 3D d'un objet sous différents points de vue. Ensuite, ces données sont superposées pour construire l'objet 3D en calculant les transformations 3D rigides entre les différentes acquisitions. La figure 4.3 illustre ce processus de modélisation 3D d'un artefact par recalage de deux nuages de points 3D.

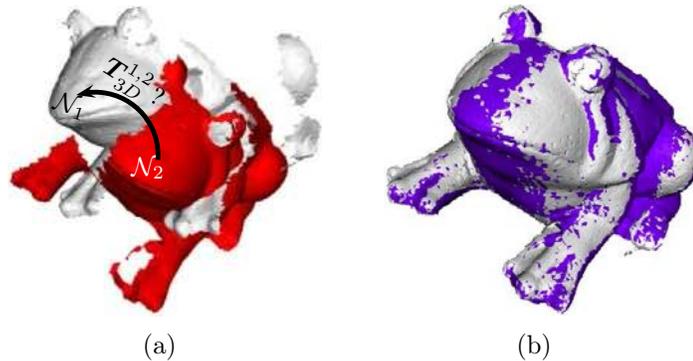


FIGURE. 4.3 – Exemple de modélisation d'un artefact sous forme de grenouille à partir de deux nuages de points 3D : \mathcal{N}_1 (nuage gris) et \mathcal{N}_2 (nuage rouge) ©[Salvi *et al.*, 2007]. (a) Position initiale de \mathcal{N}_1 et \mathcal{N}_2 . (b) Résultat du recalage 3D de \mathcal{N}_2 sur \mathcal{N}_1 en utilisant la méthode de [Besl and McKay, 1992]. L'objet 3D final est alors formé par l'union de \mathcal{N}_1 (nuage gris) et la transformation de \mathcal{N}_2 par $\mathbf{T}_{3D}^{1,2}$ (nuage bleu).

Dans [Salvi *et al.*, 2007], les auteurs proposent une synthèse des méthodes classiques de recalage de données 3D. Parmi les méthodes les plus populaires, citons [Besl and McKay, 1992; Chen and Medioni, 1992; Brett *et al.*, 1999; Chow *et al.*, 2004]. Ces travaux reposent sur l'appariement d'informations 3D homologues entre les acquisitions à recaler pour estimer les transformations 3D.

4.2.2.2 Estimation du mouvement 3D à partir d'images 2D

Lorsque les dispositifs d'acquisition ne fournissent aucune donnée 3D, le mouvement 3D doit être estimé à partir des images 2D. Les techniques appelées « *structure from*

motion » (calcul de structures 3D à partir de mouvements 2D perçus dans les images) sont très utilisées (voir section 2.2.1 du chapitre 2). Leur principe repose sur l'extraction et le suivi de primitives images dans une séquence. Le mouvement 3D de la caméra est donc estimé sous la contrainte imposée par cet appariement de primitives (voir figure 4.4). Les méthodes de référence de « *structure from motion* » sont présentées dans [Hartley and Zisserman, 2003]. Une telle approche a notamment été utilisée dans [Deguchi *et al.*, 1996; Thormahlen *et al.*, 2002] pour la reconstruction 3D de formes anatomiques 3D d'organes comme l'estomac et le côlon. Ces reconstructions ont donc été obtenues à partir de séquences d'images endoscopiques. Toutefois, les images utilisées dans ces deux travaux présentent des primitives qui peuvent être détectées de façon robuste (certaine) et précise, ce qui n'est pas notre cas.

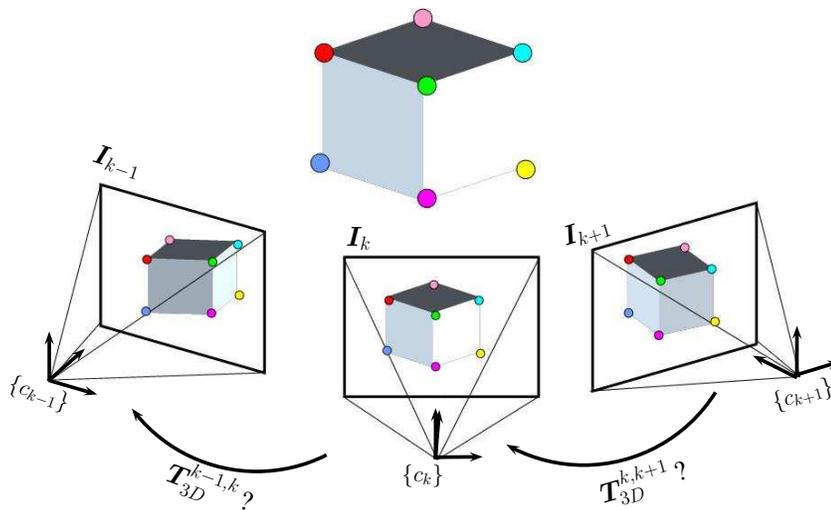


FIGURE. 4.4 – Illustration de l'estimation de mouvement 3D à partir d'images acquises pour différentes prises de vue d'une scène ©[Snavely, 2008]. Des points de référence (les coins du cube) détectés et appariés dans les différentes images permettent l'estimation des transformations 3D.

4.2.3 Discussion et choix effectué

Une méthode classique de recalage 3D requiert une quantité minimale d'informations 3D (*e.g.*, points, courbes, surfaces, *etc.*) [Salvi *et al.*, 2007]. Pour des données cystoscopiques, la difficulté majeure de la cartographie est due à la forme des portions de surfaces imagées à deux instants consécutifs. En effet, celles-ci portent très peu d'informations géométriques pour être recalées (recaler deux plans ou deux bouts de sphères sans textures est impossible puisque l'information géométrique ne permet pas de superposer des surfaces homologues). Par ailleurs, il n'existe pas de points 3D homologues entre deux acquisitions

consécutives. Rappelons que les points laser changent de positions sur la surface imagée quand l'instrument se déplace. D'autre part, les images cystoscopiques ne contiennent pas de primitives images détectables d'une manière suffisamment robuste et systématique pour appliquer les méthodes d'estimation de mouvement 3D à partir d'images 2D seulement (voir section 2.2.3 du chapitre 2).

Compte tenu de ces observations, les techniques classiques de recalage de points 3D ne sont pas adaptées à nos données. L'estimation du mouvement 3D entre deux acquisitions est impossible à réaliser soit en utilisant seulement les points 3D, soit en utilisant uniquement les images 2D. Par conséquent, nous proposons une solution qui combine toutes les données disponibles (images et points 3D/2D). Nous chercherons à estimer la transformation 3D qui conduit au meilleur recalage de images 2D consécutives. Bien qu'il n'y ait pas de points 3D homologues entre deux acquisitions, il existe des régions homologues dans deux images consécutives, car ces images se recouvrent spatialement [Miranda-Luna *et al.*, 2008; Hernández-Mier *et al.*, 2010].

4.3 Estimation du mouvement 3D guidée par recalage 2D/2D d'images

Dans cette section, nous présentons une nouvelle méthode d'estimation du mouvement 3D (calcul de $\mathbf{T}_{3D}^{k-1,k}$) guidée par le recalage des images (estimation de $\mathbf{T}_{2D}^{k-1,k}$) obtenues pour deux acquisitions consécutives. Notre méthode repose sur un système d'équations linéaires qui relie les transformations $\mathbf{T}_{3D}^{k-1,k}$ et $\mathbf{T}_{2D}^{k-1,k}$. Nous présentons d'abord le principe général de la méthode. Puis, nous détaillons le fondement mathématique sur lequel elle repose. Finalement, nous présentons l'algorithme que nous avons développé.

4.3.1 Principe général de la méthode

Le lien mathématique entre la transformation 3D rigide $\mathbf{T}_{3D}^{k-1,k}$ et la transformation 2D perspective $\mathbf{T}_{2D}^{k-1,k}$ peut être établi de par la connaissance des points $P_{3D}^{i,k}$ et $P_{2D}^{i,k}$. Nous exploiterons ce lien afin de poser le problème de l'estimation de la transformation $\mathbf{T}_{3D}^{k-1,k}$ comme un problème de recalage d'images 2D (où les inconnues sont les paramètres du mouvement 3D). Le principe général de la méthode est de remplacer un problème d'optimisation dont les paramètres sont ceux de la matrice $\mathbf{T}_{2D}^{k-1,k}$ pour le recalage d'images 2D (*e.g.*, maximisation de l'information mutuelle) par un nouveau problème d'optimisation dont les paramètres sont ceux de $\mathbf{T}_{3D}^{k-1,k}$ en effectuant un « changement de variables » (quand $\mathbf{T}_{3D}^{k-1,k}$ est connue, on sait en déduire la transformation perspective $\mathbf{T}_{2D}^{k-1,k}$ associée

de par la connaissance des points $P_{3D}^{i,k}$ et $P_{2D}^{i,k}$).

Cela nous conduira à estimer « simultanément » les transformations 3D et les transformations perspectives puisque les dernières se déduisent des premières.

4.3.2 Fondement mathématique de la méthode

4.3.2.1 Relation mathématique entre les transformations $\mathbf{T}_{3D}^{k-1,k}$ et $\mathbf{T}_{2D}^{k-1,k}$

Avant tout, rappelons que pour notre prototype expérimental, le projecteur et la caméra sont fixés l'un à l'autre d'une manière rigide. Quand l'instrument se déplace, les points laser projetés changent leurs emplacements sur la surface imagée. Il n'existe donc pas de points 3D homologues entre les acquisitions. Ainsi, la transformation $\mathbf{T}_{3D}^{k-1,k}$ place les points $P_{3D}^{i,k}$ dans le repère $\{c_{k-1}\}$ mais ne superpose pas ces points $P_{3D}^{i,k}$ sur les points $P_{3D}^{i,k-1}$. Cela se traduit par :

$$\begin{bmatrix} \widehat{P}_{3D}^{i,k-1} \\ 1 \end{bmatrix} = \begin{bmatrix} \widehat{x}_{3D}^{i,k-1} \\ \widehat{y}_{3D}^{i,k-1} \\ \widehat{z}_{3D}^{i,k-1} \\ 1 \end{bmatrix} = \mathbf{T}_{3D}^{k-1,k} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \quad (4.9)$$

La projection du point $\widehat{P}_{3D}^{i,k-1}$ dans le plan image est donnée par :

$$\widehat{z}_{3D}^{i,k-1} \begin{bmatrix} \widehat{P}_{2D}^{i,k-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{K} & 0 \\ 0 \end{bmatrix} \begin{bmatrix} \widehat{P}_{3D}^{i,k-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{K} & 0 \\ 0 \end{bmatrix} \mathbf{T}_{3D}^{k-1,k} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \quad (4.10)$$

Rappelons que la relation entre deux images consécutives \mathbf{I}_k et \mathbf{I}_{k-1} est modélisée par une transformation perspective $\mathbf{T}_{2D}^{k-1,k}$. Cette transformation ramène les pixels de l'image \mathbf{I}_k sur leurs homologues dans l'image \mathbf{I}_{k-1} . Elle déplace les points $P_{2D}^{i,k}$ (projetés des points $P_{3D}^{i,k}$ dans l'image \mathbf{I}_k) dans le repère $\{I_{k-1}\}$:

$$\beta^{i,k-1} \begin{bmatrix} \widehat{P}_{2D}^{i,k-1} \\ 1 \end{bmatrix} = \mathbf{T}_{2D}^{k-1,k} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} \quad (4.11)$$

Les points $P_{3D}^{i,k}$ et $P_{2D}^{i,k}$ font référence au même emplacement sur la surface imagée. Leurs transformées dans les repères de l'acquisition $k-1$, respectivement $\widehat{P}_{3D}^{i,k}$ et $\widehat{P}_{2D}^{i,k}$, font aussi référence au même emplacement sur la surface imagée. Par conséquent, en combinant les équations (4.10) et (4.11), on obtient une relation qui relie les transformations $\mathbf{T}_{2D}^{k-1,k}$ et

$\mathbf{T}_{3D}^{k-1,k}$:

$$\frac{1}{\beta^{i,k-1}} \mathbf{T}_{2D}^{k-1,k} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} = \frac{1}{\hat{z}_{3D}^{i,k-1}} \begin{bmatrix} 0 \\ \mathcal{K} & 0 \\ 0 \end{bmatrix} \mathbf{T}_{3D}^{k-1,k} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \quad (4.12)$$

Nous remarquons que :

- La matrice caméra \mathcal{K} est fixe car elle est indépendante du point de vue.
- Le projeté $P_{2D}^{i,k}$ d'un point $P_{3D}^{i,k}$ est considéré comme connu : $P_{2D}^{i,k}$ a été segmenté dans l'image \mathbf{I}_k
- À un point $P_{2D}^{i,k}$ donné et pour une transformation perspective $\mathbf{T}_{2D}^{k-1,k}$ donnée, ne peut correspondre qu'une seule valeur de $\beta^{i,k-1}$ (voir l'équation (4.6)).
- Pour un point $P_{3D}^{i,k}$ donné, la valeur de $\hat{z}_{3D}^{i,k-1}$ (troisième coordonnée de $\hat{P}_{3D}^{i,k-1}$) est inconnue mais elle se déduit directement de $\mathbf{T}_{3D}^{k-1,k}$ d'après (4.5).

En résumé, pour une acquisition k et un nuage de points $\{P_{3D}^{i,k}, i = 1, \dots, N\}$ mesurés, seules la transformation $\mathbf{T}_{3D}^{k-1,k}$ et la transformation perspective $\mathbf{T}_{2D}^{k-1,k}$ sont inconnues. L'équation (4.12) est satisfaite par les 8 paires de points $(P_{3D}^{i,k}, P_{2D}^{i,k})$.

4.3.2.2 Calcul de la transformation perspective \mathbf{T}_{2D} associée à une transformation \mathbf{T}_{3D}

Nous présentons dans cette section le détail du calcul d'une transformation perspective \mathbf{T}_{2D} à partir une transformation \mathbf{T}_{3D} dans l'équation (4.12). Par souci de simplicité, nous omettons les indices $k-1$ et k lorsqu'il n'y a pas d'ambiguïté. Les développements qui suivent nous conduisent à un système d'équations linéaires dont les inconnues sont les éléments a_{mn} de la matrice \mathbf{T}_{2D} .

Réécrivons l'équation (4.12) pour la $i^{\text{ème}}$ paire de points $(P_{3D}^{i,k}, P_{2D}^{i,k})$. Nous obtenons l'équation (4.13) où les quantités au numérateur et au dénominateur sont toutes des scalaires (les indices 1, 2, et 3, sont utilisés pour désigner la 1^{ère}, la 2^{ème} et la 3^{ème} ligne d'une matrice).

$$\left\{ \begin{array}{l} \frac{\mathbf{T}_{2D,1.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix}}{\mathbf{T}_{2D,3.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix}} = \frac{[\mathcal{K}_{1.,0}] \mathbf{T}_{3D} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix}}{\mathbf{T}_{3D,3.} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix}} \\ \\ \frac{\mathbf{T}_{2D,2.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix}}{\mathbf{T}_{2D,3.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix}} = \frac{[\mathcal{K}_{2.,0}] \mathbf{T}_{3D} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix}}{\mathbf{T}_{3D,3.} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix}} \end{array} \right. \quad (4.13)$$

Pour simplifier ces expressions, on désigne par $\mathcal{C}_{1,i}^{k-1,k}$ et $\mathcal{C}_{2,i}^{k-1,k}$ les termes de droite dans le système (4.13) :

$$\mathcal{C}_{1,i}^{k-1,k} = \left([\mathcal{K}_{1.,0}] \mathbf{T}_{3D} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \right) / \left(\mathbf{T}_{3D,3.} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \right) \quad (4.14)$$

$$\mathcal{C}_{2,i}^{k-1,k} = \left([\mathcal{K}_{2.,0}] \mathbf{T}_{3D} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \right) / \left(\mathbf{T}_{3D,3.} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \right) \quad (4.15)$$

Le système (4.13) devient :

$$\left\{ \begin{array}{l} \mathbf{T}_{2D,1.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} = \mathcal{C}_{1,i}^{k-1,k} \mathbf{T}_{2D,3.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} \\ \\ \mathbf{T}_{2D,2.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} = \mathcal{C}_{2,i}^{k-1,k} \mathbf{T}_{2D,3.} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} \end{array} \right. \quad (4.16)$$

On obtient finalement un système de deux équations linéaires par rapport aux paramètres a_{mn} de \mathbf{T}_{2D} :

$$\left\{ \begin{array}{l} x_{2D}^{i,k} a_{11} + y_{2D}^{i,k} a_{12} + a_{13} - \mathcal{C}_{1,i}^{k-1,k} x_{2D}^{i,k} a_{31} - \mathcal{C}_{1,i}^{k-1,k} y_{2D}^{i,k} a_{32} = \mathcal{C}_{1,i}^{k-1,k} \\ \\ x_{2D}^{i,k} a_{21} + y_{2D}^{i,k} a_{22} + a_{23} - \mathcal{C}_{2,i}^{k-1,k} x_{2D}^{i,k} a_{31} - \mathcal{C}_{2,i}^{k-1,k} y_{2D}^{i,k} a_{32} = \mathcal{C}_{2,i}^{k-1,k} \end{array} \right. \quad (4.17)$$

en remarquant que $a_{33} = 1$. Les équations (4.17) sont satisfaites pour tous les points $P_{2D}^{i,k}$, $i \in \{1, \dots, N\}$ à l'instant k . Nous obtenons donc un système de $2N$ équations à N

inconnues que nous récrivons sous forme matricielle :

$$\begin{bmatrix}
 x_{2D}^{1,k} & y_{2D}^{1,k} & 1 & 0 & 0 & 0 & -\mathcal{C}_{1,1}^{k-1,k} x_{2D}^{1,k} & -\mathcal{C}_{1,1}^{k-1,k} y_{2D}^{1,k} \\
 0 & 0 & 0 & x_{2D}^{1,k} & y_{2D}^{1,k} & 1 & -\mathcal{C}_{2,1}^{k-1,k} x_{2D}^{1,k} & -\mathcal{C}_{2,1}^{k-1,k} y_{2D}^{1,k} \\
 \vdots & \vdots \\
 x_{2D}^{N,k} & y_{2D}^{N,k} & 1 & 0 & 0 & 0 & -\mathcal{C}_{1,N}^{k-1,k} x_{2D}^{N,k} & -\mathcal{C}_{1,N}^{k-1,k} y_{2D}^{N,k} \\
 0 & 0 & 0 & x_{2D}^{N,k} & y_{2D}^{N,k} & 1 & -\mathcal{C}_{2,N}^{k-1,k} x_{2D}^{N,k} & -\mathcal{C}_{2,N}^{k-1,k} y_{2D}^{N,k}
 \end{bmatrix}
 \begin{bmatrix}
 a_{11} \\
 a_{12} \\
 a_{13} \\
 a_{21} \\
 a_{22} \\
 a_{23} \\
 a_{31} \\
 a_{32}
 \end{bmatrix}
 =
 \begin{bmatrix}
 \mathcal{C}_{1,1}^{k-1,k} \\
 \mathcal{C}_{2,1}^{k-1,k} \\
 \vdots \\
 \mathcal{C}_{1,N}^{k-1,k} \\
 \mathcal{C}_{2,N}^{k-1,k}
 \end{bmatrix}
 \quad (4.18)$$

Pour une transformation \mathbf{T}_{3D} donnée et pour $N = 8$, on résout donc un système surdéterminé de 16 équations à 8 inconnues. Il est à noter que pour $N \geq 4$, une transformation \mathbf{T}_{3D} dans (4.18) ne peut conduire qu'à une seule transformation \mathbf{T}_{2D} .

4.3.3 Algorithme développé

L'estimation de la transformation $\mathbf{T}_{3D}^{k-1,k}$ est formulée comme l'optimisation d'un critère de similarité entre deux images 2D. Il s'agit de calculer les paramètres de $\mathbf{T}_{3D}^{k-1,k}$ conduisant à une transformation perspective $\mathbf{T}_{2D}^{k-1,k}$ qui superpose au mieux l'image \mathbf{I}_k sur l'image \mathbf{I}_{k-1} .

Une matrice $\mathbf{T}_{3D}^{k-1,k}$ a six degrés de liberté, trois angles de rotation (θ , ϕ , ψ) et trois translation (t_x , t_y , t_z) (voir équation (4.4)). Nous choisissons naturellement d'estimer ces six paramètres plutôt que les 12 éléments inconnus dans la matrice $\mathbf{T}_{3D}^{k-1,k}$ (matrice de dimension 4×4 dont la dernière ligne est égale à $[0, 0, 0, 1]$) afin de réduire la dimension de l'espace de recherche à 6. Formellement, le problème d'optimisation s'écrit :

$$\begin{aligned}
 \tilde{\mathcal{V}} &= \arg \min_{\mathcal{V} \in \mathbb{R}^6} \mathcal{F}(\mathcal{V}) \\
 \mathcal{F} &: \mathbb{R}^6 \longmapsto \mathbb{R}
 \end{aligned}
 \quad (4.19)$$

où $\mathcal{V} = [\theta, \phi, \psi, t_x, t_y, t_z]^T$ est un vecteur de dimension 6 et \mathcal{F} est une fonction coût qui indique la dissimilarité entre les images \mathbf{I}_{k-1} et $\hat{\mathbf{I}}_{k-1} = \mathbf{T}_{2D}(\mathbf{I}_k)$, résultat de la transformation de l'image \mathbf{I}_k .

Nous utilisons un algorithme d'optimisation local itératif pour résoudre le problème (4.19). Les principaux traitements à effectuer par itération sont les suivants.

- Calculer les 12 paramètres de la transformation \mathbf{T}_{3D} à partir du vecteur \mathcal{V} à évaluer en utilisant l'équation (4.4).

- Calculer la transformation perspective \mathbf{T}_{2D} à partir des paramètres \mathcal{V} , des N paires de points $(P_{3D}^{i,k}, P_{2D}^{i,k})$ et de la matrice caméra \mathcal{K} en inversant le système (4.18).
- Transformer l’image \mathbf{I}_k par la matrice \mathbf{T}_{2D} pour obtenir l’image transformée $\widehat{\mathbf{I}}_{k-1}$.
- Calculer la fonction coût $\mathcal{F}(\mathcal{V})$, c’est-à-dire la mesure de similarité entre les images $\widehat{\mathbf{I}}_{k-1}$ et image \mathbf{I}_{k-1} .
- Mettre à jour les paramètres de \mathcal{V} et réitérer.

Dans la suite, nous précisons les choix des éléments essentiels de notre algorithme, à savoir la fonction coût \mathcal{F} , la méthode de minimisation de \mathcal{F} , l’initialisation et les critères d’arrêt.

4.3.3.1 Choix de la fonction coût \mathcal{F}

Dans la section 2.1.3 du chapitre 2 nous avons introduit deux mesures de similarité adaptées aux images cystoscopiques : la distance quadratique entre les niveaux de gris (\mathcal{S}_{DQ}) et l’information mutuelle (\mathcal{S}_{IM}). Nos images contiennent, en plus de la texture de la paroi de la vessie, des spots laser. Il est important d’exclure les pixels appartenant aux spots laser du calcul des mesures de similarité. En effet, les spots laser ne sont pas des données homologues (ils se déplacent sur la surface entre deux acquisitions) et ne doivent donc pas contribuer dans l’estimation de la similarité des images. Pour surmonter ce problème, nous avons utilisé le résultat de la segmentation des spots laser pour construire à chaque acquisition une image binaire utilisée pour masquer les pixels qui ne doivent pas être considérés dans les calculs (voir figure 4.5). Dans un souci de simplification, nous n’introduisons pas cette notion « d’image masque » dans les explications du calcul des mesures de similarité⁹.

Nous avons implémenté les deux mesures \mathcal{S}_{DQ} et \mathcal{S}_{IM} afin de tester leur impact sur notre algorithme d’estimation de $\mathbf{T}_{3D}^{k-1,k}$. Dans la suite, nous présentons le calcul de la fonction coût \mathcal{F} à partir de ces deux mesures de similarité. Dans tous les cas, une interpolation bilinéaire est utilisée pour calculer l’intensité des pixels de l’image $\widehat{\mathbf{I}}_{k-1}$ (voir [Zitová and Flusser, 2003; Hernandez-Mier, 2007] pour plus de détails).

Mesure de similarité \mathcal{S}_{DQ} . Elle désigne la distance quadratique entre les niveaux de gris des pixels homologues dans les images \mathbf{I}_{k-1} et $\widehat{\mathbf{I}}_{k-1} = \mathbf{T}_{2D}^{k-1,k}(\mathbf{I}_k)$:

$$\mathcal{S}_{DQ}(\mathbf{I}_{k-1}, \widehat{\mathbf{I}}_{k-1}(\mathcal{V})) = \sum_{\mathbf{I}_{k-1} \cap \widehat{\mathbf{I}}_{k-1}} \|\mathbf{I}_{k-1} - \widehat{\mathbf{I}}_{k-1}\|^2. \quad (4.20)$$

9. Le rejet des points laser est un problème d’implémentation et ne contribue pas à la compréhension des algorithmes de calcul des mesures de similarité

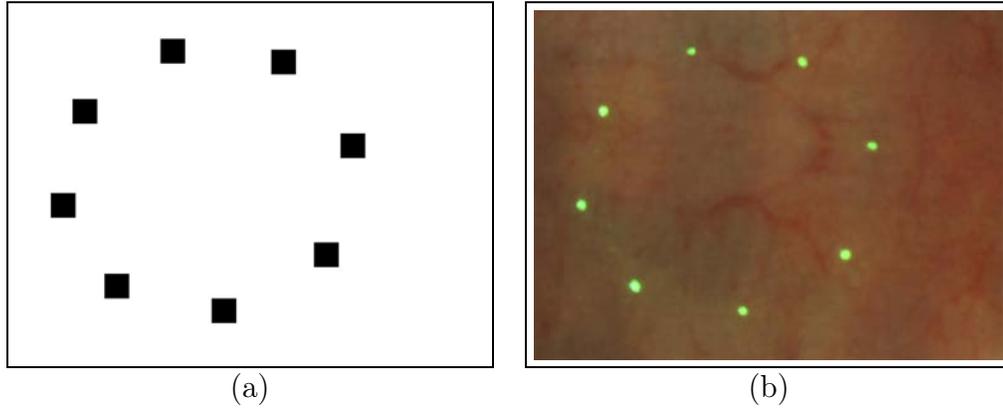


FIGURE. 4.5 – (a) Image binaire construite pour masquer les pixels qui ne doivent pas être pris en compte dans l'image (b) lors du calcul des mesures de similarité. Les correspondants de ces pixels dans l'image (a) ont une valeur de 0 alors que le reste des pixels (à prendre en compte) ont une valeur de 1.

$\mathcal{S}_{\text{DQ}}(\mathbf{I}_{k-1}, \widehat{\mathbf{I}}_{k-1}(\mathcal{V}))$ tend vers 0 lorsque les zones communes (homologues) à \mathbf{I}_{k-1} et $\widehat{\mathbf{I}}_{k-1}$ sont parfaitement superposées. La fonction coût \mathcal{F} est définie à partir de $\mathcal{S}_{\text{DQ}}(\mathbf{I}_{k-1}, \widehat{\mathbf{I}}_{k-1})$ et du nombre total de pixels $nbPix$ utilisés dans le calcul (4.20) :

$$\mathcal{F}(\mathcal{V}) = \frac{1}{nbPix} \mathcal{S}_{\text{DQ}}(\mathbf{I}_{k-1}, \widehat{\mathbf{I}}_{k-1}(\mathcal{V})) \quad (4.21)$$

Mesure de similarité \mathcal{S}_{IM} . Cette mesure a été utilisée dans les travaux de [Miranda-Luna *et al.*, 2008] pour le recalage d'images cystoscopiques. Elle indique le degré de superposition d'une image \mathbf{I}_k (image source) sur l'image \mathbf{I}_{k-1} (image cible) à l'aide de la transformation perspective \mathbf{T}_{2D} . Le calcul de l'information mutuelle de deux images \mathbf{I}_{k-1} et $\widehat{\mathbf{I}}_{k-1}$ (transformée de l'image \mathbf{I}_k par \mathbf{T}_{2D}) est basé sur la densité conjointe de probabilité des niveaux de gris des images [Collignon *et al.*, 1995; Miranda-Luna *et al.*, 2008]. L'expression de \mathcal{S}_{IM} est formulée dans l'équation (4.22).

$$\mathcal{S}_{\text{IM}}(\mathbf{I}_{k-1}, \widehat{\mathbf{I}}_{k-1}(\mathcal{V})) = \sum_{z=z_{\min}}^{z_{\max}} \sum_{z'=z_{\min}}^{z_{\max}} p_{k-1,k}(z, z') \ln \left(\frac{p_{k-1,k}(z, z')}{p_{k-1}(z) p_k(z')} \right) \quad (4.22)$$

où :

- z et z' désignent les niveaux de gris des pixels des images \mathbf{I}_{k-1} et $\widehat{\mathbf{I}}_{k-1}$,
- $z_{\min} = 0$ et $z_{\max} = 255$ sont la plus petite et la plus grande valeur des niveaux de gris.
- $p_{k-1}(z)$, $p_k(z')$ sont respectivement les fonctions de densité de probabilité des niveaux de gris de \mathbf{I}_{k-1} , de $\widehat{\mathbf{I}}_{k-1}$. $p_{k-1,k}(z, z')$ est la densité de probabilité jointe. Ces densités

de probabilité sont définies directement par les histogrammes de $\widehat{\mathbf{I}}_{k-1}$ et \mathbf{I}_{k-1} et par l'histogramme conjoint normalisé [Collignon *et al.*, 1995].

L'information mutuelle est maximale lorsque les images sont recalées. La fonction coût \mathcal{F} à minimiser est finalement définie par :

$$\mathcal{F}(\mathcal{V}) = -\mathcal{S}_{\text{IM}}(\mathbf{I}_{k-1}, \widehat{\mathbf{I}}_{k-1}(\mathcal{V})). \quad (4.23)$$

\mathcal{F} est d'autant plus faible que les images \mathbf{I}_{k-1} et $\widehat{\mathbf{I}}_{k-1}$ sont bien recalées.

4.3.3.2 Minimisation de la fonction coût \mathcal{F}

La fonction coût \mathcal{F} n'est pas analytiquement dérivable. Nous choisissons donc la méthode du *simplexe* de Nelder-Mead [Press *et al.*, 2007], une méthode de référence pour l'optimisation sans dérivation. Elle se base uniquement sur l'évaluation de la fonction coût (de 1 à 2 évaluations par itération).

Dans le cas de la minimisation de $\mathcal{F}(\mathcal{V})$, la méthode de Nelder-Mead consiste à utiliser un simplexe composé de 7 sommets notés \mathcal{V}_j . Les sommets \mathcal{V}_j sont définis dans \mathbb{R}^6 avec $j \in \{1, \dots, 7\}$. Le simplexe se déforme d'une manière itérative pour converger vers un minimum local de \mathcal{F} . Pour plus de détails sur cet algorithme, nous renvoyons le lecteur à l'annexe A.

Initialisation du simplexe. Rappelons que lors d'un examen cystoscopique, le clinicien déplace l'instrument relativement lentement (quelques millimètres par seconde) au regard de la vitesse élevée d'acquisition (25 acquisitions par seconde). Le mouvement 3D entre deux acquisitions consécutives est donc petit. Le simplexe initial est construit en tenant compte de cette observation. Nous supposons qu'entre deux acquisitions consécutives, les rotations ne dépassent pas 2° et les translations ne dépassent pas 0.6 mm. Ces seuils de rotation et de translation ont été choisis car ils conduisent à des déplacements 2D (dans les images) largement supérieurs à ceux qu'on trouve dans des séquences cystoscopiques cliniques. Les sommets du simplexe initial sont construits de la façon suivante à partir

des bornes de rotations et de translation 3D que nous avons choisies :

$$\begin{aligned}
\mathcal{V}_1 &= [0, 0, 0, 0, 0, 0] \\
\mathcal{V}_2 &= [2, 0, 0, 0, 0, 0] \\
\mathcal{V}_3 &= [0, 2, 0, 0, 0, 0] \\
\mathcal{V}_4 &= [0, 0, 2, 0, 0, 0] \\
\mathcal{V}_5 &= [0, 0, 0, 0.6, 0, 0] \\
\mathcal{V}_6 &= [0, 0, 0, 0, 0.6, 0] \\
\mathcal{V}_7 &= [0, 0, 0, 0, 0, 0.6]
\end{aligned}$$

Cette façon de construire le simplexe assure sa convexité comme la méthode de Nelder-Mead le requiert.

Critère d'arrêt de l'algorithme. Dans la littérature, le critère d'arrêt porte généralement sur la taille du simplexe [Press *et al.*, 2007] : lorsqu'elle est inférieure à un seuil pré-défini, l'algorithme s'arrête. Dans notre cas, les paramètres de la transformation 3D à retrouver doivent conduire à une superposition précise des deux images \mathbf{I}_{k-1} et \mathbf{I}_k . Nous avons utilisé un autre critère d'arrêt adapté à la spécificité de notre problème. Nous nous sommes inspirés du critère d'arrêt de l'algorithme de recalage d'images cystoscopiques proposé dans [Miranda-Luna *et al.*, 2008]. Dans ce dernier, l'algorithme d'optimisation s'arrête lorsque le changement des paramètres de la transformation perspective (l'itérée) d'une itération à la suivante conduit à un déplacement 2D (dans l'image) moyen inférieur à 1 pixel. Ce seuil est suffisamment petit pour assurer la cohérence visuelle d'une mosaïque 2D construite avec la méthode de [Miranda-Luna *et al.*, 2008]. Cependant pour notre méthode nous avons vérifié expérimentalement que ce seuil doit être réduit à 0.1 pixel pour assurer une estimation précise des paramètres de la transformation 3D.

Solution finale. Au final, la solution obtenue, notée $\tilde{\mathcal{V}}$, est utilisée pour construire les deux transformations $\tilde{\mathbf{T}}_{3D}^{k-1,k}$ avec l'équation (4.4) et $\tilde{\mathbf{T}}_{2D}^{k-1,k}$ en résolvant le système (4.18). Ces deux transformations désignent l'estimation des transformations réelles $\mathbf{T}_{3D}^{k-1,k}$ et $\mathbf{T}_{2D}^{k-1,k}$ à retrouver.

4.4 Évaluation de la méthode proposée

La méthode d'estimation des matrices $\mathbf{T}_{3D}^{k-1,k}$ guidée par recalage d'images est la partie centrale du schéma général de la construction de surfaces étendues de la paroi interne de la vessie puisqu'elle est répétée pour chaque acquisition $k = 2, \dots, K$. La qualité visuelle de la carte générée (cohérence visuelle de l'image panoramique et de la surface qui interpole le nuage de points 3D) dépend essentiellement de la précision de l'estimation des transformations $\mathbf{T}_{3D}^{k-1,k}$ et $\mathbf{T}_{2D}^{k-1,k}$ entre chaque paire d'acquisitions consécutives.

Dans cette section, nous évaluons la méthode d'estimation de ces transformations sur des données simulées à partir de fantômes numériques. Nous commençons par étudier la précision de la méthode. Ensuite, nous évaluons sa robustesse vis-à-vis d'une part, de la variabilité des textures dans les images cystoscopiques et, d'autre part, de la vitesse de mouvement pendant l'acquisition d'une séquence de données cystoscopiques (c'est-à-dire qu'on détermine les mouvements limites autorisant une estimation précise des transformations 3D rigide et 2D perspective).

Dans le présent chapitre, nous testons donc uniquement la capacité de notre algorithme à recalculer les données de deux acquisitions consécutives. Ces tests sont basés sur des données simulées pour évaluer uniquement la précision et la robustesse maximales de l'algorithme. Les données simulées sont « idéales », c'est-à-dire qu'elles sont simulées en appliquant les véritables transformations géométriques (inconnues) aux images \mathbf{I}_k et aux points $P_{3D}^{i,k}$. Le chapitre 5 sera dédié aux tests effectués avec des données réelles acquises avec des fantômes réalistes. Nous testerons non seulement la capacité de l'algorithme à bien recalculer deux points de vue consécutifs mais aussi la qualité de la cartographie globale (construction de cartes 2D et placement de tous les points 3D dans un repère global).

4.4.1 Expérimentations

Les données utilisées dans ces expérimentations sont simulées à partir d'un fantôme numérique constitué d'un modèle 3D générique de vessie (voir figure 4.6(a)). Ce modèle a été construit à partir des travaux réalisés au CRAN par Yannick Rémy [Rémy, 2006] avec l'assistance d'un urologue. La surface interne de ce modèle de vessie est tapissée avec des images 2D cystoscopiques réalistes.

Pour l'étude de précision nous avons réalisé une photographie haute résolution d'une vessie de porc qui a été incisée et mise à plat (voir figure 4.6(b)). La texture de la paroi interne de la vessie de porc est visuellement très proche d'une vessie humaine.

Pour l'étude de robustesse vis-à-vis de la vitesse de déplacement de l'instrument et de la variabilité des textures observables dans les séquences cystoscopiques, la texture

plaquée sur le modèle 3D de vessie provient d'images cystoscopiques cliniques présentant des différences très marquées au niveau visuel (voir figure 4.7).

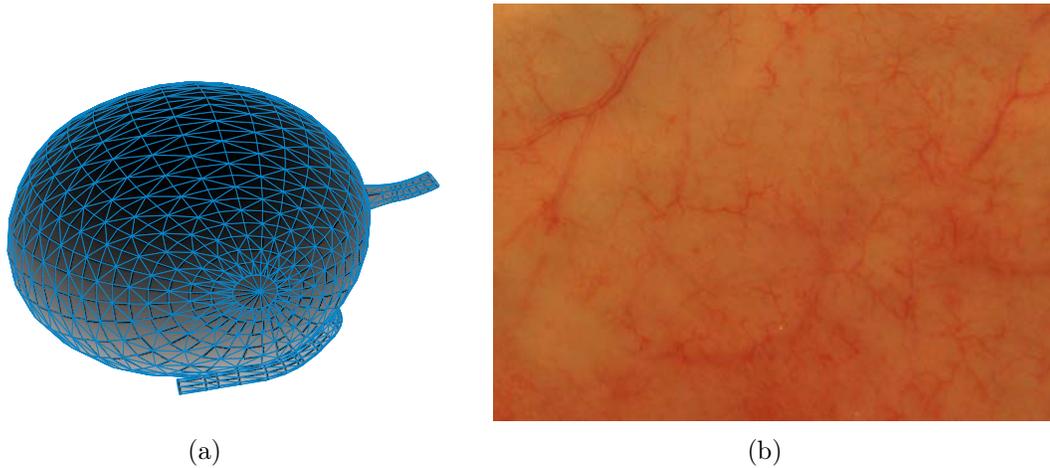


FIGURE. 4.6 – (a) Forme 3D du fantôme numérique utilisé pour simuler des données cystoscopiques 3D/2D ©[Rémy, 2006]. (b) Photographie haute résolution (4288×2848 pixels) de la paroi interne d'une vessie de porc incisée et ouverte à plat.

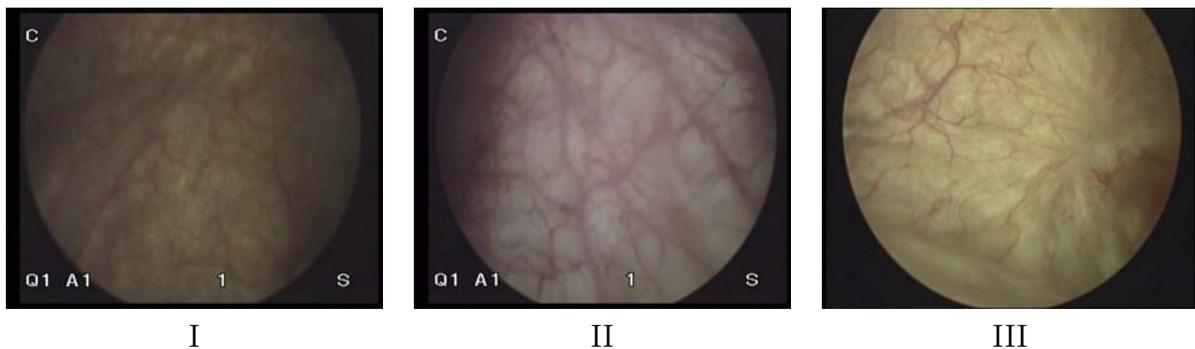


FIGURE. 4.7 – I, II et III : Images extraites de trois séquences cystoscopiques cliniques. Les images choisies présentent une forte variabilité de texture, de contraste et d'intensité lumineuse. Cette variabilité est due à la fois à la scène (variabilité inter-patients) et à l'instrument (variabilité engendrée par les cystoscopes).

4.4.1.1 Simulation des données 3D/2D.

Nous avons développé un outil qui permet de simuler l'acquisition de données 3D/2D cystoscopiques à partir du fantôme numérique. Cet « instrument virtuel d'acquisition » simule le fonctionnement des modèles caméra et la projection des huit points, détaillés dans le chapitre 3. Pour une position $k - 1$, supposée connue de l'instrument virtuel à l'intérieur du fantôme, le modèle caméra permet de construire l'image \mathbf{I}_{k-1} de taille

500 × 500 pixels (voir section 2.1.1 du chapitre 2). L'intensité des pixels de l'image est interpolée à partir de la surface texturée se trouvant dans le champ de vue de l'instrument virtuel. Pour la même pose $k - 1$, les coordonnées des huit points laser $P_{3D}^{i,k-1}$ situés sur la surface sont calculées avec le modèle projecteur (équations de 8 droites 3D dans le repère de la caméra). Les coordonnées des points $P_{2D}^{i,k-1}$ sont calculées en projetant les points $P_{3D}^{i,k-1}$ dans le plan image avec le modèle caméra.

Pour simuler la nouvelle position de l'endoscope pour l'acquisition suivante k , une transformation $\mathbf{T}_{3D}^{k,k-1}$ connue est appliquée à l'instrument virtuel, ce qui permet de simuler les $N = 8$ droites 3D et le repère caméra $\{c_k\}$. L'image \mathbf{I}_k et les points lasers $P_{3D}^{i,k-1}$ et $P_{2D}^{i,k-1}$ sont finalement générés selon la procédure décrite au paragraphe ci-dessus.

De plus, on peut calculer la transformation perspective $\mathbf{T}_{2D}^{k-1,k}$ qui relie l'image \mathbf{I}_k à l'image \mathbf{I}_{k-1} à partir de $\mathbf{T}_{3D}^{k,k-1}$, mais ce calcul n'est pas nécessaire à la simulation des données. En résumé, pour la simulation d'une paire d'acquisitions $k - 1$ et k , nous obtenons les données et les transformations suivantes :

- deux images \mathbf{I}_{k-1} et \mathbf{I}_k ,
- deux ensembles de 8 paires de points $(P_{3D}^{i,k-1}, P_{2D}^{i,k-1})$ et $(P_{3D}^{i,k}, P_{2D}^{i,k})$,
- la transformation $\mathbf{T}_{3D}^{k-1,k} = (\mathbf{T}_{3D}^{k,k-1})^{-1}$ superposant le repère $\{c_k\}$ sur le repère $\{c_{k-1}\}$,
- la transformation perspective $\mathbf{T}_{2D}^{k-1,k}$ superposant l'image \mathbf{I}_k sur l'image \mathbf{I}_{k-1} .

4.4.1.2 Application de l'algorithme sur données simulées

Notre méthode est ensuite appliquée aux données 3D/2D simulées pour estimer les transformations 3D rigide et 2D perspective reliant chaque paire d'acquisitions consécutives. Ces transformations estimées (notées respectivement $\tilde{\mathbf{T}}_{3D}^{k-1,k}$ et $\tilde{\mathbf{T}}_{2D}^{k-1,k}$) sont finalement comparées aux transformations théoriques $\mathbf{T}_{3D}^{k-1,k}$ et $\mathbf{T}_{2D}^{k-1,k}$ (connues et fournies par la simulation des données).

4.4.1.3 Critères quantitatifs de l'évaluation

La comparaison des transformations estimées et théoriques est effectuée « indirectement » par le calcul de deux mesures d'erreurs. La première, notée $\epsilon_{3D}^{k-1,k}$, quantifie l'erreur de positionnement des points $P_{3D}^{i,k}$ dans le repère $\{c_{k-1}\}$ avec la transformation $\tilde{\mathbf{T}}_{3D}^{k-1,k}$. Pour ce faire, chaque point $P_{3D}^{i,k}$ est transformé par $\mathbf{T}_{3D}^{k-1,k}$ (transformation théorique) pour générer la position $\hat{P}_{3D}^{i,k-1}$ théorique (idéale) du point. Ensuite, la transformation $\tilde{\mathbf{T}}_{3D}^{k-1,k}$ est appliquée au même point $P_{3D}^{i,k}$ pour trouver sa position estimée $\tilde{P}_{3D}^{i,k-1}$ dans le repère $\{c_{k-1}\}$. La mesure d'erreur $\epsilon_{3D}^{k-1,k}$ est définie par la moyenne des distances euclidiennes

$d(\widehat{P}_{3D}^{i,k-1}, \widetilde{P}_{3D}^{i,k-1})$ calculées entre les positions 3D théoriques et estimées de l'ensemble des $N = 8$ points 3D :

$$\epsilon_{3D}^{k-1,k} = \frac{1}{N} \sum_{i=1}^N d(\widehat{P}_{3D}^{i,k-1}, \widetilde{P}_{3D}^{i,k-1}) \quad (4.24)$$

La deuxième mesure $\epsilon_{2D}^{k-1,k}$ quantifie l'erreur de recalage des images \mathbf{I}_k et \mathbf{I}_{k-1} . Par analogie à la mesure d'erreur $\epsilon_{3D}^{k-1,k}$, les points $P_{2D}^{i,k}$ sont ramenés dans le repère de l'image \mathbf{I}_{k-1} en utilisant la transformation perspective théorique $\mathbf{T}_{2D}^{k-1,k}$ pour obtenir leurs positions théoriques $\widehat{P}_{2D}^{i,k-1}$ (voir équation (4.6)). Ensuite, la transformation $\widetilde{\mathbf{T}}_{2D}^{k-1,k}$ est appliquée aux points $P_{2D}^{i,k}$ pour trouver leurs positions estimées $\widetilde{P}_{2D}^{i,k-1}$ dans le repère $\{\mathbf{I}_{k-1}\}$. La mesure d'erreur $\epsilon_{2D}^{k-1,k}$ est définie comme la moyenne des distances euclidiennes calculées entre les positions 2D théoriques et estimées pour tous les $N = 8$ points 2D :

$$\epsilon_{2D}^{k-1,k} = \frac{1}{N} \sum_{i=1}^N d(\widehat{P}_{2D}^{i,k-1}, \widetilde{P}_{2D}^{i,k-1}) \quad (4.25)$$

Dans le cas idéal, c'est-à-dire lorsque les transformations estimées tendent vers les transformations théoriques, les deux mesures d'erreurs $\epsilon_{3D}^{k-1,k}$ et $\epsilon_{2D}^{k-1,k}$ tendent toutes les deux vers 0.

4.4.2 Étude de la précision inhérente à la méthode

L'évaluation de la précision de la méthode est effectuée en utilisant le modèle 3D générique de vessie associé à une texture de vessie de porc (voir figure 4.6(b)). Nous avons simulé une séquence de 65 acquisitions de données 3D/2D selon un scénario pré-défini de combinaisons de déplacements 3D entre les acquisitions consécutives de la séquence. Les six paramètres des transformations théoriques appliqués durant la séquence sont détaillés dans le tableau 4.1.

Il est important de noter que les variations de translation (0.33 et 0.25 *mm*) et de rotation (1°) sont plus importantes que celles usuellement observées dans des séquences cystoscopiques réelles. Cette remarque se base sur la comparaison des transformations $\mathbf{T}_{2D}^{k-1,k}$ simulées (et calculés à partir des $\mathbf{T}_{3D}^{k-1,k}$) avec les transformations 2D perspectives calculées pour des séquences cystoscopiques réelles.

Nous avons calculé les transformations $\widetilde{\mathbf{T}}_{3D}^{k-1,k}$ et $\widetilde{\mathbf{T}}_{2D}^{k-1,k}$ entre chaque paire d'acquisition consécutive en utilisant l'information mutuelle (\mathcal{S}_{IM}) puis la distance quadratique (\mathcal{S}_{DQ}). $\epsilon_{3D}^{k-1,k}$ et $\epsilon_{2D}^{k-1,k}$ ont donc été calculés deux fois pour quantifier la précision de chacune des mesures de similarité. La figure 4.8 résume les différentes erreurs obtenues. Comme le témoigne la figure 4.8, la méthode que nous avons proposé estime d'une façon

TABLEAU. 4.1 – Valeurs des paramètres des transformations 3D utilisées pour la simulation des acquisitions. θ , ϕ et ψ désignent trois angles de rotation en degré et t_x, t_y et t_z représentent les trois translations en mm (voir équation (4.4) pour plus de détails sur ces paramètres).

Acquisitions (k)	θ	ϕ	ψ	t_x	t_y	t_z
1 ... 17	0	0	0	0.33	0	0
17 ... 21	0	0	1	0	0.33	0
21 ... 25	0	0	1	0	0.33	0
25 ... 29	0	0	0	0	0.33	0.25
29 ... 33	0	0	0	0	0.33	0.25
33 ... 41	1	0	0	0.33	0	0
41 ... 49	1	0	0	0.33	0	0
49 ... 65	0	0	0	0	0.33	0

précise les transformations 3D et 2D et les résultats sont comparables pour les deux mesures de similarité \mathcal{S}_{IM} et \mathcal{S}_{DQ} . En effet, les profils d'évolution des courbes d'erreurs 3D et 2D ont une forme similaire. En d'autres termes, lorsque $\epsilon_{3D}^{k-1,k}$ augmente (diminue), $\epsilon_{2D}^{k-1,k}$ augmente (diminue). En moyenne, l'erreur 3D est égale à 0.058 mm pour l'utilisation de \mathcal{S}_{IM} contre 0.059 mm pour \mathcal{S}_{DQ} . L'erreur 2D moyenne vaut respectivement 0.15 et 0.16 pixels pour les mesures \mathcal{S}_{IM} et \mathcal{S}_{DQ} . Pour $\epsilon_{3D}^{k-1,k}$ on remarque que les deux mesures de similarité affichent une grande précision lorsque le mouvement est une pure translation (t_x, t_y et t_z). Les erreurs 3D pour ce type de mouvement sont autour de 0.04 mm. Par contre, les erreurs 3D augmentent pour les deux mesures de similarité lorsque le mouvement est une combinaison de rotations et de translations (acquisitions 17 à 25, et acquisitions 33 à 49). Nous remarquons également qu'il y a un léger avantage en faveur de l'information mutuelle lorsque des translations (t_x) sont combinées à des rotations (ψ) (acquisition 17 à 25).

Remarques et bilan. Lors d'un examen cystoscopique, le clinicien cherche généralement à positionner l'instrument d'une manière perpendiculaire à la paroi de la vessie pour assurer un meilleur rendu visuel. Ensuite, le clinicien balaye une zone d'intérêt avec une trajectoire plus ou moins lisse (sans discontinuité). Le mouvement entre deux acquisitions est dominé essentiellement par des translations (pour le balayage de surfaces) combinées à de très faibles rotations (pour le positionnement perpendiculaire à la surface imagée). C'est précisément pour ce type de transformations que notre méthode affiche les meilleures performances.

Cette étude de la précision inhérente à la méthode proposée nous a permis d'abord de choisir la mesure de similarité \mathcal{S}_{IM} pour le calcul de la fonction coût \mathcal{F} , car elle fournit

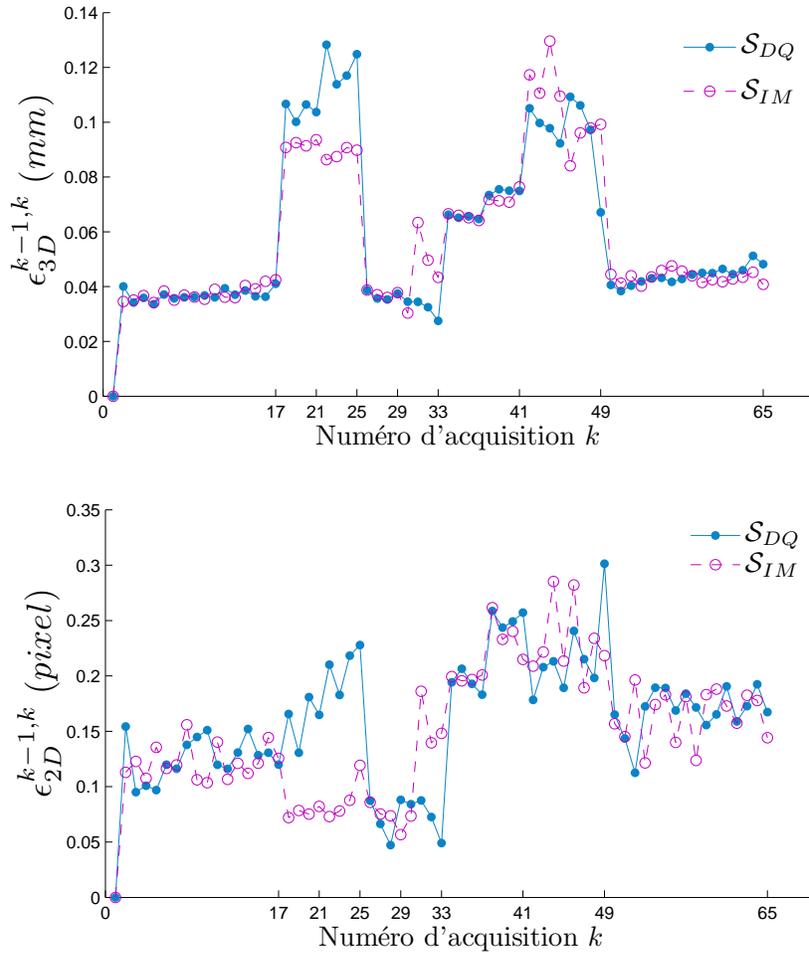


FIGURE. 4.8 – Erreurs $\epsilon_{3D}^{k-1,k}$ et $\epsilon_{2D}^{k-1,k}$ obtenues lors de l'estimation des transformations 3D et 2D entre chaque paire d'images consécutives de la séquence simulée.

de meilleurs résultats que \mathcal{S}_{DQ} si des translations sont associées à des rotations. \mathcal{S}_{IM} est donc la mesure de similarité retenue pour le reste des expérimentations de ce document. Les résultats de cette étude nous ont également permis de vérifier la précision de la méthode que nous avons proposée pour des données idéales. C'est donc la précision maximale de la méthode. Cette précision est de l'ordre de 0.06 mm lorsqu'on ramène un point 3D du repère k au repère $k-1$ et de l'ordre de 0.15 pixel lorsqu'on recalcule une image I_k sur l'image I_{k-1} .

4.4.3 Étude de la robustesse inhérente à la méthode

Cette étude de robustesse est effectuée par rapport à la variabilité des textures des images cystoscopiques et par rapport à la vitesse de déplacement de l'instrument dans la

vessie. Pour ce faire, nous avons utilisé à nouveau le modèle 3D de vessie tapissé cette fois-ci par trois images (I, II et III) extraites de différentes séquences cystoscopiques cliniques (voir figure (4.7)). Ainsi, nous obtenons trois fantômes de vessie présentant des textures très différentes.

La robustesse de la méthode par rapport à la vitesse de déplacement de l'instrument consiste à délimiter les intervalles de valeur des paramètres de la transformation 3D pour lesquelles l'estimation des paramètres des transformations est effectué avec une précision acceptable. Cette précision est définie par des erreurs 3D et 2D valant approximativement 0.1 mm et 0.1 pixel . Pour chaque fantôme, nous réalisons une première acquisition ($k = 1$) en utilisant l'outil de simulation que nous avons développé. L'instrument virtuel d'acquisition est positionné à une distance de 4 cm de la surface interne du fantôme. Ensuite, nous appliquons un déplacement connu sur l'instrument virtuel pour effectuer une deuxième acquisition ($k = 2$). Nous ne varions qu'un seul paramètre à la fois pour les transformations $\mathbf{T}_{3D}^{2,1}$ simulées. Cette façon d'opérer permet de déterminer la valeur maximale pour chaque paramètre qui conduit à une estimation précise des transformations 3D et 2D. Nous n'avons fait varier dans cette étude de robustesse que les valeurs des paramètres t_x , t_z , ϕ et ψ , car l'impact sur les données simulées des paramètres t_y et θ sur les images est similaire à celui des paramètres t_x et ϕ . t_x et t_y impliquent des translations dans le plan image, alors que la translation t_z (selon l'axe optique de la caméra) conduit à un changement de facteur d'échelle dans l'image. Les paramètres de rotation θ et ϕ (autour respectivement les axes \vec{x}_c et \vec{y}_c) impliquent des changements de perspective dans les images. Quant à la rotation ψ selon l'axe optique de la caméra, l'impact qui en résulte est une rotation dans le plan image.

Paramètre t_x . Le premier paramètre que nous avons étudié est t_x . Les valeurs à retrouver et appliquées lors de la simulation sont données dans la deuxième colonne du tableau 4.2. Les valeurs à retrouver pour les autres paramètres de $\mathbf{T}_{3D}^{1,2}$ doivent être nulles. Le tableau montre que pour les translations inférieures à 1.2 mm , l'estimation des paramètres de transformation 3D est assez précise. Non seulement les valeurs de t_x sont proches des valeurs théoriques (celles qui sont entre parenthèses), mais les valeurs des autres paramètres sont quasi nulles. Pour une translation de 1.5 mm , les valeurs de t_x trouvées sont quasiment égales aux valeurs théoriques. Par contre, les valeurs du reste des paramètres s'écartent sensiblement de 0. Cela rend le résultat de la méthode moins précis ($\epsilon_{3D}^{1,2} \approx 0.19\text{ mm}$ et $\epsilon_{2D}^{1,2} \approx 0.14\text{ pixel}$ pour $t_x = 1.5\text{ mm}$).

Paramètre t_z . Le deuxième paramètre que nous considérons dans cette étude de robustesse est t_z . Il correspond à une translation selon l'axe \vec{z}_c du repère caméra. Les valeurs de t_z testées sont représentées entre parenthèses dans la deuxième colonne du tableau 4.3. Les résultats montrent que la valeur maximale autorisant une estimation précise des paramètres de la transformation 3D est égale à 1.2 mm. Pour cette valeur de t_z , les erreurs $\epsilon_{3D}^{1,2}$ et $\epsilon_{2D}^{1,2}$ sont respectivement autour de 0.1 mm et 0.1 pixel.

Paramètre ϕ . Il correspond à une rotation en degré selon l'axe \vec{y}_c du repère caméra. Les valeurs testées de ϕ varient de 0.5° à 2.5° (voir la deuxième colonne du tableau 4.4). Pour $\phi \leq 1^\circ$, notre algorithme conduit à des erreurs faibles ($\epsilon_{3D}^{1,2} \approx 0.6$ mm et $\epsilon_{2D}^{1,2} \approx 0.12$ pixel). Pour $\phi > 1^\circ$, les erreurs 3D et 2D augmentent considérablement pour atteindre approximativement 0.2 mm et 0.4 pixel.

Paramètre ψ . Le dernier paramètre testé est l'angle de rotation ψ . De même que pour ϕ , nous avons varié sa valeur de 0.5° à 2.5° (voir tableau 4.5). Les résultats que nous obtenons sont similaires à ceux trouvés pour l'étude du paramètre ϕ . Notre algorithme affiche une meilleure précision quand l'angle ψ est inférieur à 1. Les erreurs $\epsilon_{3D}^{1,2}$ et $\epsilon_{2D}^{1,2}$ pour cette valeur de rotation sont approximativement égales à 0.1 mm et 0.1 pixel.

Remarques et bilan. Les expérimentations effectuées avec des données simulées montrent que les résultats obtenus avec les trois fantômes sont très similaires. On peut donc conclure que la méthode proposée est potentiellement robuste vis-à-vis de la variabilité de texture des images cystoscopiques. Cette robustesse est étroitement liée au choix de l'information mutuelle comme mesure de similarité dans le processus d'estimation des transformations 3D et 2D.

Dans notre approche de construction de surfaces étendues, la qualité du placement des points 3D dans un repère global dépend de la qualité de l'estimation des transformations 3D. Plus l'estimation de ces transformations est robuste et précise plus la surface 3D à reconstruire est fidèle à la forme réelle de la surface imagée. La robustesse et la précision des transformations 3D estimées a aussi un impact sur la qualité visuelle de l'image panoramique générée. Par conséquent, les valeurs limites des paramètres que nous retenons sont celles qui assurent une estimation très précise de la transformation 3D, même si au-delà de ces limites, notre méthode fournit une estimation qui reste précise qualitativement. Les limites sur les paramètres du déplacement 3D entre chaque paire d'acquisitions consécutives sont définies comme suit :

$$\begin{aligned}
 -1^\circ &\leq \theta \leq 1^\circ \\
 -1^\circ &\leq \phi \leq 1^\circ \\
 -1^\circ &\leq \psi \leq 1^\circ \\
 -1.2 \text{ mm} &\leq t_x \leq 1.2 \text{ mm} \\
 -1.2 \text{ mm} &\leq t_y \leq 1.2 \text{ mm} \\
 -0.9 \text{ mm} &\leq t_z \leq 0.9 \text{ mm}
 \end{aligned}$$

Les bornes ainsi définies correspondent à des déplacements de cystoscope (entre deux acquisitions consécutives) qui sont plus grandes que ceux usuellement rencontrés dans des examens réels.

TABLEAU. 4.2 – Valeurs des paramètres θ , ϕ , ψ , t_x , t_y et t_z obtenues pour les différentes paires d’acquisitions simulées en variant la valeur de (t_x) . Les valeurs théoriques de t_x à trouver sont données dans la deuxième colonne du tableau alors que les valeurs estimées sont fournies dans la sixième colonne. Les colonnes 9 et 10 donnent les erreurs 3D et 2D relatives à chaque résultat de traitement d’une paire d’acquisitions. Le tableau présente aussi le nombre d’itérations effectuées aussi que la valeur finale de \mathcal{F} .

fantômes	(t_x)	θ	ϕ	ψ	t_x	t_y	t_z	$\epsilon_{3D}^{1,2}$	$\epsilon_{2D}^{1,2}$	\mathcal{F}	itér.
I	(0.3)	-0.002	0.004	0.005	0.298	-0.002	0.001	0.033	0.083	-3.002	45
	(0.6)	-0.004	0.038	0.070	0.563	0.001	-0.002	0.063	0.159	-3.101	54
	(0.9)	-0.002	0.044	-0.050	0.876	-0.006	0.034	0.094	0.117	-3.239	55
	(1.2)	0.023	-0.009	0.034	1.205	0.017	0.067	0.054	0.065	-2.889	83
	(1.5)	0.082	0.155	0.013	1.374	0.073	-0.004	0.120	0.114	-3.437	135
II	(0.3)	-0.001	0.004	0.005	0.300	-0.002	-0.001	0.035	0.053	-3.412	47
	(0.6)	0.029	0.015	0.014	0.593	0.029	-0.003	0.046	0.104	-3.529	55
	(0.9)	-0.024	0.089	0.037	0.833	-0.024	0.022	0.050	0.046	-3.687	54
	(1.2)	-0.093	0.051	-0.010	1.160	-0.082	0.005	0.050	0.103	-3.315	78
	(1.5)	0.129	0.185	0.013	1.348	0.116	-0.014	0.191	0.157	-3.785	122
III	(0.3)	-0.002	0.004	0.010	0.299	-0.001	0.000	0.030	0.087	-3.666	54
	(0.6)	0.022	0.020	0.004	0.590	0.021	-0.003	0.059	0.101	-3.757	57
	(0.9)	-0.049	0.052	0.037	0.859	-0.047	0.044	0.052	0.053	-3.727	46
	(1.2)	-0.109	0.091	0.140	1.127	-0.082	-0.004	0.015	0.038	-3.251	71
	(1.5)	-0.011	0.322	0.161	1.217	0.016	-0.102	0.221	0.141	-3.390	67

TABLEAU. 4.3 – Valeurs des paramètres θ , ϕ , ψ , t_x , t_y et t_z obtenues pour les différentes paires d'acquisitions simulées pour les trois différents fantôme en variant la valeur de (t_z). Les valeurs théoriques de t_z à trouver sont données dans la deuxième colonne du tableau alors que les valeurs estimées sont détaillées dans la huitième colonne.

fantômes	(t_z)	θ	ϕ	ψ	t_x	t_y	t_z	$\epsilon_{3D}^{1,2}$	$\epsilon_{2D}^{1,2}$	\mathcal{F}	itér.
I	(0.3)	0.049	-0.032	0.009	0.003	0.021	0.318	0.040	0.065	-3.067	48
	(0.6)	0.043	-0.011	0.010	-0.035	-0.005	0.629	0.073	0.086	-3.026	65
	(0.9)	0.001	0.040	0.001	-0.101	-0.066	0.942	0.113	0.094	-3.008	49
	(1.2)	-0.002	0.048	0.006	-0.129	-0.090	1.258	0.150	0.064	-3.008	59
	(1.5)	-0.116	0.027	-0.007	-0.134	-0.215	1.589	0.200	0.104	-3.003	73
II	(0.3)	0.023	0.010	0.000	-0.031	-0.002	0.313	0.038	0.050	-3.483	53
	(0.6)	0.017	0.015	0.039	-0.060	-0.027	0.630	0.070	0.081	-3.408	59
	(0.9)	0.053	-0.013	0.014	-0.054	-0.018	0.954	0.108	0.104	-3.420	53
	(1.2)	-0.025	0.092	0.003	-0.170	-0.109	1.252	0.148	0.104	-3.418	59
	(1.5)	-0.026	0.069	0.003	-0.172	-0.134	1.590	0.198	0.116	-3.430	81
III	(0.3)	0.031	-0.016	0.006	-0.009	0.006	0.321	0.070	0.094	-3.707	52
	(0.6)	0.038	-0.008	-0.002	-0.036	-0.011	0.626	0.089	0.063	-3.670	48
	(0.9)	-0.024	0.024	0.002	-0.087	-0.089	0.941	0.118	0.112	-3.650	67
	(1.2)	-0.203	0.120	-0.014	-0.194	-0.269	1.231	0.173	0.080	-3.582	67
	(1.5)	-0.040	0.598	0.014	-0.647	-0.144	1.507	0.224	0.113	-3.456	94

TABLEAU. 4.4 – Paramètres de transformation 3D obtenues pour chaque paire d'acquisitions simulées en variant la valeur de (ϕ). Les valeurs théoriques et estimées de ce paramètre sont données respectivement dans les colonnes 2 et 4. Les valeurs du reste des paramètres doivent tendre vers 0.

fantômes	(ϕ)	θ	ϕ	ψ	t_x	t_y	t_z	$\epsilon_{3D}^{1,2}$	$\epsilon_{2D}^{1,2}$	\mathcal{F}	itér.
I	(0.5)	-0.006	0.473	-0.003	0.027	-0.005	0.027	0.049	0.103	-2.940	49
	(1)	-0.006	0.964	-0.006	0.021	-0.010	0.026	0.069	0.118	-2.980	57
	(1.5)	-0.028	1.346	0.107	0.134	-0.013	-0.094	0.193	0.584	-2.897	86
	(2)	-0.037	1.746	0.038	0.211	-0.040	-0.025	0.130	0.269	-2.828	86
	(2.5)	-0.149	1.593	0.050	0.789	-0.116	-0.224	0.085	0.302	-2.832	88
II	(0.5)	0.002	0.477	-0.003	0.021	-0.006	0.017	0.048	0.112	-3.322	51
	(1)	-0.003	0.964	-0.013	0.023	-0.013	0.026	0.069	0.103	-3.315	55
	(1.5)	-0.061	1.397	0.037	0.084	-0.071	0.091	0.154	0.476	-2.895	51
	(2)	-0.169	1.765	0.096	0.189	-0.154	-0.031	0.102	0.267	-3.106	63
	(2.5)	-0.339	1.231	0.065	1.115	-0.293	-0.114	0.185	0.326	-3.022	99
III	(0.5)	-0.008	0.462	0.015	0.029	-0.014	0.010	0.034	0.113	-3.494	38
	(1)	0.101	1.018	-0.024	-0.014	0.085	-0.066	0.054	0.115	-3.633	77
	(1.5)	-0.057	1.346	0.086	0.126	-0.066	0.020	0.157	0.463	-3.076	59
	(2)	-0.105	1.890	0.047	0.094	-0.082	-0.197	0.179	0.336	-3.506	77
	(2.5)	-0.091	1.578	0.032	0.799	-0.068	-0.153	0.185	0.407	-3.317	100

TABLEAU. 4.5 – Résultats obtenus pour chaque variation de l'angle de rotation ψ et pour chaque fantôme. Les valeurs théoriques et estimé de ce paramètre sont données respectivement dans les colonnes 2 et 5.

fantômes	(ψ)	θ	ϕ	ψ	t_x	t_y	t_z	$\epsilon_{3D}^{1,2}$	$\epsilon_{2D}^{1,2}$	\mathcal{F}	itér.
I	(0.5)	-0.046	-0.035	0.486	0.002	-0.011	0.002	0.054	0.057	-3.009	40
	(1)	-0.041	-0.075	0.990	0.000	0.017	0.087	0.092	0.100	-2.927	52
	(1.5)	-0.076	-0.124	1.496	0.002	0.010	0.168	0.142	0.154	-2.732	53
	(2)	0.131	-0.224	2.002	0.074	0.230	0.033	0.181	0.166	-3.012	95
	(2.5)	0.102	0.034	2.512	-0.181	0.233	0.009	0.214	0.140	-3.027	104
II	(0.5)	-0.074	-0.033	0.487	-0.000	-0.039	0.023	0.054	0.085	-3.402	43
	(1)	-0.066	-0.074	0.980	-0.004	-0.010	0.109	0.092	0.084	-3.227	49
	(1.5)	-0.079	-0.116	1.457	-0.001	0.003	0.175	0.140	0.153	-3.032	53
	(2)	-0.007	-0.035	2.004	-0.089	0.112	0.007	0.178	0.151	-3.445	81
	(2.5)	0.087	-0.113	2.494	-0.051	0.220	0.016	0.216	0.161	-3.430	98
III	(0.5)	-0.007	-0.036	0.496	0.002	0.024	-0.002	0.056	0.088	-3.654	58
	(1)	-0.068	-0.078	0.957	0.005	-0.020	0.097	0.101	0.104	-3.365	56
	(1.5)	-0.052	-0.122	1.460	0.011	0.015	0.159	0.145	0.147	-3.147	65
	(2)	0.081	-0.385	1.984	0.216	0.181	0.069	0.185	0.166	-3.562	109
	(2.5)	-0.026	-0.019	2.503	-0.135	0.122	-0.002	0.215	0.139	-3.670	134

4.5 Conclusion

Nous avons présenté dans ce chapitre l'approche générale adoptée pour la construction de cartes 3D relatives aux surfaces 3D étendues de la paroi interne de la vessie à partir d'une séquence de données 3D/2D cystoscopiques. Cette approche nécessite l'estimation de transformations 3D et 2D entre chaque paire d'acquisitions consécutives. Cette estimation est l'élément central de la méthode de construction de cartes 3D.

Nous avons proposé une nouvelle méthode d'estimation simultanée d'une transformation 3D rigide et d'une transformation 2D perspective adaptée aux spécificités des données 3D/2D dont nous disposons. Cette méthode repose sur un système linéaire qui relie les paramètres des deux transformations. Les résultats de l'étude de précision inhérente à la méthode ont montré que même pour des conditions sévères, la méthode permet une grande précision d'estimation des transformations 3D et perspective (les erreurs 3D et 2D moyennes valent respectivement 0.058 mm et 0.15 pixel). Une deuxième étude a montré que la méthode proposée est robuste vis-à-vis la grande variabilité de textures rencontrée dans les séquences d'images cystoscopiques. Dans cette seconde étude, nous avons aussi montré que notre méthode assure une grande précision même lorsque le déplacement 3D est important par rapport à ceux rencontrés en conditions cliniques.

Ces évaluations ont été faites sur des données simulées à partir de fantômes numériques de vessie. Dans le chapitre qui suit, nous proposons une évaluation de la méthode complète de cartographie appliquée à des fantômes réels en utilisant notre prototype expérimental.

Contributions

- Développement d'une nouvelle méthode d'estimation simultanée de transformations 3D et perspective entre une paire d'acquisitions de données 3D/2D cystoscopiques.

Publications

- [Ben-Hamadou *et al.*, 2010a], *IEEE International Conference on Image Processing (ICIP)*.
- [Daul *et al.*, 2010], *IEEE International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*.

5

Évaluation de l’algorithme de cartographie 3D

Dans le chapitre 4, nous nous sommes intéressés à l’évaluation de la performance de notre méthode d’estimation du mouvement 3D entre deux acquisitions cystoscopiques consécutives. Il s’agissait d’évaluer la précision et la robustesse du placement du nuage de points 3D reconstruits pour l’acquisition k dans le repère de l’acquisition numéro $k - 1$. Nous proposons dans ce chapitre de compléter cette évaluation par des expérimentations visant à quantifier la qualité du placement des points 3D de toute une séquence (points reconstruits pour tous les points de vue) dans un repère global commun.

Un banc optique incluant un cystoscope est actuellement en cours de développement au CRAN. L’objectif de ce banc est de valider le principe de la reconstruction 3D de points par vision active dans le cas de la cystoscopie¹⁰. Cependant, nous ne disposons pas actuellement d’un cystoscope 3D utilisable en condition clinique. Nous ne sommes donc pas en mesure de tester nos algorithmes sur des données patients. Pour cette raison, nous avons conçu des fantômes de différentes formes, constitués de surfaces sur lesquelles sont plaquées des textures de vessies très réalistes. L’objectif de ces fantômes est de vérifier la capacité de notre algorithme à retrouver la forme 3D de surfaces imagées dans une séquence d’acquisitions.

Nous présentons d’abord les conditions d’expérimentation, à savoir la description des

10. Ainsi présenté dans le chapitre 2, le principe de la vision active a été validé dans la littérature dans le cas général de l’endoscopie [Chan *et al.*, 2003], mais n’a pas été testé dans le cas particulier de la cystoscopie. Dans [Chan *et al.*, 2003], la lumière structurée laser est projetée par un canal et l’image est acquise à travers un deuxième canal. Dans le cas d’un cystoscope (instrument à un canal), il faut deux chemins optiques différents dans le même canal (le chemin optique de la lumière laser doit être différent du chemin optique de l’acquisition de l’image).

fantômes utilisés, les différents scénarios de trajectoires de notre prototype expérimental et les critères d'évaluation de la méthode de cartographie 3D. Ensuite, nous détaillons les résultats obtenus pour chaque fantôme et pour chaque scénario de trajectoire. Une discussion des résultats obtenus conclut le chapitre.

5.1 Description des expérimentations

5.1.1 Objectif des expérimentations

Dans le chapitre 4 nous avons quantifié l'erreur minimale effectuée lors du placement des points 3D du repère k dans le repère de l'acquisition numéro $k - 1$. L'objectif de ce chapitre est de mettre en lumière l'effet de l'accumulation des erreurs lors de la cartographie 3D. Il s'agit donc de vérifier si le placement des points (reconstruits pour tous les points de vue) dans un repère global unique (celui de la première acquisition) conduit à des surfaces qui ont une géométrie comparable (en termes de tailles et de formes) à celles des surfaces des fantômes. Pour ce faire, nous avons utilisé trois fantômes 3D formés de supports de différentes formes qui ont été tapissés avec une texture de vessie de porc¹¹. La dimension de chaque fantôme est connue au millimètre près.

Les expérimentations consistent à acquérir des séquences de données pour les fantômes en utilisant notre prototype expérimental du chapitre 3. Les données 3D/2D (une image couleur enrichie d'un ensemble de 8 paires de points 3D et 2D pour chaque point de vue) de chaque séquence sont calculées en utilisant les outils et méthodes présentés dans le chapitre 3. Nous utilisons notre méthode d'estimation de mouvement 3D pour construire une image panoramique 2D et placer l'ensemble des points 3D d'une séquence dans le repère caméra de la première acquisition. La forme du nuage de points 3D dans le repère global sera finalement comparée à la forme 3D théorique (connue) du fantôme utilisé.

5.1.2 Fantômes utilisés

Les trois fantômes correspondent à différentes formes géométriques (voir figure 5.1). Nous les présentons du plus simple au plus complexe.

11. Cette texture est obtenue en photographiant une vessie de porc incisée et mise à plat (photographie à haute définition). La texture des vessies de porc est très similaire à celle des vessies humaines.

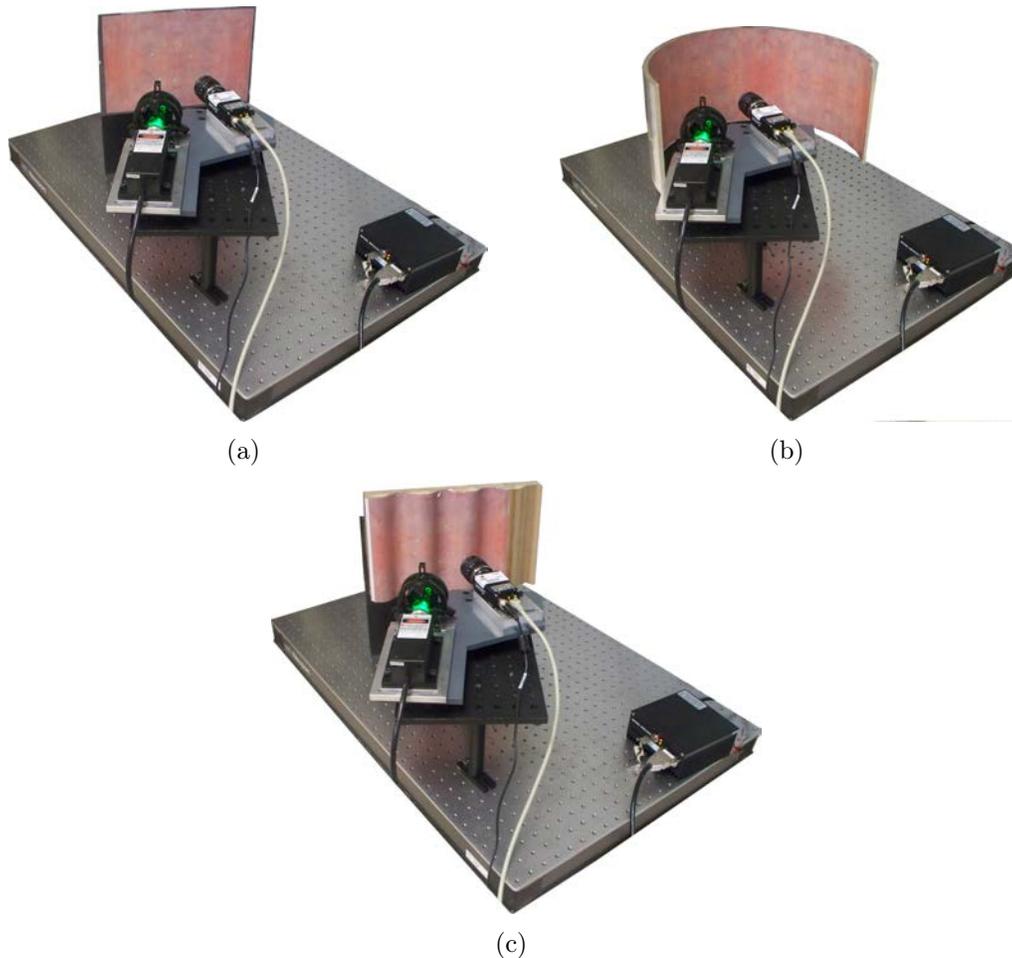


FIGURE. 5.1 – Configuration expérimentale pour l’acquisition de données avec notre prototype à vision active. (a) Fantôme plan (\mathcal{X}_1). (b) Fantôme cylindrique (\mathcal{X}_2). (c) Fantôme formé de vagues (\mathcal{X}_3).

5.1.2.1 Fantôme plan (\mathcal{X}_1)

Le premier fantôme, appelé \mathcal{X}_1 , est le plus simple des trois fantômes. C’est un plan tapissé par une photographie de vessie de porc (voir figure 5.1(a)). C’est le fantôme le moins réaliste car il n’inclut aucune courbure et ne correspond donc pas à la géométrie des surfaces internes de la vessie. Cependant, il permet de vérifier la performance de notre approche sur une surface relativement simple qui simule une situation idéale : la surface imagée est parfaitement plane. En effet, notre algorithme de cartographie 3D est guidé par le recalage d’images 2D (voir chapitre 4). L’hypothèse à la base de cet algorithme est que les images \mathbf{I}_k et \mathbf{I}_{k-1} à recaler visualisent des surfaces quasi-planes étant donné que le cystoscope est proche de la paroi. Dans ce cas, la transformation qui recale les deux images est une perspective.

5.1.2.2 Fantôme cylindrique (\mathcal{X}_2)

La vessie a une forme approximativement ellipsoïdale en supposant qu'elle n'est pas déformée par un autre organe en contact avec elle. Pour des raisons technologiques, nous ne disposons pas d'un fantôme ellipsoïdal (un morceau de sphère qui approxime une partie d'ellipsoïde) que depuis très récemment. C'est pourquoi nous avons considéré dans un premier temps un fantôme cylindrique qui est plus simple à usiner. De toute évidence, ce fantôme ne représente pas idéalement une vessie. Mais il comporte, en comparaison avec le fantôme plan \mathcal{X}_1 , des courbures qui conduisent à une difficulté de traitement de données proche de celle liée à une forme ellipsoïdale. Le fantôme est un demi-cylindre dont le rayon de la face interne vaut 190 *mm* (voir figure 5.1(b)). La face interne du cylindre est tapissée par une texture de vessie de porc. Les expérimentations utilisant ce fantôme consistent à placer le prototype expérimental à l'intérieur de \mathcal{X}_2 et à réaliser des acquisitions en tout déplaçant le fantôme ou le prototype selon le scénario de mouvement à simuler (voir section 5.1.3). Ce fantôme permet de tester si notre approche de cartographie 3D reproduit fidèlement des formes 3D avec des courbures qui peuvent être proches de celles de la vessie. Le fantôme cylindrique permet donc d'évaluer notre algorithme de cartographie lorsque les surfaces vues dans une image ne sont plus quasiment planes.

5.1.2.3 Fantôme en forme de vagues (\mathcal{X}_3)

Le fantôme \mathcal{X}_3 est le plus complexe en terme de géométrie et de courbures. Constitué d'un assemblage de demi-cylindres, sa forme approche une sinusoïde orientée suivant le plan horizontal (voir la figure 5.1(c)). La période de la sinusoïde vaut 80 *mm* et la différence entre les profondeurs extrêmes vaut 10 *mm*. Comparé au demi-cylindre dont la courbure est relativement limitée, le fantôme \mathcal{X}_3 nous place dans des conditions relativement « sévères » car la courbure de la surface est beaucoup plus grande, ce qui fait que les surfaces imagées ne sont plus quasiment planes (c'est l'hypothèse faite lors du recalage des images \mathbf{I}_k et \mathbf{I}_{k-1} avec une transformation perspective). Ce fantôme permet d'éprouver la robustesse de notre algorithme de cartographie lorsqu'un organe déforme la vessie en exerçant une pression sur les parois de l'organe (voir l'illustration de la figure 5.2). Dans ce cas de figure, le fantôme \mathcal{X}_3 simule de telles déformations de la vessie.

5.1.3 Scénarios testés

Rappelons que lors d'un examen cystoscopique, le clinicien essaye de positionner la tête du cystoscope plus ou moins perpendiculairement à la paroi de la vessie afin d'assurer un rendu visuel optimal. Ensuite, avec une trajectoire plus ou moins lisse, le clinicien

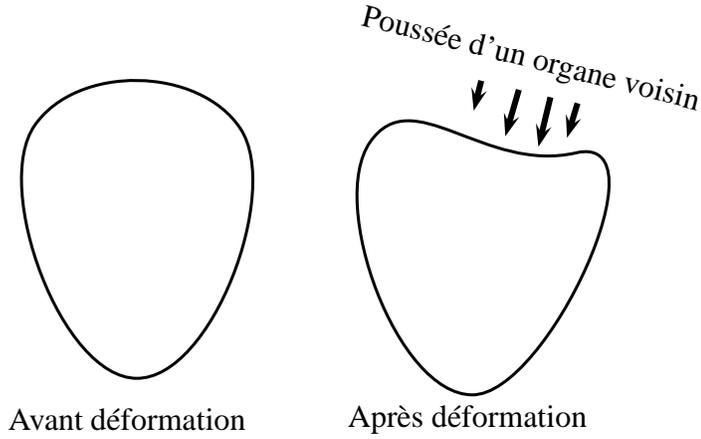


FIGURE. 5.2 – Apparition d’une zone non convexe sur la paroi interne de la vessie suite à la poussée d’un organe comme le rectum contre la paroi externe de la vessie.

couvre une zone d’intérêt. Le mouvement 3D est composé de translations combinées à des rotations plus ou moins fortes. Nous avons tenu compte de cette observation dans le choix des scénarios pour tester des déplacements réalistes. Lorsque la tête du cystoscope est à un ou deux centimètres de la paroi interne de la vessie, la surface vue dans les images mesure entre 1 et quelques centimètres carrés. Pour les différents scénarios, nous avons essayé de respecter, à une échelle près, le ratio entre cette distance et l’aire de la surface observée.

Dans nos expériences, le système des coordonnées $(O_{c_1}, \vec{x}_1, \vec{y}_1, \vec{z}_1)$ décrit la position du repère $\{c_1\}$ lors de sa première acquisition. Dans la plupart des tests la caméra est placée de sorte que son axe optique (qui correspond à \vec{z}_1) est approximativement perpendiculaire à l’orientation locale des surfaces des fantômes.

Dans ce qui suit, les mouvements 3D des différents scénarios sont effectués par rapport aux axes du repère $\{c_k\}$ qui a une position donnée au moment de la $k^{\text{ème}}$ acquisition. Les différents scénarios testés sont décrits ci-dessous.

Scénario I. Le déplacement est formé essentiellement de translations dans la direction de l’axe \vec{x}_k ou \vec{y}_k du repère caméra (les rotations sont très faibles).

Scénario II. Il comprend essentiellement des rotations autour de l’axe \vec{x}_k ou \vec{y}_k . Ce mouvement 3D permet typiquement de positionner l’instrument perpendiculairement aux surfaces imagées, notamment dans le cas du fantôme cylindrique.

Scénario III. Le mouvement combine des translations selon les axes \vec{x}_k , \vec{y}_k et \vec{z}_k du repère caméra et des déplacements importants selon l’axe \vec{z}_k de la caméra. Ce dernier

déplacement permet de zoomer ou dézoomer une région (translations selon \vec{z}_k) tout en balayant une surface (déplacements selon \vec{x}_k et \vec{y}_k).

Scénario IV. Ce dernier scénario est le plus complet et donc le plus réaliste. Il est formé d'une combinaison des mouvements des trois premiers scénarios. C'est le mouvement idéal pour parcourir les surfaces d'un organe creux comme la vessie. Il est composé de translations selon \vec{x}_k ou/et \vec{y}_k pour déplacer le cystoscope vers une région d'intérêt, de rotations autour des mêmes axes pour un positionnement optimal de la tête du cystoscope et d'une translation selon \vec{z}_k pour pouvoir se rapprocher des surfaces et avoir plus de résolution en termes de textures, ou à l'inverse pour s'éloigner de la paroi, pour avoir une vue d'ensemble d'une région plus grande.

5.1.4 Critères d'évaluation

Construction du nuage de points 3D. Soit une séquence de K acquisitions réalisées avec un fantôme \mathcal{X}_λ , avec $\lambda \in \{1, 2, 3\}$. Nous utilisons notre méthode de reconstruction 3D du chapitre 3 pour calculer les données 2D/3D relatives à chaque acquisition $k \in \{1, \dots, K\}$ de la séquence. Ensuite, nous exploitons la méthode d'estimation de mouvement 3D décrite dans le chapitre 4 pour placer l'ensemble des points 3D dans le repère global commun, à savoir celui de la première acquisition. Nous notons $P_{3D}^{i,k}$ le $i^{\text{ème}}$ point 3D de l'acquisition numéro k exprimé dans le repère global $\{c_1\}$.

Préparation des surfaces de référence. L'évaluation de la qualité de la construction d'un nuage de points 3D est effectuée en le comparant à une surface de référence que nous calculons pour chaque expérimentation. La forme analytique des surfaces des deux premiers fantômes est connue (en particulier, le rayon du cylindre est connu au millimètre près). Elle est aussi presque connue pour le troisième fantôme (les rayons des cylindres assemblés pour former le fantôme sont connus). Dans le cadre de notre application (pour laquelle il ne faut pas faire de la mesure dimensionnelle), ces surfaces sont de loin suffisamment précises pour servir de référence et être assimilées à des formes connues exactement. Les informations géométriques de ces surfaces nous ont permis de construire pour chaque fantôme \mathcal{X}_λ une forme 3D notée $\{\mathcal{B}_\lambda\}$ représentant fidèlement ces surfaces. Chaque surface $\{\mathcal{B}_\lambda\}$ est représentée par un maillage dense de points 3D.

Par ailleurs, pour pouvoir comparer le nuage de points 3D à la surface $\{\mathcal{B}_\lambda\}$, il faut retrouver la position de $\{\mathcal{B}_\lambda\}$ dans le repère de la première acquisition (celui du nuage de points 3D). Pour ce faire, nous superposons d'une manière optimale la surface théorique

$\{\mathcal{B}_\lambda\}$ sur les points 3D $P_{3D}^{i,1}$ de la première acquisition¹². Ce choix s'explique par le fait que les positions des points 3D de la première acquisition ne sont affectées que par les erreurs de reconstruction 3D qui sont très faibles (voir chapitre 3), alors que le reste des points sont affectés en plus par les erreurs de placement (dans le repère global) que nous voulons quantifier. Ce positionnement des $\{\mathcal{B}_\lambda\}$ sur les 8 points $P_{3D}^{i,1}$ est effectué moyennant la boîte à outils¹³ de recalage de données 3D proposée par [Salvi *et al.*, 2007].

Mesures d'erreur. L'erreur de positionnement d'un point 3D par rapport à la surface maillée théorique, notée $\varepsilon_{i,k}$, est définie comme la plus courte distance euclidienne entre le point 3D $P_{3D}^{i,k}$ et la surface $\{\mathcal{B}_\lambda\}$:

$$\varepsilon_{i,k} = d(P_{3D}^{i,k}, \{\mathcal{B}_\lambda\}). \quad (5.1)$$

De plus, on définit l'erreur moyenne de positionnement pour le point de vue k :

$$\bar{\varepsilon}_k = \frac{1}{N} \sum_{i=1}^N \varepsilon_{i,k}. \quad (5.2)$$

5.2 Résultats

Dans cette section, les résultats sont présentés par fantôme et par scénario testé. Une synthèse globale de l'ensemble de ces résultats est proposée à la section 5.3.

La simulation des mouvements 3D de tous les différents scénarios est effectuée manuellement, car nous ne disposons pas de tables de translation et d'axes de rotation motorisés. Une alternative aurait été d'utiliser des tables et des axes positionnés manuellement avant chaque acquisition. Cela aurait permis de contrôler parfaitement le mouvement, mais les acquisitions n'auraient pas été effectuées avec un instrument en mouvement d'où une qualité d'image supérieure à celle obtenue dans des conditions réelles avec une caméra en mouvement.

12. Estimation d'une transformation 3D rigide qui positionne $\{\mathcal{B}_\lambda\}$ sur les 8 points $P_{3D}^{i,1}$ en minimisant la distance euclidienne entre chaque point $P_{3D}^{i,1}$ et $\{\mathcal{B}_\lambda\}$.

13. Cette boîte à outils est téléchargeable à partir de l'adresse suivante « <http://eia.udg.es/~cmatabos/research.htm> »

5.2.1 Résultats obtenus pour le fantôme plan

5.2.1.1 Scénario I

L'axe principal de la caméra est approximativement perpendiculaire au plan du fantôme. 95 acquisitions sont effectuées pour un mouvement du prototype composé essentiellement de translations selon l'axe \vec{x}_k de la caméra. En plus du nuage de points 3D générés après traitement de la séquence, la figure 5.3 montre le parcours¹⁴ du prototype et son orientation par rapport au nuage de points. On remarque que le parcours estimé est composé principalement de translations selon les axes \vec{x}_k et \vec{y}_k de la caméra. Les points noirs sur cette figure indiquent les 8 points 3D reconstruits pour la première acquisition. Les huit « traînées » vertes correspondent aux points reconstruits pour le reste des acquisitions ($k \in [2, 95]$) et exprimés dans le repère $\{c_1\}$ de la première acquisition. La figure 5.4 montre le positionnement du plan théorique $\{\mathcal{B}_1\}$ qui passe par les 8 points 3D de la première acquisition. Nous notons que le nuage de points représente bien une forme plane. Nous remarquons également que les points 3D s'éloignent légèrement du plan théorique $\{\mathcal{B}_1\}$ vers la fin de la séquence. La figure 5.5 met en lumière cet éloignement en représentant l'évolution de l'erreur $\bar{\varepsilon}_k$ tout au long de la séquence. On constate que cette erreur moyenne augmente graduellement pour atteindre 2.5 mm à la fin de la séquence. La valeur moyenne et la valeur maximale de l'erreur $\varepsilon_{i,k}$ de positionnement des points 3D valent respectivement 1.30 mm et 2.75 mm. Ces erreurs sont très petites et représentent moins de 2% de la distance de mesure qui vaut approximativement 120 mm (la distance entre le capteur et les surfaces imagées varie de quelques millimètres au cours de l'expérience).

D'autre part, la mosaïque calculée pour cette séquence est montrée dans la figure 5.6. Cette mosaïque a une taille de 442×914 pixels. La qualité visuelle est très bonne car il n'y a pas de discontinuité au niveau des textures.

14. Le mouvement 3D est mis en lumière par la représentation de quelques positions des repères caméra sur la trajectoire 3D.

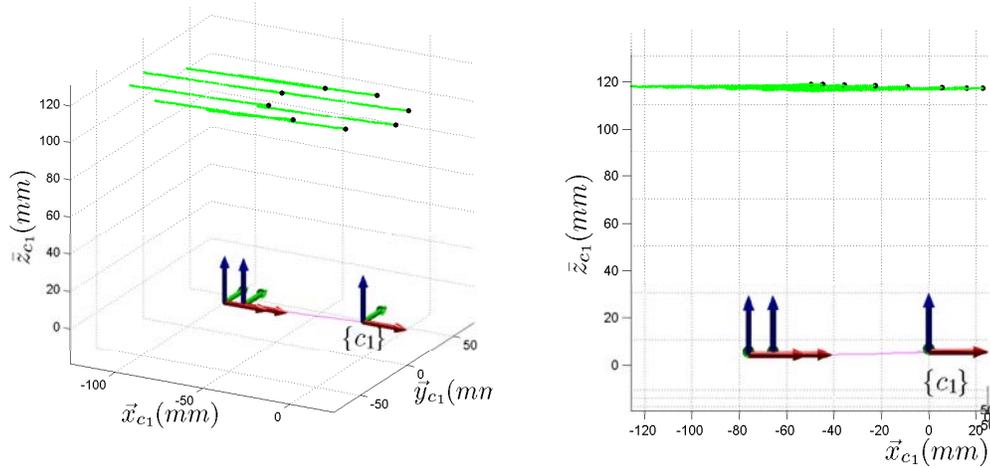


FIGURE. 5.3 – Affichage sous deux points de vue du nuage de points 3D obtenu pour la séquence d’acquisitions réalisées avec le fantôme \mathcal{X}_1 selon le scénario I. Ce nuage est composé de 960 points 3D tous ramenés (exprimés) dans le repère de la première acquisition. L’origine de ce repère a les coordonnées $(0, 0, 0)$. Les repères indiqués avec des flèches représentent le parcours du prototype d’acquisition tout au long de la séquence. On remarque que la trajectoire estimée ne correspond pas exactement à une droite 3D. Cependant, la translation selon l’axe \vec{x}_k représente l’essentiel du mouvement.

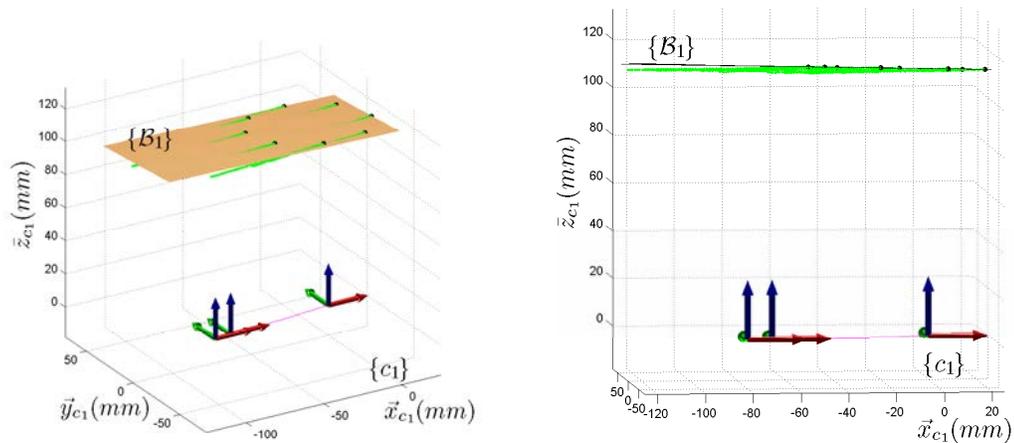


FIGURE. 5.4 – Positionnement de la surface théorique $\{B_1\}$ calculée à partir des 8 points 3D de la première acquisition (points noirs). La surface $\{B_1\}$ est affichée en couleur marron alors que le nuage de points 3D est en vert. Sur la vue de gauche, les « traînées » vertes sont situées légèrement en dessous du plan marron.

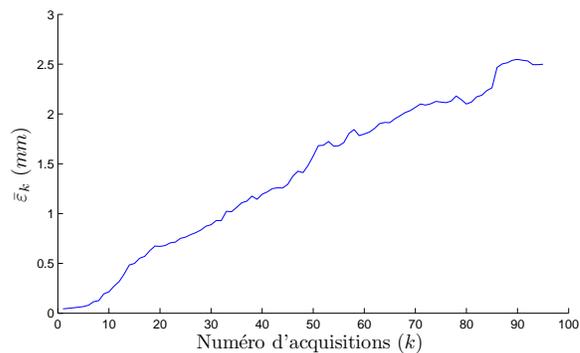


FIGURE. 5.5 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ (erreurs calculées pour le scénario I et le fantôme \mathcal{X}_1) en fonction du point de vue k . Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.

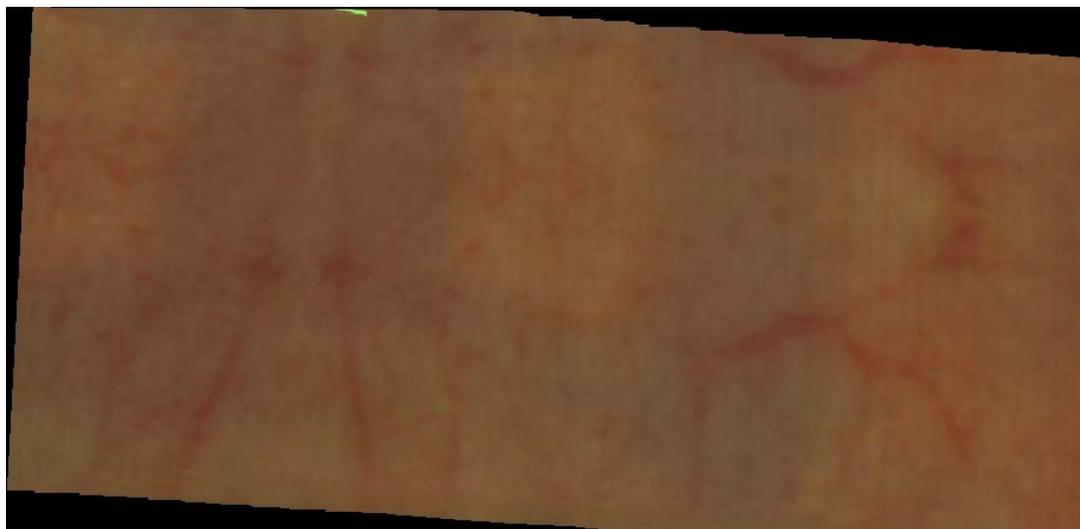


FIGURE. 5.6 – Mosaïque obtenue pour la séquence de 95 acquisitions réalisées avec le fantôme \mathcal{X}_1 selon le scénario I.

5.2.1.2 Scénario II

Pour ce deuxième scénario, nous avons réalisé une séquence de 43 acquisitions avec le fantôme \mathcal{X}_1 . Le mouvement 3D est composé essentiellement de rotations autour de l'axe \vec{y}_k de la caméra du prototype. Le nuage de points 3D obtenu pour cette séquence est représenté sur la figure 5.7. La trajectoire calculée pour le prototype obtenue montre que le mouvement 3D estimé est effectivement composé de rotations autour de l'axe \vec{y}_k . Ce résultat est donc cohérent avec le déplacement 3D appliqué au prototype à vision active pour l'acquisition de la séquence.

La figure 5.8 montre que le nuage de points 3D retrouve avec précision la forme plane du fantôme \mathcal{X}_1 . Cette observation se confirme dans la figure 5.9 qui montre l'évolution de l'erreur $\bar{\varepsilon}_k$ tout au long de la séquence. Le cumul des erreurs de positionnement des points 3D reste faible au cours de la séquence puisque $\bar{\varepsilon}_k$ vaut 0.17 mm pour le positionnement dans le repère $\{c_1\}$ des points 3D de la deuxième acquisition et atteint 0.7 mm pour l'acquisition 43. Globalement, les valeurs moyenne et maximale de l'erreur $\varepsilon_{i,k}$ valent respectivement 0.46 mm et 1.2 mm .

Cette précision se traduit aussi au niveau de la qualité visuelle de la carte 2D calculée (voir figure 5.10). Cette mosaïque, de taille 603×1041 pixels, ne présente aucune discontinuité de texture. Sur cette mosaïque nous remarquons la présence d'un spot laser qui n'a pas pu être éliminé car aucune autre image de la séquence ne couvre le même emplacement sur la surface imagée. En effet, les pixels verts (correspondant aux spots laser) ne peuvent pas être remplacés par la couleur des pixels homologues d'autres images qui ne sont pas touchés par le spot.

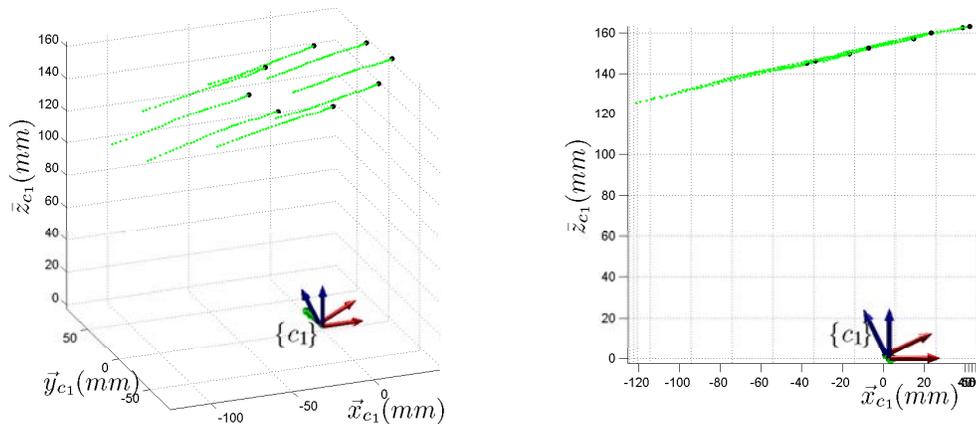


FIGURE. 5.7 – Affichage sous deux différents points de vue du nuage de points 3D obtenu pour le fantôme \mathcal{X}_1 et le scénario II. Ce nuage est composé de 344 points. Les repères 3D représentent à nouveau le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

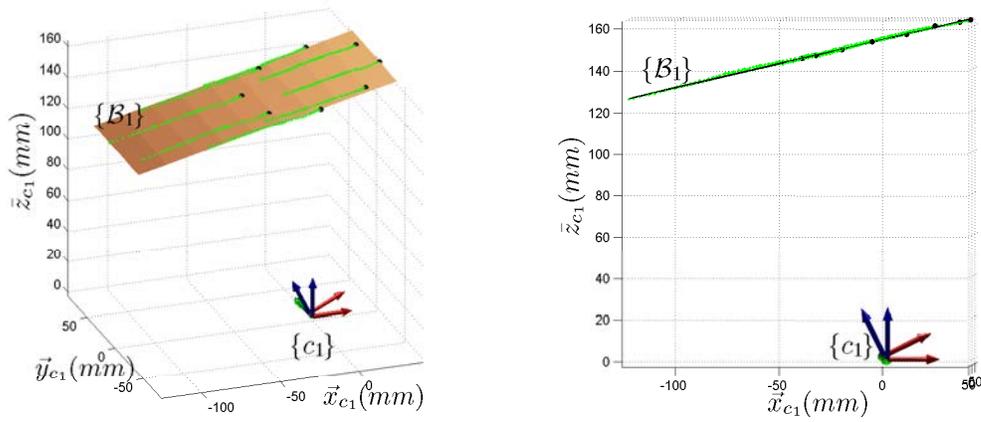


FIGURE. 5.8 – Positionnement de la surface théorique $\{B_1\}$ sur les 8 points 3D de la première acquisition. $\{B_1\}$ est affiché en couleur marron alors que celui le nuage de points 3D est en vert.

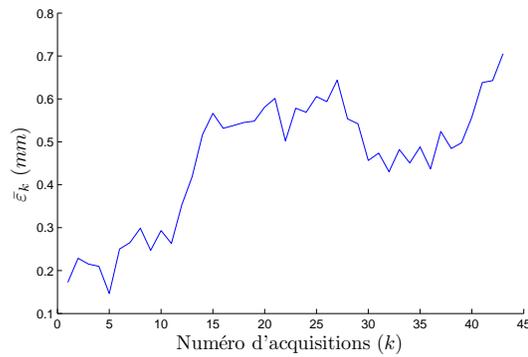


FIGURE. 5.9 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ en fonction du numéro k du point de vue. Ces erreurs sont calculées pour \mathcal{X}_1 et le scénario II. Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.

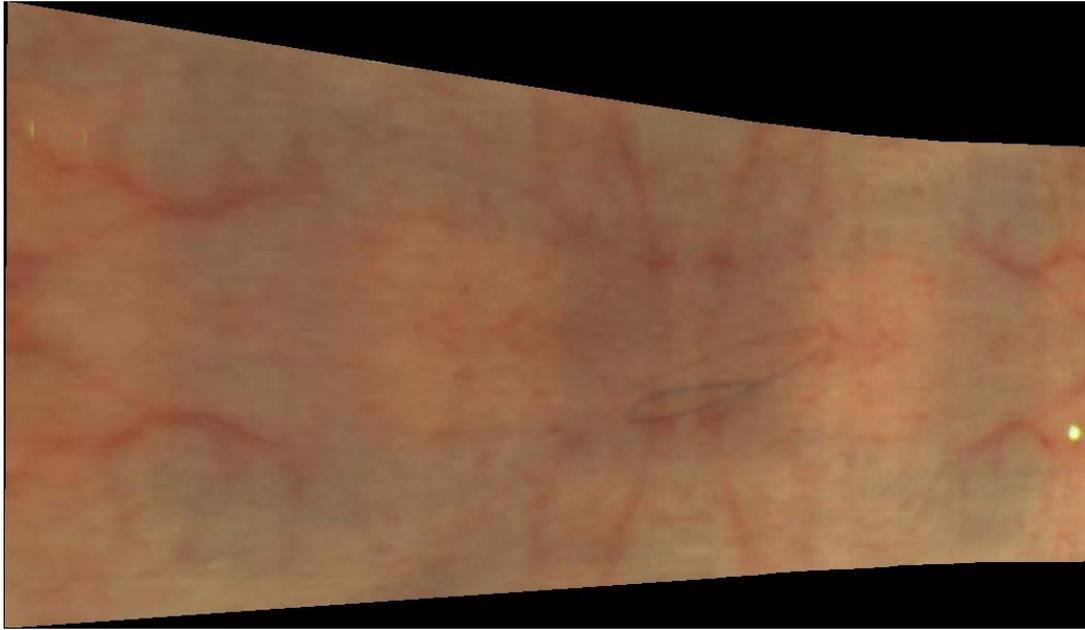


FIGURE. 5.10 – Mosaïque obtenue pour la séquence de 43 acquisitions réalisées avec le fantôme \mathcal{X}_1 selon le scénario II.

5.2.1.3 Scénario III

Ce scénario consiste en l'application de translations selon l'axe \vec{z}_k du repère $\{c_1\}$ de la caméra du prototype pour produire un effet de zoom dans les images. La séquence d'acquisition réalisée avec le fantôme \mathcal{X}_1 selon ce scénario comporte 35 images. Nous avons fait varier la distance entre le prototype et la surface de \mathcal{X}_1 approximativement entre 120 mm et 200 mm . La surface totale couverte par cette séquence est de l'ordre de 130×90 mm^2 . Le résultat du placement de l'ensemble des points 3D dans le repère global est donné dans la figure 5.11. Dans cette figure, nous constatons que le mouvement 3D estimé est effectivement composé essentiellement de translations selon l'axe \vec{z}_k de la caméra avec de faibles translations selon les axes \vec{x}_k et \vec{y}_k . Dans la figure 5.12, nous remarquons que les points 3D sont très proches de la surface théorique $\{\mathcal{B}_1\}$. Ils représentent donc bien une forme plane. La qualité de la construction est confirmée par la figure 5.13 qui fournit l'évolution de la mesure d'erreur $\bar{\varepsilon}_k$ (quantification de l'erreur moyenne de positionnement des points 3D dans le repère $\{c_1\}$). $\bar{\varepsilon}_k$ varie entre 0.08 mm et 0.35 mm . Ramenées à la distance de mesure (de 110 à 160 mm), ces valeurs correspondent à une erreur qui va de 0.4 à 0.9%. D'autre part, la valeur moyenne de l'erreur $\varepsilon_{i,k}$ de positionnement des points 3D dans le repère global est de l'ordre de 0.2 mm alors que sa valeur maximale est approximativement égale à 0.6 mm .

La mosaïque obtenue pour cette séquence est montrée dans la figure 5.14. À l'instar des mosaïques calculées pour les précédents scénarios, cette mosaïque est de bonne qualité visuelle car n'incluant aucune discontinuité de texture.

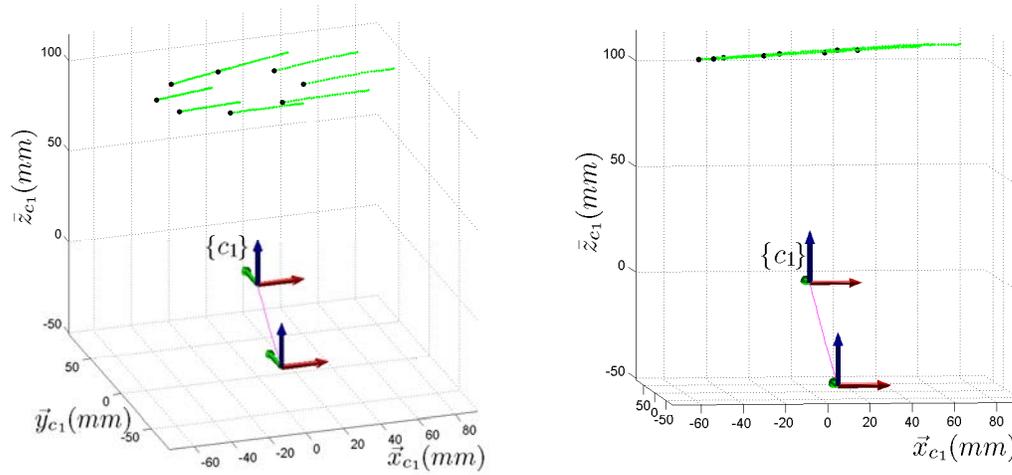


FIGURE. 5.11 – Représentation sous deux différents points de vue du nuage de points 3D calculés pour la séquence d'images acquise pour le fantôme \mathcal{X}_1 selon le scénario III. Ce nuage est composé de 280 points. Les repères 3D représentent le parcours du prototype tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

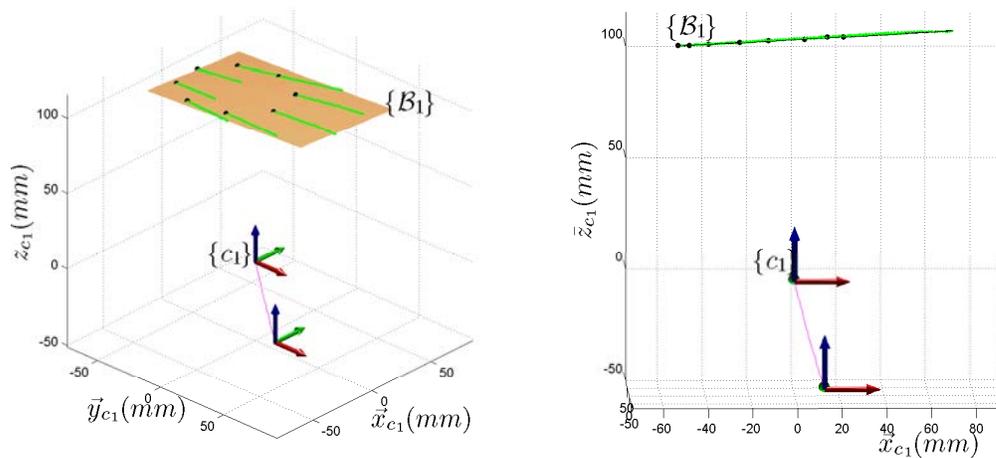


FIGURE. 5.12 – Positionnement de la surface théorique $\{B_1\}$ sur les 3D en se référant aux 8 points 3D de la première acquisition (points noirs). $\{B_1\}$ est représentée en marron alors que le nuage de points 3D est en vert. La surface calculée a bien la forme d'un plan.

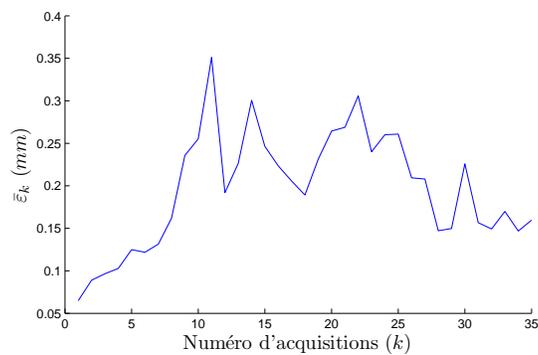


FIGURE. 5.13 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ en fonction du numéro k de l'acquisition (fantôme \mathcal{X}_1 , scénario III). Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.



FIGURE. 5.14 – Mosaïque obtenue pour la séquence de 35 acquisitions réalisées avec le fantôme \mathcal{X}_1 selon le scénario III.

5.2.1.4 Scénario IV

La dernière expérience réalisée avec le fantôme \mathcal{X}_1 consiste en l'acquisition d'une séquence de 24 images selon le scénario IV. Ce dernier correspond à un déplacement 3D du prototype formé d'une combinaison des trois premiers scénarios, à savoir des translations selon les axes \vec{x}_k et \vec{z}_k (zoom) et de faibles rotations autour de l'axe \vec{y}_k de la caméra du prototype. Après traitement de la séquence acquise, nous obtenons le nuage de points 3D présenté dans la figure 5.15. Sur cette figure, nous remarquons que le parcours estimé est effectivement en cohérence avec le mouvement 3D du scénario IV. Cependant, en comparant le nuage de points 3D obtenu avec la surface théorique de la figure 5.16, nous constatons qu'il y a effectivement des erreurs de positionnement de points 3D vers la fin de la séquence. Ces erreurs, qui restent cependant faibles, sont aussi visibles dans la figure 5.17 qui montre les différentes valeurs de l'erreur $\bar{\varepsilon}_k$ obtenues tout au long de la séquence. La valeur maximale de cette erreur vaut 1.9 mm et est mesurée pour le positionnement dans le repère $\{c_1\}$ de l'acquisition 24 (fin de la séquence). D'autre part, la valeur moyenne de l'erreur $\varepsilon_{i,k}$ de positionnement des points 3D par rapport à la surface théorique \mathcal{B}_1 est égale à 1.1 mm . La valeur maximale de cette mesure d'erreur est égale à 4 mm .

La figure 5.18, montre la carte 2D obtenue pour l'expérience réalisée avec le fantôme \mathcal{X}_1 selon le scénario IV. Cette carte 2D est formée de 401×642 pixels. Elle représente une bonne qualité visuelle, néanmoins inférieure à celles des mosaïques obtenues pour les précédents scénarios. En effet, il y a des endroits dans la carte 2D pour lesquels il n'y a pas une continuité parfaite de texture. Il est à noter aussi que lorsque la carte 2D est visuellement moins cohérente, le placement des points 3D dans le repère global se trouve légèrement affecté. Ce résultat est prévisible pour l'algorithme présenté dans le chapitre 4.

La première conclusion qui peut être tirée de l'ensemble des tests réalisés avec le fantôme \mathcal{X}_1 est que **notre algorithme de reconstruction de surfaces a systématiquement été capable de reconstruire la forme d'un plan**. L'objet de la section suivante est de tester notre méthode en présence de courbures dans les surfaces imagées.

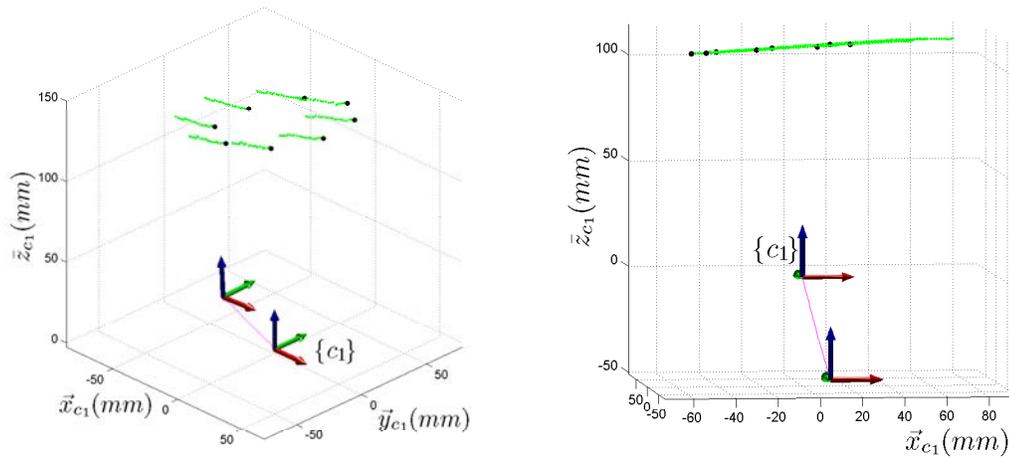


FIGURE. 5.15 – Affichage sous deux différents points de vue du nuage de points 3D pour le fantôme \mathcal{X}_1 et le scénario IV. Ce nuage est composé de 192 points 3D. Les repères 3D représentent le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

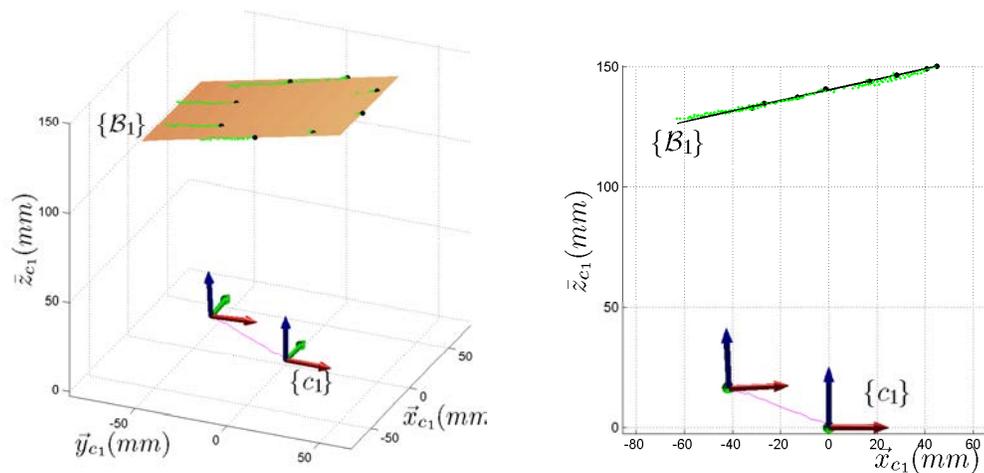


FIGURE. 5.16 – Positionnement de la surface théorique $\{\mathcal{B}_1\}$ sur les 8 points 3D de la première acquisition. $\{\mathcal{B}_1\}$ est représenté en couleur marron alors que le nuage de points 3D est en vert.

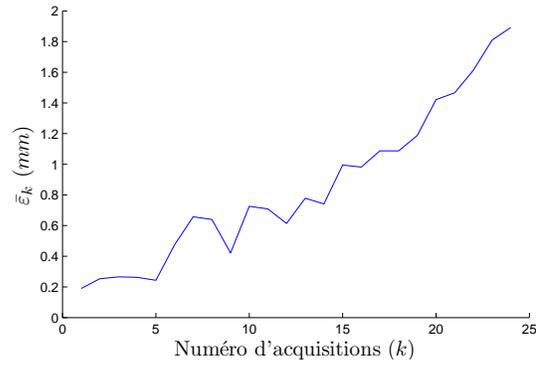


FIGURE. 5.17 – Évolution de la valeur de l'erreur $\bar{\varepsilon}_k$ en fonction du numéro k de l'acquisition (fantôme \mathcal{X}_1 , scénario IV). Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.



FIGURE. 5.18 – Mosaïque obtenue pour la séquence de 25 acquisitions réalisées avec le fantôme \mathcal{X}_1 selon le scénario IV.

5.2.2 Résultats obtenus pour le fantôme cylindrique

5.2.2.1 Scénario I

Pour ce scénario, nous avons réalisé une séquence de 120 acquisitions du demi-cylindre \mathcal{X}_2 en appliquant un mouvement 3D composé essentiellement de translations selon l'axe \vec{x}_k de la caméra du prototype à vision active (voir figure 5.19). Après traitement de la séquence, un nuage de points 3D est généré en plaçant tous les points 3D dans le repère caméra de la première acquisition. En plus du nuage de points 3D, la figure 5.19 montre le parcours du prototype et son orientation par rapport au nuage de points 3D. Il est visible sur cette figure que le parcours estimé pour le prototype correspond essentiellement à de fortes translations selon la direction de l'axe \vec{x}_k de la caméra auxquelles s'ajoute de très faibles translations selon les axes \vec{y}_k et \vec{z}_k de cette même caméra. Nous remarquons aussi que les axes des repères 3D restent approximativement parallèles ce qui signifie que mouvement n'inclut pas de rotations importantes.

Les points noirs de la figure 5.19 symbolisent, comme pour le fantôme plan, les 8 points 3D reconstruits pour le premier point de vue du prototype ($k = 1$). Les huit trajectoires vertes correspondent aux points 3D reconstruits pour le reste des acquisitions ($k \in [2, 120]$) et exprimés dans le repère $\{c_1\}$ de la première acquisition. On constate visuellement sur la figure 5.20 que la surface couverte par les points verts peut effectivement correspondre à la surface d'une portion de cylindre.

La figure 5.20 montre le positionnement optimal de la surface théorique $\{\mathcal{B}_2\}$ correspondant au fantôme \mathcal{X}_2 par rapport aux 8 points 3D de la première acquisition. Ces points 3D sont représentés en noir et le reste des points 3D du nuage est affiché en vert. Nous remarquons que l'ensemble des points 3D représente bien la forme cylindrique du fantôme, même si par moments certains points « décollent » légèrement de la surface de $\{\mathcal{B}_2\}$. Ces erreurs de placements de points 3D dans $\{c_1\}$ sont plus perceptibles sur la figure 5.21 qui représente l'évolution de l'erreur $\bar{\varepsilon}_k$ tout au long de la séquence. Rappelons que l'erreur $\bar{\varepsilon}_k$ est calculée avec les équations (5.1) et (5.2).

Sur cette figure nous remarquons que les erreurs de positionnement dans le repère $\{c_1\}$ n'évoluent pas systématiquement dans un même sens. En effet, du début des acquisitions jusqu'à l'acquisition 55, les points 3D s'éloignent de plus en plus de la surface de $\{\mathcal{B}_2\}$. Cela se traduit par une augmentation de la valeur de $\bar{\varepsilon}_k$ qui culmine approximativement à 6.5 mm. Cependant, après ces acquisitions, la valeur de $\bar{\varepsilon}_k$ commence à baisser, que les points 3D se rapprochant à nouveau de la surface théorique. La valeur moyenne et maximale de l'erreur $\varepsilon_{i,k}$ de placement des points 3D par rapport à la surface théorique sont respectivement de l'ordre de 4 mm et de 8 mm. Pour toute la séquence d'images, la

distance entre la caméra et la surface à reconstruire va de 120 *mm* à 180 *mm* environ. L'erreur moyenne de positionnement de 4 *mm* est donc à ramener à ces distances. La distance entre les points 3D et la surface théorique correspond à 2 ou 3 % de la distance de mesure.

D'autre part, la mosaïque 2D générée avec les 120 images de la séquence est présentée dans la figure 5.22. Elle est de taille 570 × 1100 pixels. Elle a été construite avec les régions centrales (de 500 × 500 pixels) des 120 images de la séquence. La mosaïque affiche une bonne cohérence visuelle. En effet, les textures visibles dans la mosaïque (correspondant par exemple à des vaisseaux sanguins) semblent cohérentes puisqu'elles sont sans discontinuités (des images mal recalées sont à l'origine de discontinuités de texture dans une carte 2D).

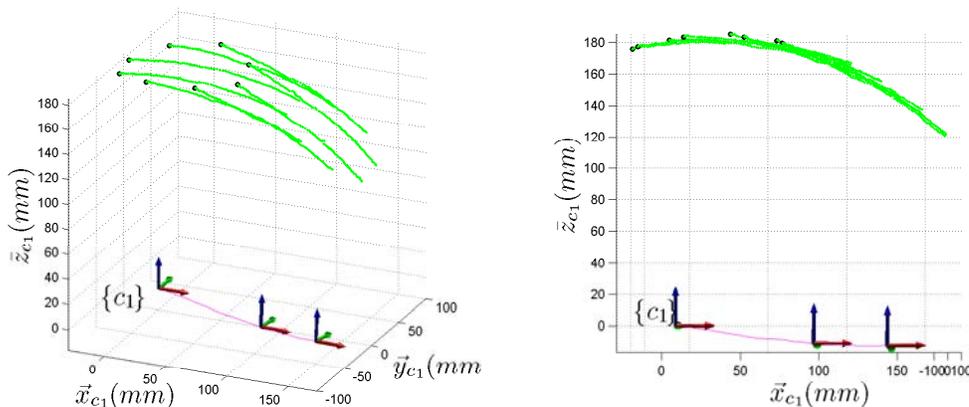


FIGURE. 5.19 – Affichage sous deux différents points de vue du nuage de points 3D obtenu pour la séquence d'acquisitions réalisées avec le fantôme \mathcal{X}_2 selon le scénario I. Ce nuage est composé de 960 points 3D tous ramenés (exprimés) dans le repère de la première acquisition. L'origine de ce repère a les coordonnées (0,0,0). Les repères 3D représentent le parcours du prototype d'acquisition tout au long de la séquence. On remarque que la trajectoire du prototype n'est pas exactement sur une droite 3D. Cependant, la translation selon l'axe \vec{x}_k de la caméra représente l'essentiel du mouvement.

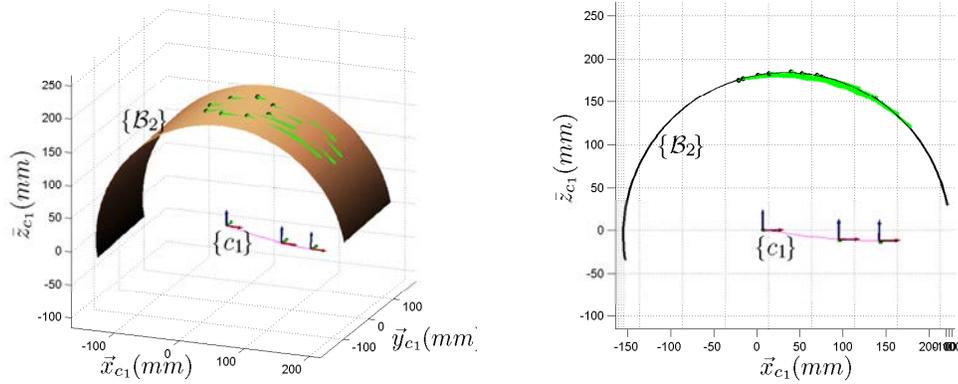


FIGURE. 5.20 – Positionnement de la surface théorique $\{B_2\}$ sur les 8 points 3D de la première acquisition (points noirs). La surface $\{B_2\}$ est en marron, alors que le nuage de points 3D est en vert.

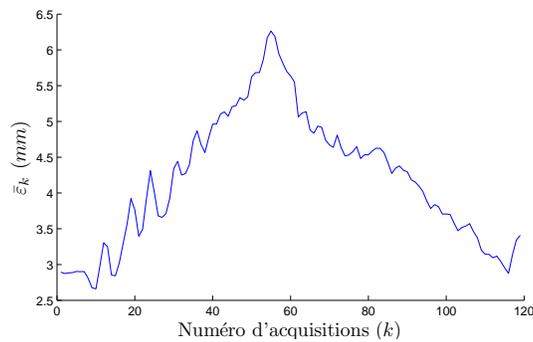


FIGURE. 5.21 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ (erreurs calculées pour le scénario I et le fantôme \mathcal{X}_2) en fonction du numéro k du point de vue. Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.

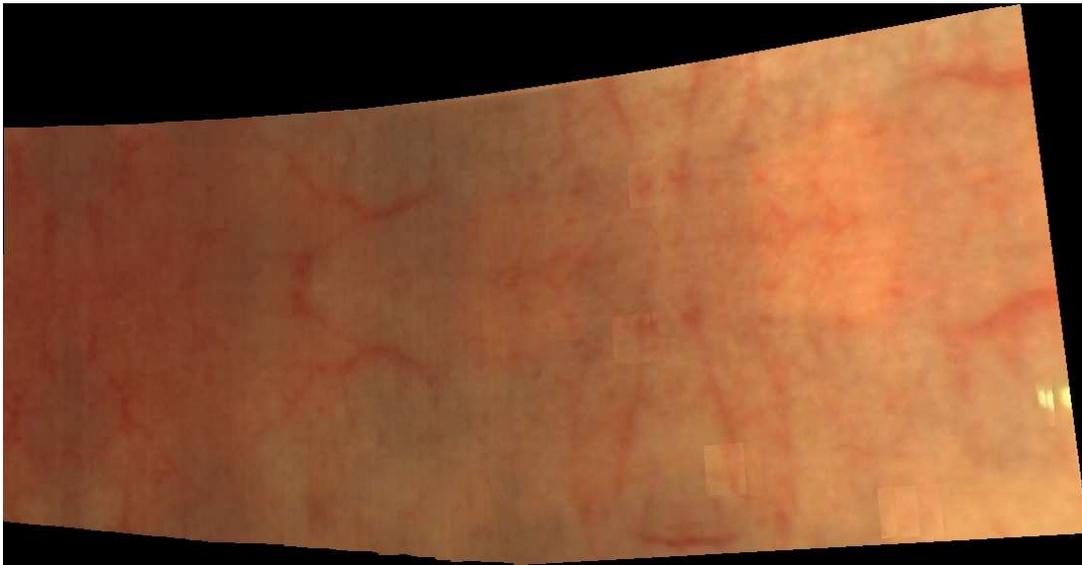


FIGURE. 5.22 – Mosaique obtenue pour la séquence de 120 acquisitions réalisées avec le fantôme \mathcal{X}_2 selon le scénario I.

5.2.2.2 Scénario II

Pour tester ce scénario sur le fantôme \mathcal{X}_2 , nous avons réalisé une séquence de 85 acquisitions couvrant une surface supérieure à un quart de cylindre. Le prototype à vision active est placé à l'intérieur du fantôme. Ensuite nous avons appliqué au prototype un mouvement composé essentiellement de rotations autour de l'axe \vec{y}_k de la caméra. Le nuage de points 3D obtenu pour cette séquence est montré dans la figure 5.23. Nous remarquons que les origines des repères caméra (positions du centre optique de la caméra) ont peu bougé. Les translations effectuées pour cette séquence sont donc effectivement très faibles. Par contre, on remarque qu'il y a principalement des rotations autour de l'axe \vec{y}_k de la caméra. Visuellement, les points noirs et verts placés dans le repère de la première acquisition semblent à nouveau être situés sur une portion de cylindre.

La figure 5.24 montre le positionnement optimal de la surface théorique $\{\mathcal{B}_2\}$ sur les 8 points 3D de la première acquisition (points noirs). Même si les points 3D commencent à s'éloigner de $\{\mathcal{B}_2\}$ après un certain nombre d'acquisitions, l'allure générale de la surface 3D reste cylindrique. L'évolution de la valeur de la mesure d'erreur $\bar{\varepsilon}_k$ est donnée dans la figure 5.25 qui traduit l'augmentation de la distance entre les points 3D et $\{\mathcal{B}_2\}$. Pour cette séquence, l'erreur cumulée atteint 7 mm au bout des 85 acquisitions. La valeur moyenne et la valeur maximale de la mesure d'erreur $\varepsilon_{i,k}$ valent respectivement 3.9 mm et 9.5 mm.

La mosaïque générée pour la séquence est donnée dans la figure 5.26. Elle a une taille de 532×1056 pixels. Le spot laser que nous remarquons dans cette mosaïque provient de la première image. Les spots lasers peuvent être automatiquement enlevés de la carte 2D en les remplaçant par des pixels de régions homologues d'autres images qui ne sont pas illuminées par les spots laser. De telles régions n'existant pas pour ce spot de la première image, celui-ci n'a pas pu être « dissimulé ». À l'exception de ce détail, la mosaïque présente une très bonne qualité visuelle, car n'incluant pas de discontinuités de texture décelables à l'œil nu.

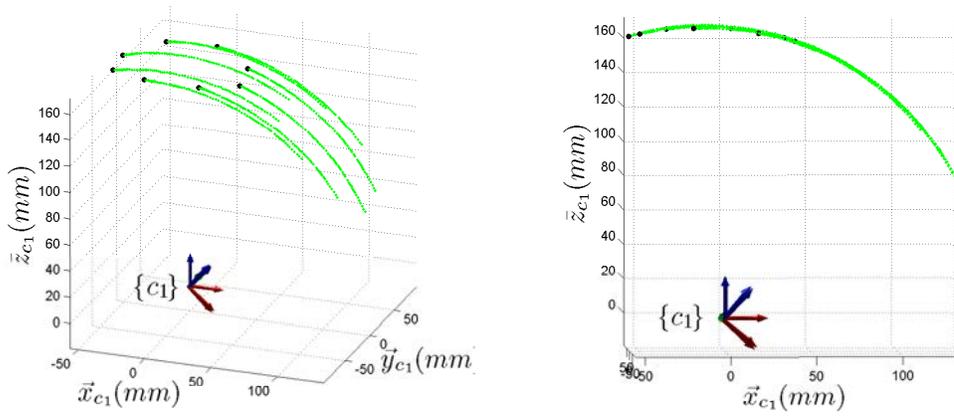


FIGURE. 5.23 – Affichage sous deux différents points de vue du nuage de points 3D calculé pour le fantôme \mathcal{X}_2 selon le scénario II. Ce nuage est composé de 680 points 3D. Les repères 3D représentent à nouveau le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

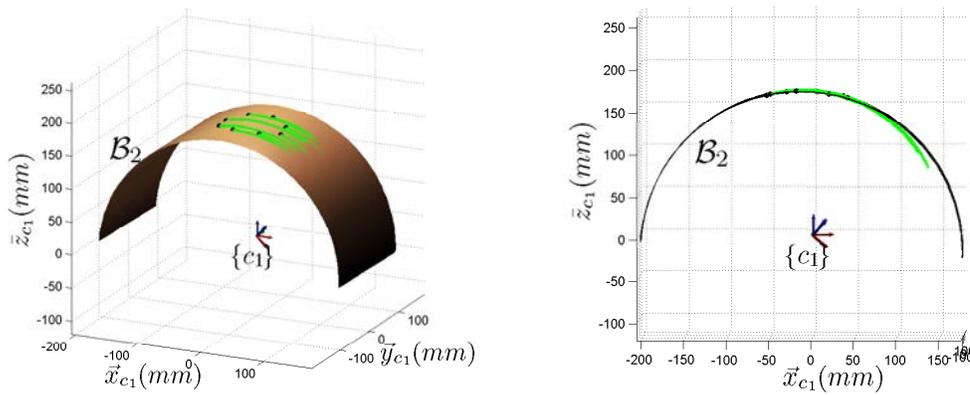


FIGURE. 5.24 – Positionnement de la surface théorique $\{B_2\}$ sur les 8 points 3D de la première acquisition. $\{B_2\}$ est représenté en marron alors que le nuage de points 3D est en vert.

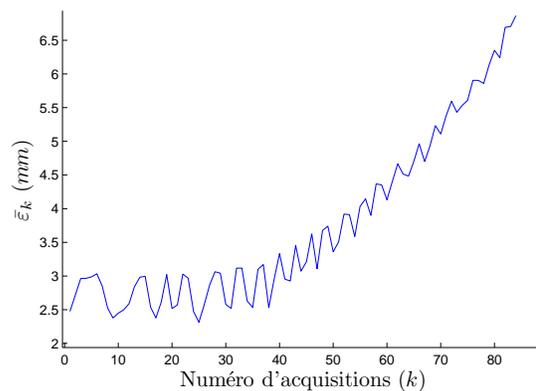


FIGURE. 5.25 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ en fonction du numéro k du pont de vue. Ces erreurs sont calculées pour \mathcal{X}_2 et le scénario II. Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.

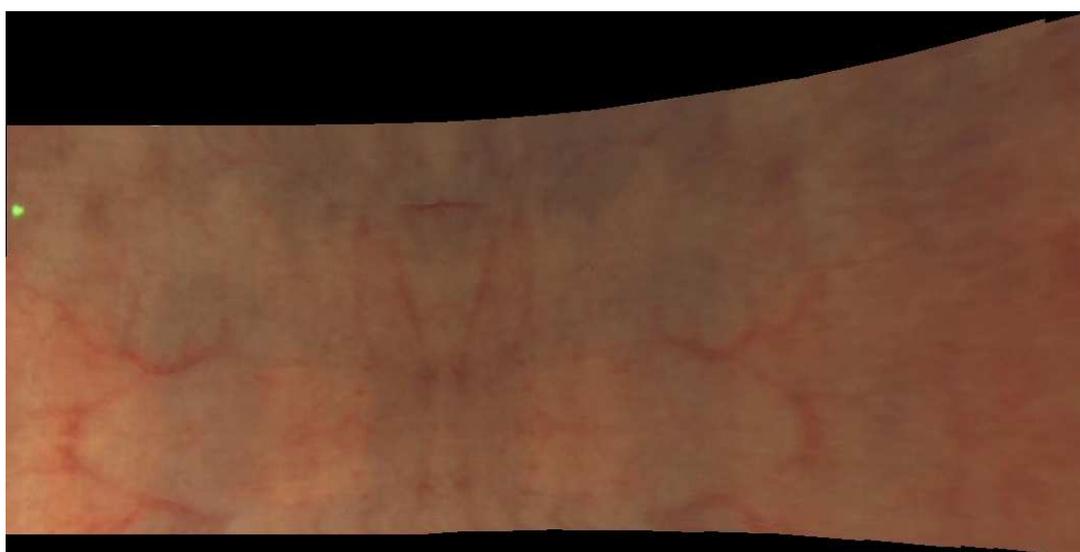


FIGURE. 5.26 – *Mosaïque obtenue pour la séquence de 85 acquisitions réalisées avec le fantôme \mathcal{X}_2 selon le scénario II.*

5.2.2.3 Scénario III

Le scénario III consiste en l'application au prototype d'un déplacement 3D composé essentiellement de translations selon l'axe optique de la caméra (axe \vec{z}_k). Pour réaliser une séquence de 56 acquisitions avec le fantôme \mathcal{X}_2 , nous avons placé le prototype à une distance approximative de 130 *mm* de la face interne du fantôme. Ensuite, nous avons éloigné le prototype de la surface en le déplaçant selon l'axe optique de la caméra jusqu'à atteindre une distance de 200 *mm*. La surface totale imagée dans la séquence est de l'ordre de 140 × 100 *mm*. Il est à noter que le mouvement est aussi constitué de translations selon les axes \vec{x}_k et \vec{y}_k de la caméra. Le nuage de points 3D obtenu pour cette séquence est présenté dans la figure 5.27. Cette figure montre que le chemin emprunté par le prototype est effectivement composé essentiellement de translations selon l'axe \vec{z}_k .

Les erreurs $\bar{\varepsilon}_k$ obtenues pour cette séquence sont données dans la figure 5.29. Les erreurs moyenne et maximale de ε_k sont respectivement égales à 3.7 *mm* et 6.2 *mm* pour des distances de mesure allant de 130 à 200 *mm*.

Pour ce scénario il est difficile de vérifier si le nuage de points 3D recouvre une forme cylindrique ou pas, car la surface couverte est très petite (voir figure 5.28). Cependant, les erreurs obtenues sont faibles relativement si on considère distance de mesures (l'erreur moyenne varie entre 2.8% et 1.85% de la distance de mesure). D'autre part, la mosaïque générée affiche une cohérence visuelle, car nous pouvons à peine distinguer les bords des images collées dans la mosaïque (voir figure 5.30). Le fait qu'il n'y ait pas de discontinuité dans les structures visuelles (*e.g.*, vaisseaux sanguins) est en général un indicateur positif en ce qui concerne la qualité de la construction de la mosaïque et par conséquent de la surface 3D.

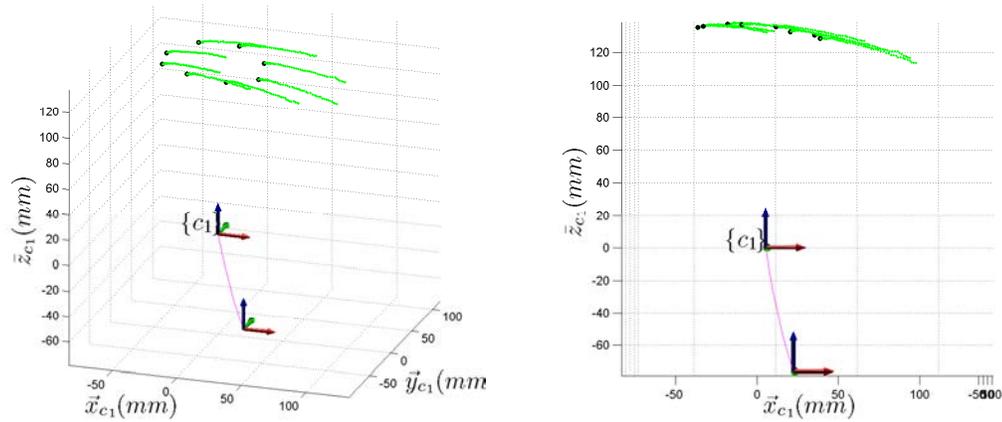


FIGURE. 5.27 – Affichage sous deux différents points de vue du nuage de points 3D pour le fantôme \mathcal{X}_2 et pour le scénario I. Ce nuage est composé de 448 points 3D. Les repères 3D représentent le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

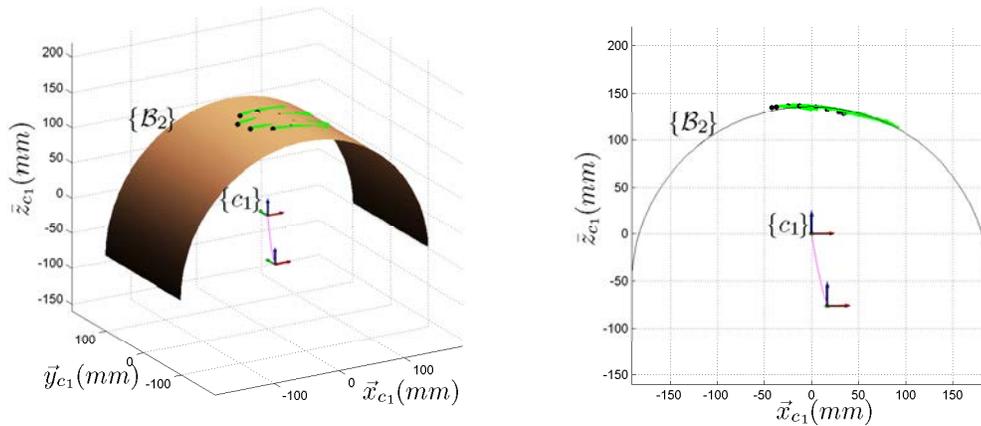


FIGURE. 5.28 – Positionnement de la surface théorique $\{\mathcal{B}_2\}$ sur les 8 points 3D de la première acquisition. $\{\mathcal{B}_2\}$ est représenté en marron alors que celui le nuage de points 3D est en vert.

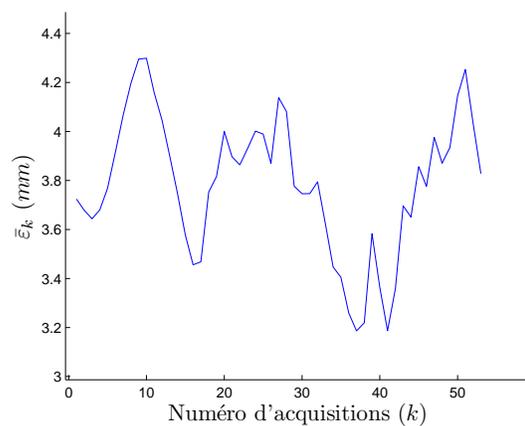


FIGURE. 5.29 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ en fonction du numéro k de l'acquisition (fantôme \mathcal{X}_2 , scénario III). Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.



FIGURE. 5.30 – Mosaique obtenue pour la séquence de 56 acquisitions réalisées avec le fantôme \mathcal{X}_2 selon le scénario III.

5.2.2.4 Scénario IV

La dernière expérience réalisée avec le fantôme \mathcal{X}_2 consiste à tester un déplacement 3D qui combine les mouvements des trois premiers scénarios, à savoir des translations, des rotations et des zooms. La séquence réalisée comporte 150 acquisitions et couvre un quart de cylindre. Le prototype est placé face à la surface interne du fantôme et est déplacé manuellement de telle sorte à obtenir un mouvement composé essentiellement de translations selon \vec{x}_k et \vec{z}_k et de rotations autour de \vec{y}_k . Le nuage de points 3D obtenu pour cette expérience est illustré dans la figure 5.31. Le parcours du prototype estimé par notre méthode est en accord avec le déplacement appliqué manuellement lors de la réalisation de la séquence. En effet, les changements de position symbolisés par les repères de la figure 5.31 font apparaître des translations selon les axes \vec{x}_k et \vec{z}_k et une rotation autour de l'axe \vec{y}_k . La figure 5.31 indique que les traînées des points verts semble effectivement être le support d'une forme cylindrique. La figure 5.32 met en lumière la différence géométrique entre le nuage de points 3D et la surface théorique $\{\mathcal{B}_2\}$. Cette figure confirme que le nuage de points 3D a approximativement une allure cylindrique, même si vers la fin de la séquence les points 3D s'éloignent de la surface théorique pour atteindre une erreur $\bar{\varepsilon}_k$ égale à 8.4 mm (voir l'évolution de cette mesure d'erreur dans la figure 5.33). La valeur moyenne et maximale de l'erreur $\varepsilon_{i,k}$ sont respectivement égale à 4.3 mm et 15.7 mm.

L'image panoramique qui correspond à cette séquence a une taille de 1501×758 pixels (voir figure 5.34). Sur cette mosaïque on peut remarquer que la qualité visuelle est bonne dans les régions construites avec les 100 premières images (il n'y a pas de discontinuité de texture pour ces 100 premières images). Cependant, la qualité visuelle de la carte 2D se dégrade sensiblement quand il s'agit des régions construites avec les dernières images de la séquence. Ce fait est lié au cumul des erreurs de l'estimation des transformations 3D et perspectives.

La première conclusion qui peut être tirée à partir des tests avec le fantôme \mathcal{X}_2 est que, **même si l'algorithme de construction de surfaces (guidé par le recalage d'images consécutives) est basé sur l'hypothèse d'une transformation perspective, il permet aussi de reconstruire la forme de surfaces avec des courbures non nulles.**

Dans cette section nous avons vérifié la capacité de notre algorithme à traiter des surfaces convexes. Dans la section suivante, nous appliquons notre méthode à des surfaces plus complexes.

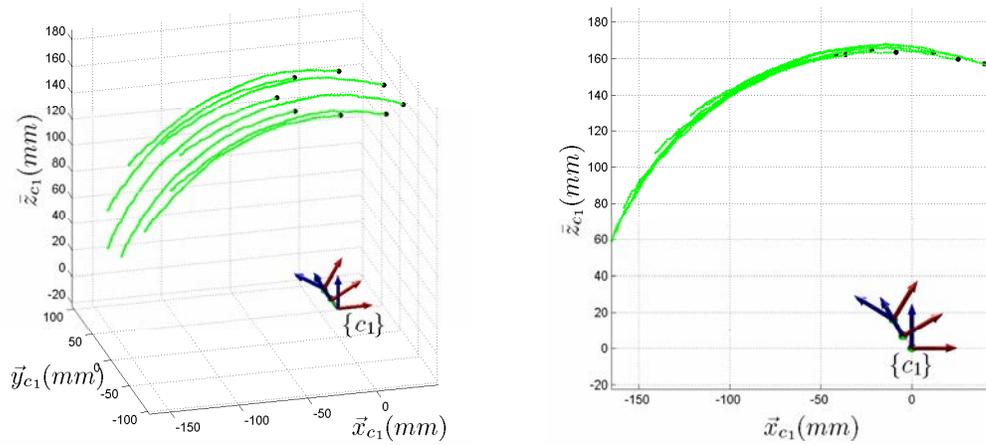


FIGURE. 5.31 – Affichage sous deux différents points de vue du nuage de points 3D calculé pour la séquence d'images acquises avec le fantôme \mathcal{X}_2 selon le scénario IV. Ce nuage est composé de 1200 points 3D. Les repères 3D représentent le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

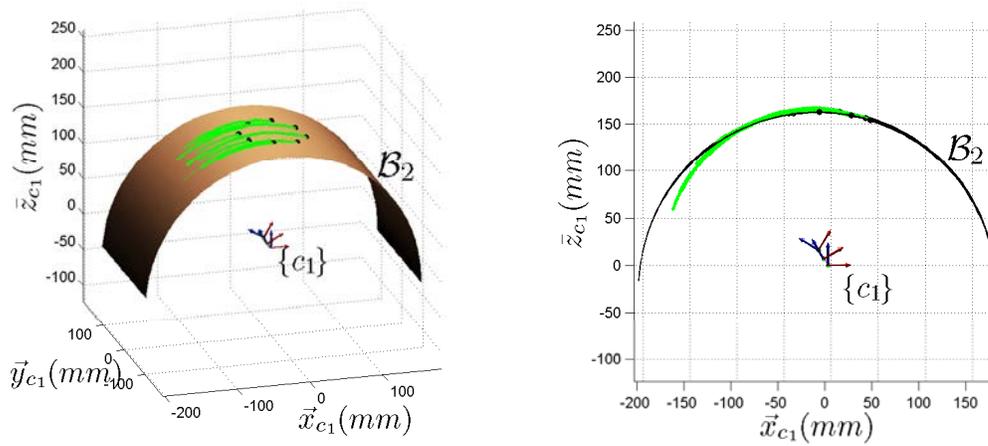


FIGURE. 5.32 – Positionnement de la surface théorique $\{\mathcal{B}_2\}$ sur les 8 points 3D de la première acquisition. $\{\mathcal{B}_2\}$ est présenté marron alors que le nuage de points 3D est en vert.

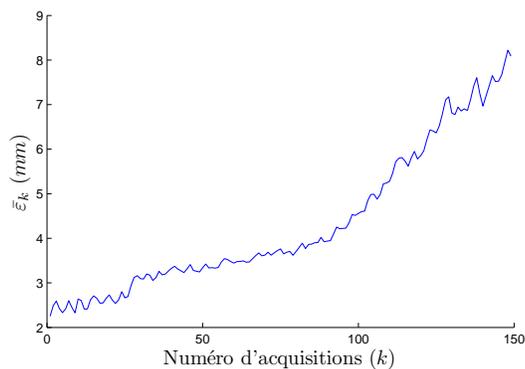


FIGURE. 5.33 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ en fonction du numéro k de l'acquisition (fantôme \mathcal{X}_2 , scénario IV). Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.

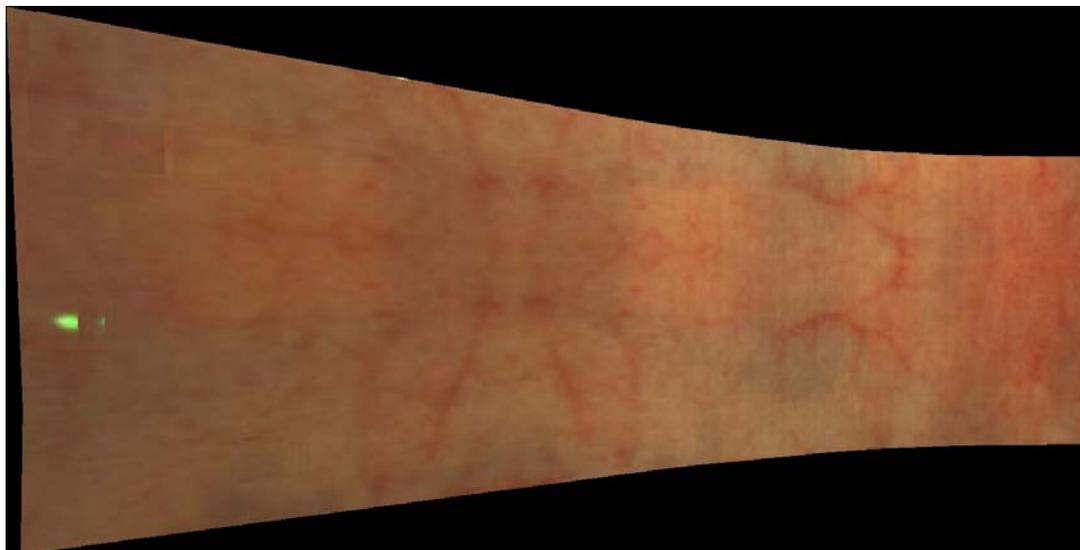


FIGURE. 5.34 – Mosaique obtenue pour la séquence de 150 acquisitions réalisées avec le fantôme \mathcal{X}_2 selon le scénario IV.

5.2.3 Résultats obtenus pour le fantôme en forme de vagues

5.2.3.1 Scénario I

Nous réalisons pour cette expérience une séquence de 145 acquisitions avec le fantôme \mathcal{X}_3 selon le scénario I. Comme pour les précédentes expérimentations réalisées selon ce scénario, le fantôme est placé à une distance approximative de 100 *mm* en face du prototype expérimental et le mouvement effectué est formé essentiellement de translations selon l'axe \vec{x}_k de la caméra. La surface imagée à cette distance est de l'ordre de $60 \times 60 \text{ mm}^2$ et couvre approximativement un creux et une crête (donc une période de sinusoïde, voir figure 5.35). Cette figure montre le nuage de points 3D qui résulte du traitement de la séquence. **On constate que le champ de vue a doublé puisque la surface totale couverte par le nuage de points 3D est de l'ordre de $80 \times 160 \text{ mm}^2$. On remarque également que le parcours du prototype (voir les repères 3D de la figure 5.35) est composé essentiellement de translation selon l'axe \vec{x}_k de la caméra.** Le parcours obtenu correspond donc bien au mouvement 3D que nous avons appliqué au prototype pour l'acquisition de cette séquence.

La figure 5.36 montre la superposition de la surface théorique $\{\mathcal{B}_3\}$ qui correspond au fantôme \mathcal{X}_3 sur les 8 points 3D de la première acquisition. Ces points 3D sont représentés en noir alors que le reste des points du nuage est visualisé en vert. **Nous remarquons que les points 3D préservent globalement la forme « sinusoïdale » du fantôme.** Nous quantifions l'éloignement entre les points 3D et la surface théorique $\{\mathcal{B}_3\}$ avec la mesure d'erreur $\varepsilon_{i,k}$ qui représente la distance entre chaque point $P_{3D}^{i,k}$ et la surface théorique (voir équation (5.1)). La valeur moyenne et la valeur maximale de cette mesure valent respectivement 1.8 *mm* et 4.4 *mm*. Pour la configuration testée, l'erreur moyenne vaut donc 1.75% de la distance de mesure (100 *mm*). L'erreur $\bar{\varepsilon}_k$ quantifie le cumul des erreurs de positionnement des points $P_{3D}^{i,k}$ dans le repère $\{c_1\}$ du premier point de vue (voir équation (5.2)). L'évolution de la valeur de cette erreur est illustrée dans la figure 5.37. Nous constatons que les erreurs ne varient pas de façon monotone. Elles augmentent jusqu'à l'acquisition 98 et atteignent 2.58 *mm*. Ensuite, elles diminuent jusqu'à l'acquisition 124 et atteignent 1.3 *mm*.

La mosaïque 2D générée avec les 135 images de la séquence est donnée dans la figure 5.38. Elle a pour taille 432×1111 pixels. La mosaïque affiche une bonne cohérence visuelle. En effet, les textures visibles dans la mosaïque sont cohérentes car sans discontinuités.

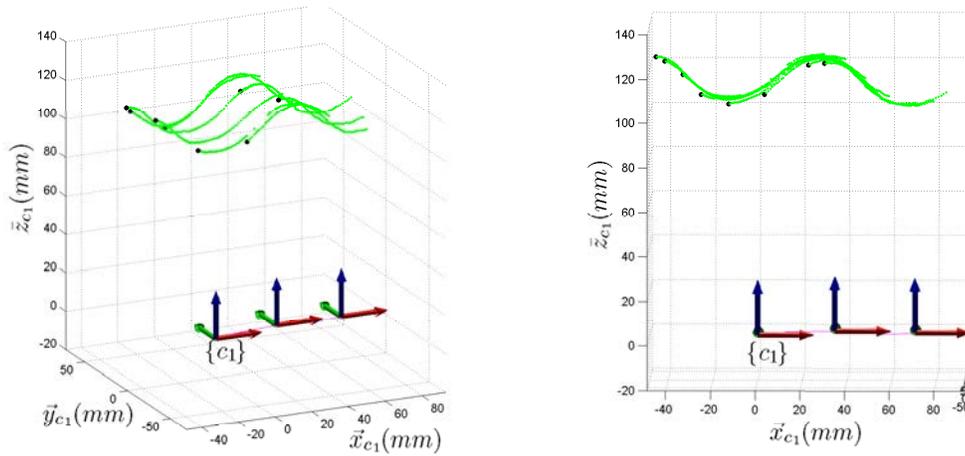


FIGURE. 5.35 – Affichage sous deux différents points de vue du nuage de points 3D calculés pour le fantôme \mathcal{X}_3 selon le scénario I. Ce nuage est composé de 1160 points 3D tous ramenés (exprimés) dans le repère de la première acquisition. L'origine de ce repère a les coordonnées $(0,0,0)$. Les repères 3D représentent le parcours du prototype d'acquisition tout au long de la séquence. On remarque que la trajectoire du prototype n'est pas exactement sur une droite 3D. Cependant, la translation selon l'axe \vec{x}_k représente l'essentiel du mouvement.

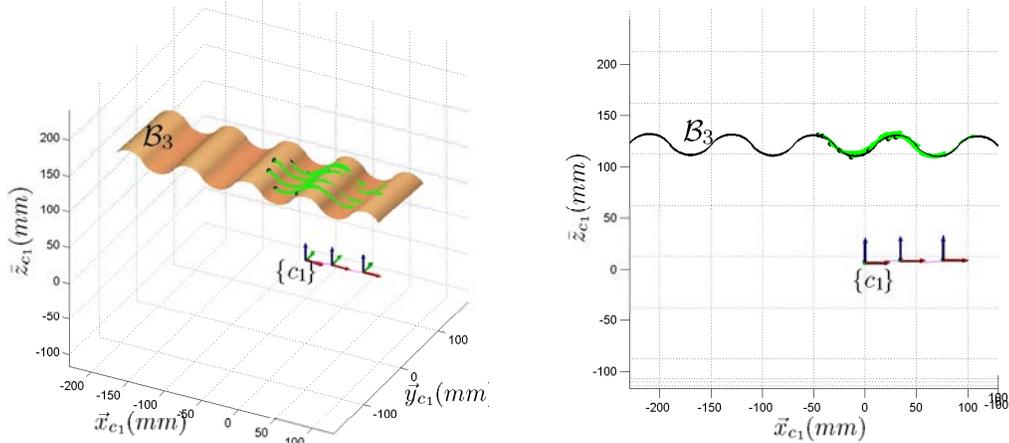


FIGURE. 5.36 – Positionnement de la surface théorique $\{\mathcal{B}_3\}$ sur les 8 points 3D de la première acquisition (points noirs). La surface $\{\mathcal{B}_3\}$ est en marron alors que le nuage de points 3D est en vert.

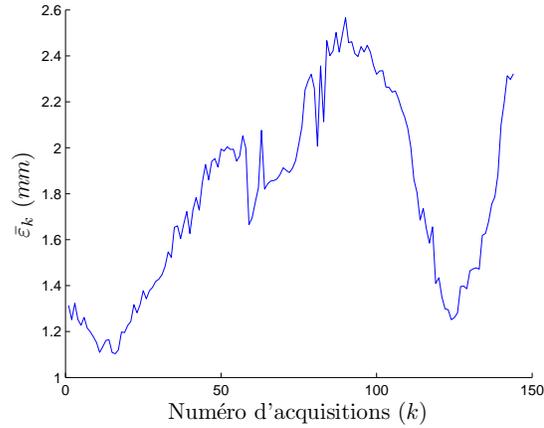


FIGURE. 5.37 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ (erreurs calculées pour le scénario I et le fantôme \mathcal{X}_3) en fonction du numéro k du point de vue. Cette courbe représente l'erreur moyenne lorsqu'on ramène les points de la $k^{\text{ème}}$ acquisition dans le repère de la première acquisition.

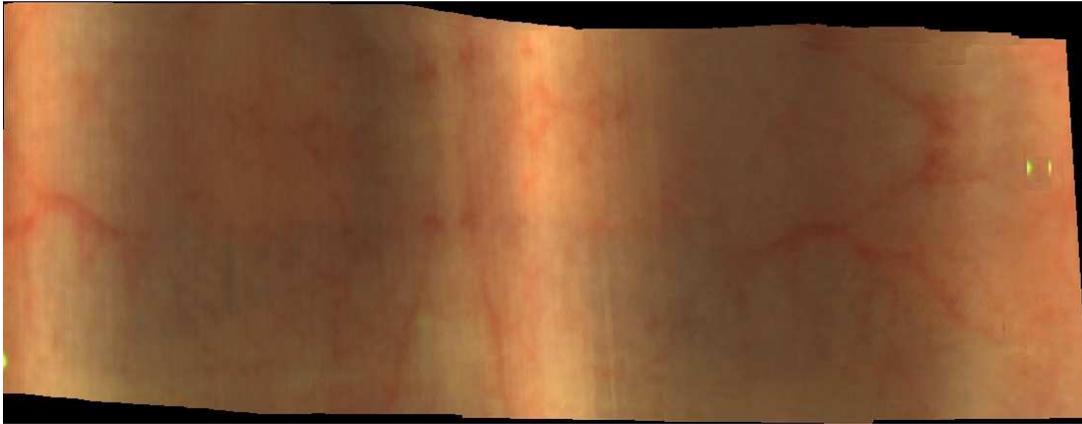


FIGURE. 5.38 – Mosaïque obtenue pour la séquence de 160 acquisitions réalisées avec le fantôme \mathcal{X}_3 selon le scénario I.

5.2.3.2 Scénario II

Pour tester ce scénario sur le fantôme \mathcal{X}_3 , nous avons réalisé une séquence de 85 acquisitions. Le déplacement 3D appliqué à cette séquence est composé essentiellement de rotations autour de l'axe \vec{y}_k de la caméra. Le nuage de points 3D obtenu pour cette séquence est montré dans la figure 5.39. Visuellement, la forme 3D du points 3D est cohérente avec la forme en sinusoïde du fantôme. D'autre part, les repères visualisés sur cette même figure, montre que le mouvement 3D estimé pour la séquence correspond effectivement à succession de rotation autour de la l'axe \vec{y}_k de la caméra. La figure 5.40 montre la différence de forme entre le nuage de points 3D généré et la surface théorique $\{\mathcal{B}_3\}$. Cependant, ces différences sont faibles. Les erreurs de placement de points 3D par rapport la surface référence $\{\mathcal{B}_3\}$ sont données par la mesure d'erreur $\varepsilon_{i,k}$. La valeur moyenne et maximale sont respectivement égales à 2.3 mm et 7.3 mm . L'erreur moyenne de placement des points dans le repère $\{c_1\}$ vaut donc 1.3% de la distance de mesure. D'autre part, l'évolution de la mesure d'erreur $\bar{\varepsilon}_k$ est donnée dans la figure 5.40. Cette erreur s'accroît pour atteindre sa valeur maximale 3.4 mm au niveau de l'acquisition 65 de la séquence.

La mosaïque générée pour la séquence est donnée dans la figure 5.42. Elle a une taille de 451×788 pixels. Le spot laser que nous remarquons dans cette mosaïque provient de la première image et n'a pu être « dissimulé » avec les pixels d'autre images. La mosaïque présente une très bonne qualité visuelle car n'incluant pas de discontinuités de texture décelables à l'œil nu.

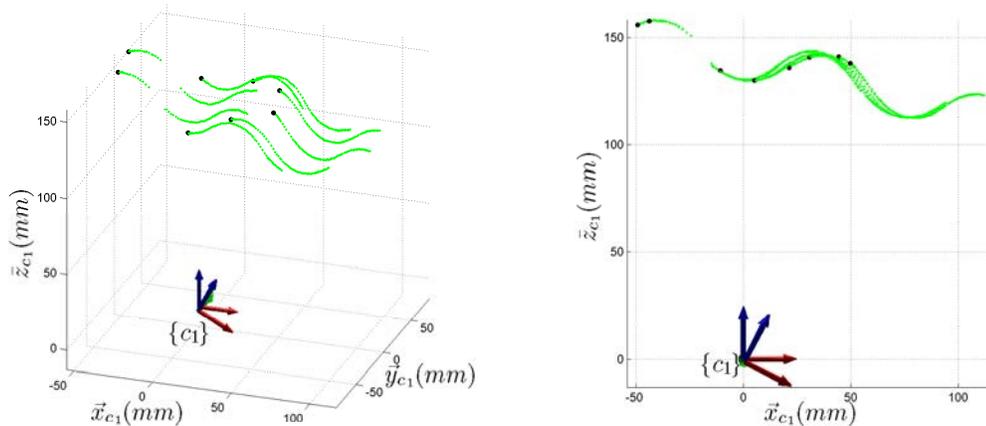


FIGURE. 5.39 – Affichage sous deux différents points de vue du nuage de points 3D calculé pour le fantôme \mathcal{X}_3 selon le scénario II. Ce nuage est composé de 680 points 3D. Les repères 3D représentent à nouveau le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

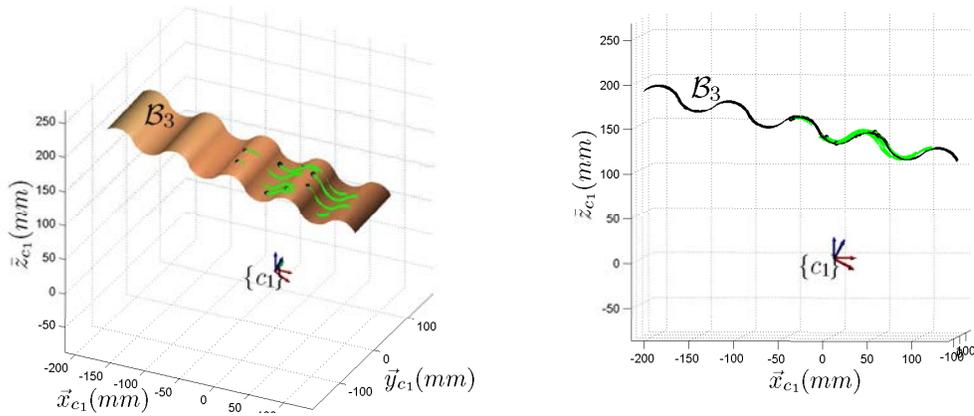


FIGURE. 5.40 – Positionnement de la surface théorique $\{B_3\}$ sur les 8 points 3D de la première acquisition. $\{B_3\}$ est en marron alors que le nuage de points 3D est en vert.

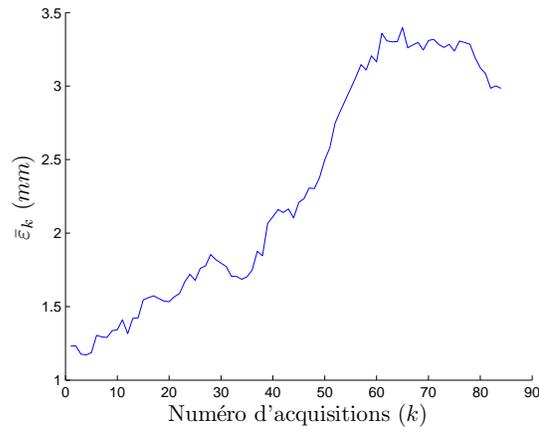


FIGURE. 5.41 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ en fonction du numéro k du pont de vue. Ces erreurs sont calculées pour \mathcal{X}_3 et le scénario II.

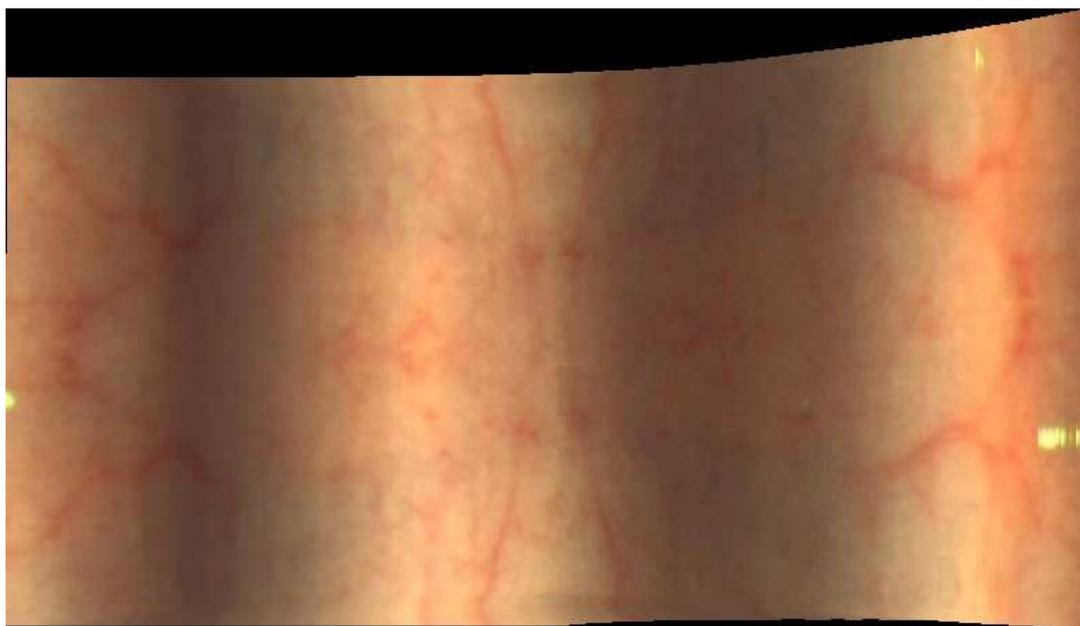


FIGURE. 5.42 – *Mosaïque obtenue pour la séquence de 85 acquisitions réalisées avec le fantôme \mathcal{X}_3 selon le scénario II.*

5.2.3.3 Scénario III

La troisième expérience réalisée pour le fantôme \mathcal{X}_3 consiste en l'approximation d'un déplacement 3D composé essentiellement de translations selon \vec{z}_k (l'axe optique de la caméra) pour effectuer des zooms. Pour réaliser une séquence de 40 acquisitions avec le fantôme \mathcal{X}_3 , nous avons placé le prototype à une distance approximative de 200 mm de la face texturée du fantôme. Ensuite, nous avons rapproché le prototype de la surface en le déplaçant selon l'axe optique de la caméra jusqu'à atteindre une distance de 130 mm. Le nuage de points 3D obtenu pour cette séquence est présenté dans la figure 5.43. Cette figure montre que le chemin emprunté par le prototype est effectivement composé essentiellement de translations selon l'axe \vec{z}_k .

Les erreurs $\bar{\varepsilon}_k$ obtenues sont représentées sur la figure 5.45. L'erreur moyenne et le maximum des erreurs $\varepsilon_{i,k}$ valent respectivement 2.3 mm et 8.3 mm pour des distances de mesure allant de 130 à 200 mm. En pourcentage de la distance de mesure, l'erreur moyenne vaut entre 1.7% et 1.15%.

Il est visible sur la figure 5.44, que certains points 3D du nuage s'éloignent de façon sensible de la surface de référence $\{\mathcal{B}_3\}$. Même si le parcours estimé semble cohérent avec le scénario appliqué, le cumul d'erreurs de positionnement est très important (voir figure 5.45). Ces erreurs affectent sensiblement la qualité visuelle de la carte 2D générée et montrée dans la figure 5.46.

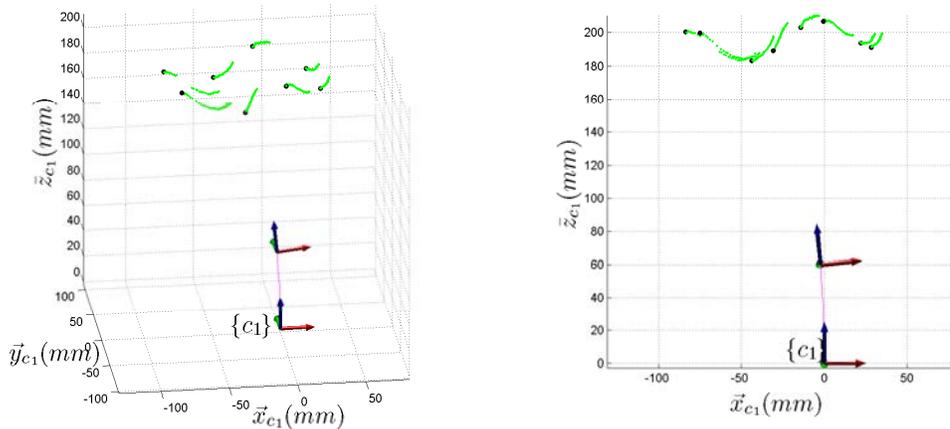


FIGURE. 5.43 – Affichage sous deux différents points de vue du nuage de points 3D obtenu par le traitement de la séquence d'images acquises pour le fantôme \mathcal{X}_3 selon le scénario III. Ce nuage est composé de 320 points 3D. Les repères 3D représentent le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

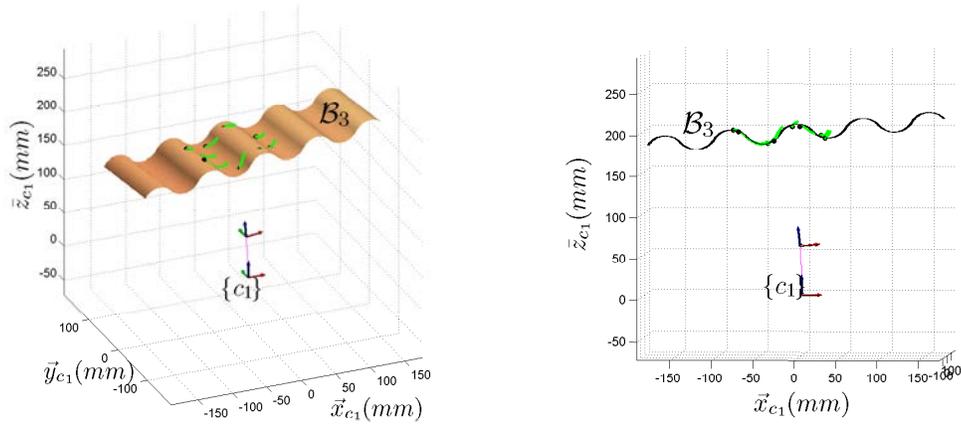


FIGURE. 5.44 – Positionnement de la surface théorique $\{B_3\}$ sur le nuage de points 3D en se référant aux 8 points 3D de la première acquisition. $\{B_3\}$ est affiché en couleur marron alors que celui le nuage de points 3D est en vert.

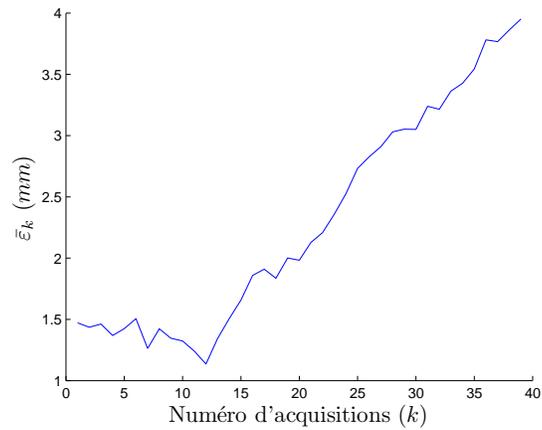


FIGURE. 5.45 – Évolution de la valeur de l'erreur \bar{e}_k en fonction du numéro k de l'acquisition (fantôme \mathcal{X}_3 , scénario III).



FIGURE. 5.46 – Mosaïque obtenue pour la séquence de 40 acquisitions réalisées avec le fantôme \mathcal{X}_3 selon le scénario III.

5.2.3.4 Scénario IV

La dernière expérience réalisée avec le fantôme \mathcal{X}_3 consiste à tester un déplacement 3D qui combine les mouvements des trois premiers scénarios, à savoir des translations, des rotations et des zooms. La séquence réalisée comporte 46 acquisitions. Le prototype est placé face au fantôme et est déplacé manuellement de telle façon à obtenir un mouvement composé essentiellement de translations selon \vec{x}_k et \vec{z}_k et de rotations autour de \vec{y}_k . Le nuage de points 3D obtenu pour cette expérience est donné dans la figure 5.47. Le parcours du prototype estimé par notre méthode est en accord avec le déplacement appliqué manuellement lors de la réalisation de la séquence. En effet, les changements de position symbolisés par les repères de la figure 5.47 font apparaître des translations selon les axes \vec{x}_k et \vec{z}_k et de faibles rotations autour de l'axe \vec{y}_k . La figure 5.48 met en lumière la différence de forme entre le nuage de points 3D et la surface théorique $\{\mathcal{B}_3\}$. Sur cette figure, on remarque que les points 3D commencent à s'éloigner de la surface de référence au bout de quelques acquisitions (approximativement à partir de la 20^{ème} acquisition d'après la figure 5.49). Ces erreurs sont également traduites par la mesure $\bar{\varepsilon}_k$ qui augmente tout au long de la séquence pour atteindre 12.34 mm. La valeur moyenne et la valeur maximale de l'erreur $\varepsilon_{i,k}$ sont respectivement égales à 5.32 mm et 19.27 mm. L'image panoramique qui correspond à cette séquence a une taille de 500 × 944 pixels (voir figure 5.50). Sur cette mosaïque on peut remarquer qu'il y a de légères discontinuités dans les textures de l'image. Ce fait est lié au cumul des erreurs de l'estimation des transformations 3D et perspectives.

Les premiers enseignements à tirer des tests réalisés avec le fantôme \mathcal{X}_3 sont que **les déplacements simples (translations ou rotations) permettent de reconstruire avec précision la forme sinusoïdale du fantôme. Cette reconstruction de surface est plus imprécise dans le cas de zoom et surtout pour un mouvement libre qui combine tous les types de déplacement. Ce fantôme est cependant un cas extrême, la vessie étant typiquement déformée par un seul creux (du à la poussée d'un autre organe contre les parois de la vessie) et non pas une succession de creux. Les résultats obtenus pour les translations et les rotations restent encourageants pour ce cas extrême.**

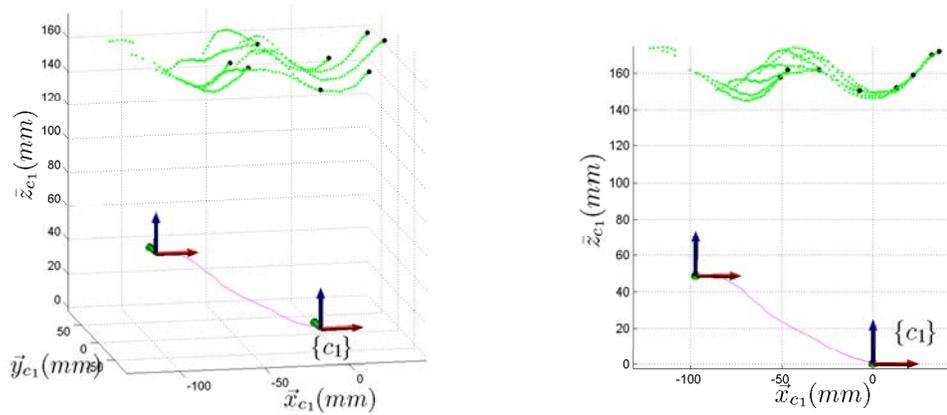


FIGURE. 5.47 – Affichage sous deux différents points de vue du nuage de points 3D calculé pour la séquence d'images acquises avec le fantôme \mathcal{X}_3 selon le scénario IV. Ce nuage est composé de 368 points 3D. Les repères 3D représentent le parcours du prototype d'acquisition tout au long de la séquence. Les points noirs sont ceux de la première acquisition.

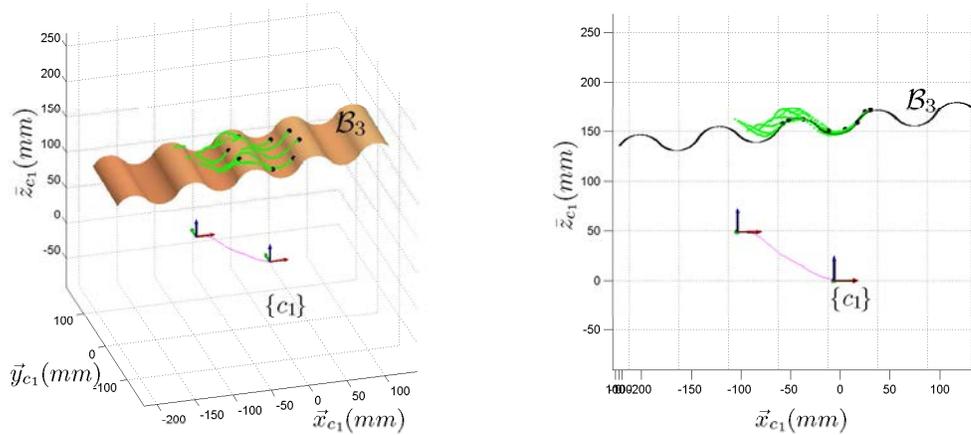


FIGURE. 5.48 – Positionnement de la surface théorique $\{\mathcal{B}_3\}$ sur les 8 points 3D de la première acquisition. $\{\mathcal{B}_3\}$ est affiché en couleur marron alors que celui le nuage de points 3D est en vert.

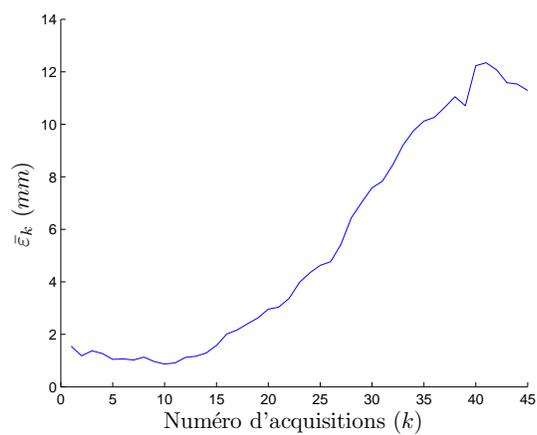


FIGURE. 5.49 – Évolution de la valeur de l'erreur $\bar{\epsilon}_k$ en fonction du numéro k de l'acquisition (fantôme \mathcal{X}_3 , scénario IV).



FIGURE. 5.50 – Mosaïque obtenue pour la séquence de 46 acquisitions réalisées avec le fantôme \mathcal{X}_3 selon le scénario IV.

5.3 Synthèse des résultats

Cette section a pour but d'analyser les résultats obtenus pour les différents fantômes et scénarios.

Synthèse générale. En termes de cartographie 3D, les résultats obtenus sont globalement satisfaisants. En effet, dans le cas de la vessie il s'agit de reconstruire une surface avec une forme réaliste, l'objectif n'étant pas de reconstruire des points d'une surface avec une grande précision. Excepté la reconstruction du fantôme en forme de vagues avec les données acquises suivant les scénarios III et IV, la forme des surfaces calculées est non simplement réaliste, mais de plus, les points 3D sont placés avec précision sur la carte 3D. Cela conduit à des surfaces précises dans la plupart des cas testés. Hormis pour les scénarios III et IV appliqués au fantôme en forme de vagues, les erreurs calculées varient entre 1.5% et 4% des distances de mesure (distances entre la caméra du prototype et les surfaces imagées). Cette précision va au-delà des exigences requises pour l'application de la vessie.

En termes de cartographie 2D, nous avons constaté qu'à l'exception des deux tests réalisés avec le fantôme en forme de vagues selon les scénarios III et IV, les images panoramiques générées affichent une cohérence visuelle car elles n'incluent pas de discontinuités de textures perceptibles à l'œil nu. Notons de plus que l'utilisation d'un projecteur de spots laser n'a pas dégradé la qualité des images panoramiques générées. En effet, les spots laser sont segmentés et masqués automatiquement dans les mosaïques (les images consécutives se recouvrant, les couleurs vertes des pixels illuminés par des spots laser peuvent être remplacées par les couleurs de la paroi de la vessie des pixels homologues d'autres images).

Planéité des surfaces imagées. Le degré de courbure des surfaces imagées est un aspect important dans notre étude car c'est l'hypothèse émise pour représenter mathématiquement le mouvement 2D entre deux images consécutives par un modèle perspective. Plus la surface imagée par acquisition est plane, plus le modèle perspectif est pertinent. Parmi les trois fantômes utilisés, le fantôme plan justifie parfaitement l'utilisation du modèle perspectif : quelle que soit la distance de mesure, la surface acquise reste plane. Par contre, selon la distance de mesure entre le prototype et le fantôme en forme de vagues, les surfaces imagées ne sont plus planes. Cela justifie la différence de précision entre les expériences de ces deux fantômes. L'erreur moyenne de construction du nuage de points 3D varie entre 0.46 *mm* et 1.80 *mm* pour le fantôme \mathcal{X}_1 , et entre 1.79 *mm* et 5.32 *mm* pour le fantôme en forme de vagues. Pour le fantôme cylindrique, qui est un cas moins

extrême et d'une difficulté plus proche de la réalité, la dimension de la surface imagée dépend de la distance entre le prototype et le fantôme. Plus cette distance est grande plus la surface imagée est grande et par conséquent moins elle est plane. Pour les expériences réalisées avec ce fantôme, nous avons essayé de respecter le rapport entre la surface imagée et la distance de la mesure (pour une vessie, si la distance de mesure varie de 1 à quelques centimètres, la surface imagée varie de 1 à quelques centimètres près). En respectant à une échelle près cette géométrie, nous nous sommes placés dans des conditions réalistes de mesures. En effet, dans cette situation nous avons simulé des courbures du même ordre de grandeur que dans la vessie. Les résultats obtenus avec le fantôme cylindrique montrent que l'approximation des surfaces imagées par un plan est assez précise pour que les mouvements 3D estimés garantissent le calcul d'un nuage de points 3D conduisant à une surface de forme réaliste.

Mouvements 3D. Pour évaluer notre approche de cartographie 3D, nous avons considéré quatre scénarios plus ou moins réalistes pour simuler les mouvements qu'applique un clinicien au cystoscope lors d'un examen. Nous avons constaté qu'indépendamment du fantôme avec lequel les acquisitions sont réalisées, les scénarios formés de mouvements simples (scénario I par exemple) conduisent aux meilleures performances en terme de précision de construction des nuages de points. En effet, nous remarquons sur les figures 5.4, 5.20 et 5.36 que les nuages de points 3D sont très proches des surfaces théoriques des fantômes. En revanche, plus on augmente la complexité (combinaisons de mouvement) du déplacement du prototype, plus les nuages de points 3D s'éloignent des surfaces théoriques (voir figure 5.16, 5.32 et 5.48). Cependant, la vessie étant un organe mou et déformable, la précision de la reconstruction 3D obtenue pour le scénario IV reste largement suffisante car conduisant toujours à des formes représentatives de la surface.

Temps de calcul. À ce stade de l'étude, le temps nécessaire pour la construction d'une carte 3D est le moins important des critères d'évaluation. En effet, pour une application visant un archivage « efficace »¹⁵ des données d'examen cystoscopiques, il n'y a pas de contrainte de temps de calcul. Le calcul des cartes peut se faire d'une manière « hors-ligne », c'est-à-dire après la fin des consultations. Cependant, pour que les cartes 3D servent à un second diagnostic durant la consultation même, la construction des vues étendues doit être réalisée une quinzaine de minutes qui suivent l'intervention. Pour le moment, l'algorithme proposé traite deux acquisitions consécutives en 35 secondes en

15. Il est rappelé que les cartes 3D autorisent l'exploitation des données archivées par tous les urologues (et non seulement par le clinicien qui a réalisé l'examen). Ceci facilite le suivi d'un patient.

moyenne¹⁶. Ce temps de calcul correspond à une soixantaine d'itérations de l'algorithme du simplexe. Toutefois, les temps de calcul peuvent être diminués dans de futurs travaux en améliorant le processus d'optimisation de notre méthode. Une solution serait de réaliser ce processus d'optimisation en deux temps. D'abord, une première solution est estimée avec une méthode d'optimisation qui se base sur le calcul numérique de gradients (par exemple la méthode du gradient conjugué). Cette première solution sera ensuite utilisée pour initialiser la méthode du simplexe à fin d'obtenir une solution précise. Nous pensons également que nous pouvons encore gagner en terme de temps de calcul en optimisant le code d'implémentation de nos algorithmes par exemple avec une programmation « *Multithreading* ».

5.4 Conclusion et travaux restant à faire

Nous avons présenté dans ce chapitre une évaluation de la méthode de cartographie détaillée dans le chapitre 4. Cette évaluation consiste à tester quantitativement si notre méthode retrouve les formes des surfaces imagées. Nous avons utilisé un ensemble de trois fantômes plus ou moins réalistes. Les formes et les dimensions de chacun des fantômes sont connues au millimètre près. Les expérimentations entreprises sont réalisées selon des scénarios de déplacements 3D qui simulent les mouvements du cystoscope lors d'un examen clinique. Les résultats obtenus pour les différents fantômes et scénarios montrent que notre méthode est potentiellement adaptée à la cartographie des surfaces internes de la vessie¹⁷. En effet, pour la plupart des expériences réalisées, notre méthode a généré des nuages de points situés sur la surface des fantômes (les erreurs de placement des points sur la surface varient entre 1.5 % et de 4% de la distance de mesure). Il est à noter que dans le cas de la vessie, nous ne cherchons pas à réaliser une reconstruction extrêmement précise de ses parois. Nos résultats sont acceptables et très prometteurs. Notre travail est le premier qui a pour objectif la reconstruction de carte 3D pour les parois internes de la vessie.

Implémentation logicielle. Pour que notre travail soit totalement finalisé il faudra cependant encore traiter les points suivants :

- Calculer, à partir des traînées des 8 points laser, une surface dense de points 3D. Une solution serait par exemple de faire une interpolation bicubique des points ramenés

16. La machine sur laquelle nous avons effectué les expériences est équipée de 2Go de mémoire vive et d'un processeur Intel de 2.4Ghz.

17. Des tests sur des données patient devraient confirmer ces résultats

dans le repère de la carte.

- Projeter les textures sur les surfaces ainsi reconstruites. Ce dernier point est sans difficultés, car la position des caméras dans le repère de la carte est connue pour chaque acquisition. Cette information peut être utilisée pour projeter les couleurs des images leur correspondant sur la surface interne reconstruite.

Ces deux points ne posent pas de difficulté méthodologique mais relèvent d'une implémentation logicielle finalisée.

Défis méthodologiques. Du point de vue du traitement de données, nous savons reconstruire des parties de surfaces étendues de vessie qui peuvent contenir un repère anatomique et les lésions. Cependant, à l'instar de la cartographie 2D d'autres problèmes restent à résoudre, *a fortiori* lorsque toute ou presque toute la paroi de la vessie doit être reconstruite. Ces problèmes viennent du fait que le cystoscope peut former des trajectoires qui se recoupent lorsque le clinicien revient à des mêmes points de vue dans la vessie. Dans ce cas, les résultats de l'algorithme conduisent à des discontinuités de surfaces 3D et donc de textures. Pour atténuer (lisser) ces discontinuités, on pourra développer un algorithme d'optimisation dont le critère sera lié à la fois à la contrainte de forme des surfaces et à la continuité des textures. Cet algorithme devra « gommer » les discontinuités tout en préservant une forme réaliste de surface.

Conclusion et perspectives

Les travaux réalisés dans le cadre de cette thèse représentent un premier pas vers la cartographie tridimensionnelle (3D) des parois internes de la vessie. Une représentation 3D de l'intérieur de la vessie va faciliter l'interprétation des données acquises lors d'un examen cystoscopique et permettra la « navigation » virtuelle dans l'organe si celui-ci est reconstruit en entier. Reconstruire des champs de vue larges des parois peut conduire à de nouveaux critères de diagnostic comme des indices sur la distribution spatiale des lésions qui sont très utiles pour cerner l'évolution des lésions d'un examen à un autre (examens réalisés pour un même patient à plusieurs semaines d'intervalle).

Contributions majeures de la thèse

Un travail en cours au CRAN réside dans le développement d'un banc optique devant valider la reconstruction de points 3D avec un cystoscope à vision active. Ce type d'instrument fournira, pour chaque point de vue et en complément des images 2D, quelques points situés à la surface interne de l'organe. Les travaux réalisés dans cette thèse complètent ce développement instrumental et s'articulent principalement autour de deux axes. Une partie de nos travaux a porté sur le calibrage du système à vision active pour la reconstruction de points 3D. La deuxième partie de cette thèse concernait l'exploitation des données disponibles (image 2D et un ensemble de paires de points 3D et 2D) pour la construction d'un nuage de points 3D d'une part, et d'une image panoramique 2D d'autre part. Ce nuage de points 3D et cette image panoramique correspondent respectivement à la surface 3D et la texture de l'ensemble de la surface imagée.

Calibrage du système à vision active et reconstruction 3D. Nous avons développé une nouvelle méthode de calibrage de système à vision active composé d'une caméra et d'un projecteur de points lasers (dans notre cas, huit spots formant un cercle et projeté à la périphérie des images cystoscopiques). La méthode proposée modélise la trajectoire de chaque spot laser dans l'espace par une droite 3D. L'avantage de notre méthode réside

dans sa flexibilité. En effet, la méthode ne nécessite qu'un simple plateau de calibrage constitué d'une mire imprimée sur une feuille fixée sur un plan (un plateau en PVC dans nos expériences). Il est à préciser qu'aucun équipement de positionnement précis n'est nécessaire pour notre méthode. Le plateau de calibrage est déplacé manuellement selon l'axe optique de la caméra pour acquérir au minimum deux images. Nous avons montré que notre méthode permet une reconstruction de points 3D avec une grande précision (l'erreur de reconstruction 3D vaut approximativement 0.07% de la distance de mesure). Un autre avantage de la méthode réside dans l'adaptabilité à des projecteurs de points formant différents motifs. Dans notre cas, les points du motif lumineux forment un cercle, mais la méthode fonctionne aussi pour un ensemble de points non régulièrement espacés, voire situés aléatoirement dans le plan. Il est à noter également que même si la méthode a été conçue en premier lieu pour le calibrage d'un système cystoscopique à vision active, elle pourrait être utilisée pour le calibrage d'autres systèmes à vision active dans autres domaines d'application utilisant la reconstruction 3D par vision active (par exemple pour la mesure dimensionnelle de pièces industrielles).

La méthode de calibrage que nous avons proposée fournit les paramètres nécessaires à la reconstruction de la position 3D du centre de chaque spot laser détecté dans les images. Les données 3D/2D disponibles à chaque instant de la séquence sont donc formées d'une image couleur et d'un ensemble de huit paires de points 2D (centres des spots laser) et 3D (localisés sur la surface). Il est à souligner que les données 2D/3D fournies pour une acquisition donnée sont exprimées dans les repères « locaux » liés à la position du cystoscope au moment de l'acquisition en question. Il s'agit précisément du repère 2D image et du repère 3D lié à la caméra. L'ensemble de ces points 2D/3D disponibles tout au long d'une séquence doit être ramené dans des repères « globaux » (repères image et caméra de la première acquisition) pour représenter, en une seule fois, la totalité de la portion de surface imagée dans la séquence. C'est l'objet de la deuxième contribution majeure de cette thèse.

Cartographie 3D de surfaces étendues de la paroi interne de la vessie. La construction d'une carte 3D nécessite à la fois le calcul d'un nuage de points 3D situés sur la surface de la paroi interne de la vessie et la détermination d'une image panoramique 2D représentant la texture de l'ensemble de la surface imagée. La reconstruction du nuage de points 3D permettra (par l'interpolation des points par des splines bicubiques) de générer une surface 3D sur laquelle l'image panoramique 2D pourra être projetée.

Nous avons proposé une nouvelle méthode d'estimation simultanée d'une transformation 3D rigide ($\mathbf{T}_{3D}^{k-1,k}$) et d'une transformation 2D perspective ($\mathbf{T}_{2D}^{k-1,k}$) (reliant géométri-

quement les données des acquisitions aux instants k et $k - 1$) adaptée aux spécificités des données 3D/2D dont nous disposons. Cette méthode repose sur un système d'équations linéaires qui relie les paramètres des deux transformations. Les transformations $\mathbf{T}_{3D}^{k-1,k}$ estimées entre chaque paire d'acquisitions consécutives permettent de construire le nuage de points 3D en plaçant l'ensemble de points reconstruits pour toute une séquence dans un repère global (celui de la caméra à la première acquisition dans cette thèse). À l'instar de la construction du point 3D, les transformations perspectives $\mathbf{T}_{2D}^{k-1,k}$, reliant les régions homologues d'images consécutives, sont utilisées pour construire une image panoramique en ramenant les pixels de toutes les images acquises d'une séquence dans le repère image de la première acquisition.

À ce jour, le principe de la vision active a été validé dans le cas général de l'endoscopie [Chan *et al.*, 2003], cependant il n'a pas été testé dans le cas particulier de la cystoscopie. En l'absence de données cliniques, nous avons éprouvé notre méthode de cartographie 3D avec des données respectivement simulées et acquises à partir de fantômes numériques et réels de vessies. Nos fantômes numériques sont construits à partir d'un modèle 3D générique de vessie. Les fantômes réels (physiques) sont construits à partir de supports de différentes formes géométriques dont les dimensions sont connues au millimètre près. L'ensemble de ces fantômes numériques et réels a été tapissé avec des textures réalistes obtenues à partir de séquences cystoscopiques cliniques ou à partir de photographies des parois internes d'une vessie de porc.

Les expériences effectuées avec les fantômes numériques ont mis en lumière la grande précision de notre méthode d'estimation des transformations 3D et perspective entre paires d'acquisitions consécutives. En effet, les erreurs de positionnement des points 3D, reconstruits pour une acquisition k , dans le repère 3D de l'acquisition $k - 1$ valent approximativement 0.058 mm . D'autre part, les erreurs d'estimation des transformations perspectives entre chaque paire d'images consécutives valent environ 0.15 pixel. Les expériences réalisées avec les fantômes numériques ont démontré la robustesse de la méthode vis-à-vis de la grande variabilité de textures rencontrée dans les séquences d'images cystoscopiques. Nous avons aussi constaté que notre méthode témoigne d'une grande précision, même lorsque le déplacement 3D de l'instrument entre deux acquisitions est plus important que ceux rencontrés classiquement en conditions cliniques.

Les expériences menées avec les fantômes réels sont réalisées selon des scénarios de déplacements 3D définis qui simulent des mouvements de cystoscopes lors d'examen cystoscopiques. Les résultats obtenus pour les différents fantômes et scénarios montrent que notre méthode est potentiellement adaptée à la cartographie des surfaces internes de la vessie. En effet, pour la plupart des expériences réalisées, notre méthode a généré des

nuages de points 3D situés à la surface des fantômes (les erreurs de placement des points sur la surface varient entre 1.5 % et de 4% de la distance de mesure). Il est à noter que dans le cas de la vessie, nous ne cherchons pas à réaliser une reconstruction extrêmement précise de ses parois. L'important est plutôt de reconstruire des formes qui correspondent à une réalité. Cet objectif a été atteint avec nos fantômes. Les résultats obtenus à ce stade de l'étude sont acceptables et très prometteurs. D'autre part, la cohérence visuelle des images panoramiques 2D générées dans la plupart des expériences témoigne de la précision de notre méthode. De plus, même si les spots laser sont visibles sur les images cystoscopiques, ils sont automatiquement dissimulés dans les images panoramiques pour garder le maximum d'informations de textures sur les parois explorées. Pour finaliser le travail, il faut maintenant passer du nuage de points 3D à une surface représentée par un maillage dense de points sur laquelle sont projetées les textures des images 2D. À ce stade du travail nous sommes capable de représenter des surfaces étendues qui ne couvrent pas la totalité de l'organe, mais qui peuvent contenir l'ensemble d'une lésion.

Perspectives

Les travaux réalisés dans le cadre de cette thèse peuvent être poursuivis dans plusieurs directions. La première direction concerne l'amélioration de la méthode de cartographie 3D. À ce stade du travail nous avons développé une méthode pour reconstruire des parties de surfaces étendues de la vessie. Ces surfaces peuvent contenir un repère anatomique et la lésion. Cependant, d'autres problèmes doivent être résolus, notamment lorsque la totalité ou la quasi-totalité de la paroi de la vessie doit être cartographiée. Dans ce cas, la trajectoire parcourue par le cystoscope se croise ou forme des boucles fermées, notamment aux endroits avec des repères anatomiques auxquels le clinicien revient pour se repérer dans l'organe. Aux endroits avec des croisements, des discontinuités de surfaces et de textures apparaîtraient. Ces discontinuités sont dues au cumul des erreurs lors de la cartographie 3D et 2D. Une solution possible serait de développer un algorithme d'optimisation dont le critère sera lié à la fois à des contraintes de forme des surfaces et à la continuité des textures. Un tel algorithme devra limiter l'effet des discontinuités tout en préservant une forme réaliste des surfaces reconstruites.

En ce qui concerne le temps de calcul, notre méthode d'estimation des paramètres des transformations perspectives nécessite un temps relativement important pour le traitement d'une paire d'acquisitions (35 secondes en moyenne)¹⁸. Nous pensons que les temps de calcul de notre méthode d'optimisation peuvent être diminués. En effet, la méthode

18. Le critère de temps n'est pas le facteur le plus important qui a guidé notre travail.

du simplexe utilisée est précise mais néanmoins lente. Nous pouvons réduire le temps de calcul de la méthode en réalisant une optimisation en deux temps. En premier temps, une première solution est estimée avec une méthode d'optimisation qui se base sur le calcul numérique du gradient (par exemple la méthode du gradient conjugué). Cette première solution, qui nous placerait près de l'optimum, est ensuite utilisée pour initialiser la méthode du simplexe à fin d'obtenir la solution finale. Nous pensons également que nous pouvons encore diminuer les temps de calcul en optimisant le code d'implémentation de nos algorithmes.

La dernière direction concerne le développement d'un outil de visualisation de carte 3D des parois internes de la vessie. En effet, nos travaux permettent de construire un nuage de points 3D localisés sur la surface imagée. Il s'agit maintenant de calculer, à partir de ce nuage de points, une surface dense de points 3D. La solution que nous pouvons envisager est de réaliser une interpolation bicubique du nuage de points 3D. La surface dense à calculer représentera le support final de la carte 3D sur lequel nous pouvons projeter la texture des parois imagées. Cette dernière étape de projection est sans difficulté (purement technique) puisque les positions de la caméra dans repère de la carte 3D sont connues. Les couleurs des images de la séquence peuvent donc être projetées sur le support 3D.



Algorithme d'optimisation de Nelder-Mead

Cette annexe présente un développement détaillé de l'algorithme du simplexe utilisé à deux reprises dans cette thèse. Dans le chapitre 3, l'algorithme du simplexe est utilisé pour optimiser la position 3D du centre optique du projecteur de points laser. Dans le chapitre 4, il a été utilisé pour l'estimation du mouvement 3D entre deux acquisitions cystoscopiques consécutives.

$$\begin{aligned}\tilde{\mathcal{L}} &= \arg \min_{\mathcal{L} \in \mathbb{R}^M} \mathcal{G}(\mathcal{L}) \\ \mathcal{G} &: \mathbb{R}^M \mapsto \mathbb{R}\end{aligned}\tag{A.1}$$

Soit le problème d'optimisation décrit dans l'équation (A.1). \mathcal{L} est le vecteur de paramètres à optimiser selon une fonction coût \mathcal{G} . \mathcal{L} est défini dans \mathbb{R}^M . Pour résoudre ce problème, l'algorithme du simplexe consiste à utiliser un simplexe composé de $M+1$ sommets notés \mathcal{L}_j avec $j \in \{1, \dots, M+1\}$. Le simplexe se déforme d'une manière itérative pour converger vers un minimum local de \mathcal{G} . Les règles essentielles de déformation (déplacement des sommets du simplexe) sont communément appelées : « réflexion », « expansion », « contraction » et « rétraction ». Une succession de ces étapes assure systématiquement la convergence du simplexe vers un minimum de la fonction coût [Press *et al.*, 2007]. Une représentation schématique de ces règles est donnée à titre d'exemple dans la figure A.1 pour un simplexe de 4 sommets.

L'algorithme de la figure A.2 détaille toutes les opérations à effectuer dans l'algorithme que nous avons implémenté en s'inspirant de [Press *et al.*, 2007]. L'algorithme dépend des quatre paramètres α , β , γ et δ qui sont liés respectivement aux règles de

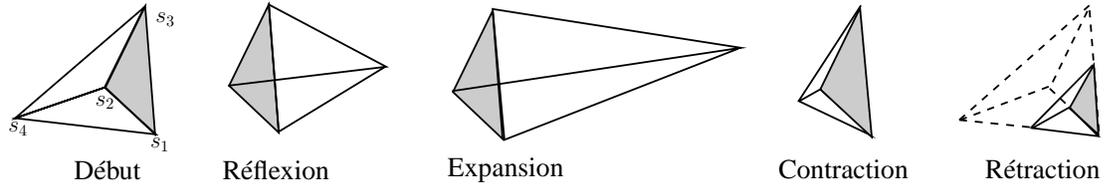


FIGURE. A.1 – Règles de base pour la déformation d'un simplexe à 4 sommets (nombre des paramètres à estimer égal à 3). Les règles sont la réflexion, l'expansion et la contraction et la rétraction. s_1 et s_4 représentent respectivement le meilleur et le pire des sommets du simplexe (avec la meilleure et la plus mauvaise valeur de la fonction coût). Les trois sommets (s_1 , s_2 et s_3) du triangle en gris sont ceux utilisés pour calculer le quatrième et nouveau sommet remplaçant s_4 .

réflexion, de contraction, d'expansion et de rétraction. Les formules qui permettent de calculer ces déformations sont données dans la figure A.2. Les valeurs choisies pour α , β , γ et δ sont les valeurs standards de [Nelder, 2009]. Elles sont respectivement égales à 1, $\frac{1}{2}$, 2 et $\frac{1}{2}$.

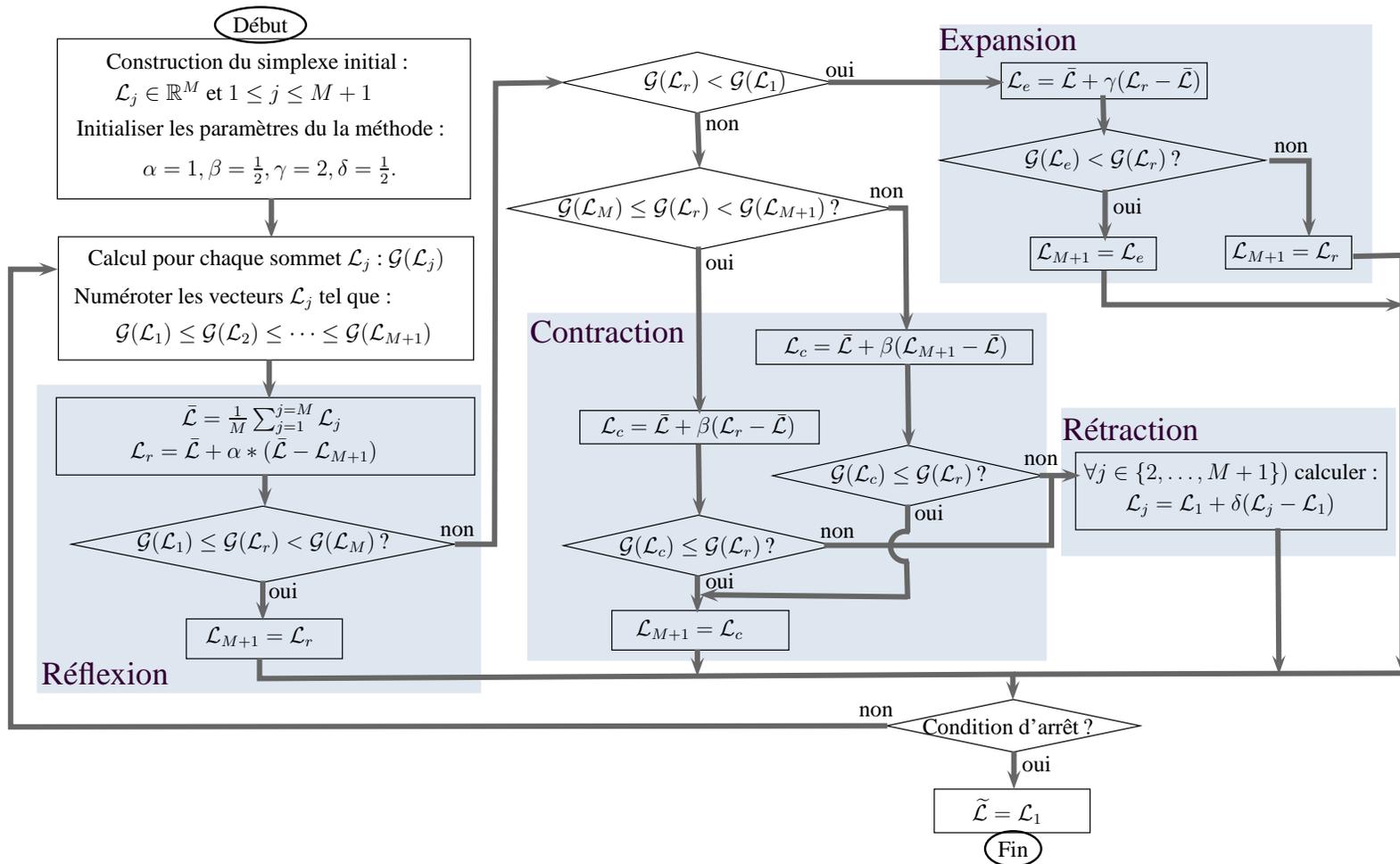


FIGURE. A.2 – Algorithme de l'algorithme de Nelder-Mead qui estime la solution minimale $\tilde{\mathcal{L}}$ qui minimise une fonction coût \mathcal{G} .

Bibliographie

- [Abelev and Kopnin, 2000] G. I. Abelev and B. P. Kopnin. Biochemistry of cancer : An introduction to this issue. *Biochemistry*, 65(1) :3–4, 2000.
- [Alberts *et al.*, 2005] B. Alberts, I. Mowszowicz, and F. Wright. *L'essentiel de la biologie cellulaire*. Flammarion médecine-sciences, 2005.
- [Baker and Matthews, 2004] S. Baker and I. Matthews. Lucas-Kanade 20 years on : A unifying framework. *International Journal of Computer Vision*, 56(3) :221–255, 2004.
- [Behrens *et al.*, 2009] A. Behrens, T. Stehle, S. Gross, and T. Aach. Local and global panoramic imaging for fluorescence bladder endoscopy. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, page 6990, 2009.
- [Behrens *et al.*, 2010] A. Behrens, M. Guski, T. Stehle, S. Gross, and T. Aach. Intensity based multi-scale blending for panoramic images in fluorescence endoscopy. In *Biomedical Imaging : From Nano to Macro, 2010 IEEE International Symposium on*, pages 1305–1308, 2010.
- [Behrens, 2008] A. Behrens. Creating panoramic images for bladder fluorescence endoscopy. *Acta Polytech : J Adv Eng*, 3 :50–54, 2008.
- [Ben-Hamadou *et al.*, 2009a] A. Ben-Hamadou, C. Soussen, W. Blondel, C. Daul, and D. Wolf. Comparative study of image registration techniques for bladder video-endoscopy. In *SPIE European Conferences on Biomedical Optics*, pages 737118–737118–7, Munich, Germany, 2009.
- [Ben-Hamadou *et al.*, 2009b] A. Ben-Hamadou, C. Soussen, W. Blondel, C. Daul, and D. Wolf. Mosaiquage d'images cystoscopiques : étude comparative de méthodes de recalage. In *XXIIe colloque GRETSI (traitement du signal et des images), Dijon (FRA), 8-11 septembre 2009*, 2009.
- [Ben-Hamadou *et al.*, 2010a] A. Ben-Hamadou, C. Daul, C. Soussen, A. Rekik, and W. Blonde. A novel 3D surface construction approach : Application to three-dimensional endoscopic data. In *2010 IEEE International Conference on Image Processing*, pages 4425–4428, Hong Kong, Hong Kong, 2010.

- [Ben-Hamadou *et al.*, 2010b] A. Ben-Hamadou, C. Soussen, C. Daul, W. Blondel, and D. Wolf. Flexible projector calibration for active stereoscopic systems. In *2010 IEEE International Conference on Image Processing*, pages 4241–4244, Hong Kong, 2010.
- [Besl and McKay, 1992] P. J Besl and N. D McKay. A method for registration of 3-D shapes. *IEEE Transactions on pattern analysis and machine intelligence*, pages 239–256, 1992.
- [Brett *et al.*, 1999] A. D. Brett, A. Hill, and C. J. Taylor. A method of 3D surface correspondence and interpolation for merging shape examples. *Image and Vision Computing*, 17(8) :635–642, 1999.
- [Brown, 1992] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4) :325–376, 1992.
- [Butime *et al.*, 2010] J. Butime, L. Galo, and I. Gutierrez. *Application of Computer Vision to 3D Reconstruction : A Survey of Reconstruction methods*. VDM Verlag Dr. Müller, 2010.
- [Can *et al.*, 2002] A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum. A Feature-Based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3) :347, 2002.
- [Chan *et al.*, 2003] M. Chan, W. Lin, C. Zhou, and J. Y Qu. Miniaturized three-dimensional endoscopic imaging system based on active stereovision. *Appl. Opt.*, 42 :1888–1898, 2003.
- [Chen and Kak, 1987] CH Chen and AC Kak. Modelling and calibration of a structured light scanner for 3D robot vision. In *in IEEE Conference on Robotics and Automation*, pages 807–815, 1987.
- [Chen and Li, 2003] S. Chen and Y.F. Li. Self-recalibration of a colour-encoded light system for automated three-dimensional measurements. *Measurement Science and Technology*, 14(1) :33–40, 2003.
- [Chen and Medioni, 1992] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. In *Robotics and Automation, Proceedings, IEEE International Conference on*, pages 2724–2729, 1992.
- [Chen *et al.*, 2007] X. Chen, J. Xi, Y. Jin, and J. Sun. Accurate calibration for a camera-projector measurement system based on structured light projection. *Optics and Lasers in Engineering*, 47(3-4) :310–319, 2007.

-
- [Chou *et al.*, 1997] J. S Chou, J. Z Qian, Z. Wu, and H. F Schramm. Automatic mosaic and display from a sequence of peripheral angiographic images. In *Proceedings of SPIE*, volume 3034, page 1077, 1997.
- [Chow *et al.*, 2004] C. K Chow, H. T Tsui, and T. Lee. Surface registration using a dynamic genetic algorithm. *Pattern recognition*, 37(1) :105–117, 2004.
- [Collignon *et al.*, 1995] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In *Information Processing in Medical Imaging*, pages 263–274, 1995.
- [Daul *et al.*, 2010] C. Daul, W. Blondel, A. Ben-Hamadou, R. Miranda-Luna, C. Soussen, and D. Wolf. From 2D towards 3D cartography of hollow organs. In *In IEEE 7th International Conference on Electrical Engineering, Computing Science and Automatic Control CCE*, 2010.
- [Deguchi *et al.*, 1996] K. Deguchi, T. Sasano, H. Arai, and H. Yoshikawa. 3-D shape reconstruction from endoscope image sequences by the factorization method. *IEICE Transactions On Information And Systems E Series D*, 79 :1329–1336, 1996.
- [Devernay *et al.*, 2001] F. Devernay, F. Mourgues, and E. Coste-Manière. Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery. *Proceedings of Medical Imaging and Augmented Reality*, pages 16–20, 2001.
- [Dipanda *et al.*, 2003] A. Dipanda, S. Woo, F. Marzani, and J.M. Bilbault. 3-D shape reconstruction in an active stereo vision system using genetic algorithms. *Pattern Recognition*, 36(9) :2143–2159, 2003.
- [Draréni *et al.*, 2009] J. Draréni, S. Roy, and P. Sturm. Projector calibration using a markerless plane. volume 2, pages 377–382, 2009.
- [Faugeras and Toscani, 1986] O. Faugeras and G. Toscani. The calibration problem for stereo. In *Proceedings, CVPR '86 (IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, 22-26, 1986)*, pages 15–20. IEEE, 1986.
- [Faugeras, 1993] O. Faugeras. *Three-dimensional computer vision : a geometric viewpoint*. MIT Press, 1993.
- [Fischler and Bolles, 1981] M. A Fischler and R. C Bolles. Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6) :381–395, 1981.
- [Forster and Tozzi, 2000] C. H.Q Forster and C. L Tozzi. Towards 3D reconstruction of endoscope images using shape from shading. In *13th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)*, 2000.

- [Forsyth and Ponce, 2003] D. A. Forsyth and J. Ponce. *Computer Vision : A Modern Approach*. Prentice Hall Professional Technical Reference, 2003.
- [Fuchs *et al.*, 1998] H. Fuchs, M. A. Livingston, R. R., D. Colucci, K. Keller, A. State, J. R. Crawford, P. Rademacher, S. H. Drake, and A. A. Meyer. *Augmented Reality Visualization for Laparoscopic Surgery*, page 934. 1998.
- [GAO *et al.*, 2008] W. GAO, L. WANG, and Z. Y HU. Flexible calibration of a portable structured light system through surface plane. *Acta Automatica Sinica*, 34(11) :1358–1362, 2008.
- [Gattengo and Chopin, 2001] B. Gattengo and D. Chopin. Progrès en urologie. volume 11, pages 993–1019. Urofrance, 2001.
- [Guerchouche and Coldefy, 2008] R. Guerchouche and F. Coldefy. Camera calibration methods evaluation procedure for images rectification and 3d reconstruction. In *Proceedings 16th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2008 (WSCG'2008), Plzen, Czech Republic*, pages 205–210, 2008.
- [Guisser *et al.*, 2000] L. Guisser, R. Payrissat, and S. Castan. PGSD : an accurate 3D vision system using a projected grid for surface descriptions. *Image and Vision Computing*, 18(6-7) :463–491, 2000.
- [Hall *et al.*, 1982] E. L. Hall, J. B. K. Tio, C. A. McPherson, and F. A. Sadjadi. Measuring curved surfaces for robot vision. *Computer Journal*, 15(12) :42–54, 1982.
- [Harris and Stephens, 1988] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50, 1988.
- [Hartley and Zisserman, 2003] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2003.
- [Hasegawa and Sato, 2001] K. Hasegawa and Y. Sato. Endoscope system for high-speed 3D measurement. *Systems and Computers in Japan*, 32(8) :30–39, 2001.
- [Hasegawa *et al.*, 2002] Kazuhide Hasegawa, Kazufusa Noda, and Yukio Sato. Electronic endoscope system for shape measurement. In *Pattern Recognition, International Conference on*, volume 1, page 10792, Los Alamitos, CA, USA, 2002. IEEE Computer Society.
- [Hayashibe *et al.*, 2002] M. Hayashibe, N. Suzuki, A. Hattori, and Y. Nakamura. Intraoperative fast 3D shape recovery of abdominal organs in laparoscopy. In *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part II*, pages 356–363. Springer-Verlag, 2002.

-
- [Hayashibe *et al.*, 2006] M. Hayashibe, N. Suzuki, and Y. Nakamura. Laser-scan endoscope system for intraoperative geometry acquisition and surgical robot safety management. *Medical Image Analysis*, 10(4) :509–519, 2006.
- [Heikkila and Silven, 1997] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, CVPR '97, pages 1106–, Washington, DC, USA, 1997. IEEE Computer Society.
- [Hernandez-Mier, 2007] Y. Hernandez-Mier. *Construction rapide d'images panoramiques applicables à l'exploration cystoscopique et à l'endoscopie fluorescence en cancérologie*. PhD thesis, Institut National Polytechnique de Lorraine, 2007.
- [Hernández-Mier *et al.*, 2010] Y. Hernández-Mier, W.C.P.M. Blondel, C. Daul, D. Wolf, and François Guillemin. Fast construction of panoramic images for cystoscopic exploration. *Computerized Medical Imaging and Graphics*, 34(7) :579–592, 2010.
- [Horn and Brooks, 1989] B. K.P Horn and M. J. Brooks. *Shape from shading*. MIT press Cambridge Massachusetts, 1989.
- [Horn, 1989] B. K. P. Horn. Obtaining shape from shading information. In *Shape from shading*, pages 123–171. MIT Press, 1989.
- [Huynh *et al.*, 1999] D.Q. Huynh, R.A. Owens, and P.E. Hartmann. Calibrating a structured light stripe system : A novel approach. *International Journal of Computer Vision*, 33(1) :73–86, 1999.
- [Iyengar *et al.*, 2001] Arun K. S. Iyengar, Hiroatsu Sugimoto, David B. Smith, and Michael S. Sacks. Dynamic in vitro quantification of bioprosthetic heart valve leaflet motion using structured light projection. *Annals of Biomedical Engineering*, 29(11) :963–973, 2001.
- [Jalink *et al.*, 1996] A. Jalink, J. McAdoo, G. Halama, and H. Liu. CCD mosaic technique for large-field digital mammography. *IEEE transactions on medical imaging*, 15(3) :260–267, 1996.
- [Jean-Denis, 2009] E. Jean-Denis. *Géométrie analytique classique*. Calvage et Mounet, 2009.
- [Keller and Ackerman, 2000] K. Keller and J. D Ackerman. Real-time structured light depth extraction. In *Proceedings of SPIE*, volume 3958, pages 11–18, 2000.
- [Koschan and Abidi, 2008] A. Koschan and M. Abidi. *Digital Color Image Processing*. Wiley-Interscience, 2008.

- [Li *et al.*, 2009] L. Li, Y. Ke, and K. Jiang. Surface reconstruction based on computer stereo vision using structured light projection. In *2009 International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2009*, volume 2, pages 451–454, 2009.
- [Lin and Lin, 1999] C.-F. Lin and C.-Y. Lin. A new approach to high precision 3-D measuring system. *Image and Vision Computing*, 17(11) :805–814, 1999.
- [Liu *et al.*, 2006] Q. Liu, R. J. Sclabassi, N. Yao, and M. Sun. 3D construction of endoscopic images based on computational stereo. In *Bioengineering Conference*, page 69–70, 2006.
- [Lowe, 2004] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2) :91–110, 2004.
- [Maintz and Viergever, 1998] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2 :1–36, 1998.
- [Markelj *et al.*, 2010] P. Markelj, D. Tomazevic, B. Likar, and F. Pernus. A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis*, 2010.
- [Marzani *et al.*, 2002] F. S. Marzani, Y. Voisin, L. F.C.L.Y. Voon, and A. Diou. Calibration of a three-dimensional reconstruction system using a structured light source. *Optical Engineering*, 41 :484, 2002.
- [Miranda-Luna *et al.*, 2004] R. Miranda-Luna, W.C.P.M. Blondel, C. Daul, Y. Hernandez-Mier, R. Posada, and D. Wolf. A simplified method of endoscopic image distortion correction based on grey level registration. In *Proceedings - International Conference on Image Processing, ICIP*, volume 2, pages 3383–3386, 2004.
- [Miranda-Luna *et al.*, 2008] R. Miranda-Luna, C. Daul, W.C.P.M. Blondel, Y. Hernandez-Mier, D. Wolf, and F. Guillemin. Mosaicing of bladder endoscopic image sequences : Distortion calibration and registration algorithm. *IEEE Transactions on Biomedical Engineering*, 55(2) :541–553, 2008.
- [Miranda-Luna, 2005] R. Miranda-Luna. *Contribution au mosaïquage d’images endoscopiques d’organes creux pour l’aide au diagnostic du cancer en phase précoce*. PhD thesis, Institut National Polytechnique de Lorraine, 2005.
- [Mosnier *et al.*, 2009] J. Mosnier, F. Berry, and O. Ait-Aider. A new method for projector calibration based on visual servoing. In *Machine Vision Application, MVA2009, Japan*, 2009.

-
- [Mourgues *et al.*, 2001] F. Mourgues, F. Devernay, and E. Coste-Maniere. 3D reconstruction of the operating field for image overlay in 3D-endoscopic surgery. In *Proceedings of International Symposium on Augmented Reality*, 2001.
- [Nelder and Mead, 1965] JA Nelder and R. Mead. A simplex method for function minimization. *The computer journal*, 7(4) :308, 1965.
- [Nelder, 2009] J. Nelder. Nelder-Mead algorithm. *Scholarpedia*, 4(7) :2928, 2009.
- [Okatani and Deguchi, 1997] T. Okatani and K. Deguchi. Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center. *Computer Vision and Image Understanding*, 66(2) :119–131, 1997.
- [Ouellet and Hébert, 2008] J.-N. Ouellet and P. Hébert. Precise ellipse estimation without contour point extraction. *Machine Vision and Applications*, 21(1) :59–67, 2008.
- [Pluim *et al.*, 2003] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images : A survey. *IEEE Transactions on Medical Imaging*, 22(8) :986–1004, 2003.
- [Posada *et al.*, 2007] R. Posada, C. Daul, D. Wolf, and P. Aletti. Towards a non invasive intracranial tumor irradiation using 3-D optical imaging and multimodal data registration. *International Journal of Biomedical Imaging, Special Issue on Multimodality Imaging and Hybrid Scanners*, 2007 :14, 2007.
- [Press *et al.*, 2007] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes 3rd Edition : The Art of Scientific Computing*. Cambridge University Press, 3 edition, 2007.
- [Rémy, 2006] Y. Rémy. Reconstruction volumique d’une vessie à partir d’images endoscopiques mosaïquées. Rapport de stage d’ingénieur, 2006.
- [Salvi *et al.*, 1998] J. Salvi, J. Batlle, and E. Mouaddib. A robust-coded pattern projection for dynamic 3D scene measurement. *Pattern Recognition Letters*, 19(11) :1055–1065, 1998.
- [Salvi *et al.*, 2002] J. Salvi, X. Armangué, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern recognition*, 35(7) :1617–1635, 2002.
- [Salvi *et al.*, 2004] J. Salvi, J. Pages, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4) :827–849, 2004.
- [Salvi *et al.*, 2007] J. Salvi, C. Matabosch, D. Fofi, and J. Forest. A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, 25(5) :578–596, 2007.

- [Sansoni *et al.*, 2009] G. Sansoni, M. Trebeschi, and F. Docchio. State-of-The-Art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation. *Sensors*, 9 :568–601, 2009.
- [Schubert and Müller, 1998] M. Schubert and A. Müller. Evaluation of endoscopic images for 3-dimensional measurement of hollow biological organs. *Biomedizinische Technik. Biomedical engineering*, 43 Suppl :32–3, 1998.
- [Snavely, 2008] K. N Snavely. *Scene Reconstruction and Visualization from Internet Photo Collections*. PhD thesis, University of Washington, 2008.
- [Stoyanov *et al.*, 2005] D. Stoyanov, A. Darzi, and G. Z Yang. A practical approach towards accurate dense 3D depth recovery for robotic laparoscopic surgery. *Computer Aided Surgery*, 10(4) :199–208, 2005.
- [Svoboda, 2005] T. Svoboda. A software for complete calibration of multicamera systems. In *SPIE-IS & T Electronic Imaging, SPIE 2005*, volume 5658, pages 115–128, 2005.
- [Tamas *et al.*, 2009] L. Tamas, G. Lazea, M. Popa, I. Szoke, and A. Majdik. Laser based localization techniques for indoor mobile robots. In *Proceedings - 2009 Advanced Technologies for Enhanced Quality of Life, AT-EQUAL 2009*, pages 169–170, 2009.
- [Thormahlen *et al.*, 2002] T. Thormahlen, H. Broszio, and P. N. Meier. Three-dimensional endoscopy. In *Medical imaging in gastroenterology and hepatology*, 2002.
- [Tsai, 1987] R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of robotics and Automation*, 3(4) :323–344, 1987.
- [Tyczynski and Parkin, 2003] J. E. Tyczynski and D. M. Parkin. Bladder cancer in europe. *ENCR Cancer Fact Sheets. European Network of Cancer Registries. International Agency for Research on Cancer*, 3, 2003.
- [Vercauteren *et al.*, 2005] T. Vercauteren, A. Perchant, X. Pennec, and N. Ayache. Mosaicing of confocal microscopic in vivo soft tissue video sequences. In *Medical Image Computing and Computer - Assisted Intervention-MICCAI*, pages 753–760, 2005.
- [Viola and Wells III, 1997] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2) :137–154, 1997.
- [Wang *et al.*, 2009] G. Wang, J. Han, and X. Zhang. Three-dimensional reconstruction of endoscope images by a fast shape from shading method. *Measurement Science and Technology*, 20 :125801, 2009.
- [WHO, 2009] WHO. *World Health Organization. Fact sheet N° 297*. WHO Media Center, 2009.

-
- [Wu *et al.*, 2007] C. Wu, S. G. Narasimhan, and B. Jaramaz. Shape-from-shading under near point lighting and partial views for orthopedic endoscopy. In *Workshop on Photometric Analysis For Computer Vision*, 2007.
- [Wu *et al.*, 2010] C. Wu, S. G. Narasimhan, and B. Jaramaz. A Multi-Image Shape-from-Shading framework for Near-Lighting perspective endoscopes. *Int. J. Comput. Vision*, 86(2-3) :211–228, 2010.
- [Yamauchi *et al.*, 2008] K. Yamauchi, H. Saito, and Y. Sato. Calibration of a structured light system by observing planar object from unknown viewpoints. In *Pattern Recognition, ICPR . 19th International Conference on*, pages 1–4, 2008.
- [Yeung *et al.*, 1999] S. Y. Yeung, H. T. Tsui, and A. Yim. Global shape from shading for an endoscope image. In *Proceedings of the Second International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 318–327. Springer-Verlag, 1999.
- [Zhang *et al.*, 1999] R. Zhang, P. S. Tsai, J. E. Cryer, and M. Shah. Shape from shading : a survey. *IEEE transactions on pattern analysis and machine intelligence*, 21(8) :690–706, 1999.
- [Zhang, 2000] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11) :1330–1334, 2000.
- [Zhou and Zhang, 2005] F. Zhou and G. Zhang. Complete calibration of a structured light stripe vision sensor through planar target of unknown orientations. *Image and Vision Computing*, 23(1) :59–67, 2005.
- [Zitová and Flusser, 2003] B. Zitová and J. Flusser. Image registration methods : a survey. *Image and Vision Computing*, 21(11) :977–1000, 2003.

Résumé

La cystoscopie est actuellement l'examen clinique de référence permettant l'exploration visuelle des parois internes de la vessie. Le cystoscope (instrument utilisé pour cet examen) permet d'acquérir une séquence vidéo des parois épithéliales de la vessie. Cependant, chaque image de la séquence vidéo ne visualise qu'une surface réduite de quelques centimètres carrés de la paroi. Les travaux réalisés dans le cadre de cette thèse ont pour objectif de construire une carte 3D reproduisant d'une manière fidèle les formes et les textures des parois internes de la vessie. Une telle représentation de l'intérieur de la vessie permettrait d'améliorer l'interprétation des données acquises lors d'un examen cystoscopique. Pour atteindre cet objectif, un nouvel algorithme flexible est proposé pour le calibrage de systèmes cystoscopiques à vision active. Cet algorithme fournit les paramètres nécessaires à la reconstruction précise de points 3D sur la portion de surface imagée à chaque instant donné de la séquence vidéo cystoscopique. Ainsi, pour chaque acquisition de la séquence vidéo, un ensemble de quelques points 3D/2D et une image 2D est disponible. L'objectif du deuxième algorithme proposé dans cette thèse est de ramener l'ensemble des données obtenues pour une séquence dans un repère global pour générer un nuage de points 3D et une image panoramique 2D représentant respectivement la forme 3D et la texture de la totalité de la paroi imagée dans la séquence vidéo. Cette méthode de cartographie 3D permet l'estimation simultanée des transformations 3D rigides et 2D perspectives liant respectivement les positions du cystoscope et les images de paires d'acquisitions consécutives. Les résultats obtenus sur des fantômes réalistes de vessie montrent que ces algorithmes permettent de calculer des surfaces 3D reproduisant les formes à retrouver.

Mots-clés: Cartographie 3D/2D, recalage 2D, stéréovision active, reconstruction 3D, aide au diagnostic, cancer de la vessie.

Abstract

Cystoscopy is currently the reference clinical examination for visual exploration of the inner walls of the bladder. A cystoscope (instrument used in this examination) allows for video acquisition of the bladder epithelium. Nonetheless, each frame of the video displays only a small area of few squared centimeters. This work aims to build 3D maps representing the 3D shape and the texture of the inner walls of the bladder. Such maps should improve and facilitate the interpretation of the cystoscopic data. To reach this purpose, a new flexible algorithm is proposed for the calibration of cystoscopic active vision systems. This algorithm provides the required parameters to achieve accurate reconstruction of 3D points on the surface part imaged at each given moment of the video cystoscopy. Thus, available data for each acquisition are a set of few 3D points (and their corresponding 2D projections) and a 2D image. The aim of the second algorithm described in this work is to place all the data obtained for a sequence in a global coordinate system to generate a 3D point cloud and a 2D panoramic image representing respectively the 3D shape and the texture of the bladder wall imaged in the video. This 3D cartography method allows for the simultaneous estimation of 3D rigid transformations and 2D perspective transformations. These transformations give respectively the link between cystoscope positions and between images of consecutive acquisitions. The results obtained on realistic bladder phantoms show that the proposed method generates 3D surfaces recovering the ground truth shapes.

Keywords: 3D/2D cartography, 3D registration, active stereovision, 3D reconstruction, diagnostic aid, bladder cancer.

AUTORISATION DE SOUTENANCE DE THESE
DU DOCTORAT DE L'INSTITUT NATIONAL
POLYTECHNIQUE DE LORRAINE

o0o

VU LES RAPPORTS ETABLIS PAR :

Monsieur David FOFL, Professeur, IUT IE Creusot - Le2i, LE CREUSOT

Monsieur Su RUAN, Professeur, LITIS, Université de Rouen

Le Président de l'Institut National Polytechnique de Lorraine, autorise :

Monsieur BEN-HAMADOU Achraf

à soutenir devant un jury de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE,
une thèse intitulée :

**"Contribution à la cartographie 3D des parois internes de la vessie par cystoscopie à
vision active. "**

en vue de l'obtention du titre de :

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Spécialité : « Automatique, Traitement du Signal et des Images, Génie Informatique »

Fait à Vandoeuvre, le 05 septembre 2011

Le Président de l'IN.P.L.

F. LAURENT

