

Titre: Système de visualisation pour l'assistance des chirurgies
minimalement invasives du rachis

Auteur: Martin Bisson
Author:

Date: 2010

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Bisson, M. (2010). Système de visualisation pour l'assistance des chirurgies
minimalement invasives du rachis [Master's thesis, École Polytechnique de
Montréal]. PolyPublie. <https://publications.polymtl.ca/385/>

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

URL de PolyPublie: <https://publications.polymtl.ca/385/>
PolyPublie URL:

**Directeurs de
recherche:** Farida Cheriet, & Stefan Parent
Advisors:

Programme: Génie informatique
Program:

UNIVERSITÉ DE MONTRÉAL

SYSTÈME DE VISUALISATION POUR L'ASSISTANCE DES CHIRURGIES
MINIMALEMENT INVASIVES DU RACHIS

MARTIN BISSON
DÉPARTEMENT DE GÉNIE INFORMATIQUE ET GÉNIE LOGICIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE INFORMATIQUE)
AOÛT 2010

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

SYSTÈME DE VISUALISATION POUR L'ASSISTANCE DES CHIRURGIES
MINIMALEMENT INVASIVES DU RACHIS

présenté par : BISSON Martin

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

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

M. LANGLOIS J. M. Pierre, Ph.D., président

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

M. PARENT Stefan, MD, membre et codirecteur de recherche

M. OZELL Benoît, Ph.D., membre

REMERCIEMENTS

Je tiens à d'abord remercier chaleureusement ma directrice de recherche, Dr Farida Cheriet, pour toute l'aide qu'elle m'a apportée au cours de ma maîtrise. Grâce à sa patience et son ouverture d'esprit, elle a su me guider à travers mon cheminement quelque peu non linéaire, ayant toujours à cœur le succès de ses étudiants dans un sens plus large que le contexte scolaire. Lorsque j'étais intimidé par le chemin qu'il restait à faire, elle me rappelait tout le chemin parcouru et restait toujours extrêmement positive. Je ne peux tout simplement pas imaginer un meilleur directeur de maîtrise.

Je veux également remercier mon co-directeur de recherche, Dr Stefan Parent. L'intérêt et la passion qu'il porte pour l'avancement de la recherche, malgré son emploi du temps plus que chargé, sont des sources d'inspiration constantes.

Merci également à Philippe Debanné, pour son aide sur de nombreuses questions et pour sa disponibilité de tous les instants.

Je tiens aussi à souligner la merveilleuse communauté formée par les étudiants du LIV4D. L'incroyable quantité d'énergie intellectuelle présente dans ce laboratoire constituait une source inépuisable d'échanges qui m'ont aidé à continuer à progresser lorsque je me butais à des impasses. Les discussions avec les membres du laboratoire, autant techniques que non techniques, ont contribué à faire de ma maîtrise une expérience plus que scolaire. Bien que tous les membres du laboratoire se soient avérés extraordinaires, je tiens à remercier spécialement Olivier Dionne, qui a su me garder en forme par ses « *kill shots* » et qui a enrichi mon expérience de maîtrise par de nombreuses discussions et par son amitié.

Finalement, je tiens donner un immense, chaleureux et sincère remerciement à mes parents pour leur support tout au long de ma maîtrise, à la fois financier, alimentaire, moral, psychologique et même parfois locomoteur. Ils ont vraiment tout fait ce qui était en leur pouvoir pour m'aider à tous les niveaux, s'en voulant même parfois de ne pouvoir m'aider sur des aspects qui de toute façon étaient bien en dehors de leur contrôle. Je ne pourrai jamais les remercier assez, mais je vais quand même essayer :

Merci. Mille fois.

RÉSUMÉ

La scoliose est une déformation tridimensionnelle complexe de la colonne vertébrale qui peut nécessiter un traitement chirurgical. Les chirurgies minimalement invasives, par opposition aux chirurgies conventionnelles, sont réalisées en pratiquant de petites incisions par lesquelles les instruments chirurgicaux et une caméra endoscopique peuvent être introduits à l'intérieur du corps. Ce type de chirurgie possède de nombreux avantages reconnus, notamment une réduction du volume de perte sanguine, de la douleur postopératoire, du séjour aux soins intensifs, du séjour à l'hôpital, des complications, du temps de convalescence et du coût total du traitement. Par contre, il amène de nouveaux défis pour les chirurgiens. La visualisation du site opératoire est faite de manière indirecte via un moniteur présentant les images provenant de la caméra endoscopique, ce qui peut s'avérer désorientant. Comme cette vue est monoculaire et que la lentille de la caméra est très près du site opératoire, la perception de la profondeur est perdue. Cette proximité de la lentille entraîne également une perte de la perception de la forme globale de la colonne entière, puisque seule une petite partie est visible à la fois. Ces difficultés font des chirurgies minimalement invasives des procédures très complexes, spécialement dans le cas de chirurgies de la colonne vertébrale, où de nombreux organes critiques sont situés à proximité du site opératoire.

Des systèmes d'assistance chirurgicale guidés par l'image sont de plus en plus disponibles dans différents domaines tels que la cardiologie, la neurologie et l'orthopédie. Pour les chirurgies de la colonne, on a souvent recours à des modalités d'imagerie intra-opératoires ionisantes comme la tomodensitométrie ou la fluoroscopie qui impliquent des radiations à la fois pour le patient et pour l'équipe chirurgicale. Aussi, ces systèmes sont généralement conçus pour les chirurgies conventionnelles qui impliquent de grandes incisions, ce qui facilite l'accès aux structures d'intérêt. Ainsi, des marqueurs métalliques peuvent être vissés sur les niveaux vertébraux d'intérêt, ce qui constitue une procédure invasive difficilement praticable dans le contexte minimalement invasif. Certaines approches récentes d'assistance de chirurgies minimalement invasives tentent de fournir des indices de profondeur au chirurgien uniquement à partir du contenu des images endoscopiques. Les méthodes utilisant un endoscope stéréoscopique produisent des résultats mitigés à cause de la faible disparité entre les deux vues acquises et des complications de mise en correspondance de primitives sur les deux vues entraînées par la présence de fumée, de sang et de réflexions spéculaires. Dans

le même ordre d'idée, les approches basées sur les algorithmes de structure à partir de l'ombrage produisent à ce jour des résultats acceptables dans des environnements contrôlés, mais ne performant pas aussi bien avec des images réelles.

L'objectif de ce projet de recherche est donc de proposer un outil de visualisation 3D pour l'assistance des chirurgies minimalement invasives de la colonne vertébrale. Cet outil réalise la fusion multimodale des images provenant de la caméra endoscopique et d'un modèle 3D préopératoire obtenu par IRM ayant été préalablement segmenté puis recalé dans un repère lié au site opératoire, présentant ainsi au chirurgien davantage d'information que celle présente dans les images endoscopiques seules.

Afin de réaliser cet objectif, un modèle de formation de l'image endoscopique est développé à partir de modèles proposés dans la littérature. Un modèle de caméra à projection perspective est retenu. Toutefois, la rotation relative qu'il est possible d'effectuer entre la tête de la caméra endoscopique et le tube de l'endoscope afin d'explorer l'intérieur du corps entraîne un degré de liberté supplémentaire qui doit également être modélisé. Pour ce faire, une variante du modèle de Yamaguchi est utilisée pour représenter cette transformation supplémentaire, alors qu'une variante de la méthode de Wu est utilisée pour simplifier le modèle de Yamaguchi sans toutefois perdre de flexibilité et pour obtenir l'angle de rotation sans avoir recours à un encodeur de rotation.

Ensuite, un processus de calibrage hors-ligne des paramètres du modèle est développé. Ce dernier est basé sur l'algorithme de calibrage développé par Zhang tel qu'implémenté par la librairie OpenCV. Une identification complètement automatique des points de la grille de l'objet de calibrage est proposée pour minimiser le besoin d'intervention de l'utilisateur. L'objet de calibrage est conçu de manière à être simple et peu coûteux à produire, en plus de permettre une grande précision et d'éviter le besoin d'avoir recours à un numériseur 3D pour le mesurer.

Puis, une méthode de mise à jour d'une partie des paramètres qui varient au cours de la chirurgie, soit les paramètres extrinsèques de la caméra et l'angle de la rotation relative entre la tête de la caméra et le tube de l'endoscope, est proposée. Cette méthode repose sur l'utilisation d'un système de suivi optique, soit le MicronTracker de Claron Technology Inc. Ce dernier permet de suivre en temps réel la position et l'orientation de marqueurs faciles à concevoir et à produire. Un de ces marqueurs est posé sur le tube de l'endoscope de manière

à pouvoir suivre en temps réel la position de la caméra, alors qu'un autre est posé sur la tête de la caméra afin de pouvoir évaluer l'angle de rotation de la tête par rapport au tube, tel que proposé par Wu.

Par la suite, deux modes de visualisation sont proposés pour présenter l'information combinée des images endoscopiques et du modèle 3D préopératoire d'une manière efficace. On propose d'abord une visualisation de réalité augmentée dans laquelle la réalité, soit les images endoscopiques, est augmentée à l'aide d'information provenant du modèle préopératoire telle que la projection du modèle sur l'image. On propose ensuite une visualisation de virtualité augmentée dans laquelle le modèle préopératoire obtenu par IRM est augmenté en y ajoutant la réalité, c'est-à-dire les images endoscopiques positionnées adéquatement selon la position de l'endoscope par rapport à la colonne vertébrale.

Différentes parties de la méthode de calibrage sont validées à l'aide de simulations et d'expériences avec des images endoscopiques réelles. Les simulations, au cours desquelles les données utilisées pour le calibrage ont été artificiellement bruitées, montrent que la méthode est tolérante au bruit. Les expériences de validation sur les images réelles ont permis de valider le modèle de formation d'images proposé ainsi que la répétabilité des résultats. Elles confirment également que les vues calibrées à l'aide de la méthode proposée peuvent être utilisées pour mesurer des distances avec une précision millimétrique à l'aide de vues multiples d'un segment à mesurer. Les vues proposées ont été également approuvées par un expert chirurgical qui confirme la pertinence de ces dernières.

Une intégration de l'ensemble des composantes du système de visualisation 3D reste à être effectuée afin d'obtenir une plateforme comprenant la méthode de calibrage hors-ligne, la mise à jour en ligne de l'ensemble des paramètres intrinsèques et extrinsèques de l'endoscope, la segmentation et le recalage du modèle préopératoire obtenu par IRM et la fusion des images endoscopiques avec le modèle de la colonne vertébrale. Une fois cette plateforme développée, elle pourra être validée dans un contexte chirurgical réel au cours d'une chirurgie réalisée sur un modèle animal comme le mini-porc.

ABSTRACT

Scoliosis is a complex three-dimensional deformity of the spine that may require surgical treatment. Minimally invasive surgeries, as opposed to conventional surgeries, are done by performing small incisions through which surgical instruments and an endoscopic camera can be inserted within the body. This type of surgery offers numerous known benefits, such as reduction of blood loss, of post-surgery pain, of intensive care unit and hospital stay lengths, of complications, of recovery time and of overall treatment costs. However, it brings new challenges for surgeons. Indirect visualization of surgical sites via a monitor that shows images from the endoscopic camera can be disorienting. Depth perception is lost as the view is monocular and the camera lens is very close to the surgical site. This lens proximity also causes a loss in perception of the global spine shape since only a small part of it is visible at one time. These difficulties make minimally invasive surgeries very complex procedures, especially in the case of spine surgeries, where several critical organs are located near the surgical site.

Image-guided surgical assistance tools are increasingly available in various fields such as cardiology, neurology and orthopaedics. For spine surgeries, ionizing imaging modalities such as computed tomography or fluoroscopy, involving radiation for both the patient and the surgical team, are often used. These surgical navigation systems are also generally designed for conventional surgeries requiring long incisions to facilitate access to targeted structures. Thus, metal markers can be screwed into specific vertebral levels, a procedure that is both invasive and hard to perform in a minimally invasive context. Recent minimally invasive surgical assistance approaches try to provide depth cues to the surgeon using only the content of endoscopic images. Methods using a stereoscopic endoscope give unsatisfactory results due to the weak disparity between the two acquired views and to the complexity of feature matching in the two views, caused by the presence of smoke, blood and specular reflections. By the same token, approaches based on shape-from-shading algorithms give, to date, acceptable results in controlled environments but do not perform as well with real images.

Hence, this research project's objective is to propose a 3D visualization tool for minimally invasive spine surgery assistance. This tool realizes multimodal fusion of the images acquired from the endoscopic camera with a preoperative 3D model acquired by MRI that has been

previously segmented and then registered in a coordinate system linked to the surgical site, in order to display more information than what is available in the endoscopic images alone.

To achieve this objective, we first develop a model of endoscopic image formation from models available in the literature. We choose a perspective projection camera model for this purpose. However, since the endoscopic camera head can be rotated relative to the endoscope tube in order to explore the body cavity, this additional degree of freedom must also be modeled. To do so, we use a variant of Yamaguchi's method to model this extra transformation, while a variant of Wu's method allows us to simplify Yamaguchi's model without losing any flexibility and to obtain the rotation angle without using a rotary encoder.

Then, we develop an offline calibration process for the model parameters. It is based on the calibration algorithm developed by Zhang as implemented in the OpenCV library. We propose a fully automatic identification of calibration object grid points to minimize the need for user intervention. The calibration object is designed to be simple and inexpensive to produce, in addition to allowing high precision and avoiding the need to measure it using a 3D scanner.

Next, we propose a method for online update of a subset of parameters that vary during surgery, namely the camera's extrinsic parameters and the relative rotation angle between the camera head and the endoscope tube. This method relies on the use of an optical tracking system, the MicronTracker by Claron Technology Inc. This system is able to track in real time the position and orientation of markers which can be easily designed and produced. One such marker is placed on the endoscope tube to track in real time the camera's location, while another is placed on the camera head to measure its rotation angle with respect to the tube, as proposed by Wu.

Subsequently, we propose two visualization modes to efficiently display the combined information from endoscopic images and the preoperative 3D model. The first of these is an *augmented reality visualization* mode in which reality, i.e. the endoscopic images, is augmented with information from the preoperative model such as its projection on the images. The second mode is an *augmented virtuality visualization* in which the preoperative model acquired from MRI is augmented by adding reality, i.e. the endoscopic images, which must be properly positioned according to the endoscope's location relative to the spine.

We validate different parts of the calibration method using simulations as well as experiments with real endoscopic images. Simulations, in which the calibration data is artificially noised, show that our method is tolerant to noise. Validation experiments on real images allow us to demonstrate the proposed image formation model as well as results reproducibility. They also confirm that views calibrated using the proposed method can be used to measure distances with millimetre precision using multiple views of a segment to be measured. The proposed display modes have also been approved by an expert surgeon who confirms their relevance.

A full integration of all the 3D visualization system components has to be carried out in order to obtain a framework featuring the offline calibration method, the online update of all intrinsic and extrinsic endoscope parameters, the segmentation and registration of the preoperative model obtained by MRI and finally the fusion of endoscopic images with the *a priori* spine model. Once this framework is developed, it will be possible to validate it in a real surgical setting during a surgery performed on an animal model such as a miniature swine.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
RÉSUMÉ	iv
ABSTRACT	vii
TABLE DES MATIÈRES	x
LISTE DES TABLEAUX	xiv
LISTE DES FIGURES	xv
LISTE DES SIGLES ET ABRÉVIATIONS	xxv
LISTE DES ANNEXES	xxvi
INTRODUCTION	1
CHAPITRE 1 REVUE DE LA LITTÉRATURE	8
1.1 Recalage de modèles préopératoires pour l'assistance chirurgicale	9
1.1.1 Reconstruction des modèles	10
1.1.2 Recalage des modèles	14
1.2 Visualisation 3D	22
1.2.1 Technologies d'affichage	22
1.2.2 Paradigmes de visualisation	25
1.2.3 Autres outils d'aide à la visualisation 3D	33
1.2.4 Endoscopes stéréoscopiques	33
1.2.5 Structure à partir de l'ombrage	34
1.2.6 Images intra-opératoires pour l'assistance de la chirurgie du rachis	36
1.3 Description générale de l'endoscope	36
1.4 Modèle de formation de l'image endoscopique	40
1.5 Calibrage du modèle	42
1.5.1 Calibrage par la méthode <i>DLT</i>	42
1.5.2 Calibrage par la méthode de Tsai	44

1.5.3	Calibrage par la méthode de Zhang.....	45
1.5.4	Calibrage par la méthode de Girard.....	47
1.6	Mise à jour des paramètres extrinsèques.....	48
1.6.1	Technologies de suivi.....	49
1.7	Extensions spécifiques à la caméra endoscopique.....	58
1.7.1	Méthode de Yamaguchi.....	58
1.7.2	Méthode de Wu.....	60
1.7.3	Méthode de De Buck.....	65
1.8	Limites des systèmes existants.....	67
1.9	Objectifs de recherche.....	67
CHAPITRE 2 MÉTHODOLOGIE.....		69
2.1	Matériel utilisé.....	69
2.1.1	Système endoscopique.....	69
2.1.2	Système de suivi.....	69
2.2	Modèle de formation des images endoscopiques.....	70
2.2.1	Transformation 1 : paramètres extrinsèques.....	70
2.2.2	Transformation 2 : rotation relative de la tête de la caméra par rapport au tube.....	72
2.2.3	Transformation 3 : distorsion de lentille.....	76
2.2.4	Transformation 4 : numérisation.....	78
2.3	Méthode de calibrage hors-ligne.....	79
2.3.1	Choix d'implémentation.....	79
2.3.2	Conception de l'objet de calibrage.....	80
2.3.3	Acquisition des vues de calibrage.....	84
2.3.4	Calibrage sans rotation.....	87
2.3.5	Calibrage avec rotation.....	92
2.4	Méthode de mise à jour en ligne.....	96
2.4.1	Mise à jour des paramètres extrinsèques.....	97
2.4.2	Mise à jour de l'angle de rotation.....	97
2.5	Résumé du calibrage.....	98
2.6	Génération des vues.....	100

2.6.1	Vue #1 : réalité augmentée	100
2.6.2	Vue #2 : virtualité augmentée.....	101
2.7	Validation	103
2.7.1	Expériences de simulation	104
2.7.2	Expériences de validation.....	105
2.7.3	Validation des vues.....	108
CHAPITRE 3 RÉSULTATS		111
3.1	Expériences de simulation	111
3.1.1	Simulation #1 : bruit dans les images	113
3.1.2	Simulation #2 : bruit sur l'objet de calibrage	114
3.1.3	Simulation #3 : bruit sur le suivi	116
3.2	Expériences de validation sur des images endoscopiques réelles	118
3.2.1	Expérience #1 : construction de l'objet de calibrage.....	118
3.2.2	Expérience #2 : répétabilité des résultats	119
3.2.3	Expérience #3 : position du centre de distorsion.....	127
3.2.4	Expérience #4 : modèle de compensation de la rotation	128
3.2.5	Expérience #5 : mesure de distances	131
3.3	Validation des vues.....	133
CHAPITRE 4 DISCUSSION GÉNÉRALE		134
4.1	Système de suivi	134
4.2	Manipulations pour le calibrage	135
4.3	Méthode de validation	136
4.4	Défis pour l'intégration en salle de chirurgie	137
4.4.1	Recalage des modèles préopératoires.....	137
4.4.2	Suivi des instruments.....	139
4.4.3	Système de suivi	139
4.4.4	Interaction avec le système.....	141
4.4.5	Définition du protocole de calibrage	141
4.4.6	Compensation de la rotation	142
4.4.7	Synchronisation des multiples sources de données	143
4.4.8	Système endoscopique du LIV4D	144

CHAPITRE 5	CONCLUSIONS ET RECOMMANDATIONS.....	146
5.1	Recommandations	148
5.1.1	Éléments de visualisation	148
5.1.2	Ajouts intéressants.....	150
RÉFÉRENCES	152
ANNEXES	161

LISTE DES TABLEAUX

Tableau 1.1. Comparaison des modalités de suivi de pose (adapté de [76]).....	50
Tableau 2.1. Résumé du modèle de formation d'image.....	99
Tableau 3.1. Résultats de la simulation #3	117
Tableau 3.2. Comparaison des objets de calibrage.....	119
Tableau 3.3. Résultats partiels de l'expérience #2.....	121
Tableau 3.4. Résultats partiels de l'expérience #2.....	123
Tableau 3.5. Résultats de l'expérience #5.....	132

LISTE DES FIGURES

- Figure 1. Exemple de radiographies de colonne en santé (gauche, adaptée de Wikimedia Commons, "File:DualEnergyAbsorptiometryofChestTissueSubtracted.png," *Wikimedia Commons*, 2007. [En ligne]. Disponible: <http://commons.wikimedia.org/wiki/File:DualEnergyAbsorptiometryofChestTissueSubtracted.png>. [Consulté le 2010/06/24]. ©2007 Kieran Maher. Dans le domaine public.) et de colonne scoliotique (droite, patient de l'Hôpital Ste-Justine)..... 1
- Figure 2. Cicatrice résultant d'une thoracotomie (gauche, adaptée de E. Popovic, "Thoracic disc prolapse," *Cervical spinal surgery, injury symptoms, neck pain treatment*. [En ligne]. Disponible: http://www.popovic.com.au/surgery_cervical.html. [Consulté le 2010/07/12]. ©2010 Emil Popovic. Reproduit avec permission.) et d'une thoracoscopie (droite, adaptée de Y. Qiu, B. Wang, et F. Zhu, "Comparison of the curative effects of video assisted thoracoscopic anterior corrected and small incision, thoracotomic anterior correction for idiopathic thoracic scoliosis," *Chinese Medical Journal*, vol. 121, no. 15, pp. 1369-1373. [En ligne]. Disponible: <http://www.cmj.org/Periodical/images/200807/2008730144442674.jpg>. [Consulté le 2010/05/05]. Sous licence Creative Commons.)..... 2
- Figure 3. Vue d'ensemble du système d'assistance chirurgicale (©2010 LIV4D. Reproduit avec permission)..... 5
- Figure 1.1. Étapes du système de réalité augmentée 8
- Figure 1.2. Appareil de tomodensitométrie par rayons X (Tiré de Radiological Society of North America, Inc., "CAT Scan (CT) – Spine," *Images and Videos, CAT Scan (CT) – Spine quipment/procedure photograph*. 2010. [En ligne]. Disponible: <http://www.radiologyinfo.org/en/photocat/gallery3.cfm?image=philip25.jpg&pg=spinect>. [Consulté le 2010/06/30]. ©2010 Radiological Society of North America, Inc. Reproduit avec permission.)..... 10

- Figure 1.3. De gauche à droite : radiographie PA-0, PA-20, LAT; reconstruction vue de derrière, de gauche (patient de l'Hôpital Ste-Justine)..... 11
- Figure 1.4. Exemple de coupe sagittale de volume IRM (patient de l'Hôpital Ste-Justine)..... 13
- Figure 1.5. Plans axial, coronal et sagittal..... 14
- Figure 1.6. Exemple de marqueur optique fixé à une vertèbre (adaptée de J. S. Hott, N. Theodore, J. D. Klopfenstein, S. Papadopolous, C. A. Dickman, et V. K. H. Sonntag, "Image guidance in spinal surgery for trauma: Iso-C 3D," *Operative Techniques in Neurosurgery*, vol. 7, no. 1, pp. 42-47, 2004. ©2004 Elsevier. Reproduit avec permission.)..... 16
- Figure 1.7. Acquisition d'images avec la sonde à ultrasons et son marqueur attaché (tirée de S. Winter, I. Pechlivanis, C. Dekomien, C. Igel, et K. Schmieder, "Toward Registration of 3D Ultrasound and CT Images of the Spine in Clinical Praxis: Design and Evaluation of a Data Acquisition Protocol," *Ultrasound in Medicine & Biology*, vol. 35, no. 11, pp. 1773-1782, 2009. ©2009 Elsevier. Reproduit avec permission.)..... 21
- Figure 1.8. Résultat du recalage par ultrasons. Gauche : tranche sagittale de la tomodensitométrie où la surface des os est identifiée par des lignes noires. Centre : image ultrasonique correspondante. Droite : superposition de la surface identifiée, après recalage, et de la tranche ultrasonique. (Tirée de B. Brendel, S. Winter, A. Rick, M. Stockheim, et H. Ermert, "Bone registration with 3D CT and ultrasound data sets," *International Congress Series*, vol. 1256, pp. 426-432, 2003. ©2003 Elsevier. Reproduit avec permission.) 22
- Figure 1.9. Casque d'affichage optique (gauche, tirée de W. Birkfellner, M. Figl, K. Huber, F. Watzinger, F. Wanschitz, J. Hummel, R. Hanel, W. Greimel, P. Homolka, R. Ewers, and H. Bergmann, "A head-mounted operating binocular for augmented reality visualization in medicine — Design and initial evaluation," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 8, pp. 991–997, Aug. 2002. ©2002 IEEE. Reproduit avec permission.) et vidéo (droite, tirée de C. Bichlmeier. "AW: AW: Permission to use an

- image," Courriel personnel (2010/07/01). ©2010 Christoph Bichlmeier. Reproduit avec permission.).....23
- Figure 1.10. Fenêtre de réalité augmentée (gauche, tirée de M. Blackwell, C. Nikou, A. DiGioia, et T. Kanade, "An image overlay system for medical data visualization," in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, 1998, pp. 232-240. ©1998 Springer. Reproduit avec permission. Droite, tirée de G. Fichtinger, A. Deguet, K. Masamune, E. Balogh, G. S. Fischer, H. Mathieu, R. H. Taylor, S. J. Zinreich, and L. M. Fayad, "Image overlay guidance for needle insertion in ct scanner," *Biomedical Engineering, IEEE Transaction on*, vol. 52, no. 8, pp. 1415–1424, Aug. 2005. ©2005 IEEE. Reproduit avec permission.) ...24
- Figure 1.11. Système d'assistance chirurgicale adapté du StealthStation (tirée de Y. R. Rampersaud et K. T. Foley, "Image-guided spinal surgery," *Operative Techniques in Orthopaedics*, vol. 10, no. 1, pp. 64-68, 2000. ©2000 Elsevier. Reproduit avec permission.).....26
- Figure 1.12. Exemple de visualisation proposée par Traub et al. (tirée de J. Traub, T. Sielhorst, S. M. Heining, et N. Navab, "Advanced Display and Visualization Concepts for Image Guided Surgery," *Journal of Display Technology*, vol. 4, no. 4, pp. 483-490, 2008. ©2008 IEEE. Reproduit avec permission.).....27
- Figure 1.13. Méthode de visualisation hybride in-situ (adaptée de C. Bichlmeier, F. Wimmer, S. M. Heining, et N. Navab, "Contextual Anatomic Mimesis Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality," in *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 1-10. ©2007 IEEE. Reproduit avec permission.).....28
- Figure 1.14. Exemple de données denses à représenter (tirée de M. Burns, M. Haidacher, W. Wein, I. Viola, et M. E. Gröller, "Feature Emphasis and Contextual Cutaways for Multimodal Medical Visualization," in *Data Visualization - EuroVis 2007, Norköpping, Sweden, 2007*, pp. 275-282. ©2007 IEEE. Reproduit avec permission.)29

- Figure 1.15. Impression de profondeur créée par l'ombre (tirée de D. Kersten, P. Mamassian, et D. C. Knill, "Moving cast shadows induce apparent motion in depth," *Perception*, vol. 10, no. 2, pp. 171-192, 1997. ©1997 Pion. Reproduit avec permission)30
- Figure 1.16. Pertinence de l'ombre pour estimer la distance de l'outil à l'organe (tirée de M. Nicolaou, A. James, B. P. L. Lo, A. Darzi, et G.-Z. Yang, "Invisible Shadow for Navigation and Planning in Minimal Invasive Surgery," in *Proceedings of the 8th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2005, pp. 25-32. ©2005 Springer. Reproduit avec permission.)31
- Figure 1.17. Impression de profondeur créée par la réflexion (adaptée de C. Bichlmeier, S. M. Heining, M. Feuerstein, et N. Navab, "The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 9, pp. 1498-1510, 2009. ©2009 IEEE. Reproduit avec permission.)32
- Figure 1.18. Perception de la profondeur provenant de l'ombre (gauche) et de la réflexion (droite) (tirée de C. Bichlmeier, S. M. Heining, M. Feuerstein, et N. Navab, "The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 9, pp. 1498-1510, 2009. ©2009 IEEE. Reproduit avec permission.)32
- Figure 1.19. Exemple d'endoscope stéréoscopique (adaptée de C. A. Dickman, D. J. Rosenthal, et N. I. Perin, *Thoracoscopic Spine Surgery*, New York: Thieme, 1999. ©1999 Thieme Medical Publishers, Inc. Reproduit avec permission.)33
- Figure 1.20. Illustration de l'ambiguïté concave/convexe (tirées de P. Doherty, "crater_illusion," *Crater Illusion*. [En ligne]. Disponible: http://www.exo.net/~pauld/summer_institute/summer_day1perception/crater_illusion.html. [consulté le 2010/07/13]. ©1999 NASA. Reproduit avec permission.)34
- Figure 1.21. Exploration par rotation du tube endoscopique38

Figure 1.22. Rotation possible du tube et de la caméra (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission)	39
Figure 1.23. Exemple du mouvement de rotation l'endoscope.....	40
Figure 1.24. Objet de calibrage utilisé par Girard (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission).....	47
Figure 1.25. Système Polaris® (tirée de Northern Digital Inc., "Passive Polaris Spectra System," <i>Passive Polaris Spectra System</i> . [En ligne]. Disponible: http://www.ndigital.com/medical/polarisfamily-spectrasystempassive.php . [Consulté le 2010/06/24]. ©2008 Northern Digital Inc. Reproduit avec permission.)	51
Figure 1.26. Instrument adapté avec un marqueur optique (adaptée de J. S. Hott, N. Theodore, J. D. Klopfenstein, S. Papadopolous, C. A. Dickman, et V. K. H. Sonntag, "Image guidance in spinal surgery for trauma: Iso-C 3D," <i>Operative Techniques in Neurosurgery</i> , vol. 7, no. 1, pp. 42-47, 2004. ©2004 Elsevier. Reproduit avec permission.).....	51
Figure 1.27. Système MicronTracker (tirée de LIV4D, "Nos installations," <i>Laboratoire d'Imagerie et de Vision 4D (LIV4D)</i> , 2010. [En ligne]. Disponible: http://www.polymtl.ca/liv4d/installation/equipement.php . [Consulté le 2010/03/16]. ©2010 LIV4D. Reproduit avec permission.)	52
Figure 1.28. <i>Xpoint</i> du MicronTracker (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," <i>Manuel d'utilisation</i> , 2009. ©2009 Claron Tech Inc. Reproduit avec permission.).....	53
Figure 1.29. Vecteur du MicronTracker (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," <i>Manuel d'utilisation</i> , 2009. ©2009 Claron Tech Inc. Reproduit avec permission.).....	53
Figure 1.30. Système de coordonnées du MicronTracker (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," <i>Manuel d'utilisation</i> , 2009. ©2009 Claron Tech Inc. Reproduit avec permission.).....	55

- Figure 1.31. Système de coordonnées d'un marqueur (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," *Manuel d'utilisation*, 2009. ©2009 Claron Tech Inc. Reproduit avec permission.) 56
- Figure 1.32. Représentation du modèle de rotation du tube de l'endoscope développé par Yamaguchi (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.) 59
- Figure 1.33. Placement des marqueurs pour l'estimation de l_s (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.) 60
- Figure 1.34. Comparaison du calibrage total de Yamaguchi à un calibrage partiel où seulement certains des paramètres sont calibrés alors que d'autres sont assumés respectant le cas idéal (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.) 62
- Figure 1.35. Erreur de reprojection en fonction de l'angle selon que le calibrage de l_h soit fait (gauche) ou non (droite) (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.) 63

Figure 1.36. Erreur de reprojection de la méthode de Wu comparée à aucune considération de la rotation (adaptée de C. Wu et B. Jaramaz, "An easy calibration for oblique-viewing endoscopes," in <i>IEEE International Conference on Robotics and Automation</i> , 2008, pp. 1424-1429. ©2008 IEEE. Reproduit avec permission.)	65
Figure 2.1. Transformation du repère du marqueur vers celui de la caméra. À noter que l'axe Z est représenté comme sortant de la caméra pour plus de clarté, alors qu'en réalité il entre dans cette dernière.	71
Figure 2.2. Position des marqueurs sur l'endoscope.....	73
Figure 2.3. Cercle décrit par un point du marqueur de la tête dans le repère du tube	74
Figure 2.4. Repère défini par rapport au cercle et décrivant ce dernier	75
Figure 2.5. Cercle permettant de déterminer l'angle associé à chacune des positions	75
Figure 2.6. Patron de l'objet de calibrage.....	80
Figure 2.7. Objet de calibrage réel	81
Figure 2.8. Objet de calibrage avec une bande noire plutôt que des bandes colorées.....	82
Figure 2.9. Outil d'enregistrement des images de calibrage.....	85
Figure 2.10. Construction permettant de trouver la couleur moyenne de la bande.....	86
Figure 2.11. Illustration du système de coordonnées sphériques (tirée de Wikimedia Commons, " File:Spherical Coordinates.svg," <i>Wikimedia Commons</i> , 2007. [En ligne]. Disponible: http://commons.wikimedia.org/wiki/File:Spherical_Coordinates.svg . [Consulté le 2010/07/10]. ©2007 Cronholm144. Dans le domaine public.)..	96
Figure 2.12. Vue en réalité augmentée (©2010 LIV4D. Reproduit avec permission)....	101
Figure 2.13. Augmentation du modèle virtuel par positionnement de l'image endoscopique	102
Figure 2.14. Vue augmentée avec instrument et sous-image endoscopique	103
Figure 2.15. MicroScribe utilisé pour numériser l'objet de calibrage	106
Figure 3.1. Exemple de vue générée pour les simulations	113
Figure 3.2. Graphique de l'erreur de reprojection en fonction de l'écart-type du bruit 2D ajouté.....	114
Figure 3.3. Graphique de l'erreur de reprojection en fonction de l'écart-type du bruit 3D ajouté.....	116

Figure 3.4. Graphique de l'erreur de reprojection en fonction de l'écart-type du bruit ajouté aux points de marqueurs	117
Figure 3.5. Erreur de reprojection dans le meilleur cas (haut à gauche), le pire cas (haut à droite) et le cas moyen (bas).....	126
Figure 3.6. Résultats sur la position du centre de distorsion	128
Figure 3.7. Comparaison des méthodes de compensation de la distorsion. De haut en bas : méthode de Wu sans calibrage, méthode de Wu avec calibrage, méthode proposée sans calibrage, méthode proposée avec calibrage	130
Figure 3.8. Comparaison de l'erreur de reprojection pour les deux méthodes utilisant un calibrage.....	131
Figure 3.9. Exemple de paire de vues pour la reconstruction 3D.....	132
Figure 4.1. Décalage dans les différentes sources de données en temps réel (tirée de T. Sielhorst, "New Methods for Medical Augmented Reality," Technische Universität München (TUM), 2008. ©2008 Tobias Sielhorst. Reproduit avec permission.)	143
Figure 4.2. Différentes approches au problème de synchronisation de sources multiples de données en temps réel (tirée de T. Sielhorst, "New Methods for Medical Augmented Reality," Technische Universität München (TUM), 2008. ©2008 Tobias Sielhorst. Reproduit avec permission.)	144
Figure I.1. Projection orthographique	162
Figure I.2. Projection perspective faible.....	163
Figure I.3. Projection perspective.....	163
Figure I.4. Dérivation des équations de projection perspective (tirée de R. Hartley et A. Zisserman, <i>Multiple View Geometry in Computer Vision</i> , 2 ^e éd.: Cambridge University Press, 2004. ©2004 Cambridge University Press. Reproduit avec permission.)	164
Figure I.5. Effet des pixels non carrés : $\alpha = 1$ à gauche, $\alpha \neq 1$ à droite	165
Figure I.6. Effet des pixels obliques : $s = 0$ à gauche, $s \neq 0$ à droite	166
Figure I.7. Centre optique par rapport à l'origine (tirée de R. Hartley et A. Zisserman, <i>Multiple View Geometry in Computer Vision</i> , 2 ^e éd.: Cambridge University Press, 2004. ©2004 Cambridge University Press. Reproduit avec permission.)	167

- Figure II.1. Deux types de distorsion (a : barillet; b : coussinet) (adaptée de J. Weng, P. Cohen, et M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965-980, 1992. ©1992 IEEE. Reproduit avec permission.) 171
- Figure II.2. Distorsion en barillet (gauche) et en coussinet (droite) (adaptée de Wikipedia, "Distorsion (optique)," *Distorsion (optique)*. [En ligne]. Disponible: [http://fr.wikipedia.org/wiki/Distorsion_\(optique\)](http://fr.wikipedia.org/wiki/Distorsion_(optique)). [Consulté le 2010/03/16] . Sous licence Creative Commons.) 172
- Figure II.3. Représentation de la signification géométrique des paramètres de la distorsion tangentielle..... 174
- Figure II.4. Grille de calibrage hexagonale non distordue et l'équivalent distordu permettant d'estimer le centre de distorsion par l'intersection des droites (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission) 176
- Figure II.5. Exemple de motifs originaux (adaptée de R. Miranda-Luna, W. C. P. M. Blondel, C. Daul, Y. Hernandez-Mier, R. Posada, et D. Wolf, "A simplified method of endoscopic image distortion correction based on grey level registration," in *International Conference on Image Processing*, vol. 5, 2004, pp. 3383-3386. ©2004 IEEE. Reproduit avec permission.) 179
- Figure II.6. Exemple de motifs avec distorsion (adaptée de R. Miranda-Luna, W. C. P. M. Blondel, C. Daul, Y. Hernandez-Mier, R. Posada, et D. Wolf, "A simplified method of endoscopic image distortion correction based on grey level registration," in *International Conference on Image Processing*, vol. 5, 2004, pp. 3383-3386. ©2004 IEEE. Reproduit avec permission.) 179
- Figure II.7. Grille de calibrage de Girard (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École

Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission).....	181
Figure II.8. Comparaison entre l'effet d'un déplacement du centre de distorsion et celui de la non-perpendicularité de la grille de calibrage	182
Figure IV.1. Image de départ.....	191
Figure IV.2. Image de tons de gris	192
Figure IV.3. Image après l'égalisation de l'histogramme	193
Figure IV.4. Image après la binarisation	194
Figure IV.5. Image après la dilatation	195
Figure IV.6. Image après l'identification des contours	196
Figure IV.7. Image après l'élimination de régions.....	197
Figure IV.8. Image avec le bon ensemble de régions connexes.....	198
Figure IV.9. Image avec les coins identifiés et ordonnés.....	199
Figure IV.10. Localisation précise du coin de la grille (tirée de G. Bradski et A. Kaehler, <i>Learning OpenCV</i> , Sebastopol: O'Reilly Media, 2008. ©2008 O'Reilly Media. Reproduit avec permission.).....	200
Figure IV.11. Résultat final après raffinement.....	201
Figure V.1. Grille avant (gauche) et après (droite) la rotation pour aligner la grille	202
Figure VI.1. Premier ensemble de 20 images pour les simulations	204
Figure VI.2. Second ensemble de 20 images pour les simulations	205

LISTE DES SIGLES ET ABRÉVIATIONS

2D	Deux dimensions, ou bidimensionnel.
3D	Trois dimensions, ou tridimensionnel.
AR	<i>Augmented reality</i> (Réalité augmentée)
AV	<i>Augmented virtuality</i> (Virtualité augmentée)
CCD	<i>Charge-coupled Device</i>
CMI	Chirurgie minimalement invasive.
CT	<i>Computed tomography</i> (Tomodensitométrie ou TDM)
DLT	<i>Direct linear transformation</i>
HMD	<i>Head-mounted display</i> (Casque d’affichage)
IRM	Imagerie à résonnance magnétique (<i>Magnetic resonance imaging</i> ou <i>MRI</i>)
LIV4D	Laboratoire d’Imagerie et de Vision 4D
MRI	<i>Magnetic resonance imaging</i> (Imagerie à résonnance magnétique ou IRM)
RA	Réalité augmentée (<i>Augmented reality</i> ou <i>AR</i>)
RMS	<i>Root mean square</i>
SVD	<i>Singular value decomposition</i>
TDM	Tomodensitométrie (<i>Computed tomography</i> ou <i>CT</i>)
VA	Virtualité augmentée (<i>Augmented virtuality</i> ou <i>AV</i>)

LISTE DES ANNEXES

ANNEXE I	DÉVELOPPEMENT MATHÉMATIQUE DU MODÈLE DE FORMATION DE L'IMAGE.....	161
ANNEXE II	DÉFORMATION DE LENTILLE.....	171
ANNEXE III	OUTILS MATHÉMATIQUES.....	184
ANNEXE IV	DÉTECTION AUTOMATIQUE DES COINS DE LA GRILLE	191
ANNEXE V	DÉTERMINATION DU CENTRE DE DISTORTION	202
ANNEXE VI	VUES GÉNÉRÉES POUR LES SIMULATIONS	204
ANNEXE VII	DÉTERMINATION DU CENTRE DE ROTATION.....	206

INTRODUCTION

La scoliose constitue une déformation tridimensionnelle de la colonne vertébrale pouvant entraîner certains problèmes de santé, notamment des douleurs au dos. Cette déformation, se propageant à la cage thoracique, peut aller jusqu'à nuire à la respiration des patients atteints et même entraîner des complications lors de l'accouchement [1]. De plus, chez les adolescents et adolescentes, étant à une période de leur vie où l'image est très importante, l'effet de la déformation scoliotique sur la surface externe du tronc peut causer une baisse d'estime de soi entraînant diverses conséquences psychologiques telles des pensées suicidaires ou une consommation d'alcool plus élevée [2].

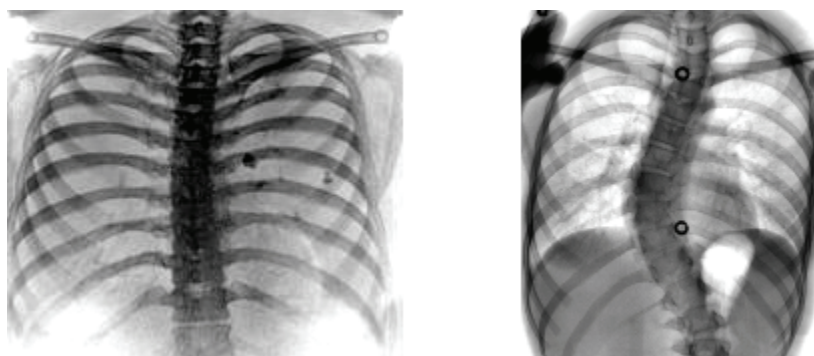


Figure 1. Exemple de radiographies de colonne en santé (gauche, adaptée de Wikimedia Commons, "File:DualEnergyAbsorptiometryofChestTissueSubtracted.png," *Wikimedia Commons*, 2007. [En ligne]. Disponible: <http://commons.wikimedia.org/wiki/File:DualEnergyAbsorptiometryofChestTissueSubtracted.png>. [Consulté le 2010/06/24]. ©2007 Kieran Maher. Dans le domaine public.) et de colonne scoliotique (droite, patient de l'Hôpital Ste-Justine)

Cette pathologie doit parfois être traitée de manière chirurgicale, mais la stratégie opératoire est déterminée par le chirurgien en fonction de la sévérité de cette pathologie ainsi que du type de correction envisagée. Le lecteur intéressé peut consulter le mémoire de maîtrise de Benboujja [3] pour une description plus détaillée des différentes procédures chirurgicales utilisées pour corriger la scoliose. Dans le cadre de ce projet, on se limite à la procédure de discectomie qui consiste en l'ablation partielle ou totale d'un disque intervertébral, utilisée pour redonner de la mobilité à la colonne vertébrale afin de pouvoir arriver à une plus grande correction de la déformation. À l'aide d'un instrument chirurgical, on retire petit à petit le disque intervertébral en totalité ou en partie. Cette intervention est délicate, car le chirurgien doit éviter de toucher des structures critiques avoisinantes telle que la moelle épinière ou

l'aorte, ce qui pourrait entraîner de graves complications. Elle peut être réalisée à l'aide d'une chirurgie classique, la thoracotomie, c'est-à-dire l'ouverture chirurgicale du thorax, pour laquelle on doit couper des tissus sains afin d'atteindre le site opératoire. Or, lors de ce type de chirurgie, le dommage fait pour atteindre le site chirurgical est souvent davantage responsable de la douleur, de l'inconfort et des autres problèmes postopératoires que la procédure en soi [4].



Figure 2. Cicatrice résultant d'une thoracotomie (gauche, adaptée de E. Popovic, "Thoracic disc prolapse," *Cervical spinal surgery, injury symptoms, neck pain treatment*. [En ligne]. Disponible: http://www.popovic.com.au/surgery_cervical.html. [Consulté le 2010/07/12]. ©2010 Emil Popovic. Reproduit avec permission.) et d'une thoracoscopie (droite, adaptée de Y. Qiu, B. Wang, et F. Zhu, "Comparison of the curative effects of video assisted thoracoscopic anterior corrected and small incision, thoracotomic anterior correction for idiopathic thoracic scoliosis," *Chinese Medical Journal*, vol. 121, no. 15, pp. 1369-1373. [En ligne]. Disponible: <http://www.cmj.org/Periodical/images/200807/2008730144442674.jpg>. [Consulté le 2010/05/05]. Sous licence Creative Commons.)

Grâce aux progrès récents dans le domaine de l'imagerie vidéo et dans le but de réduire le trauma causé aux tissus sains, une approche minimalement invasive peut désormais être utilisée pour effectuer une discectomie. Les chirurgies minimalement invasives (CMI), contrairement aux chirurgies classiques, ne nécessitent pas de grandes incisions. On utilise plutôt des petites ouvertures pour accéder au site opératoire avec les instruments chirurgicaux et pour insérer une caméra vidéo permettant de visualiser indirectement, via un moniteur, le site opératoire. Cette caméra est couramment nommée endoscope et les procédures l'utilisant chirurgies endoscopiques. Pour la discectomie, on procède à une thoracoscopie durant

laquelle on dégonfle temporairement un poumon afin de laisser un espace de travail permettant d'accéder à la partie thoracique de la colonne vertébrale. D'autres procédures sont également réalisées désormais de manière minimalement invasive, mais en utilisant plutôt une approche postérieure.

La réduction de la taille des incisions nécessaires pour effectuer la procédure entraîne de nombreux avantages reconnus pour les CMI [5, 6] : une réduction du volume de perte sanguine, de la douleur postopératoire, du séjour aux soins intensifs, du séjour à l'hôpital, des complications, du temps de convalescence et du coût total du traitement. Ces bénéfices sont obtenus sans pour autant compromettre les résultats de correction qui sont comparables à ceux obtenus à l'aide d'une approche par thoracotomie.

Bien que les nombreux gains redevables aux CMI encouragent le développement de ce type de chirurgie, plusieurs problèmes freinent malheureusement son déploiement généralisé. En effet, du point de vue des manipulations à effectuer :

- L'utilisation forcée d'instruments relativement longs limite la dextérité et les sensations tactiles
- Le point d'insertion de l'instrument ou de l'endoscope limite le mouvement
- Les gestes sont amplifiés par le principe de levier

De plus, au niveau de la visualisation :

- La visualisation du site opératoire se fait de manière indirecte via le retour visuel de l'endoscope sur un moniteur, ce qui peut s'avérer désorientant
- La vision monoculaire et la quasi-absence d'indices de distances entraînent une perte de la notion de profondeur
- La visibilité du site opératoire est réduite à cause de la proximité des lentilles (de l'ordre de 5 cm)
- La globalité de la perception est perdue puisque seule une petite partie du site opératoire est visible à la fois

Ces défis inhérents aux CMI augmentent de beaucoup le degré de difficulté de ces chirurgies. Elles sont extrêmement délicates à effectuer, spécialement lorsque le site opératoire est situé

très près de structures anatomiques critiques, comme lors d'une discectomie au cours de laquelle la proximité de la moelle épinière et de l'aorte, entre autres, est un facteur crucial. La courbe d'apprentissage est donc très longue. Ainsi, les avantages et les défis amenés par les CMI ont justifié le développement d'outils d'assistance pour ces procédures. Ces outils contribueront donc non seulement à la popularisation de ce type de chirurgies, mais amélioreront également son efficacité en facilitant l'apprentissage ainsi qu'en réduisant le risque d'erreur et le temps d'opération, qui constituent encore des défis de taille en clinique.

Afin de palier à l'insuffisance du contenu des images intra-opératoires, les systèmes d'assistance chirurgicale existants utilisent souvent un modèle préopératoire des structures d'intérêt comme complément à ces images en effectuant une fusion multimodale. Dans le contexte de ce projet, l'idée générale consiste donc à :

- obtenir, avant la chirurgie, de l'information provenant d'une modalité permettant de reconstruire un modèle 3D de la colonne vertébrale incluant le disque intervertébral.
- mettre à jour le modèle préopératoire afin qu'il soit cohérent avec le positionnement du patient dans la salle d'opération. Cette étape requiert généralement un certain calibrage du système.
- traiter l'information intra-opératoire afin de la fusionner avec le modèle préopératoire recalé au fur et à mesure de son acquisition tout au long de la chirurgie. Cette étape implique l'acquisition des images endoscopiques, mais requiert aussi généralement un système de suivi de l'endoscope, et parfois même un système de mise à jour du modèle préopératoire en fonction de la configuration géométrique de la colonne suite aux manœuvres chirurgicales, le tout dans le but d'assurer une fusion correcte, bien recalée, des images et des modèles préopératoires.

Le travail proposé s'inscrit donc dans le cadre du développement d'un système de réalité augmentée pour l'assistance chirurgicale pour les chirurgies minimalement invasives du rachis. Ce système, intégré à la salle d'opération et résumé à la Figure 3, permettra de combiner les images endoscopiques intra-opératoires avec un modèle 3D de la colonne préopératoire afin d'obtenir une vue fusionnée de ces deux modalités pour ensuite l'afficher au chirurgien en temps réel.

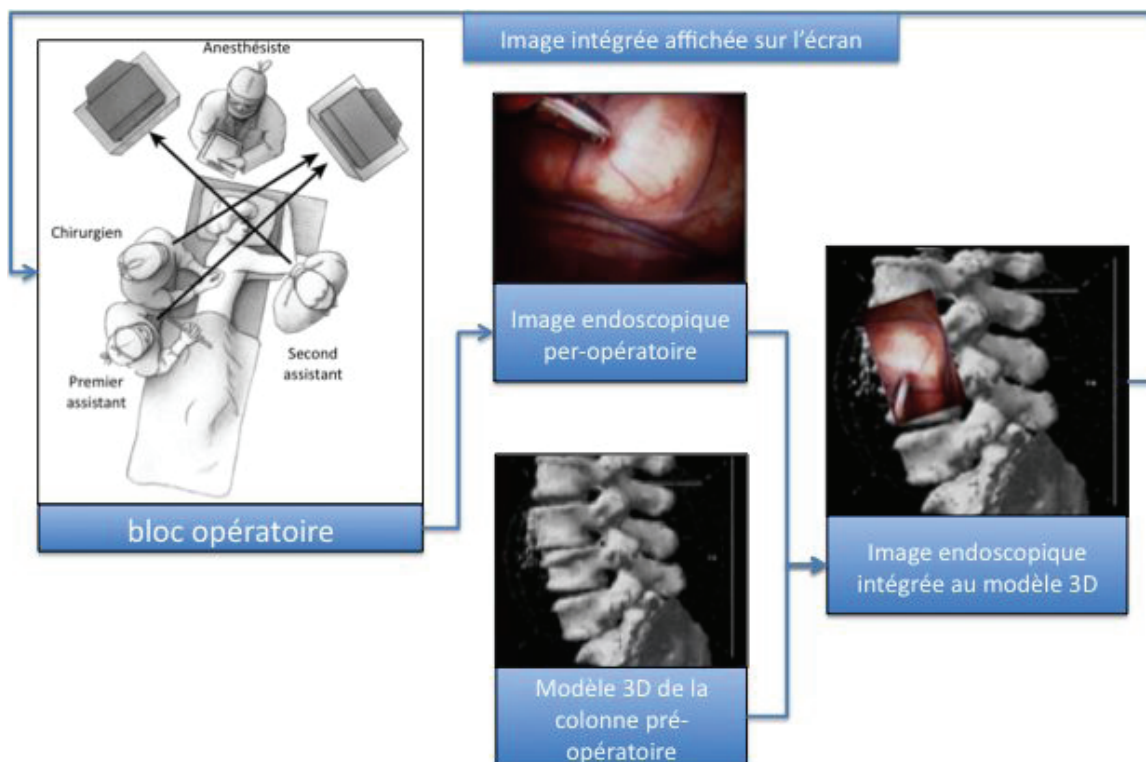


Figure 3. Vue d'ensemble du système d'assistance chirurgicale (©2010 LIV4D. Reproduit avec permission)

Grâce à ce procédé, les images endoscopiques peuvent être fusionnées au modèle préopératoire et ce, de différentes façons. On distingue deux types de vues qui peuvent être générées à partir des informations précédentes. Tout d'abord, on considère une vue comme étant de la réalité augmentée lorsqu'elle est générée en ajoutant de l'information, comme des modèles préopératoires projetés sur l'image ou des annotations, sur les images endoscopiques. Ainsi, les images endoscopiques, constituant la réalité, sont « augmentées » par l'ajout de l'information préopératoire. Puis, on considère une vue comme étant de la virtualité augmentée lorsqu'elle est générée en ajoutant l'image endoscopique à un monde complètement virtuel composé des modèles préopératoires vus d'un point de vue arbitraire. Ainsi, le monde virtuel, soit la virtualité, est « augmenté » par l'ajout de l'information intra-opératoire, soit les images endoscopiques.

Le système proposé, cherchant à être le moins invasif possible, vise :

- une acquisition du modèle préopératoire la moins invasive possible : on vise à minimiser les acquisitions ionisantes autant au niveau de l'acquisition des modèles

préopératoires que du recalage entre l'information préopératoire et intra-opératoire. Ainsi, une modalité non ionisante, soit l'imagerie à résonance magnétique (IRM), est utilisée pour reconstruire le modèle préopératoire et une seule radiographie intra-opératoire est utilisée pour la mise à jour du modèle.

- une intégration au protocole opératoire utilisé de routine durant les chirurgies minimalement invasives du rachis : on ne veut pas causer de trauma supplémentaire au patient par l'insertion d'instrumentation nécessaire au seul fonctionnement du système et non à l'intervention pratiquée; la procédure de calibrage se doit d'être la plus simple possible; etc.
- l'utilisation d'équipements médicaux couramment disponibles dans la majorité des hôpitaux plutôt que d'avoir recours à du matériel spécialisé coûteux.

Le développement de ce système a déjà été entamé dans le cadre de travaux réalisés par des membres du laboratoire de l'auteur, soit le LIV4D (Laboratoire d'Imagerie et de Vision 4D). Saint-Pierre [7] a développé une méthode de détection et de correction des réflexions spéculaires dans les images endoscopiques afin d'éliminer ces artéfacts problématiques pour de nombreux autres algorithmes d'analyse et de traitement de ces images. Windish [8] s'est attardé à la segmentation des instruments visibles dans les images endoscopiques, alors que Benboujja [3] a plutôt travaillé sur le suivi de ces instruments au cours d'une séquence d'images endoscopiques. Ces instruments ainsi suivis constituent des données utilisées en entrée par la méthode d'auto-calibrage développée par Girard [9] pour mettre à jour, et ce uniquement à partir du contenu naturel des images endoscopiques, les paramètres intrinsèques de la caméra qui varient durant la chirurgie. Par ailleurs, Chevretil propose une technique [10] de segmentation des disques intervertébraux dans les images IRM afin d'en obtenir une reconstruction 3D, ainsi qu'une technique [11] de recalage de ce modèle 3D à l'aide une radiographie intra-opératoire. Le présent travail de recherche s'attarde plutôt au calibrage hors-ligne de la caméra endoscopique qui doit être réalisé avant la chirurgie, à la mise à jour des paramètres variables au cours de la chirurgie et à la visualisation des données préopératoires et intra-opératoires fusionnées, tous nécessaires à l'implémentation du système d'assistance chirurgicale proposé.

Le présent mémoire rapporte le travail effectué pour la réalisation de ce système et est divisé en cinq chapitres. Le premier chapitre fait un survol des différentes méthodes existantes dans

la littérature s'attaquant aux problèmes découlant de la fusion multimodale, notamment le calibrage d'une caméra endoscopique, le suivi de cette dernière ainsi que le recalage des modèles 3D. Le second chapitre détaille la méthode de calibrage ainsi que les caractéristiques des paradigmes de visualisation proposés. Il décrit également une méthode de validation quantitative pour le calibrage et une méthode de validation qualitative pour les techniques de visualisation. Le troisième chapitre présente les résultats obtenus par l'application de la méthode proposée et discute ces derniers. Puis, le quatrième chapitre discute des limites de la méthode. Finalement, le cinquième et dernier chapitre conclut en formulant des recommandations pour les travaux à venir.

CHAPITRE 1 REVUE DE LA LITTÉRATURE

Ce chapitre vise à exposer les différents travaux antérieurs au présent travail de recherche en lien avec ce dernier. La présentation de ces travaux est structurée en fonction des différentes parties nécessaires au fonctionnement du système de réalité augmentée présenté à la Figure 1.1.

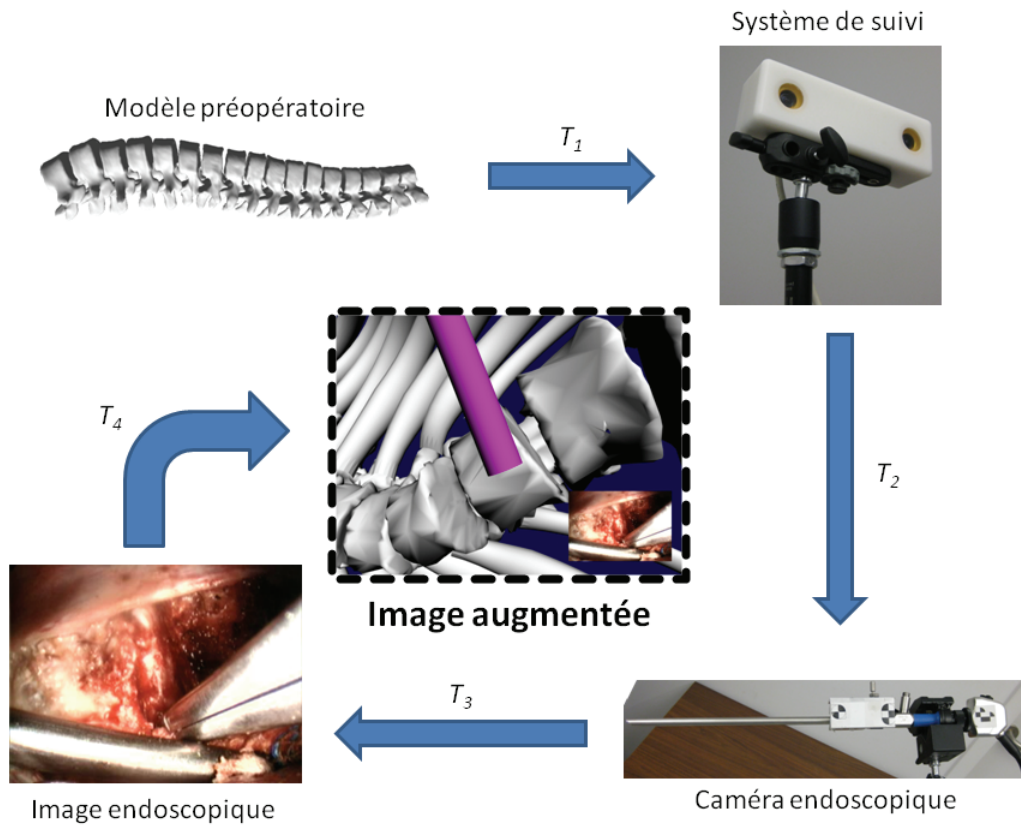


Figure 1.1. Étapes du système de réalité augmentée

En résumé, pour augmenter l'image endoscopique à l'aide d'un modèle préopératoire, on doit d'abord obtenir ce modèle préopératoire. Ce dernier est exprimé dans un repère quelconque. Il faut toutefois être en mesure de le situer dans l'espace de la salle d'opération afin de le positionner en 3D dans cet espace. Cette opération, représentée par la transformation T_1 , se nomme recalage des modèles préopératoires. Ensuite, un système de suivi de mouvement doit permettre de positionner la caméra endoscopique dans l'espace considéré comme référentiel global. Le suivi de la caméra permet donc d'obtenir la transformation T_2 , qui positionne la

caméra dans le repère global. Puis, la détermination du modèle de formation d'image par la caméra endoscopique permet d'exprimer la projection du modèle dans l'espace 2D de l'image endoscopique. Cette transformation vers l'image 2D endoscopique constitue la projection des modèles représentée par la transformation T_3 . Finalement, la visualisation de l'information résultant de cette fusion multimodale des images endoscopiques et du modèle préopératoire permet de guider le chirurgien dans son travail.

Ce chapitre présente donc les différentes techniques nécessaires à la réalisation de chacune des parties du système. D'abord, les méthodes de recalage des modèles préopératoires, correspondant à la transformation T_1 , sont présentées à la section 1.1. Ces modèles constituent des données que l'on doit fournir en entrée au système. Puis, la section 1.2 discute des techniques concernant la visualisation de modèle en tant qu'outil d'assistance chirurgicale, ce qui correspond à la transformation T_4 , et donc à la sortie du système. Ensuite, les caractéristiques d'une caméra endoscopique sont présentées brièvement à la section 1.3. Puis, le modèle général de formation de l'image correspondant à la transformation T_3 est défini à la section 1.4. Les véritables valeurs des paramètres de ce modèle peuvent ensuite être obtenues à l'aide d'un éventail de techniques présentées à la section 1.5. Par la suite, un survol des différentes technologies de suivi permettant de trouver la transformation T_2 est effectué à la section 1.6. La section 1.7 présente les différentes extensions nécessaires au modèle d'une caméra classique pour tenir compte des degrés de liberté d'une caméra endoscopique. Les limites des systèmes existants sont résumées à la section 1.8. Finalement, les objectifs de recherche sont présentés à la section 1.9.

1.1 Recalage de modèles préopératoires pour l'assistance chirurgicale

Afin d'être en mesure de fusionner les modèles préopératoires avec les images obtenues au cours de la chirurgie, une étape de recalage doit être faite. Cette dernière permet de faire correspondre le repère quelconque dans lequel sont exprimés les modèles préopératoires à un autre repère connu correspondant au patient au moment même de la chirurgie. Cette information permet donc par la suite de faire correspondre les modèles 3D avec les véritables structures anatomiques dans les images intra-opératoires.

Un bref survol des méthodes d'obtention des modèles préopératoires est d'abord fait avant de présenter les méthodes de recalage.

1.1.1 Reconstruction des modèles

L'obtention de modèles préopératoires dépasse le cadre du présent mémoire. Toutefois, afin de situer le lecteur, les principales méthodes sont brièvement résumées.

1.1.1.1 Tomodensitométrie

La tomodensitométrie permet de visualiser l'intérieur d'une coupe du corps humain en prenant une série d'acquisitions 2D, de rayons X par exemple, en tournant autour d'un axe perpendiculaire à cette coupe. Il est donc possible d'obtenir un volume en superposant plusieurs couches. En utilisant des rayons X, les structures osseuses apparaissent de manière très évidente. Il est donc par la suite facile de segmenter les vertèbres et d'obtenir un modèle 3D, par exemple en appliquant l'algorithme des « *marching cubes* » [12]. De nombreux travaux [13-19] rapportent l'utilisation de cette modalité pour l'obtention de leurs modèles préopératoires et l'appareil permettant d'acquérir ces images est représenté à la Figure 1.2.



Figure 1.2. Appareil de tomodensitométrie par rayons X (Tiré de Radiological Society of North America, Inc., "CAT Scan (CT) – Spine," *Images and Videos, CAT Scan (CT) – Spine quipment/procedure photograph*. 2010. [En ligne]. Disponible: <http://www.radiologyinfo.org/en/photocat/gallery3.cfm?image=philip25.jpg&pg=spinect>. [Consulté le 2010/06/30]. ©2010 Radiological Society of North America, Inc. Reproduit avec permission.)

Cette modalité permet d'obtenir une grande qualité de modèles, de par sa précision et sa résolution, ce qui explique sa popularité dans le domaine de la recherche. Par contre, elle expose le patient à imager à une grande quantité de radiations ionisantes, d'où son utilisation

plus fréquente sur un segment de la colonne plutôt que sur son ensemble. Ces problématiques sont acceptables en recherche sur des modèles fantômes ou des sujets cadavériques, mais les principes sous-jacents aux interventions minimalement invasives poussent à s'orienter vers d'autres modalités.

1.1.1.2 Radiographies multi-plans

Toujours basée sur les rayons X, une autre méthode de reconstruction de modèles 3D de la colonne implique l'utilisation de deux (ou davantage) radiographies. Cette méthode est présentement utilisée de manière routinière sur les patients scoliotiques pour reconstruire en 3D leur colonne vertébrale [20]. On identifie des primitives 2D comme des points (de 6 à 14 points par vertèbre) ou des courbes (pour la reconstruction des côtes) que l'on apparie sur les différentes radiographies, dont le modèle de formation a été déterminé par un certain calibrage. Il est alors possible de procéder à une triangulation permettant d'obtenir la position 3D des primitives 2D identifiées sur plusieurs images. Les trois vues les plus utilisées sont la vue PA0 (postéro-anterieur à 0 degré), PA20 (postéro-anterieur avec une angulation de 20 degrés, afin d'avoir une certaine disparité permettant la reconstruction 3D) et LAT (latéral), tel que représenté à la Figure 1.3. Elles permettent d'obtenir une reconstruction comme celle de la Figure 1.3.

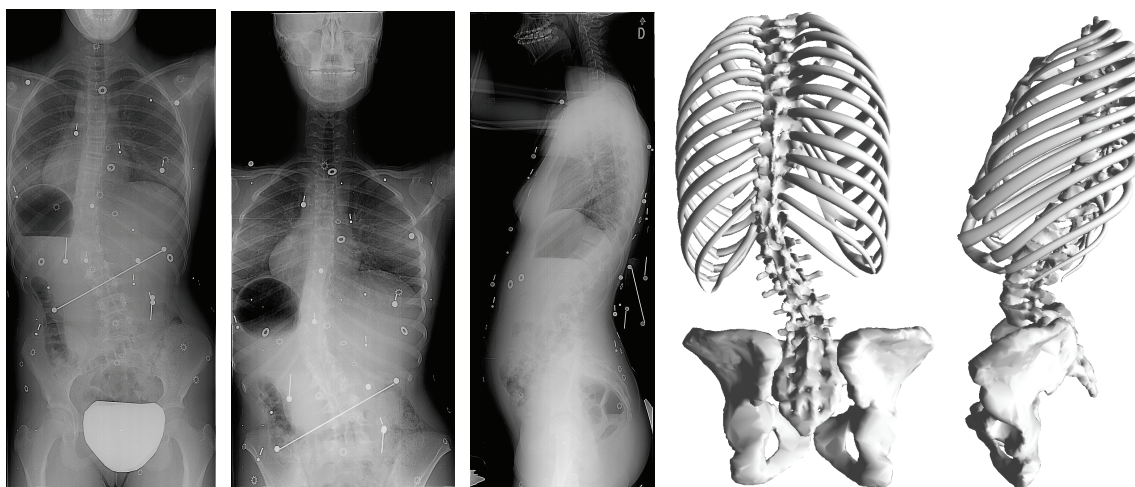


Figure 1.3. De gauche à droite : radiographie PA-0, PA-20, LAT; reconstruction vue de derrière, de gauche (patient de l'Hôpital Ste-Justine)

Des travaux récents ont été faits par rapport à ce type de reconstruction, notamment la possibilité d'utiliser des modèles de vertèbres a priori ainsi qu'une contrainte d'alignement

afin d'arriver à reconstruire la colonne vertébrale à partir d'une seule radiographie [21]. Un modèle statistique de variabilité de la colonne a également été développé [22] pour être en mesure de reconstruire plus facilement la colonne, même en présence de données partielles ou bruitées [23]. Aussi, des améliorations impliquant le protocole d'acquisition des radiographies, notamment le développement d'un nouvel objet de calibrage sous la forme d'un gilet [24], ont été réalisées. L'usage de plus en plus courant des radiographies multi-plans dans les différents centres hospitaliers poussent d'autres équipes indépendantes à tenter d'améliorer cette technique. On cherche notamment à aider l'expert réalisant la reconstruction en lui fournissant des outils nécessaires pour automatiser davantage cette tâche fastidieuse et réduire le temps nécessaire à cette dernière [25].

Bien que les radiographies multi-plans exposent le patient à un niveau de radiations moindre que la tomodensitométrie, une certaine exposition est toutefois nécessaire. Comme les patients atteints de scoliose sont souvent diagnostiqués à l'adolescence et suivis jusqu'à la maturité osseuse, le nombre de radiographies qu'ils devront subir est non négligeable. On cherche donc toujours à s'orienter vers des modalités encore moins invasives.

1.1.1.3 Imagerie à résonnance magnétique (IRM)

Cette méthode n'utilise pas de radiations ionisantes telles que les rayons X, ce qui la rend particulièrement intéressante dans un contexte minimalement invasif. Elle utilise plutôt des champs magnétiques qui alignent les atomes d'hydrogène présents dans l'eau que l'on retrouve dans les différents tissus. En simplifiant, on peut considérer que l'image est produite à partir du temps de relaxation, donc du retour à l'alignement naturel, des atomes d'hydrogène. Comme ce temps varie en fonction de la nature du tissu, il est donc possible d'obtenir des images présentant un excellent contraste entre les différents tissus mous du corps humain. Cette modalité est très utilisée pour l'imagerie du cerveau [26, 27].

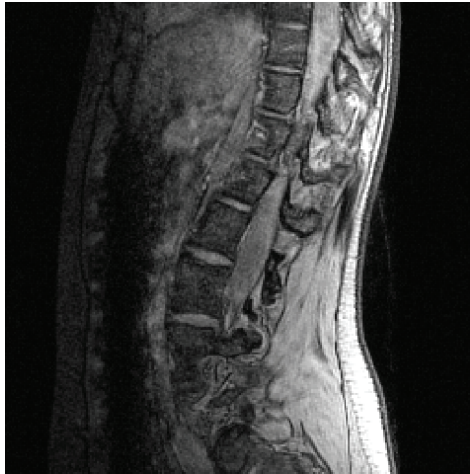


Figure 1.4. Exemple de coupe sagittale de volume IRM (patient de l'Hôpital Ste-Justine)

À l'instar de la tomodensitométrie, l'imagerie à résonnance magnétique donne l'image d'une coupe du corps, ce qui permet d'obtenir un volume en superposant les différentes tranches obtenues. Les voxels, généralement obtenus selon le plan axial et empilés, peuvent être visualisés à l'aide de coupes virtuelles selon d'autres plans, sagittal ou coronal par exemple, comme le montre la Figure 1.5. On peut distinguer sur ces images à la fois les os, mais également les tissus mous tels que les disques intervertébraux, qu'il est donc possible de segmenter automatiquement [10, 28] pour éventuellement les reconstruire en 3D. Des travaux sont en cours pour adapter ces méthodes de segmentation aux corps vertébraux, fournissant ainsi un modèle comportant à la fois les vertèbres et les disques intervertébraux.

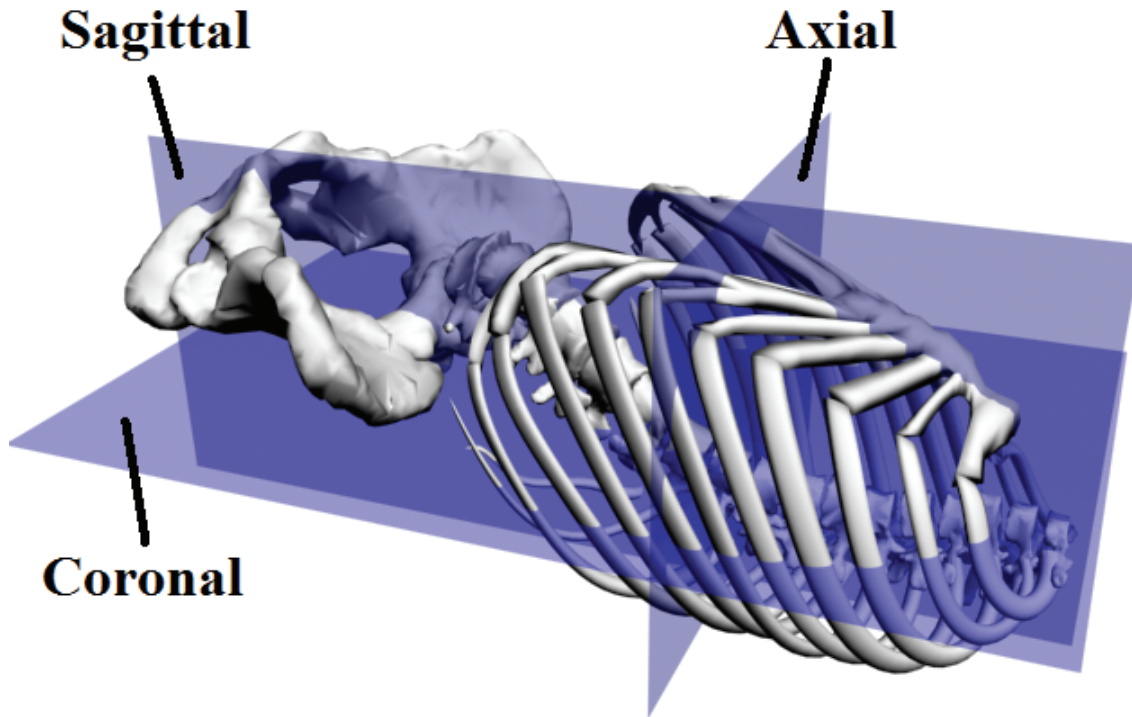


Figure 1.5. Plans axial, coronal et sagittal

Il est à noter que les objets ferromagnétiques introduisent une distorsion dans les champs magnétiques, alors que cette technique suppose la présence d'un champ magnétique uniforme. Ainsi, la présence d'instrumentation chirurgicale telle que des vis ou des tiges métalliques rendrait inefficace cette technique d'imagerie. Ce désavantage n'empêche pas l'utilisation de cette modalité afin de planifier la chirurgie. L'imagerie à résonance magnétique constitue donc une modalité très intéressante.

1.1.2 Recalage des modèles

Les modèles sont généralement reconstruits dans un repère quelconque correspondant à la méthode utilisée pour les obtenir. Or, afin de les fusionner adéquatement avec l'image endoscopique, il est important d'établir une relation entre ce repère quelconque arbitraire et celui dans lequel sont exprimés les paramètres de la caméra, que l'on choisit généralement en fonction du système de suivi.

1.1.2.1 Recalage point-à-point

La première stratégie de recalage consiste à identifier une série de points sur le modèle 3D (dans le repère du modèle), puis une série de points correspondants sur le véritable patient (dans le repère de la salle de chirurgie), et d'ensuite trouver la transformation rigide (une translation et une rotation) qui permet de faire passer le premier jeu de points au second. Mathématiquement, ce problème est considéré comme résolu [29].

Cette méthode tire son origine de la chirurgie de type stéréotaxique, au cours de laquelle un système de coordonnées est fixé au corps humain afin de planifier plus précisément des interventions délicates. Ainsi, le système de coordonnées est soumis à la même acquisition de données (tomodensitométrie, radiographie, IRM, etc.) que les organes, ce qui permet d'identifier par rapport aux images une stratégie opératoire, qu'il est possible de reproduire ensuite sur le patient à l'aide du système de coordonnées commun à la fois aux images et au patient [18]. Le même principe peut s'appliquer à la réalité augmentée : un marqueur du système de suivi définit un système de coordonnées; ce marqueur est imagé en même temps que l'organe à examiner; comme les points du marqueur peuvent être exprimés à la fois dans le système de coordonnées du marqueur et celui du modèle reconstruit, il est facile de calculer la transformation rigide reliant un repère à l'autre. Cette méthode implique donc de fixer solidement un marqueur à l'anatomie du patient, par exemple au niveau du crâne [30].

La même méthode peut être appliquée à la colonne vertébrale [15, 16, 31, 32]. Il est en effet facile de remplacer une structure anatomique rigide comme le crâne par une structure également rigide comme une vertèbre. Cette procédure est utilisée lors de chirurgies ouvertes conventionnelles, où la fixation du repère représente un trauma supplémentaire de bien moindre importance que l'intervention elle-même. Cette approche est quand même très invasive et ajoute de la complexité au protocole opératoire. Le patient est ouvert chirurgicalement, ensuite on fixe le repère sur une vertèbre à l'aide d'une vis ou d'une pince, puis on soumet le patient à l'acquisition de données. Pour les chirurgies minimalement invasives, une telle approche est difficilement envisageable, puisque l'on n'a pas accès aussi facilement aux structures sur lesquelles on voudrait poser un marqueur.

De plus, cette méthode est davantage adaptée aux gros os, tels ceux du bassin ou des jambes. La colonne vertébrale, quant à elle, est composée de plusieurs vertèbres articulées, pouvant donc se déplacer les unes par rapport aux autres. Une structure comme la colonne vertébrale

pourrait donc nécessiter de fixer un marqueur sur chacune des vertèbres considérées pour suivre le mouvement relatif des vertèbres les unes par rapport aux autres. Ainsi, même si d'autres systèmes existent pour le cerveau par exemple, ils ne sont pas disponibles pour la colonne à cause de la complexité de cette structure [33]. En effet, la colonne vertébrale est une structure à la fois rigide au niveau des corps vertébraux et flexible par le mouvement relatif des vertèbres, ce qui lui confère un grand nombre de degrés de liberté. On a parfois plutôt recours à un seul marqueur pour l'ensemble de la colonne, négligeant ainsi le mouvement relatif possible d'une vertèbre à l'autre, ce qui peut quand même s'avérer une approximation acceptable dans certains cas (voir Figure 1.6).

De plus, dépendamment de la position du patient, la relation géométrique entre la colonne vertébrale et la surface externe de la peau varie. Aussi, le type de visualisation doit être adapté à l'intervention à pratiquer, ce qui implique que le transfert d'un système existant à une application différente n'est pas direct.

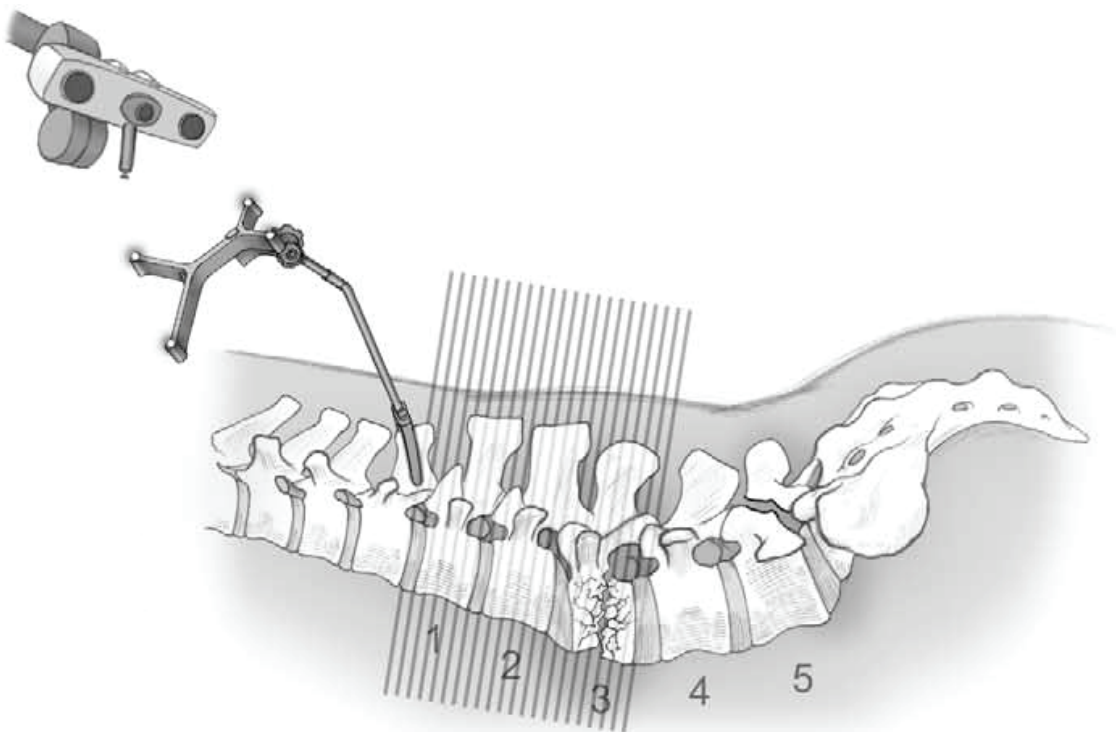


Figure 1.6. Exemple de marqueur optique fixé à une vertèbre (adaptée de J. S. Hott, N. Theodore, J. D. Klopfenstein, S. Papadopolous, C. A. Dickman, et V. K. H. Sonntag, "Image guidance in spinal surgery for trauma: Iso-C 3D," *Operative Techniques in Neurosurgery*, vol. 7, no. 1, pp. 42-47, 2004. ©2004 Elsevier. Reproduit avec permission.)

Il est également possible d'identifier, dans les deux systèmes de coordonnées, des repères anatomiques naturels plutôt que d'avoir recours aux marqueurs d'un système de suivi. Ainsi, en identifiant les points à la fois dans le modèle 3D virtuel et sur le patient à l'aide d'un pointeur 3D sur lequel est généralement fixé un marqueur de suivi, il est possible d'effectuer le recalage. Toutefois, cette méthode ne permet pas une mise à jour de la position du modèle en temps réel comme il est possible de le faire avec un marqueur suivi. De plus, dans le cas des chirurgies minimalement invasives, le chirurgien n'a accès qu'aux corps vertébraux qui ne présentent pas, comme le derrière de la colonne où l'on retrouve les processus transverses et épineux, de repères anatomiques facilement identifiables.

Ces méthodes cherchent à associer des points 3D du modèle avec des points 3D de l'espace de la salle d'opération. Il s'agit donc de méthodes « points 3D à points 3D ». D'autres travaux [14, 34] rapportent plutôt l'utilisation de marqueurs cutanés qui sont à la fois facilement segmentés automatiquement sur les images de tomодensitométrie, et à la fois identifiables automatiquement par une caméra calibrée. Les marqueurs cutanés sont donc présents dans le repère du modèle 3D et dans le repère de la salle d'opération (via la caméra calibrée), permettant ainsi le recalage point-à-point du modèle. Ces méthodes sont donc plutôt « points 3D à points 2D ».

1.1.2.2 Recalage surface à surface

Il n'est parfois pas trivial d'identifier clairement des points à la fois sur le modèle préopératoire et sur le véritable organe. Si l'organe est relativement lisse, il est parfois plus naturel de chercher à recaler une surface avec une autre surface.

Ce problème a été traité par Nakamoto et al. [35, 36] pour le recalage du poumon. Lorsqu'une tumeur au poumon doit être opérée de manière minimalement invasive, le poumon est dégonflé afin de permettre les manipulations, ce qui a pour effet de déplacer la tumeur par rapport aux images du patient prises lorsque celui-ci respirait normalement. En procédant par chirurgie minimalement invasive, le chirurgien perd la possibilité de palper le poumon pour trouver précisément la tumeur. Il est donc souhaitable de recaler le modèle préopératoire, où le poumon est normal, avec le patient, dont le poumon est dégonflé. Ceci est fait en utilisant l'algorithme ICP [37] pour recaler des points numérisés sur le poumon avec ceux du modèle.

Cette numérisation est faite soit à l'aide d'un numériseur 3D [35] ou d'un laser [36]. Dans le premier cas, un instrument chirurgical possédant un marqueur correspondant à un système de suivi est balayé sur le poumon, permettant ainsi d'enregistrer une série de points 3D sur la surface du poumon. Dans le second cas, le laser est suivi et calibré de telle sorte que la position 3D de la droite du faisceau peut être déterminée en tout temps. Le laser est ensuite balayé sur le poumon à partir de l'extérieur, et le point est identifié dans l'image endoscopique. L'intersection de la droite du laser avec celle correspondant à la projection inverse du point dans l'image permet de trouver les coordonnées 3D du point lumineux situé sur la surface du poumon, réalisant ainsi une numérisation de la surface sans contact.

Ces deux méthodes de numérisation donnent respectivement des erreurs de 4.6 ± 2.0 mm et 3.6 ± 1.4 mm. Ces erreurs sont acceptables pour ce type d'intervention chirurgicale dans laquelle le déplacement du poumon est grand, mais sont trop élevées pour une application aux chirurgies de la colonne.

D'autres techniques [38] cherchent à recalibrer des images à résonnance magnétique avec de l'ultrason 3D en utilisant un critère de similitude à la fois basé sur l'intensité et le gradient des images. Toutefois, cette méthode est utilisée pour le cerveau, qui ne pose pas les mêmes problèmes que la colonne vertébrale. En effet, les os sont opaques aux ultrasons et causent beaucoup de réflexion spéculaire des ondes acoustiques, rendant inapplicable la méthode proposée. Cette modalité peut toutefois s'avérer utile pour effectuer le recalage, tel que discuté ci-après.

1.1.2.3 Recalage surface à contour

Une autre stratégie de recalage consiste à appliquer au modèle à recalibrer la transformation correspondant au recalage, puis à projeter ce modèle sur une image acquise à l'aide d'une autre modalité, image sur laquelle sont définis des contours associables au modèle à recalibrer. L'opération de recalage consiste ensuite à trouver la transformation donnant la meilleure correspondance entre le modèle projeté et les contours identifiés sur l'image.

Lavallée [39] développe une méthode basée sur ce principe de projection de la surface d'un modèle préopératoire en un contour sur une radiographie prise juste avant la chirurgie. Afin de trouver la transformation rigide recherchée, il procède à une minimisation non linéaire dont le critère est basé sur des cartes de distance 3D. Il définit en effet, pour chaque point de

l'espace du modèle, une distance à la surface du modèle. Cette distance est égale à la plus petite des distances du point de l'espace à chacun des points du modèle. Elle est considérée négative si le point est situé à l'intérieur du modèle. Il définit également une distance, elle aussi signée, d'une droite à la surface, qui est tout simplement la distance minimale de tous les points de la droite.

À chacun des pixels d'un contour identifié sur l'image 2D correspond une droite dans l'espace 3D de la caméra, qu'il est possible de transporter dans l'espace 3D du modèle grâce à la transformation recherchée. Si le carré de la distance signée de cette droite à la surface est nul, la droite est tangente à la surface et fait par conséquent bel et bien partie du contour. Le critère de minimisation est donc le carré de la distance à la surface des droites associées aux pixels faisant partie du contour. Ainsi, il suffit d'identifier (automatiquement ou non) sur l'image 2D quelques points appartenant au contour du modèle 3D à recalculer pour permettre à l'algorithme de rechercher la transformation qui fait en sorte que les droites associées à ces pixels soient bien tangentes au modèle. Cette méthode, développée il y a presque 20 ans, propose également de nombreuses stratégies visant à optimiser l'exécution de l'algorithme afin d'obtenir des temps de calcul acceptables à l'époque.

Weese [40] propose également une méthode permettant de recalculer un modèle 3D (obtenu à partir d'une tomodensitométrie) avec une radiographie en cherchant la transformation qui donnera la radiographie possédant le moins de structures une fois qu'on ait retiré de l'image le modèle projeté. Le modèle 3D est recalculé avec la transformation à évaluer, puis projeté de manière à simuler la radiographie, pour que le résultat soit finalement retiré de la radiographie originale. La transformation optimale est celle qui donne l'image finale dans laquelle le maximum de structures a été retiré. Cette méthode peut être utilisée pour effectuer le recalage d'un modèle préopératoire avec celui d'une radiographie prise du patient sur la table de chirurgie, en ajoutant à celle-ci des repères qui seront identifiables à la fois par le système de suivi et sur la radiographie. L'utilisation de tomodensitométrie implique toutefois une irradiation importante du patient. De plus, le temps de calcul des radiographies simulées est non négligeable, même si celles-ci commencent à être implantées sur processeur graphique.

Dans le même ordre d'idée, Chevrefils [11] développe présentement une méthode permettant de recalculer des modèles 3D préopératoires obtenus par IRM et représentés à l'aide de 6 points par vertèbre, avec une radiographie unique dans laquelle on identifie également 6 points par

vertèbre. Ce recalage est rendu possible grâce à l'application d'une contrainte d'alignement sur les vertèbres [21] ainsi que de l'utilisation d'un modèle de variabilité de la colonne vertébrale [23].

Ces méthodes ne nécessitent qu'une seule radiographie pour effectuer le recalage, ce qui constitue déjà une réduction des radiations auxquelles le patient est exposé par rapport, par exemple, à la réalisation d'une tomodensitométrie du patient avec un jeu de points facilement identifiables.

Tel que mentionné plus tôt, l'ultrason peut s'avérer une modalité intéressante pour effectuer le recalage, car il s'agit d'une modalité non invasive. En effet, en calibrant adéquatement une sonde ultrasonique, il est possible de recaler le modèle préopératoire avec les images de cette sonde, qui sont exprimables dans un repère connu puisque la sonde est calibrée.

Brendel et al. ont développé une méthode complète pour effectuer le recalage de modèles tomodensitométriques avec des images ultrasoniques. Cette méthode implique d'abord un processus de minimisation [41, 42] dans lequel les caractéristiques des ondes acoustiques sont considérées : étant donné la très haute spécularité de la réflexion des ondes acoustiques sur les os, on ne projette sur les images ultrasoniques que les parties du modèle étant presque perpendiculaires à l'onde incidente, car ces parties seront les seules clairement visibles sur l'ultrason. Les zones où les os sont présents apparaissant très claires dans l'ultrason, on cherche à trouver la transformation permettant au modèle de correspondre aux endroits les plus clairs, ce qui est exprimé par une fonction objective à minimiser. La méthode décrit ensuite une procédure d'initialisation de l'algorithme de minimisation avec un estimé initial acceptable [43], nécessaire au bon fonctionnement du recalage surface-contour et obtenu à l'aide d'un recalage grossier point-à-point. Finalement, on effectue des recommandations sur le protocole d'acquisition des images ultrasoniques [44] afin de s'assurer d'une qualité suffisante pour le bon fonctionnement de l'algorithme. La Figure 1.7 montre la sonde à ultrasons lors d'une acquisition, ainsi que le marqueur nécessaire au système de suivi optique.

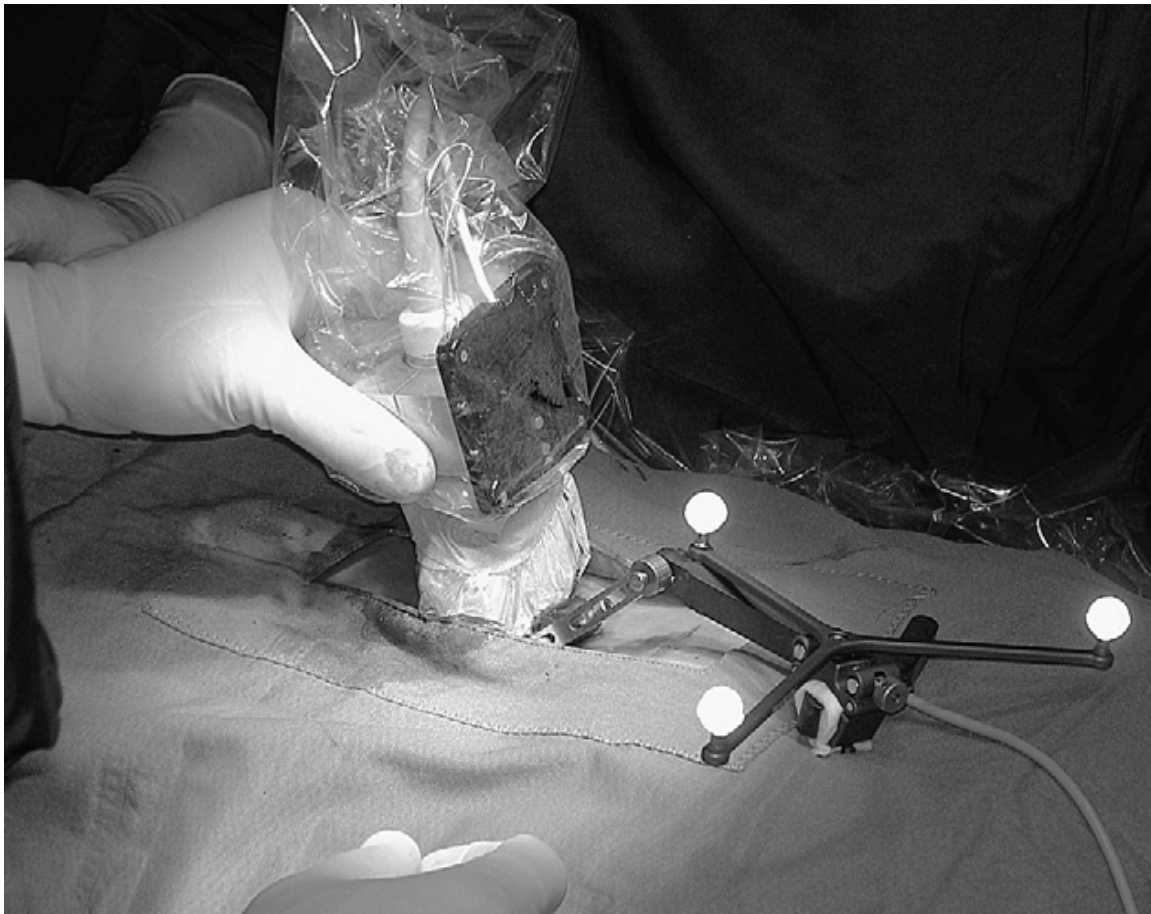


Figure 1.7. Acquisition d'images avec la sonde à ultrasons et son marqueur attaché (tirée de S. Winter, I. Pechlivanis, C. Dekomien, C. Igel, et K. Schmieder, "Toward Registration of 3D Ultrasound and CT Images of the Spine in Clinical Praxis: Design and Evaluation of a Data Acquisition Protocol," *Ultrasound in Medicine & Biology*, vol. 35, no. 11, pp. 1773-1782, 2009. ©2009 Elsevier. Reproduit avec permission.)

Les résultats de cette méthode, présentés à la Figure 1.8, sont très encourageants. Le recalage par ultrasons n'utilisant pas une modalité ionisante, il peut être effectué à plusieurs reprises au cours de la chirurgie. Ainsi, cette méthode pourrait éventuellement permettre de mettre à jour le modèle préopératoire pour tenir compte des changements de position de la colonne et des vertèbres les unes par rapport aux autres au cours de la chirurgie.

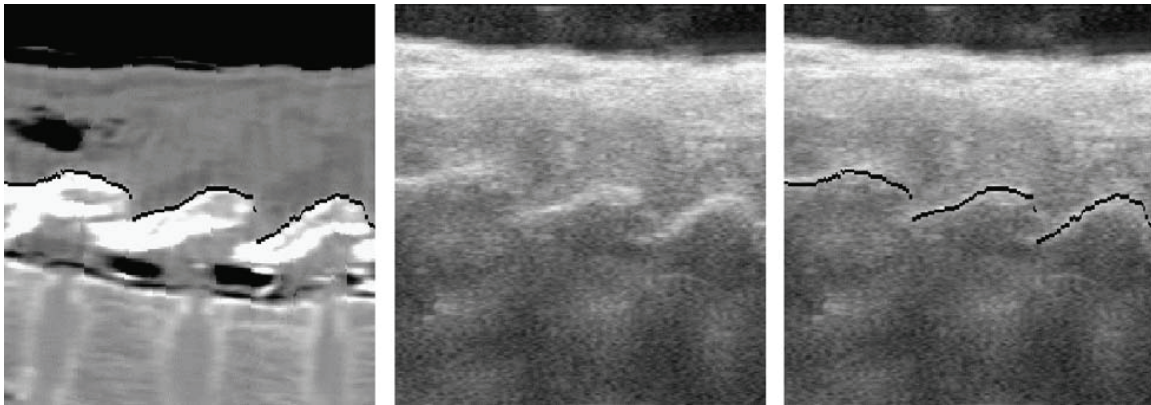


Figure 1.8. Résultat du recalage par ultrasons. Gauche : tranche sagittale de la tomodensitométrie où la surface des os est identifiée par des lignes noires. Centre : image ultrasonique correspondante. Droite : superposition de la surface identifiée, après recalage, et de la tranche ultrasonique. (Tirée de B. Brendel, S. Winter, A. Rick, M. Stockheim, et H. Ermert, "Bone registration with 3D CT and ultrasound data sets," *International Congress Series*, vol. 1256, pp. 426-432, 2003. ©2003 Elsevier. Reproduit avec permission.)

1.2 Visualisation 3D

Le choix de représentation visuelle pour l'assistance chirurgicale n'est pas trivial et a une importance cruciale sur l'utilité et la valeur ajoutée du système. De nombreux travaux ont donc été réalisés dans le but d'étudier différents points influençant l'efficacité de la visualisation de données 3D.

1.2.1 Technologies d'affichage

La chirurgie assistée par différentes méthodes d'imagerie a connu un énorme progrès au cours des dernières années, si bien qu'il a été jugé nécessaire d'effectuer un tour d'ensemble de ce domaine dans différents articles [45, 46]. Plus spécifiquement, les différentes technologies d'affichage destinées à la réalité augmentée ont également évolué très rapidement [47]. Le lecteur intéressé à une revue plus exhaustive des récentes avancées est invité à consulter ces travaux.

On a assisté à l'introduction du casque d'affichage, qui consiste en un dispositif d'affichage porté sur la tête [27]. On distingue principalement deux types de casques d'affichage permettant l'augmentation de la réalité : le casque d'affichage optique et le casque d'affichage

vidéo¹. Dans le premier cas, l'augmentation de l'image telle que vue naturellement par l'œil se fait à l'aide de « lentilles » du casque composées de miroirs semi-transparents. Ceux-ci permettent à l'utilisateur de voir à travers l'image qu'il verrait sans casque, mais permettent également de projeter une image virtuelle augmentée qui s'ajoute à cette vue naturelle. L'autre type de casque utilisé place devant les yeux de l'utilisateur deux écrans indépendants sur lesquels sont affichées les images provenant de deux caméras vidéo placées sur le casque. Ce type d'affichage donne l'avantage de simplifier la fusion des images réelles avec les parties virtuelles qui produisent l'augmentation. De plus, le délai entre l'augmentation et les images réelles, qui est très présent dans les casques optiques, est énormément réduit par le contrôle complet du contenu de l'image affichée sur les deux écrans.



Figure 1.9. Casque d'affichage optique (gauche, tirée de W. Birkfellner, M. Figl, K. Huber, F. Watzinger, F. Wanschitz, J. Hummel, R. Hanel, W. Greimel, P. Homolka, R. Ewers, and H. Bergmann, "A head-mounted operating binocular for augmented reality visualization in medicine — Design and initial evaluation," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 8, pp. 991–997, Aug. 2002. ©2002 IEEE. Reproduit avec permission.) et vidéo (droite, tirée de C. Bichlmeier. "AW: AW: Permission to use an image," Courriel personnel (2010/07/01). ©2010 Christoph Bichlmeier. Reproduit avec permission.)

Le casque d'affichage peut s'avérer encombrant et son poids peut entraîner une certaine fatigue lors d'une utilisation prolongée. Une autre technique [48] d'augmentation repose plutôt sur un miroir semi-transparent placé entre le chirurgien et le patient à opérer, comme on peut le voir sur la Figure 1.10. Ce système a l'avantage de ne pas nécessiter de suivi du point de vue du chirurgien. Les complications liées à la vision stéréoscopique du chirurgien

¹ « Casque d'affichage optique » est la traduction de « *optical see-through head-mounted display* » et « casque d'affichage vidéo » est la traduction de « *video see-through head-mounted display* ».

peuvent toutefois poser problème, tout comme l'encombrement que peut représenter la présence du miroir entre le chirurgien et le patient.

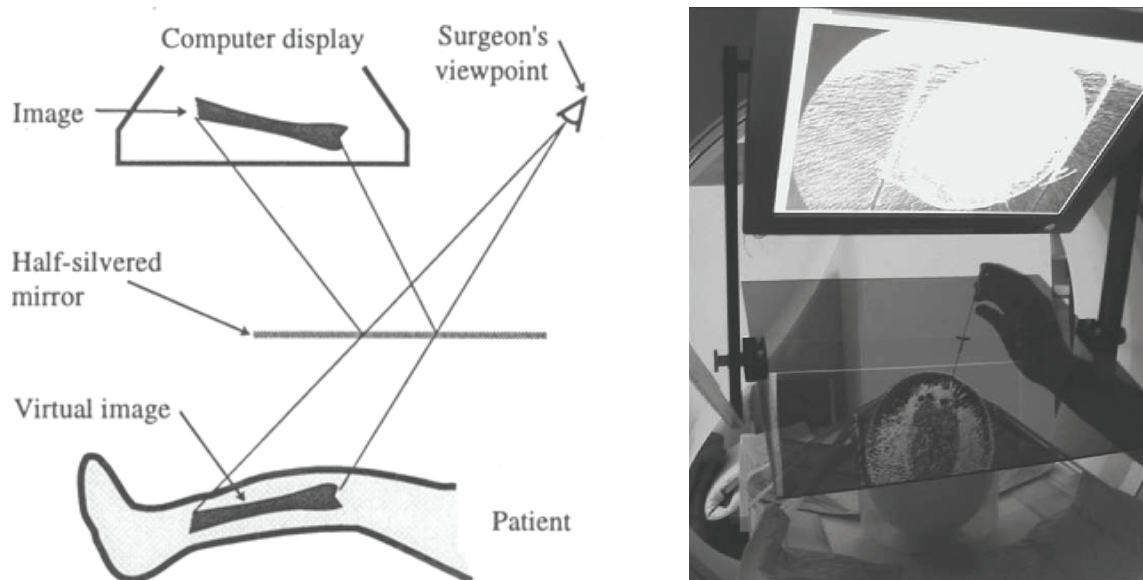


Figure 1.10. Fenêtre de réalité augmentée (gauche, tirée de M. Blackwell, C. Nikou, A. DiGioia, et T. Kanade, "An image overlay system for medical data visualization," in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, 1998, pp. 232-240. ©1998 Springer. Reproduit avec permission. Droite, tirée de G. Fichtinger, A. Deguet, K. Masamune, E. Balogh, G. S. Fischer, H. Mathieu, R. H. Taylor, S. J. Zinreich, and L. M. Fayad, "Image overlay guidance for needle insertion in ct scanner," *Biomedical Engineering, IEEE Transaction on*, vol. 52, no. 8, pp. 1415–1424, Aug. 2005. ©2005 IEEE. Reproduit avec permission.)

Les deux technologies précédentes s'appliquent surtout aux chirurgies conventionnelles au cours desquelles le chirurgien a un accès visuel direct au site opératoire. En effet, dans les deux cas, on cherche à augmenter ce que le chirurgien voit normalement directement avec ses yeux. Le principal bénéfice est donc que ce dernier continue de regarder au même endroit. Or, cet avantage ne s'applique pas vraiment aux chirurgies minimalement invasives, au cours desquelles le principal accès visuel au site opératoire se fait via une caméra endoscopique dont l'image est affichée sur un moniteur. La vue du corps externe n'apporte que très peu d'information utile à l'exécution de la procédure chirurgicale.

On utilise donc plutôt un troisième type de technologie d'affichage pour la réalité augmentée au cours de ce type de chirurgie, soit tout simplement le moniteur. En effet, l'image de la caméra endoscopique sera augmentée, plutôt que celle de la vision directe du chirurgien. Cette technologie nécessite toutefois l'utilisation d'un système de suivi de la caméra, en plus de ne pas permettre directement la vision stéréoscopique.

1.2.2 Paradigmes de visualisation

Les premiers paradigmes de visualisation amenés en salle d'opération ont été basés sur les mêmes principes que pour l'analyse d'images médicales à des fins diagnostiques, comme on peut le voir dans les vues choisies pour certains systèmes d'assistance de chirurgies conventionnelles ouvertes. Par exemple, le système d'assistance chirurgicale StealthStation, de Medtronic, généralement utilisé pour les chirurgies du cerveau, a été adapté pour les chirurgies de la colonne vertébrale [17]. L'affichage multi-plans adopté dans cette adaptation est très couramment utilisé pour l'exploration préopératoire d'images empilées en volume comme les tomodensitométries ou les images à résonnance magnétique. Toutefois, un système de suivi de l'outil chirurgical permet d'obtenir une coupe du volume perpendiculaire à la tête de la perceuse. On peut également obtenir des plans parallèles à l'axe d'une trajectoire planifiée ainsi qu'une représentation 3D du modèle préopératoire et de l'instrument. Bien que ce type de visualisation ne fusionne pas les images provenant de différentes modalités, il est intéressant de mentionner l'affichage multi-plans à cause de sa simplicité, de son efficacité et de son utilisation déjà courante.

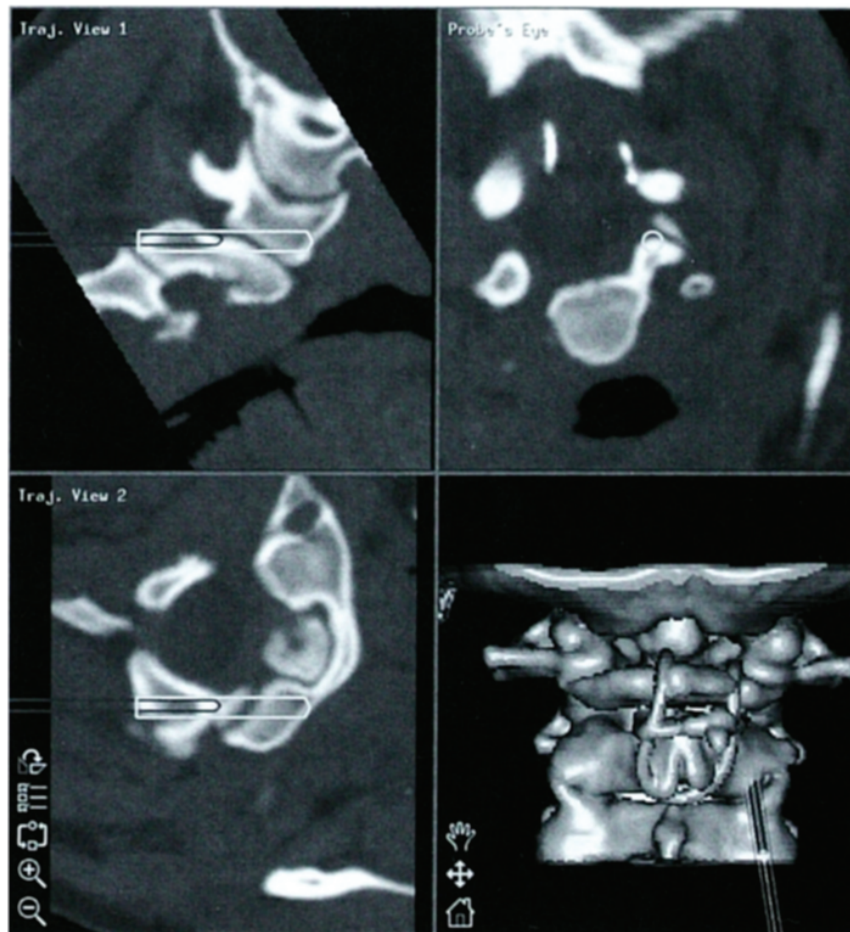


Figure 1.11. Système d'assistance chirurgicale adapté du StealthStation (tirée de Y. R. Rampersaud et K. T. Foley, "Image-guided spinal surgery," *Operative Techniques in Orthopaedics*, vol. 10, no. 1, pp. 64-68, 2000. ©2000 Elsevier. Reproduit avec permission.)

Le présent travail visant à fusionner l'information provenant de modèles préopératoires à celle provenant des images endoscopiques, la réalité augmentée constitue un paradigme de visualisation de choix. En effet, cette méthode permet de profiter de l'information provenant des deux modalités à la fois. La remise en un contexte commun de même que la possibilité de voir ensemble des informations présentes sur seulement une des deux modalités contribue à l'intérêt pour la réalité augmentée [47]. Une variante de la réalité augmentée a été introduite [26] sous le nom de virtualité augmentée. En virtualité augmentée, l'image produite se trouve dans l'espace virtuel du modèle auquel on ajoute, par la transformation appropriée, l'image réelle, alors qu'en réalité augmentée, l'image produite se trouve dans l'espace réel de l'image provenant de la caméra à laquelle on ajoute le modèle virtuel.

Il est important de comprendre que la visualisation la plus efficace ne dépend pas seulement des données à afficher, mais également de la nature de la tâche spécifique à accomplir à l'aide de ces données, tel que souligné par Traub et al. [19]. Dans ce travail, ils comparent des méthodes de visualisation pour une tâche impliquant l'utilisation d'une perceuse pour l'atteinte de points précis situés à l'intérieur des os, donc invisibles à l'œil nu et uniquement visibles sur les images tomодensitométriques. Ils comparent également l'utilisation d'un moniteur classique avec celle d'un casque d'affichage. La Figure 1.12 montre une des visualisations proposées. L'affichage multi-plans est repris, mais les images sont intégrées à l'image réelle provenant des caméras du casque d'affichage vidéo. Des tranches orthogonales provenant des images tomодensitométriques sont placées à la fois à l'endroit où la perceuse se trouve et en recouvrement de l'image totale. On ajoute aussi une coupe perpendiculaire à l'axe de la perceuse.

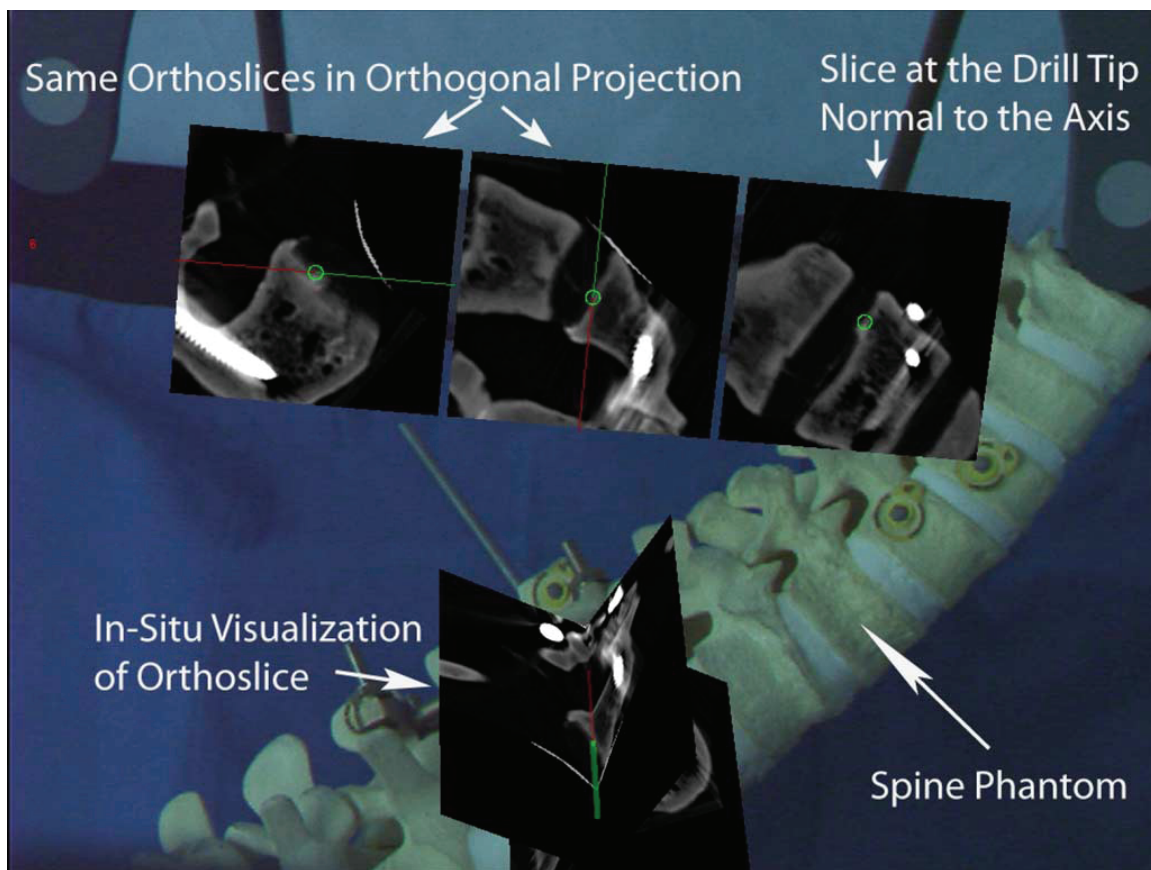


Figure 1.12. Exemple de visualisation proposée par Traub et al. (tirée de J. Traub, T. Sielhorst, S. M. Heining, et N. Navab, "Advanced Display and Visualization Concepts for Image Guided Surgery," *Journal of Display Technology*, vol. 4, no. 4, pp. 483-490, 2008. ©2008 IEEE. Reproduit avec permission.)

Ils concluent que malgré le fait que le casque fournit une représentation instinctive selon les chirurgiens utilisateurs, il n'aide pas ces derniers autant qu'un affichage plus simple et plus précis sur un moniteur. Le type de visualisation doit donc être adapté à la tâche à accomplir.

Bichlmeier et al. ont également développé une méthode de visualisation utilisant un casque d'affichage : la méthode de visualisation hybride in-situ [14]. Cette méthode, représentée à la Figure 1.13, permet de retrouver une certaine « vision rayons X » afin de voir les organes internes au travers de la surface externe. Pour ce faire, l'image réelle est affichée avec une zone centrale transparente permettant de voir les modèles 3D recalés au bon endroit, donnant ainsi l'illusion d'être situés « derrière » l'image réelle. Plusieurs facteurs, soit la courbure de la surface, l'angle d'incidence du regard ainsi que la distance à la surface, influencent le niveau de transparence. Il est donc possible de visualiser à la fois les organes internes et les instruments chirurgicaux, qui sont également intégrés à la visualisation.

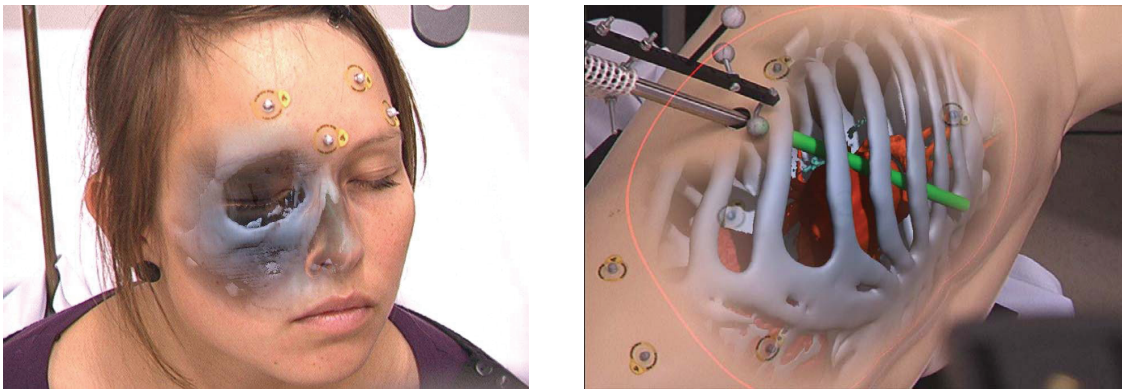


Figure 1.13. Méthode de visualisation hybride in-situ (adaptée de C. Bichlmeier, F. Wimmer, S. M. Heining, et N. Navab, "Contextual Anatomic Mimesis Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality," in *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 1-10. ©2007 IEEE. Reproduit avec permission.)

La manière de visualiser dépend également du type de données présentes. En effet, lorsque l'on est en présence de données très denses, il peut être difficile d'isoler les structures pertinentes à travers la surabondance d'information. Il est toutefois possible d'adapter les indices visuels en fonction des structures d'intérêts [49]. La Figure 1.14 représente bien cette technique, où une multitude d'organes est représentée en même temps, mais où le contraste est manipulé de manière à donner une grande importance aux os. De plus, ces derniers sont

placés devant le plan de l'image ultrasonique intégré à ce volume afin de permettre à l'observateur de situer géométriquement le plan par rapport aux os.

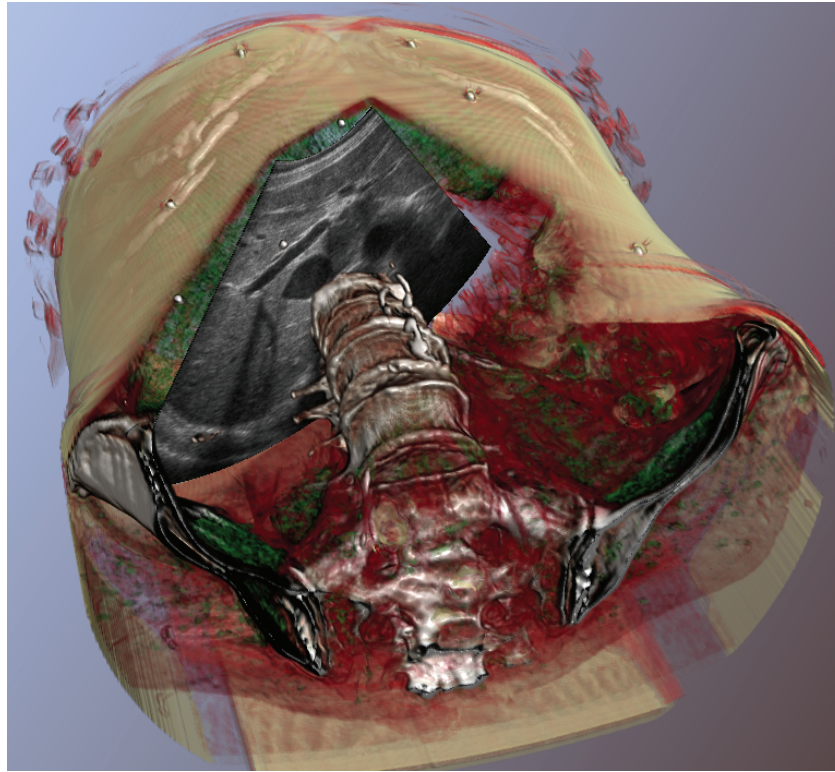


Figure 1.14. Exemple de données denses à représenter (tirée de M. Burns, M. Haidacher, W. Wein, I. Viola, et M. E. Gröller, "Feature Emphasis and Contextual Cutaways for Multimodal Medical Visualization," in *Data Visualization - EuroVis 2007, Norköpping, Sweden*, 2007, pp. 275-282. ©2007 IEEE. Reproduit avec permission.)

L'ordre de représentation des objets virtuels dans une image réelle est effectivement délicat, car l'occlusion est un indice important de la perception de la profondeur. Il est difficile d'estimer la profondeur des objets dans l'image réelle, si bien que parfois les objets virtuels sont affichés par-dessus les objets réels, alors que ces derniers devraient cacher les objets virtuels. Cette incohérence peut créer de la confusion et une désorientation spatiale chez l'observateur et constitue donc un facteur auquel il faut faire attention lors d'une représentation de réalité augmentée.

L'occlusion ne constitue qu'un des indices visuels permettant d'obtenir une perception de la profondeur. Cutting et al. [50] explorent en détail les différents indices de profondeur. Ils mentionnent entre autres l'ombrage des objets, qui est un facteur pouvant être exploité afin de redonner des indices de perception 3D. Une discussion complète faite à cet égard [51, 52]

donne plusieurs exemples d'indices aidant le cerveau humain à estimer la profondeur à partir de l'ombre, comme celui de la Figure 1.15. On y voit bien l'importance de l'ombre pour la perception de la position 3D des objets.



Figure 1.15. Impression de profondeur créée par l'ombre (tirée de D. Kersten, P. Mamassian, et D. C. Knill, "Moving cast shadows induce apparent motion in depth," *Perception*, vol. 10, no. 2, pp. 171-192, 1997. ©1997 Pion. Reproduit avec permission)

Or, avec l'utilisation d'un endoscope, la source lumineuse est pratiquement située au même endroit que la caméra, ce qui résulte en une quasi-absence d'ombres dans les images endoscopiques. Certains travaux [53] ajoutent une ombre simulée, comme on peut le voir sur la Figure 1.16, afin de permettre d'estimer plus facilement la position relative de l'instrument chirurgical par rapport au reste du site opératoire. Dans la méthode hybride in-situ mentionnée précédemment [14], on projette également une ombre produite par les instruments sur les modèles 3D afin d'aider à la perception des distances.

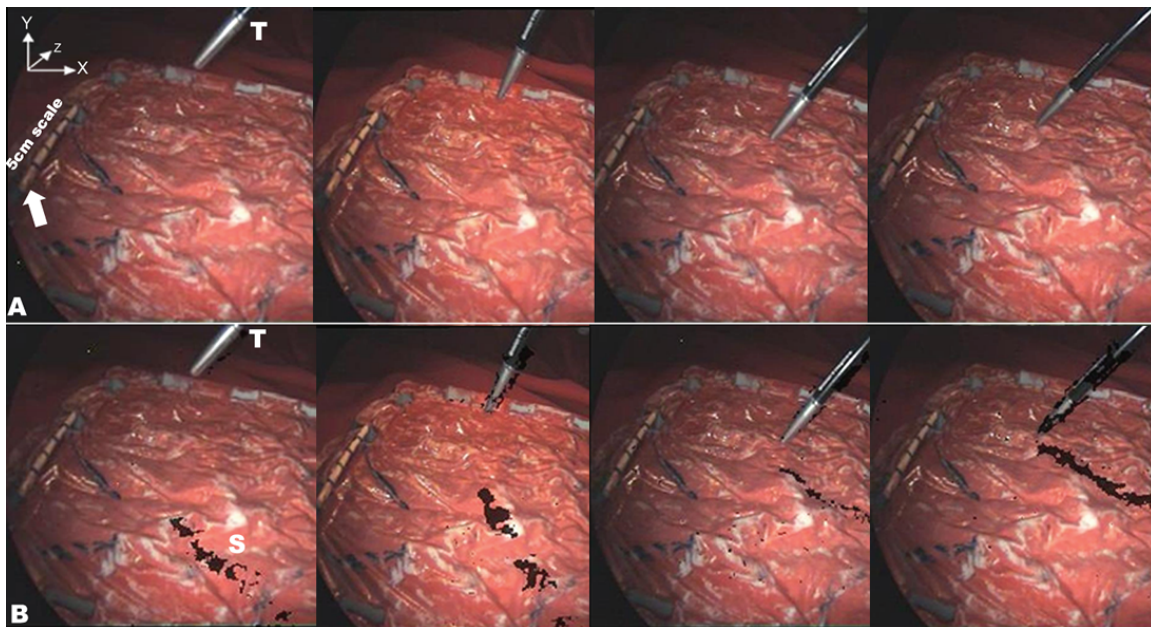


Figure 1.16. Pertinence de l'ombre pour estimer la distance de l'outil à l'organe (tirée de M. Nicolaou, A. James, B. P. L. Lo, A. Darzi, et G.-Z. Yang, "Invisible Shadow for Navigation and Planning in Minimal Invasive Surgery," in *Proceedings of the 8th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2005, pp. 25-32. ©2005 Springer. Reproduit avec permission.)

L'ombrage permet de redonner une perception 3D de la scène, car il permet d'établir une relation entre la position de l'objet causant l'ombrage et celle de l'objet sur lequel l'ombre est projetée. La réflexion d'un objet sur un autre permet d'obtenir un effet similaire, comme on peut le voir sur la Figure 1.17. Sur la scène à la gauche de cette figure, par la taille relative des objets, on a l'impression que la théière de droite est située devant celle de gauche. Toutefois, l'ajout des réflexions dans la scène de droite donne plutôt l'impression de l'inverse. Ainsi, le système de perception humaine se sert également de la réflexion des objets les uns sur les autres pour établir la relation spatiale entre les objets.

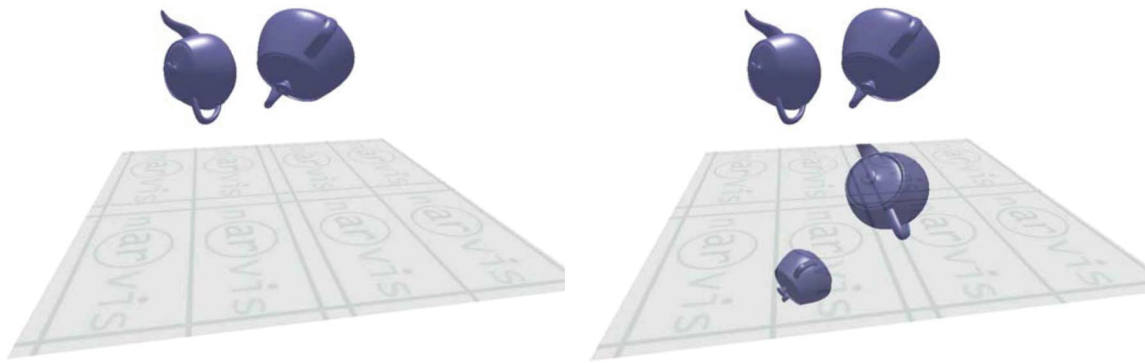


Figure 1.17. Impression de profondeur créée par la réflexion (adaptée de C. Bichlmeier, S. M. Heining, M. Feuerstein, et N. Navab, "The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 9, pp. 1498-1510, 2009. ©2009 IEEE. Reproduit avec permission.)

À partir de cette observation, le paradigme du miroir virtuel a été développé [13] et même évalué par une équipe de douze chirurgiens [54]. Le miroir virtuel est en fait un outil virtuel, donc visible uniquement à travers un casque d'affichage (le même que pour la méthode de visualisation hybride in-situ [14]), semblable au miroir utilisé par un dentiste pour examiner l'intérieur de la bouche. Il est manipulé d'une manière semblable afin d'être en mesure d'aller voir, toujours de manière virtuelle, derrière les objets considérés. Cette possibilité de positionner le miroir permet au chirurgien de placer et déplacer celui-ci afin d'obtenir des indices sur la structure 3D. La Figure 1.18 montre deux manières de les obtenir : en observant l'ombre virtuelle créée par le miroir et en observant la réflexion des objets virtuels dans ce miroir également virtuel.

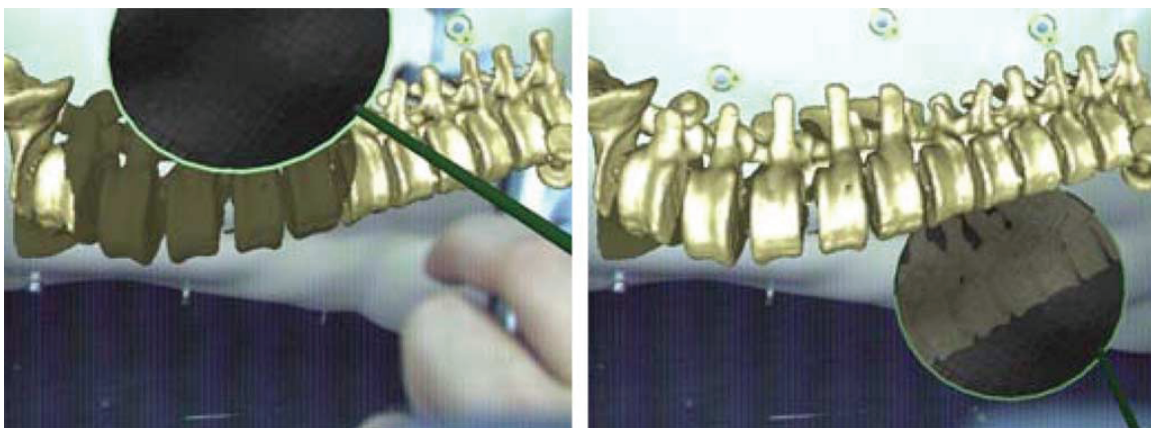


Figure 1.18. Perception de la profondeur provenant de l'ombre (gauche) et de la réflexion (droite) (tirée de C. Bichlmeier, S. M. Heining, M. Feuerstein, et N. Navab, "The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 9, pp. 1498-1510, 2009. ©2009 IEEE. Reproduit avec permission.)

Cet outil pourrait être adapté pour augmenter les images endoscopiques. En effet, il pourrait être attaché virtuellement à un outil chirurgical inséré dans le corps et visible par la caméra endoscopique, plutôt que par les caméras du casque d’affichage vidéo.

1.2.3 Autres outils d’aide à la visualisation 3D

Par ailleurs, certaines équipes ont étudié récemment la possibilité de visualiser les structures d’intérêt en 3D en se servant uniquement des séquences d’images intra-opératoires acquises par des systèmes endoscopiques.

1.2.4 Endoscopes stéréoscopiques

D’autres techniques se basent sur l’information disponible dans les images endoscopiques afin de retrouver des indices de profondeur sans avoir recours à une modalité supplémentaire. L’utilisation d’endoscopes stéréoscopiques a été proposée pour permettre de reconstruire des surfaces 3D de l’intérieur du corps à partir d’une paire d’images endoscopiques synchronisées, chacune étant prise d’un point de vue légèrement différent [55]. Toutefois, de par la configuration de l’endoscope, les deux points de vue sont très rapprochés l’un de l’autre, ce qui entraîne une faible disparité entre les images stéréoscopiques et donc de nombreux problèmes lors de la reconstruction. De plus, la présence de nombreux éléments tels le bruit vidéo, le sang, la vapeur, la buée et les réflexions spéculaires à même les images endoscopiques rend encore plus difficile la tâche de mise en correspondance d’un nombre minimal de primitives entre les images stéréoscopiques. L’application dans un contexte clinique est donc extrêmement ardue.

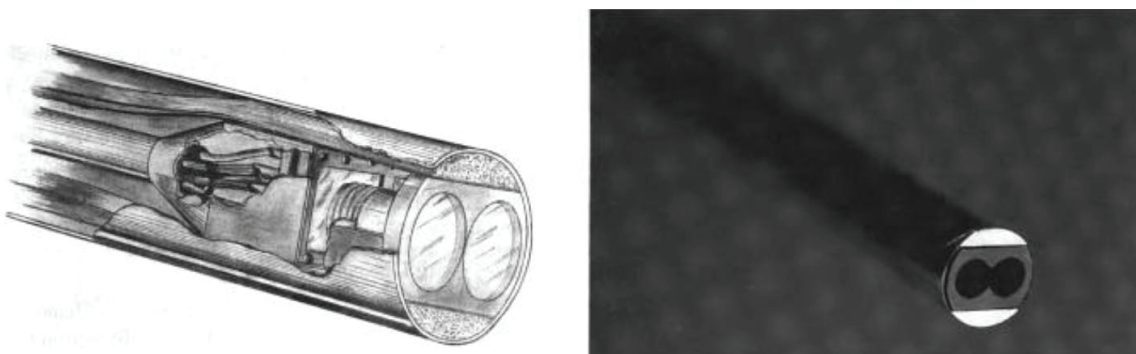


Figure 1.19. Exemple d'endoscope stéréoscopique (adaptée de C. A. Dickman, D. J. Rosenthal, et N. I. Perin, *Thoracoscopic Spine Surgery*, New York: Thieme, 1999. ©1999 Thieme Medical Publishers, Inc. Reproduit avec permission.)

1.2.5 Structure à partir de l'ombrage

Par ailleurs, les algorithmes de structure à partir de l'ombrage² constituent une autre technique basée uniquement sur le contenu des images endoscopiques. Cette méthode relativement récente de reconstruction 3D se base sur la variation de l'ombrage sur une surface donnée. En effet, comme la quantité de lumière réfléchiée par une surface de fini mat dépend beaucoup de l'angle entre la surface et la lumière incidente, il est possible d'obtenir de l'information sur le relief d'une forme visible dans une image grâce aux variations d'intensité lumineuse dans cette dernière. La Figure 1.20 montre une surface de laquelle on peut estimer le relief seulement à partir de la variation de l'ombrage. Il est intéressant de noter que la forme perçue dépend beaucoup de l'illumination perçue. Ainsi, comme on peut le voir sur la Figure 1.20, la même image peut être perçue comme possédant des creux ou des bosses, selon le sens observé. Ce phénomène, connu sous le nom d'ambiguïté concave/convexe, est dû à la perception de la position apparente de la lumière qui n'est pas la même sur chaque image.

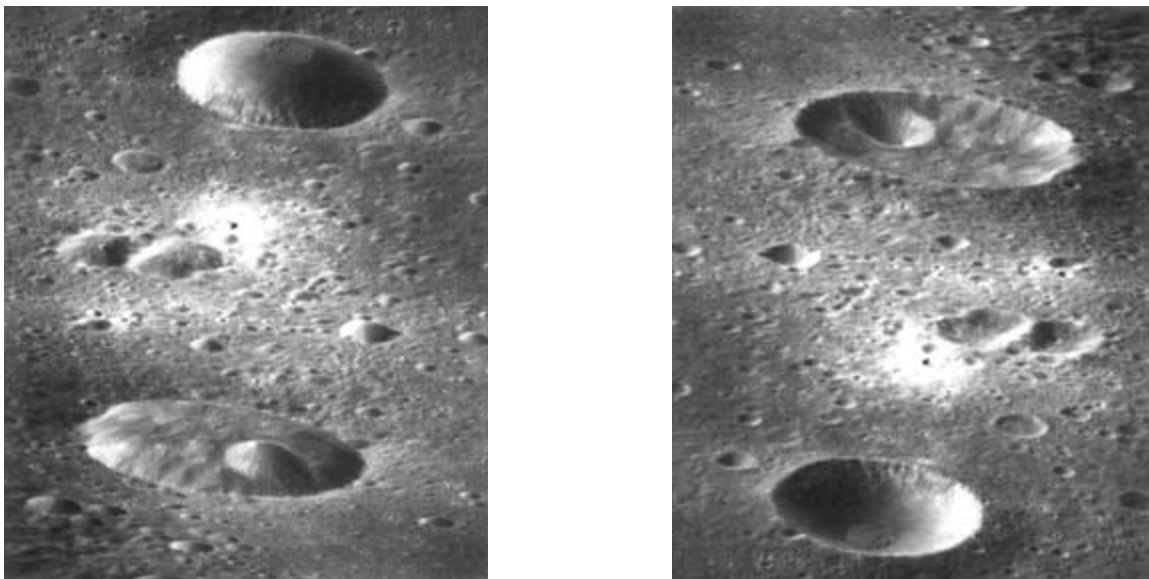


Figure 1.20. Illustration de l'ambiguïté concave/convexe (tirées de P. Doherty, "crater_illusion," *Crater Illusion*. [En ligne]. Disponible: http://www.exo.net/~pauld/summer_institute/summer_day1perception/crater_illusion.html. [consulté le 2010/07/13]. ©1999 NASA. Reproduit avec permission.)

Ce phénomène montre que la tâche de reconstruction de surfaces 3D uniquement à partir de l'information de l'ombrage est un problème difficile. Pour cette raison, le modèle de

² « Structure à partir de l'ombrage » est la traduction francophone de l'expression anglophone « *shape from shading* » plus utilisée dans la littérature.

formation de l'image supposé est généralement relativement simple. Beaucoup de techniques se basent sur un modèle d'éclairage lambertien, qui malheureusement constitue une sur-simplification de la réalité des images endoscopiques. Certaines techniques tentent de généraliser le modèle d'éclairage assumé afin d'obtenir de meilleurs résultats sur les images réelles. En effet, des travaux [56] ont été effectués dans le but de supporter d'une manière générique plusieurs modèles de réflectance. Toutefois, même les techniques qui tentent d'utiliser un modèle plus complexe, comprenant des réflexions spéculaires par exemple, n'arrivent pas à modéliser correctement la formation d'images endoscopiques :

- Les réflexions spéculaires sont souvent si lumineuses qu'elles saturent les pixels auxquels correspondent les parties de la surface sur lesquelles elles se trouvent, entraînant ainsi une perte de l'information potentielle concernant la structure de la zone saturée.
- Les inter-réflexions sont extrêmement difficiles à modéliser avec exactitude. Un rayon provenant de la source lumineuse n'éclaire pas seulement la première surface qu'il touche; il rebondit pour aller éclairer un autre point de la surface, spécialement dans un milieu petit, fermé et luisant comme celui des thoracoscopies.
- La source lumineuse n'est pas en réalité une source ponctuelle, mais possède plutôt une dimension non infinitésimale.
- La surface à l'intérieur du corps n'est pas uniforme, et donc ne réfléchit pas la lumière de manière uniforme.

Le modèle choisi est généralement traduit en une équation à résoudre, où les inconnues représentent la forme de la surface à reconstruire. À titre d'exemple, on peut choisir de représenter la forme comme une carte de hauteur, c'est-à-dire que l'on recherche la profondeur correspondant à chacun des pixels d'une image. Ainsi, pour une image de 320 pixels par 240 pixels, on parle de plus de 75 000 inconnues. Bien que cet exemple ne soit qu'à titre indicatif, le problème reste bien présent : le nombre d'inconnues est très élevé. La méthode de résolution du système d'équations différentielles découlant du modèle est donc cruciale pour l'obtention d'une solution correcte. Les différents algorithmes de reconstruction de la structure à partir de l'ombrage se distinguent généralement les uns des autres soit par le choix du modèle d'éclairage, soit par la technique de résolution de l'équation résultante, ou

les deux. La méthode de résolution de l'équation est donc très adaptée au problème et est difficile à généraliser à des modèles plus complexes, mais plus réalistes.

Par conséquent, on se retrouve en présence d'algorithmes fonctionnant bien sur des images synthétiques générées avec le modèle d'éclairage assumé [57], fonctionnant bien en environnement contrôlé [58, 59] ou donnant des résultats inacceptables dans un contexte clinique [60]. La technique de reconstruction de la structure à partir de l'ombrage n'est donc pas, à ce jour, adaptée aux images endoscopiques réelles.

1.2.6 Images intra-opératoires pour l'assistance de la chirurgie du rachis

Une modalité d'imagerie fréquemment utilisée pour l'assistance des chirurgies de la colonne vertébrale est la fluoroscopie. Cette dernière, basée sur les rayons X, permet d'obtenir des images internes au cours de la chirurgie afin de guider le chirurgien [61, 62]. Toutefois, cette modalité invasive expose le patient et l'équipe à une quantité non négligeable de radiations, malgré l'effort de certains travaux [63] pour réduire l'exposition aux radiations. Il existe également des systèmes basés sur la tomodensitométrie (TDM) [64] pour l'imagerie au cours de la chirurgie, mais le problème de radiation reste présent, en plus de nécessiter, dans certains cas, de l'équipement très particulier pour assurer un bon positionnement du patient.

Dans le cadre de ce projet, les images intra-opératoires proviennent d'une caméra endoscopique, ce qui amène à présenter les particularités de ce type de caméra, le modèle de formation de ce type d'images ainsi que les procédures de calibrage dans les sections suivantes.

1.3 Description générale de l'endoscope

Un endoscope est un instrument médical permettant d'examiner l'intérieur du corps humain sans avoir à pratiquer de très grandes incisions. Il comporte généralement un tube, rigide ou flexible, contenant un système optique permettant d'amener l'image d'une extrémité (celle à l'intérieur du corps) à l'autre (d'où l'image sera observée). Il possède habituellement un système d'illumination dont la source, située à l'extérieur du corps, transmet la lumière à l'intérieur du corps par un système de fibre optique souvent à même le tube. Finalement, ce système optique est généralement couplé à une caméra numérique permettant de visualiser l'image endoscopique sur un moniteur externe.

L'endoscope moderne doit satisfaire plusieurs exigences afin de s'avérer un outil chirurgical efficace [5] :

- Produire une image en couleurs, peu bruitée et de haute résolution
- Posséder un système d'éclairage ajustable
- Permettre de contrôler le facteur d'agrandissement de l'image
- Permettre de mettre au point l'image
- Offrir un jeu de manipulations facilitant l'exploration visuelle de l'intérieur du corps

Certaines de ces caractéristiques donnent aux chirurgiens des possibilités qui auront une grande influence sur le modèle de formation de la caméra endoscopique à développer ainsi que sur les méthodes de calibrage.

Certaines régions du corps humain, notamment la voie gastro-intestinale, sont beaucoup plus facilement accessibles à l'aide d'un endoscope flexible. Ce type d'endoscope, composé d'un tube de fibre optique capable de transporter l'image d'une extrémité à l'autre malgré la courbure, est relativement difficile à modéliser. En effet, il est ardu d'effectuer le suivi de la position du bout de l'endoscope, procédure nécessaire à l'établissement des paramètres empiriques du modèle de formation de l'image d'un endoscope.

On utilise donc plutôt les endoscopes de type rigide avec les systèmes d'assistance chirurgicale. Le tube de ces derniers, étant rigide plutôt que flexible, contient un système de lentilles acheminant les images d'une extrémité à l'autre. Si l'axe optique³ de la caméra est aligné avec l'axe du tube, on qualifie l'endoscope de droit. Le champ de vision est situé directement « devant » l'endoscope. Par contre, l'introduction d'un angle entre l'axe de la caméra et le tube permet d'effectuer des manipulations intéressantes sur l'endoscope pour l'exploration visuelle de l'intérieur du corps.

En effet, pour illustrer la différence entre les deux types d'endoscope au cours d'une chirurgie, on peut comparer l'endoscope droit à un télescope inséré dans une ouverture circulaire. La rotation du télescope autour de son axe ne change pratiquement pas la région visible par l'observateur. Par contre, l'utilisation de périscope, comparable à un endoscope oblique, permet d'explorer beaucoup plus facilement l'environnement par une rotation du

³ Une définition formelle de l'axe optique est donnée à la section I.1.3.

tube du périscope. En réalité, l'angle d'un endoscope oblique n'est pas droit comme celui d'un périscope, mais varie plutôt entre 0 et 60 degrés [9]. Le déplacement de la vue rendu possible par cette inclinaison est démontré dans la Figure 1.21.

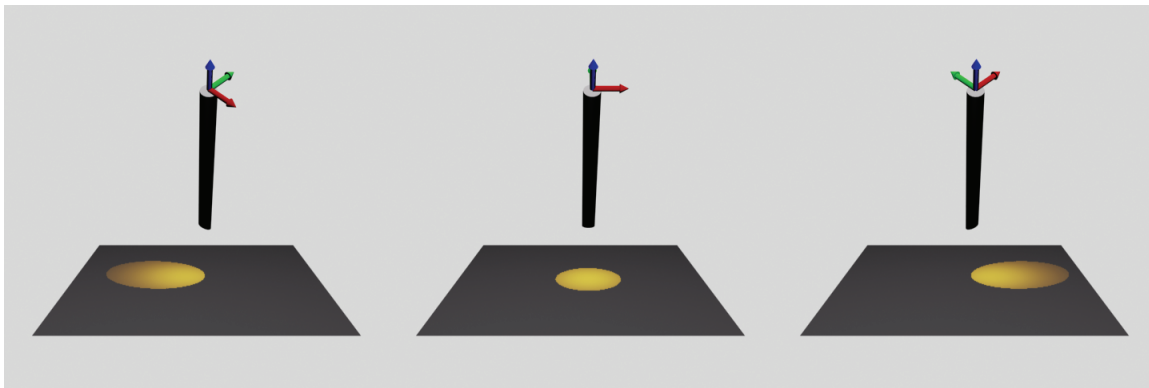


Figure 1.21. Exploration par rotation du tube endoscopique

Cette rotation entraîne, en plus du déplacement de la vue, une rotation de l'image endoscopique elle-même. Afin d'éviter que le chirurgien n'ait à se tourner la tête pour retrouver le sens instinctif de l'image, le tube de l'endoscope peut tourner par rapport à la caméra, comme on peut le voir sur la Figure 1.22. L'opération normale de l'endoscope consiste donc à maintenir la tête de la caméra relativement fixe dans la main, afin que l'orientation de l'image soit toujours pratiquement la même sur le moniteur vidéo. Puis, avec l'autre main, on peut manipuler le tube pour explorer, c'est-à-dire le faire tourner (toujours en maintenant fixe la caméra, ce qui fait que seul le tube tourne dans l'espace et par rapport à la caméra).

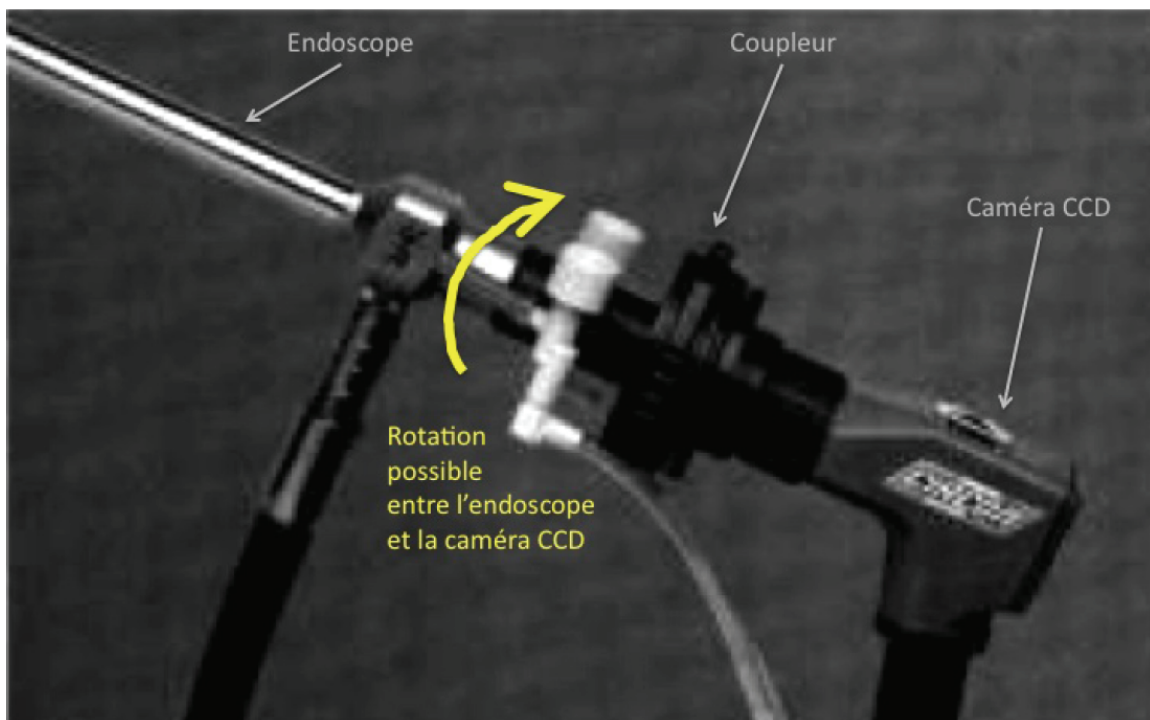


Figure 1.22. Rotation possible du tube et de la caméra (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission)

La Figure 1.23 montre un exemple de mouvement de l'endoscope. La colonne de gauche montre la position initiale de l'endoscope, permettant de visualiser la page gauche d'un livre. Puis, on effectue une simple rotation de tout l'endoscope, c'est-à-dire à la fois le tube et la tête de caméra sans rotation relative de l'un par rapport à l'autre, pour obtenir la configuration de la colonne du centre. On remarque donc que cette simple rotation (sans translation) qui est réalisable lorsque la caméra est insérée à l'intérieur du corps humain, permet maintenant de visualiser la page droite du livre. Cette exploration est rendue possible grâce à l'angulation de l'endoscope. Toutefois, on remarque que l'image de l'endoscope subit également cette rotation et se retrouve donc dans un sens peu instinctif. Pour corriger ce problème, on effectue une rotation de la tête de caméra seulement, en maintenant le tube fixe, pour obtenir la troisième configuration. Cette configuration est équivalente à la première à partir de laquelle on aurait effectué une rotation du tube de l'endoscope seulement, sans bouger la tête de la caméra, ce qui constitue l'utilisation normale de ce degré de liberté. Le chirurgien maintient la tête fixe en tournant le tube pour explorer le site opératoire. Cette manipulation permet donc de toujours garder l'image orientée d'une manière instinctive, car le maintien de

la tête de la caméra dans la même position assure que le haut de l'image endoscopique sera toujours le même.

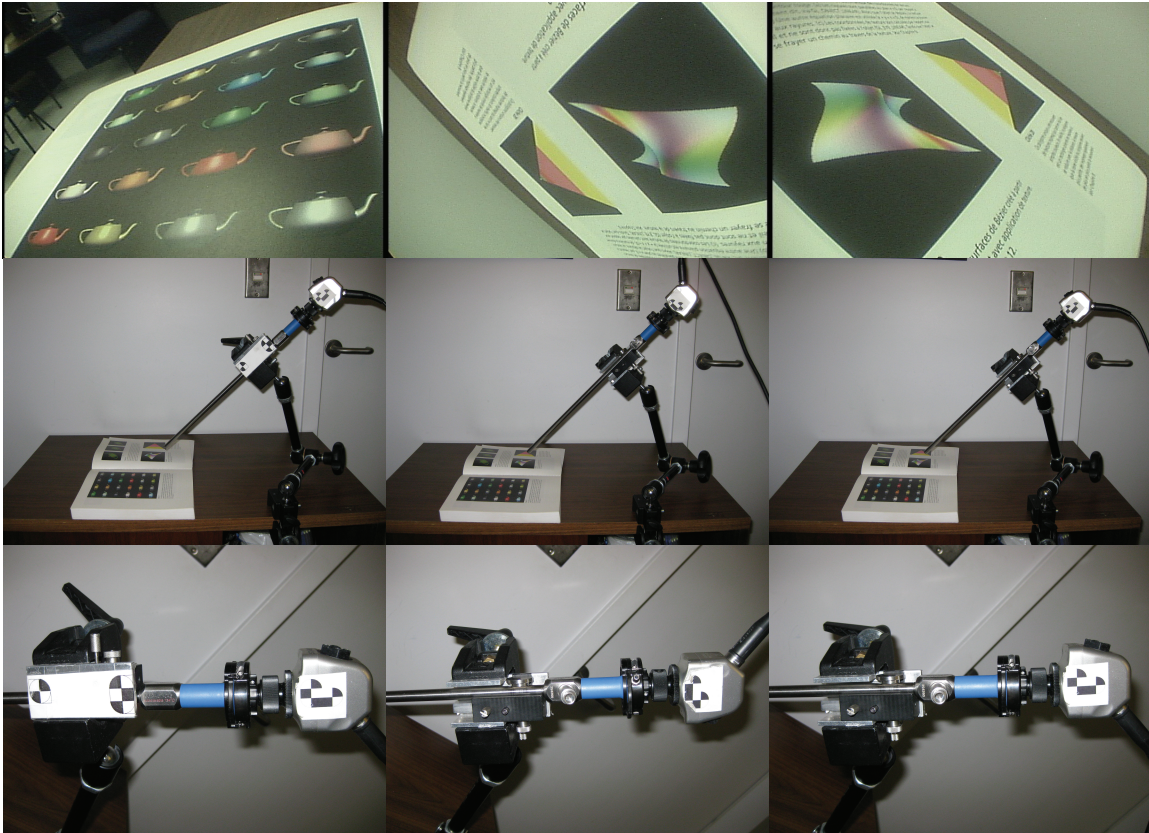


Figure 1.23. Exemple du mouvement de rotation l'endoscope

Ce degré de liberté supplémentaire, soit la rotation relative entre le tube et la tête de la caméra, ne se retrouve pas dans les caméras ordinaires et doit donc être pris en compte lors de l'établissement d'un modèle mathématique du modèle de formation d'images de l'endoscope, qui est développé dans la section suivante.

1.4 Modèle de formation de l'image endoscopique

Le modèle de formation de l'image permet de comprendre la manière dont une scène 3D quelconque visualisée à l'aide de la caméra en question est transformée pour former l'image 2D. D'un point de vue mathématique, il s'agit d'exprimer la transformation (ou la série de transformations) qui permet de transformer un point 3D de l'espace visualisé en un point 2D sur l'image produite. Le développement mathématique de ce modèle est fait en annexe (voir ANNEXE I) et ce dernier peut être résumé par l'équation suivante, reprise de l'annexe :

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K E_{c \leftarrow w} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1.1)$$

Ce modèle intègre deux types de paramètres, soit les paramètres intrinsèques et les paramètres extrinsèques.

Les paramètres intrinsèques présents sont :

- Les distances focales, généralement représentées soit par deux valeurs de distance focale, f_x et f_y , ou alors par une seule valeur ainsi que le rapport d'une valeur à l'autre.
- La position du centre optique, en pixels, dans l'image, soit x_0 et y_0 .
- Le facteur d'obliquité s , correspondant à un angle non droit entre l'axe horizontal et vertical des pixels de l'image.

Ces paramètres intrinsèques, regroupés dans la matrice K de l'équation 1.1, sont relativement fixes, car ils ne changent que lorsqu'une manipulation explicite est faite, par exemple une mise au point ou une modification du facteur de zoom. Il est donc généralement nécessaire de déterminer une seule fois la valeur de ces paramètres à l'aide d'une procédure de calibrage, et de conserver ces valeurs fixes.

Par contre, les paramètres extrinsèques, regroupés dans la matrice $E_{c \leftarrow w}$, correspondent plutôt à la position et à l'orientation de la caméra en 3D. Ainsi, le simple fait de déplacer la caméra endoscopique entraîne une modification des paramètres extrinsèques. Il sera donc nécessaire de mettre à jour ces paramètres en temps réel selon les déplacements de la caméra endoscopique. Les systèmes de suivi pouvant être utilisés sont décrits à la section 1.6.

La représentation matricielle (en coordonnées homogènes) de ce modèle est très pratique et efficace. Elle ne permet toutefois pas de représenter les transformations non linéaires impliquées dans le processus de formation de l'image endoscopique. La principale transformation manquante est la distorsion causée par l'imperfection des lentilles, qui est décrite en annexe (voir ANNEXE II). Cette déformation est souvent modélisée par une fonction polynomiale du rayon d'un point de l'image exprimée en coordonnées polaires :

$$r_d = \sum_{i=0}^n k_i r^i \quad (1.2)$$

où r est le rayon d'un point de l'image avant la déformation et r_d est le rayon du même point après l'application du modèle de déformation de lentille.

Cela conclut le survol du modèle de formation d'images généralement associé à une caméra réelle. La section suivante décrit les méthodes permettant de calibrer ce modèle.

1.5 Calibrage du modèle

Le modèle mathématique de la formation de l'image ayant été survolé, on s'attarde maintenant aux méthodes permettant de calculer les paramètres de ce modèle, c'est-à-dire d'assigner des valeurs numériques aux inconnues énoncées. On appelle calibrage d'une caméra cette procédure d'estimation des valeurs du modèle d'une caméra. De nombreuses méthodes existent pour différentes parties du modèle, certaines en chevauchant plusieurs. Les algorithmes de correction de la distorsion de lentilles sont couverts séparément en annexe (voir section II.2), car ceux-ci peuvent être effectués en prétraitement de façon complètement indépendante du reste du modèle. La section suivante couvre plutôt les méthodes complètes de calibrage.

On distingue principalement deux types de méthodes de calibrage, soit le calibrage implicite et le calibrage explicite. Dans le cas du calibrage implicite, on cherche à trouver implicitement les paramètres du modèle en estimant d'abord la matrice de transformation de l'espace 3D à l'espace 2D, puis en décomposant cette dernière afin de retrouver les paramètres du modèle. Pour le calibrage explicite, on recherche directement, explicitement, les valeurs des paramètres du modèle.

1.5.1 Calibrage par la méthode *DLT*

L'algorithme *DLT* [65] peut être effectué pour réaliser un calibrage implicite. Il s'agit de rechercher la matrice M qui permet de projeter un ensemble de points 3D dont on connaît la position vers un ensemble de points 2D pour lesquels on a identifié la position dans l'image. Les détails mathématiques de cette méthode sont reproduits en annexe (voir section III.3).

Ainsi, à partir d'un objet de calibrage dont les positions 3D de certains points, facilement identifiables par la suite sur une image, sont connues, on peut trouver la matrice de calibrage

M . Il est possible de décomposer la matrice M afin de retrouver les paramètres explicites du calibrage. En effet, on a :

$$\begin{aligned}
 M &= \begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix} \\
 &= \begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & t'_0 \\ 0 & 1 & 0 & t'_1 \\ 0 & 0 & 1 & t'_2 \end{bmatrix}
 \end{aligned} \tag{1.3}$$

avec :

$$\begin{bmatrix} t_0 \\ t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} t'_0 \\ t'_1 \\ t'_2 \end{bmatrix} \tag{1.4}$$

Ainsi, la partie gauche 3x3 de la matrice M correspond au produit d'une matrice triangulaire supérieure avec celui d'une matrice orthonormale. Il est donc possible d'utiliser la décomposition RQ pour obtenir, à partir de cette sous-matrice 3x3, la matrice triangulaire supérieure K correspondant aux paramètres intrinsèques de la caméra ainsi que la matrice orthonormale R correspondant à la rotation décrivant l'orientation de la caméra. Il ne reste que le paramètre $[t_0 \ t_1 \ t_2]^T$ à trouver :

$$(KR)^{-1} M = \begin{bmatrix} 1 & 0 & 0 & t'_0 \\ 0 & 1 & 0 & t'_1 \\ 0 & 0 & 1 & t'_2 \end{bmatrix} \tag{1.5}$$

On trouve donc $[t'_0 \ t'_1 \ t'_2]^T$ à partir de l'équation 1.5, puis $[t_0 \ t_1 \ t_2]^T$ à partir de l'équation 1.4. La matrice KR sera toujours inversible dans le cas de caméras finies, par opposition aux caméras situées à l'infini qui ne décrivent pas les caméras réelles [65].

Il est à noter que les points de l'objet de calibrage ne doivent pas être coplanaires pour que le calibrage par DLT fonctionne. Une telle configuration entraîne une dégénérescence du système d'équations produit, qui ne pourra être résolu convenablement.

1.5.2 Calibrage par la méthode de Tsai

Insatisfait des méthodes de calibrage disponibles, Tsai a développé sa méthode [66] afin de satisfaire des critères d'autonomie du calibrage (sans besoin d'intervention de l'utilisateur), de précision, d'efficacité dans le processus de calcul, de polyvalence pour s'adapter à divers contextes d'utilisation et finalement de n'avoir besoin que de caméras et de lentilles disponibles couramment plutôt que de nécessiter du matériel spécial utilisé dans le domaine de la photogrammétrie.

La méthode tente d'abord d'estimer le plus grand nombre possible de paramètres à l'aide d'équations linéaires de régression par moindres carrés, qui s'avèrent rapides et efficaces à résoudre. Puis, les autres paramètres sont obtenus par une méthode d'optimisation non linéaire qui tente de faire correspondre le mieux possible les points sur l'image à ceux de l'objet de calibrage projetés avec le modèle estimé [67]. Ainsi, le maximum de paramètres est trouvé de manière directe, puis on raffine l'estimation de ces derniers, tout en trouvant la valeur des autres paramètres, en minimisant l'erreur de reprojection des points connus sur les images.

L'algorithme de Levenberg-Marquardt [68] est très populaire pour l'optimisation non linéaire. Pratiquement tous les cas mentionnés dans ce mémoire procédant à une minimisation non linéaire l'utilisent.

Le modèle utilisé par cette méthode ne considère pas le facteur d'obliquité du modèle représenté par l'équation 1.1. De plus, la méthode fonctionne sur des objets de calibrage dont les points sont non coplanaires, mais propose également une variante pour les objets de calibrage planaires. Cette dernière ne permet toutefois pas d'estimer le paramètre nommé s_x , qui correspond au rapport f_y / f_x . De plus, si l'objet planaire est placé perpendiculairement à l'axe optique de la caméra, il ne sera pas possible de déterminer individuellement la distance focale f et la troisième composante de la translation, mais uniquement leur rapport. Or, cette configuration peut facilement être évitée et ne constitue pas une contrainte aussi restrictive que celle de devoir placer une grille de calibrage perpendiculaire à l'axe optique par exemple. Finalement, le modèle de distorsion proposé par la méthode originale n'est pas très complet, puisqu'il ne suppose qu'un seul coefficient variable. Ce modèle peut toutefois être étendu relativement facilement sans modification au reste de la méthode.

Cette méthode est très populaire pour les différents systèmes nécessitant le calibrage de caméra.

1.5.3 Calibrage par la méthode de Zhang

Une autre technique de calibrage très utilisée est la méthode de Zhang [69]. Cette dernière utilise au moins deux images prises d'une grille de calibrage dont les points sont coplanaires afin d'effectuer le calibrage.

La première partie de la méthode est très apparentée aux techniques d'auto-calibrage, car elle ne se sert que du contenu des images, c'est-à-dire les coordonnées des points appartenant à l'objet planaire. On trouve d'abord une homographie représentée par une matrice 3×3 reliant les coordonnées des points de l'objet de calibrage (dans un repère « quelconque » relié au plan dans lequel les points auront tous une coordonnée Z égale à 0) aux points de l'image. Cette homographie peut être obtenue par l'algorithme DLT présenté en annexe (voir la section III.3) à partir d'un nombre de primitives qui doivent être au moins quatre points, quatre lignes, trois points et une ligne, ou alors trois lignes et un point [9]. Zhang suggère de raffiner l'estimation de cette homographie par une optimisation non linéaire du maximum de vraisemblance.

La contrainte d'orthogonalité entre les vecteurs de la matrice de rotation du modèle de la caméra permet d'obtenir deux équations homogènes par image, c'est-à-dire deux par homographie. Ensuite, afin de retrouver la matrice des paramètres intrinsèques, on transforme cette dernière en une matrice symétrique 3×3 B , possédant 6 inconnues. Il est donc possible de créer un système d'équations pour trouver B , à partir de laquelle on trouve les paramètres intrinsèques directement. Ce système devient surdéterminé dès que l'on a plus de trois images. Avec seulement deux images, on peut imposer comme contrainte que le facteur d'obliquité est nul et donc obtenir une cinquième équation, ce qui est suffisant pour résoudre, à un facteur d'échelle près, le système homogène de 6 inconnues. Une fois les paramètres intrinsèques trouvés ainsi, les paramètres extrinsèques peuvent également être trouvés directement pour chaque image.

Les paramètres de la distorsion sont également déterminés par la méthode de Zhang en comparant les points (x, y) identifiés dans l'image avec les points (x_p, y_p) obtenus en projetant les points 3D à l'aide des paramètres intrinsèques et extrinsèques déterminés

précédemment. En ne supposant que de la distorsion radiale exprimée par le polynôme suivant et en normalisant les coordonnées en fonction des paramètres intrinsèques, on a :

$$\begin{aligned} x &= x_p + x_p \left(k_1 (x_p^2 + y_p^2) + k_2 (x_p^2 + y_p^2)^2 \right) \\ y &= y_p + y_p \left(k_1 (x_p^2 + y_p^2) + k_2 (x_p^2 + y_p^2)^2 \right) \end{aligned} \quad (1.6)$$

Il est donc possible d'obtenir, pour chaque association de points, deux équations linéaires dont les inconnues sont k_1 et k_2 . La résolution du système surdéterminé donne donc les valeurs des coefficients de distorsion.

Bien que cette extension ne fasse pas partie de la méthode originale de Zhang, il est possible d'ajouter au système d'équations linéaires précédent des inconnues correspondant aux paramètres de la distorsion tangentielle, tel que suggéré par Brown [70]. Le système reste linéaire et peut être résolu de la même façon. Cette approche est celle employée par OpenCV [71]) pour l'implémentation du calibrage de caméra avec estimation et correction de la distorsion.

La seconde partie de la méthode de Zhang consiste à utiliser l'estimation des paramètres intrinsèques, des paramètres extrinsèques et des paramètres de distorsion trouvés par la méthode directe comme point de départ pour une minimisation non linéaire de l'erreur de reprojection.

Cette méthode permet donc d'effectuer le calibrage à partir de plusieurs images, ce qui peut éventuellement réduire l'erreur d'estimation des paramètres en multipliant les prises de données que constituent les images. De plus, elle n'oblige pas de placer la grille de calibrage perpendiculairement à l'axe optique de la caméra. Par contre, il est à noter que si la rotation de la caméra est la même pour deux images, i.e. il n'y a eu qu'une translation de la caméra (ou du plan) entre ces images, la seconde homographie n'ajoutera pas d'information à la résolution du système d'équations et donc à la procédure de calibrage. Finalement, cette méthode est attrayante de par la possibilité d'utiliser un objet de calibrage simple, facilement produit par une simple imprimante par exemple.

1.5.4 Calibrage par la méthode de Girard

La méthode de calibrage développée par Girard [9] est en fait utilisée par ce dernier pour initialiser sa procédure d'auto-calibrage et a été développée dans le même cadre d'application que le présent travail.

Elle commence par estimer les paramètres de distorsion à l'aide de la méthode présentée en annexe à la section II.2.4. Puis, à partir de primitives suivies sur une série d'images obtenues en effectuant une rotation pure de la caméra autour de son axe optique, elle trouve les coordonnées du point principal comme étant le centre de rotation des primitives suivies.

Par la suite, à l'aide d'un objet de calibrage en « L » présenté à la Figure 1.24, il procède à un calibrage *DLT* pour obtenir les paramètres intrinsèques restant et extrinsèques. Finalement, une optimisation non linéaire raffine les paramètres extrinsèques et intrinsèques, à l'exception du point principal qui est fixé à la valeur obtenue par la procédure de rotation.

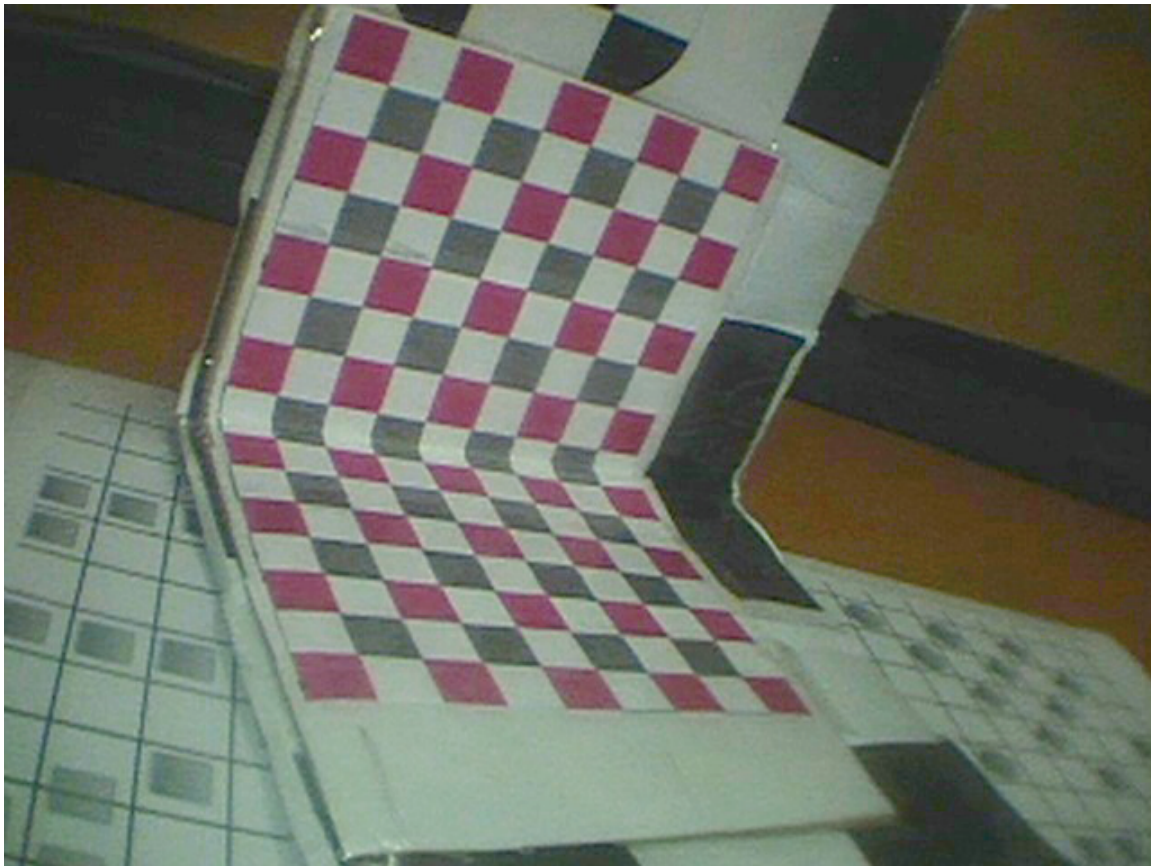


Figure 1.24. Objet de calibrage utilisé par Girard (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission)

La méthode de Girard possède toutefois certaines lacunes :

- Tel que mentionné en annexe à la section II.2.4, la procédure de calibrage de la distorsion pose des difficultés. La contrainte d'avoir à placer la grille perpendiculaire à l'axe optique est difficile à appliquer et la présence du centre de distorsion dans les paramètres de la minimisation peut entraîner un positionnement erroné de celui-ci. L'inclusion de la distorsion dans le calibrage total donne moins de contraintes sur le protocole de calibrage de la distorsion qui doit être effectué fréquemment, i.e. à chaque stérilisation de l'endoscope [72].
- La procédure de détermination de la position du centre optique par rotation de la caméra considère que l'axe de rotation est aligné avec l'axe optique de la caméra, ce qui n'est pas toujours le cas à cause de légères erreurs lors de l'assemblage. L'hypothèse que le centre optique correspond au centre de rotation de la caméra est une approximation difficile à vérifier dans la réalité.
- L'utilisation d'un objet de calibrage en « L » nécessite la numérisation précise de la position 3D des points de l'objet de calibrage, ce qui complique la production d'objets de calibrage. De plus, cette configuration complique l'identification des points dans les images de la caméra, qui doit être faite par l'utilisateur, ce dernier étant quand même aidé par une procédure semi-automatique.

1.6 Mise à jour des paramètres extrinsèques

Les méthodes de calibrage présentées précédemment permettent d'obtenir à la fois les paramètres intrinsèques et extrinsèques de la caméra pour la ou les vue(s) utilisée(s) pour effectuer le calibrage. Les paramètres intrinsèques de la caméra sont les mêmes d'une vue à l'autre⁴, alors que les paramètres extrinsèques vont varier selon la position de la caméra par rapport à l'objet de calibrage. De plus, aussitôt que la caméra est déplacée, les paramètres extrinsèques fournis par le calibrage ne sont plus valides, alors que les paramètres intrinsèques restent les mêmes. Les paramètres extrinsèques de la caméra, c'est-à-dire sa position et son orientation, devront être déterminés en temps réel afin de pouvoir mettre à jour le modèle. Pour ce faire, on a généralement recours à un système de suivi de poses qui

⁴ Cette affirmation est vraie tant que le chirurgien ne procède pas à des opérations particulières qui changent les paramètres intrinsèques, notamment le changement de facteur de zoom et la mise au point. L'usage courant consiste à déplacer l'endoscope, ce qui n'entraîne pas de variation des paramètres intrinsèques.

permet d'effectuer le suivi d'objets et d'obtenir de l'information concernant leur pose. On doit adapter les instruments à suivre afin d'y intégrer la technologie de suivi, ce qui peut être fait avec plus ou moins de difficultés selon la technologie de suivi utilisée.

1.6.1 Technologies de suivi

Plusieurs stratégies existent quant au moyen de suivre en temps réel la pose d'un objet dans l'espace :

- *Mécanique* : L'utilisation d'un système mécanique [18, 73] se fait à l'aide d'un bras articulé. Ce dernier permet d'obtenir les angles de chacun de ses joints, ce qui permet ensuite de calculer la pose d'un objet fixé à l'extrémité de ce bras robotisé. Toutefois, ces systèmes sont généralement encombrants et nécessitent une adaptation assez complexe afin de les intégrer aux instruments à suivre.
- *Acoustique* : L'ultrason⁵ peut être utilisé pour localiser un objet en 3D. À partir d'un émetteur de brefs signaux ultrasoniques et d'un ensemble d'au moins trois récepteurs calculant le temps de propagation des ondes acoustiques depuis l'émetteur, il est possible de trouver la position de l'émetteur [74]. La précision de tels systèmes, de l'ordre du millimètre, laisse à désirer et de par la nature de la modalité utilisée, il est nécessaire de combiner plusieurs émetteurs afin d'obtenir l'information d'orientation.
- *Optique* : Les ondes lumineuses peuvent également être utilisées pour faire le suivi d'objets à l'aide d'une ou plusieurs caméra(s) détectant des diodes électroluminescentes (marqueurs actifs), des sphères réflectorisantes (camera active) ou des patrons reconnaissables (système passif). Puis, ces objets sont habituellement mis en correspondance entre les différentes caméras pour éventuellement obtenir les coordonnées 3D par triangulation. Un tel système est relativement peu dispendieux et donne une excellente précision. Toutefois, les marqueurs utilisés doivent en tout temps être visibles par un minimum de caméras sans quoi le suivi ne peut être effectué.
- *Magnétique* : Les systèmes magnétiques utilisent un émetteur de champ magnétique ainsi que des récepteurs se servant du gradient du champ magnétique en question pour

⁵ Les ondes acoustiques audibles pourraient également être utilisées avec le même principe, mais pour des raisons évidentes de discrétion du système, on préfère l'ultrason, inaudible par l'oreille humaine.

déterminer leur position et orientation. Ces systèmes ont l'avantage de ne pas avoir de problème de ligne de visée, car les champs magnétiques passent au travers des objets. Par contre, le champ magnétique peut être déformé par des objets ferromagnétiques qui peuvent entraîner des erreurs RMS allant jusqu'à 8.4 mm pour la position et 166° pour l'orientation [75], ce qui restreint son potentiel d'application en salle d'opération.

Le Tableau 1.1 résume les différentes caractéristiques de chacune des modalités de suivi de la pose présentées précédemment.

Tableau 1.1. Comparaison des modalités de suivi de pose (adapté de [76])

	Mécanique	Optique	Magnétique	Acoustique
Précision	0,1 – 2,5 mm	0,1 – 0,5 mm	~5 mm	~1 mm
Résolution	N.D.	~0,01 mm au mieux	N.D.	~0,1 mm
Rafraîchissement	> 3000 Hz	100 – 2500 Hz	20 – 100 Hz	500 – 1000 Hz
Sources d'interférence	Occlusion physique	Chaleur, occlusion	Objets ferreux, champs magnétiques	Température, humidité, occlusion
Exemples	Faro Arm, NeuroNavigator	Optotrak 3020, Flashpoint	Polhemus, Flock of Birds	Sonic Wand
Contact / Sans contact	Contact direct	Contact avec cibles	Contact avec cibles	Contact avec cibles

Grâce à leur grande précision, leur coût relativement faible et la facilité avec laquelle il est possible de les intégrer, les systèmes optiques sont les plus utilisés. Les systèmes Polaris® et Optotrak® de Northern Digital Inc. sont très répandus. La Figure 1.25 montre la caméra utilisée par le système Polaris®, alors que la Figure 1.26 montre un exemple d'un instrument sur lequel a été posé un marqueur optique, composé de quatre sphères, permettant d'effectuer le suivi de la position et de l'orientation de l'instrument.



Figure 1.25. Système Polaris® (tirée de Northern Digital Inc., "Passive Polaris Spectra System," *Passive Polaris Spectra System*. [En ligne]. Disponible: <http://www.ndigital.com/medical/polarisfamily-spectrasystempassive.php>. [Consulté le 2010/06/24]. ©2008 Northern Digital Inc. Reproduit avec permission.)

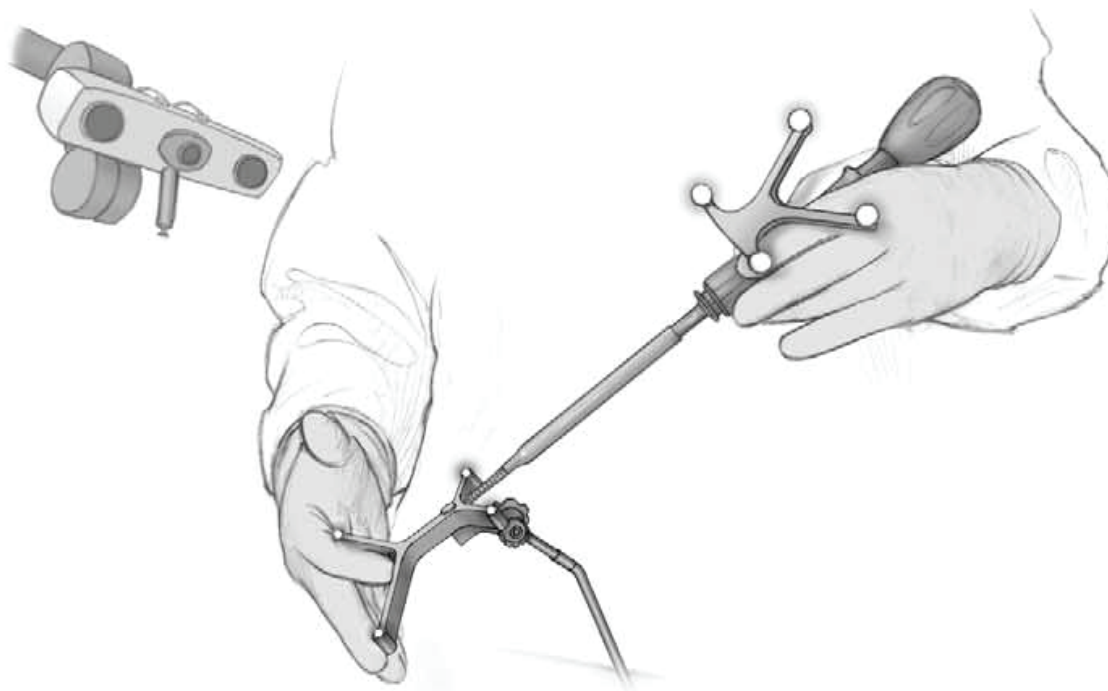


Figure 1.26. Instrument adapté avec un marqueur optique (adaptée de J. S. Hott, N. Theodore, J. D. Klopfenstein, S. Papadopolous, C. A. Dickman, et V. K. H. Sonntag, "Image guidance in spinal surgery for trauma: Iso-C 3D," *Operative Techniques in Neurosurgery*, vol. 7, no. 1, pp. 42-47, 2004. ©2004 Elsevier. Reproduit avec permission.)

Un autre système de suivi, moins populaire mais complètement passif, est le MicronTracker de Claron Technology Inc. Ce système, utilisant également une caméra binoculaire représentée sur la Figure 1.27, repère des marqueurs pouvant être produits très facilement, en les imprimant sur des surfaces quelconques. Sa flexibilité, sa facilité d'utilisation et son faible coût quant à la génération de marqueurs en font un choix très intéressant pour la recherche.

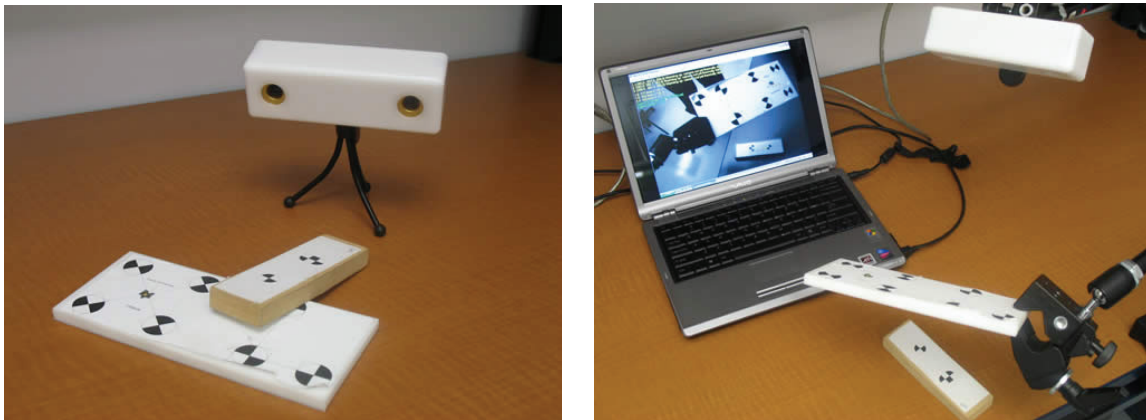


Figure 1.27. Système MicronTracker (tirée de LIV4D, "Nos installations," *Laboratoire d'Imagerie et de Vision 4D (LIV4D)*, 2010. [En ligne]. Disponible: <http://www.polymtl.ca/liv4d/installation/equipement.php>. [Consulté le 2010/03/16]. ©2010 LIV4D. Reproduit avec permission.)

1.6.1.1 Système de suivi MicronTracker

Le système de suivi MicronTracker est basé sur la reconnaissance de marqueurs qu'il est possible de concevoir à l'aide de matériel informatique couramment disponible, tel qu'une imprimante classique permettant d'imprimer sur du papier ou sur des autocollants par exemple. La terminologie de ce système de suivi est ici détaillée afin de comprendre la conception de marqueurs reconnaissables par le MicronTracker [77].

Terminologie

Tout d'abord, la Figure 1.28 représente l'élément de base des marqueurs MicronTracker, soit le *Xpoint* qui est une intersection de deux droites perpendiculaires ou non et qui forme un patron de damier, c'est-à-dire qu'en partant d'un segment formé des droites et partant du point d'intersection, une rotation autour de ce point se fait en parcourant des régions blanches et noires en alternance. Les marqueurs utilisent des droites, nommées *Xlines*, généralement perpendiculaires bien qu'elles n'aient pas à l'être. Chacune des droites est également

qualifiée en fonction de la manière dont elle sépare les couleurs noire et blanche : la ligne *WB* possède le blanc à gauche et le noir à droite (d'où les lettres *WB* pour *White-Black*), alors que la ligne *BW* possède le noir à gauche et le blanc à droite (*BW* pour *Black-White*). Chacun des *Xpoints* définit également une *Xnormal*, qui est un vecteur unitaire perpendiculaire aux deux *Xlines* et orienté de manière à sortir de la surface planaire vers l'observateur.

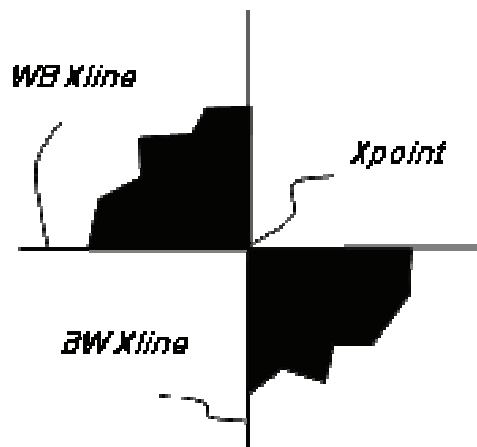


Figure 1.28. *Xpoint* du MicronTracker (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," *Manuel d'utilisation*, 2009. ©2009 Claron Tech Inc. Reproduit avec permission.)

Lorsque deux *Xpoints* sont alignés de manière à ce que la *BW Xline* de l'un corresponde à la même droite que la *WB Xline* de l'autre, les deux points définissent un vecteur orienté du point alignant la ligne *BW* vers le point alignant la ligne *WB*, tel que représenté par la Figure 1.29.

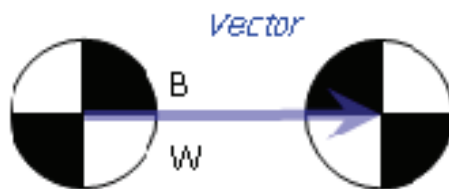


Figure 1.29. Vecteur du MicronTracker (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," *Manuel d'utilisation*, 2009. ©2009 Claron Tech Inc. Reproduit avec permission.)

Deux vecteurs, formés de trois ou quatre *Xpoints* (un *Xpoint* peut faire partie de deux vecteurs différents, car il possède deux *Xlines*), peuvent former une facette à condition qu'un vecteur

soit plus long que l'autre, pour que le système puisse les distinguer, et que l'angle entre les vecteurs soit situé entre 8 et 172 degrés, afin que les vecteurs ne paraissent pas colinéaires. Cette facette constitue l'élément le plus simple permettant de définir une pose, c'est-à-dire une position (représentée par une translation) et une orientation (représentée par une rotation). Les deux vecteurs permettent de définir un système de coordonnées propre à chaque facette, et la pose constitue la transformation rigide (translation et rotation) permettant de passer du système de coordonnées du monde du MicronTracker, défini par rapport à la caméra MicronTracker, au système de coordonnées de la facette, défini à l'aide des vecteurs.

Toujours selon la terminologie du MicronTracker, un marqueur est un ensemble d'une ou plusieurs facettes pouvant ainsi être suivi par le système. Cet ensemble de facettes peut être enregistré comme un patron de marqueur définissant une certaine géométrie. Ainsi, suite à la conception d'un nouveau marqueur, ce dernier doit être enregistré dans un fichier texte décrivant sa géométrie, afin de pouvoir être reconnu par le MicronTracker. Lors du suivi des marqueurs, le système identifie des vecteurs qu'il essaie de faire correspondre à des facettes parmi les patrons de marqueurs qu'il connaît. Même si certaines facettes d'un marqueur sont cachées, le système est en mesure de déterminer la pose de ce marqueur si au moins une facette est visible. Cette pose est quand même exprimée par rapport à la première facette du patron de ce marqueur, même si celle-ci n'est pas visible.

Systèmes de coordonnées

La pose est définie comme la transformation rigide (translation et rotation) permettant de passer du système de coordonnées du monde, c'est-à-dire de la caméra MicronTracker, au système de coordonnées du marqueur, ou vice-versa. De manière plus rigoureuse, comme il est possible de définir un système de coordonnées pour chaque facette, on définit le système de coordonnées d'un marqueur comme étant celui de la première facette du patron représentant le marqueur. Comme tous les marqueurs utilisés dans le cadre de ce travail ne possèdent qu'une seule facette et afin de simplifier l'écriture, les termes « système de coordonnées de la facette » et « système de coordonnées d'un marqueur » seront utilisés de manière interchangeable.

On désigne le repère MicronTracker par l'indice w et celui du marqueur par l'indice m . Le système de coordonnées du monde est celui dans lequel sont données les positions des marqueurs retournées par le MicronTracker et est illustré à la Figure 1.30.

- L'origine du système de coordonnées est située environ au milieu des caméras.
- L'axe Z est orienté approximativement dans la direction de l'axe optique des caméras et est dirigé vers l'extérieur des caméras.
- L'axe Y est dirigé vers le bas, perpendiculairement à l'axe Z et à la ligne imaginaire reliant les deux caméras.
- l'axe X est perpendiculaire à l'axe Z et à l'axe Y , c'est-à-dire parallèle à la ligne imaginaire reliant les deux caméras et orienté dans le sens cohérent avec les deux autres axes pour donner une base directe, c'est-à-dire respectant la règle de la main droite.

Il est à noter que ce système de coordonnées est très approximatif et ne correspond à aucune caractéristique de la configuration physique des caméras. Il dépend du résultat du calibrage de ces dernières, fait en usine.

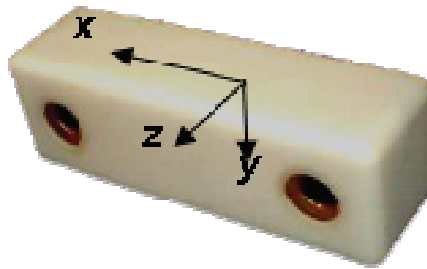


Figure 1.30. Système de coordonnées du MicronTracker (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," *Manuel d'utilisation*, 2009. ©2009 Claron Tech Inc. Reproduit avec permission.)

Le système de coordonnées d'un marqueur (ou plutôt d'une facette) est défini par rapport aux *Xpoints* le composant.

- L'origine du système est définie comme le point milieu du plus long vecteur de la facette.
- L'axe X est défini comme le vecteur unitaire parallèle au vecteur le plus long de la facette et étant orienté dans la même direction.

- L'axe Z est défini comme étant perpendiculaire aux deux vecteurs de la facette et étant orienté dans la direction opposée aux $Xnormal$ s de la facette (et donc par conséquent dans le même sens que l'axe Z du système de coordonnées du MicronTracker)
- L'axe Y est tout simplement défini comme étant perpendiculaire aux deux autres et orienté de manière à donner une base directe, c'est-à-dire à respecter la règle de la main droite. En pratique, l'axe Y correspond au vecteur le plus court auquel on a soustrait sa projection sur le vecteur le plus long (de manière à ce qu'il y soit perpendiculaire).

La Figure 1.31 démontre les systèmes de coordonnées définis par les facettes d'un unique marqueur.

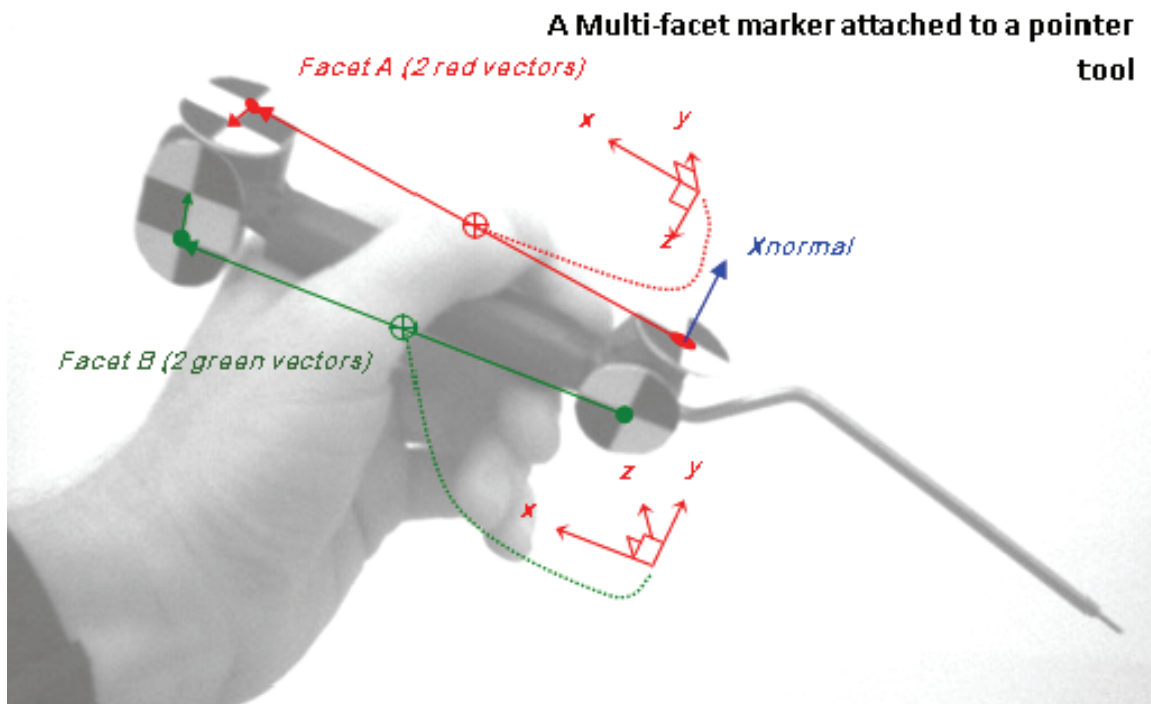


Figure 1.31. Système de coordonnées d'un marqueur (tirée de Claron Tech Inc., "MicronTracker Developer's Manual, MTC 3.0," *Manuel d'utilisation*, 2009. ©2009 Claron Tech Inc. Reproduit avec permission.)

Données fournies par le système

L'information concernant la pose du marqueur est donnée sous la forme d'une translation et d'une rotation. La translation T_i retournée par le système de suivi correspond à la position du marqueur, ou plus précisément à la position de l'origine du système de coordonnées. La

rotation R_i retournée, quand à elle, correspond à la rotation à appliquer au système de coordonnées du monde MicronTracker pour l'aligner avec celui du marqueur. Cette dernière est donnée sous deux formes : une matrice de rotation et un quaternion. La matrice de rotation est représentée en mémoire en format rangée, c'est-à-dire que si elle est contenue dans un tableau a , les éléments seront organisés de la sorte⁶ :

$$\begin{bmatrix} a[0] & a[1] & a[2] \\ a[3] & a[4] & a[5] \\ a[6] & a[7] & a[8] \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1.7)$$

La représentation est importante, car considérer que la matrice utilise la représentation colonne revient à prendre la transposée de la matrice. Or, dans le cas d'une rotation, la matrice la représentant est une matrice orthonormale, donc sa transposée est son inverse. Ainsi, le fait de ne pas interpréter la représentation en mémoire de la matrice de la bonne façon revient à utiliser la rotation inverse et change complètement le sens de l'interprétation des données.

La translation permet de passer du repère marqueur au repère monde, en ajoutant la valeur de T_i à n'importe quel point de l'espace marqueur pour le transformer dans l'espace monde. Elle est donc notée $T_{w \leftarrow m}$. La rotation, quant à elle, permet de faire passer un point de l'orientation monde à l'orientation du marqueur. Elle est donc notée $R_{m \leftarrow w}$. Ainsi, la transformation rigide $M_{w \leftarrow m}$ permettant de passer du système de coordonnées du marqueur à celui du monde est :

$$M_{w \leftarrow m} = T_{w \leftarrow m} R_{w \leftarrow m} = \begin{bmatrix} R_{m \leftarrow w}^{-1} & T_{w \leftarrow m} \end{bmatrix} = \begin{bmatrix} R_i^T & T_i \end{bmatrix} \quad (1.8)$$

De la même façon, la transformation rigide $M_{m \leftarrow w}$ permettant de passer du repère monde au repère marqueur est :

$$\begin{aligned} M_{m \leftarrow w} &= (M_{w \leftarrow m})^{-1} = (T_{w \leftarrow m} R_{w \leftarrow m})^{-1} = R_{w \leftarrow m}^{-1} T_{w \leftarrow m}^{-1} = R_{m \leftarrow w} (-T_{w \leftarrow m}) \\ &= \begin{bmatrix} R_{m \leftarrow w} & -R_{m \leftarrow w} T_{w \leftarrow m} \end{bmatrix} = \begin{bmatrix} R_i & -R_i T_i \end{bmatrix} \end{aligned} \quad (1.9)$$

Le MicronTracker permet donc d'obtenir, pour ces marqueurs faciles à concevoir, leur pose en temps réel, à une fréquence de 15 Hz.

⁶ Ce détail d'implémentation a été déterminé expérimentalement, mais comme il n'apparaît nulle part dans la documentation, on le mentionne ici pour le lecteur intéressé à l'implémentation.

1.7 Extensions spécifiques à la caméra endoscopique

Les méthodes présentées jusqu'à maintenant s'appliquent à la majorité des caméras numériques. Il est en effet possible de mettre à jour en temps réel le modèle de la caméra à l'aide d'un marqueur posé sur cette dernière. Or, tel que mentionné à la section 1.3, l'endoscope possède un degré de liberté que les caméras classiques ne possèdent pas : la possibilité d'effectuer une rotation de la tête de la caméra par rapport au tube de l'endoscope. Cette rotation doit donc être incluse dans le modèle de formation de l'image et suivie en temps réel.

1.7.1 Méthode de Yamaguchi

Yamaguchi [78] a étendu la méthode de Tsai pour développer à la fois un modèle spécifique à l'endoscope oblique dont le tube peut tourner par rapport à la tête de caméra et également une procédure de calibrage pour les nouveaux éléments de ce modèle.

Il fixe un marqueur optique sur la tête de la caméra et ajoute un encodeur de rotation entre la caméra et le tube, ce qui permet de mesurer en temps réel la rotation relative entre les deux.

Ainsi, il ajoute cette rotation aux paramètres extrinsèques de la caméra. Il décompose en fait la rotation en deux rotations, représentées à la Figure 1.32, s'appliquant au système de coordonnées S_c de la caméra. Tout d'abord, une rotation d'un angle θ de l'ensemble du système autour d'un axe l_s , sans rotation relative entre la tête et le tube est considérée. Dans le cas d'un assemblage parfait, l'axe l_s passerait exactement par l'origine de S_c et posséderait un angle avec l'axe Z de S_c équivalent à l'angulation de l'endoscope. Ensuite, cette rotation est suivie d'une rotation d'un angle $-\theta$ de la tête de caméra pour la ramener à son emplacement original. L'axe $l_h(\theta)$ de cette rotation est en fait l'axe Z de S_c et dépend donc de θ , puisque l'axe Z de S_c varie effectivement en fonction de θ .

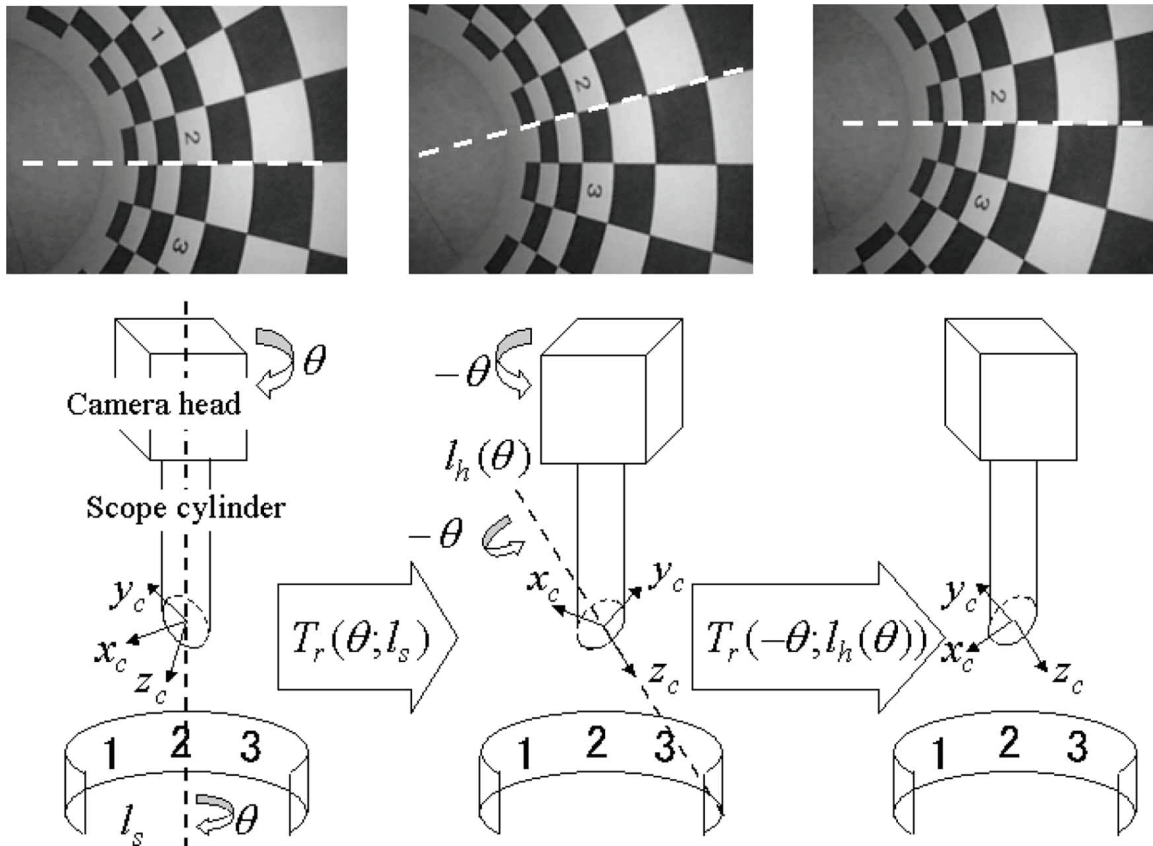


Figure 1.32. Représentation du modèle de rotation du tube de l'endoscope développé par Yamaguchi (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.)

Afin de simplifier le développement mathématique des équations en évitant de faire dépendre l_h de θ , il est préférable de considérer la rotation de la tête en premier, l'axe l_h étant ainsi toujours le même : celui associé à $\theta = 0$. Il reste donc à estimer les axes (donc un point et une orientation) l_s et l_h .

l_s est estimé en attachant un marqueur optique sur le tube de l'endoscope, comme représenté à la Figure 1.33. Lorsque l'on effectue une rotation du tube, la position du marqueur décrit un cercle par rapport à la tête de la caméra. La normale au plan de ce cercle représente l'axe de rotation et le centre un point de cet axe. L'axe l_h est déterminé en identifiant dans un certain nombre d'images, prises avec un angle de rotation différent, un certain nombre de points dont on connaît les coordonnées 3D. On trouve ensuite les paramètres manquants du modèle, soit

le point et la direction de l'axe l_h , en effectuant une optimisation non linéaire de l'erreur de reprojection obtenue en appliquant tout le modèle aux points 3D.

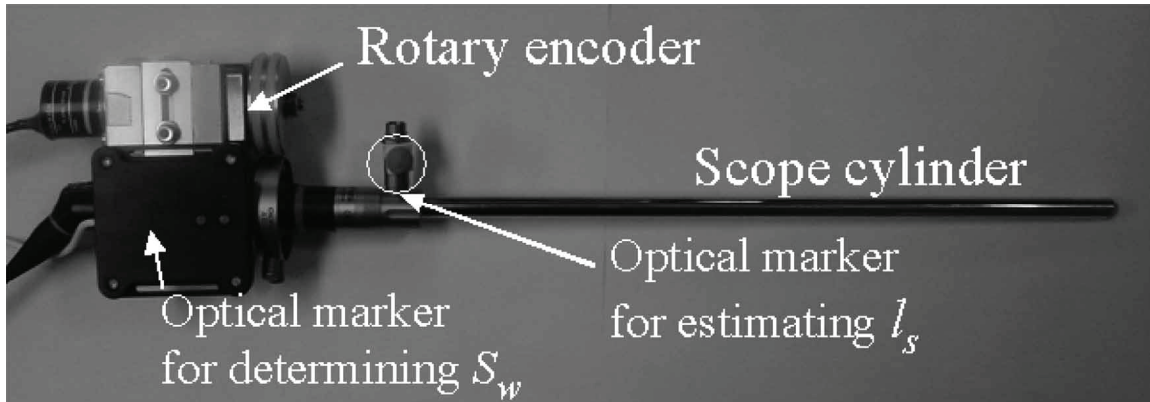


Figure 1.33. Placement des marqueurs pour l'estimation de l_s (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.)

Une comparaison a été faite entre l'utilisation de la valeur idéale, soit celle théorique en présence d'un assemblage parfait, de chacun des paramètres du modèle, soit le point et la direction de chacun des axes l_s et l_h , et l'utilisation de la valeur obtenue par le calibrage de la méthode proposée. Les résultats obtenus démontrent la pertinence d'effectuer le calibrage pour chacun des paramètres, l'erreur augmentant beaucoup en fonction de l'angle de rotation en l'absence d'un ou l'autre des calibrages, mais restant relativement constante pour le calibrage complet.

1.7.2 Méthode de Wu

La méthode de Wu [79] vise à simplifier le modèle de Yamaguchi en réduisant le nombre de paramètres à calibrer. En effet, plutôt que de mettre le marqueur optique sur la tête de la caméra, on le place sur le tube de l'endoscope, ce qui permet d'éviter d'avoir à calculer l'axe l_s . Ensuite, on considère que la rotation se fait dans le plan image, autour du centre optique. Ainsi, le seul paramètre à estimer est l'angle de rotation, plutôt que l'angle et deux axes de rotations (direction et point par lequel passe l'axe).

Pour estimer l'angle, on peut utiliser un encodeur de rotation, comme Yamaguchi, ou alors un second marqueur optique placé sur la tête de la caméra, tel que proposé par Wu. Il suffit

ensuite de choisir un point dans le référentiel du marqueur de la caméra (l'origine, par exemple) et de l'exprimer dans le référentiel du marqueur du tube et ce, pour différentes valeurs de l'angle de rotation entre la tête de caméra et le tube. Ce point décrivant un cercle dans le référentiel du tube, on détermine le centre du cercle par construction géométrique de l'intersection des bissectrices des côtés d'un triangle formé par trois points, combiné à un algorithme RANSAC afin d'utiliser tous les points (plus de trois) et de trouver le centre qui minimise la variance de la distance par rapport à chacun des points. L'angle est également déterminé par une construction géométrique soumise à l'algorithme RANSAC.

Il est important de jeter un regard critique à ce travail. Tout d'abord, l'innovation de poser le marqueur sur le tube de l'endoscope plutôt que sur la tête de caméra contribue véritablement à simplifier le modèle de Yamaguchi en éliminant le besoin de déterminer l'axe de rotation l_s . Par contre, le tube étant l'objet subissant la rotation, le marqueur doit également subir les mêmes rotations, ce qui peut gêner le travail du chirurgien et entraîner des problèmes de ligne de visée obstruée pour le système de suivi. Ces problèmes sont moins présents lorsque le marqueur est posé sur la tête de caméra, que le chirurgien maintient presque toujours dans la même orientation.

De plus, le reste de la méthode repose sur l'hypothèse que la rotation de la caméra entraîne une rotation dans le plan de l'image seulement. Pour utiliser la notation de Yamaguchi, cela revient à poser que l'axe de rotation l_h est l'axe Z du système de coordonnées de la caméra, passant par le centre optique. Certaines affirmations de Wu, notamment « *Misalignment will be included in the estimation of intrinsic and extrinsic parameters. Thus we do not need to estimate other parameters as Yamaguchi [sic] et al.* » et « *Since the image plane is in the camera head, the rotation only affects the image plane* » ne sont valides que si cette hypothèse est vérifiée.

Ainsi, l'utilisation de la méthode de Wu revient à ne pas calibrer l'axe l_h de la méthode de Yamaguchi et le poser égal au cas idéal où l'alignement du système de lentilles et de la caméra est parfait. Or, les résultats de Yamaguchi montrent clairement que l'erreur est beaucoup plus grande lorsque l'on fait cette hypothèse plutôt que de calibrer correctement l_h . Notamment, il montre que l'erreur est pratiquement constante lorsque le calibrage est fait, alors qu'elle augmente presque linéairement en fonction de l'angle lorsque le calibrage n'est pas fait, tel que représenté à la Figure 1.34. Visuellement, on peut remarquer ce phénomène à

la Figure 1.35. Ainsi, l'hypothèse sur laquelle se base une bonne partie de la méthode de Wu semble invalidée par les résultats de Yamaguchi.

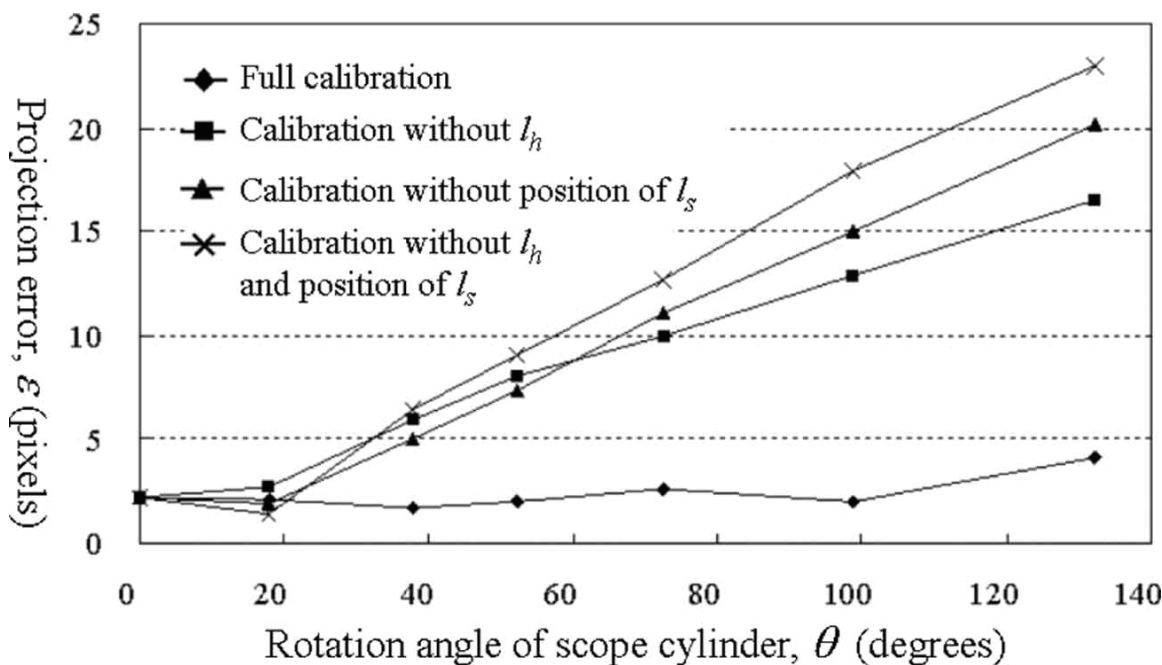


Figure 1.34. Comparaison du calibrage total de Yamaguchi à un calibrage partiel où seulement certains des paramètres sont calibrés alors que d'autres sont assumés respectant le cas idéal (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.)

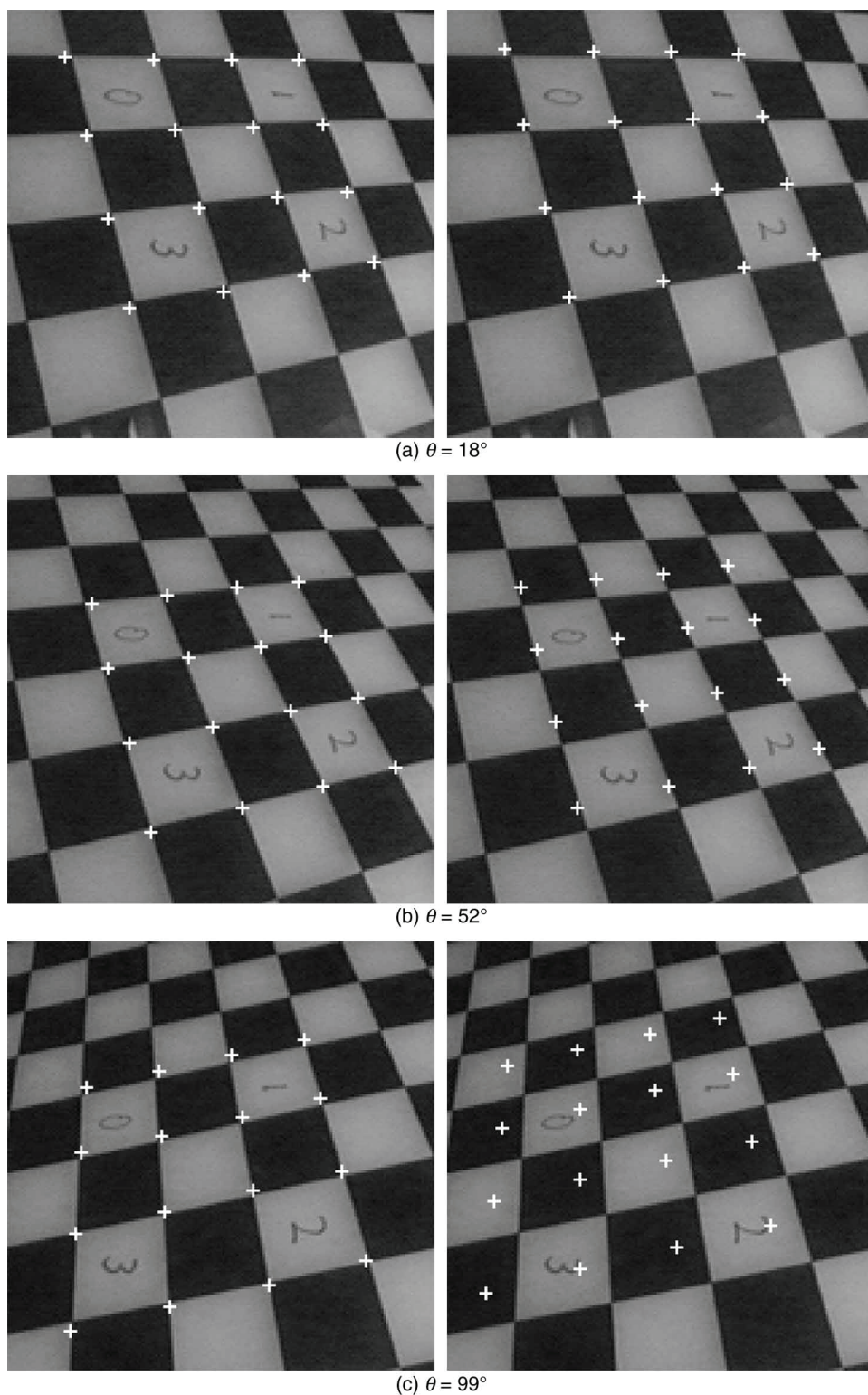


Figure 1.35. Erreur de reprojection en fonction de l'angle selon que le calibrage de l_h soit fait (gauche) ou non (droite) (tirée de T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005. ©2005 John Wiley & Sons. Reproduit avec permission.)

Aussi, la rotation dans le plan image n'est pas exactement équivalente à la rotation ajoutée aux paramètres extrinsèques. Pour que ces rotations soient équivalentes, le facteur de rapport d'aspect de la caméra doit être égal à 1. Autrement, les deux distances focales doivent être égales, ce qui n'est généralement pas vérifié. Dans le cas où l'échantillonnage des lignes entraîne la présence d'un facteur différent de 1, la rotation dans le plan image se fera dans un système de coordonnées inadéquat.

De plus, la validation effectuée par Wu n'est pas de très grande valeur. En effet, tel que montré sur la Figure 1.36, la comparaison des résultats de sa méthode est faite à partir de l'erreur de reprojection comparée à l'erreur obtenue en ne prenant pas en compte du tout de la rotation, ce qui entraîne évidemment une erreur plus grande. Par comparaison, Yamaguchi étudie plutôt l'effet de l'utilisation ou non de certaines parties de son modèle afin de justifier la pertinence de chacune d'elles. De plus, on peut voir sur cette même figure que l'erreur de reprojection de la méthode de Wu augmente en fonction de l'angle de rotation, tel que l'on aurait pu le prédire en observant les résultats de Yamaguchi. Cette augmentation continue de l'erreur en fonction de l'angle laisse supposer un problème dans la modélisation de la rotation qui est de plus en plus visible lorsque l'angle de rotation augmente.

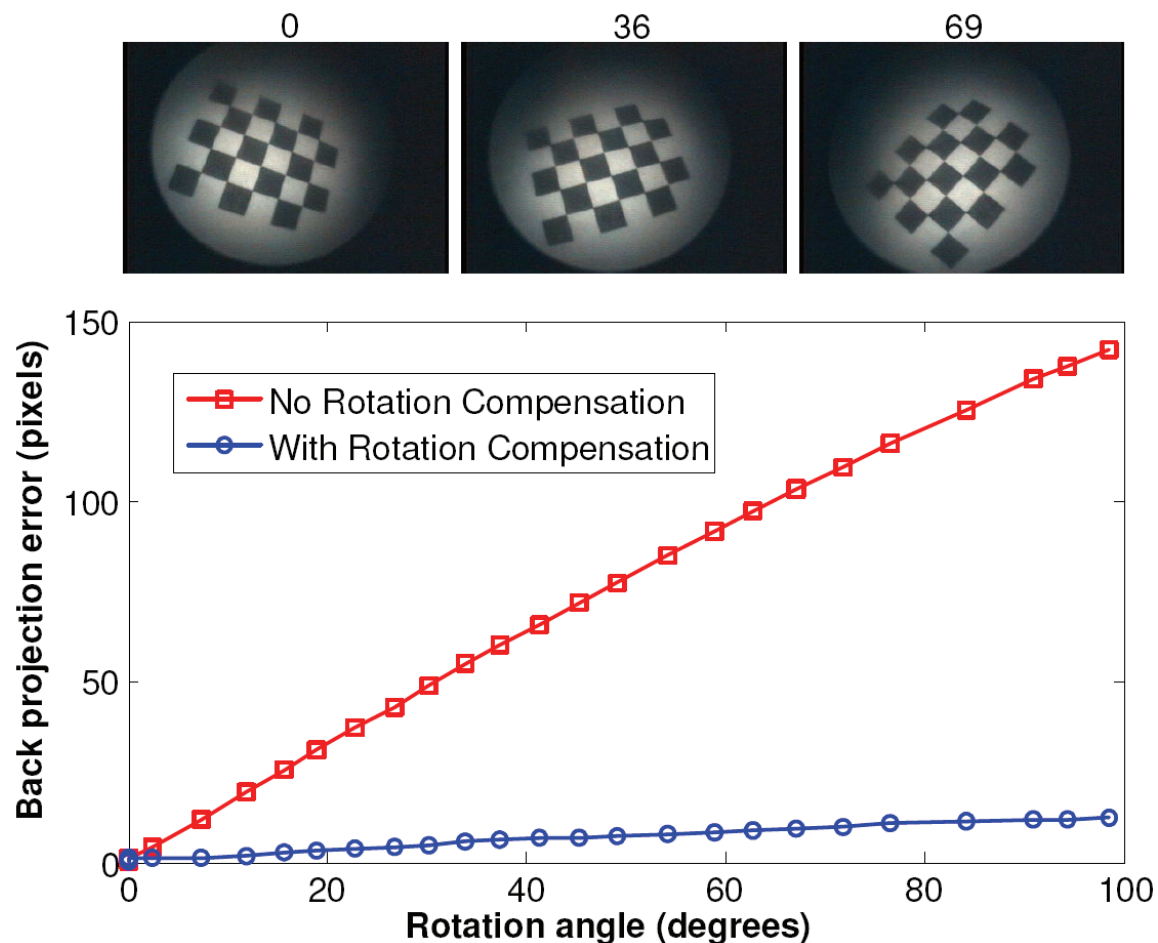


Figure 1.36. Erreur de reprojection de la méthode de Wu comparée à aucune considération de la rotation (adaptée de C. Wu et B. Jaramaz, "An easy calibration for oblique-viewing endoscopes," in *IEEE International Conference on Robotics and Automation*, 2008, pp. 1424-1429. ©2008 IEEE. Reproduit avec permission.)

Ainsi, la méthode de Wu apporte une contribution intéressante par le placement du marqueur sur le tube plutôt que sur la tête de caméra, ainsi qu'en introduisant l'utilisation d'un second marqueur optique plutôt que celle d'un encodeur de rotation, mais l'hypothèse à la base de l'élimination de l'axe l_h de Yamaguchi n'est pas justifiée et la simplification qui en résulte n'est en fait que l'assignation d'un cas idéal à un paramètre à calibrer.

1.7.3 Méthode de De Buck

Contrairement à Yamaguchi et à Wu, De Buck [80] ne considère pas que la rotation de la tête de la caméra résulte en une simple rotation à effectuer sur les paramètres extrinsèques ([78])

ou sur les coordonnées du point dans le plan image ([79]). En effet, il considère également une modification potentielle apportée aux paramètres intrinsèques de la caméra.

Ainsi, au lieu de n'ajouter qu'une matrice de rotation au modèle classique de la caméra à sténopé, on ajoute plutôt une homographie, c'est-à-dire une matrice 3x3. Ainsi, l'équation 1.1 deviendrait plutôt :

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = KH(\theta) E_{c \leftarrow w} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1.10)$$

$$= \begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{00}(\theta) & h_{01}(\theta) & h_{02}(\theta) \\ h_{10}(\theta) & h_{11}(\theta) & h_{12}(\theta) \\ h_{20}(\theta) & h_{21}(\theta) & h_{22}(\theta) \end{bmatrix} \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Malgré les 9 inconnues, l'homographie ne possède que 8 degrés de liberté, car elle est définie à un facteur d'échelle près. Ainsi, dans la procédure de calibrage, on détermine l'homographie pour différentes valeurs de θ . Ensuite, l'homographie correspondant à une nouvelle valeur de θ est trouvée en interpolant entre les homographies trouvées lors du calibrage.

Il est à noter que cette méthode modélise également l'effet de la rotation sur la position du centre de la caméra par l'ajout d'un terme de translation, dépendant de θ , au point 3D. Toutefois, les résultats ne justifient pas l'intégration d'un tel terme au modèle.

L'intérêt de cette méthode est donc de permettre, par l'utilisation d'une transformation projective (homographie), d'inclure des facteurs qu'une simple transformation rigide pourrait négliger. L'erreur obtenue est comparable à celle de Yamaguchi, même si le domaine des angles utilisés est toutefois plus restreint, probablement à cause du choix d'utiliser des marqueurs optiques plutôt qu'un encodeur de rotation, ce qui limite les angles accessibles sans perte du suivi des marqueurs. La modélisation de la rotation par une homographie donne donc de bons résultats, mais l'interpolation entre les homographies va induire des erreurs d'approximation si les angles pour lesquels une homographie est calculée sont trop distancés les uns des autres. Dans ce cas, les matrices interpolées risquent de ne pas représenter

adéquatement la transformation causée par la rotation relative de la tête de la caméra par rapport au tube.

1.8 Limites des systèmes existants

Les systèmes existants possèdent certaines limites qui les rendent difficilement applicables au contexte de la chirurgie minimalement invasive du rachis.

En effet, certains systèmes d'assistance chirurgicale déjà développés et utilisés pour d'autres régions du corps, notamment le cerveau, ne peuvent être appliqués directement au rachis dont la structure semi-rigide complique la modélisation. D'autres systèmes sont utilisés pour les chirurgies conventionnelles où le chirurgien a un plein accès au site opératoire. Dans un contexte minimalement invasif par contre, il n'est pas aussi facile de fixer les marqueurs sur les structures d'intérêt, opération qui est en soi invasive et qui va un peu à l'encontre des principes de ce type de chirurgie.

Par ailleurs, d'autres systèmes sont basés sur des modalités ionisantes comme la tomodensitométrie ou la fluoroscopie, ce qui entraîne une exposition aux radiations à éviter à la fois pour le patient et pour l'équipe chirurgicale. De plus, ce genre de méthodes exige du matériel coûteux qui n'est pas couramment disponible. Ce problème est également le cas des endoscopes stéréoscopiques, qui sont toujours difficilement applicables dans un contexte clinique. Une autre méthode basée sur le contenu des images endoscopiques consiste à appliquer des algorithmes de structure à partir de l'ombrage pour retrouver la forme 3D des surfaces observées par un endoscope classique; cette méthode repose toutefois sur des hypothèses qui ne sont pas vérifiées dans le contexte des chirurgies minimalement invasives.

Ces différentes limites doivent être considérées dans l'établissement des objectifs de ce travail de recherche.

1.9 Objectifs de recherche

L'objectif de ce projet est de développer un outil de visualisation 3D pour l'assistance de chirurgies minimalement invasives de la colonne vertébrale. Afin d'atteindre cet objectif général, on doit :

- développer une méthode de calibrage de l'endoscope, à exécuter avant la chirurgie, afin d'effectuer la fusion entre les modèles 3D et les images endoscopiques, qui se doit d'être précise tout en ne nécessitant qu'un minimum d'intervention de l'utilisateur
- proposer deux alternatives de visualisation de ces informations combinées facilitant la réalisation de chirurgies minimalement invasives du rachis
- valider quantitativement la précision de la technique de calibrage
- valider qualitativement la pertinence et l'utilité des deux méthodes de visualisation

CHAPITRE 2 MÉTHODOLOGIE

Le présent chapitre décrit les méthodes qui ont été implantées dans le cadre de ce travail de recherche. Le matériel employé est d'abord présenté. Puis, le modèle de formation de l'image endoscopique utilisé dans ce travail de recherche est détaillé. La méthode de calibrage de ce modèle est ensuite décrite, suivie des différentes vues augmentées qui sont proposées. Finalement, les méthodes de validation des différentes parties de ce travail de recherche sont exposées.

2.1 Matériel utilisé

La présente section décrit le matériel utilisé dans le cadre de ce travail de recherche. Plus précisément, elle fait part du système endoscopique ainsi que du système de suivi utilisés. Bien que cette méthode soit aisément adaptable à d'autres systèmes endoscopiques ou d'autres systèmes de suivi, il est pertinent de mentionner ceux utilisés dans le cadre spécifique de ce travail de recherche.

2.1.1 Système endoscopique

Le système utilisé est le « *Stryker Endoscopy 988 3 CHIP Digital Camera* ». Il est composé d'une caméra endoscopique couleur dont la résolution est 752 pixels par 582 ainsi que d'une boîte de capture possédant différentes sorties, notamment une sortie VGA 1024 par 768 et une sortie FireWire 720 par 480. La caméra endoscopique est jumelée à un tube endoscopique oblique dont l'angle est 45 degrés et la distance focale est de 10 mm selon le manufacturier. On peut également joindre à ce tube le câble provenant d'un système d'illumination, soit le Stryker X6000 Light Source, qui permet de fournir une source lumineuse lorsque l'endoscope est situé à l'intérieur du corps.

2.1.2 Système de suivi

Le système de suivi utilisé est le système MicronTracker de Claron Technology Inc. Ce système complètement passif permet de facilement créer de nouveaux marqueurs adaptés à des besoins spécifiques, ce qui le rend un choix intéressant pour l'exploration nécessaire dans le domaine de la recherche. Le système est décrit plus en détails à la section 1.6.1.1.

2.2 Modèle de formation des images endoscopiques

Le modèle de formation de l'image endoscopique consiste en la représentation des transformations appliquées à un objet imagé par la caméra endoscopique afin que celui-ci se retrouve dans l'image produite par la caméra. Afin de présenter ce modèle, on peut tout simplement décrire les différentes transformations appliquées à un point 3D appartenant à l'objet imagé, dont les coordonnées sont en unités du système de coordonnées du monde c'est-à-dire en millimètres dans le présent système, afin d'obtenir la position 2D de ce point sur l'image produite, dont les coordonnées sont en pixels. Cette série de transformations pouvant être appliquée à chacun des points de l'objet, elle décrit donc la manière dont une image endoscopique est formée à partir d'un objet quelconque et représente ainsi le modèle de formation de l'image.

Certains paramètres de ce modèle mathématique sont déterminés par une étape de calibrage effectuée avant la chirurgie (calibrage hors-ligne), alors que d'autres paramètres de ce modèle doivent être mis à jour au cours de la chirurgie (mise à jour en ligne). Comme tous ces paramètres sont nécessaires à la modélisation de la formation de l'image endoscopique, ils sont tous présentés dans les prochaines sous-sections décrivant le modèle de formation de l'image endoscopique.

La détermination de la valeur de ces paramètres est toutefois différente selon qu'ils soient obtenus par calibrage hors-ligne ou par mise à jour en ligne. Cette détermination des valeurs est donc abordée de manière distincte dans la section 2.3.

Les prochaines sous-sections présentent donc le modèle de formation de l'image endoscopique en entier.

2.2.1 Transformation 1 : paramètres extrinsèques

Soit P_0 un point 3D de l'objet imagé exprimé dans les coordonnées du monde. La première transformation du modèle de formation de l'image consiste à transformer ce point de manière à ce qu'il soit exprimé dans un système de coordonnées correspondant à la caméra, c'est-à-dire au bout de la lentille du tube de l'endoscope. Ce système possède l'axe X orienté vers la droite, l'axe Y orienté vers le haut et l'axe Z orienté vers l'intérieur de la caméra. Cette transformation repose donc sur les paramètres extrinsèques de la caméra endoscopique, soit sa position et son orientation, afin de pouvoir convertir le point en question des coordonnées

monde aux coordonnées caméra. Elle permet donc de tenir compte des paramètres extrinsèques dans le processus de formation de l'image.

La mise à jour des paramètres extrinsèques de la caméra correspond à retrouver sa position et son orientation en tout temps, ce qui peut être fait à l'aide du système de suivi. Toutefois, ce dernier donne une position et une orientation qui ne sont pas directement celles de la caméra.

En effet, le système de suivi fournit la position et l'orientation d'un marqueur situé sur le tube. Il permet donc d'obtenir une transformation vers le système de coordonnées du marqueur représenté à la gauche de la Figure 2.1. Comme le marqueur est fixé sur le tube, toute transformation rigide appliquée au tube, c'est-à-dire les déplacements effectués par le chirurgien, est appliquée à la fois au système du marqueur et à celui de la caméra. Ainsi, la transformation reliant le système du marqueur à celui de la caméra reste constante tant que le marqueur reste fixé au tube. Il est donc possible, à partir de cette transformation constante, de retrouver le système de coordonnées de la caméra.

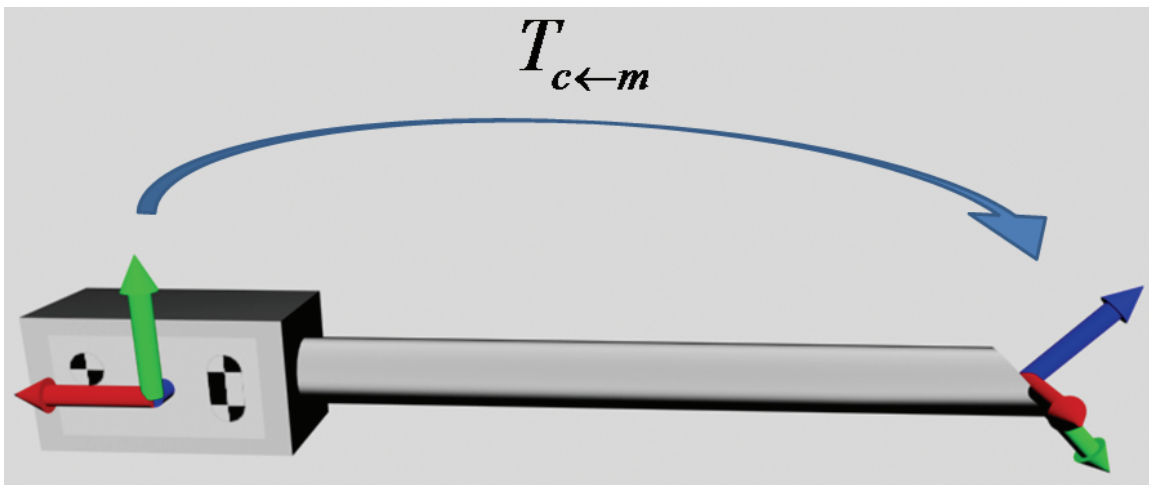


Figure 2.1. Transformation du repère du marqueur vers celui de la caméra. À noter que l'axe Z est représenté comme sortant de la caméra pour plus de clarté, alors qu'en réalité il entre dans cette dernière.

Ainsi, en obtenant par calibrage la transformation $T_{c \leftarrow m}$ qui relie le système de coordonnées du marqueur placé sur le tube à celui de la caméra (situé au bout du tube de l'endoscope), il devient facile de calculer les paramètres extrinsèques de la caméra à partir de l'information fournie par le système de suivi. Si $M_{m \leftarrow w}$ est la matrice décrivant la pose du marqueur telle

que définie par l'équation 1.9, la matrice $E_{c \leftarrow w}$ des paramètres extrinsèques peut être obtenue simplement :

$$E_{c \leftarrow w} = T_{c \leftarrow m} M_{m \leftarrow w} \quad (2.1)$$

Si P_0 est le point considéré exprimé en coordonnées monde et P_1 le même point exprimé en coordonnées caméra selon la position de cette dernière, cette transformation peut être résumée par l'équation suivante :

$$P_1 = T_{c \leftarrow m} M_{m \leftarrow w} P_0 \quad (2.2)$$

Ainsi, l'application de la transformation associée aux paramètres extrinsèques nécessite le calibrage de la transformation rigide $T_{c \leftarrow m}$ entre le marqueur et la caméra ainsi que le suivi du marqueur placé sur le tube de la caméra.

2.2.2 Transformation 2 : rotation relative de la tête de la caméra par rapport au tube

Comme la tête de la caméra peut tourner par rapport au tube, une transformation correspondant à une rotation doit être appliquée ensuite au point exprimé dans les coordonnées de la caméra. L'application de cette rotation correspond à faire tourner le point en coordonnées caméra autour d'un axe de rotation. Cet axe est en fait une droite représentée par un point et une direction, également exprimés en coordonnées caméra. Dans des conditions parfaites, l'axe de rotation correspondrait à l'axe Z du système de coordonnées de la caméra; toutefois, dans les assemblages de lentilles réels, l'axe de rotation doit être calibré. Cet axe, représenté par le point p_r et le vecteur directeur d_r , doit donc être calibré pour appliquer la transformation décrivant la rotation de la tête de la caméra par rapport au tube de l'endoscope.

Bien que cet axe de rotation soit obtenu grâce au calibrage hors-ligne, l'angle de la rotation à effectuer constitue un paramètre qui doit être mis à jour au cours de l'utilisation de l'endoscope. Cet angle ne peut être obtenu directement à partir du système de suivi. En effet, un paramètre supplémentaire doit être calibré hors-ligne afin de permettre la mise à jour en ligne de l'angle de rotation, tel que décrit dans la sous-section suivante.

2.2.2.1 Détermination de l'angle de rotation

Le calibrage de l'axe ayant été fait lors du processus initial de calibrage, l'angle de rotation, quant à lui, doit être déterminé en temps réel. On peut déterminer cet angle à l'aide d'un

second marqueur placé sur la tête de la caméra, tel que suggéré par Wu [79] et comme on peut le voir sur la Figure 2.2.

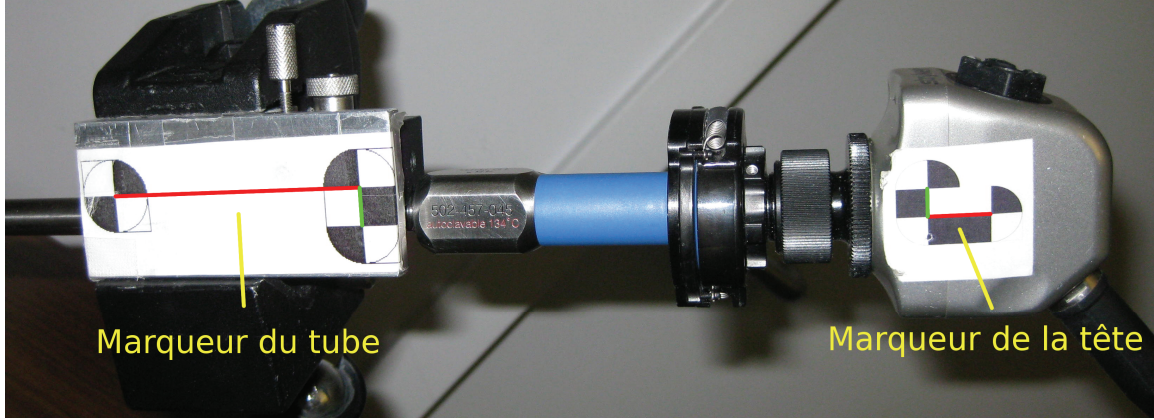


Figure 2.2. Position des marqueurs sur l'endoscope

Il est à noter que le modèle mathématique retenu pour la transformation associée à la rotation de la tête de la caméra par rapport au tube est une variante de celui de Yamaguchi [78]. Une variante de la méthode de Wu [79] est utilisée pour déterminer l'angle de rotation de la tête par rapport au tube. De plus, le placement d'un marqueur sur le tube proposé par Wu plutôt que d'utiliser un encodeur de rotation comme Yamaguchi permet d'éliminer un axe à calibrer dans la méthode de Yamaguchi sans perte de généralité ou de degré de liberté.

Lorsque la tête est tournée par rapport au tube, tous les points de la tête effectuent une rotation relative au repère du marqueur du tube. On peut donc prendre un point P quelconque exprimé dans le repère du marqueur de la tête, par exemple l'origine, et exprimer ce point dans le repère du marqueur du tube. Soit $M_{m1 \leftarrow w}$ la matrice représentant la transformation (obtenue par le système de suivi) du repère du monde vers celui du marqueur du tube, $M_{w \leftarrow m2}$ la matrice représentant la transformation (obtenue par le système de suivi) du repère du marqueur de la tête vers celui du monde, P_{m2} le point choisi dans le repère du marqueur de la tête et P_{m1} le même point exprimé dans le repère du marqueur du tube. On peut exprimer P_{m1} en fonction du point choisi et de l'état du système de suivi :

$$P_{m1} = M_{m1 \leftarrow w} M_{w \leftarrow m2} P_{m2} \quad (2.3)$$

Lorsque la tête est tournée, le point P_{m2} , fixe dans le repère de la tête, décrit un cercle dans le repère du tube, tel que représenté à la Figure 2.3.

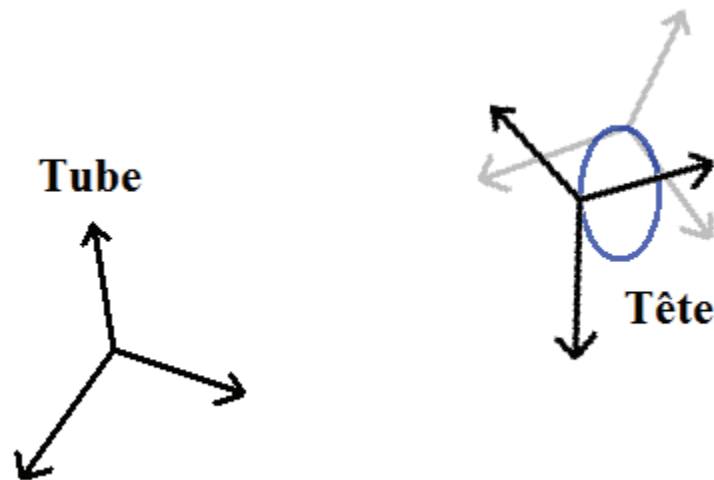


Figure 2.3. Cercle décrit par un point du marqueur de la tête dans le repère du tube

Il est possible de définir un repère par rapport à ce cercle :

- L'origine du repère est le centre du cercle.
- L'axe Z est la normale au cercle, orienté de manière à ce que le signe des angles soit cohérent.
- L'axe X est déterminé comme étant le vecteur unitaire, situé dans le plan du cercle, correspondant au vecteur reliant le centre du cercle à un point P_{ml} associé à un angle de rotation nul.
- L'axe Y est déterminé à partir de l'axe X et de l'axe Z de manière à obtenir une base directe.

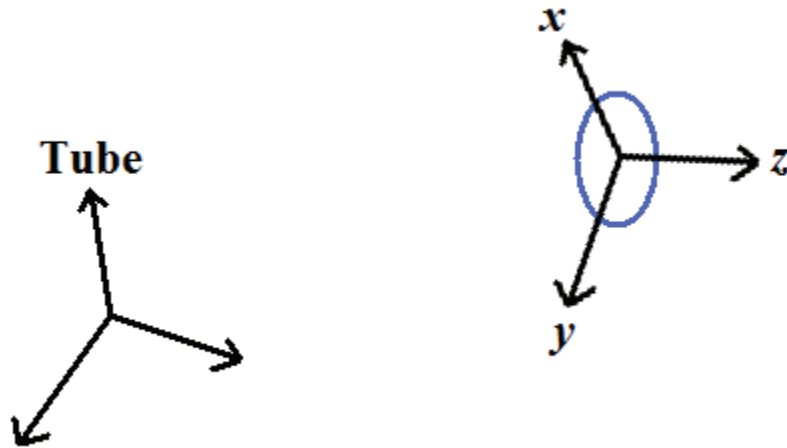


Figure 2.4. Repère défini par rapport au cercle et décrivant ce dernier

Ce repère peut être représenté à l'aide d'une transformation rigide (translation et rotation) T_C transformant un point exprimé dans le repère du tube en un point exprimé dans le repère du cercle. Les détails du calibrage nécessaire à l'évaluation du cercle et donc de la transformation sont donnés plus loin. Convertir un point P_{mI} pour qu'il soit exprimé dans le repère du cercle revient à projeter le point P_{mI} sur le cercle. Il est ensuite aisé de déterminer l'angle dans le cercle à partir des coordonnées X et Y du point.

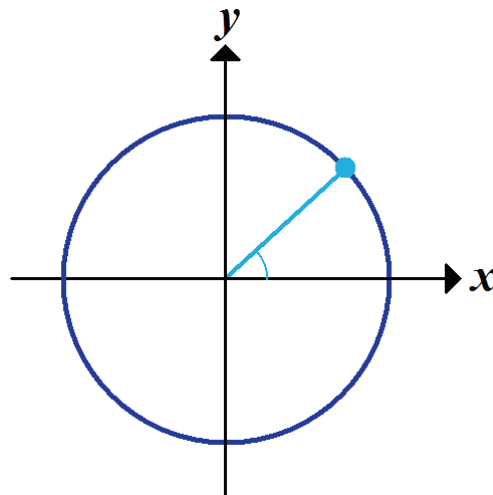


Figure 2.5. Cercle permettant de déterminer l'angle associé à chacune des positions

En résumé, pour appliquer la transformation correspondant à la rotation de la tête de la caméra par rapport au tube, deux paramètres doivent être calibrés : l'axe de rotation autour

duquel le point 3D à transformer est tourné, exprimé par un point p_r et un axe u_r , ainsi que la transformation rigide associée au cercle décrit par un point de la tête dans le repère du marqueur du tube, exprimée par une matrice M_C .

Une fois ce calibrage effectué, l'angle de rotation peut être estimé en temps réel à partir de l'état $S(m1, m2)$ du système de suivi :

$$p(S) = M_C S(m1, m2) P_{m2} = M_C M_{m1 \leftarrow w} M_{w \leftarrow m2} P_{m2} \quad (2.4)$$

où $p(S)$ est la position du point P_{m2} projeté dans le cercle pour un état S du système de suivi. L'angle de rotation est ensuite déterminé simplement :

$$\theta(S) = \tan^{-1} \left(p(S)_x, p(S)_y \right) \quad (2.5)$$

Une fois cet angle déterminé, la matrice M_R correspondant à la rotation peut finalement être déterminée :

$$M_R = T(p_r) R(u_r, \theta) T(-p_r) \quad (2.6)$$

où $T(p)$ est une matrice correspondant à une translation amenant l'origine au point p et $R(n, \theta)$ est une rotation d'un angle θ autour de l'axe n .

Donc, si P_1 est le point considéré exprimé en coordonnées caméra tel que défini précédemment et P_2 le même point après la transformation correspondant à la rotation relative entre la tête de caméra et le tube, cette transformation peut être résumée par l'équation suivante :

$$P_2 = T(p_r) R(d_r, \theta) T(-p_r) P_1 \quad (2.7)$$

2.2.3 Transformation 3 : distorsion de lentille

À cette étape, le point 3D considéré ayant été transformé en coordonnées caméra et ayant subi la rotation de la tête de caméra peut être interprété comme un point 2D en coordonnées homogènes. En effet, le point $P_2 = [x \ y \ z]^T$ peut être vu comme le point 2D $P_2' = [x' \ y'] = [x/z \ y/z]^T$, ce qui revient à effectuer la même projection que celle considérée par le modèle de la caméra à sténopé (voir section I.1.3 en annexe) en considérant une distance focale unitaire pour avoir une coordonnée normalisée. Ainsi, considérer le point P_2 comme un point 2D en

coordonnées homogènes plutôt qu'un point 3D revient à le projeter sur un plan image situé à 1 unité de l'origine.

On peut être tenté d'appliquer la transformation correspondant à la distorsion à la toute dernière étape des transformations représentant le modèle de formation de l'image, c'est-à-dire après avoir converti le point P_2 en coordonnées image exprimées en pixels. Toutefois, afin de modéliser plus adéquatement le phénomène physique de distorsion de lentille, la transformation correspondant à la distorsion doit être appliquée avant la numérisation faite par le CCD. En effet, la distorsion radiale causée par les lentilles tend à rapprocher ou à éloigner les points 2D du centre de projection dans les coordonnées de la caméra, étant donné que la lentille causant la distorsion agit dans ce repère.

La numérisation de l'image qui se produit après la lentille comporte généralement un ré-échantillonnage des lignes du capteur entraînant des pixels non carrés, tel que décrit à la section I.2.1 en annexe. Lorsque les pixels sont carrés, ce qui revient à affirmer que leur rapport d'aspect est de 1, que les distances focales sont les mêmes en X et en Y ou que la résolution spatiale de la numérisation est la même en X et en Y , l'application de la transformation correspondant à la distorsion peut être faite de manière équivalente en coordonnées caméra ou en coordonnées image. Toutefois, ce n'est pas le cas lorsque les distances focales ne sont pas les mêmes en X et en Y . En effet, dans un tel cas, un ensemble de points représentant un cercle dans un des repères représente plutôt une ellipse dans l'autre repère. Ainsi, comme la distorsion est radiale et donc uniquement une fonction de la distance au centre de distorsion, l'application de la transformation dans l'un ou l'autre des systèmes de coordonnées n'est pas équivalente.

Comme le phénomène de distorsion de lentille est complètement indépendant des paramètres de numérisation tel que la résolution ou le rapport d'aspect des pixels, il est plus approprié d'appliquer la transformation correspondant à ce phénomène avant celle correspondant à la numérisation.

Les paramètres k_1 et k_2 sont les paramètres de distorsion radiale à calibrer. Il s'agit des seuls paramètres à considérer, puisque la distorsion tangentielle est ignorée. Ainsi, à partir du calcul du rayon du point ($r^2 = x'^2 + y'^2$), on peut exprimer les nouvelles coordonnées après l'application de la transformation associée à la distorsion :

$$\begin{aligned} x'' &= x' (1 + k_1 r^2 + k_2 r^4) \\ y'' &= y' (1 + k_1 r^2 + k_2 r^4) \end{aligned} \quad (2.8)$$

Le point $P_3 = [x'' \ y'' \ 1]^T$ représente donc le point après l'application de la transformation associée à la distorsion.

2.2.4 Transformation 4 : numérisation

Les coordonnées du point après l'application de la transformation correspondant à la distorsion de lentille sont toujours en coordonnées caméra, dans un plan situé à 1 unité du point principal sur l'axe optique. Les coordonnées image, en pixels, sont obtenues par l'application de la transformation représentée par la matrice K des paramètres intrinsèques qui contient les distances focales en plus des coordonnées du centre optique. Cette transformation, comprenant une mise à l'échelle et une translation, représente la numérisation de l'image par le *CCD* et permet de passer en coordonnées image, exprimées en pixels, c'est-à-dire en unités de numérisation. Ainsi, le point $[u \ v]^T$ en coordonnées image est donné par :

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} f_x & 0 & -x_0 \\ 0 & f_y & -y_0 \\ 0 & 0 & -1 \end{bmatrix} P_3 = KP_3 \quad (2.9)$$

Le choix d'utiliser la version négative de la troisième colonne plutôt que celle de l'équation 1.1 permet de simplifier la correspondance avec le modèle de caméra de OpenGL en gardant les mêmes définitions pour les axes du repère lié à la caméra et les mêmes conventions pour l'image inversée ou non (voir section I.1.3 en annexe).

Donc, si P_3 est le point considéré après l'application de la transformation correspondant à la distorsion de lentille et que P_4 est le même point après la numérisation et exprimé en coordonnées homogènes dans le repère image dont l'unité est le pixel, cette transformation peut être résumée par l'équation suivante :

$$P_4 = KP_3 \quad (2.10)$$

Ainsi, les paramètres à calibrer pour cette étape de l'application du modèle sont les deux distances focales f_x et f_y ainsi que les coordonnées du centre optique x_0 et y_0 .

Cela complète le modèle de formation d'images assumé pour la caméra endoscopique. La prochaine section décrit la méthode hors-ligne pour obtenir les valeurs des différents paramètres à calibrer.

2.3 Méthode de calibrage hors-ligne

La présente section énonce la méthode de calibrage hors-ligne utilisée avant la chirurgie pour déterminer les valeurs des paramètres du modèle défini à la section 2.2.

2.3.1 Choix d'implémentation

La méthode de calibrage de Zhang [69] tel qu'implémentée dans la librairie OpenCV [71] a été retenue pour l'implémentation du calibrage proposé par cette méthode pour plusieurs raisons.

Tout d'abord, la méthode de Zhang permet d'utiliser des objets de calibrage planaires, beaucoup plus simples à fabriquer que des objets tridimensionnels. De plus, en incluant la distorsion dans la procédure de calibrage, elle ne nécessite pas de manipulation particulière pour le positionnement de l'objet de calibrage par rapport à la caméra, ce qui simplifie le processus de calibrage.

Par ailleurs, le choix d'utiliser l'implémentation fournie par OpenCV provient du désir de vérifier si une solution déjà existante et couramment utilisée peut répondre au besoin de calibrage de l'endoscope. Cette solution déjà implémentée dans une librairie de type logiciel libre est plus efficace à explorer que la réalisation d'une implémentation maison. Par contre, si elle ne s'avère pas suffisante, elle constituera quand même un point de départ qui pourra être remplacé par une méthode sur mesure grâce au développement modulaire proposé par ce travail de recherche. Finalement, l'implémentation de OpenCV est utilisée, et donc testée, par une grande communauté, ce qui donne un bon niveau de confiance en la qualité de la solution.

La méthode principale utilisée pour effectuer le calibrage est la fonction `cvCalibrateCamera2()`. Sans trop entrer dans les détails des paramètres de cette fonction, qui peuvent être retrouvés dans la documentation de OpenCV [71], cette fonction prend un certain nombre d'entrées pour fournir un certain nombre de sorties. En entrée, on note d'abord les positions 3D des points de l'objet de calibrage utilisé, qui doivent former un plan. On y trouve également les positions 2D de ces mêmes points de l'objet de calibrage telles que

vues sur les images utilisées pour le calibrage. Il est possible de fournir plusieurs vues de l'objet de calibrage pour effectuer le calibrage. En sortie, la fonction fournit les paramètres intrinsèques de la caméra tels que calculés à partir des points 3D et 2D passés en paramètres. Ces paramètres sont retournés sous la forme d'une matrice de paramètres intrinsèques et de coefficients de distorsion qui sont les mêmes pour toutes les vues. Elle retourne également les paramètres extrinsèques sous la forme d'une rotation et d'une translation pour chacune des vues de l'objet de calibrage passées en paramètres. En effet, même si les paramètres intrinsèques sont communs à toutes les vues, les paramètres extrinsèques ne le sont généralement pas puisque la caméra (ou l'objet de calibrage) est généralement déplacée d'une vue à l'autre.

2.3.2 Conception de l'objet de calibrage

L'objet de calibrage choisi est représenté à la Figure 2.6.

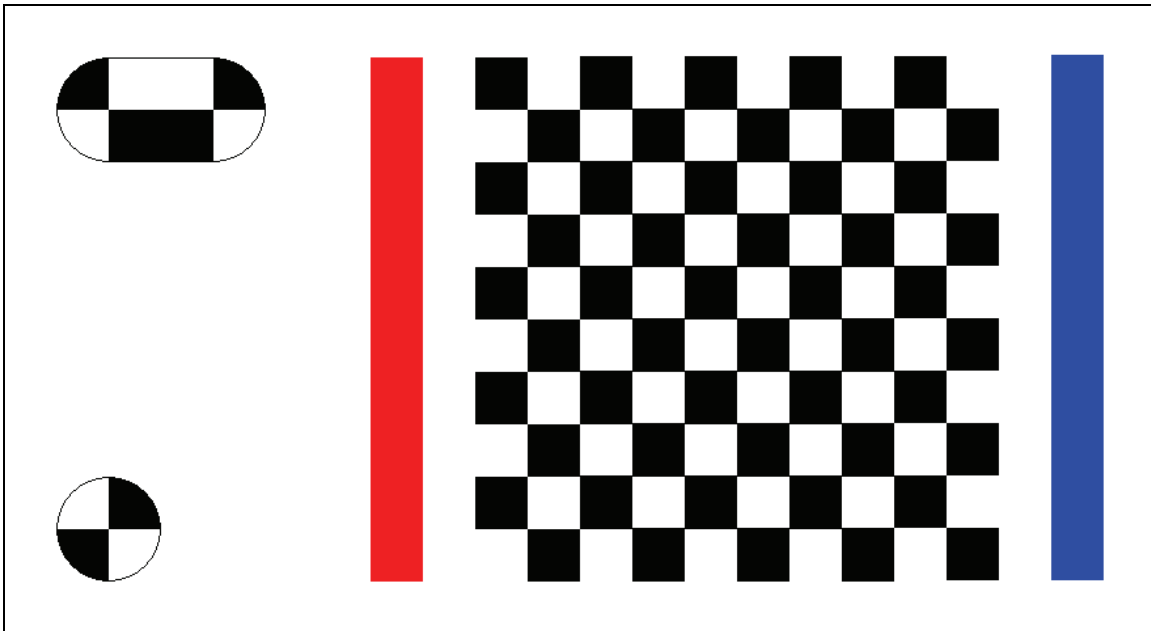


Figure 2.6. Patron de l'objet de calibrage

Cet objet de calibrage a été imprimé puis laminé afin d'obtenir un objet rigide et résistant, représenté à la Figure 2.7.

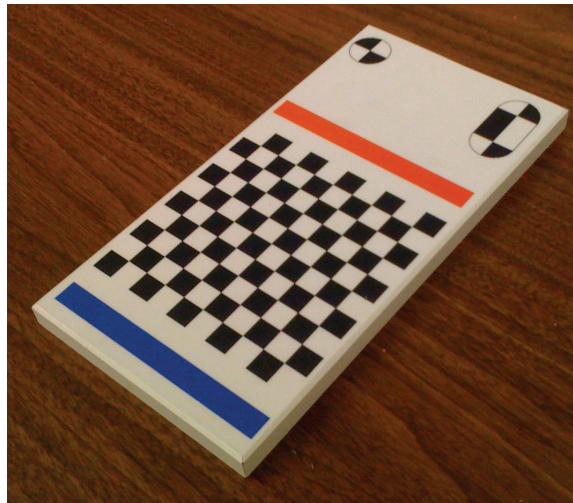


Figure 2.7. Objet de calibrage réel

La conception de cet objet de calibrage est basée sur trois éléments clés : la grille de calibrage, les bandes colorées et le marqueur MicronTracker.

2.3.2.1 Grille de calibrage

Bien qu'il existe plusieurs stratégies pour la conception d'une grille de calibrage (utilisation de lignes perpendiculaires, utilisation de cercles, etc.), le patron d'échiquier a été retenu. Ce dernier peut être localisé automatiquement et avec précision grâce à l'algorithme présenté en annexe (voir ANNEXE IV). L'automatisation de l'identification des points de la grille de calibrage est un avantage majeur, car il élimine tout besoin d'identification manuelle, ce qui permet d'accélérer énormément la procédure de calibrage, facilitant ainsi l'utilisation d'un plus grand nombre d'images et donc l'obtention d'un meilleur calibrage.

On a choisi une taille de grille de 10 carrés par 10 carrés, ce qui fournit 9 coins intérieurs par 9 coins intérieurs, pour un total de 81 points. Le côté de chaque carré mesure 8,00 mm, ce qui fait que l'objet de calibrage dans son ensemble mesure 96,0 mm par 160 mm. Cette taille permet à l'objet de calibrage d'occuper la majeure partie de l'image endoscopique lorsqu'il est situé à une distance comparable à celle des organes qui sont visibles au cours d'une chirurgie, soit environ 15 cm.

La grille choisie pour l'objet de calibrage permet donc d'obtenir facilement et automatiquement les coordonnées 2D des points de la grille dans les vues utilisées pour le calibrage.

2.3.2.2 Bandes colorées

L'identification automatique des points de la grille peut poser une certaine ambiguïté. En effet, dans le cas d'une grille carrée, il existe quatre positions de l'objet de calibrage qui positionneront la grille au même endroit⁷. L'association entre les coordonnées 2D des points identifiées sur l'image et les coordonnées 3D de l'objet de calibrage ne peut être faite à partir de la grille seulement (i.e. on ne sait pas à quel coin de l'objet de calibrage 3D correspond un coin donné de la grille identifiée en 2D).

Afin de résoudre cette ambiguïté, on ajoute deux bandes colorées aux extrémités de la grille. Ces bandes peuvent également être identifiées facilement sur les vues, ce qui permet d'éliminer l'ambiguïté.

Il serait également possible, voire même souhaitable d'utiliser une seule bande noire plutôt que des bandes colorées, comme on peut le voir à la Figure 2.8. Cette configuration permet également de résoudre l'ambiguïté et ne nécessite pas l'utilisation d'imprimante couleur.

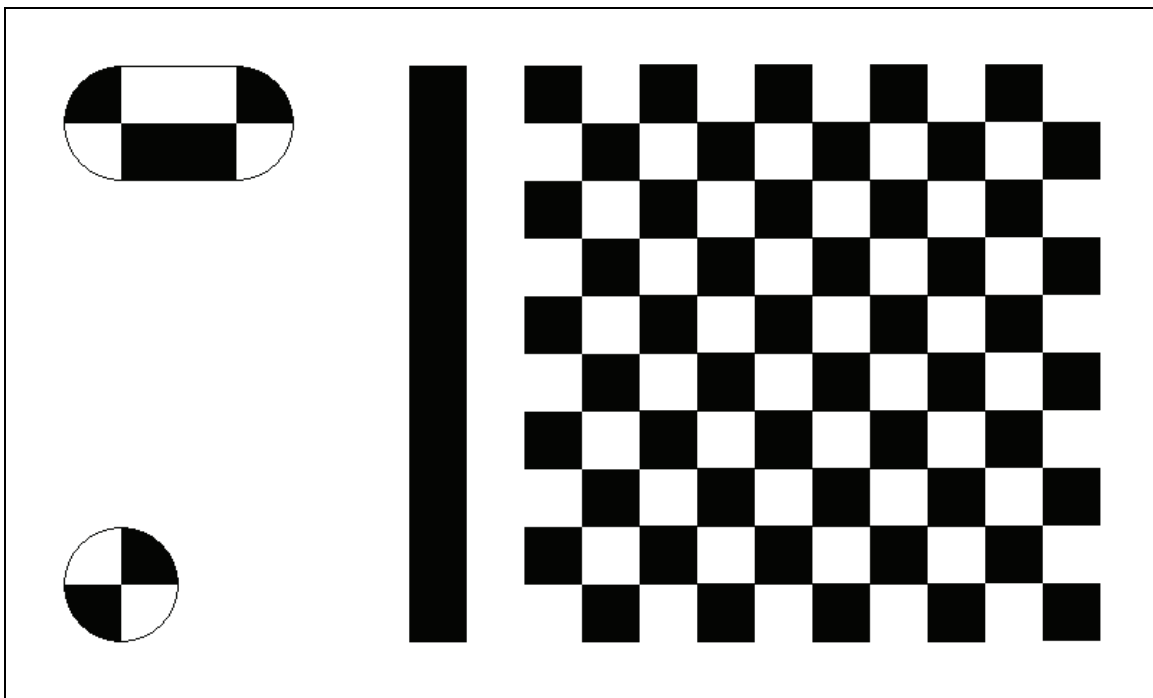


Figure 2.8. Objet de calibrage avec une bande noire plutôt que des bandes colorées

⁷ Le nombre de configurations ambiguës est réduit à deux si la grille est rectangulaire plutôt que carrée.

Les bandes colorées ont toutefois été utilisées dans le cadre de ce travail de recherche. Elles permettent donc d'éliminer l'ambiguïté dans l'association des points de la grille 2D avec les points 3D de l'objet de calibrage.

2.3.2.3 Marqueur MicronTracker

Un marqueur MicronTracker est également intégré à l'objet de calibrage afin d'être en mesure d'obtenir les positions 3D des points de la grille de calibrage. Ce marqueur est nécessaire pour pouvoir obtenir les positions 3D des points de l'objet de calibrage. En effet, on doit être en mesure d'exprimer ces points dans l'espace MicronTracker afin que les paramètres extrinsèques retournés par la procédure de calibrage soit également en coordonnées MicronTracker. Même si on utilisait un outil de numérisation 3D comme le MicroScribe pour effectuer la numérisation, ce marqueur serait tout de même nécessaire pour effectuer la conversion entre l'espace du numériseur et celui du MicronTracker. Il est à noter que la construction de l'objet de calibrage est conçue de manière à ne pas nécessiter le recours à un numériseur 3D puisque les coordonnées des points peuvent être déterminées à partir de la disposition géométrique des éléments de l'objet. La précision de la mesure des points de la grille de calibrage est donc reliée à la précision de l'impression⁸ plutôt qu'à la précision de la numérisation des coordonnées 3D des points de la grille.

L'objet de calibrage et son marqueur sont construits de manière à permettre une conversion facile entre les unités du MicronTracker et les unités de la grille. Par exemple, l'axe X de l'objet de calibrage de la Figure 2.6 mesure exactement 8 carrés de la grille alors que l'axe Y en mesure exactement 2. Il est facile d'exprimer les coins de la grille dans le repère du marqueur en unités de la grille. Par exemple, les coordonnées des coins de la grille du marqueur de la Figure 2.6 sont $(-4, -8)$ pour le coin supérieur gauche, $(-4, -16)$ pour le coin supérieur droit, $(4, -16)$ pour le coin inférieur droit et $(4, -8)$ pour le coin inférieur gauche. Lorsque l'on connaît la longueur de l'axe X du marqueur en unités MicronTracker c'est-à-dire en millimètres⁹, il est facile de convertir ces unités en millimètres et donc d'avoir les points exprimés dans le repère du marqueur. Le système de suivi fournissant la transformation du

⁸ À titre indicatif, les imprimantes d'entrée de gamme d'aujourd'hui ont une résolution d'au moins 1200 points-par-pouce, ce qui correspond à 1 pouce / 1200 points ou 0,0212 mm / point.

⁹ Cette longueur est connue puisqu'elle est nécessaire à la reconnaissance du marqueur par le MicronTracker. Elle est enregistrée dans le patron du marqueur.

repère du marqueur vers le monde, on a donc les coordonnées 3D des coins de la grille de calibrage dans l'espace MicronTracker.

Ainsi, l'inclusion du marqueur MicronTracker dans l'objet de calibrage permet d'obtenir les positions 3D des points de la grille nécessaires au calibrage.

2.3.3 Acquisition des vues de calibrage

Les acquisitions nécessaires à la réalisation du calibrage se font à l'aide d'un outil développé à l'interne qui permet d'enregistrer plusieurs états, chaque état comprenant :

- L'image de la caméra endoscopique.
- La position 2D des coins de la grille identifiés automatiquement sur l'image de la caméra endoscopique.
- Les marqueurs visibles par le système de suivi avec leur position et leur orientation.

La Figure 2.9 montre une capture d'écran de l'outil d'enregistrement. On y voit la grille identifiée automatiquement, tout comme les bandes de couleur qui permettent de localiser le coin inférieur droit, identifié par une paire d'axes rouge et bleu.

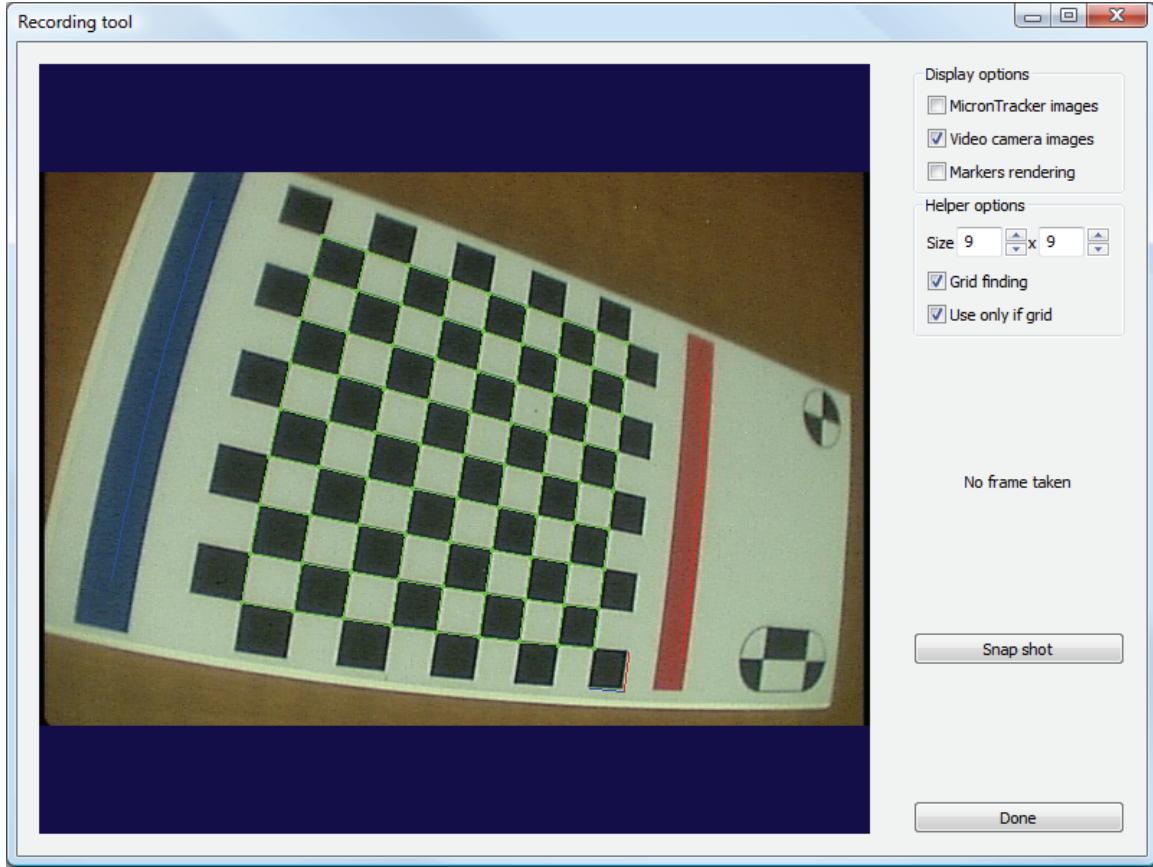


Figure 2.9. Outil d'enregistrement des images de calibrage

2.3.3.1 Algorithme d'identification des coins de la grille

La première partie de l'algorithme, qui consiste à identifier les coins de la grille sur l'image, est présentée à l'ANNEXE IV. Une fois cette première étape accomplie, il reste toujours une ambiguïté quand à l'orientation de cette grille.

Soit $p(i, j)$ le point de la grille appartenant à la i^{e} colonne et à la j^{e} ligne, où $0 \leq i < M$, M étant la dimension horizontale de la grille, et $0 \leq j < N$, N étant la dimension verticale de la grille.

La grille fournie par l'algorithme de l'ANNEXE IV est de telle sorte que les axes de la grille correspondent à un « système direct ». Autrement dit, on peut définir des axes tels que :

$$\begin{aligned}\vec{x} &= p(i+1, j) - p(i, j) \\ \vec{y} &= p(i, j+1) - p(i, j)\end{aligned}\tag{2.11}$$

La grille respecte toujours la condition que le produit vectoriel donne une composante positive :

$$\vec{x}_x \vec{y}_y - \vec{x}_y \vec{y}_x > 0 \quad (2.12)$$

Par contre, le point $p(0, 0)$ peut correspondre à n'importe quel des quatre coins de l'objet de calibrage. Les bandes de couleurs sont utilisées pour résoudre cette ambiguïté.

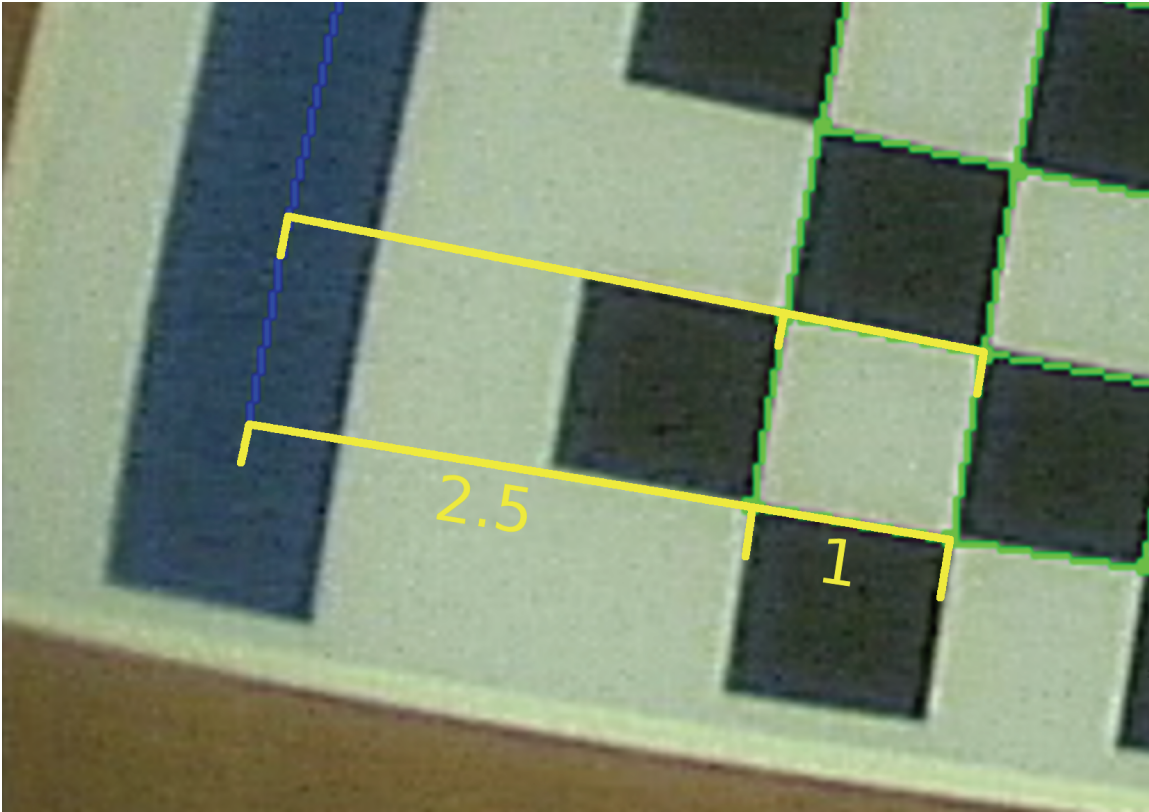


Figure 2.10. Construction permettant de trouver la couleur moyenne de la bande

Par construction, comme on peut le voir sur la Figure 2.10, la ligne centrale des bandes de couleurs est située à 2,5 unités des côtés. On utilise ce fait pour calculer la couleur moyenne à 2,5 unités de chacun des côtés de la grille. Soit $C(p)$ la couleur de l'image au point p . On définit les couleurs moyennes NORD, SUD, EST et OUEST :

$$\begin{aligned}
c_{NORD} &= \frac{1}{M} \sum_{i=0}^{M-1} C(p(i,0) + 2,5(p(i,0) - p(i,1))) \\
c_{SUD} &= \frac{1}{M} \sum_{i=0}^{M-1} C(p(i,N-1) + 2,5(p(i,N-1) - p(i,N-2))) \\
c_{EST} &= \frac{1}{N} \sum_{j=0}^{N-1} C(p(0,j) + 2,5(p(0,j) - p(1,j))) \\
c_{OUEST} &= \frac{1}{N} \sum_{j=0}^{N-1} C(p(M-1,j) + 2,5(p(M-1,j) - p(M-2,j)))
\end{aligned} \tag{2.13}$$

Les points dont les coordonnées tombent en dehors de l'image ne sont pas comptés dans la moyenne des couleurs. Ensuite, les points de la grille sont réordonnées (par une rotation de position) de manière à ce que c_{NORD} soit le plus près possible du rouge pur, c'est-à-dire (1, 0, 0) en format RGB, et que c_{SUD} soit le plus près possible du bleu pur. La distance utilisée est la distance euclidienne dans l'espace RGB. Bien que cette méthode soit suffisante pour les besoins de l'application, il serait facile d'utiliser plutôt l'espace HSV par exemple, ou alors d'utiliser une bande noire comme celle de la Figure 2.8, si le besoin d'améliorer la méthode se présentait.

2.3.4 Calibrage sans rotation

Cette partie du calibrage hors-ligne nécessite l'acquisition d'au moins une vue de l'objet de calibrage. Idéalement, en plus de la grille identifiée automatiquement, on a pour chacune des vues la position de deux marqueurs clés : celui sur l'objet de calibrage ainsi que celui sur le tube de la caméra. Pour cette étape, le marqueur situé sur la tête de la caméra n'a pas besoin d'être visible. On doit toutefois s'assurer qu'aucune rotation n'est effectuée entre la tête de la caméra et le tube entre les différentes vues prises pour le calibrage.

On a donc, pour chacune des vues :

- La position 3D des points de l'objet de calibrage : ces positions sont exprimées dans le repère du marqueur de l'objet de calibrage. Ces positions sont d'abord définies par la configuration de l'objet de calibrage puis converties en millimètres grâce à la longueur de l'axe X de l'objet de calibrage. La raison pour laquelle on utilise ce repère est qu'OpenCV s'assure que l'objet est planaire en vérifiant que la moyenne des valeurs Z des points 3D est pratiquement nulle, tout comme l'écart-type. Il faudra tenir compte de ce choix lors de l'interprétation de la matrice des paramètres extrinsèques, qui

seront exprimés dans le repère du marqueur de l'objet de calibrage. De plus, ce choix a pour conséquence que les points 3D de l'objet de calibrage seront les mêmes pour toutes les vues, peu importe si on déplace l'objet de calibrage entre les vues.

- La position 2D des points de l'objet de calibrage : ces positions sont exprimées en pixels, avec l'origine au coin inférieur gauche, l'axe X positif orienté vers la droite et l'axe Y positif orienté vers le haut.

2.3.4.1 Paramètres intrinsèques

Le calibrage par la méthode de Zhang implémentée par OpenCV donne directement les paramètres intrinsèques ainsi que les paramètres de distorsion, c'est-à-dire f_x, f_y, x_0, y_0, k_1 et k_2 .

2.3.4.2 Paramètres extrinsèques

On peut ensuite se servir des paramètres extrinsèques pour retrouver la transformation $T_{c \leftarrow m}$ entre le marqueur sur le tube et le repère de la caméra. Pour cette étape, seules les vues où le marqueur du tube et le marqueur de l'objet de calibrage sont tous deux visibles peuvent être utilisées.

Pour chacune de ces vues, on récupère les paramètres extrinsèques fournis par OpenCV sous la forme d'un vecteur de rotation r et d'un vecteur de translation t . Le modèle utilisé par OpenCV est toutefois légèrement différent de celui utilisé dans ce travail :

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r[0] & r[1] & r[2] & t[0] \\ r[3] & r[4] & r[5] & t[1] \\ r[6] & r[7] & r[8] & t[2] \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.14)$$

Or, l'équation 2.9 est construite de manière à ce que la dernière colonne de la matrice des paramètres intrinsèques soit négative. On peut donc transformer l'équation précédente :

$$\begin{aligned}
& \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r[0] & r[1] & r[2] & t[0] \\ r[3] & r[4] & r[5] & t[1] \\ r[6] & r[7] & r[8] & t[2] \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \\
& \begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} r[0] & r[1] & r[2] & t[0] \\ r[3] & r[4] & r[5] & t[1] \\ r[6] & r[7] & r[8] & t[2] \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \\
& \left(\begin{bmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \right) \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} r[0] & r[1] & r[2] & t[0] \\ r[3] & r[4] & r[5] & t[1] \\ r[6] & r[7] & r[8] & t[2] \end{bmatrix} \right) \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}
\end{aligned} \tag{2.15}$$

pour finalement obtenir :

$$\begin{bmatrix} f_x & 0 & -x_0 \\ 0 & f_y & -y_0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} r[0] & r[1] & r[2] & t[0] \\ r[3] & r[4] & r[5] & t[1] \\ -r[6] & -r[7] & -r[8] & -t[2] \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.16}$$

De plus, le calibrage donne un système de coordonnées indirect, c'est-à-dire qui respecte la règle de la main gauche plutôt que celle de la main droite. Comme les paramètres extrinsèques sont exprimés dans le repère du marqueur de l'objet de calibrage, cela revient à supposer que l'axe des Z n'« entre » pas dans le marqueur, mais « sort » de ce dernier. Cette différence vient du fait que OpenCV considère le système de coordonnées des pixels comme ayant son origine dans le coin supérieur gauche (plutôt qu'inférieur gauche) et donc l'axe des Y se trouve ainsi inversé.

Les points que l'on passe dans ce modèle doivent donc voir leur coordonnée Z être inversée :

$$\begin{aligned} \begin{bmatrix} X \\ Y \\ -Z \\ 1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ -Z \\ 1 \end{bmatrix} &= \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \end{aligned} \quad (2.17)$$

La matrice des paramètres extrinsèques devient donc :

$$\begin{aligned} \begin{bmatrix} r[0] & r[1] & r[2] & t[0] \\ r[3] & r[4] & r[5] & t[1] \\ -r[6] & -r[7] & -r[8] & -t[2] \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ -Z \\ 1 \end{bmatrix} \right) \\ \left(\begin{bmatrix} r[0] & r[1] & r[2] & t[0] \\ r[3] & r[4] & r[5] & t[1] \\ -r[6] & -r[7] & -r[8] & -t[2] \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} X \\ Y \\ -Z \\ 1 \end{bmatrix} \\ \begin{bmatrix} r[0] & r[1] & -r[2] & t[0] \\ r[3] & r[4] & -r[5] & t[1] \\ -r[6] & -r[7] & r[8] & -t[2] \end{bmatrix} \begin{bmatrix} X \\ Y \\ -Z \\ 1 \end{bmatrix} \end{aligned} \quad (2.18)$$

Ainsi, la matrice des paramètres extrinsèques qui correspond au modèle présenté dans ce travail, lorsque les paramètres extrinsèques sont exprimés par rapport au repère du marqueur de l'objet de calibrage (ce repère étant noté par l'indice g), est :

$$E_{c \leftarrow g} = \begin{bmatrix} r[0] & r[1] & -r[2] & t[0] \\ r[3] & r[4] & -r[5] & t[1] \\ -r[6] & -r[7] & r[8] & -t[2] \end{bmatrix} \quad (2.19)$$

2.3.4.3 Transformation $T_{c \leftarrow m}$

Cette matrice permet d'obtenir la transformation $T_{c \leftarrow m}$ recherchée :

$$T_{c \leftarrow m} = E_{c \leftarrow g} T_{g \leftarrow w} T_{w \leftarrow m} \quad (2.20)$$

La transformation $T_{w \leftarrow m}$ du marqueur situé sur le tube vers le repère du monde est donnée par le système de suivi, tout comme la transformation $T_{g \leftarrow w}$ du repère du monde vers celui du marqueur situé sur l'objet de calibrage. Il est donc possible d'obtenir un estimé de $T_{c \leftarrow m}$ pour chaque vue pour laquelle les deux marqueurs sont identifiés par le système de suivi.

Lorsque plus d'une vue est disponible, on effectue la moyenne des transformations rigides trouvées à l'aide de la méthode présentée en annexe à la section III.4.

2.3.4.4 Raffinement

Comme la transformation $T_{c \leftarrow m}$ est obtenue par une moyenne pouvant posséder des erreurs, on effectue une optimisation non linéaire de ce paramètre afin d'améliorer son estimation. On représente la transformation comme un vecteur de 6 paramètres, 3 des paramètres représentant la translation et 3 la rotation. Une rotation d'un angle θ autour d'un vecteur unitaire $[x, y, z]$ est représentée par les 3 valeurs θx , θy et θz .

On effectue ensuite la minimisation non linéaire de l'erreur de reprojection à l'aide de l'algorithme de Levenberg-Marquardt [68] (l'implémentation utilisée est *levmar*, disponible à <http://www.ics.forth.gr/~lourakis/levmar/>, consulté le 2010/04/28). Pour ce faire, on calcule l'erreur de reprojection pour chacun des points de l'objet de calibrage en fonction de la transformation $T_{c \leftarrow m}$ correspondant aux 6 paramètres à optimiser. L'erreur est calculée de la manière suivante :

1. La transformation $T_{c \leftarrow m}$ est reconstruite à partir des 6 paramètres actuels passés par la procédure d'optimisation.
2. Pour chacune des vues, chacun des points de l'objet de calibrage, alors exprimé en coordonnées du marqueur de l'objet de calibrage, est converti en coordonnées monde à l'aide de la transformation $T_{w \leftarrow g}$ fournie par le système de suivi.
3. Chaque point est ensuite converti en coordonnées du marqueur du tube à l'aide de la transformation $T_{m \leftarrow w}$ fournie par le système de suivi.
4. Chaque point est ensuite converti en coordonnées de la caméra à l'aide de la transformation $T_{c \leftarrow m}$ pour laquelle l'erreur est évaluée.
5. Chaque point est ensuite converti en coordonnées 2D par l'application de la transformation représentant la distorsion et de la conversion en coordonnées image à

l'aide des paramètres intrinsèques (aucune rotation due à la tête de caméra n'est appliquée puisque l'acquisition a été faite sans rotation).

6. L'erreur de chaque point correspond à la distance entre la projection calculée et la véritable position du point identifié sur l'image. Une variante consiste à prendre chacune des distances en x et en y plutôt que la distance euclidienne.

Cette minimisation, prenant la moyenne calculée comme estimé initial donne finalement la valeur calibrée du paramètre $T_{c \leftarrow m}$.

2.3.5 Calibrage avec rotation

Pour cette partie du calibrage hors-ligne, on doit acquérir un certain nombre de vues de l'objet de calibrage en effectuant une rotation relative entre la tête de la caméra et le tube de l'endoscope. La première partie du calibrage avec rotation n'a besoin que de l'information provenant du système de suivi, c'est-à-dire la pose du marqueur situé sur le tube ainsi que celle du marqueur situé sur la tête de la caméra. La seconde partie du calibrage avec rotation nécessite des vues « complètes », c'est-à-dire où la grille est visible et identifiée par la caméra endoscopique, et où les marqueurs de l'objet de calibrage, du tube et de la tête de la caméra sont reconnus par le système de suivi. Les vues complètes peuvent évidemment être utilisées pour la première partie du calibrage avec rotation.

2.3.5.1 Détermination de l'angle de rotation

Par l'équation 2.3, chaque vue où les marqueurs du tube et de la tête de la caméra sont identifiés par le système de suivi fournit un point que l'on nomme $P_i = [X_i \ Y_i \ Z_i]^T$ exprimé dans le système du repère du tube. Les points P_i décrivent un cercle par la rotation de la tête de la caméra par rapport au tube.

Pour déterminer le centre de ce cercle, on commence par déterminer le meilleur plan, au sens des moindres carrés, qui passent par tous les points, à l'aide de l'algorithme de la section III.5 présenté en annexe. Ce plan respecte donc l'équation $A_F + B_F + C_F + D_F = 0$. Le centre du cercle devrait se trouver sur ce plan. Il devrait également se trouver le plan médiateur de chaque paire de points p_i et p_j . Ce plan a pour normale le vecteur $p_i - p_j$ et passe par le point $(p_i + p_j) / 2$, ce qui donne le développement suivant :

$$\begin{aligned}
A &= p_{i,x} - p_{j,x} \\
B &= p_{i,y} - p_{j,y} \\
C &= p_{i,z} - p_{j,z} \\
-D &= A \frac{p_{i,x} + p_{j,x}}{2} + B \frac{p_{i,y} + p_{j,y}}{2} + C \frac{p_{i,z} + p_{j,z}}{2} \\
&= \frac{(p_{i,x} - p_{j,x})(p_{i,x} + p_{j,x}) + (p_{i,y} - p_{j,y})(p_{i,y} + p_{j,y}) + (p_{i,z} - p_{j,z})(p_{i,z} + p_{j,z})}{2} \\
&= \frac{p_{i,x}^2 - p_{j,x}^2 + p_{i,y}^2 - p_{j,y}^2 + p_{i,z}^2 - p_{j,z}^2}{2} \\
&= \frac{(p_{i,x}^2 + p_{i,y}^2 + p_{i,z}^2) - (p_{j,x}^2 + p_{j,y}^2 + p_{j,z}^2)}{2} \\
&= \frac{\|p_i\|^2 - \|p_j\|^2}{2}
\end{aligned} \tag{2.21}$$

pour obtenir l'équation :

$$(p_{i,x} - p_{j,x})x + (p_{i,y} - p_{j,y})y + (p_{i,z} - p_{j,z})z = \frac{\|p_i\|^2 - \|p_j\|^2}{2} \tag{2.22}$$

On choisit les paires p_i et p_{i+1} , donnant donc $N - 1$ équations pour les plans médiateurs auxquelles s'ajoute l'équation pour le meilleur plan. Le système obtenu est donc :

$$\begin{bmatrix} p_{0,x} - p_{1,x} & p_{0,y} - p_{1,y} & p_{0,z} - p_{1,z} \\ \vdots & \vdots & \vdots \\ p_{i,x} - p_{i+1,x} & p_{i,y} - p_{i+1,y} & p_{i,z} - p_{i+1,z} \\ \vdots & \vdots & \vdots \\ p_{N-2,x} - p_{N-1,x} & p_{N-2,y} - p_{N-1,y} & p_{N-2,z} - p_{N-1,z} \\ A_F & B_F & C_F \end{bmatrix} \begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix} = \begin{bmatrix} \frac{\|p_0\| - \|p_1\|}{2} \\ \vdots \\ \frac{\|p_i\| - \|p_{i+1}\|}{2} \\ \vdots \\ \frac{\|p_{N-2}\| - \|p_{N-1}\|}{2} \\ -D_F \end{bmatrix} \tag{2.23}$$

La solution de ce système fournit un estimé pour le centre du cercle. On raffine ensuite cet estimé à l'aide d'une optimisation non linéaire afin que la variance de la distance du centre à chacun des points soit minimale.

On peut ensuite définir le système de coordonnées décrit à la section 2.2.2.1 et représenté à la Figure 2.4. L'origine du système, nommée O , est le centre du cercle après le raffinement non

linéaire. L'axe Z du système, nommé Z , est donné par le vecteur (normalisé) normal au meilleur plan, soit :

$$Z = \frac{[A_F, B_F, C_F]^T}{\|[A_F, B_F, C_F]^T\|} \quad (2.24)$$

Comme on veut que l'axe X corresponde à l'angle nul, c'est-à-dire l'angle avec le premier point, on définit l'axe Y comme le produit vectoriel (normalisé) de l'axe Z et du vecteur reliant le centre au premier point (afin de garder l'orthogonalité des axes du système) :

$$Y = Z \times \frac{(p_0 - C)}{\|p_0 - C\|} \quad (2.25)$$

L'axe X est ensuite tout simplement construit pour obtenir un système direct :

$$X = Y \times Z \quad (2.26)$$

Il est donc ensuite possible de construire une matrice de transformation reliant le système de coordonnées des points p_i au système de coordonnées construit à partir du cercle décrit par les points :

$$T_C = \begin{bmatrix} X_x & X_y & X_z & 0 \\ Y_x & Y_y & Y_z & 0 \\ Z_x & Z_y & Z_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -C_x \\ 0 & 1 & 0 & -C_y \\ 0 & 0 & 1 & -C_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.27)$$

Cette transformation, nécessaire à la détermination de l'angle de rotation, est donc calibrée.

2.3.5.2 Détermination de l'axe de rotation

Le principe derrière le processus de calibrage pour la détermination de l'axe de rotation est relativement simple : on cherche tout simplement l'axe de rotation qui minimise l'erreur de reprojection des points de l'objet de calibrage. La procédure est très semblable à celle utilisée pour le raffinement de $T_{c \leftarrow m}$ présenté à la section 2.3.4.3.

De manière plus détaillée, on effectue un prétraitement pour chacune des vues complètes, c'est-à-dire où la grille a été identifiée et où les poses des marqueurs du tube, de la tête de la caméra et de l'objet de calibrage sont connues. On exprime tout d'abord les points 3D de l'objet de calibrage dans le système de coordonnées de la caméra à l'aide de la transformation

$T_{w \leftarrow g}$ entre le système de coordonnées du marqueur de l'objet et celui du monde (fournie par le système de suivi), à l'aide de la transformation $T_{m \leftarrow w}$ entre le système de coordonnées du monde et celui du marqueur du tube de la caméra (fournie par le système de suivi) et à l'aide de la transformation $T_{c \leftarrow m}$ entre le système du marqueur du tube de la caméra et celui de la caméra (fournie par le calibrage effectué précédemment). De plus, pour chacune des vues, on calcule l'angle de rotation entre la tête de la caméra et le tube à l'aide de la transformation T_C calibrée dans la section précédente.

Ensuite, pour un axe donné (tel que fourni par la procédure d'optimisation), on évalue l'erreur de reprojection obtenue en appliquant aux points exprimés dans le repère de la caméra la rotation de l'angle calculé autour de l'axe fourni. Les points ainsi obtenus sont projetés avec le reste du modèle, c'est-à-dire les paramètres intrinsèques ainsi que ceux de distorsion, puis comparés avec les points 2D identifiés sur les images. L'axe de rotation est ainsi optimisé jusqu'à l'obtention de l'erreur de reprojection minimale.

L'axe à optimiser est en fait une droite 3D, qui est souvent représentée par deux points ou par un point et un vecteur directeur, ce qui donne 6 degrés de liberté à la droite. Il est toutefois possible d'appliquer des contraintes à ces deux points sans réduire la généralité de la représentation, i.e. en pouvant toujours décrire toutes les droites possibles. Premièrement, on peut choisir le premier point appartenant à la droite comme le point de la droite situé le plus près de l'origine. De plus, on peut choisir le second point de telle manière à ce que le vecteur reliant un point à l'autre soit unitaire. Ainsi, on voit donc que la droite ne possède en fait que 4 degrés de liberté. Lors d'une optimisation non linéaire telle que celle utilisée pour calibrer cet axe, il est préférable d'utiliser une représentation minimale (i.e. sans redondance) de l'entité à optimiser.

Comme l'axe recherché est pratiquement perpendiculaire au plan XY , on choisit de représenter le premier point de la droite comme l'intersection de la droite et du plan XY , c'est-à-dire un point de la forme $(X, Y, 0)$. De plus, on choisit de représenter le vecteur directeur de la droite par les angles des coordonnées sphériques ϕ et θ , de telle sorte que le vecteur directeur est donné par $(\sin \phi \cos \theta, \sin \phi \sin \theta, \cos \phi)$.

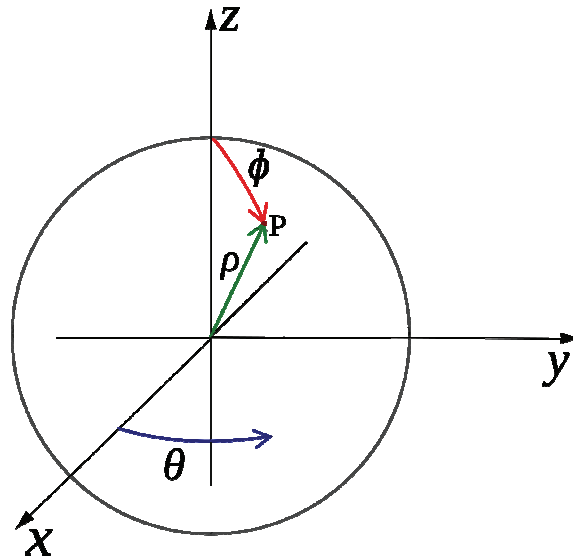


Figure 2.11. Illustration du système de coordonnées sphériques (tirée de Wikimedia Commons, "File:Spherical Coordinates.svg," *Wikimedia Commons*, 2007. [En ligne]. Disponible: http://commons.wikimedia.org/wiki/File:Spherical_Coordinates.svg. [Consulté le 2010/07/10]. ©2007 Cronholm144. Dans le domaine public.)

L'estimé initial est crucial pour le succès de l'algorithme de Levenberg-Marquardt. Ici, l'estimé est fourni par le cas idéal, car l'axe recherché devrait être très près de l'axe des Z du système de coordonnées de la caméra. Or, le sens de rotation positif ou négatif de la rotation est déterminé arbitrairement par la direction du vecteur normal du meilleur plan nécessaire au calcul de la transformation T_C . Pour résoudre cette ambiguïté, on effectue deux optimisations de l'axe de rotation : une avec comme estimé initial l'axe des Z positifs et une avec comme estimé initial l'axe des Z négatifs. Il est probable que les deux optimisations fournissent le même résultat correct si la procédure d'optimisation réussit à faire renverser l'axe initialisé dans le mauvais sens. Toutefois, pour plus de robustesse, on choisit celui dont l'erreur de reprojection résultante est la plus faible.

On obtient donc une valeur calibrée pour l'axe de rotation, représenté par un point p_r et un axe de rotation u_r .

2.4 Méthode de mise à jour en ligne

Certains paramètres du modèle de formation de l'image ne peuvent pas être évalués une seule fois avant la chirurgie à l'aide de la procédure de calibrage hors-ligne. Ils doivent être mis à

jour en temps réel au cours de la chirurgie. La présente section décrit la méthode de mise à jour en ligne de ces paramètres.

L'information obtenue en temps réel provient du système de suivi et constitue en la pose du marqueur posé sur le tube de l'endoscope et celle du marqueur posé sur la tête de la caméra. Ces poses sont suffisantes pour mettre à jour les paramètres du modèle variant au cours de la chirurgie, soit les paramètres extrinsèques et l'angle de rotation de la tête de la caméra par rapport au tube de l'endoscope.

2.4.1 Mise à jour des paramètres extrinsèques

Lorsque la caméra endoscopique est déplacée, les paramètres extrinsèques du modèle doivent être mis à jour. L'étape de calibrage hors-ligne fournit la transformation rigide, représentée par $T_{c \leftarrow m}$ dans l'équation 2.2, entre le marqueur situé sur le tube et le repère de la caméra. Cette transformation est constante tant que le marqueur reste bien fixé sur le tube.

Par contre, la pose de la caméra (et donc le repère qui lui est associé) varie lorsque le chirurgien déplace la caméra. Or, le marqueur sur la caméra permet de suivre la position de cette dernière. Comme la pose du marqueur est connue (fournie par le système de suivi) et que la transformation entre le repère du marqueur et celui de la caméra est connue, il devient facile de trouver la pose de la caméra.

Ainsi, le système de suivi fournit la pose du marqueur représentée par $M_{m \leftarrow w}$. Ce terme mis à jour, l'équation 2.2 peut être utilisée, car tous les termes sont désormais connus et à jour, ce qui complète la mise à jour des paramètres extrinsèques.

2.4.2 Mise à jour de l'angle de rotation

Lorsque la tête de la caméra subit une rotation relative par rapport au tube de l'endoscope, l'angle de rotation entre ces deux éléments doit être mis à jour. L'étape de calibrage hors-ligne fournit l'axe de rotation représenté par une ligne droite définie à l'aide du point p_r et du vecteur directeur u_r . Cet axe est constant tant que l'assemblage de la caméra endoscopique n'est pas modifié.

Par contre, l'angle entre la tête de la caméra et le tube varie lorsque le chirurgien effectue une rotation relative entre ces éléments. Il est toutefois possible de retrouver cet angle grâce au marqueur posé sur chacun de ces deux éléments à partir des équations 2.4 et 2.5. On choisit le

point P_{m2} de l'équation 2.4 comme étant l'origine du repère du marqueur posé sur la tête de la caméra, ce qui simplifie l'équation pour donner :

$$p(S) = M_C M_{m1 \leftarrow w} P_{w \leftarrow m2} \quad (2.28)$$

où M_C est la transformation vers le repère du cercle trouvée à l'aide du calibrage hors-ligne, $M_{m1 \leftarrow w}$ est donnée par la pose du marqueur du tube fournie par le système de suivi, $P_{w \leftarrow m2}$ est simplement la position, sans tenir compte de l'orientation, du marqueur de la tête de caméra.

Ainsi, cette équation, dont tous les termes sont connus et à jour, peut être utilisée avec l'équation 2.5 pour déterminer l'angle de rotation.

Cet angle peut ensuite être utilisé dans l'équation 2.7 pour appliquer la transformation correspondant à la rotation relative entre la tête de la caméra et le tube de l'endoscope, ce qui complète la mise à jour de l'angle de rotation.

2.5 Résumé du calibrage

Le modèle de formation de l'image choisi peut être résumé par les équations suivantes :

$$\begin{aligned} P_1 &= T_{c \leftarrow m} M_{m \leftarrow w} P_0 \\ P_2 &= T(p_r) R(d_r, \theta) T(-p_r) P_1 \\ P_2' &= \begin{bmatrix} \frac{P_{2,x}}{P_{2,z}} & \frac{P_{2,y}}{P_{2,z}} & 1 \end{bmatrix}^T \\ r &= \sqrt{(P_2'_x)^2 + (P_2'_y)^2} \\ P_3 &= \begin{bmatrix} P_2'_x (1 + k_1 r^2 + k_2 r^4) \\ P_2'_y (1 + k_1 r^2 + k_2 r^4) \\ 1 \end{bmatrix} \\ P_4 &= K P_3 \end{aligned} \quad (2.29)$$

Le tableau suivant résume les différents paramètres du modèle utilisé ainsi que la manière de calibrer hors-ligne ou de mettre à jour en ligne chacun d'eux.

Tableau 2.1. Résumé du modèle de formation d'image

Paramètre	Symbole	Étape	Obtention (hors-ligne VS en ligne)	Données nécessaires
Pose du marqueur du tube	$T_{m \leftarrow w}$	Convertir un point 3D des coordonnées du monde vers celles du marqueur du tube	En ligne. En temps réel par le système de suivi. Section 2.4.1.	Pose du marqueur du tube identifié par le système de suivi
Transformation rigide entre le marqueur du tube et le système de la caméra	$T_{c \leftarrow m}$	Convertir un point 3D des coordonnées du marqueur du tube vers celles de la caméra	Hors-ligne. Par moyennage des transformations de différentes vues suivi d'un raffinement non linéaire. Sections 2.3.4.3 et 2.3.4.4.	Vue(s) avec la grille identifiée dans les images et la pose des marqueurs du tube et de l'objet de calibrage
Transformation des coordonnées du marqueur du tube vers celles du cercle décrit par la tête de caméra	T_C	Calculer l'angle de rotation de la tête de la caméra par rapport au tube	Hors-ligne. Par détermination du centre de rotation du cercle et construction de la transformation projetant dans le plan du cercle. Section 2.3.5.1.	Poses des marqueurs du tube et de la tête de la caméra tous deux identifiés par le système de suivi
Angle de rotation de la tête de la caméra	θ	Calculer l'angle de rotation de la tête de la caméra par rapport au tube	En ligne. En temps réel par projection de la position d'un point de la tête de la caméra par T_C . Section 2.4.2.	Pose des marqueurs du tube et de la tête de la caméra tous deux identifiés par le système de suivi
Axe de rotation	p_r et u_r	Appliquer la rotation de la tête de la caméra par rapport au tube	Hors-ligne. Par optimisation non linéaire pour la réduction de l'erreur de reprojection. Section 2.3.5.2.	Vues avec la grille identifiée dans les images et la pose des marqueurs du tube, de la tête de la caméra et de l'objet de calibrage
Distorsion	k_1 et k_2	Appliquer la distorsion de lentille	Hors-ligne. Fournis directement par OpenCV. Section 2.3.4.1.	Vue(s) avec la grille identifiée dans les images
Paramètres intrinsèques	f_x, f_y, x_0 et y_0	Convertir en coordonnées image (finales)	Hors-ligne. Fournis directement par OpenCV. Section 2.3.4.1.	Vue(s) avec la grille identifiée dans les images

2.6 Génération des vues

Le calibrage de la caméra endoscopique présenté précédemment est nécessaire à la fusion multimodale des images endoscopiques et d'une autre modalité, soit les modèles 3D préopératoires de la colonne vertébrale. Cette combinaison comporte davantage d'information que chacune des modalités prises individuellement. Il est donc important que cette mise en commun soit faite afin de représenter l'information d'une manière utile pour le chirurgien.

Afin de développer les différentes vues potentielles pour le système d'assistance chirurgicale, des discussions ont été entreprises avec le Dr Stefan Parent afin de bien cerner les besoins des chirurgiens. De plus, l'auteur a également assisté à une discectomie minimalement invasive afin d'être exposé au contexte d'utilisation du système à réaliser ainsi que pour apprécier plus directement les défis que représentent les chirurgies minimalement invasives. Suite à ces discussions, deux types de vues ont été retenues.

2.6.1 Vue #1 : réalité augmentée

Comme son nom l'indique, la réalité augmentée consiste à « augmenter » la réalité à l'aide d'information complémentaire. Dans le cadre des chirurgies minimalement invasives, la « réalité » correspond aux images endoscopiques. On peut ajouter de l'information à ces images afin d'augmenter cette réalité. On ajoute principalement des mesures de distances, telles qu'identifiées par le chirurgien, ainsi que la projection du modèle préopératoire sur l'image endoscopique.

L'ajout de l'information de la distance permet au chirurgien d'éviter d'avoir à retirer l'instrument chirurgical pour insérer une règle. Les mesures ne sont donc pas effectuées physiquement, mais plutôt grâce à la photogrammétrie, c'est-à-dire la détermination de propriétés géométriques (distances, angles, etc.) d'objets à partir d'images de ces objets. De plus, la projection du modèle préopératoire aide le chirurgien à se réorienter dans les images endoscopiques qui sont parfois très peu instinctives.

La Figure 2.12 montre un exemple de cette vue, générée à l'aide d'outils de synthèse et non en temps réel.

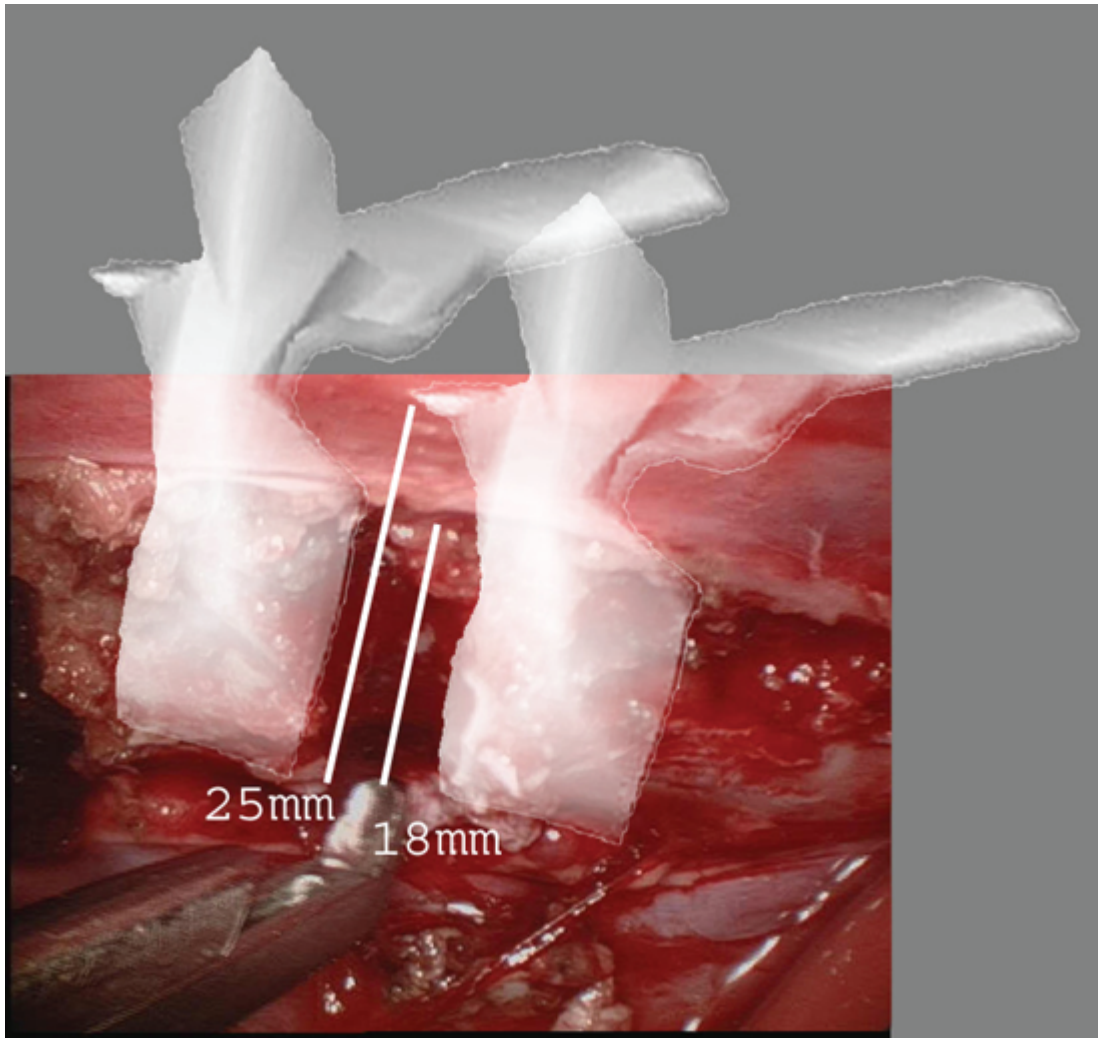


Figure 2.12. Vue en réalité augmentée (©2010 LIV4D. Reproduit avec permission)

2.6.2 Vue #2 : virtualité augmentée

La virtualité augmentée est en quelque sorte l'inverse de la réalité augmentée : au lieu de prendre des objets virtuels et de les ajouter à l'image réelle, on génère une scène complètement virtuelle à laquelle on ajoute l'image réelle.

La Figure 2.13 montre une des vues proposées de deux points de vue différents. Elle permet d'obtenir une vue d'ensemble du site opératoire et de prendre un certain recul permettant de se réorienter par rapport à l'ensemble des structures d'intérêt. Elle est principalement composée d'un rendu du modèle préopératoire de la colonne vertébrale. On ajoute également le rendu d'un cylindre représentant le tube de l'endoscope, représenté en vert, dans sa position

telle que calculée à partir des informations du système de suivi. De plus, on augmente également cette image virtuelle par l'ajout d'un plan sur lequel l'image endoscopique est plaquée. Ce plan est positionné par rapport au tube de la caméra endoscopique de manière à être aligné avec les structures imagées. Le calibrage initial de la caméra permet de localiser le plan image dans l'espace 3D et ainsi d'obtenir cette vue augmentée.

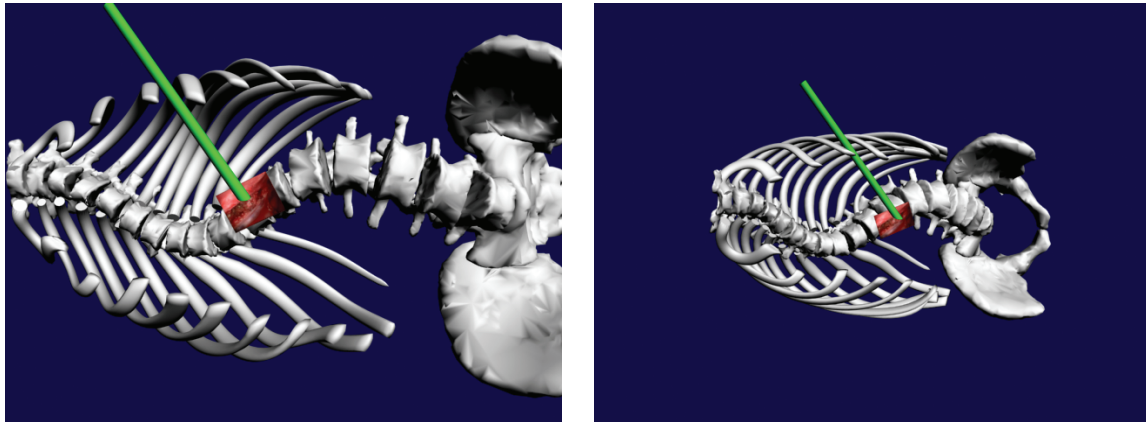


Figure 2.13. Augmentation du modèle virtuel par positionnement de l'image endoscopique

La Figure 2.14 est également une vue de virtualité augmentée. Dans cette vue, l'attention est concentrée sur le site opératoire en question. La caméra virtuelle est placée de telle sorte à être entre les plans définis par les surfaces des plateaux vertébraux. Ce point de vue permet d'apprécier l'orientation de l'instrument chirurgical, représenté en magenta, par rapport à la cavité entre les disques. La séquence d'images endoscopiques réelles est également plaquée dans le coin inférieur droit de la vue afin de permettre au chirurgien de toujours garder un œil sur l'image réelle du site opératoire. Cette possibilité de visualiser en même temps le virtuel et le réel permet au chirurgien d'apprécier la qualité de la vue en termes de précision et de délai dans le but de lui donner un meilleur niveau de confiance en l'information qu'il en tire.

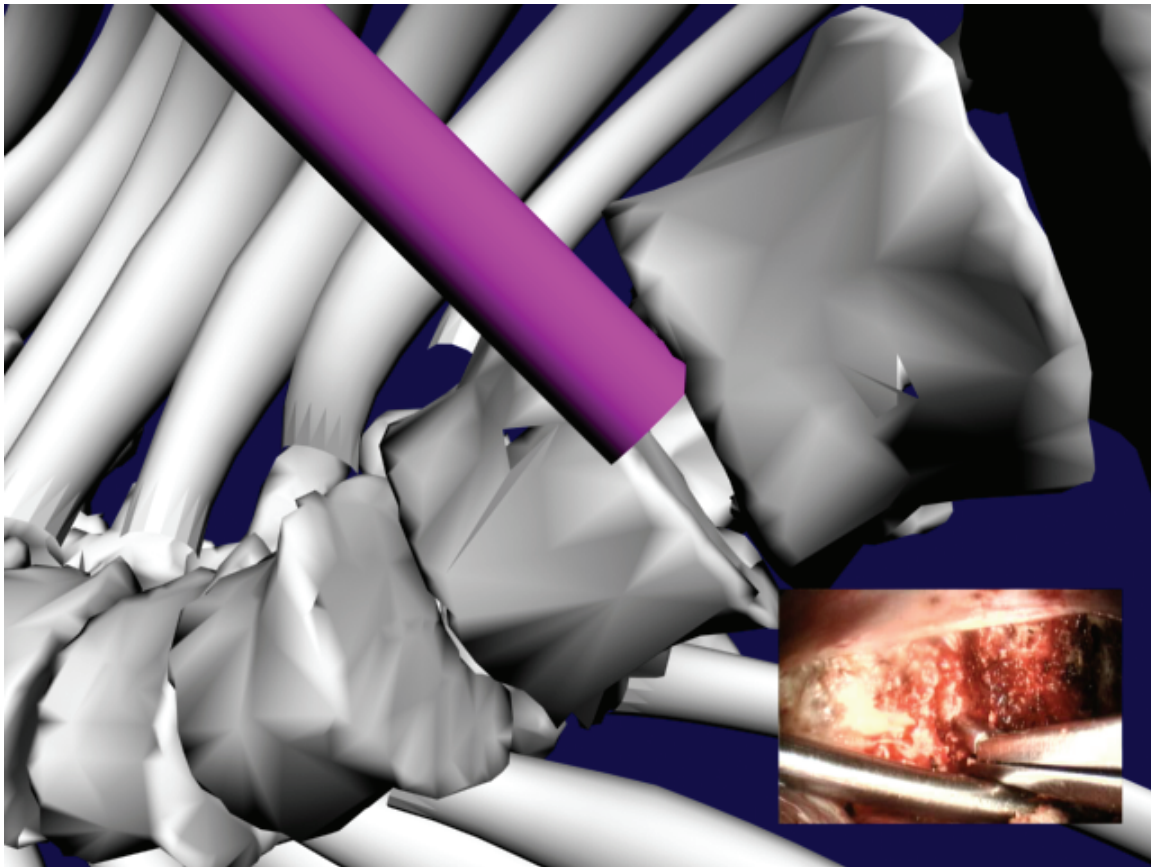


Figure 2.14. Vue augmentée avec instrument et sous-image endoscopique

Ces vues permettent donc de présenter l'information provenant de plusieurs modalités d'une manière intégrée et utile pour le chirurgien.

2.7 Validation

Un des objectifs de ce travail de recherche est de valider convenablement la méthode qui y est proposée. Cette validation est quantitative pour la méthode de calibrage, mais on se limite au niveau qualitatif pour l'évaluation des vues proposées.

L'un des outils importants pour la validation quantitative de la méthode de calibrage est l'évaluation de l'erreur de reprojection. Celle-ci permet d'évaluer si le modèle calibré permet bien de transformer une série de points 3D en leur projection 2D sur l'image. Pour chacun des points 3D, on applique les transformations du modèle calibré afin d'obtenir la position de la projection 2D du point prédite par ledit modèle. Cette position est comparée à la véritable position telle qu'identifiée dans l'image afin d'obtenir l'erreur associée à cette reprojection.

Cette erreur est moyennée pour les différents points d'une vue, et même souvent pour différentes vues, afin d'obtenir une erreur de reprojection moyenne constituant une métrique pour la précision d'un modèle et de son calibrage.

La validation quantitative se fait d'abord à l'aide de simulations effectuées sur des ensembles de données générées analytiquement, puis différents niveaux de bruit sont introduits afin de simuler les erreurs de mesure inévitables en pratique. On procède également à des expériences avec des données réelles pour valider les différentes parties du modèle. Finalement, une évaluation qualitative des deux modes de visualisation proposés est faite par un expert chirurgical.

2.7.1 Expériences de simulation

L'avantage des simulations par rapport aux expériences réelles est la grande facilité avec laquelle il est possible de générer des données pour réaliser des tests, en plus du contrôle absolu que l'on peut exercer sur ces données et de la disponibilité des valeurs acceptées (*Gold standard*), plus difficilement disponibles dans les expériences réelles. Les simulations effectuées pour valider la méthode de calibrage ont toutes la même forme : une certaine quantité de bruit est ajoutée à un ensemble de vues analytiques générées pour la simulation, un calibrage est effectué à partir de ces vues bruitées, puis l'erreur est estimée en calculant l'erreur de reprojection sur un second ensemble de vues non bruitées indépendant du premier. Ce second ensemble de vues, généré analytiquement avec les mêmes paramètres, permet de vérifier si le calibrage effectué sur le premier ensemble généralise bien à un autre ensemble de vues. En effet, on veut s'assurer que le calibrage effectué avant la chirurgie sera toujours valide pour les séquences d'images acquises après le calibrage, pendant la chirurgie.

La première simulation effectuée vise à ajouter du bruit sur l'identification de la position des coins de la grille sur les images 2D. Elle permet de s'assurer de la robustesse de la méthode de calibrage même en présence d'une certaine imprécision sur l'identification des points.

La seconde simulation ajoute plutôt du bruit sur les coordonnées des points de l'objet de calibrage. Elle simule donc une erreur dans la localisation des points 3D de l'objet de calibrage et s'assure que le calibrage peut tolérer une certaine erreur sur cette mesure.

La troisième et dernière simulation s'intéresse plutôt à l'erreur provenant du système de suivi. Elle simule l'incertitude sur l'identification de la position et de l'orientation des marqueurs afin d'évaluer son effet sur la précision du calibrage.

Les détails empiriques des conditions de réalisation de ces simulations sont donnés avec les résultats de ces dernières à la section 3.1.

2.7.2 Expériences de validation

Les expériences de validation réalisées sur des images réelles ont pour but général de s'assurer que le modèle choisi représente bien la formation des images endoscopiques et que la méthode proposée pour calibrer ce modèle est suffisamment précise. Tout comme dans le cas des expériences de simulation, deux ensembles d'images sont acquis : un pour le calibrage et un pour la validation selon l'expérience réalisée. Encore une fois, le second ensemble permet de s'assurer que les résultats ne sont pas valides uniquement pour les vues utilisées pour effectuer le calibrage, mais bien pour d'autres vues inutilisées pour faire le calibrage, toujours dans le but de vérifier que le modèle calibré peut décrire adéquatement la formation des images endoscopiques au cours de la chirurgie, après le calibrage.

La première expérience vise à évaluer que la construction de l'objet de calibrage et les mesures sur la position des points de la grille sont valides. Pour ce faire, les positions fournies par le MicronTracker et utilisées par la méthode de calibrage sont comparées avec celle d'un numériseur 3D, soit le MicroScribe que l'on peut voir à la Figure 2.15. Le modèle du MicroScribe utilisé est le G2LX, dont la précision est de 0.30 mm [81]. On compare donc la mesure de l'axe X du marqueur MicronTracker de l'objet de calibrage, c'est-à-dire la longueur du vecteur le plus long, donnée par le MicronTracker à celle donnée par le MicroScribe. On calcule également l'erreur de positionnement 3D des points assumés par la méthode de calibrage par rapport à ceux donnés par le MicroScribe. Ainsi, grâce à cette expérience, on s'assure que les mesures des positions obtenues par la construction de la grille sont comparables avec les mesures obtenues avec un numériseur 3D et donc valides. Dans le cas où la construction est suffisante pour obtenir des mesures précises, la nécessité d'utiliser un numériseur 3D pour mesurer les coordonnées des points de l'objet de calibrage avant de pouvoir l'utiliser est éliminée.



Figure 2.15. MicroScribe utilisé pour numériser l'objet de calibrage

La seconde expérience consiste à vérifier la répétabilité des résultats fournis par la méthode de calibrage. Pour ce faire, on choisit un nombre d'images à utiliser pour effectuer le calibrage. Puis, on effectue un calibrage pour différents ensembles de vues comprenant ce nombre donné d'images. Ensuite, pour chacun des ensembles, le calibrage est effectué et certaines valeurs sont retenues. Ces valeurs sont la position du centre optique, la valeur des distances focales, le rapport des distances focales, l'erreur de reprojection sur les images utilisées pour le calibrage et finalement l'erreur de reprojection sur les images de l'ensemble de validation, non utilisées pour faire le calibrage. On évalue la moyenne de chacune de ces valeurs ainsi que l'écart-type afin de quantifier la variabilité entre les différentes expériences et de s'assurer que les valeurs du modèle fournies par la méthode sont comparables peu importe l'ensemble d'images utilisé. Autrement dit, on veut s'assurer que la méthode donne les mêmes valeurs pour une même caméra.

La troisième expérience vérifie l'hypothèse que le centre de distorsion correspond au centre optique. Le processus de l'expérience est sensiblement le même que lors de la seconde expérience. Pour une multitude de sous-ensembles d'images, on calcule la position du centre

optique ainsi que la position du centre de distorsion à l'aide de la méthode présentée en annexe (voir ANNEXE V). La distance entre ces deux points est évaluée pour chacun des sous-ensembles afin de déterminer si ces deux points sont près l'un de l'autre.

La quatrième expérience vise à valider le modèle choisi pour la compensation de la rotation de la tête de la caméra par rapport au tube de l'endoscope. On compare la méthode utilisée avec la simplification proposée par Wu afin de prouver la pertinence de conserver le modèle complet de Yamaguchi. Pour ce faire, l'erreur de reprojection obtenue par l'application de chacun des modèles est comparée.

La dernière expérience vise à évaluer la précision de la reconstruction qu'il est possible d'effectuer à partir du calibrage de la caméra. Il est possible, à partir de plusieurs vues calibrées d'un même point 3D, de retrouver les coordonnées 3D de ce point à partir de sa position 2D dans chacune des images. Cette position 2D dans la i^{e} image est d'abord corrigée pour retirer les effets de la distorsion de lentille pour obtenir le point $[u_i \ v_i]^T$. La matrice M_i correspond à la matrice de projection (comprenant les paramètres intrinsèques et les paramètres extrinsèques) associée à la i^{e} vue. Chaque vue fournit donc le système d'équations suivant :

$$\lambda \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = M_i \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} m_{00}^i & m_{01}^i & m_{02}^i & m_{03}^i \\ m_{10}^i & m_{11}^i & m_{12}^i & m_{13}^i \\ m_{20}^i & m_{21}^i & m_{22}^i & m_{23}^i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.30)$$

où les seules inconnues sont les coordonnées du point 3D recherchées. Comme la troisième équation de ce système matriciel donne la valeur de λ , chaque vue fournit donc en fait deux équations utiles pour la résolution du système surdéterminé. Ainsi, à partir de deux vues, il est possible de reconstruire un point en 3D. La dernière expérience utilise ce principe pour reconstruire les deux extrémités d'un segment afin d'en déterminer la longueur et ainsi d'évaluer la précision de la mesure des distances qui peut être réalisée à partir des images endoscopiques calibrées.

Les détails empiriques des conditions de réalisation de ces expériences sont donnés avec les résultats de ces dernières à la section 3.2.

2.7.3 Validation des vues

La validation des vues proposées est faite au niveau qualitatif. Ces dernières sont présentées à un expert chirurgical, soit le Dr Stefan Parent, afin qu'il puisse juger de leur pertinence dans l'assistance chirurgicale de procédures minimalement invasives. Ce dernier fournit son appréciation ainsi que des commentaires sur les vues proposées dans le but de leur intégration dans le système final.

Les visualisations présentées sont générées à l'aide d'outils de synthèse, mais le modèle de formation d'images décrit par ce travail de recherche permet de générer ces images en temps réel lorsque le modèle préopératoire est correctement recalé.

2.7.3.1 Réalité augmentée

Pour la visualisation de réalité augmentée, après que l'image endoscopique ait été rendue en arrière-plan, l'équation 2.29 représentant le modèle de formation d'images peut être implémentée à l'aide d'une interface de programmation graphique comme OpenGL ou Direct3D pour effectuer le rendu du modèle préopératoire. Deux approches peuvent être employées pour effectuer ce rendu.

D'abord, on peut implémenter la totalité du modèle à l'aide d'un nuanceur de sommet. En effet, le modèle de formation de l'image correspond à l'étape de traitement des sommets du pipeline graphique, dans laquelle les points sont transformés de leur position 3D à leur projection 2D. Il suffit donc d'implémenter les quatre transformations à l'intérieur d'un nuanceur de sommet : la transformation #1 associée aux paramètres extrinsèques, la transformation #2 associée à la rotation relative de la tête de la caméra par rapport au tube, la transformation #3 associée à la distorsion de lentille et la transformation #4 associée à la numérisation.

Toutefois, comme le remplissage des primitives se fait après cette étape programmable de projection des sommets, les pixels ainsi remplis n'ont pas été soumis à la distorsion. Seuls les sommets subissent les transformations du modèle. En effet, la ligne reliant deux sommets dans l'espace 3D devrait être représentée par une courbe dans l'espace image; or, avec cette méthode, comme la génération de la ligne se fait après l'application de la transformation associée à la distorsion, cette ligne est dessinée droite par le système de rendu. Néanmoins,

lorsque le nombre de polygones est élevé, la distance entre chacun des points projetés est de l'ordre du pixel, si bien que cet effet est pratiquement imperceptible.

Il est possible d'éliminer cet effet à l'aide d'une seconde approche. Afin que la transformation associée à la distorsion soit appliquée aux pixels plutôt qu'aux sommets, on doit implémenter cette dernière dans le nuanceur de pixel. Pour ce faire, on implémente dans un nuanceur de sommet les transformations #1, #2 et #4. Ces transformations, sans distorsion, correspondent à la projection classique d'un point sur le plan image à l'aide d'un modèle de projection perspective; il n'est donc pas nécessaire d'avoir recours à la partie programmable du pipeline graphique puisque le pipeline standard peut se charger de ces multiplications de matrices. Le résultat de ce rendu n'est pas ajouté directement à l'image endoscopique, il est plutôt sauvegardé dans une texture temporaire. Ensuite, on implémente la distorsion à l'aide d'un nuanceur de pixel appliqué à cette texture temporaire. Ce dernier inverse tout d'abord la transformation #4 pour passer de l'espace image à celui dans lequel la distorsion doit être appliquée. Il applique par la suite la transformation #3 associée à la distorsion puis applique à nouveau la transformation #4. Ainsi, la transformation T_S implémentée par le nuanceur de sommet est :

$$T_S = T_1 T_2 T_4 \quad (2.31)$$

alors que la transformation T_P implémentée par le nuanceur de pixel est :

$$T_P = T_4^{-1} T_3 T_4 \quad (2.32)$$

Ainsi, la composée T_F de ces deux transformations, appliquée à tous les pixels plutôt qu'à tous les sommets, correspond bien au modèle de formation d'images proposé dans ce travail de recherche :

$$T_F = T_S \circ T_P = (T_1 T_2 T_4) \circ (T_4^{-1} T_3 T_4) = T_1 T_2 T_4 T_4^{-1} T_3 T_4 = T_1 T_2 T_3 T_4 \quad (2.33)$$

Il est à noter qu'il est possible de corriger la distorsion dans l'image endoscopique réelle à l'aide de la transformation inverse de celle implémentée par le nuanceur de pixel. En effet, dans le cas où l'image à laquelle on veut ajouter le modèle 3D a été corrigée par l'application de T_P^{-1} , le modèle peut être projeté sur l'image par l'utilisation du pipeline standard du matériel graphique.

Les différents paramètres du rendu, notamment ceux définissant le matériel du modèle 3D, l'éclairage ou la transparence, peuvent être ajustés pour obtenir le résultat désiré. Le rendu final du modèle n'a pas à être photoréaliste, car il est préférable pour le chirurgien de pouvoir discerner aisément les éléments réels des éléments virtuels de l'image fusionnée.

2.7.3.2 Virtualité augmentée

Pour la visualisation de virtualité augmentée, le modèle préopératoire est d'abord rendu à l'aide des méthodes classiques fournies par les interfaces de programmation graphique. Le rendu de l'image endoscopique dans son contexte se fait par l'application inverse des transformations représentant le modèle de formation de l'image.

Comme on cherche à trouver l'endroit où placer l'image endoscopique dans le monde virtuel 3D, on procède à la projection inverse des quatre coins de l'image. Chacun de ces points est d'abord reconverti dans l'espace du plan situé à une unité du centre optique par l'application de la transformation inverse #4 associée à la numérisation. Puis, la distorsion est inversée par l'application de la transformation inverse #3 associée à la distorsion. À cette étape, le point est toujours dans l'espace du plan situé à une unité du centre optique. On doit donc ici choisir la distance à laquelle le plan représentant l'image endoscopique sera affiché par rapport à la position de la caméra. Cette distance peut être choisie par l'utilisateur ou calculée en fonction de la distance de la caméra par rapport au modèle 3D.

La mise à l'échelle du plan situé à une unité de la caméra pour qu'il soit situé à la distance voulue correspond à l'inverse de la projection sur ce plan, soit la rétroprojection. Il suffit ensuite d'appliquer les transformations inverses #2 associée à la rotation relative de la tête par rapport au tube et #1 associée aux paramètres extrinsèques. Les positions des quatre coins de l'image sont donc connues en 3D dans la scène virtuelle. Il suffit donc de plaquer la texture de l'image endoscopique sur un plan passant par ces quatre points pour combiner cette image endoscopique réelle à la scène virtuelle. Cette image doit avoir été corrigée pour éliminer l'effet de la distorsion à l'aide de l'inverse de la transformation T_P représentée par l'équation 2.32.

CHAPITRE 3 RÉSULTATS

Ce chapitre présente les résultats obtenus par l'implémentation de la méthode proposée au chapitre précédent. Il présente d'abord les résultats obtenus pour les différentes simulations proposées, puis ceux des expériences de validation réalisées. Ensuite, il présente la validation qualitative qui a été effectuée pour les deux modes de visualisation proposés.

3.1 Expériences de simulation

Les différentes simulations ont été réalisées en C++ dans un environnement multi-plateforme Windows/Linux. Pour effectuer les simulations, des vues synthétiques ont été générées. Elles représentent la grille de l'objet de calibrage imagée par la caméra endoscopique positionnée à différents points de vue. Afin de réaliser des simulations réalistes, les différentes vues synthétiques générées ne peuvent être complètement aléatoires. Elles doivent être contraintes de manière à représenter des vues qu'il est possible d'obtenir lors d'un calibrage réel.

Lors de la génération des vues, la configuration choisie pour l'objet de calibrage simulé est conforme à celle de l'objet réel fabriqué pour ce travail de recherche, c'est-à-dire qu'il possède une grille de 9 points par 9 points dont le point le plus près est situé à 8 unités de grille de l'axe X du marqueur posé sur l'objet de calibrage. La longueur de l'axe X du marqueur de l'objet de calibrage, longueur qui définit la taille de l'objet, est une variable aléatoire uniformément distribuée entre 40 mm et 60 mm.

Ensuite, les paramètres intrinsèques sont calculés. Le volume de vision est calculé à partir de l'angle du champ de vision vertical choisi, soit 60 degrés. Les limites supérieure, inférieure, gauche et droite de la pyramide tronquée de vision peuvent ensuite être bruitées. Les distances focales ainsi que les coordonnées du centre optique sont ensuite calculées à partir du volume de vision.

Un des paramètres utilisés pour calculer les paramètres extrinsèques s'apparente davantage aux paramètres intrinsèques puisqu'il est propre à la caméra et ne varie pas lorsque la caméra est déplacée. Ce paramètre est la transformation $T_{c \leftarrow m}$ entre le repère du marqueur posé sur le tube et le repère de la caméra. Cette transformation rigide est générée en respectant la distance entre les deux repères que l'on choisit afin de respecter approximativement la configuration utilisée. Dans les simulations, cette distance est fixée à 30 cm.

Les paramètres intrinsèques ainsi déterminés restent les mêmes pour tout l'ensemble de vues générées. Par contre, les paramètres extrinsèques doivent être recalculés pour chaque vue, en obtenant de nouvelles valeurs pour les variables aléatoires utilisées, ce qui correspond à un point de vue différent pour la même caméra. Une position aléatoire pour la grille de calibrage est d'abord générée de manière à ce que son marqueur soit situé à l'intérieur du volume de vision du MicronTracker. Le volume de vision du MicronTracker est défini comme un prisme rectangulaire (plutôt qu'une pyramide tronquée) allant 5 m à la gauche, à la droite, en haut, en bas et devant la caméra du MicronTracker, donnant un espace possible pour les marqueurs générés de 10 m x 10 m x 5 m, ce qui est encore plus grand que l'espace réel du MicronTracker. La position du marqueur de l'objet de calibrage est une variable aléatoire uniformément distribuée à l'intérieur de ce volume.

La position de la caméra est ensuite générée de manière à ce qu'elle puisse voir toute la grille. L'angle entre l'axe optique de la caméra et la normale au plan de la grille de calibrage est une variable aléatoire uniformément distribuée entre 30 degrés et 60 degrés. Cette restriction est réaliste, car dans un contexte réel, les vues où la grille est complètement perpendiculaire ou trop angulée par rapport à la caméra sont évitées. La distance de la caméra par rapport à l'objet de calibrage ainsi que son orientation sont déterminées de manière à ce que toute la grille soit visible par la caméra simulée. Une fois la position et l'orientation de la caméra déterminée, la transformation $T_{c \leftarrow m}$ permet de calculer la position du marqueur situé sur le tube de l'endoscope.

Ainsi, tous les paramètres décrivant la configuration physique de la caméra et de la grille ayant été déterminés, on produit une vue synthétique en projetant les points de l'objet de calibrage sur le plan image à l'aide du modèle de formation présenté à la section 2.2.

Pour chaque simulation décrite dans les sous-sections suivantes, 40 vues synthétiques ont été générées. De ces 40 vues, 20 sont utilisées pour faire la simulation, c'est-à-dire qu'on leur ajoute un certain niveau de bruit correspondant aux paramètres à vérifier. Les 20 autres sont utilisées pour calculer l'effet de ce bruit sur la précision du calibrage, en calculant l'erreur de reprojection sur les 20 images du second ensemble à partir du modèle calibré sur les 20 images du premier ensemble bruité. Un exemple de vue générée est présenté à la Figure 3.1 et la totalité des vues est présentée en annexe (voir ANNEXE VI).

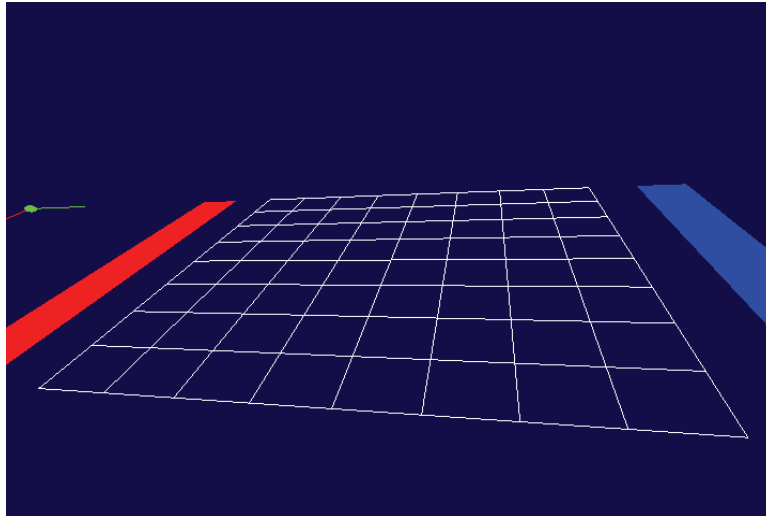


Figure 3.1. Exemple de vue générée pour les simulations

Pour ces vues générées, on connaît avec précision les valeurs du modèle de projection, les coordonnées 3D des points de l'objet de calibrage et par conséquent les coordonnées 2D de ces points projetés sur le plan image.

Il est à noter que ces simulations considèrent un bruit gaussien parfaitement symétrique autour des données bruitées. Elles ne simulent donc pas l'effet d'un biais, c'est-à-dire une erreur qui tend à pousser les données vers une certaine direction.

3.1.1 Simulation #1 : bruit dans les images

Au cours de cette simulation, on déplace les coordonnées 2D des points projetés sur le plan image afin de simuler l'erreur d'identification des points. Le bruit ajouté est un bruit gaussien dont on augmente progressivement l'écart-type afin de constater l'effet de l'augmentation du bruit. Pour chacune des valeurs d'écart-type considérées, c'est-à-dire de 0 pixel à 10 pixels par saut de 0,5 pixel, on effectue 1000 simulations. Chaque simulation consiste à perturber chaque point 2D du premier ensemble d'images à l'aide d'un bruit gaussien d'un écart-type donné, puis un calibrage du modèle est effectué à partir de cet ensemble pour finalement évaluer l'erreur de reprojection moyenne sur le second ensemble, non bruité et indépendant.

La méthode automatique d'identification des points possède une précision en-dessous du pixel et les faibles imprécisions qui en résultent sont bien inférieures à 10 pixels. Ainsi, un niveau de bruit simulé allant jusqu'à 10 pixels d'écart-type est amplement suffisant.

Le résultat de cette simulation est représenté à la Figure 3.2. L'erreur augmente pratiquement proportionnellement au niveau du bruit. On remarque que pour une erreur d'identification allant jusqu'à 10 pixels d'écart-type, l'erreur de reprojection moyenne ne dépasse pas 0,613 pixel. La méthode est donc robuste en ce qui a trait à l'erreur d'identification des coins de la grille. On peut expliquer ce résultat par l'utilisation d'un grand nombre de points, c'est-à-dire 20 vues avec 81 points par vue. Comme le bruit appliqué à chacun des points est indépendant, le phénomène de « moyennage » créé par l'utilisation d'un grand nombre de points permet d'obtenir un calibrage précis malgré la présence de bruit.

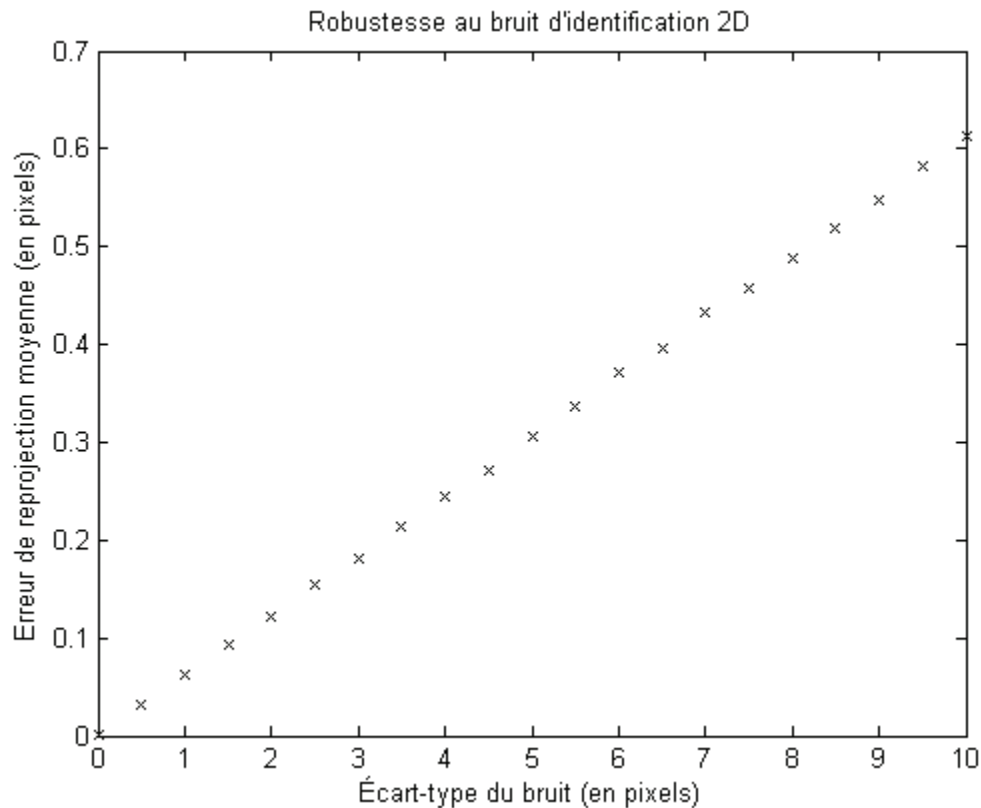


Figure 3.2. Graphique de l'erreur de reprojection en fonction de l'écart-type du bruit 2D ajouté

3.1.2 Simulation #2 : bruit sur l'objet de calibrage

Au cours de cette simulation, on ajoute du bruit aux coordonnées 3D des points de l'objet de calibrage afin de simuler l'erreur de localisation de ces points due à la construction de l'objet de calibrage. Le bruit ajouté est encore un bruit gaussien dont on augmente progressivement l'écart-type. Pour chacune des valeurs d'écart-type considérées, c'est-à-dire de 0 mm à 3 mm

par saut de 0,1 mm, on effectue 1000 simulations. Chaque simulation consiste à perturber chaque point 3D de l'objet de calibrage utilisé pour effectuer le calibrage à partir du premier ensemble d'images, pour ensuite évaluer l'erreur de reprojection moyenne de ce calibrage sur le second ensemble.

Le résultat de cette simulation est représenté à la Figure 3.3. Le résultat obtenu est comparable à celui de la première simulation. L'erreur augmente également d'une manière pratiquement proportionnelle au niveau de bruit ajouté : la meilleure droite passant par l'origine décrivant cette relation possède une pente de 0,8058 pixel/mm (coefficient de détermination de 0,9973). L'erreur de reprojection ne dépasse pas 2,52 pixels pour un niveau de bruit considérable de 3,0 mm d'écart-type. Encore une fois, l'utilisation d'un grand nombre de points aide à diminuer la propagation de l'erreur. Il est toutefois pertinent de s'assurer que le niveau de précision de la localisation des points de l'objet de calibrage réel (ici simulé jusqu'à 3,0 mm) reste dans des limites acceptables, ce qui est validé plus loin à la section 3.2.1. L'expérience #1 permettra en effet d'évaluer la précision de cette localisation, à laquelle il est désormais possible d'associer une erreur de reprojection tel que représenté à la Figure 3.3. Par exemple, si cette précision est de l'ordre de 1,0 mm, l'erreur de reprojection associée à cette imprécision est de moins de 1 pixel. Cette validation confirmera que le niveau de confiance sur les positions des points de l'objet de calibrage est suffisant pour considérer la méthode de calibrage robuste par rapport à ce bruit.

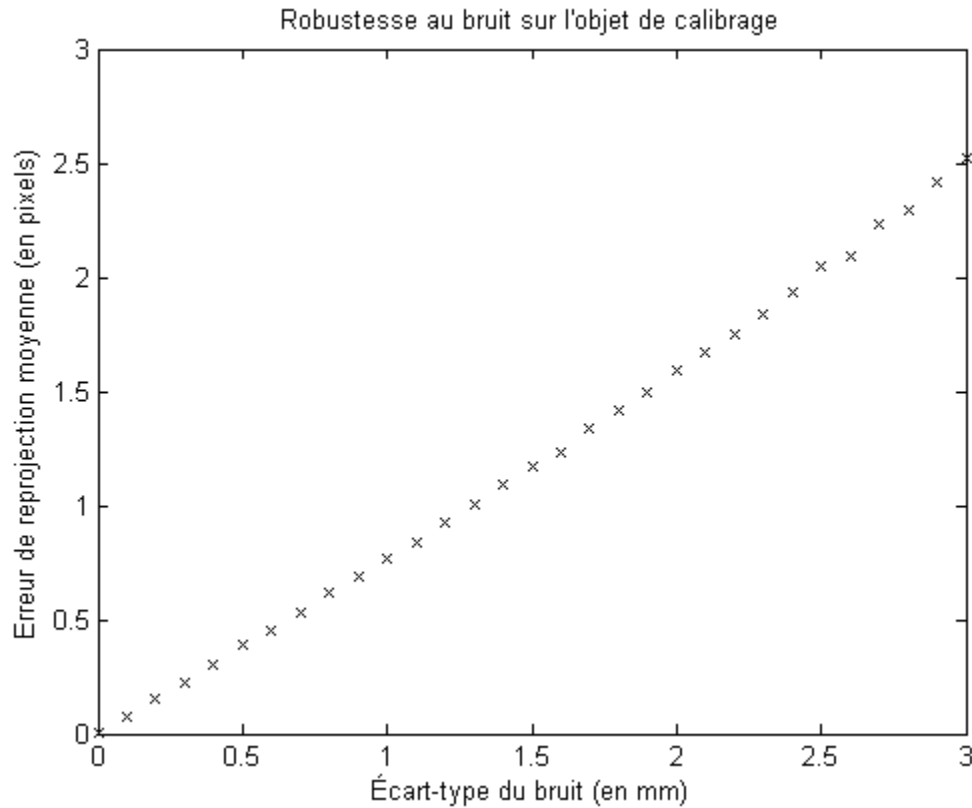


Figure 3.3. Graphique de l'erreur de reprojection en fonction de l'écart-type du bruit 3D ajouté

3.1.3 Simulation #3 : bruit sur le suivi

Au cours de cette simulation, on ajoute un bruit sur les poses des marqueurs telles qu'elles auraient été fournies par le système de suivi. Le niveau de bruit ajouté correspond aux spécifications du MicronTracker [82]. Ces dernières mentionnent deux niveaux : un niveau de précision du calibrage, soit 0,20 mm, qui correspond à l'erreur RMS de l'identification d'un seul *Xpoint*; et un niveau de précision d'une cible statique, soit 0,015 mm, qui correspond également à l'erreur RMS d'un seul *Xpoint*, celui-ci étant toutefois maintenu statique permettant ainsi aux algorithmes d'élimination du bruit d'être appliqués.

La simulation consiste donc à ajouter un bruit gaussien 3D à chacun des points définissant les marqueurs. Chaque marqueur possède trois points permettant d'identifier la pose. Ces points sont bruités, puis une nouvelle pose est calculée. Ces expériences ont été réalisées 1000 fois pour chacun des niveaux de bruit allant de 0,0 mm (bruit nul) à 1,0 mm d'écart-type par incrément de 0,005 mm. Le résultat de cette simulation est représenté sur la Figure 3.4.

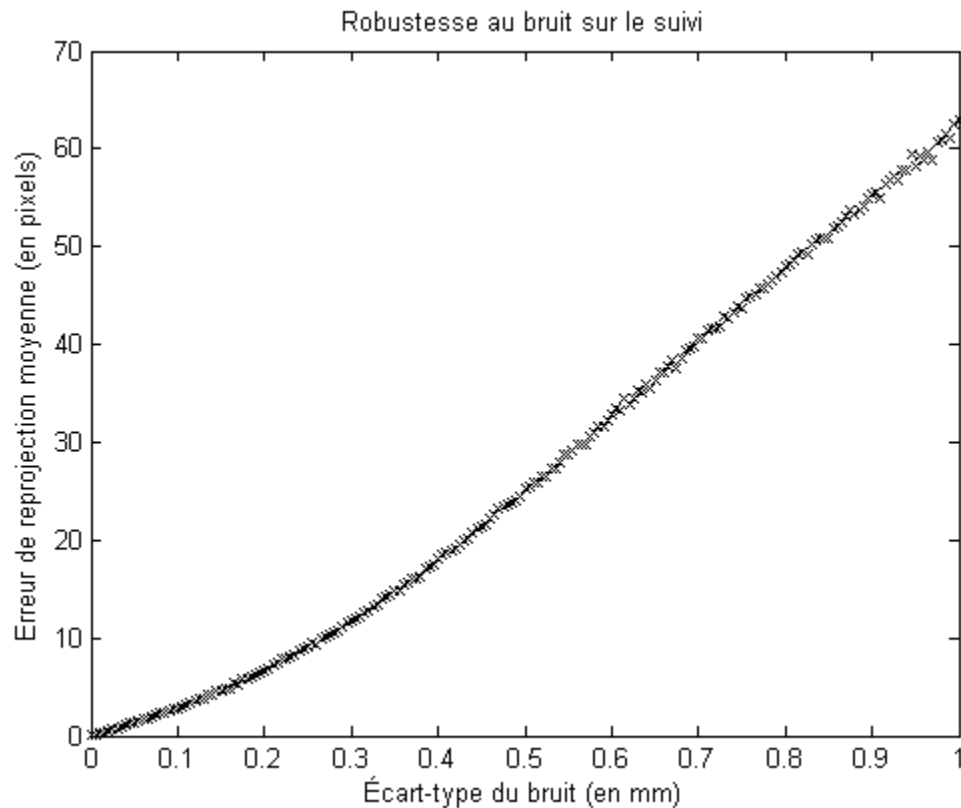


Figure 3.4. Graphique de l'erreur de reprojection en fonction de l'écart-type du bruit ajouté aux points de marqueurs

Le tableau suivant représente la valeur numérique des erreurs de reprojection associées aux niveaux de précision mentionnés par le fabricant.

Tableau 3.1. Résultats de la simulation #3

Niveau de bruit (en mm)	Erreur de reprojection moyenne (en pixels)
0,0	$6,23 \times 10^{-11}$
0,015	0,394
0,20	6,63

On voit donc que le niveau de précision obtenu pour le calibrage dépend énormément de la précision du système de suivi. En effet, si la précision du système de suivi respecte les spécifications du fabricant, on remarque que le calibrage est beaucoup plus précis lorsque les marqueurs sont statiques. Ainsi, un calibrage hors-ligne fait « à la main », c'est-à-dire en tenant manuellement la caméra endoscopique, ce qui induit quand même un certain

mouvement de cette dernière et donc des marqueurs, ne sera pas aussi précis qu'un fait avec la caméra endoscopique fixe. Pour cette raison, les expériences de calibrage hors-ligne réel réalisées dans le cadre de ce travail de recherche ont été effectuées à l'aide d'un bras ManFrotto maintenant la caméra endoscopique fixe.

De plus, l'erreur causée par le bruit sur le suivi constitue la source d'erreur la plus importante que les simulations ont permis d'identifier.

3.2 Expériences de validation sur des images endoscopiques réelles

Les expériences de validation ont été réalisées à l'aide de véritables images acquises à l'aide de la caméra endoscopique. À l'instar des ensembles de données utilisés pour les simulations, 40 images de l'objet de calibrage ont été acquises, puis séparées en deux groupes : le groupe de calibrage, utilisé pour calibrer le modèle de la caméra selon les critères de l'expérience, ainsi qu'un groupe de validation, utilisé pour vérifier la précision du modèle sur des images non utilisées pour effectuer le calibrage.

L'objet de calibrage choisi est doté d'une grille de 10 carreaux par 10 carreaux, fournissant ainsi 9 points par 9 points au calibrage. L'objet de calibrage et la caméra ont tous deux été déplacés entre les différentes vues.

3.2.1 Expérience #1 : construction de l'objet de calibrage

Au cours de cette expérience, les positions des points de deux objets de calibrage ont été mesurées. Chacun des points a été mesuré deux fois, et la totalité des mesures a été faite à deux reprises pour chaque objet.

La mesure de l'axe X du modèle utilisé pour effectuer le calibrage est celle qui est donnée par le MicronTracker et enregistrée dans le patron du marqueur de chacun des objets de calibrage. La mesure de l'axe X fournie par le MicroScribe correspond à la différence de position entre les deux points tel qu'identifiés par le MicroScribe. On effectue également un recalage rigide entre les points utilisés par la méthode (obtenus selon la structure de la construction de l'objet de calibrage) et ceux du MicroScribe afin de comparer l'erreur résiduelle sur la position 3D des points. Le tableau suivant résume les résultats.

Tableau 3.2. Comparaison des objets de calibrage

Objet	Mesure de l'axe X (en mm)	Différence (en mm)	Erreur de recalage (en mm)
Objet #1 (MicronTracker)	42,7	—	—
Objet #1 (MicroScribe #1)	42,4	0,3	0,332
Objet #1 (MicroScribe #2)	42,8	0,1	0,0762
Objet #2 (MicronTracker)	64,0	—	—
Objet #2 (MicroScribe #1)	63,8	0,2	0,195
Objet #2 (MicroScribe #2)	64,0	0,0	0,00792

On voit donc que les erreurs sont acceptables et tombent sous la limite de la précision du MicroScribe. Il est intéressant de noter que la simulation faite à la section 3.1.2 montre que les niveaux d'erreurs obtenus sur le positionnement 3D des points de l'objet de calibrage résultent en une erreur de reprojection bien inférieure à 1 pixel. La méthode de détermination des positions 3D des points de l'objet de calibrage à partir de la construction de ce dernier est donc validée.

3.2.2 Expérience #2 : répétabilité des résultats

Le but de cette expérience est de vérifier la répétabilité des résultats. On veut s'assurer qu'un calibrage hors-ligne fait sur un ensemble de vues provenant d'une caméra donnée fournira les mêmes valeurs de paramètres qu'un second calibrage hors-ligne fait sur un ensemble de vues différentes des premières, mais obtenues avec la même caméra.

Pour effectuer cette validation, on s'intéresse à la variabilité des paramètres obtenus pour la même caméra, mais à partir de vues différentes. L'écart-type est donc un indicateur intéressant de cette variabilité. Idéalement, on acquerrait un jeu de vues pour chaque calibrage à effectuer et on répéterait cette acquisition jusqu'à ce qu'on ait un nombre de calibrages suffisant pour que l'écart-type soit statistiquement intéressant. Cependant, un tel exercice représenterait un très grand nombre de vues à acquérir.

Afin d'obtenir des ensembles de vues différents permettant ainsi l'évaluation de la variabilité tout en nécessitant un nombre raisonnable de vues à acquérir, on choisit plutôt des sous-

ensembles d'un large ensemble de vues. Par exemple, à partir d'un ensemble de 20 vues, il est possible de former 1140 ensembles différents de 3 vues. Ce compromis permet donc, à partir d'un nombre raisonnable de vues, d'obtenir un nombre intéressant d'ensembles de vues, différents les uns des autres, à utiliser pour le calibrage.

Les tableaux suivants résument les résultats de l'expérience #2. Ils montrent, pour chaque nombre d'images utilisées pour former les sous-ensembles, les différentes mesures obtenues pour les calibrages effectués avec toutes les combinaisons possibles d'images.

Tableau 3.3. Résultats partiels de l'expérience #2

Nombre d'images	Nombre de combinaison	Nombre de calibrage	f_x (en pixels)	f_y (en pixels)	f_x / f_y
1	20	20	732,425 (29,0)	667,376 (27,9)	0,911176 (0,0100)
2	190	190	734,180 (6,77)	668,367 (7,54)	0,910344 (0,00310)
3	1 140	1 140	732,983 (4,43)	667,133 (4,68)	0,910159 (0,00204)
4	4 845	4 845	732,593 (3,46)	666,739 (3,52)	0,910107 (0,00158)
5	15 504	15 504	732,423 (2,89)	666,570 (2,88)	0,910089 (0,00131)
6	38 760	38 759	732,332 (2,50)	666,482 (2,46)	0,910081 (0,00113)
7	77 520	77 510	732,278 (2,21)	666,430 (2,15)	0,910077 (0,000993)
8	125 970	125 940	732,242 (1,97)	666,395 (1,91)	0,910075 (0,000884)
9	167 960	167 882	732,217 (1,77)	666,371 (1,71)	0,910073 (0,000794)
10	184 756	184 658	732,197 (1,59)	666,352 (1,53)	0,910072 (0,000715)
11	167 960	167 824	732,182 (1,43)	666,338 (1,38)	0,910071 (0,000644)
12	125 970	125 870	732,170 (1,29)	666,326 (1,24)	0,910070 (0,000580)
13	77 520	77 435	732,159 (1,15)	666,316 (1,11)	0,910070 (0,000520)
14	38 760	38 715	732,151 (1,03)	666,307 (0,985)	0,910069 (0,000463)
15	15 504	15 489	732,143 (0,904)	666,300 (0,866)	0,910068 (0,000408)
16	4 845	4 722	732,137 (0,783)	666,296 (0,750)	0,910069 (0,000354)
17	1 140	1 108	732,135 (0,655)	666,292 (0,627)	0,910066 (0,000299)
18	190	185	732,123 (0,523)	666,281 (0,501)	0,910066 (0,000236)
19	20	20	732,121 (0,366)	666,279 (0,350)	0,910066 (0,000166)
20	1	1	732,117 (0)	666,275 (0)	0,910066 (0)

La première colonne du Tableau 3.3 représente le nombre d'images utilisées dans les sous-ensembles. La seconde colonne représente le nombre théorique de combinaisons, soit :

$$\frac{n!}{k!(n-k)!} \quad (3.1)$$

avec $n = 20$ et k correspondant au nombre d'images. Toutefois, en pratique, certaines combinaisons entraînaient une assertion dans le code de OpenCV. Cette assertion correspond en fait à l'échec d'une décomposition en valeurs singulières (implémentée par Lapack) au cours de la procédure de calibrage. Afin de contourner ce bogue interne à OpenCV, les quelques cas problématiques ont été ignorés, laissant un nombre amplement suffisant de combinaisons pour obtenir des moyennes significatives. La troisième colonne rapporte le nombre de combinaisons véritablement utilisées. Les quatrième et cinquième colonnes représentent la valeur moyenne de la distance focale obtenue, respectivement en x et en y , alors que la dernière colonne représente leur rapport. Pour chacune de ces valeurs, l'écart-type de la mesure, qui est la valeur d'intérêt pour évaluer la variabilité des résultats, est exprimé entre parenthèse.

Tableau 3.4. Résultats partiels de l'expérience #2

Nombre d'images	u_0 (en pixel)	v_0 (en pixel)	Erreur de reprojection sur calibrage (en pixel)	Erreur de reprojection sur validation (en pixel)
1	391,832 (5,58)	276,078 (3,57)	0,202781 (0,0240)	29,3836 (5,32)
2	393,040 (2,69)	276,557 (2,67)	10,681 (5,37)	26,8640 (3,00)
3	393,126 (2,04)	276,591 (2,16)	12,4809 (4,32)	26,0037 (2,16)
4	393,147 (1,68)	276,608 (1,82)	13,3391 (3,64)	25,5398 (1,71)
5	393,153 (1,44)	276,620 (1,58)	13,8506 (3,15)	25,2489 (1,42)
6	393,153 (1,27)	276,628 (1,39)	14,1935 (2,77)	25,0481 (1,22)
7	393,152 (1,13)	276,635 (1,23)	14,4398 (2,46)	24,9023 (1,06)
8	393,15 (1,01)	276,640 (1,11)	14,6265 (2,21)	24,7907 (0,940)
9	393,148 (0,908)	276,644 (0,997)	14,7724 (1,99)	24,7028 (0,839)
10	393,146 (0,819)	276,648 (0,900)	14,8907 (1,79)	24,6314 (0,751)
11	393,144 (0,740)	276,650 (0,813)	14,9880 (1,62)	24,5724 (0,674)
12	393,142 (0,667)	276,652 (0,733)	15,0697 (1,46)	24,5229 (0,604)
13	393,141 (0,598)	276,654 (0,658)	15,1383 (1,31)	24,4809 (0,540)
14	393,139 (0,533)	276,656 (0,586)	15,1984 (1,17)	24,4444 (0,480)
15	393,138 (0,469)	276,657 (0,516)	15,2509 (1,023)	24,4136 (0,422)
16	393,137 (0,406)	276,660 (0,447)	15,2934 (0,887)	24,3847 (0,363)
17	393,135 (0,342)	276,665 (0,373)	15,3352 (0,745)	24,3611 (0,304)
18	393,135 (0,274)	276,664 (0,297)	15,3663 (0,596)	24,3412 (0,242)
19	393,135 (0,191)	276,661 (0,210)	15,4046 (0,416)	24,3203 (0,167)
20	393,135 (0)	276,661 (0)	15,4318 (0)	24,2952 (0)

Le Tableau 3.4 indique la suite des résultats dans un second tableau, pour des raisons d'espace. Il présente les coordonnées du centre optique ainsi que l'erreur moyenne de reprojection à la fois sur l'ensemble d'images utilisées pour le calibrage (colonne 4) ainsi que sur l'ensemble de validation (colonne 5).

On remarque que la moyenne des différents paramètres calibrés (distances focales et coordonnées du centre optique) est très comparable peu importe le nombre d'images utilisées. En effet, comme le calibrage est fait sur tous les sous-ensembles possibles, on s'attend à ce que la moyenne reste sensiblement la même. Toutefois, le point intéressant à étudier avec ce type d'expérience est la variabilité des résultats du calibrage, d'où la pertinence d'inclure l'écart-type en plus de la moyenne. On remarque que cet écart-type diminue en fonction du nombre d'images. En effet, un plus grand nombre d'images permet d'inclure plus d'information et est donc moins sujet au bruit ou aux erreurs présentes dans chaque vue individuelle. L'augmentation du nombre d'images utilisées pour le calibrage accroît donc la robustesse du calibrage. Toutefois, on remarque qu'à partir de 5 images, la réduction de la variabilité des résultats n'est plus très significative et ne vaut donc plus la peine d'ajouter des images supplémentaires. Les résultats de la présente expérience suggèrent donc d'utiliser 5 vues pour effectuer le calibrage, même si la méthode proposée fonctionne avec un nombre plus élevé ou plus faible de vues.

Il est intéressant de noter l'augmentation de l'erreur de reprojection entre l'utilisation d'une seule vue et celle de deux vues. En effet, elle passe de 0,202 781 pixel pour une vue à 10,681 pixels pour deux vues. Cette augmentation de l'erreur de reprojection est due au fait que la précision du système de suivi n'entre pas en ligne de compte pour une seule vue, alors qu'elle le fait pour plusieurs vues. Effectivement, lorsque le calibrage hors-ligne est fait à partir d'une seule vue, la transformation rigide $T_{c \leftarrow m}$ entre le repère du marqueur et celui de la caméra est calculée à partir des paramètres extrinsèques de cette vue et des poses des marqueurs fournies par le système de suivi grâce à l'équation 2.20. Lorsque l'erreur de reprojection est calculée sur cette même image, les paramètres extrinsèques sont retrouvés en effectuant exactement l'opération inverse, si bien que les paramètres extrinsèques résultants seront les mêmes que ceux fournis par la procédure de calibrage, peu importe les poses des marqueurs. Ainsi, l'erreur est pratiquement nulle, car le calibrage a été réalisé de manière à trouver les paramètres qui minimisent cette erreur sur cette unique image. Or, lorsque deux

images ou plus sont utilisées, la transformation $T_{c \leftarrow m}$ est calculée à partir d'une moyenne des transformations provenant de chacune des vues. Ainsi, lorsque le calcul de l'erreur de reprojection est effectué, les paramètres extrinsèques ne sont plus obtenus directement par une inversion qui annule l'effet de la pose des marqueurs. L'incertitude provenant du système de suivi entre donc en ligne de compte, ce qui explique l'augmentation marquée de l'erreur de reprojection sur les vues utilisées lors du calibrage lorsque le nombre de vues utilisées passe de une à deux.

Cette augmentation d'environ 10 pixels due à l'introduction du système de suivi comme source d'imprécision suggère que la précision du système de suivi est inférieure à celle annoncée par le fabricant. En effet, selon les résultats de la simulation #3 représentés à la Figure 3.4, une erreur de reprojection de 10 pixels correspond à une erreur d'environ 0,3 mm sur la localisation des points du marqueur. Il est toutefois courant d'obtenir des résultats moins précis que ceux du fabricant, car ce dernier obtient souvent ses résultats dans des conditions particulières afin d'obtenir les meilleurs possibles de manière à présenter le produit sous son meilleur jour.

Les résultats obtenus pour l'erreur de reprojection sur l'ensemble de validation méritent une attention particulière. En effet, l'erreur obtenue semble tendre vers une valeur assez élevée.

Comme la valeur moyenne des paramètres ne varie pas beaucoup selon le nombre d'images utilisées pour le calibrage, on effectue cette partie supplémentaire de l'analyse des résultats en utilisant le calibrage à partir des 20 images de l'ensemble de calibrage. Puis, afin d'identifier la source de l'erreur, on observe l'erreur de reprojection sur chacune des images, plutôt que la moyenne des images de l'ensemble de validation. La meilleure image possède une erreur moyenne de 3,89 pixels alors que la pire image possède une erreur moyenne de 95,1 pixels (l'écart-type sur les 20 images est de 21,1 pixels). On remarque que l'erreur sur chacune des images varie beaucoup. La Figure 3.5 montre différents cas d'erreur de reprojection à l'aide d'une ligne reliant chaque point de la grille au point correspondant reprojeté sur l'image à l'aide du modèle calibré.

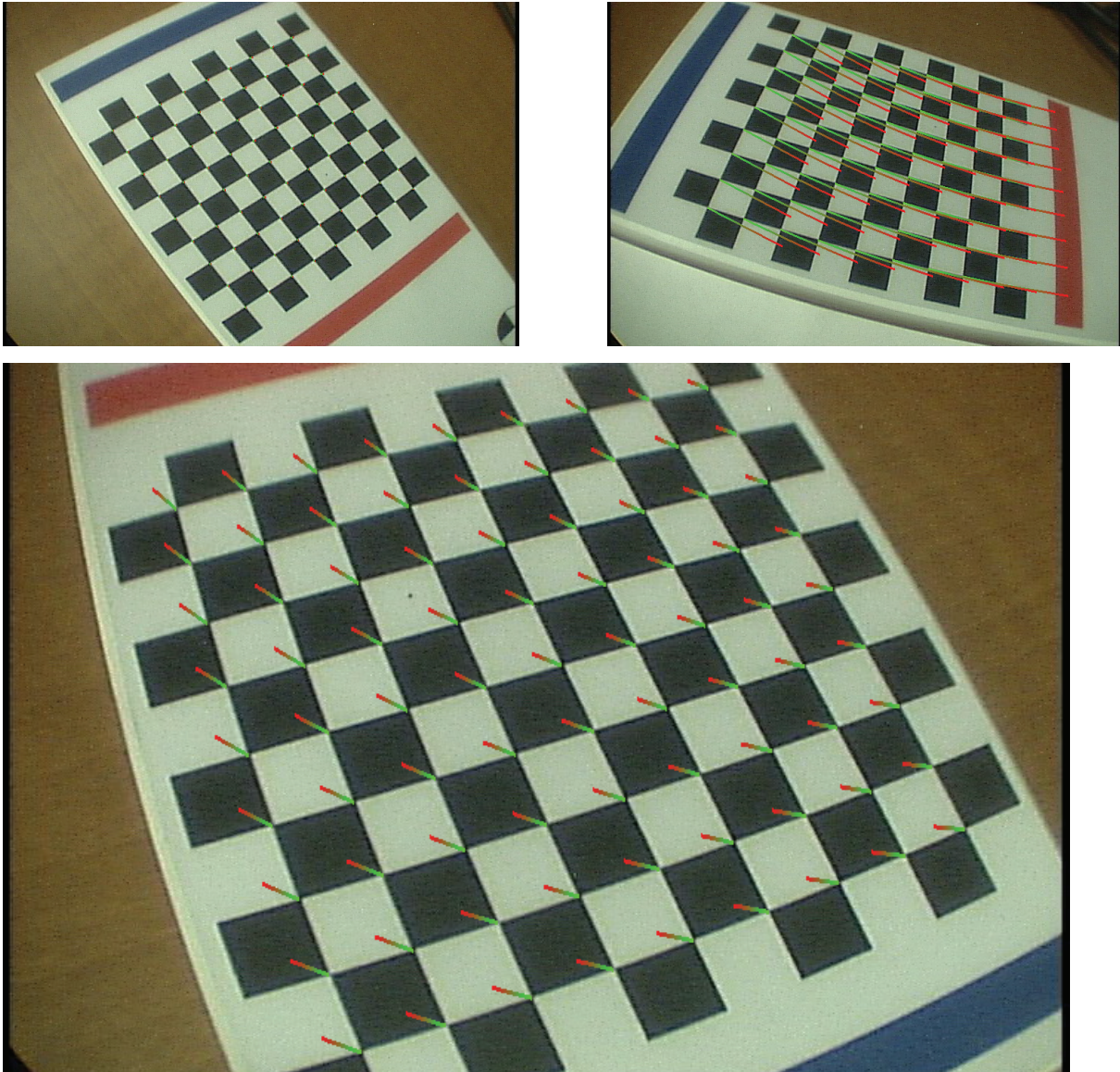


Figure 3.5. Erreur de reprojection dans le meilleur cas (haut à gauche), le pire cas (haut à droite) et le cas moyen (bas)

On remarque donc que l'erreur n'est pas aléatoirement distribuée autour des points; elle est plutôt orientée dans une direction qui a une certaine cohérence pour une image donnée. Ce phénomène est expliqué par le fait que la plus grande partie de l'erreur est due à une mauvaise estimation des paramètres extrinsèques, qui résulte donc en une estimation erronée de la position et de l'orientation de la caméra, qui résulte à son tour en un décalage cohérent de tous les points reprojetés. Les configurations des marqueurs du MicronTracker sont à l'origine des erreurs de grande amplitude : en effet, lorsque le marqueur est dans une configuration

délicate, c'est-à-dire lorsque le plan du marqueur est de plus en plus parallèle à l'axe des caméras du MicronTracker, une légère erreur dans l'identification 2D des points du marqueur par le système MicronTracker résulte en une erreur amplifiée pour la position 3D des points, et donc de la pose du marqueur, ce qui résulte en une erreur encore une fois amplifiée (par la distance entre le marqueur et le bout de l'endoscope) pour l'estimation des paramètres extrinsèques de la caméra.

Dans le même ordre d'idée, lorsque la reprojection est effectuée en temps réel sur l'image endoscopique, on observe visuellement un tremblement non négligeable des points reprojetés dû aux légères variations de la pose des marqueurs fournie par le système de suivi. Il est donc important de rester conscient de la sensibilité du système de suivi aux différentes configurations des marqueurs.

Finalement, il est important de mentionner que la méthode utilisée pour obtenir un nombre suffisant d'ensembles de vues à partir d'un nombre raisonnable de vues entraîne des conséquences sur les résultats. En effet, comme les sous-ensembles de vues comparés possèdent des vues en commun, la variabilité observée est probablement légèrement moindre que celle observée lorsque toutes les vues sont différentes. Toutefois, il s'agit d'un compromis qui permet tout de même de valider la répétabilité des résultats à partir d'un nombre raisonnable de vues.

3.2.3 Expérience #3 : position du centre de distorsion

Le protocole expérimental pour l'expérience #3 est sensiblement le même que pour l'expérience #2, à l'exception que l'on s'intéresse désormais à la comparaison entre la position du centre optique par rapport à celle du centre de distorsion obtenue à l'aide de la méthode présentée en annexe (voir ANNEXE V). Donc, pour chacun des sous-ensembles, on évalue la distance entre le centre optique obtenu par calibrage et le centre de distorsion. La Figure 3.6 montre la distance moyenne obtenue pour les sous-ensembles.

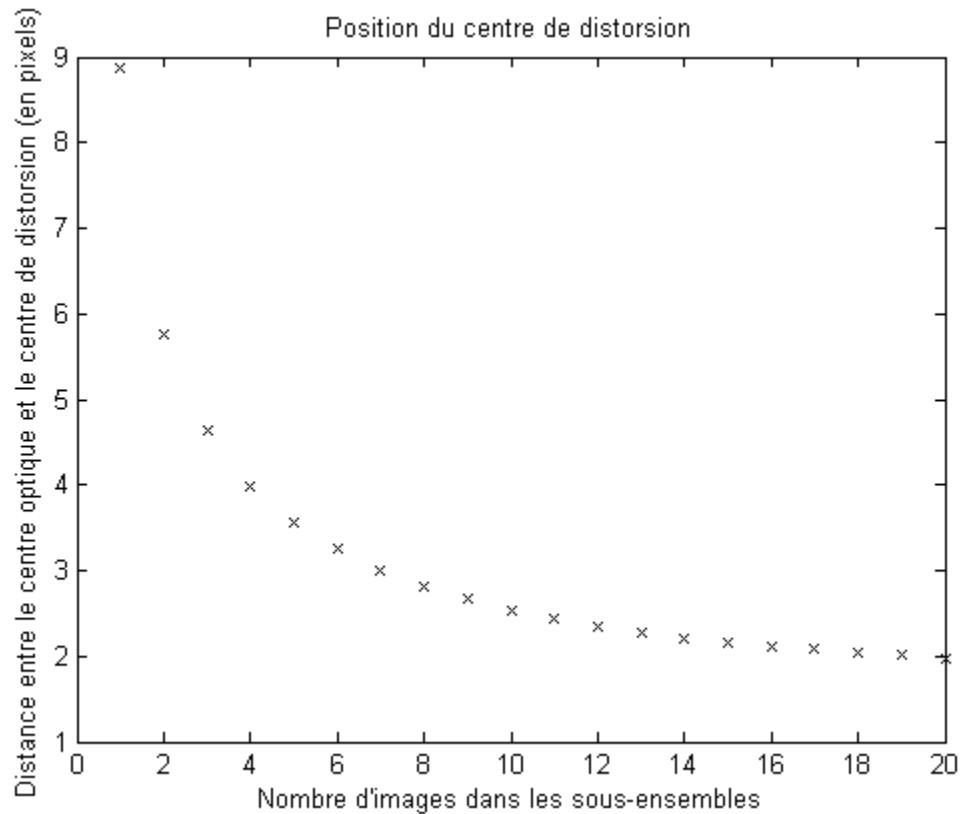


Figure 3.6. Résultats sur la position du centre de distorsion

On remarque donc que la position du centre de distorsion identifié à l'aide d'une méthode n'utilisant que les propriétés de la distorsion (voir ANNEXE V) est très près de la position du centre optique obtenu par calibrage. L'hypothèse d'utiliser le centre optique comme centre de distorsion est donc acceptable.

3.2.4 Expérience #4 : modèle de compensation de la rotation

Pour réaliser cette expérience, un calibrage est effectué à partir d'un ensemble de 8 images. Ce nombre d'images est légèrement supérieur au nombre de 5 déterminé par l'expérience #2, car le but de cette expérience n'est pas d'évaluer la partie du calibrage sans la rotation. On effectue donc cette partie du calibrage avec un plus grand nombre d'images afin d'obtenir un calibrage d'une plus grande précision pour minimiser l'impact de cette étape de calibrage sur l'expérience. Puis, le calibrage de la partie du modèle comprenant la compensation de la rotation est réalisé à partir d'un autre ensemble de 38 images dans lesquelles une rotation est effectuée entre le tube et la tête de la caméra endoscopique. Finalement, l'erreur de

reprojection est évaluée sur un troisième jeu de 47 images utilisées uniquement pour la validation et non pour le calibrage. L'angle de rotation entre la tête de la caméra et le tube endoscopique varie de $-0,689$ radians ($-39,5$ degrés) à $1,066$ radians ($61,1$ degrés) dans l'ensemble de validation. Pour des angles en dehors de cet intervalle, le système de suivi ne reconnaissait pas les deux marqueurs nécessaires à la fois. L'expérimentateur a parcouru cet intervalle en faisant tourner la tête à l'aide de 38 incréments lors du calibrage et de 47 incréments lors de la prise des vues de validation, ce qui explique le nombre de vues utilisées.

Quatre méthodes de compensation de la rotation sont comparées. La première est la méthode de Wu dans laquelle la rotation est effectuée dans le plan image avec le centre optique comme centre de rotation. La seconde effectue également la rotation dans le plan image, mais en utilisant un centre de rotation déterminé à l'aide de la méthode présentée en annexe (voir ANNEXE VII). La troisième correspond à la méthode présentée dans ce travail dans laquelle la rotation est effectuée dans le système de coordonnées de la caméra autour de l'axe Z de ce système de coordonnées, ce qui correspond à un alignement parfait du système de lentilles. La quatrième et dernière méthode est celle véritablement utilisée dans ce travail, dans laquelle l'axe de rotation est calibré plutôt que d'assumer la rotation parfaite autour de l'axe Z .

Pour chacune de ces méthodes, on évalue l'erreur de reprojection en fonction de l'angle de rotation. Il est à noter que la méthode de détermination de l'angle de rotation à partir des marqueurs optiques est la même pour les quatre méthodes, la correction apportée au modèle de projection est toutefois différente.

La Figure 3.7 montre les résultats obtenus pour les 4 méthodes. La première méthode et la troisième semblent donner des erreurs de reprojection comparables et significativement plus élevées que les deux autres méthodes. On élimine donc ces deux méthodes sans calibrage pour analyser plus en détail les deux autres qui utilisent un calibrage. En effet, dans les images utilisées pour réaliser cette expérience, il y a une différence de 48,4 pixels entre la position du centre de rotation et celle du centre optique pour les méthodes 1 et 2. Dans le même ordre d'idée, l'axe de rotation utilisé dans la méthode 4 croise le plan image à 16,6 mm de l'origine, ce qui signifie que le point physique où le système de lentille projette le centre optique est situé à 16,6 mm du point physique autour duquel s'effectue la rotation. Dans les deux cas, l'hypothèse de pouvoir considérer le cas idéal est rejetée et un calibrage est nécessaire.

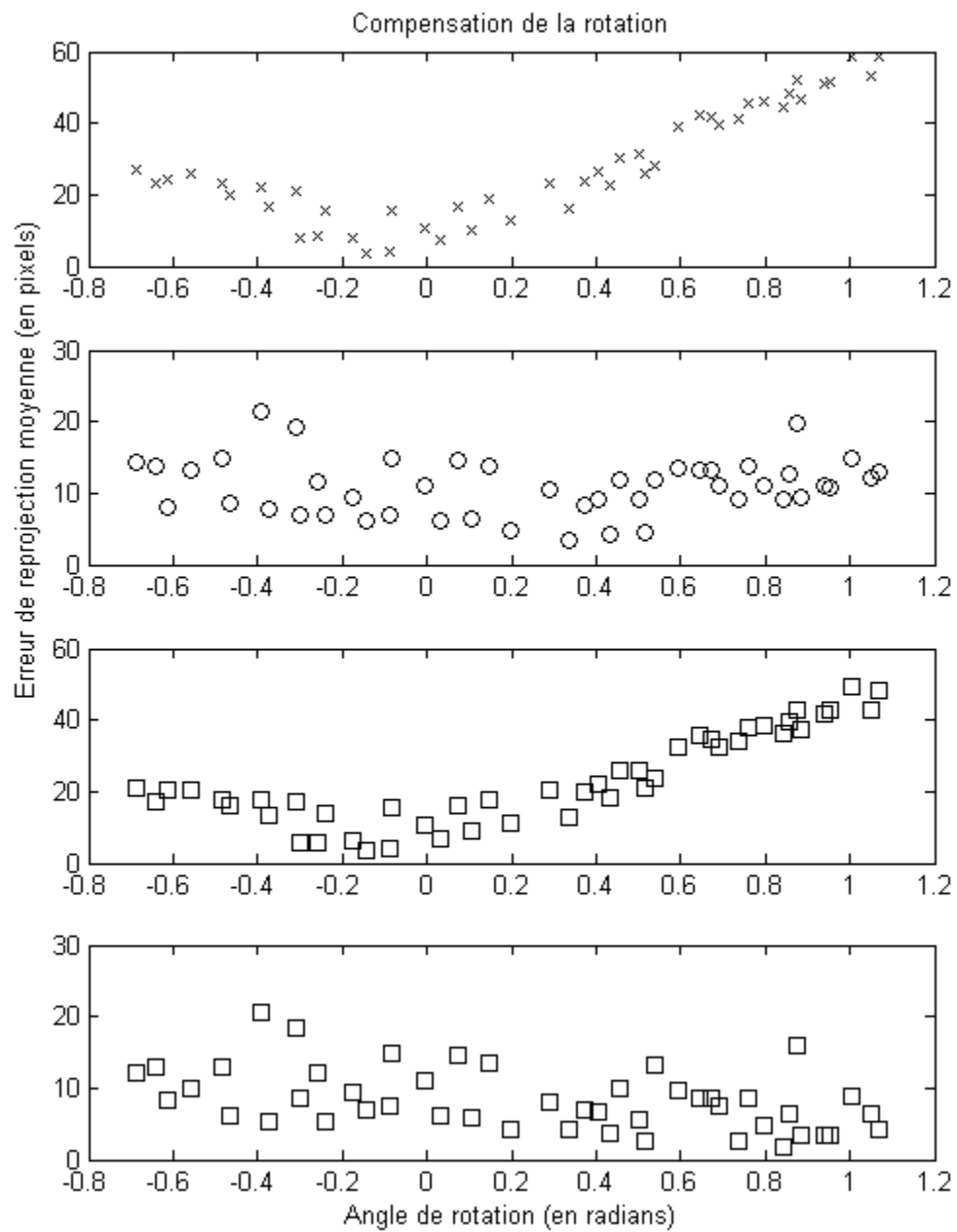


Figure 3.7. Comparaison des méthodes de compensation de la distorsion. De haut en bas : méthode de Wu sans calibrage, méthode de Wu avec calibrage, méthode proposée sans calibrage, méthode proposée avec calibrage

La Figure 3.8 montre la comparaison des deux méthodes utilisant un calibrage, soit la méthode de Wu avec calibrage du centre de rotation (représentée par les croix) ainsi que la méthode proposée par ce travail (représentée par les cercles).

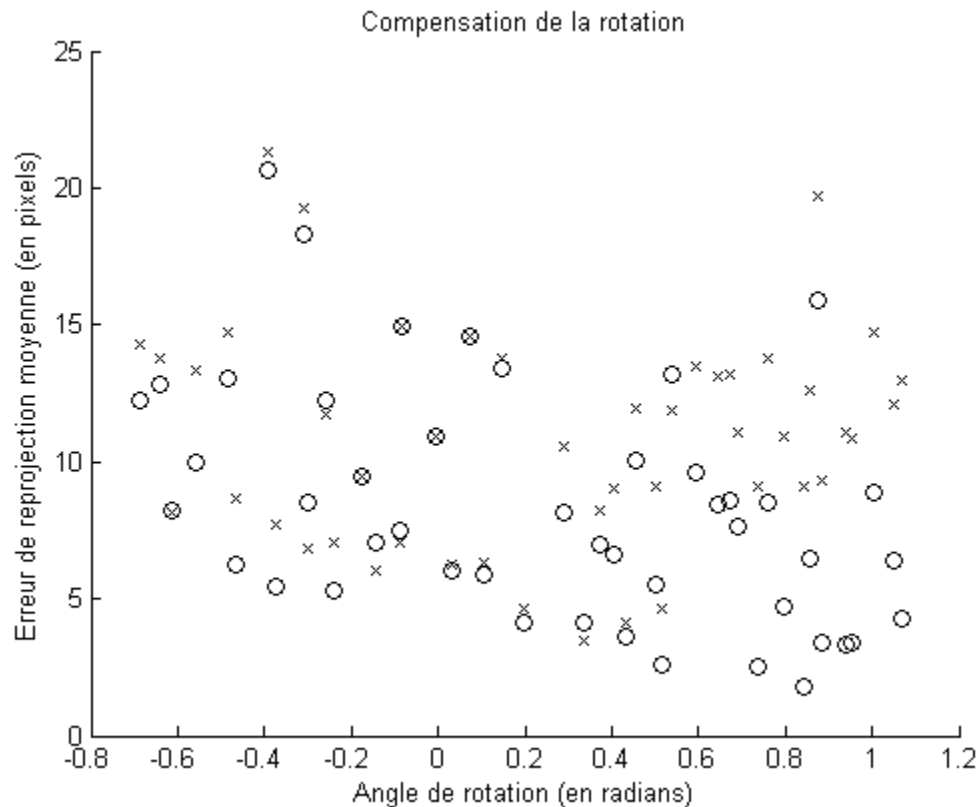


Figure 3.8. Comparaison de l'erreur de reprojection pour les deux méthodes utilisant un calibrage

On remarque que l'erreur est très semblable pour de faibles valeurs d'angles. Or, pour des angles plus grands, la méthode proposée dans ce travail fournit de meilleurs résultats.

L'utilisation du modèle plus complet de Yamaguchi pour la compensation de la rotation relative entre la tête de la caméra et le tube de l'endoscope est donc validée.

3.2.5 Expérience #5 : mesure de distances

Pour réaliser cette expérience, un calibrage est effectué à partir d'un ensemble de 5 vues, tel que déterminé par l'expérience #2. Ensuite, des images de marqueurs MicronTracker sont prises. Pour chaque marqueur, dont la longueur est disponible dans le patron du marqueur MicronTracker, trois paires de vues sont prises, une reconstruction de chacune des extrémités

de l'axe X du marqueur est effectuée et la mesure de la longueur de cet axe est comparée avec la valeur acceptée. La Figure 3.9 montre un exemple d'une des paires de vues calibrées utilisées pour la reconstruction.

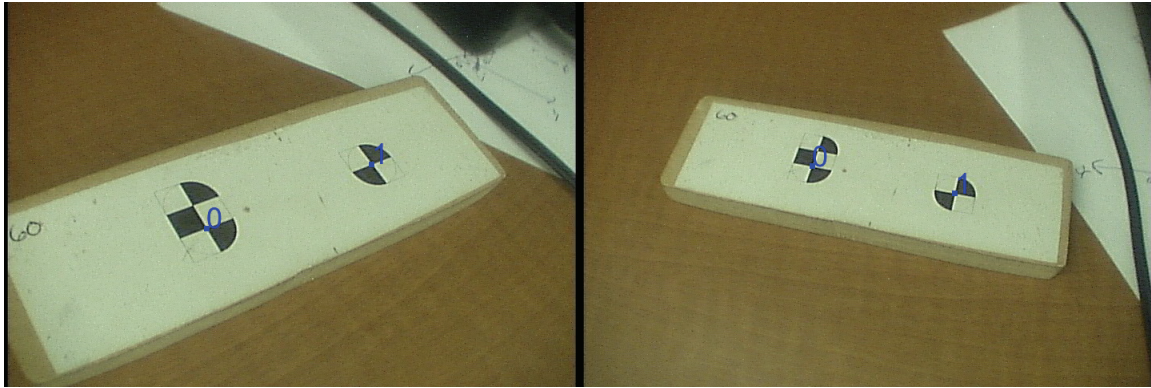


Figure 3.9. Exemple de paire de vues pour la reconstruction 3D.

Le tableau suivant résume les résultats de cette expérience :

Tableau 3.5. Résultats de l'expérience #5

Marqueur	Longueur acceptée (en mm)	Paire de vues	Longueur mesurée (en mm)	Erreur (en mm)
<i>Marker060</i>	52,1235	#1	51,3837	0,7398
		#2	51,3846	0,7389
		#3	52,8949	0,7715
		Moyenne	-	0,7501
<i>Marker080</i>	69,4298	#1	70,3830	0,9531
		#2	70,3305	0,9007
		#3	69,5805	0,1506
		Moyenne	-	0,6681
<i>Marker100</i>	87,0590	#1	85,9966	1,0624
		#2	86,3132	0,7458
		#3	85,8060	1,2530
		Moyenne	-	1,0204
<i>Marker120</i>	104,492	#1	103,773	0,7193
		#2	103,300	1,1918
		#3	102,766	1,7263
		Moyenne	-	1,2125

On remarque donc que l'erreur de reconstruction est de l'ordre du millimètre, ce qui constitue un ordre de précision acceptable pour le domaine clinique.

Afin de vérifier la précision des mesures sur des plus petites distances, on utilise les vues de validation (non utilisées pour le calibrage) afin de reconstruire en 3D les points de l'objet de calibrage. Chaque ligne de la grille $M \times N$ possède $(M - 1)$ segments, chaque colonne de la grille possède $(N - 1)$ segments, pour un total de $2MN - M - N$ segments. L'erreur de mesure de chacun de ces segments est calculée pour estimer la précision de la mesure.

La grille possède 9 points par 9 points, ce qui donne un total de 144 segments par grille. La valeur acceptée de la longueur du segment est donnée par le MicronTracker et est de 8,001 mm. 40 paires de vues ont été utilisées pour obtenir 40 reconstructions. L'erreur moyenne obtenue a été de 0,72763 mm, précision qui est également de l'ordre du millimètre.

3.3 Validation des vues

Les vues ont été présentées au Dr Stefan Parent. Ce dernier approuve la pertinence des vues proposées.

Il confirme qu'il apprécierait disposer, lors des chirurgies, des informations supplémentaires ajoutées à l'image de la caméra endoscopique dans les vues de type réalité augmentée. Les données de distances, particulièrement, constituent des informations précieuses pour aider le chirurgien à évaluer l'état de l'opération.

De plus, Dr Parent confirme l'intérêt du second type de vues, soit les vues de virtualité augmentée. Ces dernières permettent en effet de situer le site chirurgical dans un contexte un peu plus large, tout en permettant également de se concentrer sur une zone d'intérêt afin d'observer le site d'un point de vue (virtuel) différent.

Bien que les vues aient été approuvées par un expert chirurgien, d'autres expériences doivent être réalisées afin de valider la pertinence et l'utilité des vues proposées, tel que discuté dans le chapitre suivant.

CHAPITRE 4 DISCUSSION GÉNÉRALE

Ce chapitre vise à discuter de la méthode proposée et énoncer les différentes limites de cette dernière. Le système de suivi sera d'abord abordé. Puis, les différentes manipulations à effectuer pour calibrer le système sont critiquées pour ensuite discuter de la méthode de validation proposée.

4.1 Système de suivi

Le système de suivi utilisé, soit le MicronTracker, possède de nombreux avantages qui ont justifié son choix pour l'implémentation de la méthode proposée. Il offre notamment une grande flexibilité par la facilité avec laquelle il est possible de créer de nouvelles configurations de marqueurs. Certains points méritent toutefois d'être critiqués.

Tout d'abord, la précision du système pose problème dans certaines configurations critiques. En effet, les expériences ont fait ressortir qu'une grande partie de l'erreur est due à la précision des paramètres extrinsèques de la caméra, qui est fortement liée à la précision du système de suivi. Or, certaines positions des marqueurs rendent plutôt imprécise la pose du marqueur, cette imprécision étant amplifiée par le phénomène de levier entre le marqueur et la véritable position du système de coordonnées de la caméra. Il est possible de réduire cette imprécision par l'utilisation de marqueurs plus grands, mais ceux-ci encombreraient l'espace de travail du chirurgien et ne seraient donc pas vraiment utilisables dans un véritable contexte clinique.

Par ailleurs, la robustesse de la reconnaissance des marqueurs est aussi une limite. En effet, certains facteurs diminuent la capacité du système MicronTracker à reconnaître et donc suivre les marqueurs. Le système éprouve des difficultés à identifier les marqueurs en mouvement. La vitesse du mouvement doit rester relativement faible afin que la pose des marqueurs soit fournie en continue. De plus, certains algorithmes sont utilisés pour ajuster le contraste de l'image des caméras du système MicronTracker lorsqu'un marqueur est identifié, en fonction de ce marqueur. Ainsi, lorsque plus d'un marqueur sont présents, l'ajustement peut poser problème et rendre plus difficile l'identification d'un second marqueur. La robustesse de l'identification de la pose est donc dépendante du nombre de marqueurs présents et diminue en présence de marqueurs multiples, ce qui peut causer problème dans le cadre de la méthode proposée. Finalement, la présence d'un grand nombre de *Xpoints* potentiels nuit à

l'identification des marqueurs par le système. Or, l'objet de calibrage possède une grille composée d'un patron d'échiquier dont chaque point possède en fait les caractéristiques d'un *Xpoint*. Le système MicronTracker reconnaît donc tous les coins de la grille comme des *Xpoints* potentiels, ce qui nuit à la reconnaissance des véritables marqueurs. Le MicronTracker va même parfois jusqu'à reconnaître des marqueurs qui ne sont pas véritablement présents dans l'image, car la position d'une certaine combinaison des coins de la grille tombe sous la tolérance requise pour reconnaître un certain patron de marqueur. Il faut donc parfois couvrir la grille de l'objet de calibrage pour permettre au MicronTracker de reconnaître les autres marqueurs dans la scène, ce qui complique le processus de calibrage.

4.2 Manipulations pour le calibrage

Un des objectifs de la présente méthode de calibrage est qu'elle soit le plus facile possible à utiliser, d'où l'automatisation du maximum d'opérations possible, notamment l'identification des points de la grille dans les images endoscopiques. Le développement d'un outil dont la méthode de calibrage est efficace, rapide et facile à réaliser est primordial pour son intégration dans un contexte clinique.

Or, la méthode proposée requiert certaines manipulations qui pourraient être améliorées. La section précédente mentionne la difficulté que le système de suivi a à identifier les marqueurs en présence de la grille de l'objet de calibrage. Ces difficultés sont contournées par certaines manipulations peu instinctives qui consistent en l'introduction progressive des marqueurs ou alors au masquage de la grille pour les caméras du système de suivi. De telles manipulations seraient à éviter dans un contexte clinique.

Aussi, même si la méthode vise à être faite simplement en tenant la caméra de manière à visualiser un objet de calibrage, il s'avère qu'en réalité, les manipulations sont plus complexes. Il est en effet préférable de maintenir la caméra fixe à l'aide d'un appareil quelconque, un bras ManFrotto par exemple. Cette contrainte est due entre autres à certains problèmes de synchronisation qui sont détaillés dans le chapitre suivant, mais également à la précision du système de suivi qui est moindre lorsque le système est en mouvement. L'immobilité de la caméra assure une plus grande précision au processus de calibrage, mais complique ce dernier : au lieu de l'effectuer « à main levée », on doit fixer la caméra, prendre une vue, déplacer la caméra, la refixer, prendre une autre vue, etc.

Par ailleurs, la première partie du calibrage doit être effectuée sans rotation relative entre la tête de la caméra et le tube de l'endoscope. Or, ce point est relativement facile à oublier et une rotation peut aisément passer inaperçue lorsque l'on déplace la caméra entre les différentes vues utilisées pour le calibrage. Certaines manipulations peuvent être faites pour éviter le problème, comme l'utilisation d'un dispositif de fixation supplémentaire le temps du calibrage ou alors le déplacement de l'objet de calibrage plutôt que de la caméra. Il pourrait toutefois être souhaitable d'adapter la méthode de calibrage pour relaxer la contrainte de ne pas effectuer de rotation lors de la première partie du calibrage.

Finalement, le placement des marqueurs sur le tube complique la procédure de calibrage. En effet, la nécessité pour certaines étapes du calibrage que le système de suivi soit en mesure de voir à la fois le marqueur de l'objet de calibrage, celui du tube de l'endoscope ainsi que celui de la tête de la caméra, tout en permettant à la caméra endoscopique de visualiser la totalité de la grille de l'objet de calibrage, entraîne une configuration complexe des éléments du système. Cette complexité est d'autant plus augmentée par le fait que le système de suivi tend à perdre le suivi des marqueurs lorsque la grille de l'objet de calibrage se trouve dans son champ de vision.

Certains points de la méthode pourraient donc être améliorés afin d'éviter certaines manipulations plus complexes.

4.3 Méthode de validation

Étant donné l'état encore relativement préliminaire du système d'assistance chirurgicale proposé, les différents processus de validation ont dû être adaptés au travail réalisé. Pour ces raisons, la validation de la méthode de calibrage est restée au niveau de la vérification de sa précision et la validation des vues proposées est demeurée qualitative. D'autres validations doivent toutefois être effectuées avant d'être en mesure d'amener ce système dans un contexte clinique.

Tout d'abord, en ce qui a trait à la méthode de calibrage, la seule validation faite a été au niveau de la précision de la méthode lorsqu'elle est appliquée par l'auteur. Aucune mesure de la précision de la méthode lorsqu'appliquée par des techniciens médicaux n'a été effectuée. De plus, aucune mesure du temps nécessaire au calibrage n'a été effectuée, que ce soit pour évaluer la facilité avec laquelle un nouvel usager peut réaliser le calibrage ou pour évaluer la

courbe d'apprentissage de la méthode. L'état embryonnaire du système ne se prêtait pas encore à ce genre d'expériences de validation.

Dans le même ordre d'idées, les vues proposées dans le cadre de ce travail de recherche n'ont été validées que par l'opinion d'un chirurgien expert. Aucune mesure quantitative n'a été effectuée. Par exemple, aucune expérience de comparaison du nombre et du niveau des complications résultantes ou du temps nécessaire pour effectuer une tâche, lorsque la chirurgie est réalisée avec ou sans le système, n'a pu être conduite.

Ce type de validations sera nécessaire pour la justification de la pertinence du système en salle de chirurgie.

Elles pourront être effectuées lorsque le développement du système sera plus avancé. En effet, l'élaboration de ce travail de recherche s'est faite en parallèle avec le développement de la méthode de recalage qui sera utilisée pour recaler le modèle 3D préopératoire. Cette méthode étant toujours en cours de développement, il est difficile de générer des vues à partir du système réel. Cette dépendance envers la méthode de recalage a entraîné comme conséquence le recours à des outils de modélisation 3D pour générer les maquettes de visualisation et également l'impossibilité de valider le système complet dans un contexte clinique.

La prochaine section rapporte les recommandations de l'auteur quand au travail à effectuer pour poursuivre le développement de l'outil d'assistance chirurgicale dans le but de son intégration éventuelle dans les salles de chirurgies.

4.4 Défis pour l'intégration en salle de chirurgie

De nombreux points doivent être traités avant qu'il soit possible d'intégrer le système d'assistance chirurgicale à un véritable contexte clinique.

4.4.1 Recalage des modèles préopératoires

Le recalage des modèles opératoires, tel que discuté à la section 1.1.2, est toujours un problème en cours de résolution. À court terme, on se dirige vers la solution de Chevretils [11] qui consiste à recaler les modèles de la colonne à partir d'une radiographie intra-opératoire unique. Ce recalage est rendu possible sur une seule vue grâce à l'application de contraintes supplémentaires (alignement [21] ou modèle de variabilité [23]).

Bien que cette approche soit intéressante puisqu'elle ne nécessite qu'une seule radiographie, à plus long terme on cherche à éliminer ce besoin d'irradier le patient en utilisant plutôt une modalité non ionisante. Certains travaux [41-44] ont déjà été réalisés en ce sens et sont très encourageants.

Une voie qui vaudrait la peine d'être explorée est la combinaison de ces deux approches. Le recalage à l'aide des ultrasons n'est pas toujours aisé : les images sont très bruitées, cette modalité n'est pas conçue pour l'identification des os, etc. Toutefois, le fait d'ajouter l'information spécifique à la colonne, à la lueur des travaux développés à l'intérieur même du LIV4D, pourrait constituer une avenue très intéressante pour obtenir un recalage de qualité à partir d'images ultrasoniques. Ainsi, l'intégration des contraintes d'alignement ou d'un modèle de variabilité au recalage à l'aide des ultrasons constitue une piste intéressante.

Par ailleurs, une telle méthode de recalage repose sur certaines hypothèses, comme l'immobilité de la colonne par exemple. En effet, les méthodes plus agressives qui exigent de fixer un marqueur à la structure à suivre ont l'avantage d'être en mesure de suivre les mouvements des structures en temps réel une fois le calibrage initial effectué. Toutefois, cette procédure très invasive est difficilement applicable dans le contexte minimalement invasif et va à l'encontre de ses principes. Néanmoins, il serait pertinent de mesurer l'amplitude des déplacements de la colonne au cours de la chirurgie. Par exemple, il serait intéressant de vérifier l'amplitude du mouvement de la colonne vertébrale dû à la respiration du patient ou dû aux forces appliquées par le chirurgien sur le patient. Une méthode « statique », i.e. qui ne met pas à jour le recalage du modèle au cours de la chirurgie, ne permettra, dans le meilleur des cas, que d'obtenir une précision de l'ordre de ce mouvement. Il serait également intéressant de comparer ce mouvement à celui de marqueurs cutanés externes placés sur le patient qu'il serait possible de suivre de l'extérieur. Étant donné que la géométrie de la colonne n'est pas aisément prédictible à partir de l'aspect extérieur du tronc, on ne s'attend pas à être capable de prédire la modification de la forme de la colonne à partir du mouvement de ces marqueurs extérieurs. Toutefois, la détection d'un mouvement d'une amplitude non négligeable sur les marqueurs externes pourrait être un indicateur de la nécessité de faire un ajustement du recalage à l'aide de la sonde ultrasonique calibrée.

4.4.2 Suivi des instruments

Le suivi des instruments est une des étapes qui n'a pas été réalisée dans le cadre de ce projet de recherche. Que ce soit à des fins de visualisation (afficher les modèles 3D des instruments dans le monde virtuel ou les projeter sur l'image endoscopique réelle) ou pour des fins de mesure (identifier des points 3D à l'aide du bout de l'instrument), il est nécessaire d'effectuer le suivi des instruments en temps réel. Cette tâche consisterait principalement :

- À la modélisation 3D des outils : un modèle 3D, pouvant aller du simple cylindre à un modèle complexe développé à l'aide d'un outil de modélisation 3D tel que 3D Studio Max ou Maya, doit être créé.
- À la modification de l'outil pour lui fixer un marqueur identifiable par le système de suivi : on procède généralement à la fixation d'un marqueur externe à l'instrument chirurgical à suivre.
- Au recalage entre le modèle 3D et l'outil lui-même : le recalage doit en fait être effectué entre le modèle 3D et le marqueur fixé à l'outil.
- À la mise à jour en temps réel de l'état de l'outil : cette étape, moins essentielle, pourrait par exemple consister en l'ouverture ou la fermeture des pinces de l'instrument sur le modèle 3D en même temps que sur l'outil réel. À la connaissance de l'auteur, aucun système n'implémente présentement cette mise à jour en temps réel.

De nombreux travaux existent quant au calibrage nécessaire au suivi des instruments. Certains [83] le font d'une manière très ingénieuse en n'utilisant que des composantes très peu dispendieuses, comme la librairie ARToolKit, et affirment obtenir des résultats comparables à des systèmes beaucoup plus complexes. Cette dernière méthode propose entre autres une technique basée sur la formation d'une sphère pour identifier le bout des instruments ainsi qu'une méthode basée sur une segmentation très simple pour déterminer la position du manche des instruments. Ce genre de travaux peut servir de base pour l'implantation du suivi des instruments.

4.4.3 Système de suivi

Tel que mentionné dans l'analyse des résultats, le système de suivi utilisé, soit le MicronTracker, présente certains problèmes. La précision du système est un des points

importants quand au succès du suivi duquel dépend la fiabilité de l'outil d'assistance chirurgicale.

Certaines méthodes existent pour obtenir une estimation plus appropriée de l'état d'un système dont les mesures sont bruitées. Par exemple, les filtres de Kalman peuvent être utilisés pour modéliser un tel système bruité. Il serait intéressant de vérifier l'apport que pourraient amener de tels filtres quant à la continuité des mesures provenant du système de suivi afin d'éviter les problèmes de tremblement. Toutefois, le fait que les spécifications du MicronTracker [82] mentionnent une mesure de précision pour l'identification des points et une autre lorsque les cibles sont immobiles laissent croire que certains algorithmes de stabilisation ont déjà été implémentés.

Une autre manière d'augmenter potentiellement la précision, spécialement dans les configurations critiques posant problème pour le MicronTracker, consiste en l'ajout d'autres caméras binoculaires. L'interface de programmation du MicronTracker permet en effet d'interagir d'une manière assez transparente avec un nombre variable de caméras MicronTracker. Cet ajout pourrait aider à augmenter la précision et à améliorer la capacité de ne pas perdre les marqueurs suivis.

De plus, puisqu'il a été montré que le bruit provenant du système de suivi est considérable, il pourrait être intéressant de considérer une alternative pour la détermination des paramètres extrinsèques lors du calibrage. La méthode proposée dans ce travail de recherche considère que l'on connaît les positions 3D des points de l'objet de calibrage, ce qui est aisément réalisable puisque l'on utilise déjà un système de suivi optique pour suivre la caméra. Ajouter un marqueur sur l'objet de calibrage n'est donc pas un problème. Or, dans le cas où la caméra est contrôlée par un bras robotisé constituant donc le système de suivi, il n'est pas trivial d'assigner un repère à l'objet de calibrage et ainsi obtenir la position 3D des points de cet objet. Le problème de calibrage œil-main, qui consiste à trouver la transformation rigide entre le repère du système de suivi et celui de la caméra, reste toutefois le même et des méthodes [84] existent pour effectuer ce calibrage sans connaître les positions absolues des points de l'objet de calibrage. Il pourrait être intéressant d'expérimenter avec ce genre de techniques, car elles élimineraient des lectures (bruitées) de la position de l'objet de calibrage, donnée par le système de suivi et utilisée lors du calibrage, et pourraient donc éventuellement résulter en un calibrage moins soumis au bruit du système de suivi.

Par ailleurs, le système de suivi MicronTracker est constamment en développement. Les récents avancements incluent entre autres la disponibilité d'un nouveau mode de capture à domaine hautement dynamique (*High Dynamic Range* ou *HDR*). Ce dernier permet un suivi précis et fiable malgré la présence de zones où l'éclairage est très différent dans le champ de vision du MicronTracker. Ce type d'avancée pourrait régler les problèmes d'identification rencontrés lorsque plusieurs marqueurs sont présents et que le contraste est ajusté pour un seul des marqueurs. De plus, Claron Technology Inc. développe également un système de navigation nommé ClaroNav utilisant le système MicronTracker comme système de suivi. Ainsi, les améliorations constantes du système MicronTracker semblent très prometteuses comme solution aux quelques faiblesses de ce dernier.

4.4.4 Interaction avec le système

Au cours de ce travail de recherche, aucune mention n'a été faite sur l'interaction de l'utilisateur avec le système. Étant donné l'état préliminaire du projet, ce point n'est pas encore critique. Toutefois, lorsque viendra le temps d'intégrer l'outil à l'intérieur des salles de chirurgie, il faudra s'assurer que l'interaction avec le système soit adaptée à l'environnement de la salle de chirurgie ainsi qu'aux tâches que doit accomplir le chirurgien.

Sielhorst et al. [47] mentionnent ce point dans leur revue de littérature. Ils mentionnent que les éléments d'interfaces classiques, comme les fenêtres et les menus, ainsi que les périphériques standards, comme les claviers et les souris, ne sont pas adaptés à l'environnement chirurgical. Ils citent également quelques travaux qui constitueront un bon point de départ à consulter lorsque viendra le temps de définir le protocole d'interaction avec l'outil d'assistance chirurgicale.

4.4.5 Définition du protocole de calibrage

Afin d'assurer la qualité de la précision du calibrage, il sera nécessaire de bien définir les manipulations à effectuer pour réaliser le calibrage de la caméra endoscopique avant la chirurgie. En effet, de nombreux facteurs peuvent avoir une influence sur la qualité du calibrage, notamment :

- Le nombre de vues utilisées
- La taille de la grille de calibrage dans l'image endoscopique

- La position de la grille de calibrage dans l'image endoscopique
- L'angle entre la caméra et la grille de calibrage
- Etc.

Certains travaux [85] ont été réalisés pour évaluer les configurations optimales pour l'acquisition des vues fournies en entrée au calibrage. Ils pourront donc servir de base pour établir des recommandations claires sur ces différents points afin d'assurer un calibrage à la fois précis et reproductible.

4.4.6 Compensation de la rotation

Bien que les travaux de Girard [9] ont été basés sur la méthode de Wu pour la compensation de la rotation, le présent travail a démontré qu'il est préférable de revenir vers la méthode de Yamaguchi qui, bien que l'algorithme menant à son calibrage soit plus complexe à implémenter, donne de meilleurs résultats. Ainsi, d'un point de vue mathématique et algorithmique, la méthode de Yamaguchi est préférable.

De plus, au niveau du matériel, la méthode de Yamaguchi propose une configuration physique plus adaptée à un contexte d'utilisation réel. En effet, l'utilisation d'un second marqueur optique situé sur le tube de l'endoscope est pratique dans le cadre de la recherche, car il permet de réutiliser le système de suivi déjà nécessaire. L'utilisation d'un encodeur de rotation telle que proposée par Yamaguchi nécessite l'acquisition de matériel supplémentaire qui exige une adaptation, généralement plus complexe que dans le cas d'un marqueur optique, de l'endoscope pour son intégration.

Toutefois, la méthode de Yamaguchi n'exige d'avoir un marqueur que sur la tête de l'endoscope, qui est maintenue dans une orientation relativement constante par le chirurgien. Le marqueur s'y trouvant est donc sujet à beaucoup moins de problème d'encombrement et d'obstruction de ligne de vue du système de suivi, comparativement au marqueur situé sur le tube endoscopique subissant de nombreuses rotations.

Ainsi, pour des considérations d'encombrement du travail du chirurgien et de visibilité des marqueurs par le système de suivi, l'expérimentation avec un encodeur de rotation et le retour à la méthode de Yamaguchi méritent d'être explorés.

4.4.7 Synchronisation des multiples sources de données

Le système d'assistance chirurgicale proposé utilise les données provenant de plusieurs sources en temps réel. Les images endoscopiques arrivent à un débit d'environ 30 images par seconde, alors que le MicronTracker fournit les poses des marqueurs à un taux de 15 fois par seconde et ce, d'une manière aucunement synchronisée avec la caméra endoscopique. L'ajout d'un encodeur de rotation constituerait une source supplémentaire qui ne serait également pas synchronisée avec les autres. Ainsi, pour un instant donné, il y a un certain décalage entre le moment présent et la dernière obtention de données de chacun des systèmes, ce décalage étant variable pour chaque système comme le montre la Figure 4.1.

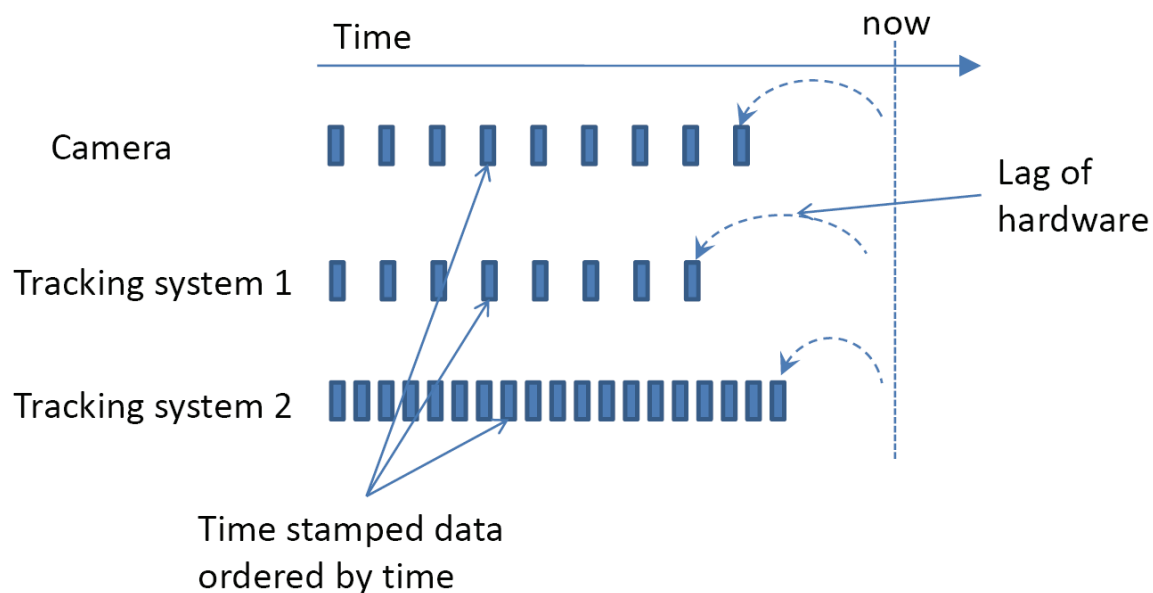


Figure 4.1. Décalage dans les différentes sources de données en temps réel (tirée de T. Sielhorst, "New Methods for Medical Augmented Reality," Technische Universität München (TUM), 2008. ©2008 Tobias Sielhorst. Reproduit avec permission.)

Dans le cadre de ce travail de recherche, les problèmes causés par l'absence de synchronisation ont été évités par l'acquisition des données alors que le système était immobile. L'immobilité entraîne que les données sont pratiquement répétées et sont donc identiques. Le décalage peut donc être ignoré puisque chaque donnée est sensiblement la même que la précédente.

Toutefois, lorsque le système sera utilisé dans un contexte clinique, il devra bien réagir même lorsque le chirurgien déplace l'endoscope ou les instruments d'un endroit à l'autre. Il existe

plusieurs stratégies pour implémenter une certaine synchronisation. Sielhorst et al. [86] en mentionnent¹⁰ quelques-unes qui sont résumées à la Figure 4.2.

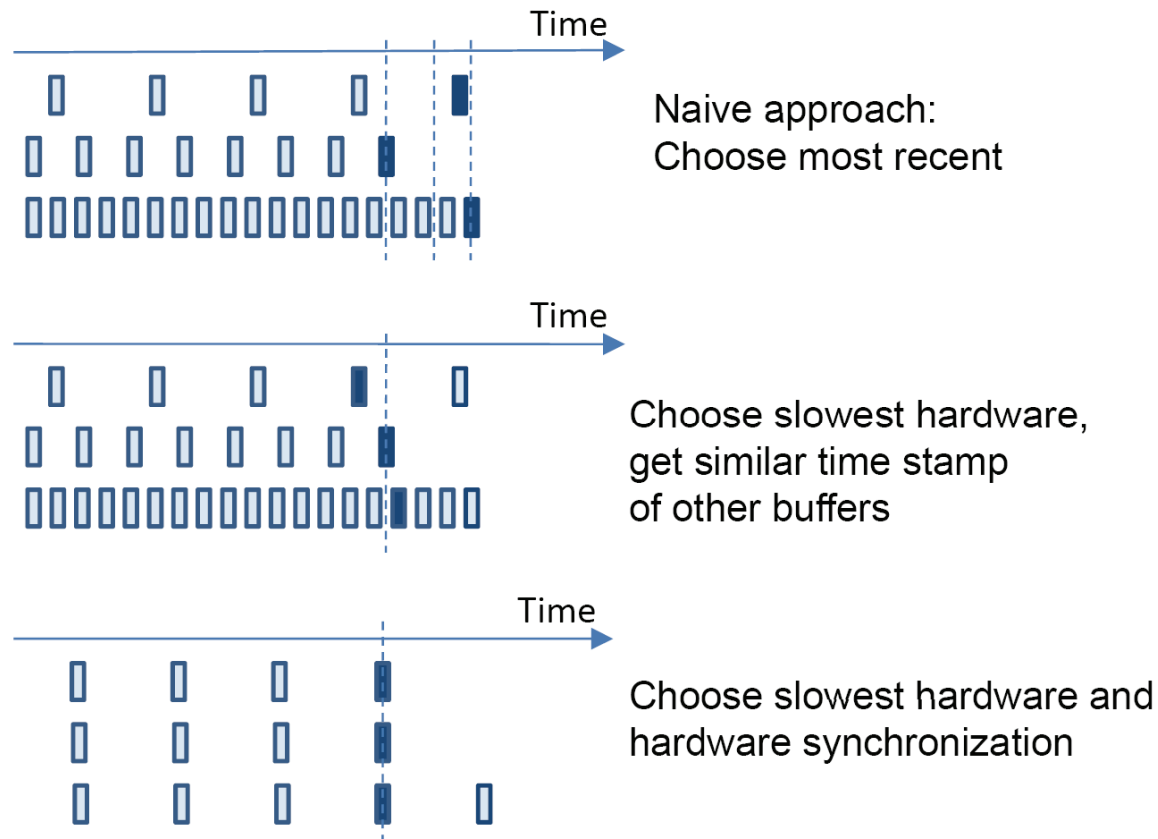


Figure 4.2. Différentes approches au problème de synchronisation de sources multiples de données en temps réel (tirée de T. Sielhorst, "New Methods for Medical Augmented Reality," Technische Universität München (TUM), 2008. ©2008 Tobias Sielhorst. Reproduit avec permission.)

Aucune des stratégies n'est parfaite et ne règle complètement le problème de la synchronisation. Il sera donc nécessaire d'explorer les différentes possibilités dans des travaux subséquents.

4.4.8 Système endoscopique du LIV4D

L'utilisation du système endoscopique présent au LIV4D dans le cadre de ce travail de recherche a permis d'identifier quelques caractéristiques de celui-ci qui méritent une certaine attention.

¹⁰ Ce travail de Sielhorst et al. décrit l'implémentation réelle d'un système de réalité augmentée. Il mentionne plusieurs points à considérer lors de l'implémentation concrète d'un tel système et constitue donc une excellente référence.

Tout d'abord, la connexion entre le système endoscopique et l'ordinateur de travail était faite à l'aide d'un port FireWire situé derrière la boîte du système endoscopique. Or, des recherches sur les caractéristiques du modèle en question, confirmées par la suite par un échange de courriel avec le service de support de Stryker, ont montré que le port FireWire ne fait pas partie des configurations supportées pour ce modèle. Autrement dit, la fonctionnalité de ce port ne fait officiellement pas partie des caractéristiques de ce modèle. Les modèles plus récents du même vendeur offrent le port FireWire comme caractéristique standard. Un port DVI est également présent sur la boîte en question et fait partie des configurations supportées. Or, ce port est généralement utilisé pour afficher directement l'image sur un moniteur externe. Bref, lorsque le système sera intégré en salle d'opération, il sera important de s'assurer que le matériel endoscopique soit utilisé d'une manière officiellement supportée.

Par ailleurs, il existe un délai entre les images provenant du port DVI (directement affichées sur un moniteur externe) et celle provenant du port FireWire. Ce délai est probablement dû à l'encodage du signal pour le transport via FireWire, suivi du décodage sur l'ordinateur de travail et de l'affichage de l'image. Ce délai est visible à l'œil nu lorsque l'on observe à la fois l'image sur le moniteur externe (port DVI) et celle sur l'ordinateur de travail (port FireWire). Selon la configuration retenue pour le système (acquisition via le port FireWire sur un modèle supporté ou via le port DVI directement), il sera nécessaire de tenir compte de ce délai, entre la réalité imagée par la caméra et l'image une fois celle-ci obtenue dans l'ordinateur contrôlant le système d'assistance chirurgicale, pour la synchronisation des différentes sources de données.

Finalement, le système utilisé est doté d'un mécanisme de correction automatique de l'intensité lumineuse qu'il n'est pas possible de désactiver. Dans l'état actuel du système où l'image de la caméra endoscopique est utilisée « directement » à des fins visuelles, cette caractéristique ne constitue pas une contrainte. Toutefois, le système complet pourrait utiliser différents algorithmes effectuant un traitement sur le flux vidéo, comme une élimination des réflexions spéculaires ou une reconstruction des surfaces grâce à des algorithmes de structures à partir de l'ombrage ou à partir du mouvement. Ces algorithmes pourraient être déstabilisés par cette correction inattendue. Il sera donc nécessaire de tenir compte du traitement fait automatiquement par le système d'acquisition des images endoscopiques lors de l'implémentation d'algorithmes plus complexes sur les images reçues.

CHAPITRE 5 CONCLUSIONS ET RECOMMANDATIONS

Le présent travail de recherche a présenté un système de réalité augmentée conçu pour l'assistance des chirurgies minimalement invasives de la colonne vertébrale.

Afin de réaliser un tel outil de visualisation 3D, un modèle de formation des images endoscopiques a été développé et validé. Ensuite, une technique de calibrage hors-ligne, effectuée avant la chirurgie, a été conçue afin de déterminer les valeurs des paramètres de ce modèle. Puis, une méthode de mise à jour en ligne, au cours de la chirurgie, a été ajoutée afin de mettre à jour en temps réel les paramètres du modèle qui varient au cours de la chirurgie. De plus, deux modes de visualisations ont été proposés afin de présenter au chirurgien la fusion des images endoscopiques et d'un modèle préopératoire recalé avec ces dernières.

Par la suite, différentes parties du modèle utilisé par la méthode de calibrage ont été validées quantitativement. Dans son ensemble, les vues calibrées par cette dernière permettent de mesurer des distances à l'aide de vues multiples d'un segment avec une précision millimétrique. De plus, les modes de visualisation proposés ont été approuvés par un expert chirurgical.

La contribution de ce travail de recherche se situe à plusieurs niveaux. D'abord, l'objet de calibrage conçu est simple et peu coûteux à produire, en plus d'être mesurable avec une grande précision sans avoir recours à un numériseur 3D. Ensuite, ce travail comprend une identification complètement automatique des points de la grille de calibrage dans les images endoscopiques, ce qui minimise l'intervention de l'utilisateur. De plus, il combine la méthode simple de Wu au modèle plus complexe mais plus précis de Yamaguchi afin de fournir un modèle mathématiquement équivalent à celui de Yamaguchi, mais nécessitant moins de calibrage. Finalement, les vues proposées sont adaptées aux chirurgies minimalement invasives de la colonne vertébrale.

Par ailleurs, un des grands avantages du système développé dans le cadre de ce travail de recherche est sa modularité ainsi que la faible dépendance entre les modules. Cette conception permet de modifier certaines parties du système en n'influençant nullement les autres parties. La liste suivante énonce quelques modifications qu'il serait possible de faire au système tout en conservant la majeure partie de ce dernier :

- Utilisation d'un système de suivi différent : Une abstraction du système de suivi a été faite dans le développement de l'outil afin de ne nécessiter du système que la pose des différents marqueurs. Il serait donc facile de remplacer le système de suivi en ne modifiant que la partie de l'outil qui obtient les poses des marqueurs¹¹.
- Utilisation d'un algorithme de calibrage différent : Bien que la méthode de Zhang ait été retenue, il serait possible d'utiliser le *DLT* ou la méthode de Tsai en ne modifiant que la partie qui obtient les paramètres intrinsèques et extrinsèques à l'aide des différentes vues de l'objet de calibrage. Cette modification très locale n'aurait aucun impact sur le reste de la procédure de calibrage, notamment la procédure de compensation de la rotation.
- Ajout de modules de prétraitement de l'image : On pourrait ajouter certains filtres qui traiteraient l'image endoscopique avant de la fournir au système. Par exemple, on pourrait effectuer un calibrage colorimétrique et effectuer une correction de couleurs. On pourrait également procéder à une réduction du bruit ou à un ajustement de l'intensité lumineuse. Tout ce prétraitement pourrait être effectué sans entraîner de modification au calibrage du système.
- Utilisation d'un endoscope différent : Tant et aussi longtemps qu'on est en mesure d'obtenir en temps réel les paramètres intrinsèques et extrinsèques de la caméra, l'outil peut utiliser cette caméra. Si un endoscope différent ne permet pas d'effectuer une rotation de la tête ou qu'il permet d'effectuer une manipulation donnant un degré de liberté supplémentaire, il est possible de tout simplement ajouter au modèle une étape tenant compte de ce degré de liberté et le reste du système restera inchangé.
- Ajout d'éléments anatomiques : Le système proposé nécessite d'avoir des modèles 3D recalés afin qu'il puisse afficher ces derniers convenablement. Il ne pose aucune hypothèse quant au contenu de ces modèles, tant que ces derniers peuvent être recalés dans l'espace d'affichage du système. Il serait donc possible d'ajouter d'autres éléments anatomiques, comme certaines artères, les poumons ou la surface externe même, sans avoir à modifier le reste de la procédure.

¹¹ À noter que selon le système de suivi choisi, il peut être nécessaire de modifier l'objet de calibrage.

- Utilisation de différentes méthodes de rendu : Encore une fois, le système ne fait pas d'hypothèse sur la méthode de rendu utilisée. Il peut utiliser autant un rendu classique de modèles formés de polygones qu'un rendu plus complexe de données volumétriques qu'il est maintenant possible d'implémenter sur les processeurs graphiques programmables modernes.

Cette liste démontre bien une des caractéristiques importantes du système proposé, soit sa modularité. Cette dernière permet non seulement d'améliorer le système en modifiant de manière indépendante certaines parties, mais elle lui donne également une grande flexibilité. En effet, cette flexibilité permettrait au système d'être adapté à différents contextes d'utilisation. Par exemple, bien que la discectomie par thoracoscopie soit la procédure ciblée par le présent travail de recherche, l'adaptation à une approche postérieure pour les chirurgies minimalement invasives de la colonne vertébrale serait très simple : la procédure de calibrage resterait la même, tout comme la méthode de recalage, permettant déjà de tenir compte du changement de position du patient entre la chirurgie et l'acquisition des données nécessaires à la reconstruction du modèle préopératoire. Toutefois, il serait probablement nécessaire de développer de nouvelles visualisations adaptées aux différentes opérations à effectuer. Cette adaptation pourrait aller de simplement placer la caméra virtuelle en un endroit différent dans les vues proposées, au développement de nouvelles vues personnalisées pour les procédures par approches postérieures.

D'autres éléments pourraient s'ajouter à la liste de modifications qu'il serait possible d'apporter au système, notamment le retour visuel de l'erreur du système de suivi présenté à la section 5.1.2.1 ainsi que la méthode d'auto-calibrage mentionnée à la section 5.1.2.2. La prochaine section fait en effet état de pistes pour les travaux de recherche à venir.

5.1 Recommandations

Le travail effectué dans le cadre de ce projet de recherche a permis de cerner de nouvelles pistes pour les travaux subséquents.

5.1.1 Éléments de visualisation

Au niveau de la visualisation, cette section discutera de la correction de la distorsion ainsi que de l'ajout d'éléments visuels enrichissant l'augmentation des vues.

5.1.1.1 Correction de la distorsion

L'approche la plus populaire pour la visualisation en ce qui concerne la distorsion de l'image consiste à corriger cette distorsion dans l'image à afficher. Pour la virtualité augmentée, cette approche est la plus logique, car afin d'ajouter l'image réelle à un monde virtuel, il est préférable d'avoir remplacé les pixels de manière à tenir compte de la distorsion causée par le système de lentilles.

Toutefois, pour la réalité augmentée, deux options s'offrent : la première, plus classique, consiste à corriger la distorsion dans l'image, puis à projeter les modèles et les autres informations supplémentaires sur l'image à augmenter. La seconde, moins fréquemment considérée, consiste à conserver l'image distordue, et à distordre de manière équivalente les modèles que l'on projette sur l'image. Cette approche repose sur l'hypothèse que les chirurgiens sont déjà habitués à travailler avec les images distordues, offrant davantage de détails au centre tout en offrant un grand champ de vision. Ainsi, plutôt que d'offrir une vue corrigée qui peut ne pas leur sembler naturelle, on conserve le type de vues auquel ils sont habitués.

Il serait donc intéressant, d'autant plus que cela n'ajoute pas beaucoup de complexité au système, d'offrir les deux options afin de permettre au chirurgien de choisir celle avec laquelle il est le plus confortable, ou du moins afin de permettre de valider laquelle est la plus appropriée pour l'assistance chirurgicale.

5.1.1.2 Effets d'ombrage

La section 1.2.2 mentionne certains indices visuels qui aident le cerveau humain à percevoir la profondeur et à s'orienter dans un espace tridimensionnel. On mentionne entre autres le miroir virtuel qui, bien qu'il soit intéressant, est plus intrusif que d'autres éléments, car il nécessite d'être ajouté (virtuellement) à un instrument. De plus, il n'est pas nécessairement très adapté à la discectomie minimalement invasive, étant donné que dans ce type de chirurgie, l'accès « derrière » les structures d'intérêt, i.e. où on voudrait placer le miroir virtuel, est bloqué.

L'ajout d'effets d'ombrage est toutefois beaucoup plus intéressant. La même section 1.2.2 mentionne différents travaux qui démontrent l'impact de l'ombrage sur la perception de la

profondeur d'une scène. De plus, cet effet peut être intégré d'une manière très subtile à la scène, fournissant ainsi des indices indirects aidant le système de perception du chirurgien.

De nombreuses techniques de rendu, bien souvent provenant de l'industrie du jeu vidéo, existent pour la génération d'ombres. Ces techniques permettant un rendu des ombres en temps réel sont très adaptées à l'outil d'assistance chirurgicale qui doit également être en mesure de réaliser l'affichage en temps réel. Une méthode [87] permettant d'afficher de très jolies ombres floues en temps réel constituerait un choix intéressant d'effet visuel à implémenter dans le système de visualisation.

5.1.2 Ajouts intéressants

Il serait possible d'effectuer certains ajouts au système proposé qui augmenterait son utilité en salle de chirurgie, notamment un retour visuel sur l'erreur du système de suivi et la possibilité d'utiliser un auto-calibrage pour mettre à jour le modèle en cas de modification des paramètres intrinsèques.

5.1.2.1 Retour visuel sur l'erreur du système de suivi

Le système de suivi possède une certaine précision qui a une grande influence sur la précision globale du système. Or, cette précision n'est pas constante. Elle varie en fonction de la configuration du système, c'est-à-dire qu'elle est non seulement dépendante de la conception des marqueurs comme telle, mais également de la position de ces derniers par rapport au système de suivi. En effet, lorsqu'un marqueur se trouve dans un angle très prononcé par rapport au plan des caméras du système de suivi, l'imprécision sur l'identification des points dans les images des caméras du système de suivi entraîne une plus grande imprécision sur la localisation 3D des points. De plus, la forme elle-même des marqueurs entraîne des niveaux de précisions variables sur l'orientation du marqueur. En effet, les axes plus courts sont sujets à une plus grande erreur au niveau de l'orientation étant donné que l'erreur de positionnement 3D est moins absorbée par la distance entre les points.

Ainsi, la précision globale du système varie dans le temps selon les positions des marqueurs. Il serait intéressant de pouvoir donner un retour visuel au chirurgien quant à la précision du système dans sa configuration courante afin qu'il puisse connaître le degré de confiance qu'il peut avoir dans l'information fournie par le système.

Certains travaux effectuent déjà ce genre d'études. Sielhorst et al. [88] étudie l'erreur en fonction de la configuration physique du monde et l'affiche en temps réel. On utilise des ellipsoïdes représentant un intervalle de confiance donné sur la position de certains éléments.

L'ajout de ce retour visuel permettrait donc au chirurgien d'apprécier visuellement la précision courante du système et lui permettrait ainsi de s'assurer que le système est toujours digne de confiance. Cette fonctionnalité constituerait donc un ajout intéressant à l'outil proposé.

5.1.2.2 Mise à jour des paramètres intrinsèques par auto-calibrage

Bien que les paramètres extrinsèques du modèle de formation de l'image soient mis à jour en temps réel grâce au système de suivi, les paramètres intrinsèques sont quant à eux obtenus grâce à la procédure de calibrage et considérés comme constants par la suite. Or, certaines manipulations effectuées par le chirurgien, comme une modification du facteur de zoom ou une mise au point, entraîne une variation des paramètres intrinsèques.

Girard [9] s'est attaqué à ce problème et propose une méthode d'auto-calibrage basée sur une rotation pure permettant ainsi de calculer les nouvelles distances focales. L'intégration de cette technique à l'outil permettrait de mettre à jour les paramètres intrinsèques du modèle sans avoir à recommencer le calibrage du système effectué avant la chirurgie. La méthode d'auto-calibrage permet en effet au chirurgien de réaliser cette dernière sans avoir à retirer l'endoscope de l'intérieur du corps du patient, ce qui constitue un avantage au niveau du protocole chirurgical.

Finalement, la méthode de Girard nécessite de calculer certains paramètres intrinsèques, notamment les paramètres de distorsion, pour chacun des niveaux discrets de zoom. Il serait donc aisé de calculer également les distances focales et tous les autres paramètres pour les niveaux discrets de zoom, éliminant ainsi la modification du niveau de zoom comme opération problématique. Comme il ne resterait que la mise au point comme opération modifiant les paramètres intrinsèques, il serait intéressant de vérifier dans quelle mesure une manipulation normale de mise au point modifie numériquement les paramètres intrinsèques.

RÉFÉRENCES

- [1] S. L. Weinstein, D. C. Zavala, et I. V. Ponseti, "Idiopathic scoliosis : long-term follow-up and prognosis in untreated patients," *The Journal of Bone & Joint Surgery*, vol. 63, no. 5, pp. 702-712, 1981.
- [2] W. K. Payne, J. W. Ogilvie, M. D. Resnick, R. L. Kane, E. E. Transfeldt, et R. W. Blum, "Does scoliosis have a psychological impact and does gender make a difference?," *Spine*, vol. 22, no. 12, pp. 1380-1384, 1997.
- [3] F. Benboujja, "Suivi automatique d'instruments dans les séquences d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009.
- [4] M. J. Mack, "Minimally Invasive and Robotic Surgery," *The Journal of the American Medical Association*, vol. 285, no. 5, pp. 568-572, 2001.
- [5] C. A. Dickman, D. J. Rosenthal, et N. I. Perin, *Thoracoscopic Spine Surgery*, New York: Thieme, 1999.
- [6] G. D. Picetti et D. Pan, "Thoracoscopic techniques for the treatment of scoliosis," *Child's Nervous System*, vol. 20, no. 11-12, pp. 802-810, 2004.
- [7] C.-A. St-Pierre, "Détection et correction des réflexions spéculaires dans les séquences d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2005.
- [8] L. Windish, "Bayesian differentiation of multi-scale line structures of model-free instrument," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2005.
- [9] F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009.
- [10] C. Chevretils, F. Cheriet, C. E. Aubin, et G. Grimard, "Texture Analysis for Automatic Segmentation of Intervertebral Disks of Scoliotic Spines From MR Images," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 4, pp. 608-620, 2009.
- [11] C. Chevretils, "2D/3D registration of pre-operative MRI models with a single radiographic image : feasibility study," in *8th International Research Society of Spinal Deformities (IRSSD)*, Montréal, 2010.
- [12] W. E. Lorensen et H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 163-169, 1987.

- [13] C. Bichlmeier, S. M. Heining, M. Feuerstein, et N. Navab, "The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 9, pp. 1498-1510, 2009.
- [14] C. Bichlmeier, F. Wimmer, S. M. Heining, et N. Navab, "Contextual Anatomic Mimesis Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality," in *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 1-10.
- [15] K. T. Foley et M. M. Smith, "Image-guided spine surgery," *Neurosurgery clinics of North America*, vol. 7, no. 2, pp. 171-186, 1996.
- [16] S. Lavallée, P. Sautot, J. Troccaz, P. Cinquin, et P. Merloz, "Computer-assisted spine surgery: A technique for accurate transpedicular screw fixation using CT data and a 3-D optical localizer," *Journal of Image Guided Surgery*, vol. 1, no. 1, pp. 65-73, 1995.
- [17] Y. R. Rampersaud et K. T. Foley, "Image-guided spinal surgery," *Operative Techniques in Orthopaedics*, vol. 10, no. 1, pp. 64-68, 2000.
- [18] H. F. Reinhardt et H. Landolt, "CT-Guided "Real Time" Stereotaxy," *Acta Neurochirica Supplementum*, vol. 46, pp. 107-108, 1989.
- [19] J. Traub, T. Sielhorst, S. M. Heining, et N. Navab, "Advanced Display and Visualization Concepts for Image Guided Surgery," *Journal of Display Technology*, vol. 4, no. 4, pp. 483-490, 2008.
- [20] J. Dansereau et I. A. F. Stokes, "Measurements of the three-dimensional shape of the rib cage," *Journal of Biomechanics*, vol. 21, no. 11, pp. 893-901, 1988.
- [21] J. Novosad, F. Cheriet, Y. Petit, et H. Labelle, "Three-dimensional (3-D) reconstruction of the spine from a single X-ray image and prior vertebra models," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 9, pp. 1628-1639, 2004.
- [22] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, et N. Ayache, "Geometric Variability of the Scoliotic Spine Using Statistics on Articulated Shape Models," *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 557-568, 2008.
- [23] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, et N. Ayache, "Articulated Spine Models for 3-D Reconstruction From Partial Radiographic Data," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 11, pp. 2565-2574, 2008.
- [24] F. Cheriet, C. Laporte, S. Kadoury, H. Labelle, et J. Dansereau, "A Novel System for the 3-D Reconstruction of the Human Spine and Rib Cage From Biplanar X-Ray Images," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 7, pp. 1356-1358, 2007.
- [25] L. Humbert, J. A. De Guise, B. Aubert, B. Godbout, et W. Skalli, "3D reconstruction of the spine from biplanar X-rays using parametric models based on transversal and

- longitudinal inferences," *Medical Engineering & Physics*, vol. 31, no. 6, pp. 681-687, 2009.
- [26] P. Paul, O. Fleig, et P. Jannin, "Augmented virtuality based on stereoscopic reconstruction in multimodal image-guided neurosurgery: methods and performance evaluation," *IEEE Transactions on Medical Imaging*, vol. 24, no. 11, pp. 1500-1511, 2005.
 - [27] F. Sauer, A. Khamene, B. Bascle, et G. J. Rubino, "A Head-Mounted Display System for Augmented Reality Image Guidance: Towards Clinical Evaluation for iMRI-guided Neurosurgery," in *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2001, pp. 707-716.
 - [28] C. Chevretils, F. Chérier, G. Grimard, et C.-E. Aubin, "Watershed Segmentation of Intervertebral Disk and Spinal Canal from MRI Images," in *Image Analysis and Recognition*, 2007, pp. 1017-1027.
 - [29] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America*, vol. 4, no. 4, pp. 629-642, 1987.
 - [30] D. Dey, D. G. Gobbi, P. J. Slomka, K. J. M. Surry, et T. M. Peters, "Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: creating stereoscopic panoramas," *IEEE Transactions on Medical Imaging*, vol. 21, no. 1, pp. 23-30, 2002.
 - [31] J. S. Hott, N. Theodore, J. D. Klopfenstein, S. Papadopolous, C. A. Dickman, et V. K. H. Sonntag, "Image guidance in spinal surgery for trauma: Iso-C 3D," *Operative Techniques in Neurosurgery*, vol. 7, no. 1, pp. 42-47, 2004.
 - [32] T. Renkawitz, M. Wegner, S. Gneiting, M. Haimerl, E. Sendtner, T. Kalteis, et J. Grifka, "Experimental validation of a pinless femoral reference array for computer-assisted hip arthroplasty," *Journal of Orthopaedic Research*, 2009. [En ligne]. Disponible: <http://dx.doi.org/10.1002/jor.21039>. [Consulté le 2010/03/05].
 - [33] H. Abbasi, S. Chin, D. Kim, G. Steinberg, et R. Shahidi, "Computerized lateral endoscopic approach to spinal pathologies," in *International Congress Series*, vol. 1230, 2001, pp. 240-247.
 - [34] M. Feuerstein, S. Wildhirt, R. Bauernschmitt, et N. Navab, "Automatic Patient Registration for Port Placement in Minimally Invasive Endoscopic Surgery," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005*, 2005, pp. 287-294.
 - [35] M. Nakamoto, N. Aburaya, K. Konishi, Y. Sato, I. Yoshino, M. Hashizume, et S. Tamura, "Estimation of intraoperative lung deformation for computer assisted thoracoscopic surgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 1, pp. 273-275, 2006.

- [36] M. Nakamoto, N. Aburaya, Y. Sato, K. Konishi, I. Yoshino, M. Hashizume, et S. Tamura, "Thoracoscopic Surgical Navigation System for Cancer Localization in Collapsed Lung Based on Estimation of Lung Deformation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, 2007, pp. 68-76.
- [37] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, vol. 13, no. 2, pp. 119-152, 1994.
- [38] A. Roche, X. Pennec, G. Malandain, et N. Ayache, "Rigid Registration of 3D Ultrasound with MR Images : a New Approach Combining Intensity and Gradient Information," *IEEE Transactions on Medical Imaging*, vol. 20, no. 10, pp. 1038-1049, 2001.
- [39] S. Lavallée, R. Szeliski, et L. Brunie, "Matching 3-D smooth surfaces with their 2-D projections using 3-D distance maps," in *Geometric Reasoning for Perception and Action*, 1993, pp. 217-238.
- [40] J. Weese, T. M. Buzug, G. P. Penney, et P. Desmedt, "2D/3D registration and motion tracking for surgical interventions," *Philips Journal of Research*, vol. 51, no. 2, pp. 299-316, 1998.
- [41] B. Brendel, S. Winter, A. Rick, M. Stockheim, et H. Ermert, "Bone registration with 3D CT and ultrasound data sets," *International Congress Series*, vol. 1256, pp. 426-432, 2003.
- [42] S. Winter, B. Brendel, et C. Igel, "Registration of bone structures in 3D ultrasound and CT data: Comparison of different optimization strategies," *International Congress Series*, vol. 1281, pp. 242-247, 2005.
- [43] K. Hensel, S. Hold, S. Winter, C. Dekomien, M. Mildenstein, B. Brendel, H. Ermert, et G. Schmitz, "Ultrasound Based Navigation System for Minimal Invasive Surgery at the Lumbar Spine within OrthoMIT," in *Advances in Medical Engineering*, 2007, pp. 224-229.
- [44] S. Winter, I. Pechlivanis, C. Dekomien, C. Igel, et K. Schmieder, "Toward Registration of 3D Ultrasound and CT Images of the Spine in Clinical Praxis: Design and Evaluation of a Data Acquisition Protocol," *Ultrasound in Medicine & Biology*, vol. 35, no. 11, pp. 1773-1782, 2009.
- [45] T. M. Peters, "Image-guided surgery: From X-rays to Virtual Reality," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 4, no. 1, pp. 27 - 57, 2001.
- [46] T. M. Peters, "Image-guidance for surgical procedures," *Physics in Medecine and Biology*, vol. 51, no. 14, pp. R505-R540, 2006.

- [47] T. Sielhorst, M. Feuerstein, et N. Navab, "Advanced Medical Displays: A Literature Review of Augmented Reality," *Journal of Display Technology*, vol. 4, no. 4, pp. 451-467, 2008.
- [48] M. Blackwell, C. Nikou, A. DiGioia, et T. Kanade, "An image overlay system for medical data visualization," in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, 1998, pp. 232-240.
- [49] M. Burns, M. Haidacher, W. Wein, I. Viola, et M. E. Gröller, "Feature Emphasis and Contextual Cutaways for Multimodal Medical Visualization," in *Data Visualization - EuroVis 2007, Norköpping, Sweden*, 2007, pp. 275-282.
- [50] J. E. Cutting et P. M. Vishton, "Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth," in *Perception of Space and Motion*, W. Epstein et S. Rogers, Éd., 1^e éd., San Diego: Academic Press, 1995, pp. 69-117.
- [51] D. Kersten, P. Mamassian, et D. C. Knill, "Moving Cast Shadows and the Perception of Relative Depth," Max Planck Institute for Biological Cybernetics, Tübingen, Rapport technique 6, 1994. [En ligne]. Disponible: <http://www.kyb.mpg.de/publications/pdfs/pdf1461.pdf>. [Consulté le 2010/03/24].
- [52] P. Mamassian, D. C. Knill, et D. Kersten, "The perception of cast shadows," *Trends in Cognitive Sciences*, vol. 2, no. 8, pp. 288-295, 1998.
- [53] M. Nicolaou, A. James, B. P. L. Lo, A. Darzi, et G.-Z. Yang, "Invisible Shadow for Navigation and Planning in Minimal Invasive Surgery," in *Proceedings of the 8th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2005, pp. 25-32.
- [54] C. Bichlmeier, S. M. Heining, M. Rustae, et N. Navab, "Laparoscopic Virtual Mirror for Understanding Vessel Structure Evaluation Study by Twelve Surgeons," in *IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 125-128.
- [55] D. Stoyanov, A. Darzi, et G.-Z. Yang, "Dense 3D Depth Recovery for Soft Tissue Deformation During Robotic Assisted Laparoscopic Surgery," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*, 2004, pp. 41-48.
- [56] A. H. Ahmed et A. A. Farag, "Shape from Shading Under Various Imaging Conditions," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-8.
- [57] E. Prados et O. Faugeras, "Shape from shading: a well-posed problem?," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, vol. 2, 2005, pp. 870-877.

- [58] C. Wu, S. Narasimhan, et B. Jaramaz, "A Multi-Image Shape-from-Shading Framework for Near-Lighting Perspective Endoscopes," *International Journal of Computer Vision*, vol. 86, no. 2, pp. 211-228, 2010.
- [59] C. Wu, S. G. Narasimhan, et B. Jamaraz, "Shape-from-Shading under Near Point Lighting and Partial views for Orthopedic Endoscopy," in *Workshop on Photometric Analysis For Computer Vision (PACV'07), in conjunction with ICCV'07*, 2007.
- [60] L. Yang et J.-q. Han, "3D shape reconstruction of medical images using a perspective shape-from-shading method," *Measurement Science and Technology*, vol. 19, no. 6, pp. 065502, 2008.
- [61] S. M. Heining, S. Wiesner, E. Euleur, et N. Navab, "Pedicule screw placement under video-augmented fluoroscopic control : first clinical application in a cadaver study," *International Journal of Computer Assisted Radiology and Surgery*, vol. 1, no. 1, pp. 189-190, 2006.
- [62] E. Jank, A. Rose, S. Huth, C. Trantakis, W. Korb, G. Strauss, J. Meixensberger, et J. Krueger, "A new fluoroscopy based navigation system for milling procedures in spine surgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 1, no. 1, pp. 196-198, 2006.
- [63] C. W. Kim, Y.-P. Lee, W. Taylor, A. Oygur, et W. Kyung Kim, "Use of navigation-assisted fluoroscopy to decrease radiation exposure during minimally invasive spine surgery," *The Spine Journal*, vol. 8, no. 4, pp. 584-590, 2008.
- [64] N. Haberland, K. Ebmeier, R. Hliscs, J. P. Grunewald, et R.-L. Kalff, "Intraoperative CT in image-guided surgery of the spine," *MedicaMundi*, vol. 43, no. 4, pp. 24-31, 1999.
- [65] R. Hartley et A. Zisserman, *Multiple View Geometry in Computer Vision*, 2^e éd.: Cambridge University Press, 2004.
- [66] R. Y. 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*, vol. 3, no. 4, pp. 323-344, 1987.
- [67] B. K. P. Horn, "Tsai's camera calibration method revisited," *Tsai's camera calibration method revisited*, 2000. [En ligne]. Disponible: http://people.csail.mit.edu/bkph/articles/Tsai_Revisited.pdf. [Consulté le 2009/10/07].
- [68] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431-441, 1963.
- [69] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 1, 1999, pp. 666-673 vol.1.

- [70] D. C. Brown, "Close-range camera calibration," *Photogrammetric Engineering*, vol. 37, no. 8, pp. 855-866, 1971.
- [71] G. Bradski et A. Kaehler, *Learning OpenCV*, Sebastopol: O'Reilly Media, 2008.
- [72] R. Miranda-Luna, W. C. P. M. Blondel, C. Daul, Y. Hernandez-Mier, R. Posada, et D. Wolf, "A simplified method of endoscopic image distortion correction based on grey level registration," in *International Conference on Image Processing*, vol. 5, 2004, pp. 3383-3386.
- [73] Y. Kosugi, E. Watanabe, J. Goto, T. Watanabe, S. Yoshimoto, K. Takakura, et J. Ikebe, "An articulated neurosurgical navigation system using MRI and CT images," *IEEE Transactions on Biomedical Engineering*, vol. 35, no. 2, pp. 147-152, 1988.
- [74] D. W. Roberts, E. M. Friets, J. W. Strohbehn, et T. Nakajima, "The Sonic Digitizing Microscope," in *Interactive Image-Guided Neurosurgery*, R. J. Maciunas, Éd.: American Association of Neurological Surgeons, 1993.
- [75] F. Poulin et L. P. Amiot, "Interference during the use of an electromagnetic tracking system under OR conditions," *Journal of Biomechanics*, vol. 35, no. 6, pp. 733-737, 2002.
- [76] D. Simon, "Intra-Operative Position Sensing and Tracking Devices," in *First Joint CVRMed / MRCAS Conference*, 1997, pp. 62-64.
- [77] C. T. Inc., "MicronTracker Developer's Manual, MTC 3.0," Rapport technique, 2009.
- [78] T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, et S. Tamura, "Development of a camera model and calibration procedure for an oblique-viewing endoscope," *Electronics and Communications in Japan (Part II: Electronics)*, vol. 88, no. 2, pp. 19-31, 2005.
- [79] C. Wu et B. Jaramaz, "An easy calibration for oblique-viewing endoscopes," in *IEEE International Conference on Robotics and Automation*, 2008, pp. 1424-1429.
- [80] S. De Buck, F. Maes, A. D'Hoore, et P. Suetens, "Evaluation of a Novel Calibration Technique for Optically Tracked Oblique Laparoscopes," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, 2007, pp. 467-474.
- [81] I. Corporation, "G2X specifications sheet," 2004. [En ligne]. Disponible: http://www.3d-microscribe.com/MSG2_0704_V1.pdf. [Consulté le 2010/05/05].
- [82] C. T. Inc., "Products," 2009. [En ligne]. Disponible: http://www.clarontech.com/measurement_products.php. [Consulté le 2010/05/05].
- [83] S. A. Nicolau, L. Goffin, et L. Soler, "A low cost and accurate guidance system for laparoscopic surgery: validation on an abdominal phantom," in *Proceedings of the*

ACM symposium on Virtual reality software and technology, Monterey, CA, USA, 2005, pp. 124-133.

- [84] R. Y. Tsai et R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 345-358, 1989.
- [85] W. Sun et J. Cooperstock, "An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques," *Machine Vision and Applications*, vol. 17, no. 1, pp. 51-67, 2006.
- [86] T. Sielhorst, "New Methods for Medical Augmented Reality," Technische Universität München (TUM), 2008.
- [87] Y. Uralsky, "Chapter 17. Efficient Soft-Edged Shadows Using Pixel Shader Branching," in *GPU gems 2 : programming techniques for high-performance graphics and general-purpose computation*, M. Pharr et R. Fernando, Eds., 1^e éd.: Addison-Wesley, 2005. Disponible: http://http.developer.nvidia.com/GPUGems2/gpugems2_chapter17.html. [Consulté le 2010/05/20].
- [88] T. Sielhorst, M. Bauer, O. Wenisch, G. Klinker, et N. Navab, "Online Estimation of the Target Registration Error for n -Ocular Optical Tracking Systems," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, 2007, pp. 652-659.
- [89] J. Weng, P. Cohen, et M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965-980, 1992.
- [90] J. P. Helferty, Z. Chao, G. McLennan, et W. E. Higgins, "Videoendoscopic distortion correction and its application to virtual guidance of endoscopy," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 605-617, 2001.
- [91] W. E. Smith, N. Vakil, et S. A. Maislin, "Correction of distortion in endoscope images," *IEEE Transactions on Medical Imaging*, vol. 11, no. 1, pp. 117-122, 1992.
- [92] D. C. Brown, "Decentering Distortion of Lenses," *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444-462, 1966.
- [93] T. A. Clarke et J. G. Fryer, "The Development of Camera Calibration Methods and Models," *Photogrammetric Record*, vol. 16, no. 91, pp. 51-66, 1998.
- [94] H. Haneishi, Y. Yagihashi, et Y. Miyake, "A new method for distortion correction of electronic endoscope images," *IEEE Transactions on Medical Imaging*, vol. 14, no. 3, pp. 548-555, 1995.

- [95] W. H. Press, S. A. Teukolsky, W. T. Vetterling, et B. P. Flannery, *Numerical Recipes in C : The Art of Scientific Computing*, 2^e éd.: Cambridge University Press, 1992.
- [96] C. Gramkow, "On Averaging Rotations," *International Journal of Computer Vision*, vol. 42, no. 1, pp. 7-16, 2001.
- [97] K. Vijayan Asari, S. Kumar, et D. Radhakrishnan, "A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation," *IEEE Transactions on Medical Imaging*, vol. 18, no. 4, pp. 345-354, 1999.

ANNEXE I DÉVELOPPEMENT MATHÉMATIQUE DU MODÈLE DE FORMATION DE L'IMAGE

Le modèle de formation de l'image permet de comprendre la manière dont une scène 3D quelconque visualisée à l'aide d'une caméra est transformée pour former l'image 2D. D'un point de vue mathématique, il s'agit d'exprimer la transformation (ou la série de transformations) qui permet de transformer un point 3D de l'espace visualisé en un point 2D sur l'image produite.

La présente annexe décrit le développement mathématique qui permet d'obtenir le modèle de la formation d'images à l'aide d'une caméra perspective.

I.1 Modèles de projection

La formation d'une image à l'aide d'une caméra consiste essentiellement à projeter la scène 3D sur un plan 2D, le plan image. Il existe plusieurs types de projections, notamment la projection orthographique, la projection perspective faible et la projection perspective.

I.1.1 Projection orthographique

La projection orthographique (Figure I.1) considère que la scène est projetée perpendiculairement au plan image, ce qui a pour conséquence que deux objets de mêmes dimensions dans la scène auront également les mêmes dimensions dans l'image, peu importe la distance entre ces objets et la caméra. Ce modèle est utilisé davantage pour générer des images non photo-réalistes, tels des plans ou des schémas techniques. Il correspond à une caméra dont la distance focale (définie plus loin) est infinie et n'est donc pas vraiment utilisé pour modéliser une caméra réelle.

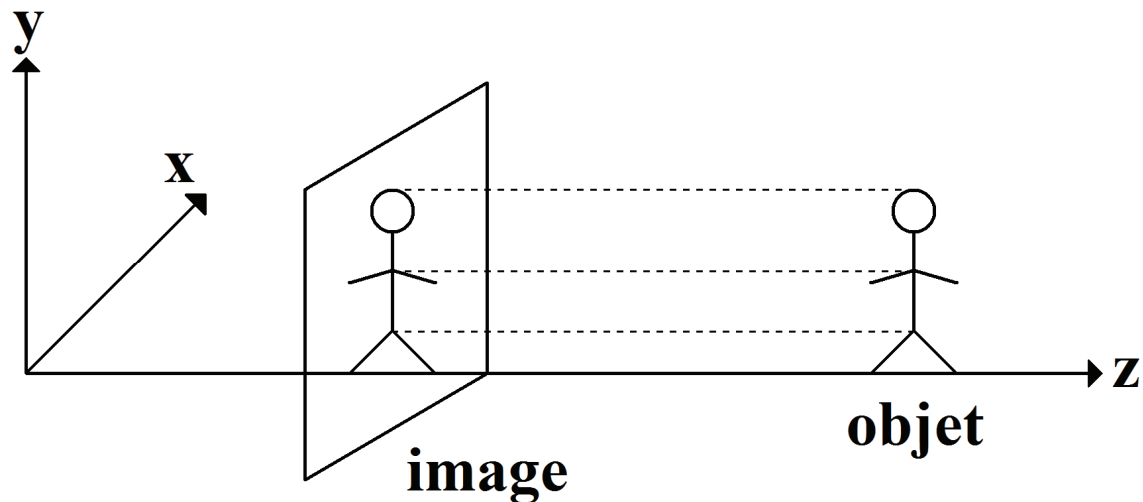


Figure I.1. Projection orthographique

I.1.2 Projection perspective faible

La projection perspective faible (Figure I.2) constitue une approximation linéaire d'une véritable projection perspective (voir la section I.1.3) qui est non linéaire. Elle consiste en une projection orthographique (transformation linéaire) suivie d'une mise à l'échelle (transformation linéaire). Cette approximation est valide si l'« épaisseur » de l'objet, c'est-à-dire sa dimension sur l'axe de projection, est petite par rapport à la distance de la caméra. En effet, l'approximation de la projection orthogonale correspond à considérer que tous les points de l'objet sont situés dans le même plan, ce qui peut être acceptable si l'objet est « mince ». Comme cette approximation n'apporte pas vraiment d'avantages dans le cas de la fusion multimodale et que les hypothèses la justifiant ne sont pas respectées dans le cas des images endoscopiques considérées, ce type de projection n'est également pas vraiment utilisé pour modéliser la formation des images endoscopiques.

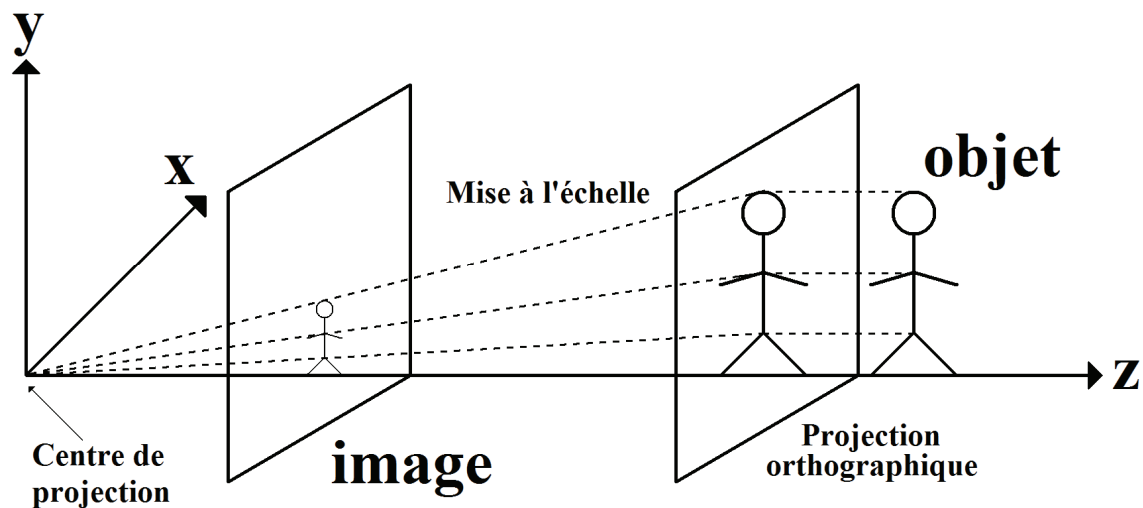


Figure I.2. Projection perspective faible

I.1.3 Projection perspective

La projection perspective (Figure I.3), souvent appelée modèle de la caméra à sténopé, est plus réaliste que les deux précédentes quand à la formation d'images à l'aide d'une véritable caméra. Elle considère que tous les rayons lumineux passent par un trou infinitésimalement petit, le sténopé, avant d'atteindre le plan image situé derrière celui-ci, produisant ainsi une image inversée. Ce type de projection est universellement utilisé pour représenter la formation d'image à l'aide d'une caméra réelle et sera donc ici mathématiquement développé [65].

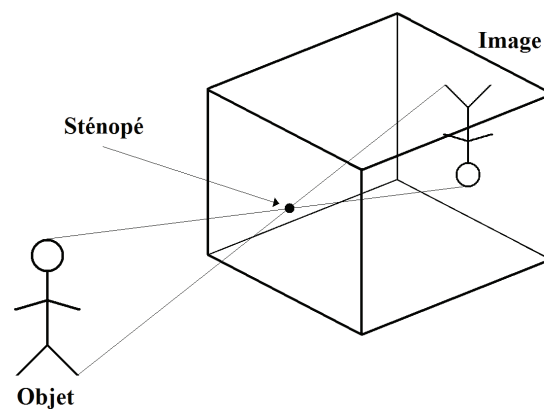


Figure I.3. Projection perspective

Tout d'abord, afin de développer le modèle mathématique, on définit arbitrairement, sans perte de généralité comme il le sera montré plus loin, un système de coordonnées dont l'origine est le centre de projection et l'axe Z , aussi connu sous le nom d'axe optique, est perpendiculaire au plan image. Ce dernier est situé à une distance f du centre de projection, f étant la distance focale de la projection. Le point d'intersection entre l'axe optique et le plan de projection constitue l'origine du repère 2D défini dans le plan image et est nommé point principal ou centre optique.

Les équations de la projection perspective peuvent être dérivées à partir de l'image inversée, tel que représenté à la Figure I.3, dans lequel cas les axes X et Y du repère 2D défini dans le plan image doivent être inversés par rapport aux axes X et Y du repère 3D défini à la caméra. Toutefois, on ramène généralement le plan de projection devant la caméra afin d'éviter le problème d'inversion d'image tel que représenté sur la Figure I.4.

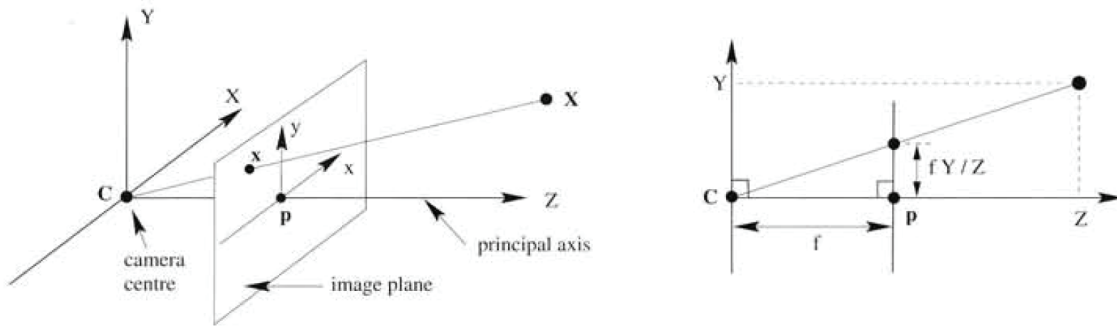


Figure I.4. Dérivation des équations de projection perspective (tirée de R. Hartley et A. Zisserman, *Multiple View Geometry in Computer Vision*, 2^e éd.: Cambridge University Press, 2004. ©2004 Cambridge University Press. Reproduit avec permission.)

Ainsi, si on nomme (x, y) la projection sur le plan image du point (X, Y, Z) , par la géométrie des triangles semblables, on obtient :

$$\begin{aligned} \frac{x}{f} &= \frac{X}{Z} \Leftrightarrow x = f \frac{X}{Z} \\ \frac{y}{f} &= \frac{Y}{Z} \Leftrightarrow y = f \frac{Y}{Z} \end{aligned} \quad (\text{I.1})$$

En utilisant les coordonnées homogènes, on peut exprimer cette équation sous forme matricielle :

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (\text{I.2})$$

où la valeur de λ est directement Z . Ceci constitue le modèle simplifié de la caméra à sténopé.

I.2 Généralisation du modèle de projection

Le modèle de projection introduit précédemment peut être généralisé par l'ajout d'autres paramètres, qui seront ici présentés, pour tenir compte de facteurs non mentionnés jusqu'à maintenant mais influençant quand même la formation de l'image.

I.2.1 Rapport d'aspect des pixels

L'échelle du système d'axes présent dans le plan image a jusqu'à maintenant été considérée comme arbitraire. Lors du traitement d'images numériques, il est généralement pratique d'utiliser le pixel comme unité de base afin de déterminer les coordonnées de points situés dans l'image. Or, lors de l'acquisition de l'image par une caméra numérique, les lignes du capteur sont souvent ré-échantillonnées, ce qui fait que les pixels ne correspondent plus nécessairement aux éléments du capteur [67]. Ce phénomène introduit un facteur de mise à l'échelle qui fait que, dans les caméras réelles, les pixels ne sont pas toujours parfaitement carrés, comme on peut le voir dans la Figure I.5.

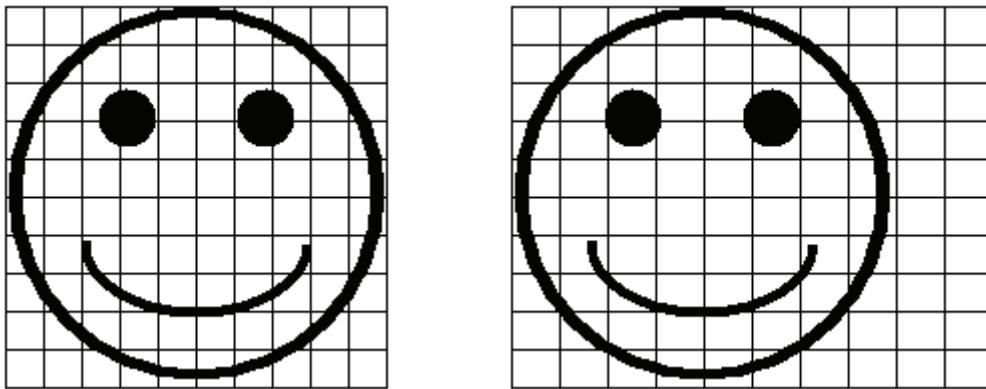


Figure I.5. Effet des pixels non carrés : $\alpha = 1$ à gauche, $\alpha \neq 1$ à droite

En effet, on voit que les coordonnées (x, y) d'un point dans la première image ne correspondent pas au même point dans la seconde image, à cause de la mise à l'échelle qui se produit sur un des axes pour tenir compte du rapport d'aspect des pixels.

Ce facteur peut être ajouté au modèle par l'introduction d'un paramètre α modélisant une mise à l'échelle de la matrice de projection sur un seul des axes :

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & \alpha f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{I.3})$$

Ainsi, au lieu d'avoir une seule distance focale, un second degré de liberté est ajouté. Ce dernier peut être représenté à l'aide d'une distance focale unique f ainsi qu'un rapport d'aspect α , ou alors à l'aide de deux distances focales f_x et f_y , avec :

$$\begin{cases} f_x = f \\ f_y = \alpha f \end{cases} \Leftrightarrow \begin{cases} f = f_x \\ \alpha = \frac{f_y}{f_x} \end{cases} \quad (\text{I.4})$$

I.2.2 Facteur d'obliquité des pixels

Le système d'axes relié au plan image a jusqu'à présent été considéré comme orthogonal. Or, encore une fois, lorsque l'on utilise les coordonnées en pixels comme système de coordonnées, il est possible de retrouver dans les caméras réelles une grille de pixels qui peut être légèrement oblique, comme on peut le voir dans la Figure I.6.

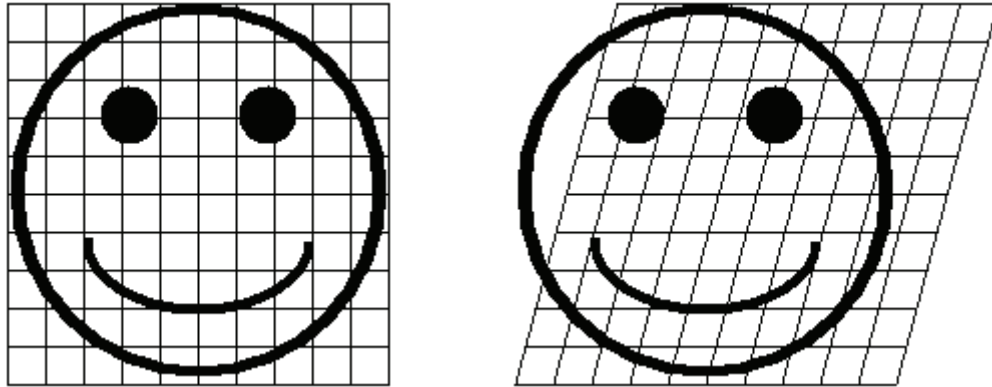


Figure I.6. Effet des pixels obliques : $s = 0$ à gauche, $s \neq 0$ à droite

En effet, on voit que les coordonnées (x, y) d'un point dans la première image ne correspondent pas au même point dans la seconde image, à cause du déplacement causé par l'obliquité des pixels.

Ce facteur peut être modélisé par l'introduction d'un paramètre s modélisant un cisaillement :

$$\begin{bmatrix} 1 & s_{fact} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & s_{fact}f_y & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & s & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (I.5)$$

Ainsi, on ajoute un facteur s (plutôt que de conserver $s_{fact}f_y$) à l'intérieur de la matrice de projection afin de pouvoir modéliser l'obliquité des pixels.

I.2.3 Centre optique

Le système d'axes présent dans le plan image a jusqu'à maintenant été considéré comme possédant comme origine le centre optique. Or, encore une fois, lorsque l'on utilise les coordonnées en pixels comme système de coordonnées, le centre optique ne correspond généralement pas à l'origine du système, qui est généralement le coin inférieur ou supérieur gauche de l'image, tel que représenté dans la Figure I.7.

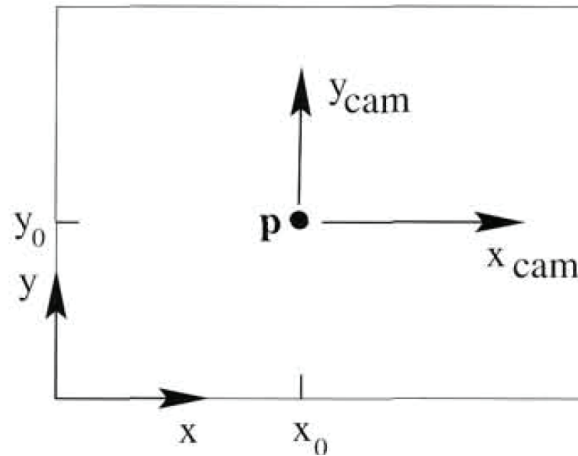


Figure I.7. Centre optique par rapport à l'origine (tirée de R. Hartley et A. Zisserman, *Multiple View Geometry in Computer Vision*, 2^e éd.: Cambridge University Press, 2004. ©2004 Cambridge University Press. Reproduit avec permission.)

Ainsi, la matrice de projection perspective définie précédemment projette les points dans un repère (x_{cam}, y_{cam}) défini par rapport au centre optique (x_0, y_0) . Il suffit donc de transformer les points projetés dans le repère du centre optique pour qu'ils soient dans le repère arbitraire choisi, ce qui peut être fait par une simple translation correspondant à la position du centre optique dans le repère choisi :

$$K = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & s & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{I.6})$$

Ainsi, on ajoute les coordonnées du centre optique à la matrice de projection perspective afin de pouvoir choisir une origine arbitraire pour le système d'axes du plan de projection.

On retrouve donc 5 paramètres dans la matrice de projection perspective K . Ces paramètres sont généralement connus sous le nom de paramètres intrinsèques, c'est-à-dire propres à la caméra elle-même et indépendants de la position de cette dernière dans un repère quelconque. Cette position correspond plutôt aux paramètres extrinsèques constituant une autre partie du modèle de formation de l'image qui sera détaillée dans la prochaine section.

I.3 Pose de la caméra

Jusqu'à présent, les points 3D éventuellement projetés sur le plan image de la caméra sont exprimés dans un système de coordonnées défini à partir de la caméra. Or, afin de donner plus de flexibilité au modèle, il est plus pratique que les points soient définis dans un autre référentiel correspondant au monde 3D visualisé à l'aide de la caméra. Il est possible de passer du référentiel monde au référentiel de la caméra à l'aide d'une transformation rigide, c'est-à-dire une rotation et une translation.

Il est à noter que l'échelle, c'est-à-dire l'unité de mesure, des axes des deux systèmes de coordonnées est généralement la même. On obtient cette correspondance tout simplement en définissant l'échelle du système de la caméra en fonction de celle du système du monde. Si l'échelle était différente, on devrait ajouter une mise à l'échelle (homothétie) à la transformation rigide reliant les repères. Cette mise à l'échelle serait ensuite compensée par un changement de la valeur de la distance focale, définie en unités du repère de la caméra. Cette mise à l'échelle ensuite annulée par une modification de la distance focale est donc futile, d'où l'utilisation de l'échelle du monde lors de la définition du repère de la caméra, évitant ainsi ces transformations inutiles et permettant d'exprimer la distance focale dans l'unité des axes du monde.

Ainsi, il est donc possible d'exprimer la pose¹² (rotation et translation) de la caméra dans le repère monde à l'aide d'une translation $T_{w \leftarrow c}$, qui correspond à la position de la caméra exprimée dans le repère monde, et d'une rotation $R_{w \leftarrow c}$. Ainsi, la transformation permettant de convertir un point exprimé dans le référentiel de la caméra en un point exprimé dans le référentiel monde est la composée $E_{w \leftarrow c}$ de ces deux transformations :

$$E_{w \leftarrow c} = T_{w \leftarrow c} R_{w \leftarrow c} = [R_{w \leftarrow c} \mid T_{w \leftarrow c}] \quad (I.7)$$

Cette transformation est exprimée à l'aide d'une matrice 3x4, où les trois premières colonnes forment une matrice orthonormale correspondant donc à la rotation, et la dernière colonne correspond à la translation.

Toutefois, lorsque les points sont projetés, ils doivent passer de l'espace monde, dans lequel ils sont naturellement exprimés, à l'espace caméra, à partir duquel ils peuvent être projetés grâce au modèle développé dans les sous-sections précédentes. Ainsi, on cherche la transformation inverse de $E_{w \leftarrow c}$, soit $E_{c \leftarrow w}$:

$$\begin{aligned} E_{c \leftarrow w} &= (E_{w \leftarrow c})^{-1} = (T_{w \leftarrow c} R_{w \leftarrow c})^{-1} = (R_{w \leftarrow c})^{-1} (T_{w \leftarrow c})^{-1} = (R_{c \leftarrow w}) (-T_{w \leftarrow c}) \\ &= [R_{c \leftarrow w} \mid -R_{c \leftarrow w} T_{w \leftarrow c}] = \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix} \end{aligned} \quad (I.8)$$

Ces paramètres, soit la pose de la caméra exprimée par une translation et une rotation représentant son orientation, sont généralement connus sous le nom de paramètres extrinsèques, c'est-à-dire extérieurs à la caméra elle-même et dépendants davantage de sa position dans un monde quelconque. La translation possédant 3 degrés de liberté, tout comme la rotation, on retrouve donc 6 degrés de liberté dans l'expression de la matrice de transformation rigide $E_{c \leftarrow w}$.

¹² La pose correspond à une position (donc une translation) et une orientation (donc une rotation). Elle peut donc être exprimée comme une transformation rigide, elle-même composée d'une translation et d'une rotation. Le terme « pose » sera donc utilisé pour désigner cette combinaison de position et d'orientation.

Le modèle général d'une caméra à sténopé, possédant 11 degrés de liberté, a donc été défini et peut être résumé par :

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = KE_{c \leftarrow w} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (I.9)$$

ANNEXE II DÉFORMATION DE LENTILLE

La majorité des systèmes de lentilles de caméra induisent un certain niveau de distorsion, souvent non négligeable, dans les images qu'ils produisent. En effet, les lentilles elles-mêmes n'ont souvent pas exactement la forme théorique parfaite produisant l'effet désiré de la lentille. On fabrique généralement des lentilles sphériques, qui constituent une approximation acceptable de la forme parabolique idéale, beaucoup plus complexe à usiner. De plus, lors de l'assemblage des différentes parties de la caméra, il reste souvent un certain désalignement, par exemple le centre de projection du système de lentille ne correspond pas au centre du *CCD* (*Charge-Coupled Device*) ou pire encore l'axe optique du système de lentille n'est pas perpendiculaire au *CCD* [89].

II.1 Modèles de déformation de lentille

On distingue principalement deux types de distorsion de lentille, tel que représenté sur la Figure II.1. La distorsion dite en barillet tend à rapprocher du centre de distorsion les points de l'image, alors que la distorsion en coussinet, ou en croissant, tend à les éloigner.

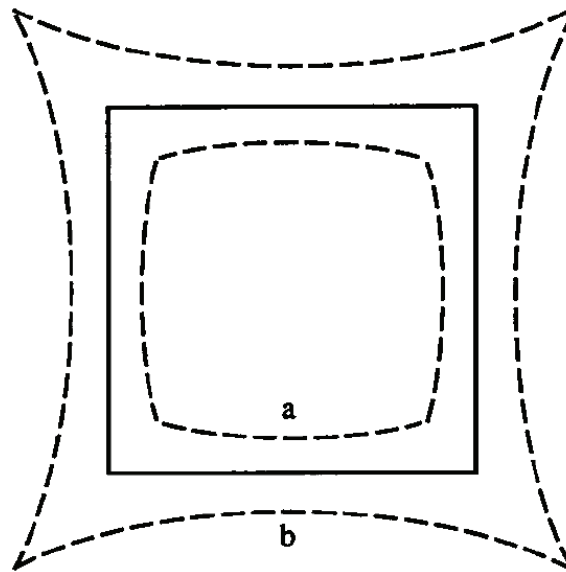


Figure II.1. Deux types de distorsion (a : barillet; b : coussinet) (adaptée de J. Weng, P. Cohen, et M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965-980, 1992. ©1992 IEEE. Reproduit avec permission.)

Comme on peut le voir sur la Figure II.2, la conséquence visuelle la plus évidente de la présence de distorsion est la transformation des lignes droites en lignes courbes. En général, dans les caméras endoscopiques, on se retrouve en présence de distorsion en barillet. Cette dernière n'est pas nécessairement involontaire : en effet, la présence de ce type de distorsion donne plus de détails au centre de l'image, tout en permettant d'avoir un angle de vue large; par contre, cet avantage rend l'estimation des distances plus difficile [90].

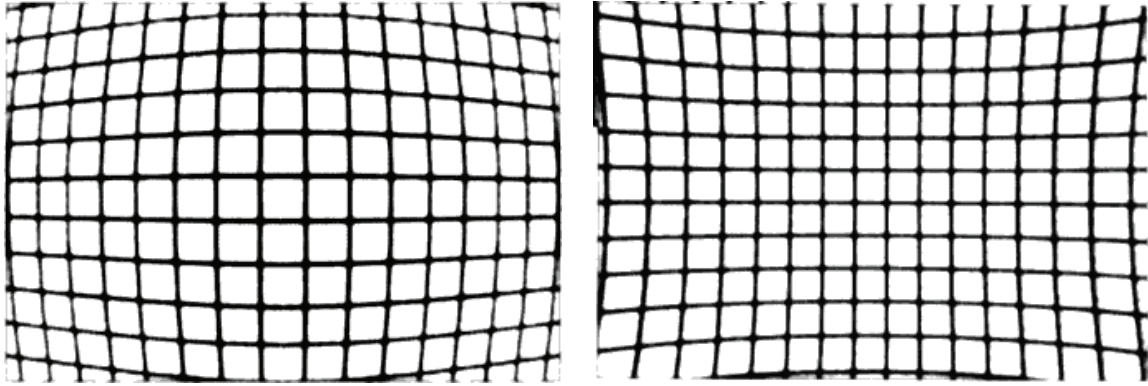


Figure II.2. Distorsion en barillet (gauche) et en coussinet (droite) (adaptée de Wikipedia, "Distorsion (optique)," *Distorsion (optique)*. [En ligne]. Disponible: [http://fr.wikipedia.org/wiki/Distorsion_\(optique\)](http://fr.wikipedia.org/wiki/Distorsion_(optique)). [Consulté le 2010/03/16] . Sous licence Creative Commons.)

II.1.1 Distorsion radiale

En pratique, les lentilles d'endoscopes typiques possèdent une symétrie circulaire à un niveau de précision de moins d'un micron, ce qui permet de considérer que la distorsion est uniquement radiale [91]. Ainsi, cela signifie que la distorsion ne fait que rapprocher ou éloigner les points d'un centre de distorsion qui, selon les auteurs, peut correspondre ou ne pas correspondre au centre optique de l'image. Dans tous les cas, il est intéressant d'exprimer les points de l'image en coordonnées polaires :

$$\begin{cases} x = c_x + r \cos(\theta) \\ y = c_y + r \sin(\theta) \end{cases} \Leftrightarrow \begin{cases} r = \sqrt{(x - c_x)^2 + (y - c_y)^2} \\ \theta = \arctan(y - c_y, x - c_x) \end{cases} \quad (\text{II.1})$$

où le point (x, y) est transformé en coordonnées polaires (r, θ) exprimées à partir du centre de distorsion (c_x, c_y) . Ensuite, l'expression de la distorsion ne devient qu'une fonction

d'une seule variable, soit r . On exprime donc le rayon r_d du point auquel la distorsion a été appliquée en fonction du rayon r du point non distordu :

$$r_d = f(r) \quad (\text{II.2})$$

Souvent, cette fonction est un polynôme en r :

$$r_d = \sum_{i=0}^n k_i r^i \quad (\text{II.3})$$

Le degré de ce polynôme ainsi que les hypothèses sur les différents coefficients des puissances, par exemple l'utilisation de puissances paires ou l'assignation de valeurs fixes telles que 0 et 1 pour k_0 et k_1 , varient selon les méthodes.

Il est à noter que ce polynôme peut être utilisé de manière directe pour corriger une image distordue. En effet, la correction consiste à générer une image non distordue dans laquelle chacun des pixels peut être déterminé en prenant sa position et en la distordant à l'aide du polynôme de distorsion afin de trouver à quel pixel il correspond dans l'image distordue.

II.1.2 Distorsion tangentielle

Un autre type de distorsion, souvent négligé, peut être produit par un système de lentilles. Il s'agit de la distorsion de décentrage, souvent associée au modèle théorique de l'effet de prisme mince (« *thin prism* »). Elle est causée par un alignement non parfait des lentilles. Brown [92] décrit analytiquement cette distorsion et cette dernière peut être modélisée à l'aide d'un terme tangentiel combiné au terme radial présenté plus haut. Ce terme tangentiel est proportionnel à la fois à une fonction du rayon nommée profil tangentiel, souvent modélisée par un polynôme et notée $P(r)$, et au cosinus de l'angle avec l'axe de distorsion tangentielle maximale (voir Figure II.3). En effet, comme il existe un axe où la distorsion tangentielle est maximale et un autre, perpendiculaire, où elle est nulle, le cosinus permet de modéliser ce comportement :

$$\Delta t(x, y) = P(r) \cos(\theta - \theta_0) \quad (\text{II.4})$$

où Δt est le déplacement tangentiel causé par la distorsion, θ est l'angle des coordonnées polaires du point considéré et θ_0 est l'angle entre l'axe de distorsion tangentielle maximale et l'axe des X . Cette équation peut être développée sous une forme plus facilement utilisable, soit :

$$\begin{aligned}\Delta x &= \left(P_1 (r^2 + 2x^2) + 2P_2 xy \right) (1 + P_3 r^2 + P_4 r^4 + \dots) \\ \Delta y &= \left(2P_1 xy + P_2 (r^2 + 2y^2) \right) (1 + P_3 r^2 + P_4 r^4 + \dots)\end{aligned}\quad (\text{II.5})$$

Il est à noter que dans la plupart des applications modernes, la distorsion tangentielle est négligée. De plus, même lorsqu'elle est considérée, on n'estime généralement que les coefficients P_1 et P_2 .

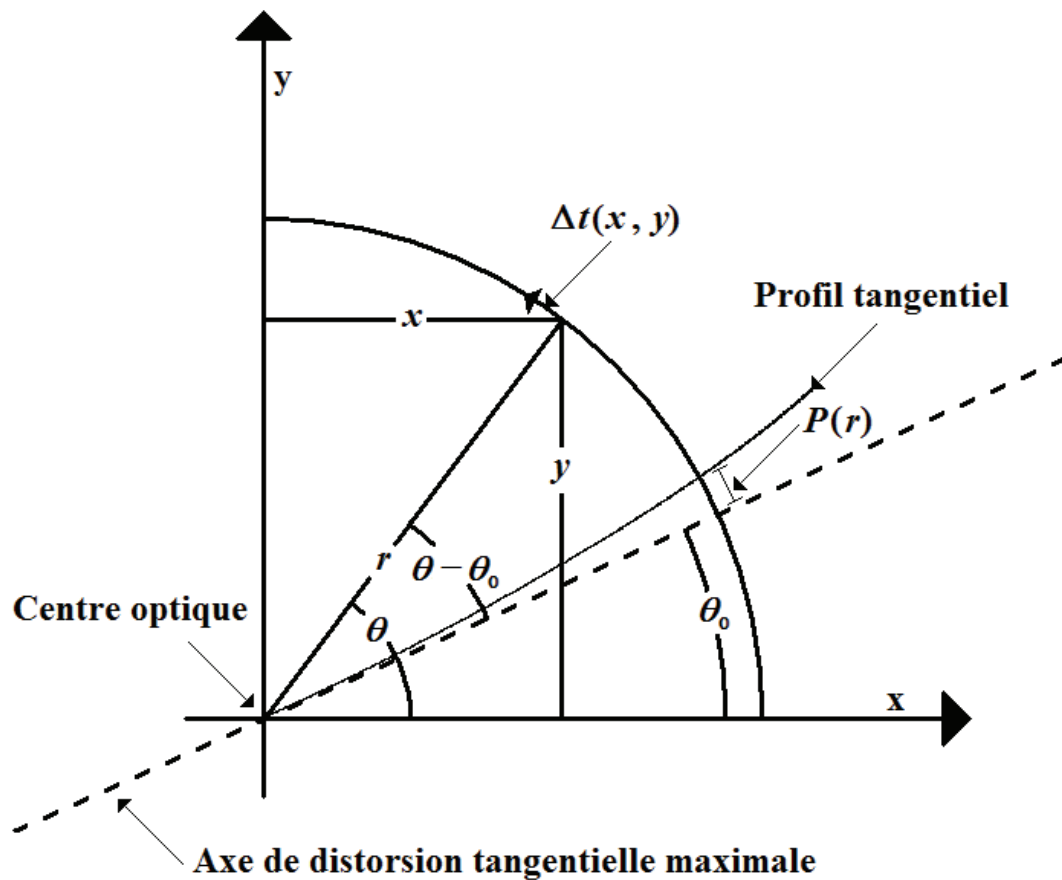


Figure II.3. Représentation de la signification géométrique des paramètres de la distorsion tangentielle

Ainsi, au modèle de projection linéaire développé aux sections I.1 à I.3 s'ajoute une transformation non linéaire modélisant la distorsion des lentilles. La section suivante décrit les méthodes permettant de calibrer le modèle de cette déformation.

II.2 Calibrage du modèle pour correction de la distorsion

Les premières applications des méthodes de correction de la distorsion ont été dans le domaine de la cartographie aérienne [93]. Au cours des années, l'avènement des mesures photogrammétriques a également motivé l'adaptation et le développement de nouvelles techniques. De nos jours, les applications de réalité augmentée nécessitent un calibrage plus complet qui implique toujours toutefois une correction de la distorsion.

De nombreuses méthodes tentent de corriger la distorsion des lentilles sans se soucier du reste du modèle de la caméra. En effet, la distorsion étant modélisée comme une transformation 2D sur l'image projetée, ce problème peut être attaqué indépendamment du reste du calibrage. On distingue les méthodes basées sur les distances radiales, celles basées sur un critère de courbure des lignes d'une grille de calibrage, celles basées sur une correspondance de motifs et celles basées sur la taille des éléments d'une grille [9].

II.2.1 Distorsion à partir des distances radiales

La première méthode présentée développée par Smith et al. [91] utilise les distances radiales des points d'une grille de calibrage afin de déterminer les coefficients du polynôme décrivant la distorsion. On estime d'abord la position du centre de distorsion en utilisant le principe suivant : comme la distorsion ne se fait que sur la coordonnée radiale d'un point exprimé en coordonnées polaires à partir du centre de distorsion, toute droite passant par le centre de distorsion restera une droite dans l'image distordue, alors que les autres droites seront toutes transformées en courbes. Ainsi, une grille hexagonale telle que celle représentée dans la Figure II.4 définit en chacun des points de la grille trois droites passant par les six (deux points par droite, un de chaque côté) voisins directs de chaque point. Lorsque la grille est distordue telle qu'elle apparaît après avoir été visualisée par la caméra, il est possible d'identifier le centre de distorsion par l'identification des lignes de points constituant quand même une droite sur l'image distordue. L'utilisation d'une grille hexagonale, fournissant trois droites, résulte en une meilleure reproductibilité de l'identification du centre de distorsion qu'avec l'utilisation d'une grille rectangulaire, ne fournissant que deux droites.

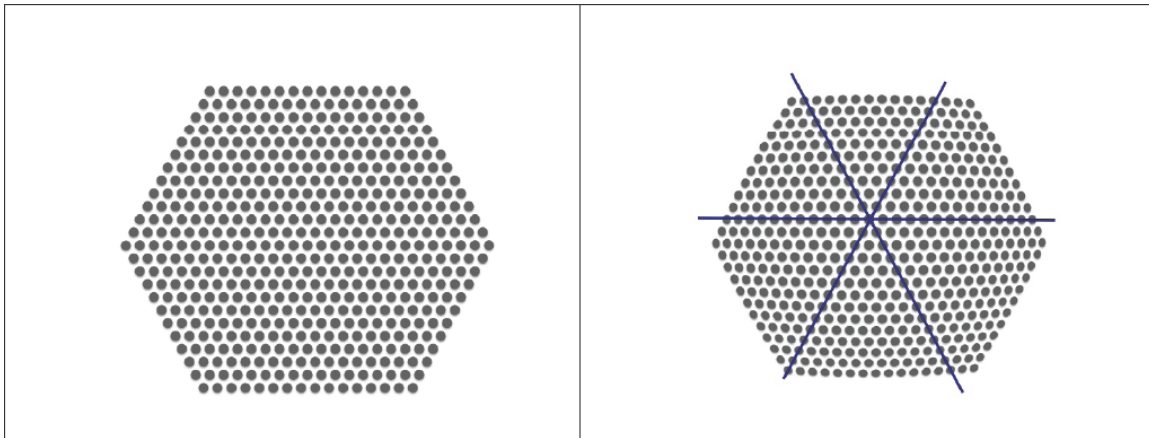


Figure II.4. Grille de calibrage hexagonale non distordue et l'équivalent distordu permettant d'estimer le centre de distorsion par l'intersection des droites (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission)

Une fois le centre déterminé, il est possible d'associer à chaque point sur un segment partant du centre une distance dans l'image non distordue. On choisit simplement l'indice de ce point lorsque l'on compte à partir du centre, distance qui devrait être linéaire dans une image non distordue puisque les points voisins sont également distancés les uns des autres. On associe également à chacun des points du segment choisi la véritable distance à partir du centre, exprimée en pixels par exemple. La détermination des paramètres de la distorsion revient donc à trouver les coefficients du polynôme transformant la distance non distordue en distance distordue. Les polynômes de Chebyshev sont utilisés pour déterminer les coefficients recherchés afin d'obtenir une solution directe (plutôt qu'itérative).

La validation de cette méthode se fait en comparant l'aire des cercles de la grille après la correction de la distorsion. Cette aire devrait être à-peu-près constante dans toute l'image, peu importe la distance au centre de distorsion. Les meilleurs résultats ont été obtenus avec un polynôme de degré 7. Toutefois, cette méthode possède deux inconvénients majeurs. Tout d'abord, les lignes servant à identifier le centre de distorsion doivent être identifiées manuellement. De plus, la grille doit être maintenue perpendiculaire à l'axe optique de la caméra pour que l'hypothèse sur les distances dans le patron non distordu, c'est-à-dire que la distance entre les points est linéaire en fonction de l'indice de ces derniers, soit respectée. Cette contrainte n'est pas simple à respecter et la vérification de la conformité à cette contrainte est également difficile.

II.2.2 Distorsion à partir de la courbure des lignes

La méthode développée par Haneishi et al. [94] évite le problème d'avoir à placer la grille de calibrage perpendiculairement à l'axe optique de la caméra. Pour ce faire, elle utilise le fait qu'une ligne droite dans la grille de calibrage devrait également être droite sur l'image. Or, en présence de distorsion, les lignes droites se retrouvent courbées dans l'image de la caméra. On définit donc un terme d'erreur augmentant en fonction de la courbure des droites. Ce terme est la somme des carrés des distances à la droite de régression de ces points. Ainsi, si la droite exprimée par l'équation :

$$ax + by + c = 0 \quad (\text{II.6})$$

est la meilleure droite passant par N points (x_i, y_i) , où i varie entre 1 et N , le terme d'erreur E devient :

$$E = \sum_{i=1}^N d_i^2 = \sum_{i=1}^N (ax_i + by_i + c)^2 \quad (\text{II.7})$$

Ce terme peut également être exprimé comme la plus petite valeur propre de la matrice M suivante :

$$M = \begin{bmatrix} \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i y_i & \sum_{i=1}^N y_i^2 & \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N y_i & N \end{bmatrix} \quad (\text{II.8})$$

On peut donc définir une fonction objective à partir de la somme de ce terme de courbure pour chacune des lignes de l'image. Cette fonction dépend à la fois du centre de distorsion et des coefficients du polynôme décrivant la distorsion; pour un jeu de paramètres, elle équivaut à trouver la courbure totale de l'image dont la distorsion a été corrigée à partir du jeu de paramètres. Il suffit donc ensuite de minimiser cette fonction pour trouver les paramètres de distorsion.

Pour ce faire, on délimite une région dans laquelle le centre de distorsion doit se trouver. Puis, pour chacun des pixels de cette région, on trouve les coefficients du polynôme qui minimise la courbure totale par descente de gradient. Finalement, on garde le centre, ainsi

que ses coefficients de polynôme, qui ont donné la courbure totale minimale parmi tous les pixels de la région considérée.

Cette méthode a donné comme résultat qu'à partir du quatrième ordre du polynôme, l'amélioration de la courbure est saturée et qu'il n'est donc pas nécessaire d'utiliser un polynôme de degré supérieur à 4.

Il est à noter que le polynôme cherché n'est pas exactement le même que celui de l'équation II.3, mais plutôt le polynôme inverse, qui transforme une distance radiale distordue en une distance radiale corrigée :

$$r = \sum_{i=0}^n k'_i r_d^i \quad (\text{II.9})$$

Ainsi, afin de pouvoir corriger la distorsion dans l'image, on a besoin de la fonction inverse de ce polynôme, qui est également approximée par un polynôme. La méthode utilisée [95] consiste à procéder à une régression multiple pour trouver les coefficients de ce polynôme, ce qui revient à utiliser le même jeu de points distordus et non distordus pour former un système d'équations linéaires où les inconnues sont les coefficients du polynôme. Ce système $AX = B$ est donc :

$$\begin{bmatrix} 1 & r_0 & r_0^2 & \cdots & r_0^j & \cdots & r_0^{n-1} & r_0^n \\ 1 & r_1 & r_1^2 & \cdots & r_1^j & \cdots & r_1^{n-1} & r_1^n \\ 1 & r_2 & r_2^2 & \cdots & r_2^j & \cdots & r_2^{n-1} & r_2^n \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 1 & r_i & r_i^2 & \cdots & r_i^j & \cdots & r_i^{n-1} & r_i^n \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 1 & r_{N-1} & r_{N-1}^2 & \cdots & r_{N-1}^j & \cdots & r_{N-1}^{n-1} & r_{N-1}^n \\ 1 & r_N & r_N^2 & \cdots & r_N^j & \cdots & r_N^{n-1} & r_N^n \end{bmatrix} \begin{bmatrix} k_0 \\ k_1 \\ k_2 \\ \vdots \\ k_j \\ \vdots \\ k_{n-1} \\ k_n \end{bmatrix} = \begin{bmatrix} r_{d0} \\ r_{d1} \\ r_{d2} \\ \vdots \\ r_{dj} \\ \vdots \\ r_{dn-1} \\ r_{dn} \end{bmatrix} \quad (\text{II.10})$$

où i est l'indice du point utilisé, j est l'indice du coefficient du polynôme, n est le degré du polynôme et N le nombre de points, avec $N \geq n$. Ce système surdéterminé peut être résolu avec la méthode de la pseudo-inverse (voir la section III.1 en annexe).

II.2.3 Distorsion à partir de recalage de niveaux de gris

Une méthode introduite par Miranda-Luna et al. [72] se base sur un motif déjà connu, comme ceux de la Figure II.5, afin trouver la transformation donnant les images distordues, comme celle de la Figure II.6.

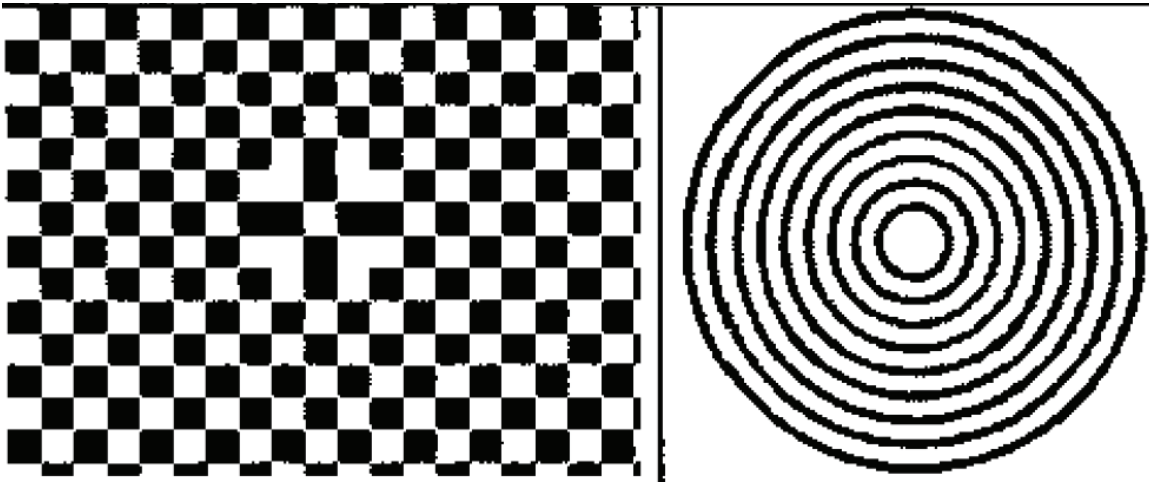


Figure II.5. Exemple de motifs originaux (adaptée de R. Miranda-Luna, W. C. P. M. Blondel, C. Daul, Y. Hernandez-Mier, R. Posada, et D. Wolf, "A simplified method of endoscopic image distortion correction based on grey level registration," in *International Conference on Image Processing*, vol. 5, 2004, pp. 3383-3386. ©2004 IEEE. Reproduit avec permission.)

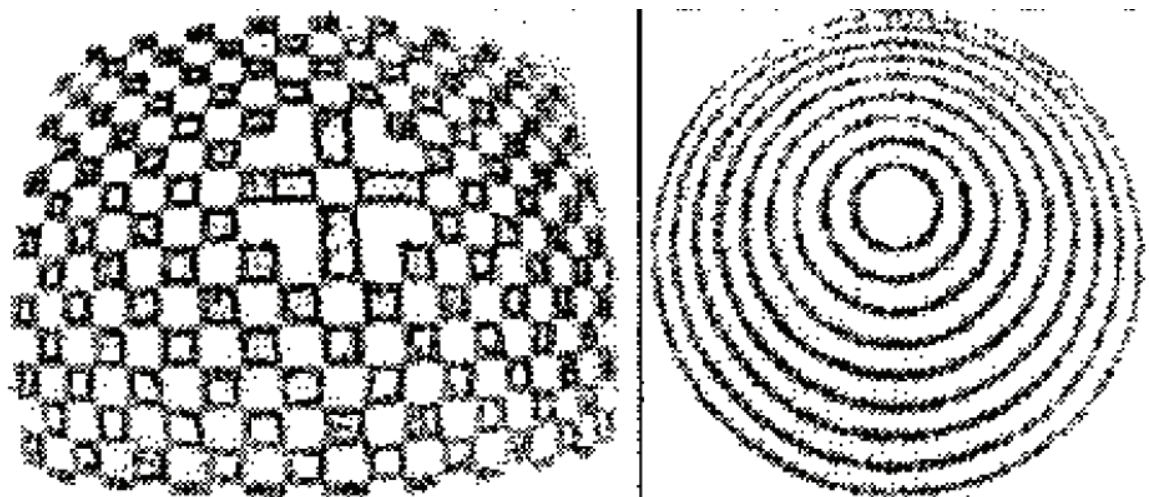


Figure II.6. Exemple de motifs avec distorsion (adaptée de R. Miranda-Luna, W. C. P. M. Blondel, C. Daul, Y. Hernandez-Mier, R. Posada, et D. Wolf, "A simplified method of endoscopic image distortion correction based on grey level registration," in *International Conference on Image Processing*, vol. 5, 2004, pp. 3383-3386. ©2004 IEEE. Reproduit avec permission.)

Afin de fonctionner même si le patron n'est pas exactement perpendiculaire à l'axe optique de la caméra, cette méthode ajoute une homographie (représentée par une matrice 3×3 appliquée aux points en coordonnées homogènes) aux paramètres à trouver. Ainsi, on cherche à trouver la transformation projective en plus des paramètres du polynôme de distorsion qui donneront le meilleur recalage entre le patron original et l'image distordue à laquelle on applique la correction de la distorsion. Le critère de similitude est basé sur l'information mutuelle entre les deux images et est défini comme la somme de l'entropie de chacune des images moins celle de leur entropie commune.

Cette méthode fixe à 5 le degré du polynôme utilisé comme bon compromis entre la précision des calculs et le temps de calcul sans donner d'autres justifications. Les résultats présentés sont satisfaisants pour la correction d'images ayant été prises avec un angle allant jusqu'à 10 degrés avec l'axe optique de la caméra. Cette méthode reste toutefois complexe à implanter et nécessite, pour son bon fonctionnement, un traitement d'image approprié.

II.2.4 Distorsion à partir de la taille des éléments d'une grille

Girard [9] présente une méthode de détermination des paramètres de la distorsion en se basant sur l'observation que dans l'image d'une grille de calibrage rectangulaire prise avec la caméra, mais dans laquelle la distorsion a été corrigée, chacun des carrés de la grille devrait avoir la même taille, cette taille étant celle du plus grand carré de l'image, situé au centre de distorsion, car la distorsion est de type barillet.

Ainsi, les paramètres recherchés sont le centre de distorsion ainsi que les coefficients du polynôme de distorsion. Tout comme avec la méthode de Miranda-Luna et al., le degré du polynôme a été fixé à 5 et les deux premiers coefficients à 0 et 1 respectivement, ce qui laisse 4 coefficients inconnus. Deux autres inconnues correspondant à la position du centre de distorsion sont ajoutées.

À partir de l'image distordue d'une grille dont les coins ont été identifiés, comme à la Figure II.7, on applique les paramètres de correction de la distorsion et on calcule une fonction objective à partir des points corrigés. Cette fonction est composée de deux termes, soit la variance des longueurs des arêtes ainsi que la moyenne des écarts entre cette longueur et la longueur du plus grand carré.

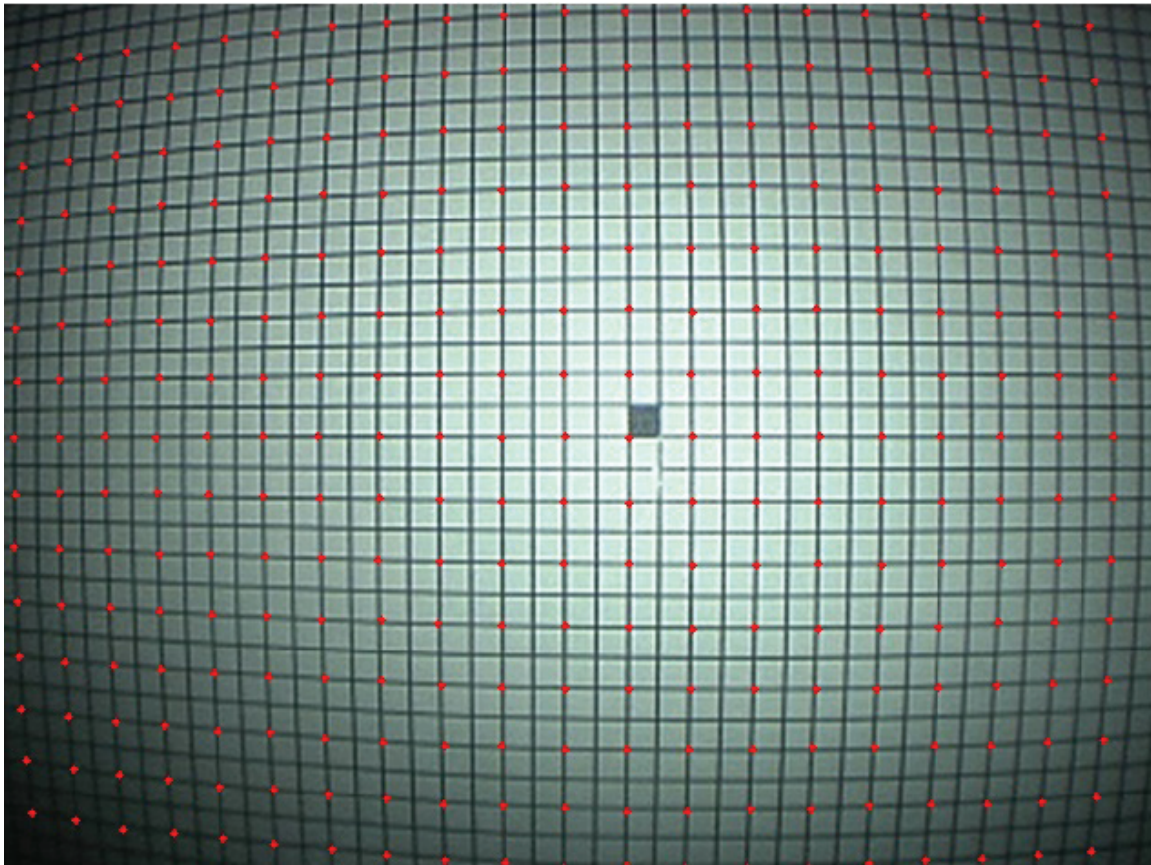


Figure II.7. Grille de calibrage de Girard (tirée de F. Girard, "Autocalibrage de séquence d'images thoracoscopiques," Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, 2009. ©2009 Fantin Girard. Reproduit avec permission)

Les paramètres de la distorsion sont donc trouvés en minimisant cette fonction. Tout comme dans le cas de la méthode de Haneishi, le polynôme inverse devra être trouvé pour procéder à la correction des images.

Cette méthode requiert également que la grille soit placée perpendiculairement à l'axe optique de la caméra, ce qui peut s'avérer problématique, spécialement compte tenu du fait que le centre de distorsion est inclus dans le processus de minimisation de la fonction objective. Il est très difficile de s'assurer visuellement, à partir d'une image distordue en plus, que l'image est belle et bien prise perpendiculairement à la caméra. Une image non perpendiculaire donnera une forme trapézoïdale à la grille, tout comme le fera un déplacement du centre de distorsion, comme on peut le voir à la Figure II.8.

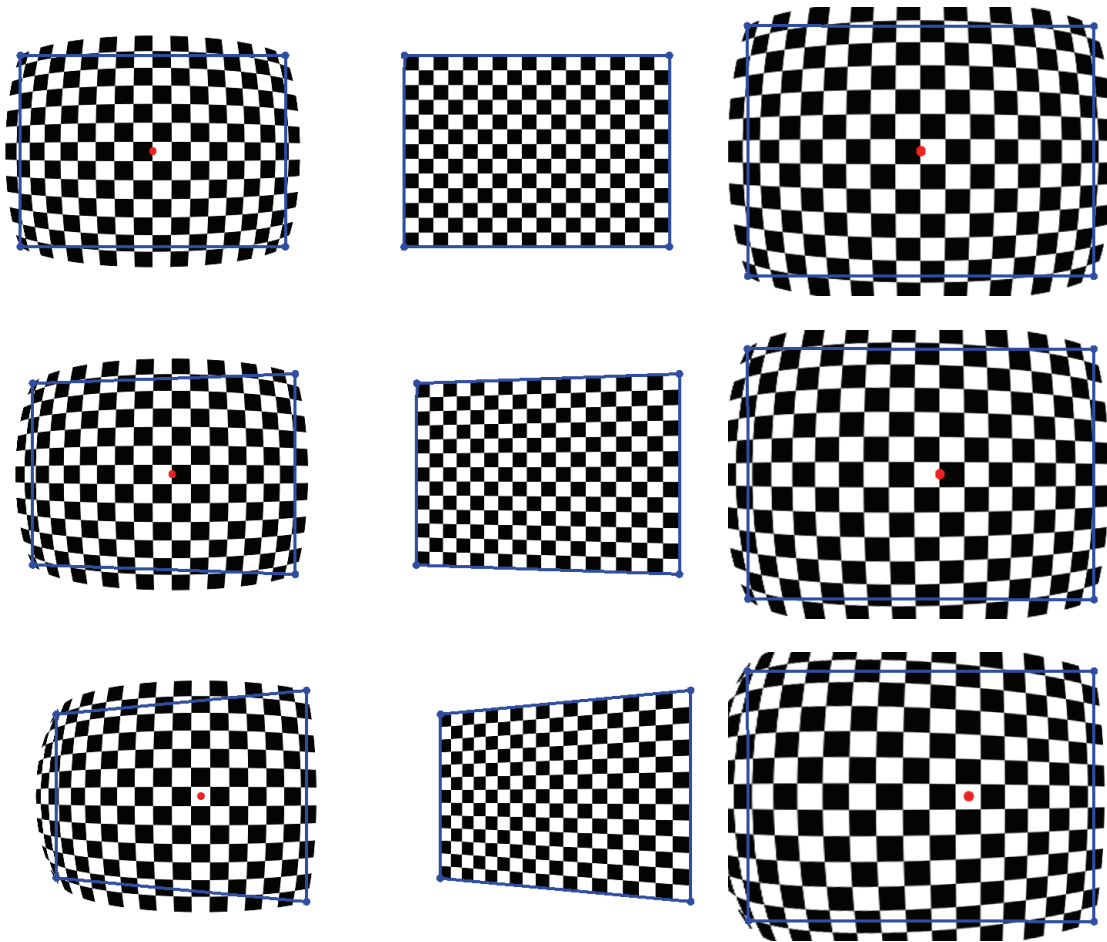


Figure II.8. Comparaison entre l'effet d'un déplacement du centre de distorsion et celui de la non-perpendicularité de la grille de calibrage

Sur cette figure, la première ligne représente le cas où le centre de distorsion est situé au centre de l'image, la seconde ligne représente le cas où le centre de distorsion est décalé de 5,0 % de la taille de l'image par rapport au centre et la troisième ligne représente le cas où le centre de distorsion est décalé de 12,5 % de la taille de l'image par rapport au centre. La colonne de gauche illustre le trapèze formé par les coins de la grille lors de l'application de la distorsion. La colonne du centre, quant à elle, illustre un trapèze identique formé par les coins de la grille à laquelle a été appliquée une transformation projective équivalente à l'angulation de la grille par rapport à la caméra. On voit donc une certaine équivalence entre l'effet d'un décentrage par rapport à l'angulation de la grille. Cette équivalence devient encore plus évidente avec la colonne de droite, où le décentrage a été combiné à une transformation projective ramenant les coins en un rectangle parfait. Il est difficile, à l'œil nu, de distinguer

la contribution de chacun des effets et d'arriver à trouver la position qui place la grille perpendiculaire à la caméra. Ainsi, la contrainte de perpendicularité de la grille par rapport à la caméra n'est pas triviale à assurer, ce qui peut rendre difficile l'utilisation de la méthode de Girard.

II.2.5 Distorsion à partir du calibrage

D'autres méthodes de correction de la distorsion combinent la détermination des paramètres de distorsion avec celle des autres paramètres du modèle et sont donc intégrées de plus près au processus de calibrage de la caméra, c'est-à-dire l'estimation des paramètres intrinsèques et extrinsèques. Comme ces méthodes estiment en même temps la position de la caméra par rapport à l'objet de calibrage, elles possèdent l'avantage de ne pas poser de restrictions sur la position de la grille par rapport à la caméra.

Ces méthodes sont présentées avec leur méthode de calibrage associée dans la section 1.5.

ANNEXE III OUTILS MATHÉMATIQUES

Cette annexe présente les différents outils mathématiques de base nécessaires à la réalisation de ce travail de recherche.

III.1 Méthode de la pseudo-inverse

Soit A une matrice possédant m lignes et n colonnes, avec $m > n$; B une matrice colonne possédant m lignes et X une matrice colonne possédant n inconnues. Le système $AX = B$ est un système d'équations linéaires surdéterminé qui peut être résolu à l'aide de la méthode de la pseudo-inverse. Cette technique consiste à transformer le système d'équations de manière à pouvoir le résoudre à l'aide d'une méthode classique. Si on multiplie chaque côté de l'équation par la transposée de la matrice A , on obtient :

$$\begin{aligned} AX &= B \\ A^T (AX) &= A^T (B) \\ (A^T A) X &= (A^T B) \end{aligned} \tag{III.1}$$

Le système ainsi obtenu peut être résolu en inversant la matrice $A^T A$, qui est une matrice carrée de n lignes et n colonnes :

$$\begin{aligned} (A^T A) X &= (A^T B) \\ (A^T A)^{-1} (A^T A) X &= (A^T A)^{-1} (A^T B) \\ X &= (A^T A)^{-1} (A^T B) \end{aligned} \tag{III.2}$$

Un algorithme classique, par exemple celui de l'élimination de Gauss-Jordan, peut ensuite être utilisé.

III.2 Résolution d'un système d'équations linéaires homogène

Soit A une matrice possédant m lignes et n colonnes, avec $m > n$; O une matrice colonne, dont toutes les valeurs sont nulles, possédant m lignes et X une matrice colonne possédant n inconnues. Le système $AX = O$ est un système d'équations linéaires surdéterminé qui possède une infinité de solutions. En effet, malgré la présence de n inconnues, il n'y a que $n - 1$ degrés de liberté, car si X est une solution, kX est également une solution. Généralement, la solution triviale où toutes les valeurs de X sont nulles n'est pas intéressante.

Il est possible de résoudre ce système en choisissant la valeur d'une des variables non nulles. Cette assignation fixe l'échelle de la matrice X . Le système résultant, à $n - 1$, n'est plus un système homogène et peut être résolu, à l'aide de la méthode de la pseudo-inverse présentée à la section III.1. Toutefois, cette technique échouera si la variable à laquelle on assigne la valeur doit être nulle. Il est alors possible de choisir une autre variable et d'essayer à nouveau.

Plus rigoureusement, il est préférable d'assigner une contrainte sur la norme de X , par exemple une norme unitaire. Il est possible de démontrer [65] que la solution de ce problème est le vecteur propre de $A^T A$ associé à la plus petite valeur propre, ou de manière équivalente le vecteur singulier de A associé à la plus petite valeur singulière. Il est donc préférable d'effectuer la décomposition en valeurs singulières (SVD) pour trouver X comme étant le vecteur singulier ayant la valeur singulière la plus petite¹³.

III.3 Algorithme *DLT*

L'algorithme *DLT* [65] permet de déterminer la matrice M reliant un ensemble de points P_i à un autre ensemble de points p_i , les points étant exprimés en coordonnées homogènes. Cet algorithme peut servir à :

- trouver l'homographie (transformation projective exprimée sous la forme d'une matrice 3x3) reliant deux ensembles de points 2D
- trouver la matrice de projection associant des points 3D à leur projection 2D, ce qui revient à faire un calibrage implicite

La seconde application sera ici utilisée pour illustrer l'algorithme, mais l'application au calcul d'une homographie est pratiquement directe.

On cherche donc la matrice M qui permet de transformer les points $P_i = [X_i \ Y_i \ Z_i]^T$ en points $p_i = [x_i \ y_i]^T$. Le point P_i est relié au point p_i par l'équation :

$$\lambda \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = M \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} = \begin{bmatrix} L_0 & L_1 & L_2 & L_3 \\ L_4 & L_5 & L_6 & L_7 \\ L_8 & L_9 & L_{10} & L_{11} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \quad (\text{III.3})$$

¹³ L'implantation de la décomposition en valeurs singulières utilisée dans le cadre de ce travail est celle du Template Numerical Toolkit (<http://math.nist.gov/tnt/index.html>).

Chaque association entre un point 3D et sa projection 2D identifiée sur l'image fournit trois équations permettant de rechercher les inconnues L_i :

$$\begin{bmatrix} \lambda x_i \\ \lambda y_i \\ \lambda \end{bmatrix} = \begin{bmatrix} L_0 X_i + L_1 Y_i + L_2 Z_i + L_3 \\ L_4 X_i + L_5 Y_i + L_6 Z_i + L_7 \\ L_8 X_i + L_9 Y_i + L_{10} Z_i + L_{11} \end{bmatrix} \quad (\text{III.4})$$

La troisième ligne donne directement la valeur de λ , qui peut être substituée dans les deux premières lignes pour finalement fournir deux équations par association de points 3D/2D. En prenant un jeu de N points, on obtient le système d'équations linéaires suivant :

$$AX = B$$

$$A = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 Z_1 & -x_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -x_2 X_2 & -x_2 Y_2 & -x_2 Z_2 & -x_2 \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -y_2 X_2 & -y_2 Y_2 & -y_2 Z_2 & -y_2 \\ & & & & & & & \vdots & & & & \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -x_i X_i & -x_i Y_i & -x_i Z_i & -x_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -y_i X_i & -y_i Y_i & -y_i Z_i & -y_i \\ & & & & & & & \vdots & & & & \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -x_N X_N & -x_N Y_N & -x_N Z_N & -x_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -y_N X_N & -y_N Y_N & -y_N Z_N & -y_N \end{bmatrix} \quad (\text{III.5})$$

$$X = [L_0 \quad L_1 \quad L_2 \quad L_3 \quad L_4 \quad L_5 \quad L_6 \quad L_7 \quad L_8 \quad L_9 \quad L_{10} \quad L_{11}]$$

$$B = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

Ce système homogène peut être résolu à l'aide de la méthode présentée à la section III.2. La matrice recherchée est donc obtenue directement à partir de l'équation III.3.

Par ailleurs, il est recommandé de normaliser les points P_i et p_i avant de résoudre le système pour éviter des problèmes de stabilité numérique [65]. Soit N_p la matrice correspondant à la normalisation (c'est-à-dire une translation et une mise à l'échelle) des points P_i , N_p la matrice correspondant à la normalisation des points p_i et L la matrice obtenue suite à la résolution du système avec les points normalisés. On peut retrouver la matrice originale M à l'aide de :

$$M = (N_p)^{-1} L N_p \quad (\text{III.6})$$

III.4 Moyenne de transformations rigides

Les transformations rigides ne sont pas simples à moyenner, car elles comprennent à la fois une rotation et une translation. Les translations seules sont simples à moyenner, car elles sont commutatives et donc leur composition est toujours la même peu importe l'ordre d'application des translations. Ce n'est pas le cas des rotations, qui ne sont pas commutatives. La méthode suivante permet d'estimer une moyenne de rotations, et donc de transformation rigides. Il est à noter que pour que cette moyenne ait un sens, il faut que les transformations soient relativement semblables, puisque la moyenne de rotations complètement différentes n'a pas vraiment de signification.

Soit A un ensemble de transformations rigides tel que :

$$A = \{M_i\} = \{T_i R_i\} \quad (\text{III.7})$$

Chacune des transformations M_i est composée d'une rotation R_i et d'une translation T_i . On définit la moyenne M de ces transformations comme étant égale à TR , où T représente la moyenne des translations et R représente la moyenne des rotations. La moyenne T est simple à calculer : il suffit d'ajouter composante à composante les vecteurs de translations puis de diviser par le nombre de vecteurs, comme une moyenne classique.

Pour les rotations, on se base sur le travail de Gramkow [96]. La moyenne des rotations devrait être la rotation qui minimise la somme de la distance à toutes les rotations. La mesure de distance utilisée entre deux rotations est simplement l'angle de la rotation permettant de transformer une rotation en l'autre rotation. À l'aide de développements de Taylor, Gramkow démontre que la moyenne algébrique des rotations, qu'elles soient représentées sous forme de matrices de rotation ou sous forme de quaternions, se rapproche beaucoup de cette moyenne idéale. On choisit donc de moyenner les rotations par la moyenne algébrique des quaternions q_i représentant les rotations R_i .

Un point doit être soulevé quand à la représentation de rotations par les quaternions : le quaternion $q_i = [w_i, x_i, y_i, z_i]$ et le quaternion $-q_i = [-w_i, -x_i, -y_i, -z_i]$ représentent tous deux la même rotation. Il est donc important, lorsque l'on fait la moyenne algébrique, de prendre la somme de quaternions orientés dans le même sens. Généralement, on compare les quaternions avec le premier : si le produit scalaire du quaternion q_0 et du quaternion q_i est positif, on ajoute q_i à la moyenne; dans le cas contraire, on ajoute $-q_i$. Finalement, comme les

quaternions représentant des rotations sont unitaires, il suffit de normaliser la somme pour obtenir la rotation moyenne.

III.5 Détermination du meilleur plan

L'algorithme permettant de déterminer l'équation $Ax + By + Cz + D = 0$ du meilleur plan passant par un ensemble de point $P_i = [x_i \ y_i \ z_i]^T$, où $0 \leq i < N$, est inspiré de la méthode présentée à <http://www.bmm.icnet.uk/people/suhail/plane.html> (consulté le 2010/05/21).

Le vecteur $[A \ B \ C]^T$ correspond au vecteur normal au plan recherché. Si on considère que ce vecteur est normalisé, la distance entre un point P_i et le plan est donnée par :

$$d_i = |Ax_i + By_i + Cz_i + D| \quad (\text{III.8})$$

Selon le critère des moindres carrés, on cherche le plan qui minimisera la somme des carrés des distances au plan. Soit S cette somme :

$$S = \sum_{i=0}^{N-1} (d_i)^2 = \sum_{i=0}^{N-1} (Ax_i + By_i + Cz_i + D)^2 \quad (\text{III.9})$$

Cette équation quadratique en fonction des variables A , B , C et D , trouvera son minimum lorsque la dérivée par rapport à chacune de ces variables s'annulera :

$$\begin{aligned} \frac{dS}{dA} &= 0 = \sum_{i=0}^{N-1} 2x_i (Ax_i + By_i + Cz_i + D) \\ \frac{dS}{dB} &= 0 = \sum_{i=0}^{N-1} 2y_i (Ax_i + By_i + Cz_i + D) \\ \frac{dS}{dC} &= 0 = \sum_{i=0}^{N-1} 2z_i (Ax_i + By_i + Cz_i + D) \\ \frac{dS}{dD} &= 0 = \sum_{i=0}^{N-1} 2(Ax_i + By_i + Cz_i + D) \end{aligned} \quad (\text{III.10})$$

L'équation de la dérivée par rapport à D peut être réécrite :

$$\begin{aligned}
 0 &= \sum_{i=0}^{N-1} 2(Ax_i + By_i + Cz_i + D) \\
 0 &= A \sum_{i=0}^{N-1} x_i + B \sum_{i=0}^{N-1} y_i + C \sum_{i=0}^{N-1} z_i + D \sum_{i=0}^{N-1} 1 \\
 0 &= A \sum_{i=0}^{N-1} x_i + B \sum_{i=0}^{N-1} y_i + C \sum_{i=0}^{N-1} z_i + DN \\
 0 &= A \frac{\sum_{i=0}^{N-1} x_i}{N} + B \frac{\sum_{i=0}^{N-1} y_i}{N} + C \frac{\sum_{i=0}^{N-1} z_i}{N} + D \\
 0 &= Ax_C + By_C + Cz_C + D
 \end{aligned} \tag{III.11}$$

où le point $P_C = [x_C \ y_C \ z_C]^T$ est le centroïde des points. Ce résultat montre donc que le plan recherché passe par le centroïde, puisque celui-ci respecte l'équation du plan. En substituant ce résultat dans l'équation III.9, on obtient :

$$\begin{aligned}
 S &= \sum_{i=0}^{N-1} (Ax_i + By_i + Cz_i + (-Ax_C - By_C - Cz_C))^2 \\
 S &= \sum_{i=0}^{N-1} (A(x_i - x_C) + B(y_i - y_C) + C(z_i - z_C))^2
 \end{aligned} \tag{III.12}$$

Encore une fois, la dérivation de ces équations donne un système d'équations linéaires, homogène cette fois-ci :

$$\begin{aligned}
 \frac{dS}{dA} &= 0 = \sum_{i=0}^{N-1} 2(x_i - x_C)(A(x_i - x_C) + B(y_i - y_C) + C(z_i - z_C)) \\
 \frac{dS}{dB} &= 0 = \sum_{i=0}^{N-1} 2(y_i - y_C)(A(x_i - x_C) + B(y_i - y_C) + C(z_i - z_C)) \\
 \frac{dS}{dC} &= 0 = \sum_{i=0}^{N-1} 2(z_i - z_C)(A(x_i - x_C) + B(y_i - y_C) + C(z_i - z_C))
 \end{aligned} \tag{III.13}$$

Ce système peut être réorganisé de la manière suivante :

$$\begin{aligned}
 0 &= A \sum_{i=0}^{N-1} (x_i - x_C)(x_i - x_C) + B \sum_{i=0}^{N-1} (x_i - x_C)(y_i - y_C) + C \sum_{i=0}^{N-1} (x_i - x_C)(z_i - z_C) \\
 0 &= A \sum_{i=0}^{N-1} (y_i - y_C)(x_i - x_C) + B \sum_{i=0}^{N-1} (y_i - y_C)(y_i - y_C) + C \sum_{i=0}^{N-1} (y_i - y_C)(z_i - z_C) \\
 0 &= A \sum_{i=0}^{N-1} (z_i - z_C)(x_i - x_C) + B \sum_{i=0}^{N-1} (z_i - z_C)(y_i - y_C) + C \sum_{i=0}^{N-1} (z_i - z_C)(z_i - z_C)
 \end{aligned} \tag{III.14}$$

Sous forme matricielle plus compacte, on obtient :

$$\begin{bmatrix} \sum_{i=0}^{N-1} (x_i - x_C)^2 & \sum_{i=0}^{N-1} (x_i - x_C)(y_i - y_C) & \sum_{i=0}^{N-1} (x_i - x_C)(z_i - z_C) \\ \sum_{i=0}^{N-1} (y_i - y_C)(x_i - x_C) & \sum_{i=0}^{N-1} (y_i - y_C)^2 & \sum_{i=0}^{N-1} (y_i - y_C)(z_i - z_C) \\ \sum_{i=0}^{N-1} (z_i - z_C)(x_i - x_C) & \sum_{i=0}^{N-1} (z_i - z_C)(y_i - y_C) & \sum_{i=0}^{N-1} (z_i - z_C)^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (\text{III.15})$$

On remarque que la matrice est en fait la matrice de covariance des points exprimés par rapport au centroïde. Comme la norme de la matrice des inconnues est unitaire, on peut directement utiliser la méthode présentée à la section III.2 pour résoudre ce système.

En résumé, l'algorithme permettant de déterminer le meilleur plan passant par un ensemble de points consiste à :

1. Calculer le centroïde des points.
2. Calculer la matrice de covariance des points auxquels on a soustrait le centroïde.
3. Obtenir les valeurs A , B et C comme étant le vecteur singulier associé à la plus petite valeur singulière de la matrice de covariance.
4. Obtenir D à l'aide du fait que $D = -Ax_C - By_C - Cz_C$.

ANNEXE IV DÉTECTION AUTOMATIQUE DES COINS DE LA GRILLE

Cette annexe présente l'algorithme de détection automatique des coins d'une grille utilisé pour automatiser le processus de calibrage. Cet algorithme est implanté à l'aide de la fonction `cvFindChessboardCorners()` de OpenCV. Les différentes étapes de l'algorithme sont expliquées à l'aide d'une image typique de grille acquise avec une caméra endoscopique, soit la Figure IV.1.

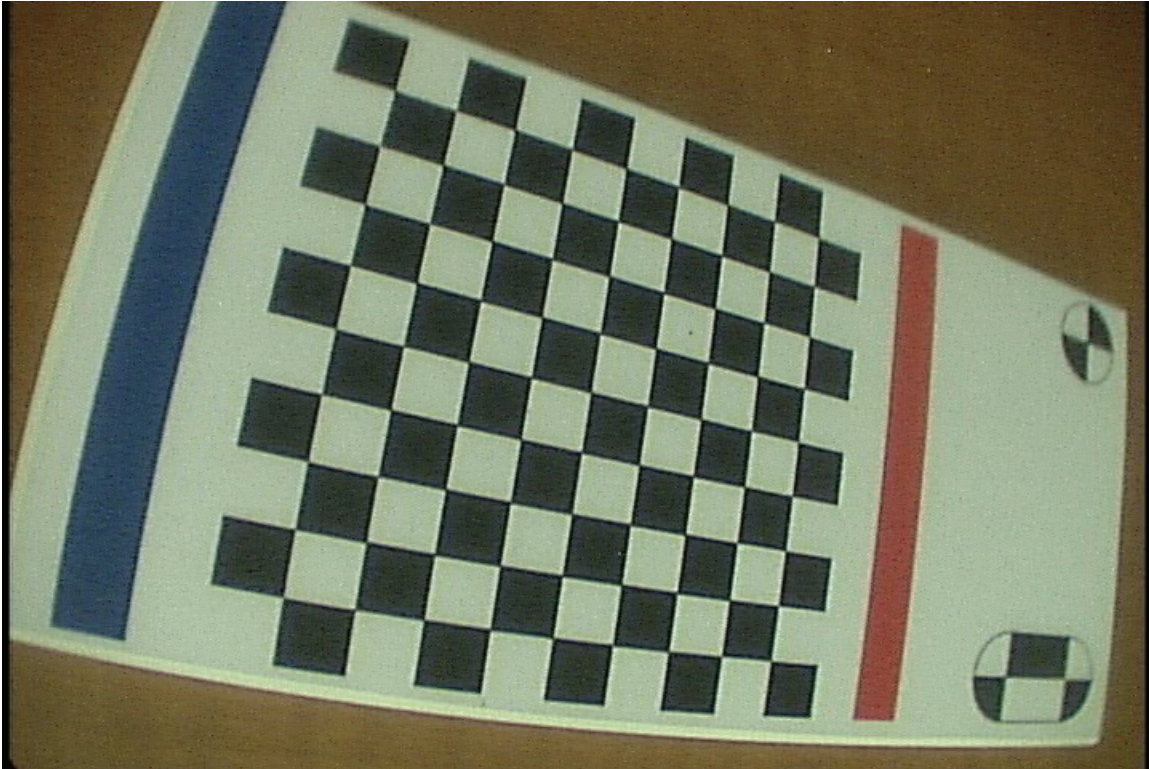


Figure IV.1. Image de départ

IV.1 Identification des coins de la grille

La première étape consiste à convertir l'image en une image de tons de gris, ce qui donne la Figure IV.2.

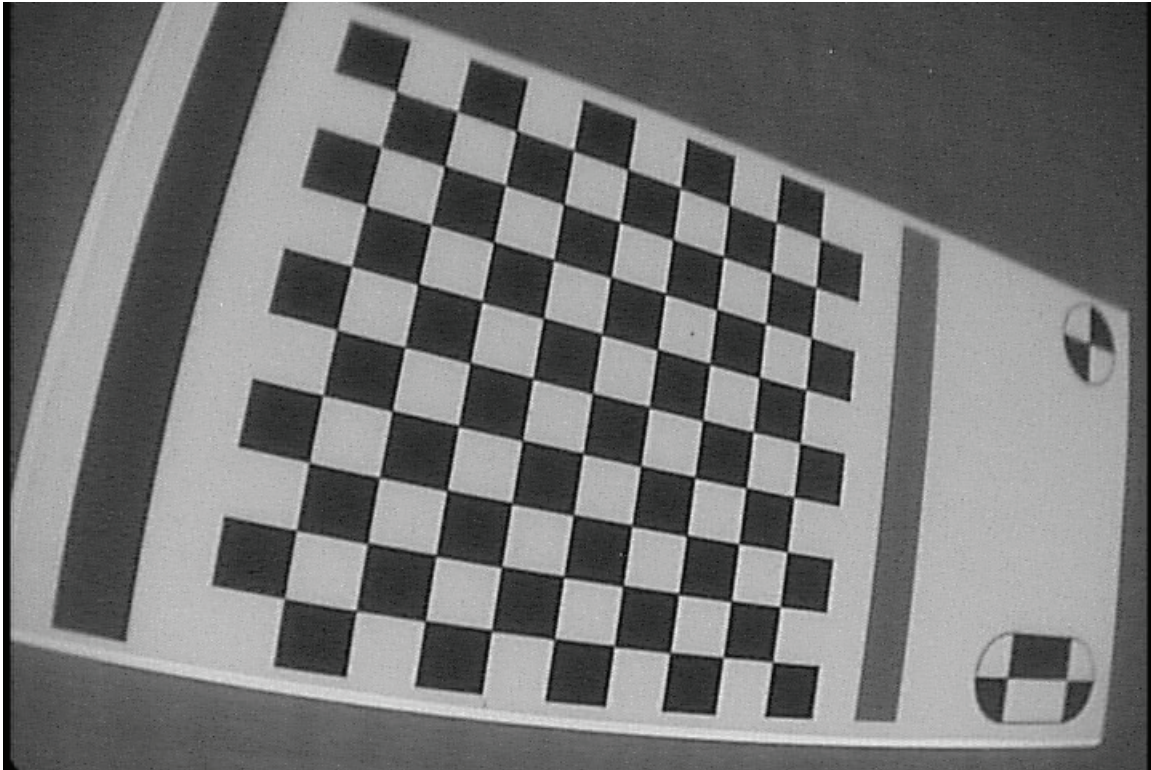


Figure IV.2. Image de tons de gris

On procède ensuite à une normalisation de l'histogramme de l'image, ce qui devrait aider à s'assurer d'avoir un contraste suffisant entre les régions noires et les régions blanches de la grille. Le résultat est présenté à la Figure IV.3.

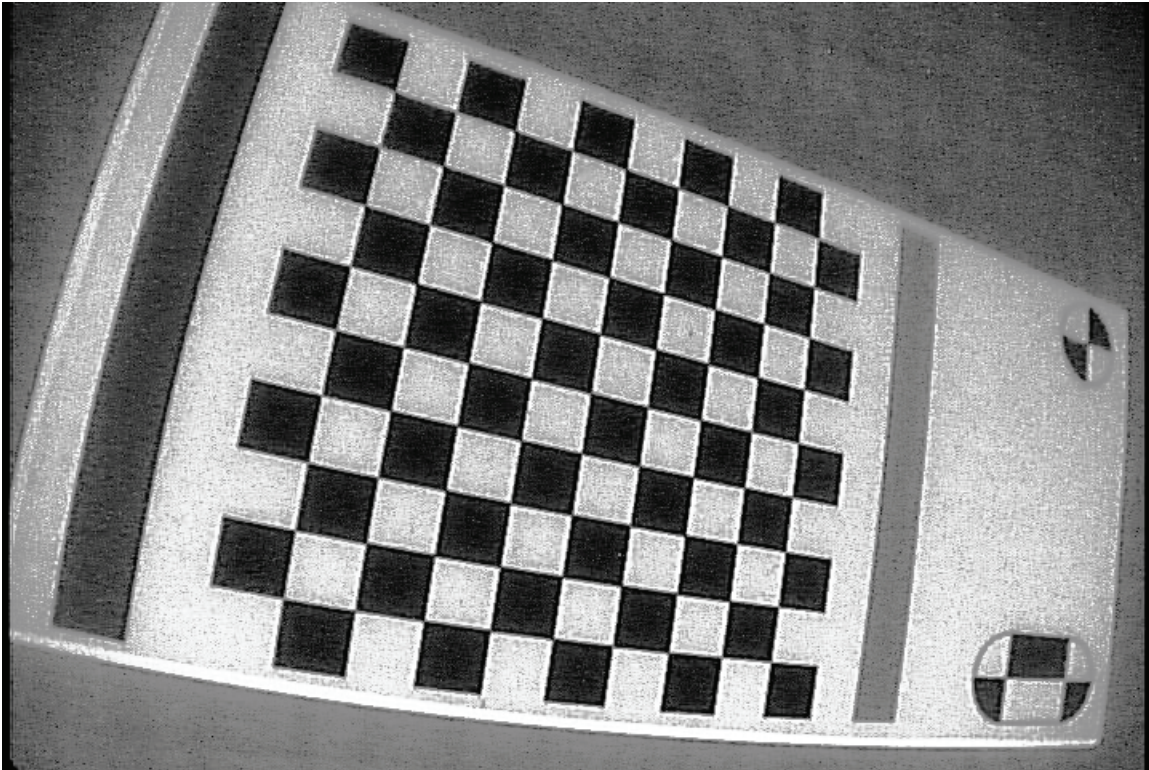


Figure IV.3. Image après l'égalisation de l'histogramme

Puis, l'image subit une binarisation à l'aide d'un seuil adaptatif. Pour un pixel donné, le seuil est calculé à partir de la moyenne des pixels membres d'une région entourant ce pixel. Le résultat de la binarisation est présenté à la Figure IV.4.

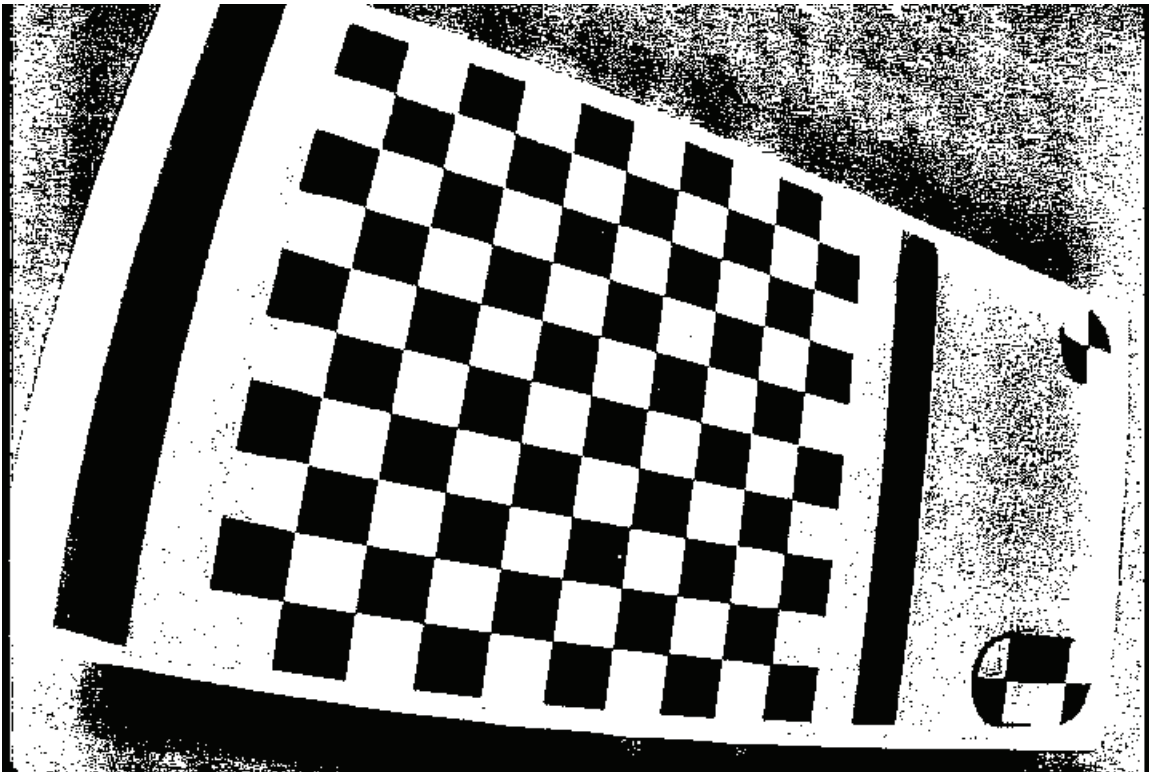


Figure IV.4. Image après la binarisation

Après la binarisation, on cherche à séparer les carrés les uns des autres. On procède donc à une dilatation morphologique du blanc (ce qui a le même effet qu'une érosion du noir), qui donne l'image de la Figure IV.5.

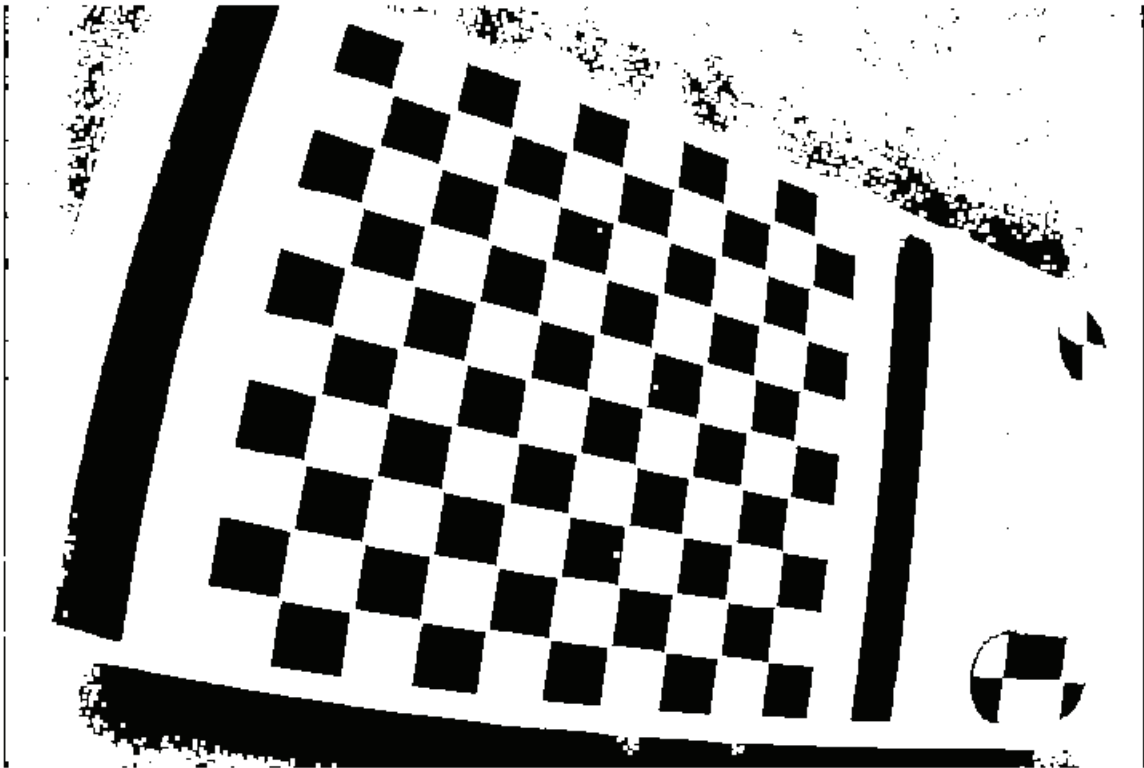


Figure IV.5. Image après la dilatation

Ensuite, l'algorithme identifie les contours sur cette image binaire. Ces derniers sont représentés sur la Figure IV.6.

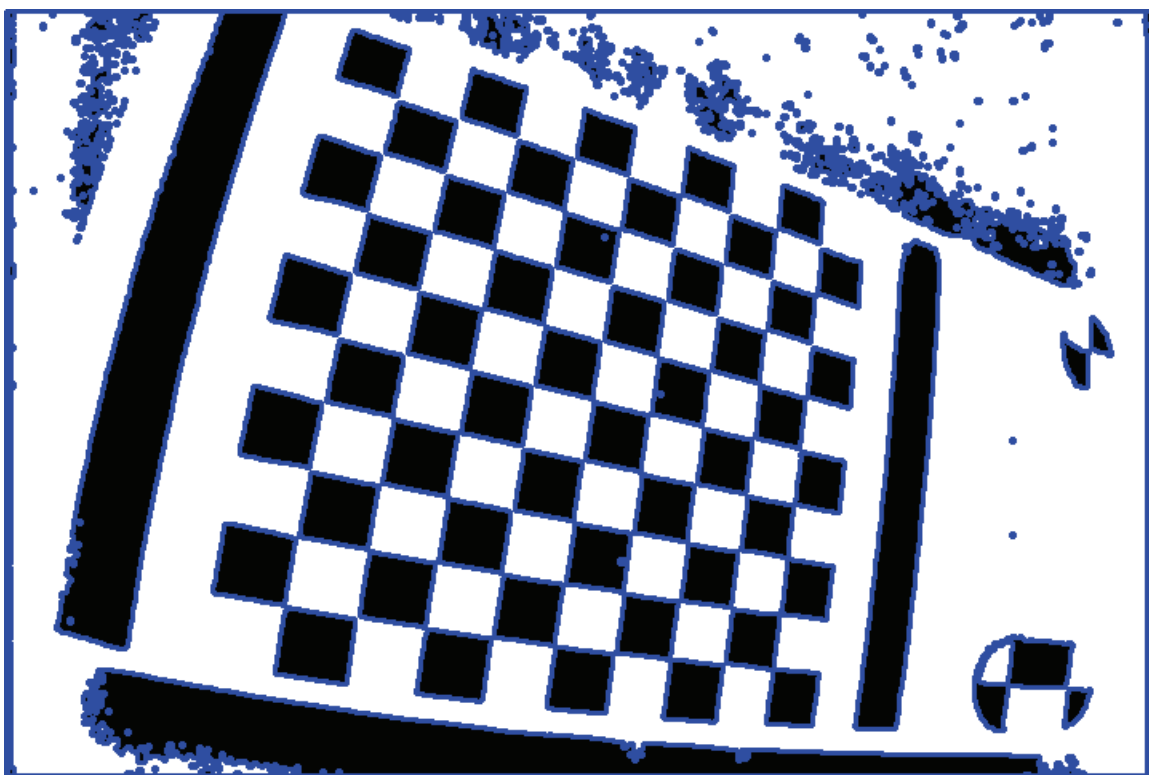


Figure IV.6. Image après l'identification des contours

On élimine par la suite les contours qui ne respectent pas certaines conditions : on rejette les contours dont le périmètre est trop petit, puis ceux qui ne sont pas formés de quatre côtés, puis finalement ceux qui ont une forme s'éloignant trop du carré. Les quadrilatères restants sont représentés à la Figure IV.7.

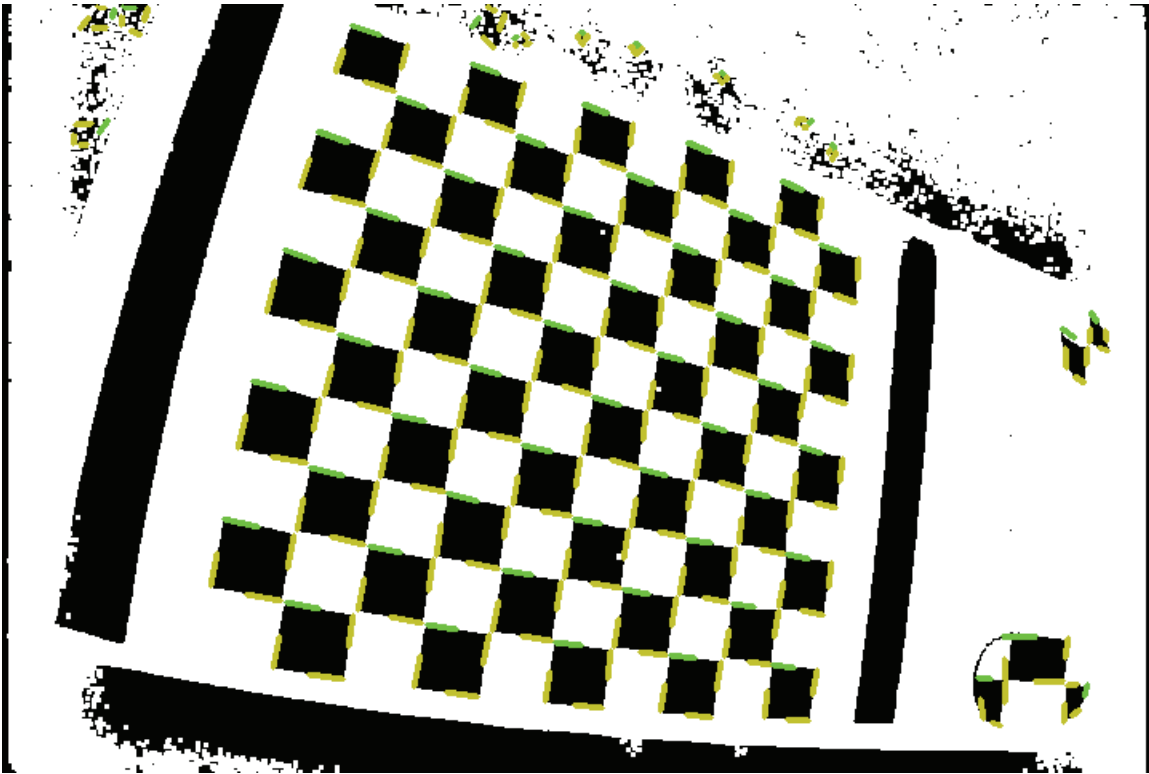


Figure IV.7. Image après l'élimination de régions

Une fois ces quadrilatères identifiés, ces derniers sont organisés en régions continues. Pour ce faire, on associe à un quadrilatère donné les voisins dont les coins chevauchent les siens. Puis, on identifie les composantes connexes du graphe afin de trouver une région contenant le nombre de quadrilatères cherché (selon la taille de la grille cherchée). Si la région contient trop de quadrilatères, on tente de retirer ceux qui ne font pas partie du patron recherché grâce à certains critères, par exemple la réduction de la taille de la boîte englobante des quadrilatères lorsque l'on en retire certains. Après que les quadrilatères aient été ordonnés d'une manière logique et reproductible, on obtient le résultat de la Figure IV.8.

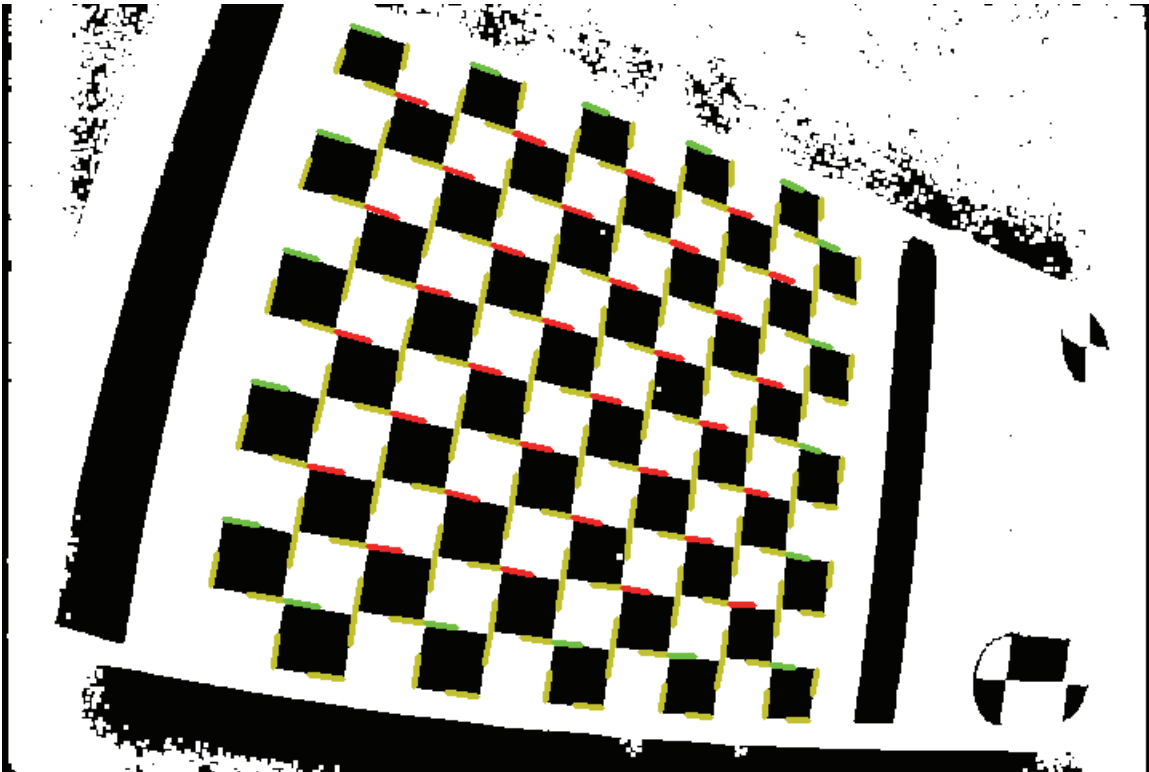


Figure IV.8. Image avec le bon ensemble de régions connexes

Finalement, les coins de la grille peuvent être identifiés et fournis dans un ordre logique de parcours de la grille, tel que représenté à la Figure IV.9.

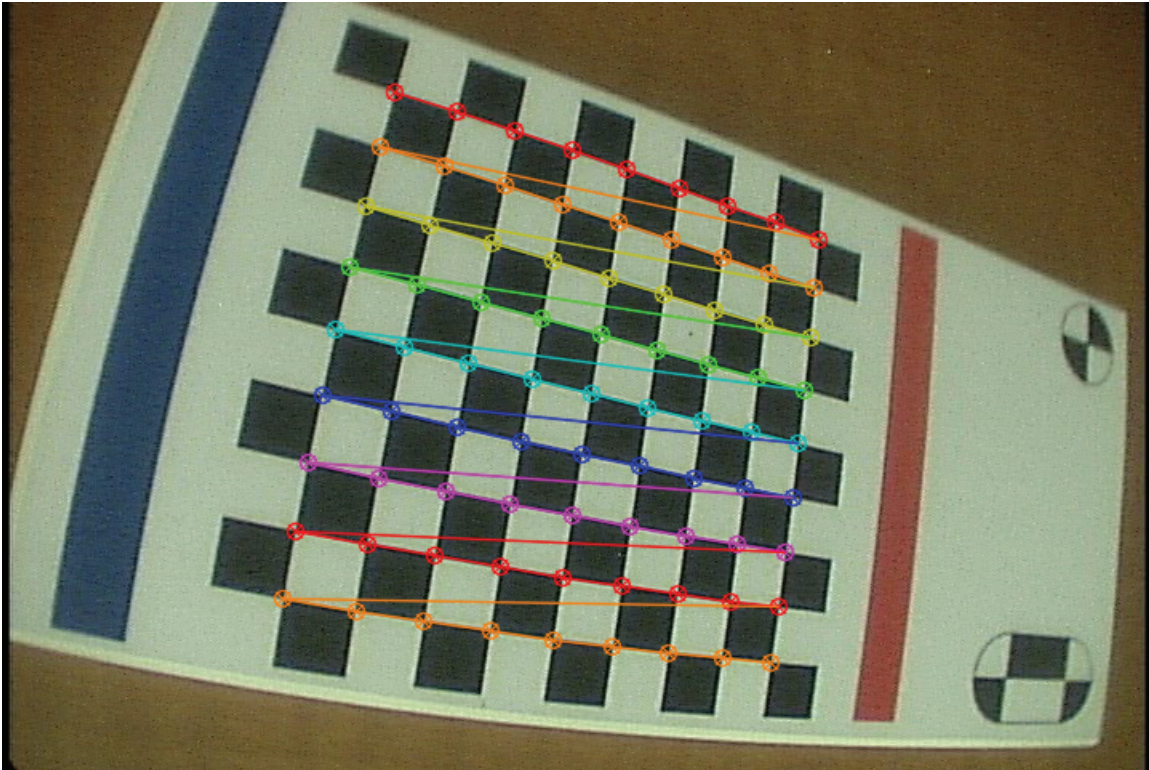


Figure IV.9. Image avec les coins identifiés et ordonnés

IV.2 Raffinement de la position des coins de la grille

Une fois les points identifiés « grossièrement » à l'aide de la méthode décrite à la section IV.1, il est possible de raffiner la position des coins identifiés afin d'obtenir une précision en-dessous du pixel.

Le principe de la méthode utilisée est résumé dans la Figure IV.10. Pour un coin donné, on cherche un point avoisinant q qui minimisera une certaine fonction objective. Cette dernière est basée sur le gradient. On observe en effet que pour tout point p situé à l'intérieur d'un carré blanc ou noir, le gradient sera nul. Par contre, sur les bords des carrés, le gradient sera non nul et perpendiculaire au côté commun des carrés. Si q est véritablement le coin de la grille, le vecteur $p - q$ sera perpendiculaire au gradient en p . Dans les deux cas (p à l'intérieur d'un carré ou p sur le côté d'un carré), le produit scalaire du gradient et du vecteur $p - q$ sera nul, si q est véritablement le coin de la grille.

On cherche donc le point q qui minimisera la somme, en valeur absolue, des produits scalaires du gradient en p et du vecteur $p - q$ pour tous les points p appartenant à un voisinage de q .

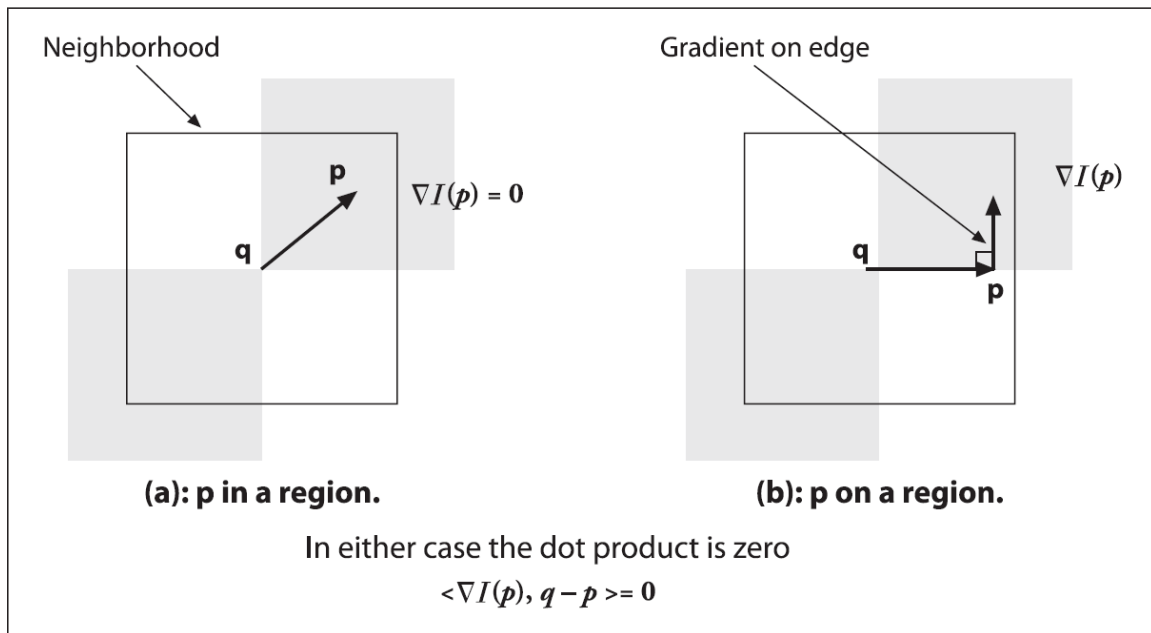


Figure IV.10. Localisation précise du coin de la grille (tirée de G. Bradski et A. Kaehler, *Learning OpenCV*, Sebastopol: O'Reilly Media, 2008. ©2008 O'Reilly Media. Reproduit avec permission.)

En appliquant ce raffinement à chacun des coins de la grille, on obtient un résultat plus précis et plus reproductible représenté à la Figure IV.11.

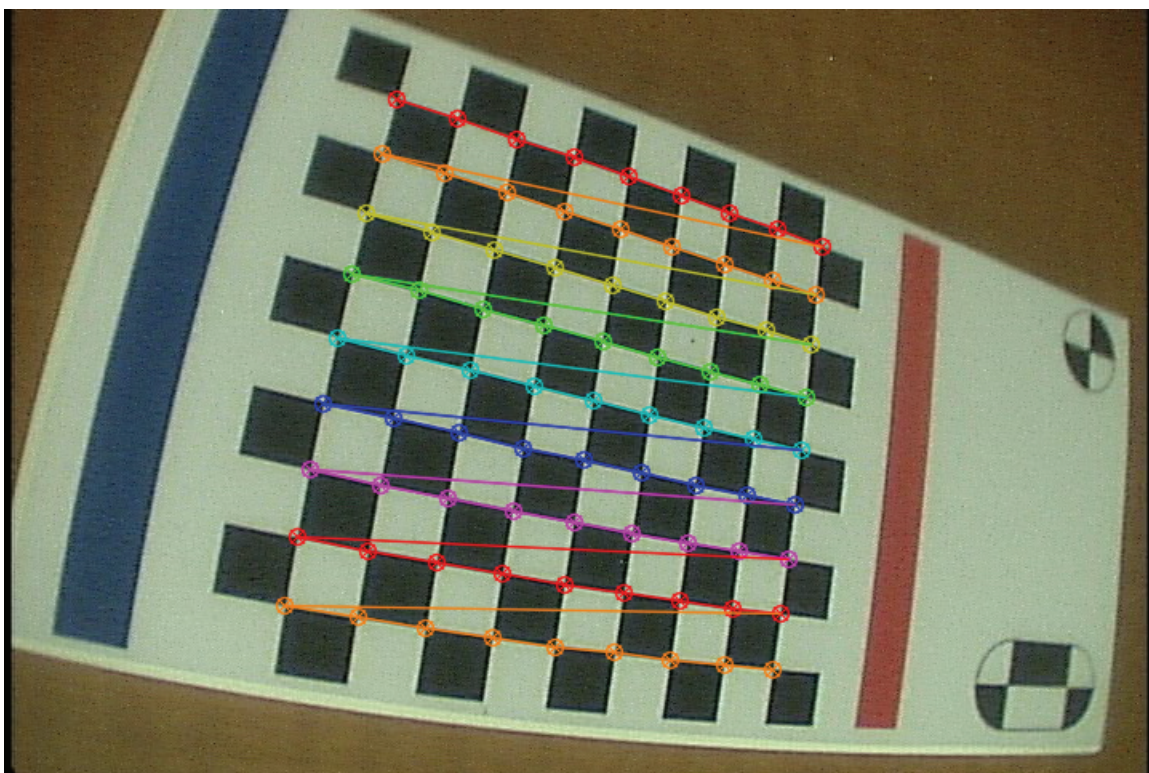


Figure IV.11. Résultat final après raffinement

ANNEXE V DÉTERMINATION DU CENTRE DE DISTORTION

La méthode présentée dans cette annexe permet de déterminer la position du centre de distorsion à partir d'une grille telle que celle extraite automatiquement de l'image de la Figure IV.11. Elle est inspirée de celle proposée par Asari et al. [97] et permet de déterminer le centre de distorsion même si la grille n'est pas perpendiculaire à l'axe optique de la caméra. Elle repose sur le fait que les lignes droites passant par le centre de distorsion resteront droites dans l'image distordue. Ainsi, en associant une mesure de courbure aux différentes lignes de la grille, on peut rechercher l'endroit où cette courbure s'annule.

La première étape consiste à appliquer une transformation rigide, c'est-à-dire une translation et une rotation, de manière à aligner la grille avec un système d'axe qui lui correspond davantage. La translation ramène le centroïde de la grille à l'origine, alors que la rotation vise à aligner la grille avec les axes, comme on peut le voir sur la Figure V.1.

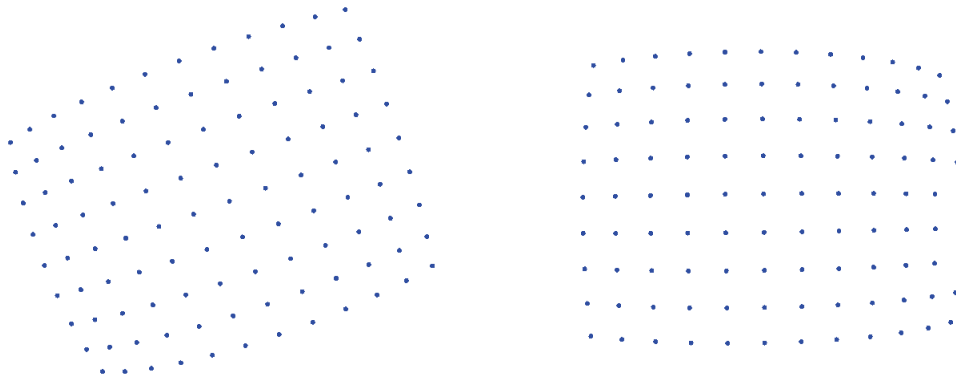


Figure V.1. Grille avant (gauche) et après (droite) la rotation pour aligner la grille

La détermination de la rotation à appliquer se fait en déterminant la meilleure droite passant par chacune des lignes « horizontales » de la grille. On nomme lignes horizontales les lignes que l'on tente d'aligner avec l'axe X et lignes verticales les lignes que l'on tente d'aligner avec l'axe Y . En effectuant la moyenne des vecteurs directeurs unitaires de chacune de ces meilleures lignes, on obtient un vecteur que l'on choisit d'aligner avec l'axe des X . Il est à noter que le vecteur u et le vecteur $-u$ décrivent la même droite. Lorsqu'on fait la moyenne des vecteurs, ceux-ci doivent être orientés dans le même sens. Généralement, on compare les vecteurs avec le premier : si le produit scalaire du vecteur u_0 et du vecteur u_i est positif, on

ajoute u_i à la moyenne; dans le cas contraire, on ajoute $-u_i$. La rotation appliquée à la grille est celle qui aligne le vecteur moyen avec l'axe des X positifs.

Soit $p_{i,j}$ le point de la i^{e} colonne et de la j^{e} ligne, où $0 \leq i < M$ et $0 \leq j < N$. On trouve les équations des N paraboles, chaque parabole passant par les points d'une ligne donnée. On cherche les paramètres A , B et C qui satisfont l'équation $y = Ax^2 + Bx + C$ à l'aide de la méthode des moindres carrés. Pour chacune des lignes, on construit un système d'équations surdéterminé pour les inconnues A_j , B_j et C_j .

Chacune de ces paraboles croise l'axe Y à la hauteur C_j , qui correspond à l'ordonnée à l'origine. On peut associer à chacune de ces ordonnées à l'origine la courbure de la parabole. Cette dernière sera positive pour les paraboles courbées vers le haut, nulle pour les droites et négative pour les paraboles courbées vers le bas. On cherche donc à trouver l'ordonnée pour laquelle la courbure sera nulle, ce qui revient à trouver le zéro de la fonction reliant les C_j au A_j . Il est possible de modéliser cette dernière de différentes façons : meilleure droite passant par le nuage de points, meilleur polynôme, spline, ou alors simplement de la modéliser comme une fonction linéaire par partie qui relie les points point-à-point. Cette dernière option a été retenue. On trouve ainsi l'ordonnée du centre de distorsion, c'est-à-dire l'ordonnée où les lignes de la grille restent droites.

On peut ensuite effectuer le même procédé avec les lignes verticales pour obtenir l'abscisse du centre à distorsion. Il suffit ensuite d'appliquer à ces coordonnées la transformation inverse de celle qui a été appliquée pour centrer et réorienter la grille pour obtenir les coordonnées du centre de distorsion dans le système de coordonnées original de la grille.

ANNEXE VI VUES GÉNÉRÉES POUR LES SIMULATIONS

Cette annexe présente les vues synthétiques générées analytiquement qui ont été utilisées pour les simulations. Le premier jeu, présenté à la Figure VI.1, est celui sur lequel le bruit est appliqué; le second jeu, illustré à la Figure VI.2, est celui à partir duquel l'erreur de reprojection est calculée.

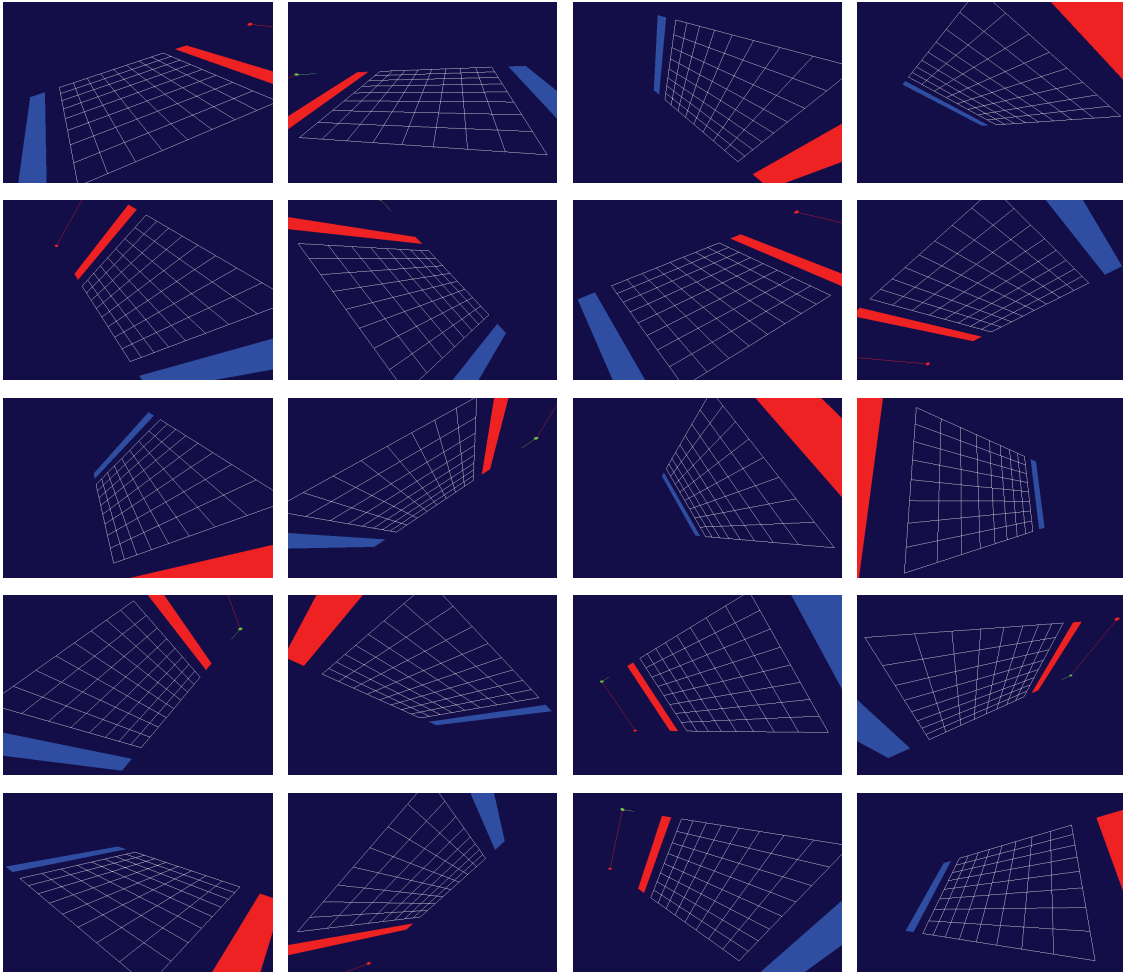


Figure VI.1. Premier ensemble de 20 images pour les simulations

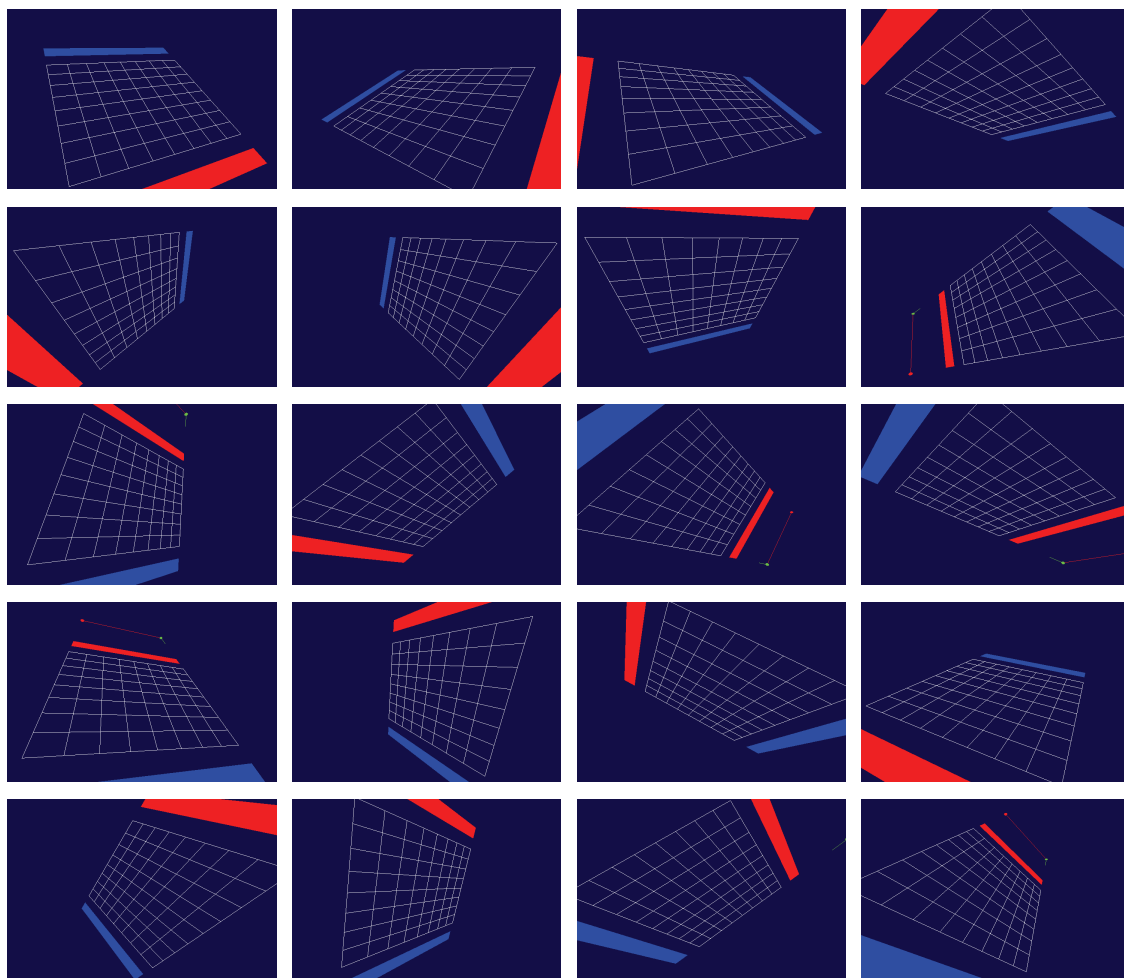


Figure VI.2. Second ensemble de 20 images pour les simulations

ANNEXE VII DÉTERMINATION DU CENTRE DE ROTATION

La méthode présentée dans cette annexe permet de déterminer la position du centre de rotation à partir d'une série de grilles telles que celle extraite automatiquement de l'image de la Figure IV.11. Elle est inspirée de celle proposée par Girard [9] et repose sur le fait que lorsqu'un point de l'image subit une rotation, la médiatrice du segment reliant le point avant la rotation au point après la rotation passe par le centre de rotation. Ainsi, en utilisant un certain nombre de points, on peut trouver le centre de rotation.

Soit p_{ij} la position du i^{e} point de l'objet de calibrage tel que vu sur la j^{e} image utilisée pour le calibrage, avec $0 \leq i < N$, où N est le nombre de points de l'objet de calibrage, et $0 \leq j < M$, où M est le nombre d'images.

Soit deux points p_{ij} et p_{ik} , où $j \neq k$, représentant donc le même point de l'objet de calibrage sur deux images différentes, les deux images étant reliées par une rotation. On recherche la droite d'équation $Ax + By = C$ représentant la médiatrice du segment reliant les deux points. Le vecteur reliant les deux points est donc perpendiculaire à cette droite médiatrice, tout comme l'est le vecteur (A, B) . On choisit donc (A, B) comme étant le vecteur reliant les deux points. Le paramètre C peut être déterminé à l'aide d'un point appartenant à la droite, soit le point milieu du segment reliant les deux points. Ainsi :

$$\begin{aligned}
 C &= Ax + By \\
 C &= (p_{i,j,x} - p_{i,k,x}) \left(\frac{p_{i,j,x} + p_{i,k,x}}{2} \right) + (p_{i,j,y} - p_{i,k,y}) \left(\frac{p_{i,j,y} + p_{i,k,y}}{2} \right) \\
 C &= \frac{p_{i,j,x}^2 - p_{i,k,x}^2 + p_{i,j,y}^2 - p_{i,k,y}^2}{2} \\
 C &= \frac{(p_{i,j,x}^2 + p_{i,j,y}^2) - (p_{i,k,x}^2 + p_{i,k,y}^2)}{2} \\
 C &= \frac{\|p_{i,j}\|^2 - \|p_{i,k}\|^2}{2}
 \end{aligned} \tag{VII.1}$$

Ainsi, il est possible de créer un système d'équations à l'aide de toutes ces droites afin de trouver le point qui appartient à toutes ces droites. Ce point, le centre de rotation, est donc la solution au système surdéterminé :

$$C = Ax + By$$

$$C = (p_{i,j,x} - p_{i,k,x}) \left(\frac{p_{i,j,x} + p_{i,k,x}}{2} \right) + (p_{i,j,y} - p_{i,k,y}) \left(\frac{p_{i,j,y} + p_{i,k,y}}{2} \right)$$

$$C = \frac{p_{i,j,x}^2 - p_{i,k,x}^2 + p_{i,j,y}^2 - p_{i,k,y}^2}{2} \quad (\text{VII.2})$$

$$C = \frac{(p_{i,j,x}^2 + p_{i,j,y}^2) - (p_{i,k,x}^2 + p_{i,k,y}^2)}{2}$$

$$C = \frac{\|p_{i,j}\|^2 - \|p_{i,k}\|^2}{2}$$

Le centre de rotation est donc déterminé en résolvant ce système surdéterminé.