

Un système de réalité augmentée pour guider les opérations du foie en radiologie interventionnelle

THÈSE

présentée et soutenue publiquement le 24 Novembre 2004

par

NICOLAU Stéphane

pour obtenir le titre de

Docteur en SCIENCES

de l'Université de Nice-Sophia Antipolis

Spécialité : Informatique

Thèse dirigée par AYACHE Nicholas

Composition du jury

<i>Président :</i>	BARLAUD Michel
<i>Rapporteurs :</i>	NAVAB Nassir BERGER Marie-Odile
<i>Directeur de thèse :</i>	AYACHE Nicholas
<i>Co-directeurs de thèse :</i>	SOLER Luc PENNEC Xavier
<i>Invité :</i>	COSTE-MANIERE Eve

Mis en page avec la classe thloria.

Remerciements

Je commencerai par remercier les membres de mon jury :

-Nicholas Ayache, mon directeur de thèse qui m'a accueilli au sein du projet Epidaure de l'INRIA Sophia-Antipolis et qui a su me pousser quand cela était nécessaire.

-Luc Soler, mon premier co-directeur de thèse, dont les conseils ont toujours été des plus avisés concernant le bon déroulement de ma thèse. Pour être franc, sans ton soutien moral, cette thèse aurait pu échouer. Vivement les prochaines expériences in vivo!

-Xavier Pennec, mon second co-directeur de thèse. Merci Xavier, tu as su m'enseigner de nombreuses choses théoriques et pratiques, tout en conservant une relation très humaine avec moi. J'espère que nous aurons d'autres occasions de travailler ensemble.

-Marie-Odile Berger et Nassir Navab, mes rapporteurs, qui m'ont fait l'honneur de relire et de juger mon travail de thèse.

-Eve Coste-Manière, qui a bien voulu faire partie de mon jury en tant qu'invitée, et qui a contribué à l'amélioration de mon manuscrit.

-Michel Barlaud, qui m'a fait l'honneur de présider mon jury. Je garderai un souvenir impérissable des questions "très techniques" que vous m'avez posées et auxquelles j'ai eu quelques difficultés à répondre.

Maintenant j'en viens aux personnes qui m'ont entouré sur mon lieu de travail.

Tout d'abord à l'INRIA :

Merci à toi Seb pour tes conseils de grand frère que j'ai pu recueillir pendant un an avant ton départ en Australie. J'espère que ma présence te fut agréable durant tous ces mois de cohabitation. Je n'oublierai jamais tous ces moments passés dans ta petite maison à Vallauris en compagnie de Nat.

Merci à vous Stab, Olivier P. et Carou. Grâce à vous j'ai pu chanter en concert une fois dans ma vie Fear of the Dark !! Sans vous, je n'aurais jamais joué de la guitare, et je n'aurais jamais rien compris au métal.

Merci à toi Isabelle S., pour tous tes efforts et ta gentillesse. L'équipe a de la chance de t'avoir.

Un clin d'oeil à JD, Alexis, Sylvain, Maxime, compagnons de maintes soirées.

A l'IRCAD :

Merci à Johan et Antoine, vous avez pris soin de moi dès mon arrivée à Strasbourg. Je me suis senti chez moi au bout de quelques semaines seulement.

Merci à Sylvain, AB, Julien, Nicolas, Marielle, Anaïs avec lesquelles j'ai passé vraiment beaucoup de bons moments durant ma première année à l'IRCAD.

Merci à ma petite bulgare Dobrina de m'avoir accompagné durant mes pauses café.

Maintenant les relations extérieures :

Evidemment, merci à benben, garstour, ren, fabio, sam et simon. Vous êtes mon noyau indestructible.... Même le temps ne parviendra pas à éroder notre amitié.

Merci aux deux frères d'or, Aurélien et Nicolas, avec lesquels j'ai partagé d'intenses émotions à Karlsruhe, Munich et Ising.

Merci à Vincent, Jean-Pascal, Sophie, Jani. La distance nous empêche de nous voir, mais l'oubli ne s'installera pas entre nous.

Merci à mon ami Mik l'osteopathe, Stéphane K. et Phetaphine. Combien de fois suis-je rentré à Grenoble, heureux à l'idée de vous revoir.

Merci à John et Marion, nous nous retrouverons, j'en suis plus que certain.

Merci à mes amis d'enfance Jean et Luc. Que serais-je devenu si je ne vous avais pas rencontrés si tôt ?

Merci à Eric et Isa. Ces trois années ont été riches en activités, en émotions, en discussions. Merci aussi de m'avoir offert le rôle de parrain de votre enfant. RRRAhhh j'crois....

Merci à Jérôme Schmid pour son travail efficace tout au long de ma dernière année. Il n'est pas courant de réussir à travailler avec quelqu'un aussi longtemps dans la bonne humeur.

Merci à mon ex co-colocataire Laurent. On a partagé deux années vraiment fertiles en discussions, en partages, en rencontres. Merci aussi de m'avoir écouté quand j'en avais vraiment besoin. Merci d'être aussi ouvert.

Merci à Pierre, Yvan, Christelle et Manu, Ariane. J'ai passé de nombreuses heures avec vous à contempler la couleur dorée de la musique.

Merci à Céline et Philippe. Des fois, quelques jours suffisent pour révéler de grandes amitiés.

Merci à Judith, ma copine qui m'a soutenu et aimé la dernière année de ma thèse (ce qui n'était pas une mince affaire).

Merci à ma grand-mère, décédée il y a maintenant trois ans, qui m'a élevé et chéri pendant plus de vingt ans. J'aurais tant souhaité que tu me voies maintenant.

Merci à mon frère Alexandre, mon unique et splendide grand frère. Les années ne nous sépareront pas. Le reste encore moins. Merci pour ta force, ta générosité, tes conseils, ta bienveillance. J'espère pouvoir un jour te rendre tout ce que tu m'as apporté.

Merci à papa et maman sans qui je ne serais pas là. Vous avez su m'éduquer et me donner les meilleures chances pour vivre et m'épanouir. Merci pour votre indéfectible amour, merci pour votre tendresse.

Table des matières

Un système de réalité augmentée pour guider les opérations du foie en radiologie interventionnelle	1
An augmented reality system to guide liver percutaneous punctures	1
1 Introduction	3
1.1 Organisation du mémoire	12
1.2 Contributions	14
2 Un système de guidage thérapeutique pour le foie	17
2.1 Réalité augmentée et opération du foie	17
2.1.1 La chirurgie ouverte	17
2.1.2 La chirurgie mini-invasive	19
2.1.3 Notre but : la radiologie interventionnelle	21
2.2 Choix de la technologie de réalité augmentée	24
2.2.1 Via un moniteur	25
2.2.2 Via un appareillage portable (Head Mounted Display)	25
2.2.3 Via un miroir semi-transparent	26
2.2.4 Par projection sur la scène réelle	27
2.2.5 Réalité augmentée via un moniteur: une technologie simple et précise	27
2.3 Objectifs	29
3 La stratégie de recalage	31
3.1 Recalage rigide ou non-rigide?	31
3.1.1 Influence de la respiration libre sur la position du foie	32
3.1.2 Cas de la respiration contrôlée par le patient	33
3.1.3 Cas d'un patient intubé	34
3.1.4 Conclusion	36
3.2 Recalage iconique, surfacique ou ponctuel?	36
3.2.1 Recalages iconiques	36

3.2.2	Recalage géométrique surfacique	38
3.2.3	Recalage géométrique ponctuel	39
3.2.4	Comparaison et choix	42
3.3	Des critères de recalage ponctuel 3D/2D inappropriés	44
3.3.1	Les critères et méthodes de recalage 3D/2D existants	44
3.3.2	Des hypothèses statistiques mal adaptées à nos conditions	45
3.4	Conclusion	46
4	Un nouveau critère de recalage 3D/2D	47
4.1	Dérivation du nouveau critère EPPC	47
4.1.1	Notations	48
4.1.2	Hypothèses du critère ISPPC sur le bruit 3D	48
4.1.3	Un critère étendu dérivé du maximum de vraisemblance	49
4.1.4	Cas de bruits 3D et 2D anisotropes	52
4.1.5	Un critère qui généralise les méthodes classiques	52
4.2	Evaluation des performances sur des données synthétiques	54
4.2.1	Les conditions d'expérimentation	56
4.2.2	Définition des mesures réalisées	56
4.3	Temps de calcul et précision	57
4.3.1	Choix du seuil d'arrêt d_{max} pour la minimisation alternée de l'EPPC	57
4.3.2	Influence de la tolérance de minimisation de Powell	58
4.3.3	Etude de la robustesse	58
4.3.4	Nouveau choix pour la tolérance de Powell	59
4.3.5	Les résultats de robustesse	60
4.3.6	Synthèse des résultats	60
4.4	Validation sur des données réelles	61
4.4.1	Méthodologie de la mesure d'erreur sans mesure étalon	61
4.4.2	Prise en compte de la distorsion dans les critères	64
4.4.3	Résultats quantitatifs et qualitatifs sur des données réelles	66
4.5	Conclusion	68
4.5.1	L'EPPC, un critère plus précis et plus robuste aux multiples applications	68
4.5.2	L'EPPC, un critère adapté à notre application médicale	68
5	Calibrage des caméras et de l'électrode	71
5.1	Deux caméras : notre consensus précision/encombrement	71
5.1.1	Les conditions expérimentales	72
5.1.2	Etude de la précision en fonction du nombre de caméras	72

5.1.3	Choix du nombre de caméras	73
5.2	Choix du calibrage des caméras	74
5.2.1	Les méthodes couramment utilisées	74
5.2.2	Le calibrage de Zhang : précis et fiable	75
5.2.3	Le modèle physique choisi	76
5.2.4	Notre calibrage stéréoscopique	79
5.3	Repérage de l'aiguille dans le repère des caméras	82
5.3.1	Amélioration de la librairie standard ARToolkit	83
5.3.2	Le calibrage de l'aiguille	85
5.3.3	Étude de la précision du positionnement de l'aiguille virtuelle	87
5.4	Conclusion	92
6	Automatisation de la localisation et de l'appariement des marqueurs	95
6.1	Extraction des marqueurs radio-opaques	95
6.1.1	Extraction dans les images vidéos	96
6.1.2	Extraction dans les images TDM	103
6.2	Mise en correspondances des marqueurs extraits	107
6.2.1	Mise en correspondances des données vidéos 2D	108
6.2.2	Mise en correspondances des données 3D et 2D	110
6.3	Suivi des marqueurs en temps réel	114
6.4	Conclusion	114
7	Assurer la précision du système	117
7.1	Étude théorique de l'incertitude	118
7.1.1	Rappel sur la propagation de l'incertitude appliquée aux critères de maximisation	118
7.1.2	Résumé algorithmique	119
7.2	Validation de la prédiction théorique	120
7.2.1	Validation de la linéarisation	121
7.2.2	Influence du calibrage	125
7.2.3	Validation sur des données entièrement réelles	126
7.3	Conclusion	131
8	Du prototype à l'application clinique	133
8.1	Validation du système sur un mannequin	133
8.1.1	Protocole expérimental	133
8.1.2	Guidage de ponctions percutanées sur des cibles à l'intérieur du mannequin	135
8.1.3	Résultats et discussion	138

8.2	Un système ergonomique et sécurisé	140
8.2.1	Un guidage optimal de la pose des marqueurs au positionnement de l'aiguille	140
8.2.2	Un système sécurisé à chaque étape	144
8.3	Evaluation du système sur des patients	145
8.3.1	Un protocole de validation sans danger pour le patient	146
8.3.2	Expériences sur trois cas cliniques	147
8.4	Conclusion	160
9	Conclusion	161
9.1	Conclusion théorique	161
9.2	Conclusion pratique	162
9.3	Conclusion médicale	163
9.4	Perspectives théoriques	163
9.5	Perspectives médicales	164
9.6	Perspectives pratiques	165
9.7	Conclusion générale	165
	Annexe	167
A	Evaluation des performances de l'EPPC	167
A.1	Temps de calcul et précision	168
A.1.1	Influence de l'amplitude du bruit	168
A.1.2	Influence de la distance objet/caméras	171
A.1.3	Influence de l'angle entre les caméras	172
A.1.4	Influence du nombre de points	173
A.2	Etude de la robustesse	174
A.2.1	T aléatoire avec une initialisation à l'identité	174
A.2.2	Initialisation aléatoire avec T fixée à l'identité	175
A.2.3	Initialisation aléatoire avec T aléatoire	176
A.3	Conclusion	176
B	Méthodes d'estimation de la pointe d'une aiguille	177
B.1	Estimation simultanée du point fixe dans les repères de référence et du corps rigide	178
B.2	Estimation du point fixe à partir des diamètres	178
C	Propagation de la covariance sur la transformation estimée à partir de l'ISPPC et de l'EPPC	181
C.1	Cas du critère ISPPC	182
C.2	Cas du critère EPPC	185

D Estimation de la position 3D relative de n points	189
D.1 Estimation à partir de mesures issues d'images stéréoscopiques	190
Bibliographie	191

En radiologie interventionnelle, le praticien se sert actuellement de plusieurs acquisitions scanner (coupes 2D) pour pouvoir atteindre sa cible au cours d'une ponction percutanée dans la zone abdominale. Notre objectif est de concevoir un système de guidage minimisant ces acquisitions et donc le temps de l'intervention. Pour cela, nous proposons un système de réalité augmentée superposant des reconstructions 3D pré-opératoires des structures abdominales du patient dans des images vidéo externes de son corps. En plus d'être précis, rapide et fiable, cet outil devra pouvoir être introduit aisément en salle d'opération.

Dans notre cas, le patient est intubé et sa ventilation contrôlée, nous pouvons donc négliger les effets de la respiration : un recalage rigide 3D/2D de marqueurs radio-opaques collés sur la peau est suffisant pour atteindre la précision requise. Les hypothèses statistiques des critères classiques n'étant pas adéquates pour notre application, nous avons dérivé un nouveau critère généralisant les approches standard. Une évaluation rigoureuse des performances démontre la supériorité de notre méthode en terme de précision et de robustesse.

Pour atteindre le temps réel en salle d'opération, nous avons ensuite développé un ensemble d'algorithmes d'extraction et de mise en correspondance des marqueurs radio-opaques dont nous avons validé la robustesse sur de nombreuses images réelles. La précision du système dépendant de nombreux paramètres (nombre de marqueurs radio-opaques, position des caméras...), elle ne peut pas être établie préalablement de manière définitive. Afin de fournir un système fiable, nous proposons donc une technique de propagation des covariances qui permet d'estimer dynamiquement l'erreur de repositionnement des modèles reconstruits. Une phase de validation méticuleuse, sur des données synthétiques et réelles, démontre que notre prédiction est fiable dans les conditions de notre application.

Après cette validation de chacun des modules, nous montrons la faisabilité et l'intérêt de notre système complet en menant une évaluation sur un mannequin : quatre chirurgiens ont réussi à atteindre des cibles en des temps dix fois inférieurs à ceux usuellement nécessaires pour ce type d'intervention et avec une précision supérieure. Finalement, plusieurs expériences cliniques sur des patients démontrent que notre système est utilisable en salle d'opération et suggèrent son utilisation en routine dans un futur proche.

Mots-clés: intervention assistée par ordinateur, réalité augmentée, radiologie interventionnelle, recalage ponctuel 3D/2D, propagation de covariance, validation, traitement d'image.

An augmented reality system to guide liver percutaneous punctures

In the interventional radiology field, the medical expert needs several CT-scan acquisitions (2D slices) to reach the target during a percutaneous puncture in the abdominal area. We aim at designing a guidance system to reduce the number of CT acquisition and thus, the intervention duration. We propose an augmented reality system that superimposes preoperative 3D reconstructions of the patient abdominal structures in external video images of his/her body. This system will not only have to be accurate, quick and reliable, but will also have to be easily introduced in the operative room.

In our case, the patient is intubated and the breathing is controlled. Therefore, we can neglect breathing motion: a rigid 3D/2D registration of radio-opaque markers is sufficient to obtain the necessary accuracy. Since statistical assumptions of existing 3D/2D registration criteria do not fit our application, we derived a new criterion generalizing classical methods. A rigorous evaluation of the performances showed the superiority of our approach in terms of accuracy and robustness.

To reach real-time processing in the operative room, we developed and validated several algorithms to automatically extract and match radio-opaque markers. Since the system accuracy depends on numerous factors (camera position, marker number...), it is not possible to calculate it beforehand. To provide a reliable system, we propose a technique of covariance propagation that enables to estimate the superimposition error of reconstructed models. A rigorous validation step on synthetic and real data showed that our prediction is correct in our conditions.

All the different parts of the system being validated, we perform an evaluation of the whole system on an abdominal phantom. Results proved the feasibility and the interest of the system: four surgeons managed to reach embodied targets with an average accuracy under 2 mm. Moreover, the time needed for the needle positioning was ten times under classical interventional duration. Finally, several clinical experiments on patients showed that our system can be used in the operating room and suggest that this system could be used routinely in a near future.

Keywords: computer-assisted surgery, augmented reality, interventional radiology, point-based 3D/2D registration, covariance propagation, image processing, 3D tool video tracking

Chapitre 1

Introduction

L'imagerie médicale : une révolution

Au XX^{ième} siècle, la médecine a connu une évolution majeure par l'essor des technologies de l'imagerie médicale, nées à la fin du siècle précédent. En effet, ces techniques offrent la possibilité de voir au travers du corps humain, et ainsi de rendre visible les structures anatomiques ou pathologiques. Une image médicale reflète, sur une ou plusieurs coupes 2D, certaines caractéristiques physiques des structures étudiées. La signification physiologique de l'image varie énormément suivant la modalité employée. La radiographie et la tomодensitométrie (TDM) fournissent une image dépendante de l'épaisseur des différents tissus et de la valeur des numéros atomiques des atomes les constituant. L'échographie permet de visualiser les différences d'impédance acoustique entre les tissus traversés. L'imagerie par résonance magnétique (IRM) visualise les temps de relaxation des atomes d'hydrogène, et l'imagerie nucléaire affiche le rayonnement gamma émis par un traceur radioactif préalablement injecté dans l'organisme. Les propriétés physiques d'organes voisins étant bien souvent différentes, le praticien est à même de reconnaître sur les images la plupart d'entre eux, parfois même les vaisseaux et les structures pathologiques.

Du diagnostic à l'intervention

La qualité et la précision de l'imagerie du corps humain permettent actuellement aux médecins d'élaborer des diagnostics très précis (en terme de pathologie et de localisation géographique dans le corps) et de planifier à l'avance des interventions cliniques qui nécessitent l'ouverture du corps du patient ou l'introduction d'outils (comme les trocarts en chirurgie coelioscopique). De plus, les systèmes les plus récents permettent, à partir de coupes scanner ou IRM, d'afficher en 3D les organes et les structures pathologiques du patient, facilitant ainsi la compréhension spatiale du praticien (cf. figure 1.1).

Néanmoins, l'exploitation de cette précieuse information, visible sur un écran, est limitée durant l'intervention. En effet, le praticien est principalement confronté à la difficulté de la mettre en correspondance avec sa vision directe du patient. Il doit donc réaliser un repositionnement mental dont la qualité va dépendre de sa capacité à se servir des repères anatomiques du patient. Pour éviter au praticien de devoir réaliser lui-même cette tâche abstraite limitant sa précision, il faudrait qu'il puisse visualiser simultanément la réalité et l'image interne du patient mises en correspondance. Pour surmonter cette difficulté, il faut résoudre deux problèmes distincts :

- le calcul de la position 3D des données internes au patient durant l'acte.

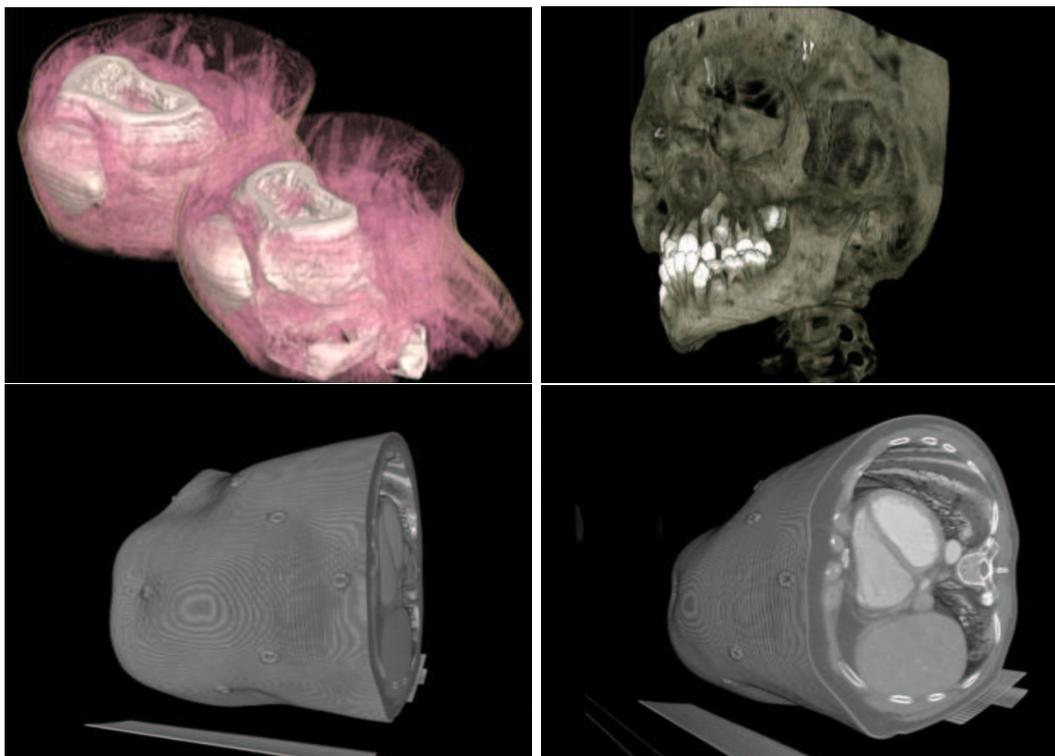


FIG. 1.1 – En haut: exemple d'images provenant du logiciel *Vworks*TM (source : <http://www.cybermed.co.kr/eclinical.html>). Rendu volumique de genoux (à gauche) et du crâne (à droite). En bas: rendu volumique du tronc d'un patient (source : <http://www.sintef.no/medtek/cx/cx-samples5/lapop0304.html>)

- la mise en place d'un système de visualisation permettant de fusionner les deux ensembles de données.

La fusion de l'information peut être réalisée selon deux manières différentes en fonction du référentiel choisi pour guider le praticien : nous distinguons les systèmes **centrés utilisateur** et les systèmes **centrés image**. Le premier mode d'interaction consiste à rajouter l'information dans le champ de vision du praticien qui regarde le patient. Le second type de système resitue l'action du praticien relativement aux données internes du patient. Dans ce cas, le praticien visualise sur son écran l'impact de ses mouvements dans un repère lié à l'image interne du patient. Selon la taxonomie introduite par [Dub01], un *système mixte* désigne l'un ou l'autre de ces systèmes. Par souci de clarté, nous remplacerons au cours de notre exposé système mixte par **système de réalité augmentée**.

Le calcul de la position spatiale de l'image interne du patient nécessite la détermination de la transformation géométrique reliant le système de coordonnées du scanner à celui du dispositif d'acquisition utilisé pour repérer le patient dans la salle d'opération. Les hypothèses de ce calcul, connu sous le nom de *recalage*, sont différentes suivant que le dispositif d'acquisition fonctionne en temps réel (cas de l'imagerie US) ou en temps différé (cas de la plupart des imageurs TDM¹ et IRM.). Dans le premier cas, le problème est fondamentalement rigide² et

1. Abréviation de Tomodensitométrie

2. Nous admettons que l'image subit un pré-traitement afin de limiter les éventuelles distorsions spatiales que

nécessite un repérage en temps réel de la source d'acquisition par rapport au patient. Dans le second cas, l'information pré-opératoire acquise en différé est différente de la réalité avec laquelle on souhaite la fusionner. Nous distinguons alors deux possibilités : soit la différence est supposée négligeable (cas où le patient est immobilisé et la pathologie statique par rapport au patient) et la mise en correspondance est rigide, soit les déformations et mouvements potentiels sont trop importants et il faut alors essayer de les modéliser afin de les compenser par une mise en correspondance non-rigide (cas où le patient a bougé significativement entre le moment de l'acquisition et de l'intervention).

Un système de guidage thérapeutique pour le foie

Le projet EPIDAURE de l'INRIA s'est donné pour but de créer, développer et étudier de nouvelles méthodologies dans le domaine du traitement informatisé de l'imagerie médicale (cf. article [Aya03]). Les domaines de prédilection de l'équipe de ce projet sont la segmentation et le recalage d'images, l'analyse de séquences d'images et la simulation d'organes.

L'IRCAD est un institut privé dont l'axe de recherche est le cancer de l'appareil digestif. Le projet Virtual-Surg développe des méthodes de traitement d'image pour la reconstruction des organes abdominaux et la détection automatique des zones tumorales du foie et du côlon. L'équipe de ce laboratoire dispose notamment d'un logiciel pouvant reconstruire automatiquement en 3D les organes d'un patient à partir d'une acquisition TDM de son abdomen [SDM⁺01].

De la synergie de ces deux projets est née l'idée de réaliser un système de réalité augmentée appliquée à la zone abdominale (et plus particulièrement au foie), permettant de faire fusionner des informations pré-opératoires et/ou per-opératoires avec le monde réel afin de guider le praticien. Comme évoqué dans la partie précédente, la difficulté de réalisation d'un tel système est double. Nous devons recalibrer les informations internes au bon endroit par rapport à la réalité puisque tout écart conduirait le praticien à se laisser guider de manière erronée, ce qui pourrait mettre en jeu la vie du patient. Parallèlement, une interface de visualisation adaptée aux besoins du praticien doit être créée pour permettre de fusionner les deux ensembles d'informations.

Des systèmes s'appuyant sur les structures osseuses...

Actuellement, il existe déjà des systèmes mettant en correspondance le monde réel et l'information virtuelle issue des images pré-opératoires. En neurochirurgie, 3D SlicerTM [GNK⁺99], ACUSTARTM [MFW⁺97] et VectorVisionTM [GWL⁺99] sont des outils centrés image permettant de visualiser la sonde manipulée par le praticien en la superposant aux données pré-opératoires (cf. figure 1.2). Les systèmes développés dans [EKM⁺00](MAGITM) et [JMF⁺02], superpose des données pré-opératoires à la vue du chirurgien au travers d'un microscope grâce à des LCD semi-transparentes intégrées au niveau des optiques (cf. figure 1.3). Dans [MMN⁺00], un système composé d'un miroir semi-transparent rattaché au scanner permet au praticien de voir à la fois le patient au travers du miroir et les coupes scanners correspondantes en superposition (cf. figure 1.4). Ces deux systèmes affichant l'information dans le champ visuel du praticien sont centrés utilisateur.

En ce qui concerne la craniotomie, dans [GEW⁺96], la tête du patient est filmée par une caméra fixe et le flux vidéo est augmenté par le modèle de la tumeur modélisée à partir d'acquisitions TDM pré-opératoires. Le chirurgien peut ainsi dessiner les zones d'intérêts sur le cuir chevelu du patient en guidant son geste grâce à l'image vidéo (cf. figure 1.5). Bien que la

pourrait provoquer le système d'acquisition.

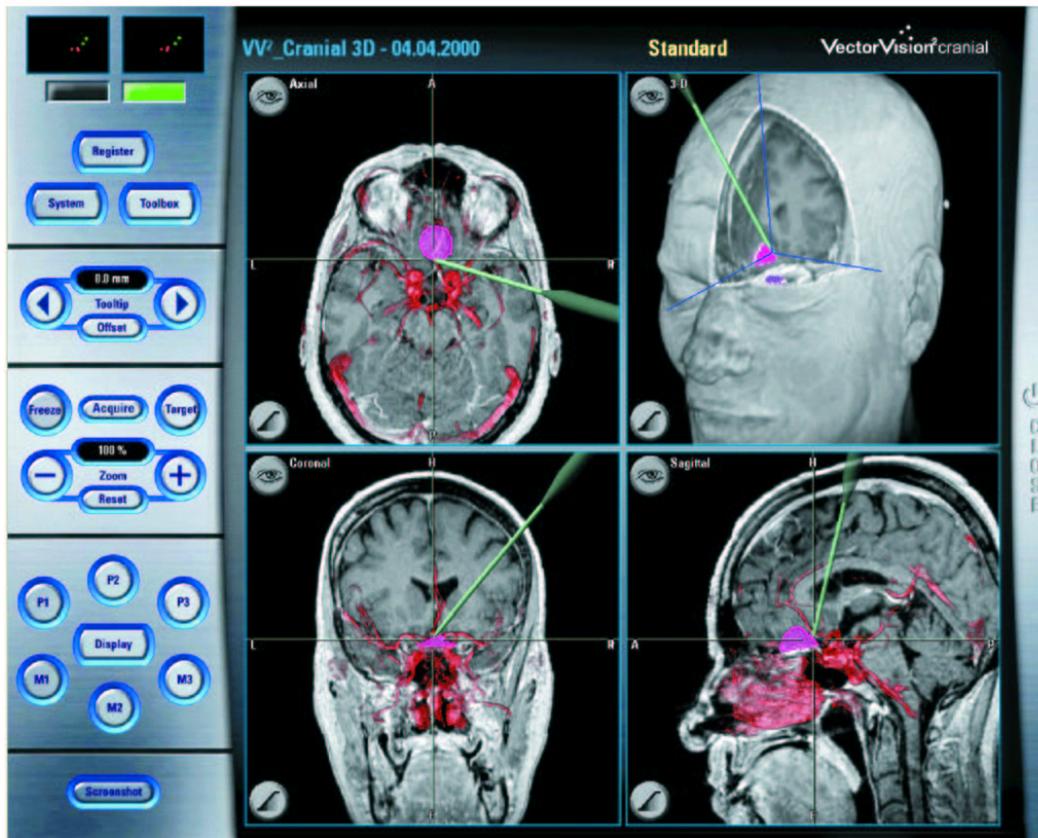


FIG. 1.2 – Le système *VectorVision™* de la société *BrainLab Inc.* [GWL⁺99] (source : <http://www.brainlab.com/>)

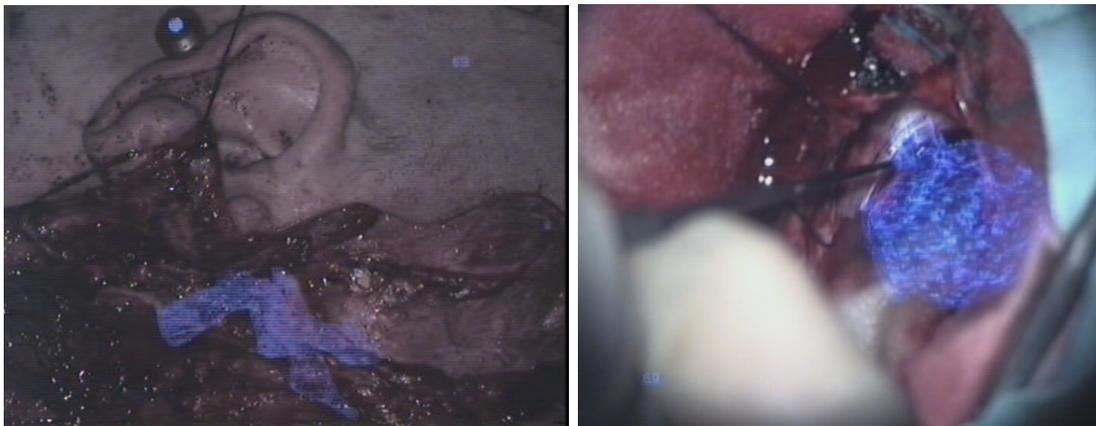


FIG. 1.3 – Le système *MAGI™* développé dans [EKM⁺00]. A gauche : superposition sur le crâne de l'arche zygomatique et de l'artère carotide. A droite : superposition du neurome acoustique (source : <http://www-ipg.umds.ac.uk/magi/>).



FIG. 1.4 – *Superposition sur un miroir semi-transparent de coupes TDM [MMN⁺ 00].*



FIG. 1.5 – *Superposition sur une image vidéo du crâne de structures d'intérêt. Application à la craniotomie [GEW⁺ 96] (source : <http://www.ai.mit.edu/projects/medical-vision/>).*

caméra se substitue au regard du praticien, ce système est centré image puisque le praticien se guide à l'aide d'un point de vue qui ne correspond pas à son point de vue naturel.

Pour finir, dans le domaine de la chirurgie orthopédique, le système centré utilisateur présenté dans Peuchot [PTE95] permet d'afficher sur un miroir semi-transparent (comme le système de Masamune) la position idéale d'une vertèbre après remise en place de la colonne vertébrale. Le système centré image dans Merloz [MLTP00] présente sur un écran vidéo la position de l'outil chirurgical par rapport à des données scanners de la vertèbre dans trois coupes différentes (cf. figure 1.6). En implantologie dentaire, le projet VirtualScope [Gra03] a pour objectif de guider le perçage des gencives en visualisant sur des données pré-opératoires le foret manipulé par le dentiste.

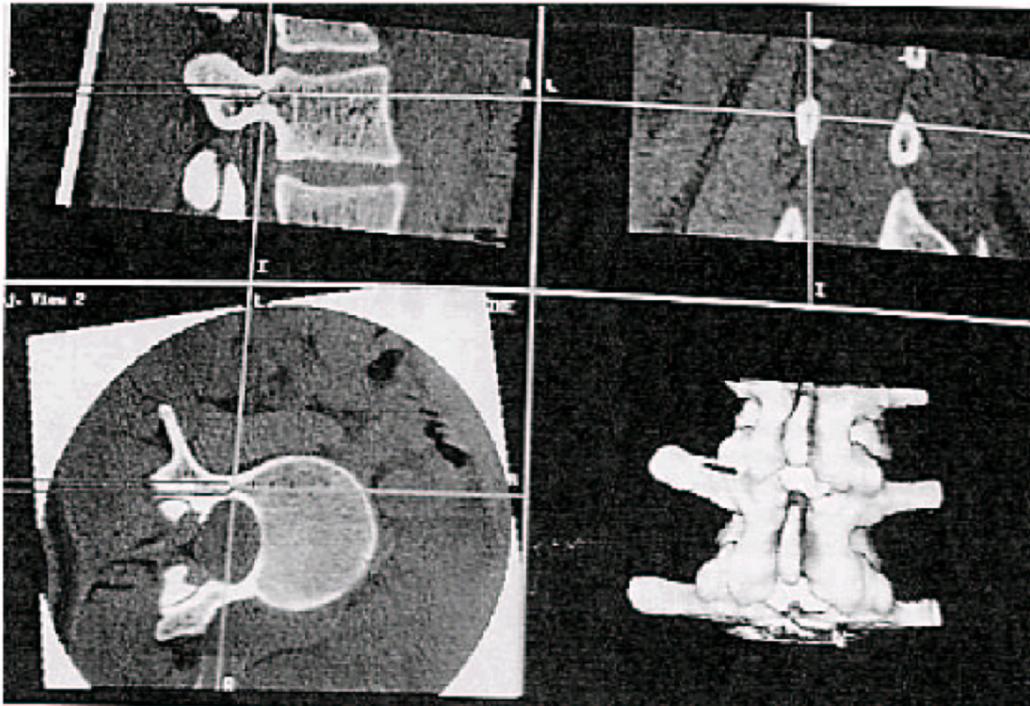


FIG. 1.6 – Informations de guidage issues du système de visée pédiculaire dans [MLTP00].

Afin de guider les biopsies de la poitrine, Rosenthal [RSL⁺02] et Sauer [SKB⁺01] utilisent un appareillage portable vidéo et rajoutent dans le champ de vision du praticien la coupe 2D acquise par la sonde US qu'il manipule ainsi que la position spatiale de l'aiguille de biopsie (cf. figure 1.8). Stetten [SSC⁺02, SCHB01, SC01], quant à lui, propose de superposer la coupe échographique sur un miroir semi-transparent rigidement lié à l'échographe (cf. figure 1.7). Navab et Mitschke [MBHN00, NBHM99] guide le positionnement d'une aiguille à l'aide d'un CAMC ("Camera Augmented Mobile C-arm") qui permet de superposer une image radiographique sur une image vidéo externe du patient.

Actuellement, tous les systèmes de réalité augmentée employés cliniquement utilisent une mise en correspondance rigide, que l'information provienne d'un système d'acquisition en temps réel ou différé. Dans le second cas, l'hypothèse d'absence de mouvement est généralement vérifiée puisque les zones de repères concernées sont des zones rigides comme les os ou la boîte



FIG. 1.7 – *Guidage d'une biopsie à l'aide d'un système centré utilisateur. Le praticien voit le plan d'acquisition de la sonde ainsi que la position de l'aiguille [RSL⁺ 02].*

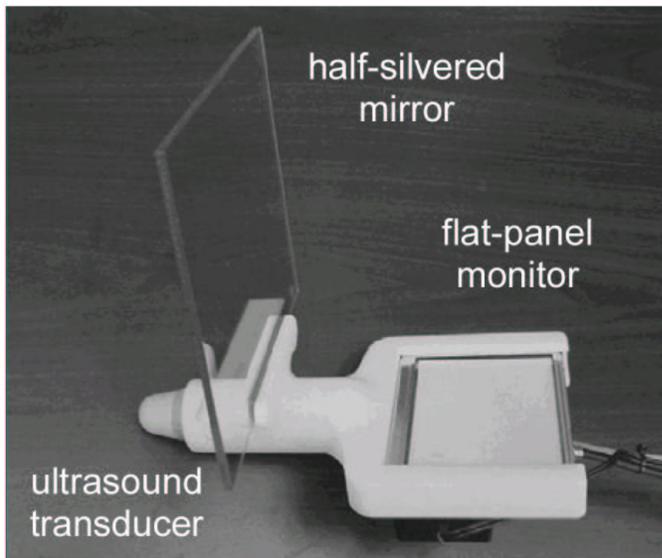


FIG. 1.8 – *Visualisation d'une coupe échographique dans le champ visuel de l'utilisateur à l'aide d'un miroir semi-transparent lié à l'échographe [SSC⁺ 02, SCHB01]. A gauche : le Sonic FlashlightTM. A droite : illustration de la vue fournie à l'utilisateur. La coupe de la main, acquise par l'échographe, est réfléchié par le miroir semi-transparent dans le champ visuel de l'utilisateur.*

crânienne³. En pratique, la marge d'erreur de ces systèmes (inférieure à 2 mm) permet de satisfaire les exigences imposées par les interventions neurochirurgicales ou orthopédiques. En revanche, il existe très peu d'études de faisabilité sur les zones thoraciques et abdominales, extrêmement pauvres en repères rigides, à cause des mouvements respiratoires. La mise en correspondance concerne, dans ce cas, des zones mobiles, cela signifie que la qualité du repositionnement du monde virtuel par rapport au monde réel risque de ne pas satisfaire les exigences de précision de l'intervention. Mourgues [Mou03], par exemple, repositionne les artères coronaires dans un contexte de chirurgie endoscopique avec une précision de l'ordre de 8 pixels (erreur due principalement au fait que le modèle virtuel à repositionner n'est pas déformable. Cf. figure 1.9). Banovac [Ba02] a réalisé un système permettant de guider les biopsies du foie en visualisant la position de l'aiguille relativement aux coupes scanners pré-opératoires. Une expérience, réalisée sur un mannequin de l'abdomen contenant un foie et des cibles, indique que sur 16 ponctions, 14 ont été effectuées correctement. Néanmoins, aucune évaluation de la précision du positionnement de l'aiguille dans la cible n'a été réalisée. Buck [DBCG⁺01] présente des résultats préliminaires en laparoscopie, cependant il n'aborde pas les problèmes liés à la respiration et considère que la zone abdominale est rigide. Une étude de faisabilité présentée dans [LYT⁺02], propose de fournir une aide visuelle au cours d'une intervention en laparoscopie en affichant la position de l'outil manipulé par le praticien dans le repère lié à l'image 3D pré-opératoire (cf. figure 1.10). Cette étude suppose aussi que l'abdomen est rigide, et un protocole de mesure non-décrit a évalué la précision de la superposition à 9 mm en moyenne.

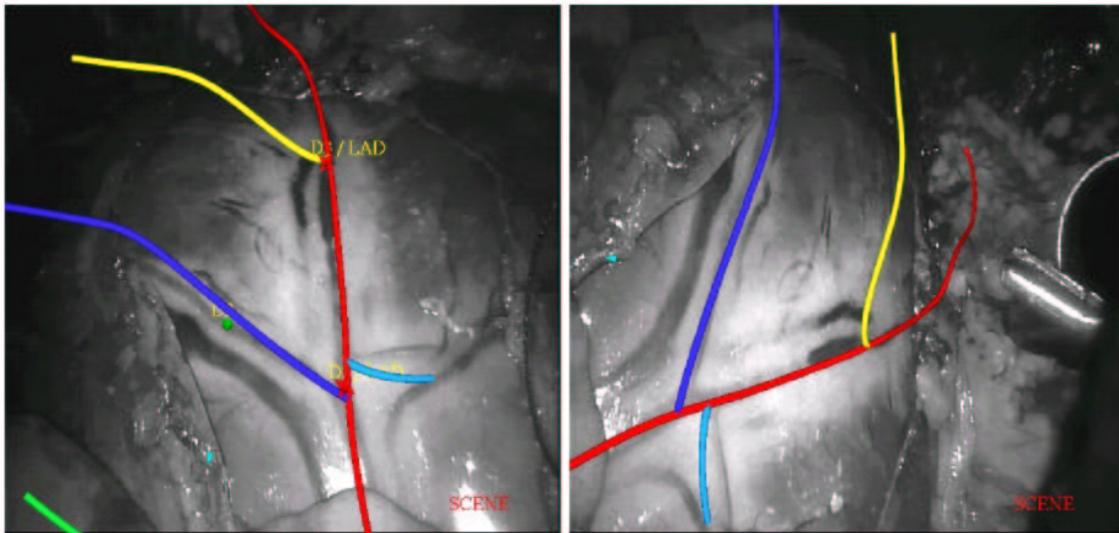


FIG. 1.9 – *Superposition d'un modèle des artères coronaires sur des images arthroscopiques. Le modèle est créé à l'aide d'une angiographie réalisée préalablement et le recalage est réalisé de manière interactive [Mou03].*

3. Les systèmes de réalité augmentée appliqués en neurochirurgie supposent l'absence de mouvement, bien que l'ouverture de la boîte crânienne provoque un effondrement local des structures cérébrales. Des travaux sont en cours afin d'étudier et modéliser ce phénomène d'affaissement (cf. [PCA03, PARC01, SND01, FSJ01]).

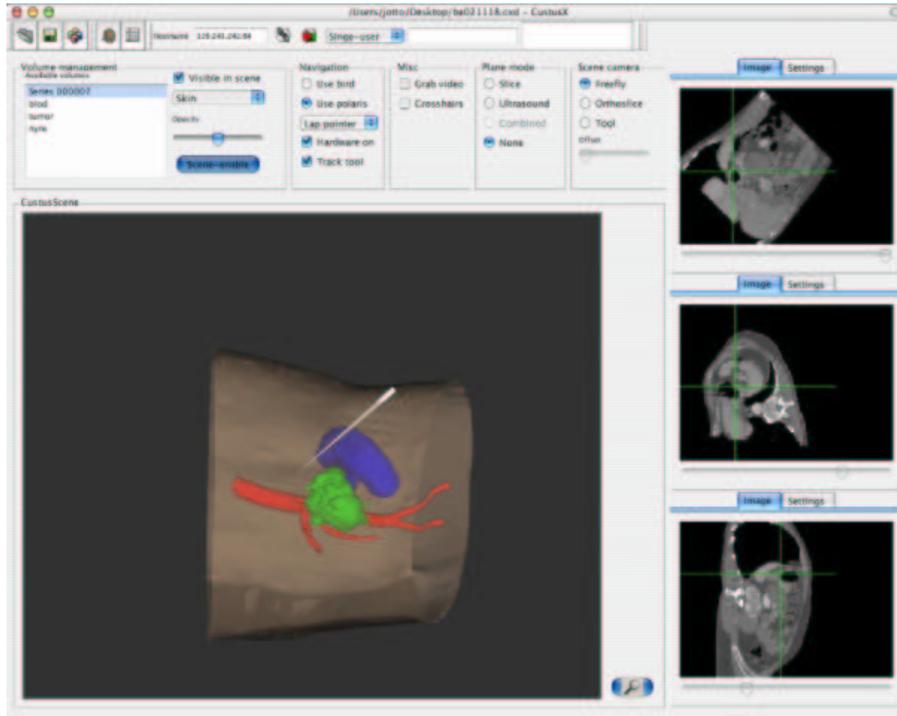


FIG. 1.10 – Informations de guidage issues du système de [LYT⁺02] en chirurgie laparoscopique. L'outil manipulé par le praticien et repéré en temps réel est affiché relativement à l'acquisition scanner du patient réalisée avant l'intervention.

Définition des objectifs

Ce bref état de l'art indique que les systèmes de réalité augmentée qui se focalisent sur les zones mobiles comme le thorax ou l'abdomen sont encore en cours d'étude et/ou d'évaluation à cause des effets difficilement prévisibles de la respiration sur la position des organes. Les quelques résultats quantitatifs obtenus par des expériences in vivo indiquent qu'une étude des effets de la respiration semble incontournable pour obtenir une précision comparable à celle des systèmes déjà existants dans les domaines adjacents (neurochirurgie...). Cette thèse tiendra compte de ce problème afin d'atteindre au mieux son objectif qui est de développer un système de réalité augmentée pour le foie. Considérant les caractéristiques techniques et pratiques des systèmes de réalité augmentée déjà utilisés en routine clinique, nous pouvons résumer en trois points les contraintes fondamentales que notre système devra surmonter.

- **Utilisable au bloc opératoire.** Le système ne doit pas gêner le praticien dans ses mouvements. Nous devons donc considérer à chaque étape de la conception l'encombrement du système.
- **Apport d'un bénéfice durant l'acte médical.** Ce bénéfice pourra être un gain de temps et/ou de précision par rapport aux conditions habituelles de l'intervention ou alors la conception d'une interface d'un type nouveau profitant à l'ergonomie des systèmes de réalité augmentée en général.
- **Prise en compte des effets de la respiration.** Une étude sera vraisemblablement nécessaire afin de quantifier l'impact de la respiration sur les déplacements des organes abdominaux.

1.1 Organisation du mémoire

Notre objectif est de réaliser un système de réalité augmentée appliqué aux opérations du foie. Les types d'opération étant nombreux, nous expliquons dans le **chapitre 2** les contraintes matérielles et temporelles qui nous conduisent à nous focaliser sur la radiologie interventionnelle plutôt que la chirurgie. Notre application étant fixée, nous pouvons définir notre cahier des charges. Afin d'apporter une aide significative, ce système doit être *précis* (erreur de recalage inférieure à 5 mm), *fiable* (probabilité négligeable de dépasser 5 mm d'erreur) et *rapide* (la procédure d'extraction et de traitement des données ne doit pas durer plus de 15 minutes). Nous choisissons ensuite de réaliser la fusion d'informations sur le flux vidéo de caméras visionnant le patient (leur nombre sera choisi dans une section ultérieure) : nous avons estimé que les autres technologies dédiées aux systèmes de réalité augmentée ne sont pas encore assez performantes en terme de précision. Finalement, nous soulignons que l'atteinte des objectifs dépend essentiellement du type de recalage retenu pour notre système. Ce problème est abordé dans le **chapitre 3**.

Le système que nous souhaitons mettre au point n'étant utilisable que si sa précision est suffisante, il est primordial d'analyser les différentes méthodes de recalage susceptibles de convenir à notre application. Nous montrons, tout d'abord, qu'un recalage rigide est un choix raisonnable dans les conditions présentes en radiologie interventionnelle. Ensuite, nous remarquons que la couleur homogène et la forme cylindrique de l'abdomen sont susceptibles de limiter la précision des recalages de type surfacique et iconique, ce qui nous conduit à choisir une méthode de recalage ponctuelle 3D/2D basée sur des marqueurs radio-opaques qui seront préalablement collés sur la peau du patient. Néanmoins, les hypothèses implicites des critères usuels de recalage 3D/2D ne tiennent pas compte de toute l'information disponible sur les caractéristiques du bruit corrompant les données. Nous dérivons donc dans le **chapitre 4** un nouveau critère 3D/2D au sens du maximum de vraisemblance généralisant les approches habituelles et intégrant l'ensemble des informations relatives au bruit d'extraction sur les données. Son évaluation sur des données synthétiques met en évidence un gain en précision allant jusqu'à 20% sur les autres critères et montre une robustesse accrue face à des initialisations éloignées de la solution recherchée. Une évaluation sur des données réelles a tout d'abord nécessité la mise en place d'un protocole de mesures original permettant de calculer l'erreur de recalage *sans mesure étalon*. Nous montrons ainsi sur des données provenant d'un mannequin de l'abdomen que nous pouvons compter sur un gain en précision de l'ordre de 9% avec notre critère.

Arrivé à ce point, nous avons choisi la méthode de recalage à employer ainsi que la méthode de fusion des informations réelles et virtuelles. Il nous faut encore aborder d'autres détails techniques comme le calibrage des caméras et le repérage de l'aiguille manipulée par le chirurgien. Ceci fait l'objet du **chapitre 5**. Pour des raisons de précision et de flexibilité, nous choisissons la méthode de Zhang afin de calibrer nos caméras et nous proposons une version améliorée d'une librairie de repérage LGPL (l'ARTkit) afin de calibrer et repérer l'aiguille en temps réel. Dans ce chapitre, nous montrons que les problèmes de calibrage des caméras et du repérage de l'aiguille peuvent être résolus relativement aisément dans un contexte médical contraignant, tout en fournissant une précision suffisante face aux objectifs. Obtenir une précision légèrement supérieure nécessiterait des moyens beaucoup plus lourds et coûteux (banc de calibration) qui compliqueraient l'introduction du système dans un bloc opératoire.

Nous disposons maintenant d'un système possédant toutes les briques nécessaires à une évaluation en bloc opératoire, néanmoins il faut encore que le praticien localise les données dans les images scanners et vidéo et les mette en correspondance. Bien qu'un système interactif soit une solution envisageable, le praticien devra perdre plusieurs minutes à indiquer la position des marqueurs dans les images selon un ordre particulier, et s'exposera au risque d'une erreur humaine (très probable si tous les marqueurs ne sont pas visibles dans les images vidéo). Nous préférons automatiser ce travail fastidieux par une procédure informatique fiable. Le **chapitre 6** présente les algorithmes de traitement d'image et de données qui permettent d'automatiser les procédures d'extraction et d'appariement des données issues des images scanner et vidéo.

L'objet du **chapitre 7** nous semble primordial puisqu'il traite de la fiabilité du système, un des points clés de nos objectifs. Nous avons développé un système dont la précision dépend de divers paramètres variant d'une intervention à l'autre (angle entre les caméras, nombre et configuration géométrique des marqueurs radio-opaques...). Ces paramètres ne pouvant être optimaux à chaque fois, la précision n'est pas forcément garantie. Il est donc crucial de fournir au chirurgien un moyen de savoir quand la précision n'est pas suffisante et comment il pourrait l'améliorer, en rajoutant des marqueurs radio-opaques par exemple. Dans ce chapitre, notre but est de fournir une prédiction fiable de l'erreur moyenne de recalage en fonction des paramètres géométriques du système et du bruit d'extraction sur les données. Pour cela, nous réalisons une étude de propagation du bruit au travers du critère d'optimisation employé. Nous obtenons ainsi une formulation analytique de la covariance sur la transformation rigide recherchée, qui nous permet de prédire théoriquement l'erreur de recalage. Cependant, le résultat obtenu découle de certaines approximations qui ne seront pas forcément vérifiées dans un contexte réel. En conséquence, une validation rigoureuse est réalisée afin de prouver pas à pas que toutes les approximations réalisées ont des effets négligeables sur la qualité de notre prédiction. Nous démontrons ainsi que nous sommes capables de prédire correctement l'erreur de recalage du modèle reconstruit du patient, mais que si nous souhaitons pouvoir prédire l'erreur de repérage de l'aiguille, une modélisation plus poussée est nécessaire.

Le **chapitre 8** clôt ce travail en présentant nos expériences de guidage pré-cliniques sur notre mannequin et avec des patients. L'évaluation de notre système de guidage, réalisée sur le mannequin, donna des résultats très encourageants. Quatre chirurgiens sont parvenus à atteindre des cibles disposées sur le foie du mannequin avec une erreur moyenne de 2.1 mm sur 10 essais. Cette expérience in vitro soulignant les lacunes de notre système au niveau de l'ergonomie et de l'interface, nous décrivons ensuite les outils et les fonctionnalités supplémentaires que nous rajoutons afin d'assurer un contrôle optimal du chirurgien durant l'intervention. Nous terminons ce chapitre en présentant nos premières expériences avec des patients qui nous permettent de comparer la précision de notre système avec celle du chirurgien. Nous montrons ainsi que notre système peut être introduit au bloc opératoire et qu'une phase d'évaluation sur des patients est déjà en cours.

1.2 Contributions

Nos contributions peuvent être regroupées dans 3 parties.

Contributions théoriques

D'un point de vue général, notre principale contribution théorique a été d'élaborer un nouveau critère de recalage ponctuel 3D/2D (baptisé EPPC) généralisant les approches classiques en introduisant la position 3D exacte des points comme une variable auxiliaire inconnue, et de fournir et valider une expression analytique de la covariance sur la transformation à l'optimum. Plus précisément, nous montrons dans le **chapitre 4** que ce nouveau critère est adapté à notre problème et qu'il est plus performant que les critères classiques 3D/2D sur des données synthétiques et réelles. En passant nous réalisons pour la première fois une étude comparative entre deux critères classiques de recalage 3D/2D (ISPPC et OSPPC) qui montre leur similitude en terme de précision. Nous proposons aussi un protocole de mesure original permettant d'évaluer l'erreur de recalage sans mesure étalon. Bien que cette idée ait été inspirée de [PT97], c'est la première fois qu'une application avec des images vidéo est envisagée. L'ensemble des contributions de ce chapitre a fait l'objet d'une publication à la conférence IS4TM⁴ [NPSA03a].

Dans le **chapitre 7**, nous proposons pour notre critère EPPC et pour le critère classique ISPPC, une expression analytique de la covariance sur la transformation, solution de l'optimisation. De plus, notre validation rigoureuse permet d'évaluer l'influence des approximations utilisées. Nous mettons ainsi en évidence le rôle de l'erreur de calibrage si l'on souhaite prédire correctement des erreurs de très faible amplitude. Ce chapitre, initialement déposé sous la forme d'un rapport de recherche [NPSA03b], a été publié et présenté oralement à la conférence ECCV⁵ [NPSA04].

Contributions pratiques

L'élaboration d'un système de réalité augmentée n'est pas envisageable sans un travail d'ingénierie considérable. La mise au point des algorithmes permettant un calibrage rapide et une automatisation des tâches fait appel à des connaissances existantes. Malgré ceci, un long travail pratique de mise au point et d'ajustement de paramètres est nécessaire pour que toutes les briques du système fonctionnent correctement une fois qu'elles sont assemblées. Les **chapitres 5 et 6** contiennent toutes les informations relatives à ce travail qui permettront à un utilisateur souhaitant reprendre notre application d'alléger les multiples essais nécessaires au fonctionnement global du système.

En outre, nous montrons dans le **chapitre 5** qu'il est possible avec des moyens très accessibles de fournir un système calibré repérant précisément un outil dans l'espace, ce qui permet de le rendre indépendant des outils coûteux dédiés au repérage spatial. Les résultats indiquant la précision du recalage en fonction de la prise en compte ou non de la distorsion et du nombre de caméras ont été présentés au colloque SFO⁶ [NPSA03c].

Dans le **chapitre 6**, nous présentons nos algorithmes d'extraction et d'appariement de marqueurs radio-opaques. Comme les marqueurs couramment utilisés en chirurgie sont systématiquement extraits interactivement, nous pensons que notre travail contribuera à l'automatisation d'une tâche inutile et fastidieuse pour le praticien. Par ailleurs, les algorithmes ont été

4. International Symposium on Surgery Simulation and Soft Tissue Modeling (IS4TM'03)

5. European Conference on Computer Vision (ECCV'04)

6. Actes du quatrième colloque francophone. Méthodes et Techniques Optiques pour l'Industrie

comparés aux performances humaines et les résultats ont démontré que ces deux algorithmes peuvent fournir une précision équivalente sinon meilleure. L'architecture globale du système et les procédures d'automatisation ont fait l'objet d'un article dans le *Journal of Computer Animation and Virtual World* [NGP⁺04].

Contributions médicales

La validité de notre travail s'est appuyée sur deux hypothèses. La première était relative à l'erreur de repositionnement du foie au cours du cycle respiratoire que nous supposions très faible. La seconde correspondait aux erreurs de recalage que nous obtiendrions dans les conditions de notre intervention : notre expérience nous suggérait qu'elles seraient suffisamment faibles pour satisfaire nos exigences de précision. Ces deux aspects ont été validés dans les **chapitres 3 et 8**. Tout d'abord, nous avons réalisé une synthèse des travaux existants sur les déplacements du foie en fonction de la respiration. Cette recherche préalable a permis de fixer des limites à l'emploi d'un recalage rigide en radiologie interventionnelle lorsque le patient respire librement ou lorsque le patient itère des mises en apnée. Nous validons ainsi, dans le chapitre 3, la pertinence d'une approche rigide, puisque nous montrons que l'erreur de repositionnement du foie est de l'ordre de 1 mm lorsque le volume d'air dans les poumons est contrôlé.

Toutes les briques du système ont été finalement assemblées et notre système de guidage par réalité augmentée a été évalué sur un mannequin de l'abdomen. Les résultats de précision et de rapidité obtenus ont permis de valider le fonctionnement couplé de toutes les parties conçues et évaluées de manière indépendante. Les résultats de l'évaluation ont été présentés au workshop AVIR⁷[NGP⁺03a], au colloque IMVIE⁸ [NGP⁺03b] et ont été publiés dans le journal francophone REE⁹ [SAN⁺04]. Nous avons présenté en dernier lieu nos premiers résultats obtenus sur des patients. Nous montrons ainsi que le système conçu peut être introduit en salle d'opération et que nous poursuivons notre travail afin d'évaluer et valider l'efficacité du système avec des patients.

7. Augmented and Virtual Reality Workshop (AVIR03)

8. Actes du colloque Imagerie pour les sciences du vivant et de l'ingénieur (IMVIE03)

9. La Revue de l'Electricité et de l'Electronique

Chapitre 2

Un système de guidage thérapeutique pour le foie

Notre but est de concevoir un système de réalité augmentée afin de guider le praticien et d'améliorer la sécurité de son geste lors d'opérations du foie. A l'IRCAD, les deux grands axes d'interventions possibles parmi ceux existants sont la chirurgie et la radiologie interventionnelle (cf. section 2.1.3). Dans le contexte chirurgical, il serait envisageable de créer un système de réalité augmentée utilisant une imagerie temps réel (sonde à ultra-sons 2D ou 3D). Néanmoins, l'IRCAD n'a pas accès à un imageur de ce type mais à un scanner à rayons X pouvant imager l'ensemble de l'abdomen en environ deux minutes.

Dans une première partie, nous détaillerons les conditions cliniques de ces deux types d'intervention. Nous verrons alors pourquoi les conditions opératoires en radiologie interventionnelle, proches de celles des domaines déjà explorés (chirurgie orthopédique, neurochirurgie), nous conduisent à focaliser notre travail sur cette intervention plutôt que sur la chirurgie.

L'application étant décidée, nous devons choisir une technologie de réalité augmentée pour faire fusionner les informations réelles et virtuelles. Dans une seconde partie, nous détaillerons les différents choix possibles et nous montrerons qu'il est préférable de superposer l'information sur une vue externe du patient retransmise par un moniteur, plutôt que de nous servir d'un appareillage portable ou d'un autre système technologique. Nous clôturons le chapitre en soulignant le rôle central que jouera le recalage pour remplir nos objectifs.

2.1 Réalité augmentée et opération du foie

2.1.1 La chirurgie ouverte

Lors d'une opération de chirurgie ouverte, la peau du patient est coupée, les côtes du patient sont écartées, et, suivant la localisation de la zone à traiter (la zone ventrale ou dorsale), le foie est légèrement soulevé ou alors sorti manuellement de la cage thoracique. Cette intervention est nécessaire lors d'une cirrhose ou d'un hépatocarcinome très volumineux. Il faut alors réaliser un transplant ou une exérèse. Dans le cas d'une exérèse, le chirurgien planifie l'opération sur les coupes scanner ou alors, lorsque cela est disponible, sur la reconstruction 3D obtenue après segmentation de l'image TDM. Il s'aide alors d'une identification des segments de Couinaud ou des réseaux sanguins pour déterminer les plans de coupes. Evidemment, en phase opératoire, toutes ces informations ne sont plus disponibles puisque cachées dans le foie, le chirurgien doit donc recalculer mentalement la position des réseaux sanguins à partir des quelques

repères externes dont il dispose (ligament falciforme, dôme...). Ensuite, le chirurgien manipule le foie afin de procéder aux découpes. Un système de réalité augmentée permettrait au chirurgien de visualiser, en regardant le foie, les réseaux sanguins et les zones fragiles, ainsi que toutes les informations qu'il aurait préalablement indiquées sur le modèle pré-opératoire (cf. figure 2.1). Cela lui permettrait d'agir plus rapidement, plus précisément et d'éviter les gestes "en aveugle" potentiellement dangereux.

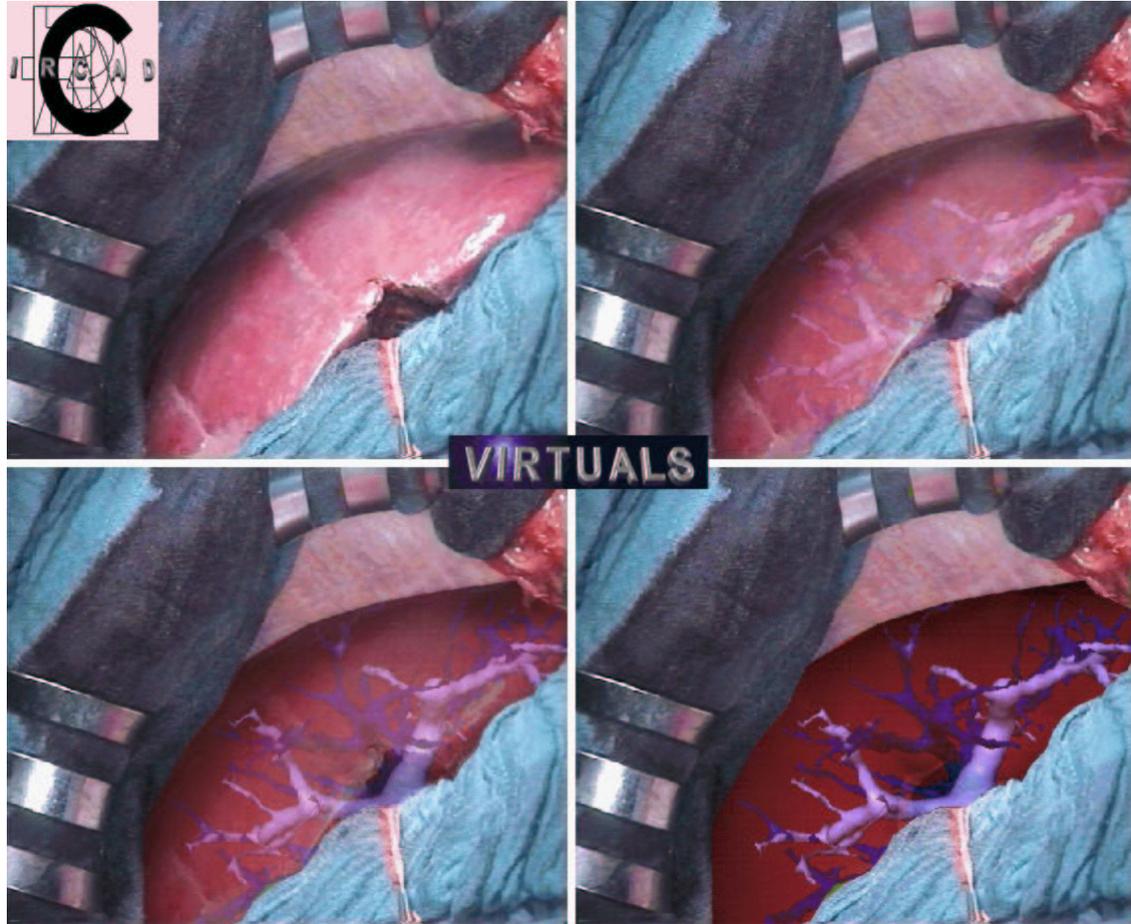


FIG. 2.1 – Exemple de vue augmentée en chirurgie ouverte : le réseau sanguin devient visible à l'intérieur du foie. Le modèle virtuel a été recalé manuellement sur l'image après l'intervention.

Concernant la chirurgie ouverte, Herline [HHS⁺00] a proposé de recalé la surface du modèle reconstruit sur un échantillon de la surface du foie obtenu à l'aide d'un capteur optique. Bien que la précision obtenue pour des cibles intérieures ait été de 2.8 mm, le foie utilisé pour l'expérience était en plastique non déformable, ce qui signifie qu'il supposait que le modèle virtuel du foie à recalé correspondait très précisément au foie réel.

Dans le contexte de la chirurgie cardiaque robotisée, Mourgues [Mou03] recalé sur des images endoscopiques un modèle 3D rigide du réseau sanguin du coeur (obtenu avant l'intervention par angiographie). Le recalage est obtenu en minimisant la distance entre des amers (bifurcations et/ou vaisseaux) relevés manuellement sur les images endoscopiques et leur correspondant sur le modèle 3D. Cette stratégie n'est pas adaptée dans le cas du foie puisqu'aucun point de repère

n'est visible à la surface du foie. De plus, l'emploi d'un recalage rigide suppose une bonne correspondance géométrique entre le modèle du foie et la réalité.

Pour obtenir un modèle reconstruit en 3D correspondant à la réalité physique, nous avons deux possibilités. La première est de nous servir d'un système d'acquisition en temps réel (sonde à ultra-sons 2D ou 3D) repéré dans l'espace, et de superposer, comme dans [RSL⁺02], la coupe acquise dans le champ de vue du praticien. L'IRCAD n'ayant pas eu accès à un dispositif de ce type au début de cette thèse (et toujours pas à l'heure actuelle), il était difficile de s'orienter vers cette option. La seconde alternative est de réaliser un scanner du patient après l'ouverture de son corps puisque le foie est manipulé, et donc déformé, par le chirurgien à ce moment (un scanner du foie réalisé avant l'ouverture ne correspondrait pas en terme de forme au foie extrait de la cavité abdominale). Cependant, cela n'est pas envisageable car il faudrait opérer le patient sur la table du scanner et donc monopoliser le scanner le temps d'une intervention qui durerait plusieurs heures. Nous nous retrouvons donc dans un contexte où l'objet réel peut être très différent de l'objet virtuel à cause des déformations entre le moment de l'acquisition et de l'intervention. Par ailleurs, un acte de chirurgie sur le foie crée deux types d'interaction. D'une part, le foie subit des déformations importantes et globales par l'intermédiaire des outils du chirurgien et de ses mains, d'autre part, le foie interagit avec tous les organes autour de lui (du fait notamment de la respiration du patient). Ainsi, la gestion des actions sur le foie nécessiterait la création d'un modèle mécanique de l'ensemble des organes en interaction. Si certains organes ont déjà fait l'objet de travaux (notamment le foie [Cot97, For03, Pic01] et le coeur [Ser03]), il n'en est pas de même pour la totalité des organes en mouvement et interagissant entre eux (poumons, diaphragme...). La conception d'un système de réalité augmentée pour la chirurgie ouverte à partir d'une image scanner nécessite donc préalablement un travail considérable sur la modélisation et la gestion en temps réel des organes de l'abdomen.

2.1.2 La chirurgie mini-invasive

Suite à la morbidité post-opératoire des interventions chirurgicales classiques qui sont très traumatisantes, la chirurgie laparoscopique s'est révélée être une alternative de premier choix pour une grande partie des interventions qui n'étaient envisageables qu'à corps ouvert. La laparoscopie est initialement une technique de diagnostic qui consiste à examiner l'intérieur des organes ou des cavités internes du corps à l'aide d'un appareil appelé endoscope. La cavité abdominale est préalablement remplie de CO_2 à l'aide d'un insufflateur automatique et l'optique d'une caméra est introduite par une incision découpée sur la peau du ventre. Cet outil de diagnostic, couplé avec de longs instruments passant par d'autres incisions (cf. figure 2.2), permet actuellement de réaliser les opérations chirurgicales ne nécessitant pas l'ablation d'un volume important du foie (qu'il serait difficile d'évacuer). Dans ce type d'intervention, le chirurgien ne regarde plus ses mains mais un écran et les informations de relief ne lui sont plus accessibles directement. De plus, il lui est difficile de s'orienter car la caméra offre un champ de vision limité et est manipulée par un assistant à qui il demande d'effectuer les mouvements désirés. Dans de telles conditions, il est facile d'imaginer tout l'intérêt que peut représenter l'apport d'un système de réalité augmentée. Tout d'abord, comme dans le cas de la chirurgie classique, le chirurgien pourrait "voir" sur son écran les structures internes du foie comme son réseau vasculaire, ainsi que les zones cibles. De plus, nous pourrions fournir un outil de laparo-navigation qui permettrait au chirurgien de savoir où ses instruments sont placés dans la cavité abdominale, sans avoir besoin de bouger la caméra pour les retrouver. D'une part, cela guiderait le chirurgien dans ses mouvements, d'autre part, cela lui permettrait d'éviter des

collisions potentiellement dangereuses entre les organes et ses outils qui n'apparaissent pas dans le champ de la caméra. Le temps de l'intervention en serait réduit et les gestes du chirurgien plus précis.

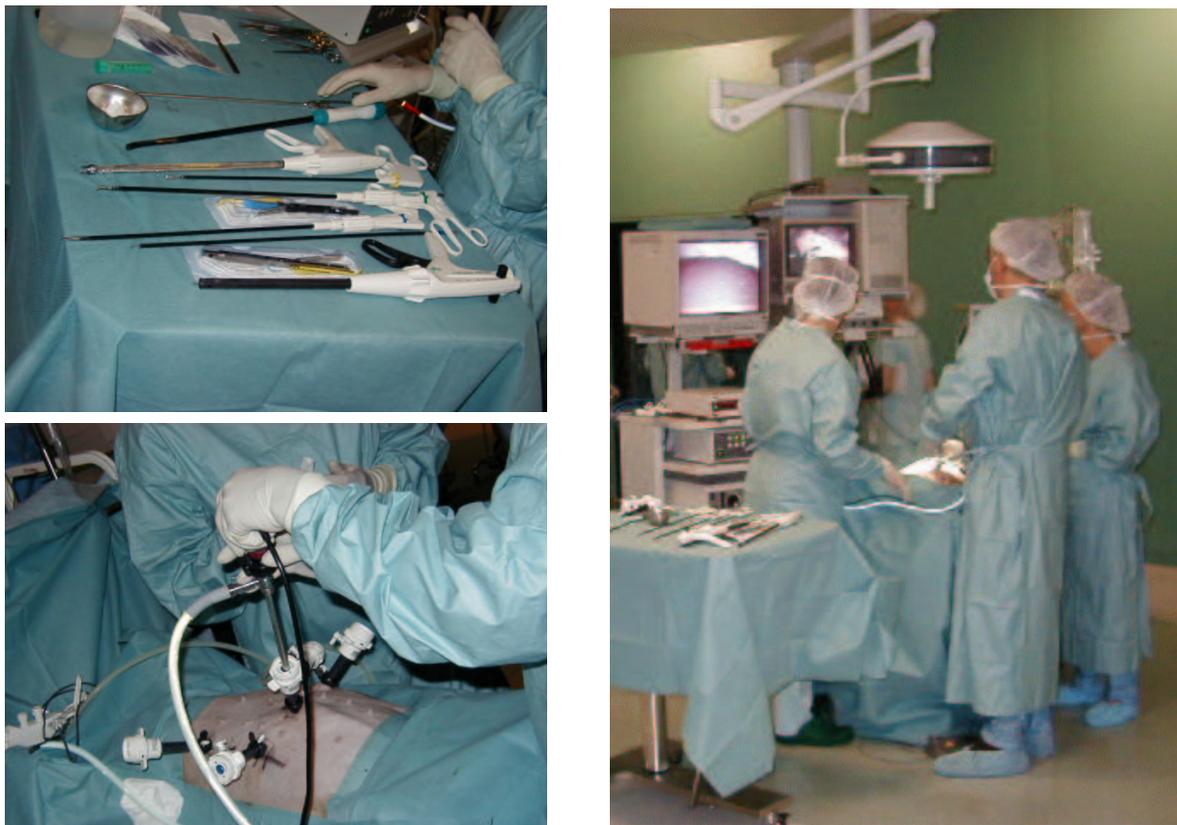


FIG. 2.2 – Exemple de cours de chirurgie laparoscopique à l'IRCAD (sur des cochons). En haut à gauche, l'ensemble des outils utilisés. En bas à gauche, on gonfle la cavité abdominale avec du CO_2 après avoir mis en place l'insufflateur. On peut observer au premier plan les deux trocars qui vont accueillir les instruments et au fond celui qui sera utilisé par la caméra. L'image de droite montre la disposition du praticien et des aides lors de l'opération.

Langø [LYT⁺02] a réalisé la première étude de faisabilité (à notre connaissance) sur un système de réalité augmentée en laparoscopie appliquée au foie. Dans son protocole, le scanner du patient est réalisé la veille de l'intervention et le modèle 3D utilisé est rigide. Cette hypothèse de rigidité est trop irréaliste pour obtenir une précision de repositionnement comparable aux systèmes de réalité augmentée employés en routine clinique. En effet, le patient se relève et bouge entre le moment de l'acquisition et celui de l'intervention. De plus, l'insufflation de CO_2 déforme les organes abdominaux (et modifie la position relative entre la peau et le foie). Ceci explique d'ailleurs que l'erreur soit évaluée à 1 cm environ.

En fait, la conception d'un système de réalité augmentée est soumise au même problème qu'en chirurgie classique. Il n'est pas possible d'obtenir un scanner de l'abdomen juste avant l'opération et les actions mécaniques exercées sur le foie nécessite au préalable la création de modèles mécaniques de l'ensemble des organes de l'abdomen ainsi que la gestion en temps réel des interactions possibles.

2.1.3 Notre but : la radiologie interventionnelle

La radiologie interventionnelle est un acte relativement récent (année 90) qui a pu se développer grâce aux progrès technologiques dans le domaine de l'imagerie médicale (notamment la résolution des images et la vitesse d'acquisition). L'opération consiste à introduire une aiguille dans une cible situés dans un organe abdominal à des fins d'analyse (biopsie) ou de destruction. En fonction de l'intervention, l'aiguille sera munie d'une électrode (dans le cas d'une thermoablation par radio-fréquence) ou bien sera creuse (dans le cas d'une biopsie ou d'un traitement d'hépatocarcinome par alcoolisation ou chimio-embolisation). Cette technique est de plus en plus utilisée en raison de son caractère très peu invasif [MDI01]. En effet, l'intervention nécessite simplement une incision au scalpel d'une dimension de 2 mm ne laissant aucune cicatrice.

Contexte et déroulement d'une opération

Actuellement, cette intervention est réalisée selon deux protocoles différents en fonction de la modalité d'imagerie utilisée par le praticien pour guider son geste : coupes échographiques (US) ou scanner. Dans les deux cas, une planification pré-opératoire est effectuée pour déterminer le chemin que prendra l'électrode pour atteindre la tumeur. En effet, il est primordial de ne pas endommager les vaisseaux et la vésicule. En général, on réalise des coupes scanner de la zone d'intérêt du patient (en l'occurrence le foie) sur lesquelles le chirurgien définira son plan d'action. Actuellement, à l'IRCAD, on obtient en 5 minutes une reconstruction tridimensionnelle du foie, de la tumeur et des organes voisins à partir des coupes TDM (cf. figure 2.3) qui offre au chirurgien une meilleure compréhension spatiale des zones à risque et du point d'entrée de l'aiguille.

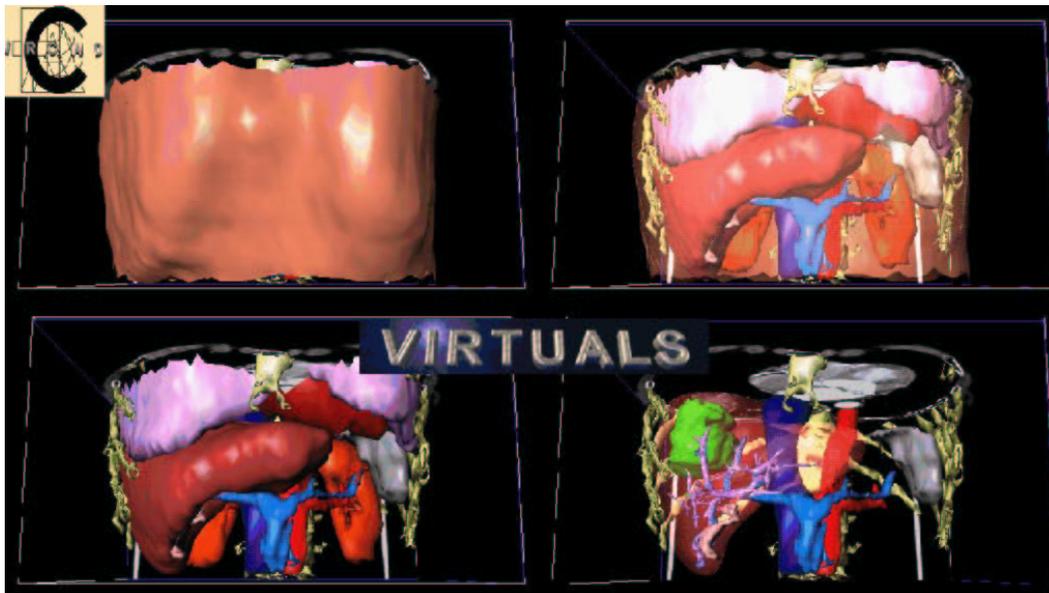


FIG. 2.3 – *Reconstruction tridimensionnelle à partir de coupes scanner. En haut à gauche : la peau abdominale est opaque. En haut à droite et en bas à gauche : la peau disparaît progressivement, laissant apparaître les organes internes du patient. En bas à droite : la surface du foie est transparente et on peut distinguer à l'intérieur la tumeur (en vert).*

Suivi par échographie : Le praticien visualise une coupe du foie grâce à une sonde échographique qu'il peut déplacer à son gré pour observer la zone tumorale. Il se sert de cette information pour guider l'électrode au centre de la tumeur (cf. figure 2.4).

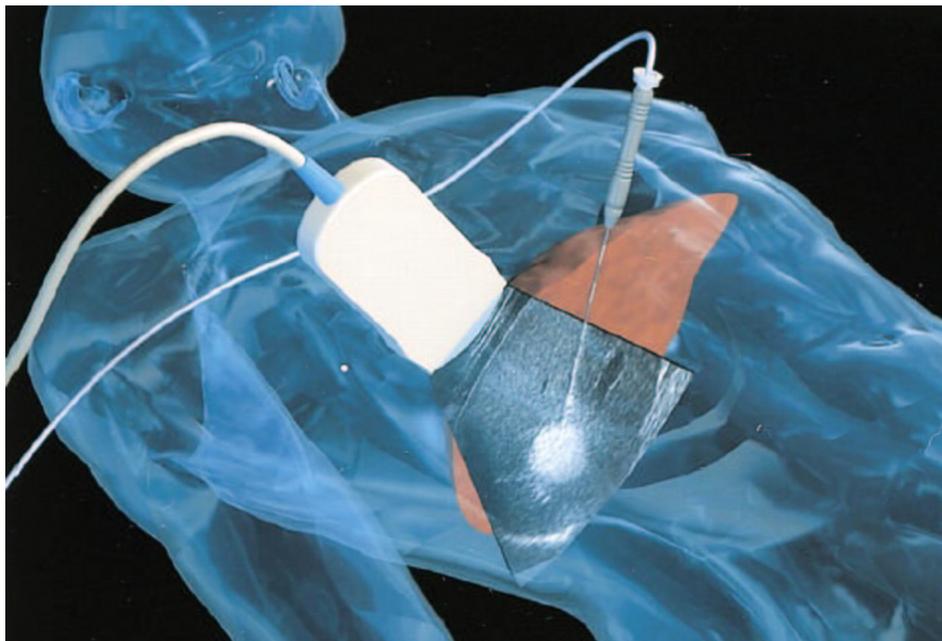


FIG. 2.4 – *Guidage échographique de l'électrode au centre de la tumeur.*

Cette méthode a l'avantage d'être non-irradiante et offre une vue en temps réel de l'évolution du déplacement de l'aiguille par rapport à la tumeur. Néanmoins, la qualité de l'image n'est pas optimale pour se guider et la méthode n'est pas utilisable dans tous les cas. En effet, il arrive souvent que la tumeur ne soit pas révélée par les images échographiques ou bien qu'elle soit située sous une côte, auquel cas l'ombre de la côte empêche la visualisation de la tumeur.

Suivi par coupes scanner : Pour guider l'électrode jusqu'au centre de la tumeur, le praticien peut se servir d'une autre modalité d'imagerie : les coupes scanner (cf. figure 2.5). Contrairement aux opérations chirurgicales, la radiologie interventionnelle guidée par coupes TDM ne dure qu'une trentaine de minutes, il est donc possible de monopoliser le scanner durant toute l'intervention. Le patient est préalablement allongé sur la table du scanner. Il est alors endormi et intubé, ce qui permet de contrôler le volume d'oxygène qu'il respire au cours de l'opération.

La première étape consiste à re-localiser la tumeur dans le foie à partir d'une première acquisition TDM complète du foie. Ensuite, le praticien commence à introduire l'aiguille puis il arrête son geste. Une série de coupes scanner est alors effectuée au niveau de la zone tumorale, il peut donc vérifier la position de l'aiguille et la modifier en conséquence pour atteindre le centre de la tumeur. Il recommence la manipulation de l'aiguille, s'arrête à nouveau, et vérifie, à l'aide d'une nouvelle série de coupes scanner, la position réelle de l'aiguille. En général, le praticien a besoin au minimum d'une quinzaine d'itérations pour parvenir jusqu'à la tumeur, et cela nécessite au moins 20 minutes. Afin de limiter le déplacement continu du foie dû à la respiration, le patient est mis sous apnée par le praticien à chaque fois qu'il déplace l'aiguille.

D'un point de vue efficacité, ces deux méthodes semblent à peu près équivalentes (selon les

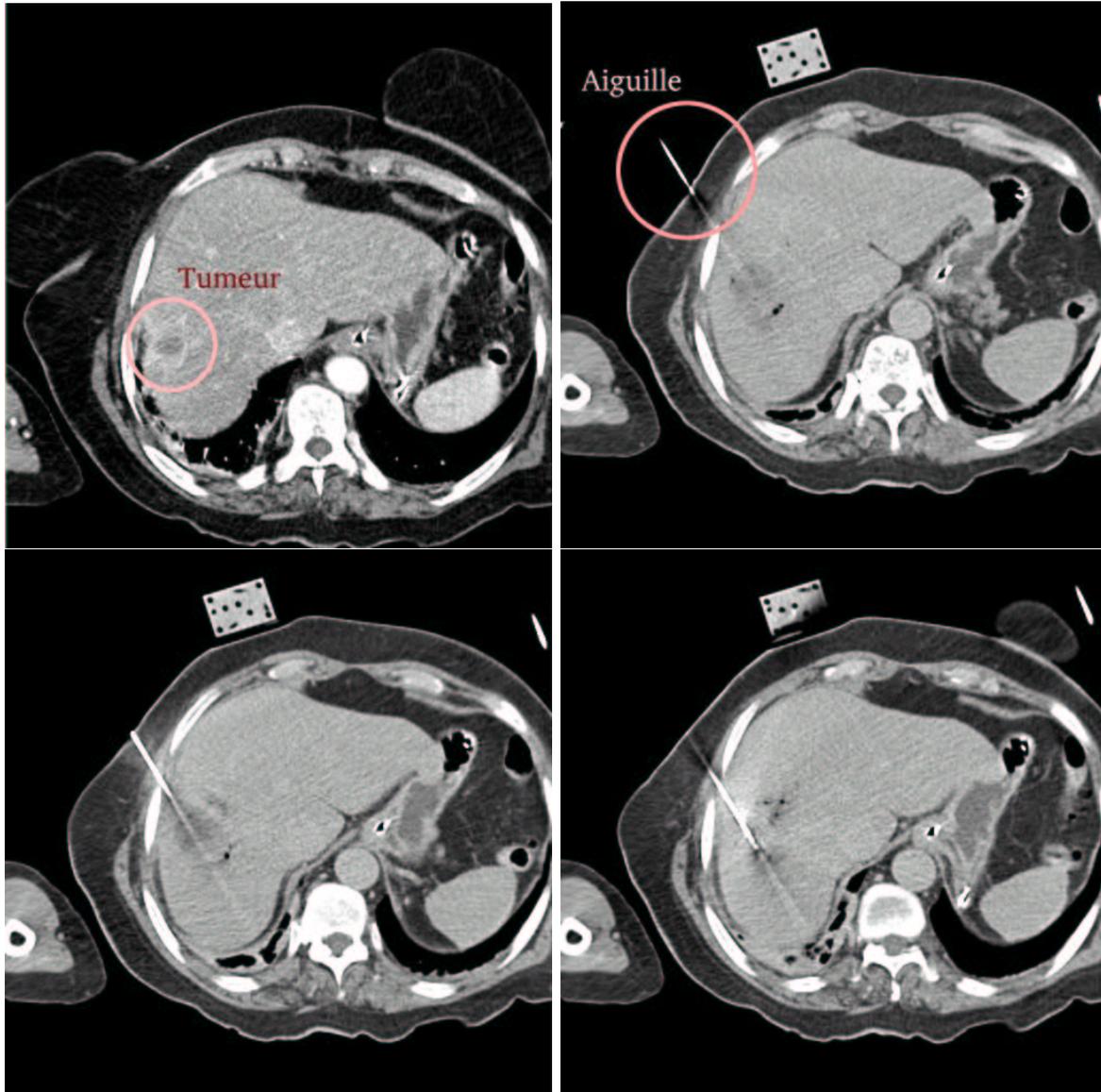


FIG. 2.5 – Exemple de coupes scanner utilisées pour guider l'aiguille au centre de la tumeur (source: IRCAD). En haut à gauche : coupe TDM indiquant la tumeur à traiter. En haut à droite et en bas : extrait d'une acquisition TDM de contrôle (3 coupes successives contenant l'aiguille manipulée par le chirurgien).

chirurgien interrogés). La modalité TDM ne permet pas au praticien de visualiser ses gestes en temps réel, contrairement au guidage par coupes US. Il n'est effectivement pas envisageable d'irradier le praticien pendant tout le temps nécessaire au positionnement correct de l'aiguille dans le foie. Néanmoins, elle permet de savoir où se situe la pointe de l'aiguille plus précisément que sur des coupes US et offre l'avantage de ne pas être sensible à la présence d'os ou de poches d'air dans le corps.

Le problème dominant pour le praticien est donc de positionner précisément la pointe de son aiguille au centre de la cible qu'il s'est fixée. Un système de réalité augmentée pourrait grandement lui faciliter la tâche en indiquant la position relative de l'aiguille par rapport à la cible, l'emplacement de structures fragiles et la qualité de l'orientation de l'aiguille avant de la faire pénétrer sous la peau. Nous pourrions même afficher le planning pré-opératoire mis en place avant l'intervention. Ce guidage permettrait de réduire considérablement le temps de l'intervention, limiterait le nombre de scanner de contrôle, qui sont nombreux tout au long de l'insertion de l'aiguille, et améliorerait la précision du positionnement de l'aiguille dans la tumeur.

Dans le cas de la radiologie interventionnelle guidée par coupes TDM, le patient est opéré sur la table du scanner, ce qui signifie que l'obtention d'une reconstruction juste avant l'opération est possible. De plus, il est anesthésié durant l'intervention et le praticien ne déplace l'aiguille que lorsqu'il est mis en apnée. Les déplacements et déformations que subissent les organes de l'abdomen sont donc limités à l'interaction locale de l'aiguille lors de son insertion. De ce fait, nous supposons que les interactions sur le foie sont suffisamment négligeables pour ne pas devoir mener le travail préliminaire de modélisation et de gestion d'organes, indispensable dans un contexte chirurgical. Nous avons donc choisi la radiologie interventionnelle comme domaine d'application de notre système.

Les conditions en radiologie interventionnelle étant très proches de celles des interventions pour lesquelles un système de réalité augmentée existe déjà (neurochirurgie, chirurgie orthopédique), nous pourrions nous appuyer sur les travaux réalisés pour concevoir notre système. Cependant, pour assurer une précision maximale du système, il faudra tenir compte des effets de la respiration sur la repositionnement du foie à chaque mise en apnée.

2.2 Choix de la technologie de réalité augmentée

D'un point de vue technologique, il existe plusieurs façons de mélanger des informations réelles et virtuelles. Ces technologies diffèrent principalement par l'outil de visualisation¹⁰. Nous avons regroupé les méthodes existantes à l'intérieur de quatre classes générales présentées ci-après et résumées dans la figure 2.6.

Remarquons que le choix de la technologie est imposé dans certains contextes médicaux. Par exemple, en chirurgie laparoscopique, le praticien se sert d'un écran vidéo pour guider ses gestes. Il semble donc naturel de rajouter l'information directement sur son écran. De même, en neurochirurgie, le praticien se guide parfois en regardant dans un microscope [JFS⁺00, JFH⁺02, EKM⁺00]. Il faut donc rajouter l'information dans le champ de vue du praticien, au niveau des optiques du microscope.¹¹

10. Les avantages et les inconvénients de la plupart des systèmes pourront être consultés par le lecteur dans [Azu97] [RF00].

11. Il est toutefois possible d'envisager un système de réalité augmentée ne se servant pas du mode de vision utilisé par le praticien pendant l'opération. Cependant, le praticien serait obligé de détourner le regard du champ opératoire, ce qui ne serait pas ergonomique, voire acceptable cliniquement.

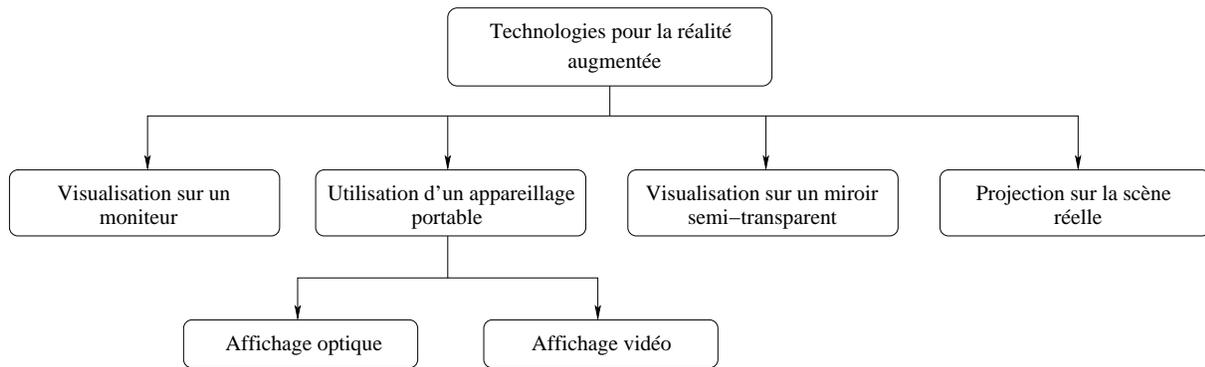


FIG. 2.6 – Les différentes technologies en réalité augmentée.

2.2.1 Via un moniteur

Dans cette configuration (cf. figure 2.7), le système est composé d'une caméra filmant la scène réelle et d'un moniteur restituant l'image vue par la caméra (donc la scène réelle) sur laquelle sont superposés les objets virtuels. Afin de fournir une vue 3D à l'utilisateur, il est possible de rajouter une seconde caméra à côté de la première (cela nécessite alors l'emploi d'une paire de lunettes chromatiques). De manière générale, l'utilisateur peut utiliser, suivant ses besoins, autant de caméras qu'il le désire. Néanmoins, cela nécessite une procédure fastidieuse afin de calibrer toutes les caméras dans un même repère.

Cette méthode est généralement employée lorsqu'il n'est pas nécessaire ou possible d'afficher l'information dans le champ de vision naturel de l'utilisateur. C'est typiquement le cas lorsque l'on souhaite faire de la télé-opération (manipulation à distance assistée par vidéo) en robotique [DBCG⁺01, Dra93, MZD93], et lorsque l'on désire bénéficier d'une assistance vidéo lors du guidage d'un instrument dans le corps d'un patient [MBHN00, GEW⁺96, MVFCM03].

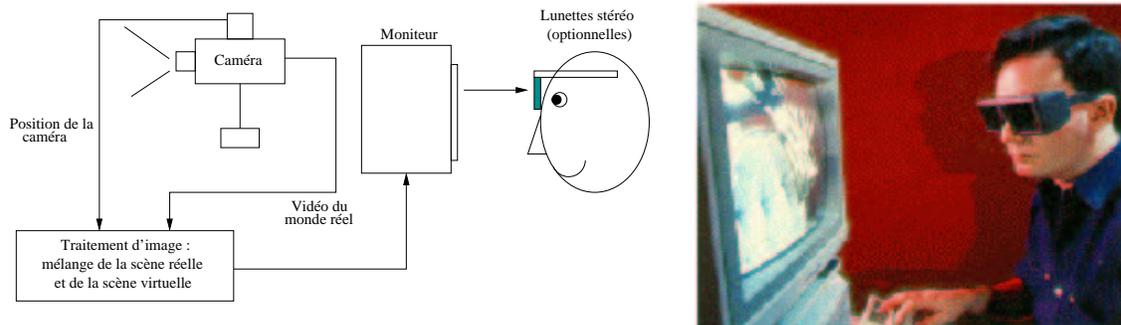


FIG. 2.7 – Principe de la réalité augmentée via un moniteur.

2.2.2 Via un appareillage portable (Head Mounted Display)

Il s'agit de la technologie la plus prisée à l'heure actuelle car la plus naturelle pour l'être humain. Le système de réalité augmentée est composé d'un casque porté par l'utilisateur qui lui permet d'avoir accès à l'information virtuelle dans son champ de vue naturel quelle que

soit sa position dans la scène réelle. Il en existe deux variétés, l'une basée sur des miroirs semi-transparents, l'autre sur des caméras.

– appareillage portable optique :

Dans ce modèle, la fusion du virtuel avec le réel est réalisée par projection sur une paire de lunettes dont les optiques sont deux miroirs semi-transparents (cf. figure 2.8). Dans un tel système, il est nécessaire de suivre le mouvement de la tête de l'utilisateur et de calibrer sa distance inter-pupillaire [TGN02, GTN02] pour afficher les données appropriées à sa localisation et son orientation. Cette technologie est prisée dans le domaine industriel [Nav04, Nav03] puisqu'elle ne réduit pas le champ visuel de l'utilisateur (c'est le cas avec un appareillage portable vidéo). En effet, dans le cadre d'une application industrielle, l'utilisateur a souvent besoin de se déplacer aisément dans son environnement.

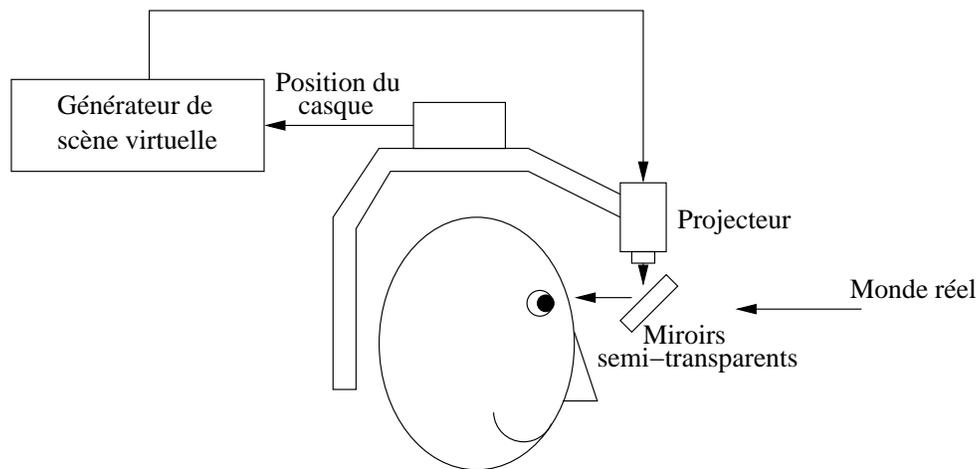


FIG. 2.8 – *Fonctionnement de l'appareillage portable optique.*

– appareillage portable vidéo :

Concernant ce dispositif, les lunettes sont remplacées par deux moniteurs, positionnés devant chaque œil, qui restituent les images filmées par deux caméras vidéo montées sur le casque. Le mélange d'informations est donc réalisé après un traitement informatique des flux vidéo (cf. figure 2.9). Avec ce système, un repérage de la tête de l'utilisateur n'est plus forcément nécessaire si les images vidéo fournies par les caméras contiennent des informations géométriques permettant de déterminer par elle-même leur position [SB02a, BWDPS99, SB02b, GRS⁺02].

2.2.3 Via un miroir semi-transparent

Ce système introduit par Peuchot [PTE95] consiste à afficher sur un miroir semi-transparent situé entre les mains et la tête du chirurgien toute l'information virtuelle (cf. figure 2.10). Dans le cas où cette information est volumique (en 3D), ce système nécessite un repérage en temps réel de la tête afin de faire apparaître les données correspondant au point de vue de l'utilisateur (à l'instar de l'appareillage portable optique). Si l'information à afficher est une image plane (coupes 2D échographique, TDM ou IRM), le repérage de l'utilisateur n'est plus nécessaire [MMN⁺00, SC01].

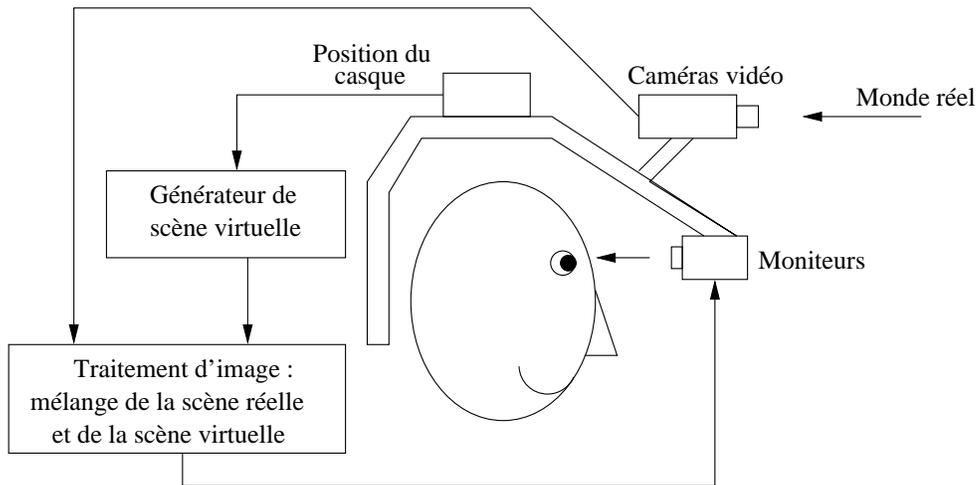


FIG. 2.9 – *Fonctionnement de l'appareillage portable vidéo.*

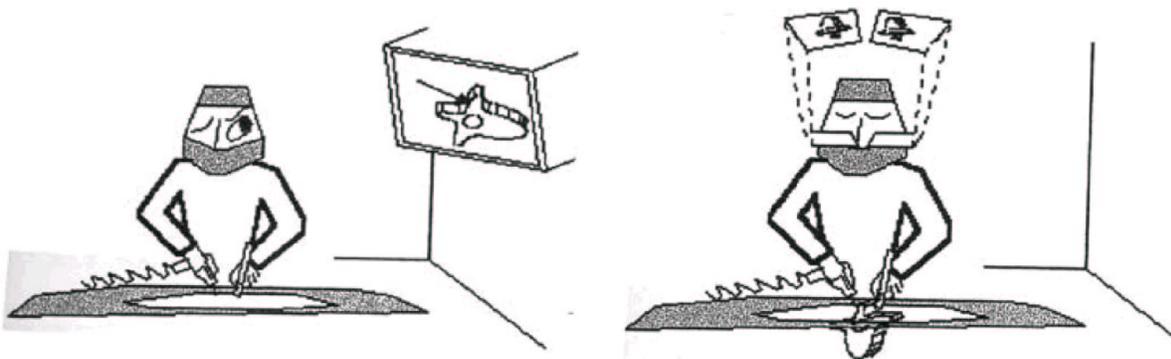


FIG. 2.10 – *Fonctionnement du système de Peuchot [PTE95]. A gauche: le chirurgien doit se détourner du champ opératoire pour tenir compte des informations virtuelles. A droite, les informations sont affichées dans le champ de vision du chirurgien sur des miroirs placés entre ses yeux et le patient.*

2.2.4 Par projection sur la scène réelle

La mise en place de cette approche dans le domaine de la réalité augmentée est très récente et les publications sont encore rares [DBCM01, WH01, GW03, CMAMB04]. L'information virtuelle est projetée sur la peau même du patient, l'utilisateur doit donc éviter de se positionner dans le champ du projecteur. Dans le cas où l'information correspond uniquement à des points de repères sur la peau, un repérage de la tête de l'utilisateur n'est pas nécessaire (cf. figure 2.11).

2.2.5 Réalité augmentée via un moniteur : une technologie simple et précise

A priori, les systèmes les plus naturellement adaptés à l'être humain sont l'appareillage portable optique, le projecteur d'information sur la scène réelle et le système basé sur un miroir semi-transparent puisque le chirurgien a directement dans son champ de vision l'information

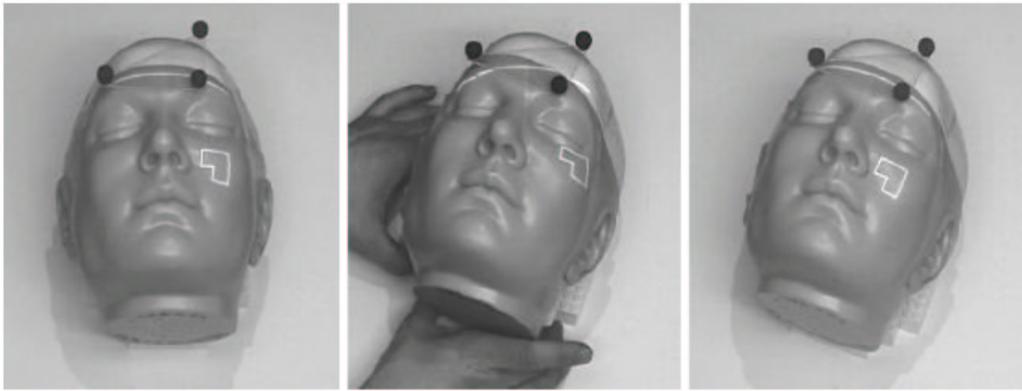


FIG. 2.11 – Exemple de système de réalité augmentée utilisant un projecteur [WH01]. A gauche : mannequin sur la joue duquel est projeté le planning pré-opératoire. Au milieu : juste après le déplacement du mannequin, les informations apparaissent à une position erronée. A droite : après un rafraîchissement des coordonnées des marqueurs, l'information est à nouveau projetée correctement.

virtuelle. Néanmoins, l'appareillage portable optique nécessite un calibrage de la distance interpupillaire et surtout une relation fixe entre le système rigide des deux globes oculaires et celui des deux projecteurs. Si jamais il y a un déplacement (glissement de casque), il y aura alors un biais systématique dont les conséquences seront un décalage de l'information virtuelle induisant une erreur de guidage immédiate. Par ailleurs, une gestion correcte de la transparence et de l'opacité nécessite un environnement faiblement éclairé, ce qui rend cette technologie difficilement exploitable. Les systèmes par *projection* et *miroir semi-transparent* sont soumis aux mêmes problèmes que l'appareillage portable optique. Etant donné qu'il faut repérer la position de la tête (des yeux en fait) afin d'afficher la bonne image 3D aux deux yeux, un calibrage des pupilles par rapport à un objet rigidement fixé au crâne ou sur la tête est aussi nécessaire. A nouveau, le système est soumis à un biais systématique si l'objet fixé sur la tête glisse. Dans le cas du système par projection, si le modèle de la surface sur laquelle est effectuée la projection n'est pas très proche de la réalité, il y aura des erreurs de visualisation supplémentaires. En conséquence, pour des questions de précision et de sécurité, un système de réalité augmentée basé sur un appareillage optique, un miroir semi-transparent ou une projection sur la surface du patient ne peut être envisagé pour notre application.

La différence entre un système basé sur des caméras fixes ou un appareillage portable vidéo provient uniquement de la nécessité de repérer l'appareillage portable en temps réel, ce qui introduit une erreur de précision supplémentaire. D'un point de vue ergonomique, l'appareillage portable vidéo est préférable puisque le champ de vue correspond approximativement au champ de vue naturel de l'utilisateur (effet de vision tunnel). Bien qu'un système basé sur des caméras fixes n'offre pas une vision augmentée dans l'axe de vue de l'utilisateur, nous pouvons rajouter une vue centrée image (cf. figure 2.12) qui permettrait de faire de la laparo-navigation, ce qui n'est pas faisable avec un appareillage portable vidéo. Cela permettrait d'offrir des points de vue difficilement envisageables autrement. En définitive, un système de réalité augmentée basé sur la présence de plusieurs caméras judicieusement placées autour du patient est plus adapté à notre application.

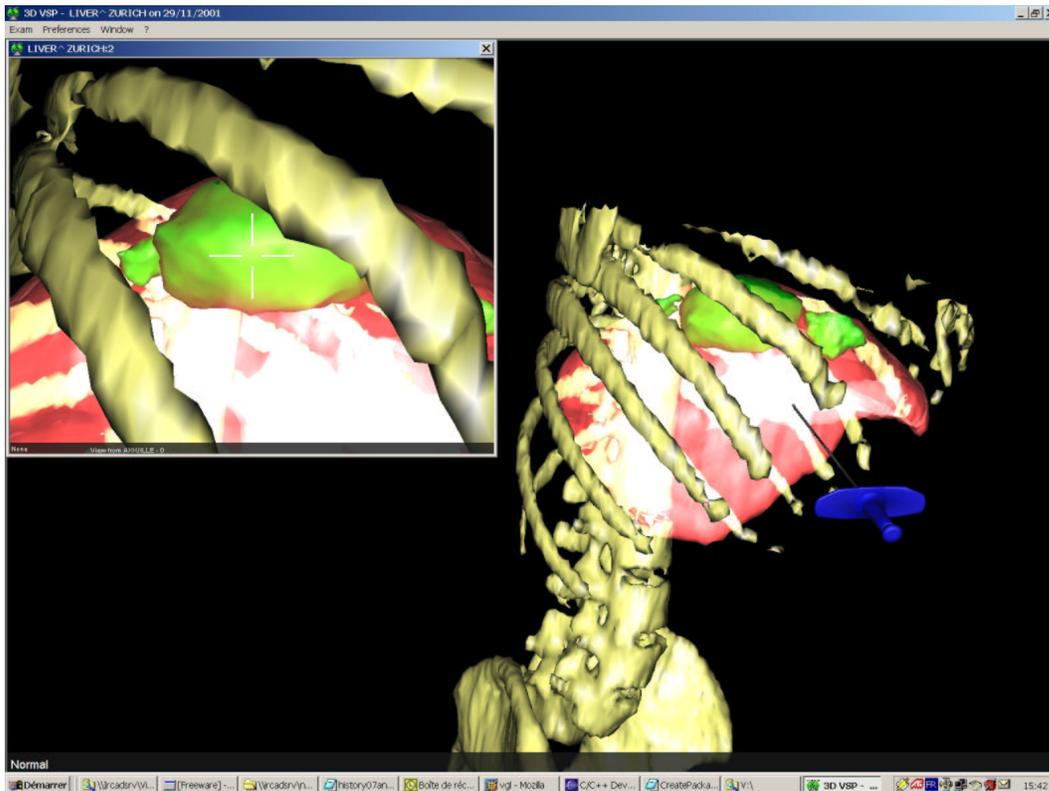


FIG. 2.12 – Exemple de vue dont disposerait le praticien sur un moniteur pour mener son aiguille au centre de la tumeur (en vert). La fenêtre en haut à gauche correspond à la vue d'une caméra virtuelle située au bout de l'aiguille et orientée suivant son axe.

2.3 Objectifs

Dans ce chapitre, nous avons décidé quel type d'opérations du foie nous allons traiter et quelle technologie nous allons employer. Nous pouvons donc déterminer plus précisément nos objectifs. En concertation avec les chirurgiens travaillant à l'IRCAD, nous avons défini le cahier des charge que notre système devait respecter afin de leur apporter une aide significative. Notre système doit être :

- **précis** : comme nous l'avons mentionné précédemment, pour que ce système apporte réellement une aide aux praticiens pendant la phase de l'insertion de l'aiguille, il doit permettre d'atteindre une cible avec une erreur maximale de 5 mm. Cela signifie que la somme des erreurs dues au recalage du modèle et du repérage de l'aiguille doit être inférieure à 5 mm.
- **fiable** : nous devons assurer une précision minimale du système. Il serait dangereux d'avoir un système précis en moyenne à 5 mm et qui soit à 10 mm dans 10% des cas. Cela signifierait que 10% des patients seraient susceptibles d'être les victimes d'un geste mal guidé. En général, la fiabilité découle d'une évaluation de la précision pendant laquelle les conditions varient à l'intérieur d'un intervalle réaliste. Nous verrons qu'il est possible d'obtenir une évaluation théorique fiable *en fonction* des paramètres.

- **rapide** : dans sa globalité, notre système devra être capable de recueillir et traiter l'information nécessaire au repositionnement du modèle en moins de 10 minutes. Durant l'intervention, le patient ne pouvant rester plus de quelques secondes en apnée, les procédures d'extraction et de recalage dans les images vidéo doivent pouvoir être réalisées dans des conditions proche de la seconde.

D'un point de vue pratique, le respect de ce cahier des charges dépend essentiellement du type et de la qualité du recalage que nous réaliserons avec notre système. En effet, la rapidité variera en fonction de la quantité d'informations utilisées pour le calculer, et la précision/fiabilité sera liée à la pertinence de ces informations. Le choix du recalage va donc s'avérer primordial pour atteindre les objectifs fixés. Cette remarque nous mène au chapitre suivant dans lequel nous réaliserons une analyse détaillée des conditions opératoires de notre application afin de déterminer quelle stratégie de recalage est susceptible de satisfaire au mieux nos contraintes.

Chapitre 3

La stratégie de recalage

Nous avons vu, dans le chapitre précédent, que le recalage est la pierre angulaire de notre système de réalité augmentée basé sur la superposition d'un objet virtuel sur son homologue réel dans l'image vidéo. Ce problème de recalage, dans le domaine de la réalité augmentée, est déjà largement traité et il existe de nombreuses méthodes permettant de fournir une solution. Néanmoins, notre système doit être précis, fiable et rapide pour être utilisable en conditions cliniques. L'objectif de ce chapitre est de déterminer quelle est la méthode de recalage la plus appropriée pour satisfaire nos contraintes (si elle existe!). Notre analyse sera principalement articulée autour de deux constats majeurs : le foie bouge en fonction de la respiration et la peau de l'abdomen présente peu de singularités (réduisant ainsi la précision potentielle de certaines méthodes de recalage).

Nous verrons tout d'abord que, dans les conditions de notre intervention (la respiration du patient est sous contrôle), un recalage rigide est compatible avec les objectifs de précision fixés (erreur de repositionnement du foie de l'ordre de 1 mm). Ensuite, nous détaillerons les méthodes de recalages rigides existantes. Pour des raisons de précision, nous mettrons de côté les méthodes surfaciques, iconiques et ponctuelles 3D/3D, au profit des ponctuelles 3D/2D.

3.1 Recalage rigide ou non-rigide?

Le foie, organe le plus volumineux du corps humain, est situé sous le diaphragme qui est en contact direct avec les poumons (cf. figure 3.1). Il est donc raisonnable de penser que sa position et sa forme (notamment son dôme) puisse être modifiée pendant la respiration du patient. Le choix du recalage (rigide ou non-rigide) dépendant de l'influence de la respiration sur la position du foie, une étude quantitative est nécessaire.

Certaines études se sont intéressées aux modes de déformations du foie en fonction de la respiration [BKG⁺01, LLS02, BGK⁺04] et suggèrent l'existence d'un lien liant les déformations du foie à celle de la peau. Comme nous cherchons à repositionner correctement le foie virtuel dans l'image vidéo à partir d'une information provenant de la peau, il semblerait logique d'opter par défaut pour un recalage non-rigide. Or, la conception d'un modèle déformable du foie d'un patient et l'étude de ses modes en fonction de la respiration nécessite de nombreux scanners de l'abdomen durant plusieurs cycles respiratoires. Cela représente une dose d'irradiation bien trop importante pour être envisagée au cours du protocole de radiologie interventionnelle classique.

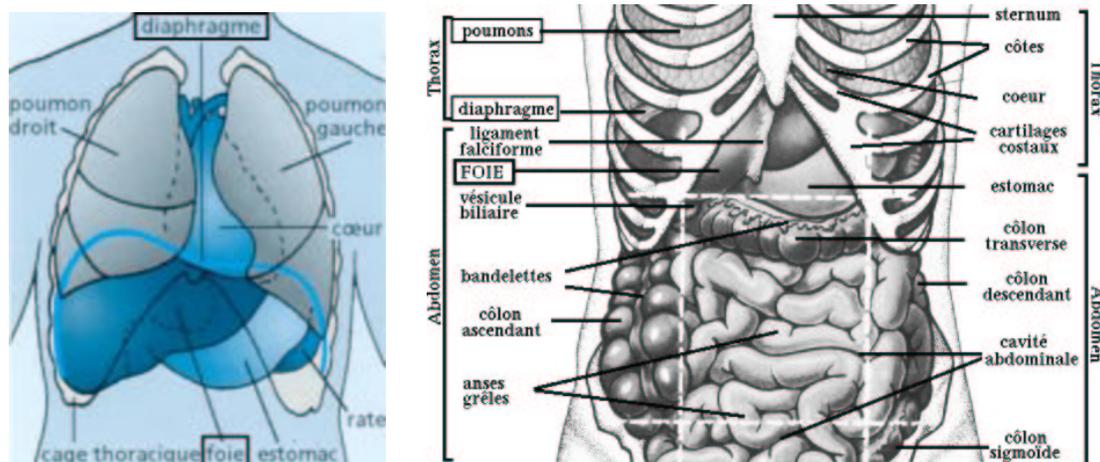


FIG. 3.1 – Anatomie de l'abdomen : le foie est au contact des poumons par l'intermédiaire du diaphragme (source : <http://www.educnet.education.fr> ; Gillot, 1996).

3.1.1 Influence de la respiration libre sur la position du foie

Nous entendons par respiration “libre” toute respiration qui n’est pas influencée par la volonté du patient, c’est à dire que le patient respire normalement sans qu’il lui soit imposé une quelconque contrainte ou un quelconque rythme. Les études menées sur les mouvements du foie avec ce type de respiration présentent des conditions expérimentales assez variées et aboutissent à des résultats plutôt cohérents entre eux. Le tableau 3.1 rassemble les informations contenues dans les articles consacrés aux mouvements du foie lors de la respiration (le lecteur intéressé pourra consulter l'article de Clifford [CBL02] résumant la majorité des résultats concernant les déplacements du foie en fonction de la respiration).

Référence	Nombre de patients	Déplacement moyen (mm) :		
		cranio-caudal	antérieur-postérieur	latéral
Harauz (1979) [HB79]	51	14		
Suramo (1984) [SPM84]	28	25		
Korin (1992) [KER ⁺ 92]	15	13	2.5	
Davies (1994) [DHH ⁺ 94]	9	10		
Balter (1996) [BTHL ⁺ 96]	6	17		
Herline (1999) [HSD99]	1	10.8		
Schimizu (1999) [SSX ⁺ 99]	1	21	8	9
Schimizu (2000) [SSA ⁺ 00]	6	10.6	4.6	5.2
Rohlfing (2001) ^a [RMOZ01]	1	12–26	1–12	1–3
Moyenne ^b	1	14.8–16.4	4.0–6.7	5.0–5.7

TAB. 3.1 – Synthèse quantitative des études menées sur le déplacement du foie en fonction de la respiration. Le mouvement du foie se fait principalement suivant l'axe cranio-caudal avec une amplitude moyenne comprise entre 14 et 16 mm.

^a Rohlfing n'indique pas de valeur moyenne dans son étude mais seulement les valeurs minimales et maximales des déplacements observées

^b La première (resp. second) valeur correspond à la moyenne calculée avec la limite basse (resp. haute) de l'étude de Rohlfing.

Ces études indiquent des amplitudes de déplacement moyen allant de 10 mm jusqu'à 25 mm. Cette différence apparemment conséquente peut s'expliquer par le fait que les mesures n'ont pas toutes été faites avec les mêmes moyens d'acquisition (IRM, TDM, fluoroscopie, caméra stéréoscopique...), dans les mêmes conditions (étude de Herline [HSD99] menée à corps ouvert) et avec les mêmes points de référence pour prendre les mesures. De plus, toutes les études ne contiennent pas le même nombre de patients, il faut donc interpréter celles ayant été réalisées sur un unique patient avec circonspection. En effet, un patient pesant 100 kilogrammes et mesurant 1m90 fournira des mesures probablement plus élevées qu'un patient de 60 kilogrammes. Les écarts entre les observations des différentes études ne sont donc pas vraiment surprenants.

En revanche, nous pouvons remarquer que l'ordre de grandeur de l'amplitude des déplacements du foie est à peu près cohérent et dépasse à chaque fois 1 cm en moyenne suivant l'axe cranio-caudal. Souhaitant obtenir une erreur de recalage inférieure à 5 mm, l'emploi d'un recalage rigide ne pourrait être justifié que si la peau et le foie se déplaçaient suivant le même axe et avec une amplitude équivalente. A ce jour, il n'existe aucune preuve que ce soit effectivement le cas (l'observation suggère même plutôt que la peau ne subit pas de translation mais plutôt une dilatation vers l'extérieur). De plus, le mouvement cranio-caudal n'est pas censé être uniforme sur le foie et induit plutôt des déformations. Utiliser un recalage rigide dans ces conditions n'est donc pas envisageable avec une acquisition en temps différé puisque les déplacements observés du foie sont trop importants par rapport à notre application.

3.1.2 Cas de la respiration contrôlée par le patient

Nous venons de voir que la respiration peut générer des déplacements très importants au regard de la précision souhaitée, rendant inadapté un recalage rigide. Afin de limiter les mouvements du foie, les praticiens utilisent communément en radiothérapie une technique appelée "respiratory gating". Elle consiste à demander au patient de se repositionner régulièrement à un même instant de son cycle respiratoire (typiquement en fin d'expiration ou d'inspiration profonde). Cette technique permet alors d'obtenir répétitivement un volume à peu près constant dans les poumons et donc une position quasi-identique du foie.

Bien que les études qualitatives ne soient pas forcément d'accord sur le repositionnement des organes [SSA⁺00, THBM97, KH96, OOA⁺89], les études quantitatives sont toutes cohérentes et appuient la certitude que l'erreur de repositionnement est proche en moyenne de 1 mm. Balter [BLM⁺98] a étudié sur 15 patients le déplacement moyen du foie pendant de multiples cycles respiratoires (typiquement 10), et a évalué que l'erreur de repositionnement était de 0.9 mm au moment de l'expiration profonde. Suramo [SPM84], contrôlant le volume inspiré à l'aide d'une barre métallique suspendue horizontalement au dessus du ventre, a évalué l'erreur de repositionnement à 2 mm en moyenne sur 28 patients. Wong [WSJ⁺99] se sert d'un appareil spécifique (proche d'un spiromètre) qui mesure le volume d'air inspiré et expiré par le patient afin d'avoir répétitivement les mêmes conditions volumétriques d'air dans les poumons. Son étude sur 2 patients montre que l'erreur de repositionnement obtenue est de 0.8 mm. Les erreurs de ces différentes études sont cohérentes et du même ordre de grandeur, bien que Suramo ait évalué l'erreur à 2 mm plutôt que 1 mm. Néanmoins, cette différence peut facilement s'expliquer par le manque de précision du matériel disponible à l'époque (en 1984) et par la technologie utilisée pour réaliser les mesures (il s'est servi de coupe US plutôt que TDM).

3.1.3 Cas d'un patient intubé

Les conditions cliniques de notre intervention sont particulières puisque le patient est intubé et sa respiration est sous contrôle volumétrique (cf. section 2.1.3). Nous pouvons donc itérer un volume identique dans les poumons à chaque cycle respiratoire en étant, cette fois-ci, indépendant du patient.

Comme il n'existe aucune expérience quantifiant l'erreur de repositionnement des organes de l'abdomen lorsque le patient est anesthésié et intubé, nous avons mené notre propre étude sur deux cas cliniques. Notre protocole expérimental consiste à obtenir deux acquisitions TDM de l'abdomen du patient en des instants différents (environ 10 minutes d'intervalle) mais avec des conditions volumétriques identiques dans les poumons (la respiration du patient est sous contrôle). Un chirurgien segmente ensuite les deux acquisitions et nous mesurons l'écart entre les deux surfaces reconstruites issues de leur segmentation. Nos résultats confirment les observations des études précédentes puisque nous avons quantifié sur 2 patients des écarts de surface à 1 mm en moyenne.

Les deux images en bas de la figure 3.3 montrent, selon deux plans de coupe en Z différents, la soustraction de deux images TDM du foie d'un patient obtenues à deux instants différents. Les voxels noir et blanc correspondent à une soustraction d'amplitude 40 HU (l'échelle des deux images originales va de -1024 à +1024). Nous constatons que le liseré de couleurs aléatoires à la frontière du foie, représentatif de son mouvement entre les deux acquisitions, a une épaisseur en dessous de 2 voxels (i.e. $2 \cdot 0.68 = 1.36$ mm). Le foie n'a donc presque pas bougé d'une acquisition à la suivante. La figure 3.2 montre la position des deux foies reconstruits dans le même repère (aucun recalage n'a été effectué). En supposant que la délimitation manuelle est parfaite, l'entrelacement des surfaces confirme visuellement l'hypothèse d'un repositionnement millimétrique.

Par ailleurs, une étude de 3 segmentations d'un même foie, réalisées par 3 chirurgiens différents, a montré une variabilité inter-utilisateur moyenne d'environ 1.4 mm (cf. tableau 3.2). Ce

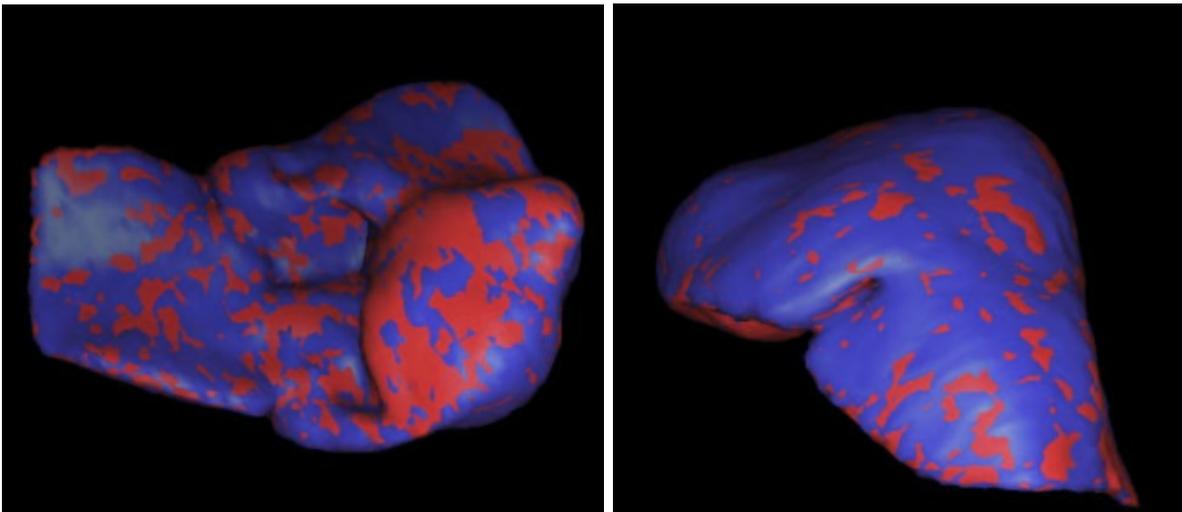


FIG. 3.2 – Visualisation simultanée dans le même repère de 2 reconstructions d'un même foie à des moments différents mais au même instant du cycle respiratoire. L'aspect persillé des couleurs indique visuellement que le foie n'a presque pas bougé.

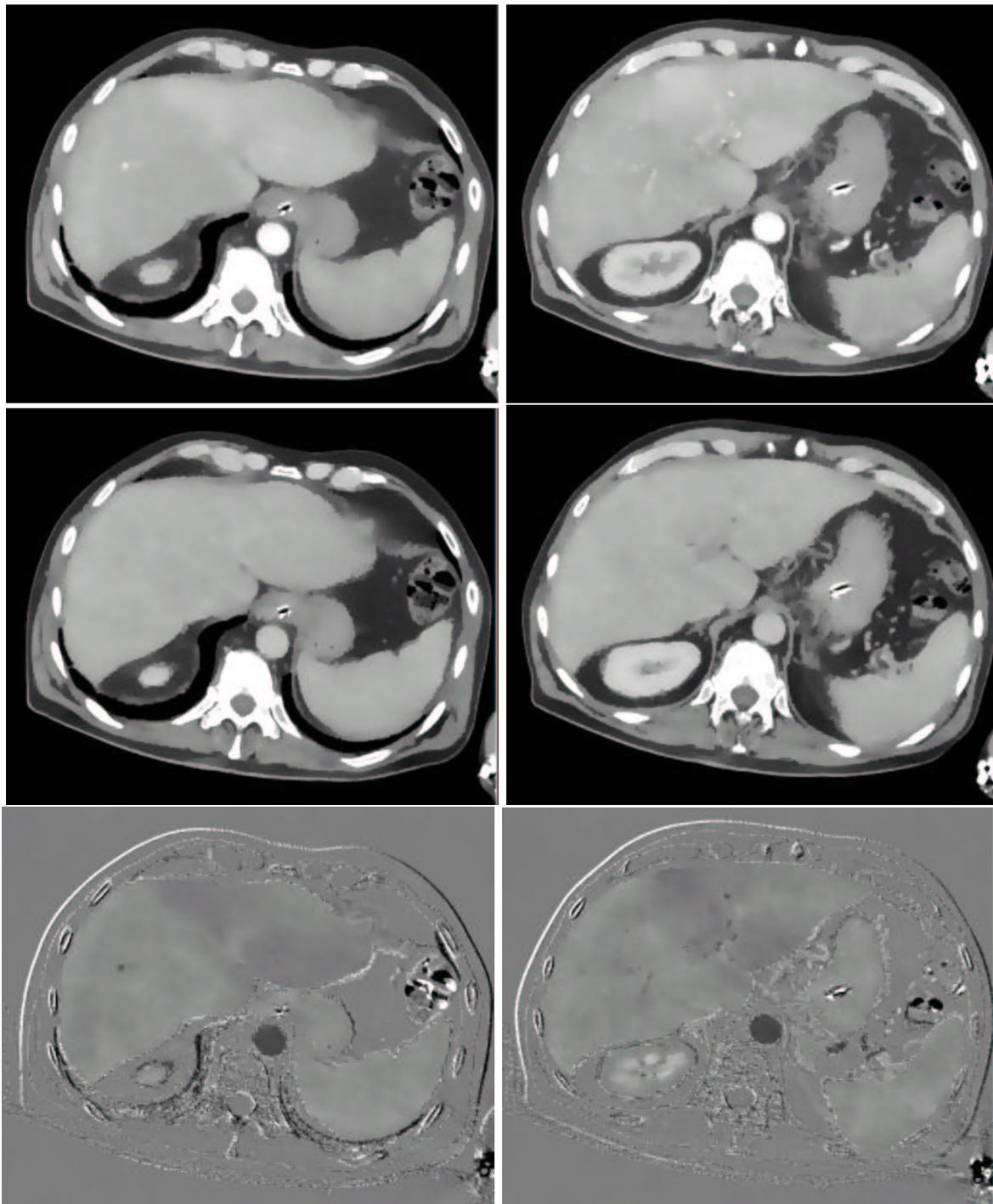


FIG. 3.3 – Vérification visuelle de la qualité du repositionnement du foie à partir de deux images TDM acquises à deux instants différents. Un filtre anisotrope conservant les contours a été préalablement appliqué aux deux images afin de réduire les effets du bruit d'acquisition. Les deux colonnes sont relatives à deux plans de coupe en Z différents. Les deux coupes en haut (resp. du milieu) sont extraites de la première (resp. de la seconde) acquisition TDM réalisée. Les deux coupes du bas correspondent au résultat de la soustraction des deux acquisitions TDM (dont l'échelle de niveaux de gris va de -1024 HU à $+1024$ HU). Dans les images du bas, les voxels blancs et noirs correspondent à une amplitude de ± 40 HU. Le fin liseré de couleurs aléatoires sur la frontière séparant le foie des autres organes est représentatif de son mouvement entre les acquisitions. En l'occurrence, son épaisseur ne dépasse pas 2 voxels (i.e. $2 \times 0.68 = 1.36$ mm).

résultat est cohérent avec celui obtenu par Anxionnat [ABK⁺03] qui estime la variabilité inter- et intra-utilisateur pour la segmentation d’image angiographique. L’erreur de repositionnement observée est donc aussi due à l’erreur de segmentation du chirurgien.

Écart observé (mm)	chirurgien 1	chirurgien 2	chirurgien 3
chirurgien 1	0	1.52	1.28
chirurgien 2	1.52	0	1.50
chirurgien 3	1.28	1.50	0

TAB. 3.2 – *Etude de la différence entre trois segmentations d’un même foie, réalisées par trois chirurgiens différents. L’erreur mesurée correspond à la moyenne des distances entre les surfaces (au sens du plus proche voisin) obtenues à partir des trois segmentations. La variabilité moyenne de la segmentation, d’un chirurgien à l’autre, est de l’ordre de 1.4 mm.*

3.1.4 Conclusion

Dans cette section, nous avons mis en évidence que nos contraintes de précision sont trop strictes pour nous permettre d’utiliser un recalage rigide si le patient respire librement. Néanmoins, les conditions en radiologie interventionnelle permettent de contrôler la respiration et de se positionner répétitivement au même point du cycle respiratoire. Il est alors possible d’obtenir une erreur de repositionnement faible ($\simeq 1$ mm) en regard des contraintes de précision demandées (5 mm). Dans de telles conditions, il est tout à fait raisonnable d’aborder le problème de recalage en optant pour une solution rigide.

3.2 Recalage iconique, surfacique ou ponctuel?

Nous devons maintenant faire un choix parmi les différentes méthodes de recalages rigides possibles. Dans cette section, nous détaillons les principales méthodes possibles, puis nous montrons pourquoi, pour des raisons de précision et de temps de calcul, un recalage ponctuel 3D/2D est préférable aux méthodes surfaciques, iconique et ponctuel 3D/3D.

3.2.1 Recalages iconiques

Maximisation de l’information mutuelle

Ce type de recalage, introduit par Viola [VW97], ne nécessite qu’une image vidéo et la reconstruction tri-dimensionnelle de la scène virtuelle que l’on souhaite recalcr dans l’image (dans notre cas la peau du patient). L’idée repose sur l’existence d’une relation statistique entre l’intensité des niveaux de gris dans l’image vidéo et les vecteurs normaux de la surface du modèle reconstruit. Plus précisément, on suppose qu’il existe une fonction F telle que:

$$v(P \star T(x)) = F(u(x), q) + \eta$$

où :

- v est la fonction qui associe à une coordonnée pixel de l’image son intensité
- u est la fonction qui associe à un point 3D du modèle son vecteur normal
- q contient les paramètres externes de la scène (généralement les conditions d’illumination)

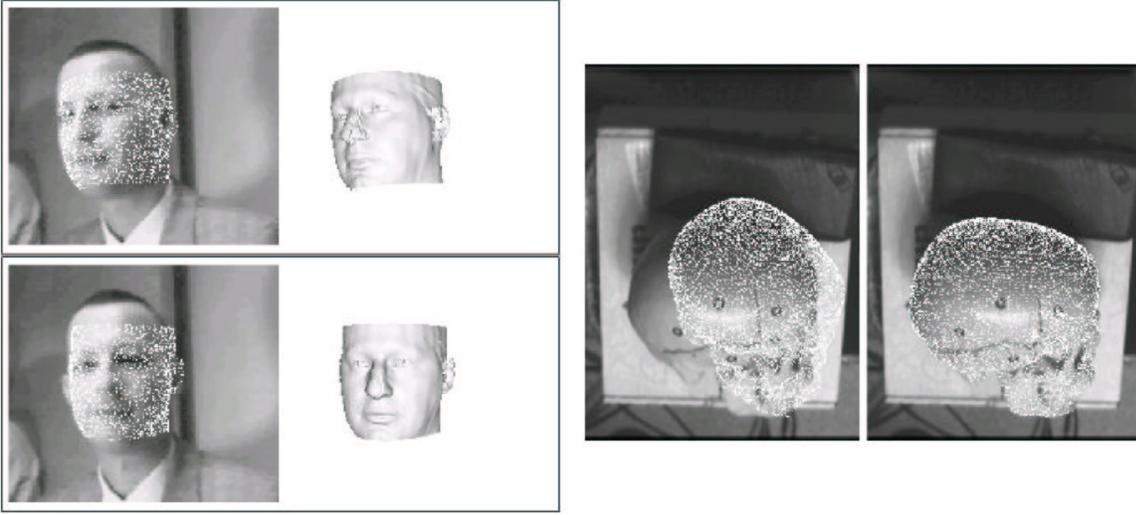


FIG. 3.4 – Illustration de recalages obtenus par optimisation de l'information mutuelle entre l'intensité des niveaux de gris dans l'image vidéo et les vecteurs normaux de la surface du modèle à recaler [VW97]. A gauche : recalage du modèle 3D d'un visage dans une image vidéo. L'orientation du modèle 3D, après recalage, correspond à l'orientation du visage de la personne dans les images vidéo. A droite : recalage du modèle 3D d'un crâne dans une image vidéo du crâne dont le modèle est issu. L'image de gauche (resp. de droite) correspond à la position du modèle 3D avant (resp. après) recalage.

- η est une variable aléatoire modélisant le bruit lors de l'acquisition des images
- T est la transformation recherchée
- P est l'opérateur de projection des points 3D en points 2D
- F est en fait une fonction modélisant les propriétés de réflectivité de la surface du modèle.

Dans ce cas, il existe de l'information mutuelle entre les fonctions $u()$ et $v()$. L'estimation de T est donc réalisée en maximisant l'information mutuelle entre $u()$ et $v()$:

$$\hat{T} = \arg \max_T \left(MI(u(x), v(P \star T(x))) \right) \quad (3.1)$$

$$= \arg \max_T \left(h(u(x)) + h(v(P \star T(x))) - h(u(x), v(P \star T(x))) \right) \quad (3.2)$$

où MI est l'information mutuelle entre le modèle et l'image et h est l'entropie d'une variable aléatoire. Trois exemples de recalage obtenus par optimisation de l'information mutuelle sont montrés sur la figure 3.4.

Maximisation de la photo-cohérence

Cette méthode de recalage récemment introduite par Clarkson [CRHH00] est basée sur l'observation que, pour une surface lambertienne, la réflectance d'un point donné de cette surface n'est dépendante que des sources lumineuses présentes et non pas de la position de la caméra visionnant la scène. Ainsi, étant donné S caméras (au moins 2), un même point 3D d'une

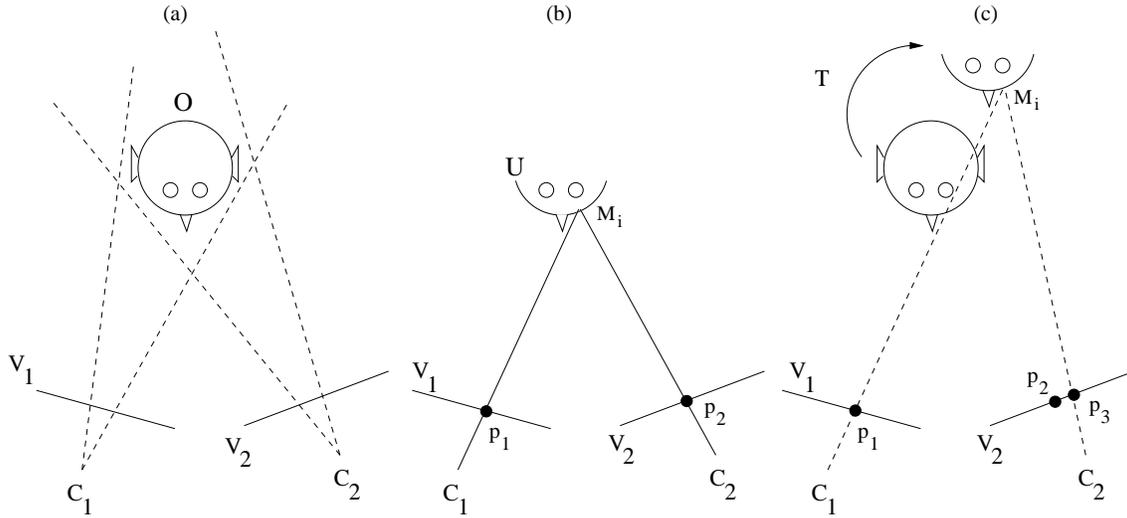


FIG. 3.5 – Illustration de l'application de la photo-cohérence au recalage : 2 caméras de centre optique C_1 et C_2 visionnent un objet réel O (cas (a)). Soit U la surface à recalcer dans l'image vidéo sur l'objet réel O . Chaque point M_i de la surface U est projeté sur les images vidéo V_1 et V_2 en p_1 et p_2 . Si M_i est correctement recalé dans l'image vidéo, alors les intensités lumineuses en p_1 et p_2 devraient être photo-cohérentes (cas (b)). Si le modèle U n'est pas bien recalé par la transformation T (cas (c)), alors les projections p_1 et p_3 de M_i devraient être moins photo-cohérentes que dans le cas (b).

surface doit avoir la même intensité lumineuse dans chacune des images vidéo. Cette propriété appelée photo-cohérence peut-être utilisée, après introduction d'une mesure de similarités pour recalcer un modèle d'une surface sur son homologue dans des images vidéo (cf. figure 3.5)

Les S caméras (de fonction de projection P_l) étant calibrées dans le même repère, on va chercher à rendre minimum la mesure de similarité définie par ¹² :

$$C_{Sim} = 1/N \sum_{i=1}^N e_i^2 \quad \text{avec} \quad e_i^2 = \sum_{l=1}^S \left(I(P_l(T \star M_i)) - \bar{I}_i \right)^2$$

où M_i correspond aux coordonnées 3D d'un des N points de la surface modèle U à recalcer, $I(u)$ est la fonction qui associe à un point de l'image vidéo son intensité lumineuse et \bar{I}_i est la moyenne des intensités lumineuses des projections du point M_i dans les S images vidéo.

3.2.2 Recalage géométrique surfacique

Cette méthode nécessite une reconstruction surfacique dans le repère de référence des caméras de l'objet à recalcer. Cette reconstruction peut être obtenue par stéréoscopie, si 2 caméras ou plus sont utilisées [Dev97, BCMDV02], ou par un dispositif calibré dans le repère des caméras (par exemple un télémètre laser). Disposant de 2 surfaces S_1 et S_2 des mêmes objets dans deux repères différents, le recalage 3D/3D est généralement assuré par la méthode ICP [BM92, Zha94] (cf. figure 3.6). L'estimation de la transformation T est réalisé par l'optimisation du critère :

¹². La mesure de similarité présentée ici correspond à la première utilisée historiquement. Il a été prouvé ensuite que d'autres mesures apportaient de meilleurs résultats en terme de robustesse [CRHH00]

$$C(T) = \sum_{M_j \in S_1} \| T \star M_j - \text{ClosestPoint}(T \star M_j) \|^2$$

où *ClosestPoint* est une fonction qui associe à un point de l'espace son point le plus proche sur S_2 . Cette minimisation s'effectue en deux étapes successives alternées. La première consiste à apparier les points selon une distance donnée, et la seconde est l'évaluation aux moindres carrés du déplacement T qui superpose au mieux les paires de points apparées. Ce critère est le premier historiquement. Nous redirigeons le lecteur vers [Fel95, Gra03] pour une description exhaustive des critères existants actuellement

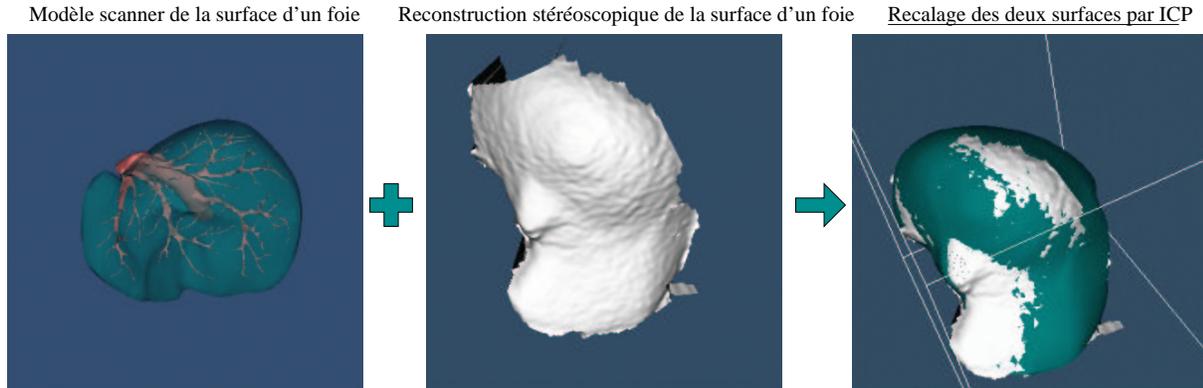


FIG. 3.6 – *Illustration d'un recalage surfacique. Les données scanners proviennent du VisibleManTM et la reconstruction stéréoscopique a été obtenue avec une reproduction en plastique du foie du VisibleManTM. A gauche : la reconstruction surfacique du foie obtenue à partir d'une segmentation d'images TDM. Au milieu : reconstruction surfacique obtenue par stéréoscopie. A droite : résultat du recalage des 2 surfaces obtenu avec un algorithme de type ICP.*

3.2.3 Recalage géométrique ponctuel

Ce recalage nécessite la présence d'amers (points de référence) visibles dans les images vidéo et TDM. La peau étant très uniforme, il n'est pas possible de définir un nombre suffisant de points particuliers que l'on retrouvera chez tous les patients. Dans ce cas de figure, il est nécessaire de se servir de marqueurs que l'on place sur la peau (cf. figure 3.7). Nous présentons ci-après les deux principales approches pour effectuer le recalage.

Approche 3D-3D

Supposons que nous ayons deux caméras (ou plus) conjointement calibrées visionnant N marqueurs et que nous connaissions l'appariement entre les marqueurs extraits des images TDM et ceux des images vidéo. Il est alors possible de les reconstruire en 3D dans le repère des caméras, et donc d'estimer la transformation T en effectuant un recalage 3D/3D ponctuel. Le critère de minimisation classique 3D/3D est :

$$C(T) = \sum_{j=1}^N \| T \star M_j - M_j' \|^2$$

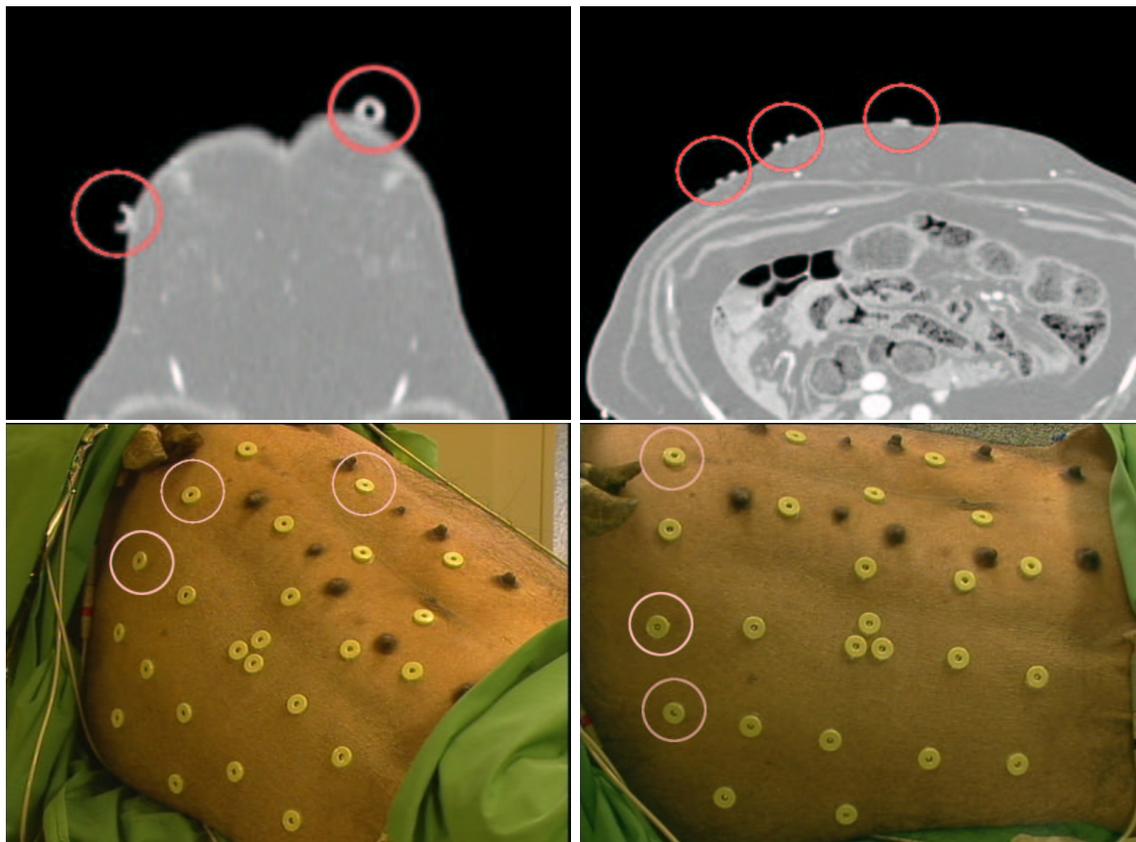


FIG. 3.7 – Les marqueurs radio-opaques sont mis en évidence dans les cercles rouges. En haut : dans les images scanner. En bas : dans les images vidéo.

Ce problème est déjà bien connu et il existe de nombreuses méthodes permettant le calcul direct d'une solution [AHB87, HBE86, FH83].

Approche 3D-2D

Nous présentons ici les deux critères classiques les plus intuitifs dont l'optimisation fournit une estimation de la transformation recherchée. Le premier consiste à recaler le modèle dans le plan projectif des caméras et le second est un recalage en 3D entre un ensemble de points et un ensemble de droites.

Minimisation dans l'espace image (ISPPC)

Dans ce cas de figure, l'emploi d'une seule caméra permet de calculer la transformation T , mais il n'est alors généralement pas possible d'obtenir une précision suffisante, surtout en présence de bruit (cf. section 5.1). Comme précédemment, nous supposons que les appariements vidéo/scanner sont déjà connus. Nous recalons la projection des marqueurs M_j du modèle scanner sur les marqueurs m_j repérés dans la ou les images vidéo (cf. figure 3.8). Le critère usuellement utilisé pour le recalage projectif dans le cas d'une caméra est de la forme :

$$C(T) = \sum_{j=1}^N \| P(T \star M_j) - m_j \|^2$$

où P est la fonction de projection de la caméra. Ce critère sera désormais appelé ISPPC pour *Image Space Projective Points Criterion*.

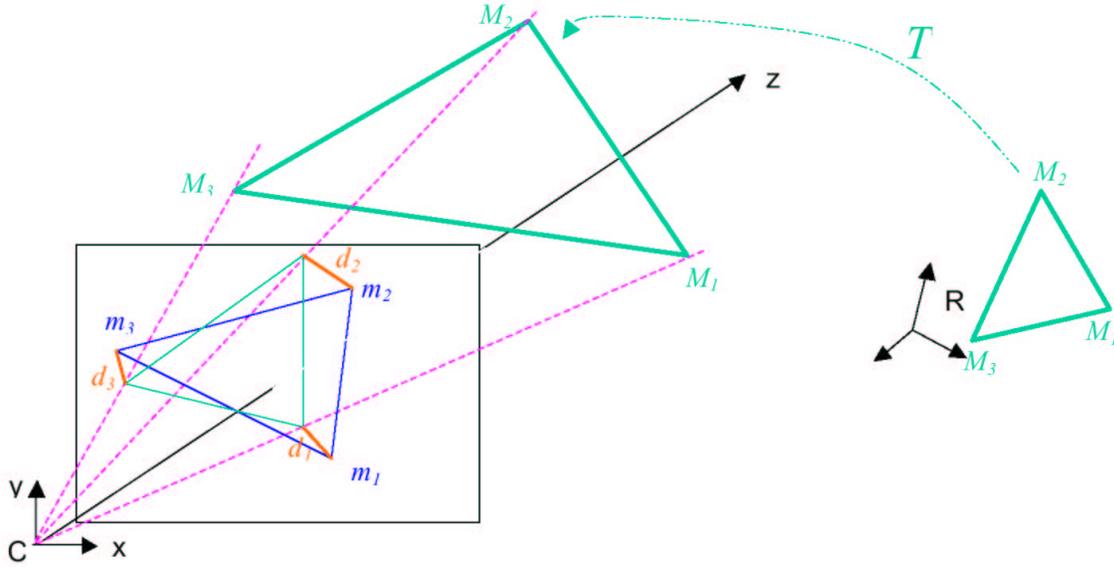


FIG. 3.8 – Illustration du principe de l'ISPPC : on cherche la transformation T qui minimise la somme des distances aux carrés séparant les m_i des projections dans l'image vidéo des M_i .

Minimisation dans l'espace objet (OSPPC)

Ce critère introduit par Lu [LHM00] minimise la somme des distances 3D au carré entre les points 3D du modèle et les demi-droites issues du centre optique C passant par les points 2D extraits des images vidéos (cf. figure 3.9). Le critère de minimisation se présente sous la forme :

$$C(T) = \sum_{j=1}^N \frac{1}{\| \vec{Cm}_i \|^2} \| \overrightarrow{C(T \star M_i)} \wedge \vec{Cm}_i \|^2$$

Ce critère sera désormais appelée OSPPC pour *Object Space Projective Points Criterion*.

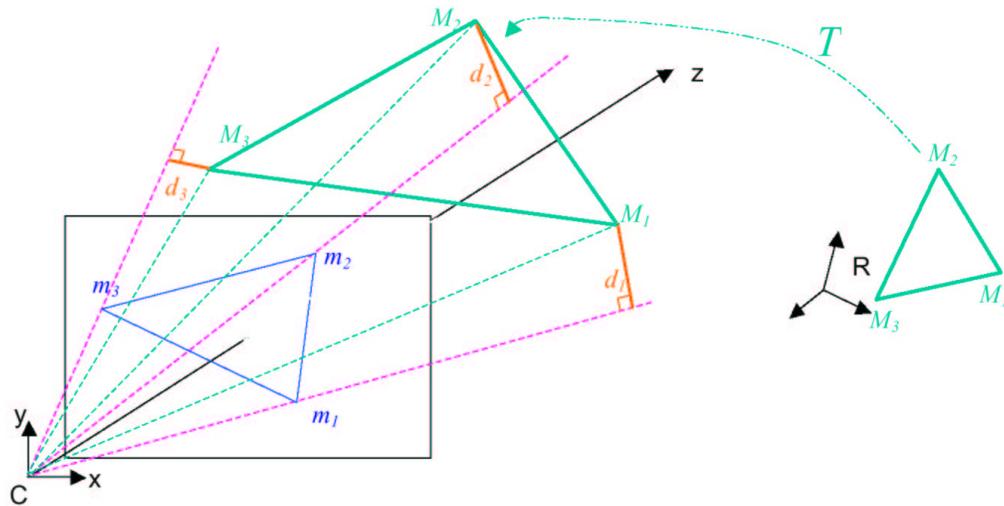


FIG. 3.9 – Illustration du principe de l’OSPPC : on cherche la transformation T qui minimise la somme des distances aux carrés séparant les points 3D M_i des demi-droites issues du centre optique C et passant par les m_i .

3.2.4 Comparaison et choix

Nous avons le choix entre trois méthodes générales différentes dont les données initiales, le temps de calcul et la précision ne sont pas identiques. Il faut donc considérer l’adaptabilité de chacune des méthodes à notre problème.

Les premiers résultats qualitatifs de recalage par maximisation d’information mutuelle ont été obtenus par Viola [Vio95, VW97] avec une application à la surface de la peau d’un visage. Les temps de calcul étaient proches de 0.5 sec. sur un PC à 1 GigaHz et les résultats visuellement satisfaisants. Leventon [Lev97, LWG97] a mesuré l’erreur de recalage à environ 3 mm (RMS) pour un objet (une voiture) mesurant 50 cm vu par deux caméras à une distance de 1 mètre et avec un angle de 70° . Plus récemment, Clarkson [CRHH99] a évalué l’erreur allant de 4.4 jusqu’à 12 mm dans un cas mono-vue et à 1.05 mm avec 5 caméras visionnant un squelette du crâne. La position relative des caméras et leur focale n’ont pas été spécifiées néanmoins, au vu des images fournies dans l’évaluation, le cône d’observation est de l’ordre de 180 stéradian et le crâne remplit l’image à chaque fois. Concernant la méthode iconique basée sur la photo-cohérence, Clarkson [CRHH00] obtient une erreur de 1.5 mm avec 4 vues et de 2.5 mm avec 2 vues (pour un angle de l’ordre de 40° entre les caméras). L’objet recalé était un visage et, comme dans le cas précédent, le visage remplissait à peu près toute l’image.

La précision obtenue par les méthodes iconiques semble donc compatible avec nos contraintes. Néanmoins, les objets 3D utilisés pour réaliser les évaluations possédaient systématiquement des courbes prononcées, contrairement à la peau abdominale qui est lisse et offre peu de singularités. D’ailleurs Clarkson [CRHH99] présente un exemple de recalage dans le cas d’une vue



FIG. 3.10 – Image type d'un cas où le recalage iconique est imprécis [CRHH99]. La surface du crâne ne présentant pas de variation significative de la courbure, le recalage est précis à 11 mm avec une initialisation à 7 mm (les billes noires servent à l'élaboration de la mesure étalon).

caméra du dessus du crâne (cf. figure 3.10) dont la précision finale est plus mauvaise que l'initialisation. Nous pensons donc que la forme cylindrique et la couleur homogène de l'abdomen humain peut conduire à de très fortes incertitudes le long de l'axe cranio-caudal. De plus ces méthodes requièrent des conditions d'illumination qui ne seront pas forcément satisfaites dans notre application [Her98]. Il semble donc difficile d'obtenir une précision suffisante avec cette méthode dans les conditions de notre intervention.

Le recalage de surface a été utilisé pour faire de la réalité augmentée dans le contexte de la neuro-chirurgie [GEW⁺96]. La surface de la peau extraite d'images IRM était recalée à une surface de points acquis sur le visage à l'aide d'un télémètre laser. L'acquisition de points du visage aurait aussi pu être réalisée à l'aide d'un système stéréoscopique (auquel cas il est préférable de projeter une texture mouchetée sur la peau pour obtenir une reconstruction dense [G.00]), ou bien en utilisant un système monoscopique et de la lumière structurée [GBDH95]. Grâce à des zones de courbures importantes (comme le nez, les oreilles), la précision mesurée était de l'ordre de 1 mm. A priori, on pourrait penser que recalcr deux surfaces contenant un nombre de points importants conduirait à une meilleure précision, proportionnelle à σ/\sqrt{n} (où σ est la variance sur le bruit d'extraction des points et n le nombre de points), mais ce n'est généralement pas le cas pour deux raisons. D'une part les mesures effectuées sur les points sont corrélées et ne représentent donc pas des marqueurs indépendants, d'autre part les points ne sont pas réellement homologues puisque deux reconstructions 3D d'une même surface contiennent des points physiquement différents. En outre, la géométrie particulière de l'abdomen détériore la précision du recalage. Comme l'abdomen n'a pas de zone fortement courbée, la translation le long de l'axe cranio-caudal sera peu contrainte, malgré une densité importante de points à recalcr.

Nous pensons donc qu'un recalage basé sur des points de référence est mieux adapté à notre problème. En effet, les temps de calculs sont généralement très bas (en dessous de 0.01

sec.) et la précision semble suffisante, même s'il y a peu d'évaluations cliniques (0.5 à 1.4 mm dans le cas d'un recalage 3D/3D de marqueurs implantés dans le crâne [MFW⁺97] et 2.0 à 2.5 mm dans le cas de marqueurs collés sur la peau [PPR⁺03]). Dans notre cas, il n'y a pas de marqueur anatomique visible, nous avons donc décidé de coller quelques marqueurs radio-opaques sur la peau du patient. Employer un recalage 3D/3D nous obligerait à ne sélectionner que les marqueurs visibles dans les deux caméras. Cela réduirait la précision potentielle (à cause de la perte d'information). De plus il faudrait calculer les covariances (anisotropes) sur les reconstructions afin de minimiser un critère tenant compte de toute l'information a priori. Nous avons donc décidé d'opter pour un recalage 3D/2D de marqueurs. Nous verrons dans le chapitre 6 comment nous extrayons automatiquement la position des marqueurs dans les images TDM et vidéo, et comment nous les apparions. Cette fois-ci, comme nos points de référence sont réellement homologues, la forme cylindrique de la peau n'est plus un problème. Il ne nous reste qu'à choisir la technique de calcul de la transformation solution, mais nous verrons dans la section suivante que les hypothèses statistiques des techniques existantes modélisent imparfaitement nos conditions expérimentales.

3.3 Des critères de recalage ponctuel 3D/2D inappropriés

Le recalage 3D/2D, connu aussi sous l'acception *estimation de pose* est un problème traité depuis plusieurs décennies, puisqu'il correspond en fait à l'étape de détermination des paramètres extrinsèques d'une caméra lors de son calibrage. De manière générale, les méthodes cherchent à estimer la solution du critère ISPPC (le critère OSPPC est très récent), leur différence provenant essentiellement de la méthode utilisée pour calculer la transformation minimisant le critère. Dans cette section, nous passerons en revue les principales méthodes existantes et ferons ressortir après analyse celles qui fournissent les meilleures précisions. Finalement, nous mettrons en avant leurs lacunes statistiques qui font qu'elles ne sont pas optimales pour notre problème.

3.3.1 Les critères et méthodes de recalage 3D/2D existants

Les travaux pionniers ne considéraient le problème qu'avec 6 points ou moins et proposaient une solution directe utilisable dans la plupart des configurations géométriques [FB81, DRL89, HYH85]. Ces méthodes sont directes mais très sensibles au bruit puisque la transformation est calculée à partir de relations géométriques exactes entre deux ensembles de points dont le nombre est minimal. Quand le nombre de correspondances de points est plus grand que le nombre d'inconnus, il est préférable de se servir de tous les points pour minimiser l'influence de l'erreur de mesure, et de prendre en compte toute l'information possible.

Ganapathy et d'autres [Gan84, LHF90] ont proposé une résolution linéaire remplaçant la rotation par une matrice 3×3 sans aucune contrainte. Cela signifie qu'ils ont approximé la transformation rigide recherchée par une transformation affine. Cette résolution linéaire présente l'avantage de fournir une solution directe au problème d'estimation. Néanmoins, elle est très sensible au bruit à cause de l'absence de contraintes sur la matrice de rotation. Plus récemment, Quan [QL99] et Ansar [AD03] présente une solution directe pour 4 points et plus, basée sur une contrainte de distance inter-points. Leur formulation du problème les conduit à un système polinomial de degrés 4 dont la résolution permet d'obtenir la profondeur par rapport à la caméra des points dans les images vidéo.

Yuan [Yua89] et Liu [LHF90] et DeMenthon [DD95] résolvent le problème en séparant la composante rotationnelle de la translationnelle. Yuan a démontré que, étant donné les correspondances 3D et 2D, les paramètres de rotation sont les racines communes de six équations quadratiques. Alors, elles sont trouvées en se servant d'une descente de gradient. Dans le cas de correspondances 3D/2D de lignes, Liu trouve itérativement la composante rotationnelle en se servant des angles d'Euler. DeMenthon propose plutôt d'approximer la *projection perspective* par une *projection orthographique*. Le résultat direct obtenu par sa formulation linéaire du problème est ensuite suivie d'un algorithme raffinant la solution initiale.

Lowe, Phong et Simon [Low91, PHT95, SB98] déterminent la pose en minimisant itérativement le critère classique aux moindres carrés. Dans ces cas, les méthodes diffèrent par la méthode de minimisation (descente de Newton, algorithme de Powell, méthode par région de confiance), et par la paramétrisation de la rotation (vecteur rotation, angle d'Euler, quaternions duaux).

Haralick [HJL⁺89] et Or [OLWK98] ont tourné le problème 3D/2D en un problème 3D/3D dans lequel ils estiment la profondeur des points observés dans les images vidéo. Ils minimisent tout deux une distance euclidienne 3D en deux étapes. Or, dérivant un critère au sens du maximum de vraisemblance, estime itérativement la transformation et la profondeur des points. Haralick fournit un calcul itératif de la profondeur basée sur des considérations géométriques. De plus chaque terme de son critère est pondéré par la qualité de l'observation. De même, Lu [LHM00] se replace dans un contexte 3D/3D en minimisant la distance entre les points 3D du modèle à recalculer et les demi-droites issues du centre optique et passant par la projection 2D des points dans les images vidéo.

3.3.2 Des hypothèses statistiques mal adaptées à nos conditions

Voyons maintenant lesquelles de ces approches semblent les plus adaptées à notre problème. L'émergence des résolutions linéaires a été motivée par les applications temps-réel et, jusque dans les années 90, les temps de calculs nécessaires pour optimiser un critère aux moindres carrés étaient prohibitifs. Actuellement, cela peut-être réalisé en moins de 0.01 sec. pour une trentaine de points avec une initialisation peu précise. En conséquence, comme la précision est cruciale dans notre application, nous ne pouvons pas nous permettre de nous servir d'une solution linéaire plus sensible aux bruits présents dans nos procédures d'extraction de données.

L'approche alternative consistant à séparer la rotation de la translation a été examinée par Kumar [Kum89], et il montre que cela conduit à une estimation des paramètres plus mauvaise que l'estimation par optimisation itérative du critère classique ISPPC. Malheureusement, il n'y a pas d'article de comparaison à propos de la méthode de Haralick et Or, il n'est donc pas possible de critiquer leurs performances en précision.

D'un point de vue précision, Ansar [AD03] montre que son approche sur les contraintes inter-points, et l'optimisation de l'OSPPC et l'ISSPC mènent à des résultats à peu près équivalents. De ces techniques restantes, nous pensons néanmoins qu'une estimation de la transformation par optimisation itérative du critère aux moindres carrés ISPPC ou OSPPC a un sérieux avantage puisqu'il est alors possible de prédire précisément l'influence du bruit sur le recalage final et donc, sur les points à recalculer [PT97]¹³. Pour notre application cela est crucial puisque cela permettra de détecter les mauvaises configurations de marqueurs ou de caméras qui conduiraient à un recalage potentiellement dangereux.

13. L'analyse des perturbations réalisé par Ansar sur sa propre méthode lui permet uniquement de fournir une surestimation grossière (d'un facteur supérieur à 5) de l'erreur moyenne sur le recalage.

Cependant, ces deux critères ne sont pas parfaitement adaptés à notre problème. En effet, de par leur formulation (minimisation dans un espace 2D *ou* 3D), le critère aux moindres carrés ISPPC considère implicitement que les données 2D sont bruitées mais que les données 3D sont exactes, et le critère OSPPC suggère à l'inverse que seules les données 3D sont bruitées. Dans notre application, ces hypothèses sont erronées puisque nous avons un bruit d'extraction présent à la fois dans nos images vidéo *et* dans l'image TDM. Le chapitre suivant étayera cette remarque et montrera qu'il est possible d'obtenir un critère généralisant ces deux approches.

3.4 Conclusion

Dans ce chapitre, nous avons analysé les méthodes de recalage pour résoudre notre problème. En effet, le système de réalité augmentée que nous souhaitons construire doit satisfaire certaines contraintes de précision et de temps de calcul.

Une étude bibliographique de la position du foie en fonction de la respiration, confirmée par notre expérimentation sur deux patients, aura permis de prouver que les mouvements du foie dans nos conditions cliniques ne sont pas suffisamment importants pour justifier l'emploi d'un recalage non-rigide. Ensuite, nous avons montré que, pour des raisons de précision, une approche recalage ponctuel 3D/2D est préférable aux approches surfaciques et iconiques. La dernière étape de l'analyse, consistant à trouver la meilleure méthode d'estimation parmi les techniques existantes, a mis en avant la précision des critères ISPPC et OSPPC par rapport aux autres méthodes. Néanmoins, nous avons fait remarqué que ces deux critères ont des lacunes au niveau de leurs hypothèses statistiques. Cela nous mène au chapitre suivant dans lequel nous développons un nouveau critère de recalage 3D/2D ponctuel.

Chapitre 4

Un nouveau critère de recalage 3D/2D

Comme nous l'avons souligné dans le paragraphe 3.3.2, nos données observées 2D *et* 3D sont corrompues par un bruit d'extraction. Les hypothèses statistiques des critères de recalage 3D/2D existants ne tenant compte du bruit que sur un type de donnée (2D ou 3D), la transformation ne sera pas estimée de manière optimale. Nous montrons dans ce chapitre comment réviser les hypothèses statistiques classiques pour dériver au sens du maximum de vraisemblance un critère de recalage adapté aux conditions de notre application. Une considération de cas limites sur les hypothèses montrera que notre nouveau critère (désormais dénommé EPPC pour *Extended Points Projective Criterion*) généralise les critères aux moindres carrés ISPPC et OSPPC (définis dans la section 3.2.3), ainsi que la méthode de recalage basée sur une reconstruction stéréoscopique suivie d'un recalage ponctuel 3D/3D. Une évaluation des performances de l'EPPC, de l'ISPPC et de l'OSPPC sur des données synthétiques proches de celles de notre application montrera ensuite que l'EPPC apporte un gain moyen de 10% sur la précision et offre une robustesse accrue pour des temps de calcul 10 à 30 fois supérieurs.

Afin de confirmer le gain de précision de l'EPPC sur l'ISPPC et l'OSPPC dans des conditions représentatives de notre application, nous nous penchons ensuite sur le problème de la validation avec des données réelles. Classiquement, nous devrions rechercher des mesures étalon afin de pouvoir comparer les performances des trois critères. Néanmoins, ces mesures ne sont pas disponibles dans le cadre de notre application (nous n'aurions pas alors développé un nouveau critère). Nous montrons finalement comment contourner cet obstacle en concevant une méthodologie d'évaluation statistique de l'erreur de recalage sans mesure étalon. Nous vérifions ainsi que l'EPPC apporte effectivement une précision supérieure aux autres critères de l'ordre de 9% sur des données provenant d'un mannequin de l'abdomen humain.

4.1 Dérivation du nouveau critère EPPC

Ce paragraphe présente les étapes menant à la formulation statistique d'un critère adapté à nos hypothèses. Nous rappelons tout d'abord comment l'estimation de la transformation recherchée au sens du maximum de vraisemblance mène au critère ISPPC, quand les hypothèses supposent uniquement la présence d'un bruit sur les données 2D. Cette première étape constitue la base du calcul théorique suivant qui introduit l'existence d'un bruit corrompant les données 3D. Finalement, nous détaillons les considérations statistiques et géométriques qui permettent

de regarder l'EPPC comme un critère généralisant les méthodes habituelles.

4.1.1 Notations

Soient $M_i = [x_i, y_i, z_i]^\top$ les points 3D qui correspondent à la position exacte dans le repère du scanner du centre des marqueurs radio-opaques disposés sur le corps du patient et $m_i^{(l)} = [u_i, v_i]^\top$ les points 2D qui correspondent à la position exacte en unité réelle du centre des marqueurs radio-opaques dans les images vidéo. Nous supposons que les correspondances sont connues i.e. que M_i est apparié à $m_i^{(l)}$. L'indice (l) indique à quelle caméra l'appariement fait référence. La prise en compte des occlusions est réalisée par l'introduction de la variable binaire ξ_i^l qui vaut 0 si M_i n'est pas visible ou n'est pas dans le champ de la caméra (l) et 1 si le point est observé.

Nous notons $T \star M$ l'action de la transformation rigide T sur les points 3D M_i . Soit P_l ($1 \leq l \leq M$) la fonction de projection 3D/2D de la caméra (l). Nous avons donc la relation suivante entre un point non bruité défini dans l'espace et son correspondant "exact" dans l'espace image de la caméra (l) :

$$P^{(l)}(T \star M_i) = m_i^{(l)} \quad (4.1)$$

Dans la suite, \hat{A} représentera une *estimation* de la donnée parfaite A , et \tilde{A} une *mesure ou observation de cette donnée*. Ainsi, les points 3D mesurés seront écrits \tilde{M}_i et les points 2D mesurés dans l'image vidéo \tilde{m}_i .

4.1.2 Hypothèses du critère ISPPC sur le bruit 3D

Supposons que nos données 3D soient exactes (i.e. que les points 3D \tilde{M}_i mesurés sont égaux aux points 3D M_i exacts), et que seuls les points 2D sont corrompus par un bruit gaussien centré isotrope η_i de matrice de covariance :

$$\Sigma_{2D} = \begin{pmatrix} \sigma_{2D}^2 & 0 \\ 0 & \sigma_{2D}^2 \end{pmatrix}$$

En nous servant de la relation géométrique reliant les points exacts 3D et 2D (Eq. 4.1), le modèle du bruit corrompant les points 2D mesurés s'écrit :

$$\tilde{m}_i^{(l)} = m_i^{(l)} + \eta_i = P^{(l)}(T \star M_i) + \eta_i \quad \text{avec} \quad \eta_i \sim N(0, \sigma_{2D})$$

En conséquence, la probabilité de mesurer la projection du point 3D $M_i = [x_i, y_i, z_i]^\top$ à la position $\tilde{m}_i^{(l)} = [u_i^{(l)}, v_i^{(l)}]^\top$ dans l'image vidéo (l), sachant les paramètres (de transformation) $\theta = \{T\}$, est donnée par :

$$p(\tilde{m}_i^{(l)} | \theta) = k_{2D} \cdot \exp \left(- \frac{\| P^{(l)}(T \star M_i) - \tilde{m}_i^{(l)} \|^2}{2\sigma_{2D}^2} \right) = G_{\sigma_{2D}} \left(P^{(l)}(T \star M_i) - \tilde{m}_i^{(l)} \right),$$

où $k_{2D} = 1/2\pi\sigma_{2D}^2$ est une constante de normalisation.

Soit χ notre vecteur de données (regroupant toutes les données mesurées, i. e. seuls les points 2D $\tilde{m}_i^{(l)}$, puisque les points 3D sont supposés exacts). La détection de chacun des points 2D

étant réalisée de manière indépendante, la probabilité des mesures observées s'écrit simplement $p(\chi | \theta) = \prod_{l=1}^M \prod_{i=1}^N p(\tilde{m}_i^{(l)} | \theta)^{\xi_i^l}$. Dans cette formule, les points 2D non-observés (pour lesquels $\xi_i^l = 0$) n'ont implicitement aucune influence sur le calcul de la probabilité.

Maintenant, la transformation cherchée au sens du *maximum de vraisemblance* est la transformation $\hat{\theta}$ qui maximise la probabilité des données observées, ou de manière équivalente, qui minimise son logarithme négatif :

$$\hat{\theta} = \arg \min_{\theta} -\log[p(\chi | \theta)]$$

En développant la fonction de vraisemblance nous obtenons :

$$-\log[p(\chi | \theta)] = -\log \left[\prod_{l=1}^M \prod_{i=1}^N p(\tilde{m}_i^{(l)} | \theta)^{\xi_i^l} \right] = \sum_{l=1}^M \sum_{i=1}^N -\log \left[p(\tilde{m}_i^{(l)} | \theta)^{\xi_i^l} \right]$$

Après simplification, nous terminons par la minimisation de :

$$C_{2D}(T) = \sum_{l=1}^M \sum_{i=1}^N \xi_i^l \cdot \frac{\|P^{(l)}(T \star M_i) - \tilde{m}_i^{(l)}\|^2}{2 \cdot \sigma_{2D}^2} - \left(\sum_{l=1}^M \sum_{i=1}^N \xi_i^l \right) \cdot \log[k_{2D}] \quad (4.2)$$

Par suite, à une constante et un facteur multiplicatif près, cette estimation au maximum de vraisemblance mène au critère classique ISPPC. Ce critère suppose qu'il n'y pas de bruit sur les données 3D, l'idée étant que le bruit 3D possible soit distribué sur les erreurs de mesure 2D. Cette hypothèse simpliste permet d'aboutir à un critère quadratique avec seulement 6 paramètres à estimer, ce qui implique évidemment de très faibles temps de calcul (en dessous de 0.01s pour une minimisation impliquant 30 points, sur un PC 1.1GHz). Néanmoins, d'un point de vue statistique, distribuer l'erreur 3D sur les mesures 2D implique leur corrélation, ce qui ne correspond pas à l'hypothèse d'indépendance utilisée pour dériver le critère d'estimation au sens du maximum de vraisemblance.

4.1.3 Un critère étendu dérivé du maximum de vraisemblance

Pour introduire la présence d'un bruit sur les données 3D, il est plus raisonnable de considérer que nous mesurons une version bruitée des points 3D exacts (nous supposons aussi un bruit gaussien isotrope):

$$\tilde{M}_i = M_i + \varepsilon_i \quad \text{avec} \quad \varepsilon_i \sim N(0, \sigma_{3D})$$

Dans ce cas, la position exacte M_i des points 3D est considérée en tant que paramètre, comme la transformation T . En statistique, cela est appelé une *variable cachée ou latente*, alors que la dénomination de *variable auxiliaire* est plus connue en vision par ordinateur. Il s'ensuit que la probabilité de mesurer un point 3D sachant $\theta = \{T, M_1, \dots, M_N\}$ est :

$$p(\tilde{M}_i | \theta) = G_{\sigma_{3D}} \left(M_i - \tilde{M}_i \right), \quad (4.3)$$

et la probabilité de mesurer un point 2D reste :

$$p(\tilde{m}_i^{(l)} | \theta) = G_{\sigma_{2D}} \left(P^{(l)}(T \star M_i) - \tilde{m}_i^{(l)} \right) \quad (4.4)$$

Une des caractéristique importante de cette modélisation statistique est que nous pouvons supposer sans crainte que toutes les mesures 3D et 2D sont indépendantes¹⁴. Nous pouvons donc ré-écrire la probabilité de notre vecteur observé $\chi = (\tilde{m}_1^1, \dots, \tilde{m}_N^1, \dots, \tilde{m}_1^M, \dots, \tilde{m}_N^M, \tilde{M}_1, \dots, \tilde{M}_N)$ comme le produit des probabilités individuelles :

$$p(\chi | \theta) = \left(\prod_{l=1}^M \prod_{i=1}^N p(\tilde{m}_i^{(l)} | \theta)^{\xi_i^l} \right) \cdot \left(\prod_{i=1}^N p(\tilde{M}_i | \theta) \right) \quad (4.5)$$

L'estimation au maximum de vraisemblance restant l'optimisation de la probabilité des données observées, nous obtenons $C(\theta) = C(T, M_1, \dots, M_N) = -\log(p(\chi | \theta))$. Après introduction des équations 4.3 et 4.4 dans 4.5, nous avons :

$$C(T, M_1, \dots, M_N) = \sum_{i=1}^N \frac{\|\tilde{M}_i - M_i\|^2}{2 \cdot \sigma_{3D}^2} + \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \frac{\|\tilde{m}_i^{(l)} - m_i^{(l)}\|^2}{2 \cdot \sigma_{2D}^2} + K(\sigma_{2D}, \sigma_{3D}) \quad (4.6)$$

où $m_i^{(l)} = P^{(l)}(T \star M_i)$ et K est une constante de normalisation dépendante de σ_{2D} et σ_{3D} .

La différence évidente entre ce critère et l'ISPPC est qu'il y a de nouvelles inconnues dans le système à estimer (les variables cachées). Il est donc nécessaire de modifier l'algorithme d'optimisation puisqu'il y a maintenant deux ensembles d'inconnues à estimer, le premier représentant la transformation rigide et le deuxième correspondant aux M_i .

Parmi les différentes possibilités envisageables, nous avons finalement opté pour une minimisation alternée du critère C . En premier lieu, nous initialisons les M_i avec les \tilde{M}_i que nous avons relevés puis nous effectuons une première minimisation sur T (à partir d'une initialisation T_0). On pourra remarquer que cette étape revient en fait à minimiser le critère ISPPC. En second lieu, nous fixons T et nous effectuons la minimisation du critère sur les M_i . Par suite, nous obtenons une estimation des M_i au sens du maximum de vraisemblance. Nous continuons ensuite les itérations en remplaçant les nouveaux \hat{M}_i estimés dans le critère et en itérant à nouveau la minimisation de C sur T . La convergence est assurée étant donné que nous minimisons à chaque étape le même critère et qu'il est à valeurs positives.

Nous arrêtons l'algorithme une fois que la transformation résiduelle $\hat{T}^{i+1} \star (\hat{T}^i)^{-1}$ a une composante rotationnelle inférieure à d_{max}^r et une composante translationnelle inférieure à d_{max}^t . Les valeurs d_{max}^r et d_{max}^t sont choisies en fonction de la précision voulue sur la transformation. Les étapes de l'algorithme sont résumées sur la figure 4.1.

En toute rigueur, nous devrions rajouter un second terme dans le critère d'arrêt qui tiendrait compte de la seconde partie des données, à savoir les coordonnées estimées des points M_i . Néanmoins, nous ne nous intéressons pas aux résultats de l'estimation des M_i , il n'est donc pas vraiment nécessaire d'en tenir compte, d'autant plus que le critère semble suffisamment lisse pour éviter toutes formes d'oscillations entre les estimations successives de T et des M_i .

Notons par ailleurs qu'il ne serait pas forcément judicieux de se servir tout simplement de la valeur du critère C pour décider quand arrêter l'algorithme. En effet, nous ne connaissons pas réellement les valeurs de σ_{3D} et de σ_{2D} , ce qui signifie que nous ne connaissons pas la vraie valeur du critère et ne pouvons donc pas dire à quel moment nous sommes suffisamment proche de la convergence de l'algorithme.

14. Plus précisément, les mesures des points 3D et 2D sont conditionnellement indépendantes étant donné les paramètres $\theta = \{T, M_1, \dots, M_N\}$.

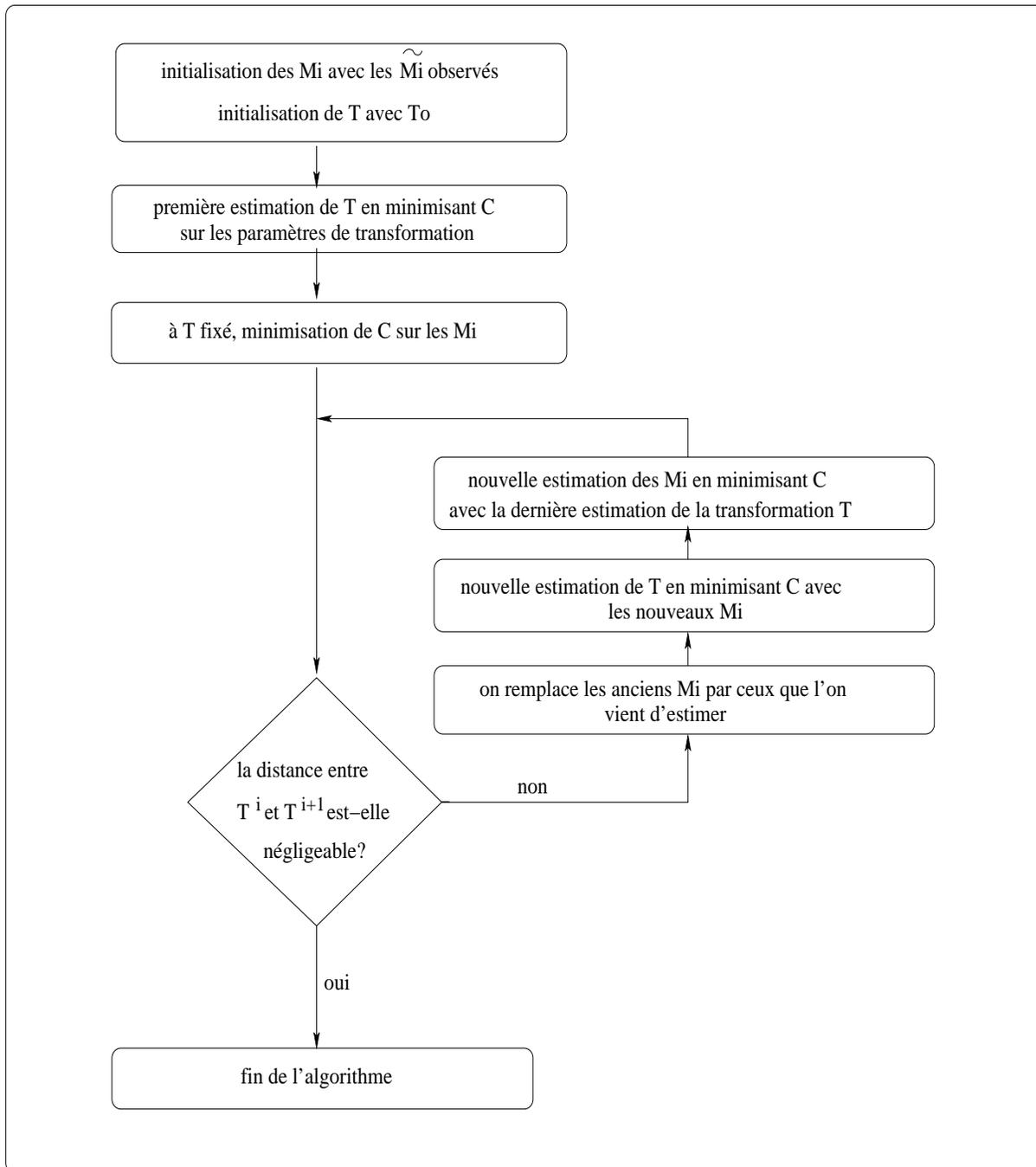


FIG. 4.1 – Schéma descriptif de l'algorithme utilisé pour optimiser l'EPPC. Nous avons opté pour une minimisation alternée des paramètres de T et des coordonnées exactes M_i .

4.1.4 Cas de bruits 3D et 2D anisotropes

Jusqu'à présent, nous avons considéré des bruits 2D et 3D isotropes. Cependant, la plupart des acquisitions TDM ne sont pas isotropes (l'épaisseur des coupes est souvent plus grande que la largeur et la longueur d'un pixel contenu dans une coupe). Dans ce cas, l'erreur de localisation 3D des marqueurs est anisotrope :

$$\varepsilon_i \sim N(0, \Sigma_{3D}) \quad \text{avec} \quad \Sigma_{3D} = \begin{pmatrix} \sigma_{3D_x}^2 & 0 & 0 \\ 0 & \sigma_{3D_y}^2 & 0 \\ 0 & 0 & \sigma_{3D_z}^2 \end{pmatrix}$$

ce qui implique quelques modifications mineures dans la formulation de notre critère : il faut remplacer dans notre critère (équation 4.6) le premier terme $\frac{\|\tilde{M}_i - M_i\|^2}{2 \cdot \sigma_{3D}^2}$ par la moitié de la distance de Mahalanobis $(\tilde{M}_i - M_i)^\top \cdot \Sigma_{3D}^{-1} \cdot (\tilde{M}_i - M_i)$. Dans le cas où le bruit 2D est aussi anisotrope, des modifications équivalentes doivent être apportées au second terme de l'équation.

4.1.5 Un critère qui généralise les méthodes classiques

Lien avec l'ISPPC

Le lien avec l'ISPPC est immédiat vu que ce critère est aussi dérivé d'une approche maximum de vraisemblance, mais avec des hypothèses moins complètes. De fait, l'EPPC se comporte comme l'ISPPC si le bruit sur les données 3D est sous-estimé ou bien très faible par rapport au bruit sur les données 2D

Lien avec la reconstruction stéréoscopique et le recalage 3D/3D

Considérons le changement de variable $M'_i = T \star M_i$. Nous mesurons alors les points 3D exacts dans le repère des caméras (à la place du repère scanner) et recherchons la transformation $T' = T^{(-1)}$ (la transformation du repère caméra au repère scanner au lieu de la transformation du scanner vers les caméras). La présence des primes étant transparente pour la compréhension, le critère EPPC peut donc se ré-écrire

$$C(T, M_1, \dots, M_N) = \sum_{i=1}^N \frac{\|\tilde{M}_i - T \star M_i\|^2}{2 \cdot \sigma_{3D}^2} + \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \frac{\|\tilde{m}_i^{(l)} - P^{(l)}(M_i)\|^2}{2 \cdot \sigma_{2D}^2} + K(\sigma_{2D}, \sigma_{3D}) \quad (4.7)$$

Comme précédemment expliqué dans le paragraphe 4.1.3, ce critère peut être optimisé itérativement en estimant successivement les coordonnées 3D et la transformation T . Notons aussi que le changement de variable n'affecte pas la transformation à calculer.

Maintenant, si nous supposons que nous réalisons une optimisation avec un σ_{3D} largement *surestimé* et une estimation correcte de σ_{2D} , nous avons autour de la transformation optimale \hat{T} :

$$\frac{1}{N} \sum_{i=1}^N \|\tilde{M}_i - T \star M_i\|^2 \simeq \frac{1}{N} \sum_{i=1}^N \|\tilde{M}_i - \hat{T} \star M_i\|^2 = \hat{\sigma}_{3D}^2 \ll \sigma_{3D}^2$$

Par conséquent, le premier terme du critère est négligeable par rapport au second terme (σ_{2D} étant correctement estimé) : l'optimisation de la position des points 3D conduit à la minimisation de

$$C_{Rec}(M_1, \dots, M_N) = \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \|\tilde{m}_i^{(l)} - P^{(l)}(M_i)\|^2$$

Ce critère est en fait l'un des plus utilisés en reconstruction. Ensuite, après la détermination des coordonnées 3D des points dans le repère caméra, l'étape suivante consiste à minimiser le critère par rapport à T . Comme le second terme ne dépend pas de T , cela revient à minimiser

$$C_{3Drec}(T) = \sum_{i=1}^N \|\tilde{M}_i - T * M_i\|^2$$

qui n'est autre que le critère standard de recalage ponctuel 3D/3D aux moindres carrés.

En conclusion, la méthode consistant à reconstruire la position des points 3D à partir de leur position image et à les recalculer ensuite en 3D peut être vue comme un cas limite de l'EPPC quand le bruit sur les données 3D est largement *suresstimé* (par rapport au bruit sur les données 2D).

Lien avec l'OSPPC

Considérons comme précédemment le changement de variable $M_i' = T * M_i$, mais positionnons l'action de T sur les \tilde{M}_i plutôt que sur les M_i . Ce choix ne modifie pas la solution du critère puisque la transformation recherchée est rigide. Le critère EPPC s'écrit cette fois-ci :

$$C(T, M_1, \dots, M_N) = \sum_{i=1}^N \frac{\|T * \tilde{M}_i - M_i\|^2}{2 \cdot \sigma_{3D}^2} + \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \frac{\|\tilde{m}_i^{(l)} - P^{(l)}(M_i)\|^2}{2 \cdot \sigma_{2D}^2} + K(\sigma_{2D}, \sigma_{3D})$$

et est identique au critère 4.7 à la différence que nous allons bien chercher T et non pas T^{-1} .

Supposons qu'il n'y ait qu'une caméra de centre optique O observant la scène et que σ_{2D} tende vers 0 (ou bien supposons que σ_{2D} soit fortement sous-estimé par rapport à σ_{3D}). Dans ce cas limite, à T fixé, la valeur du second terme du critère sera très importante par rapport à celle du premier si les points M_i ne sont pas "exactement" sur les demi-droites $[O\tilde{m}_i]$. Les M_i qui minimisent C sont donc nécessairement très proches des demi-droites correspondantes. Optimiser C revient donc à optimiser le premier terme sous la contrainte que les M_i appartiennent aux demi-droites $[O\tilde{m}_i]$. Cela signifie que la transformation solution T sera celle qui minimise la distance des \tilde{M}_i aux demi-droites $[O\tilde{m}_i]$. Nous retrouvons finalement la considération géométrique présentée dans la section 3.2.3 qui mène à la formulation du critère OSPPC.

Dans ces conditions, nous devrions donc nous attendre à obtenir des comportements similaires entre l'OSPPC et l'EPPC. En revanche, dans le cas où σ_{2D} est non-nul, nous devrions observer une supériorité de l'EPPC sur l'OSPPC. Une brève étude menée sur des données synthétiques vérifie ce comportement. En effet, pour un rapport $\sigma_{2D}/\sigma_{3D} \gg 1$, l'EPPC apporte un gain de précision de l'ordre de 4% et pour un rapport $\sigma_{2D}/\sigma_{3D} \ll 1$, leurs précisions sont comparables à 0.5% près (les conditions de cette expérience correspondent à celle décrites dans la section 4.2.1, à la différence que nous ne considérons qu'une seule caméra).

En revanche, si ce comportement est vrai pour une caméra, pour deux caméras ou plus, l'interprétation proposée ci-dessus n'est plus valable. En effet, si σ_{2D} tend vers 0 (ou bien σ_{2D}

fortement sous-estimé par rapport à σ_{3D}), la minimisation de C correspond en fait à une reconstruction stéréoscopique, suivie d'un recalage 3D/3D. Observons maintenant sur la figure 4.2 le positionnement de 2 points \tilde{M}_i qui minimisent le critère OSPPC. Il apparaît que l'OSPPC va pondérer l'importance des points en fonctions de leur distance par rapport aux caméras : le point \tilde{M}_1 va être placé à une distance d_1 de M_1 plus grande que la distance d_2 entre \tilde{M}_2 et M_2 . Cette pondération n'existant pas pour l'EPPC (on va minimiser équitablement les distances d_i), ses performances devraient être meilleures.

Finalement, nous venons de montrer que les critères et méthodes classiques utilisés pour résoudre ce problème de recalage peuvent être interprétés comme des cas limites (sur les hypothèses statistiques) du critère EPPC. Il est donc raisonnable de s'attendre à ce que l'EPPC soit plus efficace que les autres méthodes s'il y a effectivement du bruit sur les données 3D et si nous possédons une bonne estimation du rapport des variances 3D et 2D.

Remarque sur la méthode d'optimisation de Lavest

Dans le contexte du calibrage de caméra, Lavest [LVD98] estime la position des points 3D de la mire de calibrage (sans variable cachée), en plus des paramètres intrinsèques et extrinsèques de la caméra. Ses expériences démontrent l'intérêt de sa méthode dans le cas où la connaissance des points de la mire de calibrage est peu précise par rapport à celle des points 2D correspondants dans l'image vidéo.

Dans le cadre de notre application, les marqueurs 3D et 2D sont extraits avec un RSB du même ordre de grandeur (cf. section 6.1.1). La technique de Lavest n'est donc d'aucun intérêt pour nous. En effet, adapter sa méthode à notre contexte correspond à optimiser l'ISPPC sur la transformation T et les M_i . Cela revient à reconstruire par stéréoscopie la position des points 3D à partir des points 2D relevés, et à remplacer les M_i (obtenus initialement à partir de l'image scanner) par les points 3D reconstruits. Cette approche considère donc implicitement que le bruit corrompant les données 3D est infini, ce qui ne correspond pas à nos conditions réelles.

Nous en profitons pour glisser un mot à propos de l'EPPC et du calibrage de caméras. En général, la mire utilisée pour calibrer une caméra est connue avec une très grande précision (en dessous de 0.02 mm), auquel cas le rapport RSB_{2D}/RSB_{3D} est très grand, même si l'extraction des amers dans les images vidéo est sub-pixellique. Dans le cas de figure où la définition de la mire est moins bonne (ou bien l'extraction vidéo excellente), il peut être intéressant de rajouter un second terme (équivalent à celui de l'EPPC) dans le critère d'optimisation classique utilisé pour calibrer une caméra. Ce second terme permettrait de tenir compte de l'erreur 3D et devrait donc fournir un meilleur calibrage.

4.2 Evaluation des performances sur des données synthétiques

Nous proposons dans cette section une étude comparative des trois critères exposés précédemment (ISPPC, OSPPC, et EPPC). Le but est d'évaluer les performances de chacun des critères en temps de calcul, en précision et en robustesse.

L'intérêt de cette partie est motivé par les contraintes de notre application. En effet, nous avons pour objectif de réaliser un recalage en temps réel des organes internes sur les images vidéo avec une précision inférieure à 5 mm et avec un rafraîchissement d'image à 25Hz. Nous

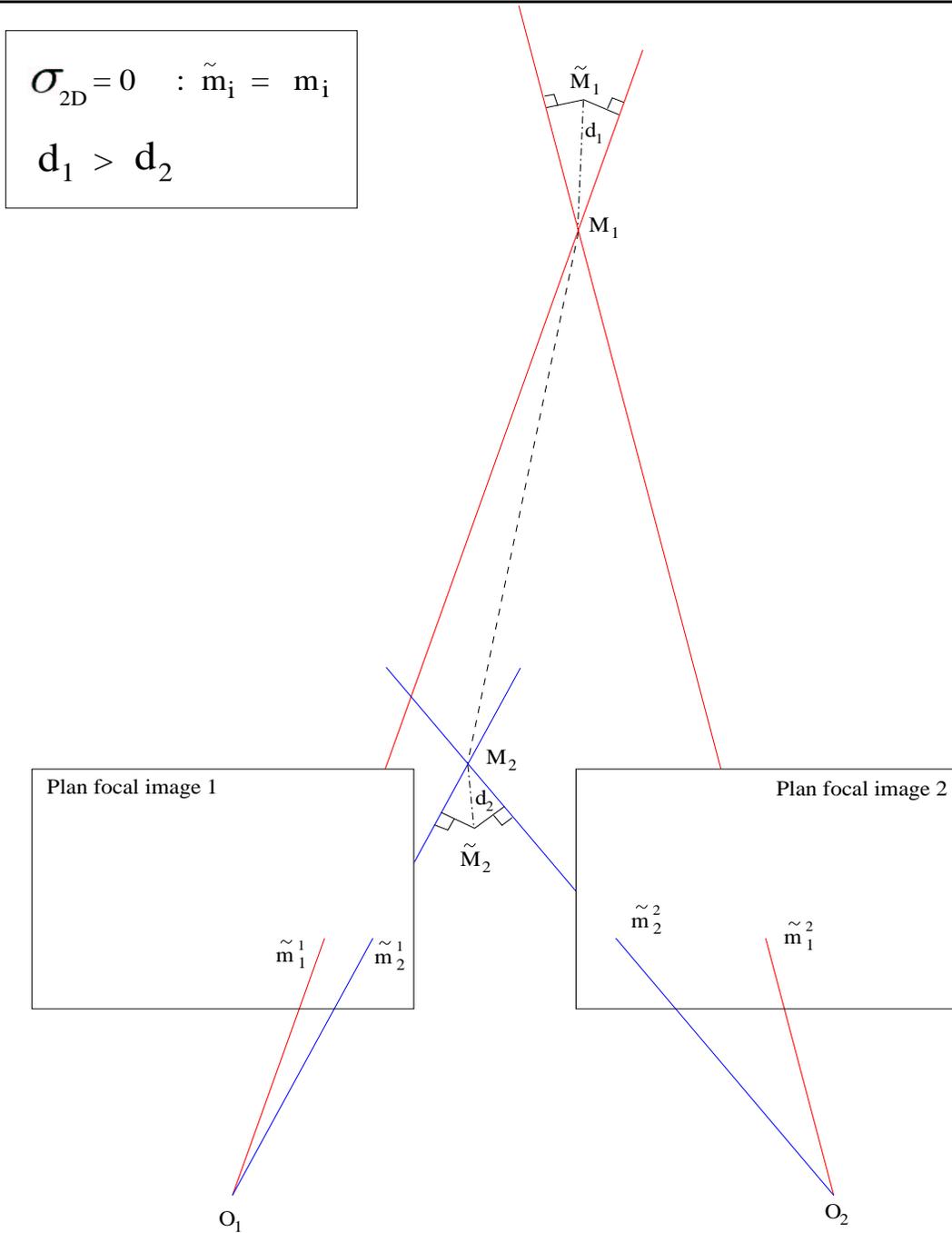


FIG. 4.2 – Comparaison de l'EPCC et de l'OSPPC avec deux caméras et deux points lorsque $\sigma_{2D} \rightarrow 0$ et $\sigma_{3D} \neq 0$. Les droites $O_1\tilde{m}_1^1$ et $O_1\tilde{m}_2^1$ (resp. $O_1\tilde{m}_1^2$ et $O_1\tilde{m}_2^2$) se coupent en M_1 (resp. M_2) étant donné que $\sigma_{2D} = 0$ (et donc $\tilde{m}_i = m_i$). Dans cet exemple, l'objet $\tilde{M}_1\tilde{M}_2$ est recalé au sens du critère OSPPC sur l'objet M_1M_2 observé par les deux caméras : la position du point \tilde{M}_1 (resp. \tilde{M}_2) minimise les distances aux droites $O_1\tilde{m}_1^1$ et $O_2\tilde{m}_2^2$ (resp. $O_1\tilde{m}_1^2$ et $O_2\tilde{m}_2^1$). On voit que la transformation solution du critère OSPPC pondère les rôles de \tilde{M}_1 et \tilde{M}_2 en fonction de leur distance aux caméras puisque $d_1 > d_2$. Dans ce cas limite, l'EPCC se comporterait comme une reconstruction stéréoscopique suivie d'un recalage 3D/3D. La transformation solution de l'EPCC minimiserait donc la somme des distance aux carrés \tilde{M}_iM_i . L'optimisation serait donc équitable quelle que soit la position des points relativement aux caméras. L'EPCC devrait donc fournir une meilleure précision que l'OSPPC.

disposons donc de 0.04 s pour réaliser le recalage, la transmission des données images et l’affichage des modèles reconstruits. Il est donc crucial de pouvoir évaluer les compromis à faire entre la précision supplémentaire apportée par l’EPPC et le coût temporel que ce gain implique. L’évaluation des performances doit évidemment être reliée aux paramètres initiaux des trois algorithmes qui sont identiques.

Nous précisons ici que cette étude sera menée dans le cas où nous disposons de 2 caméras. En effet, nous verrons dans le chapitre 5 consacré au calibrage que la précision est bien insuffisante avec une seule caméra et que le choix d’utiliser deux caméras semble être le meilleur compromis possible entre la précision supplémentaire apportée par la redondance d’information et l’encombrement que cela représente dans une salle d’opération.

4.2.1 Les conditions d’expérimentation

Les expériences sont réalisées avec deux caméras synthétiques observant le même point avec une différence d’angle de visée de 45° , et calibrées dans le même repère de référence R_f . Nous utilisons 15 points N_i dont les coordonnées 3D sont connues dans R_f , et dont les projections m_i dans les images vidéos sont précisément pré-calculées (en unité pixel). La position spatiale de ces 15 points est représentative de ce que nous aurions sur un abdomen humain moyen. Nous avons décidé de recalibrer 15 points puisqu’il s’agit du nombre vraisemblable de marqueurs que nous collerons sur l’abdomen du patient. Les deux caméras virtuelles sont disposées à une distance de 1000 mm du centre de gravité des N_i et les paramètres intrinsèques choisis correspondent à ceux obtenus dans des cas réels.

Pour simuler le fait que les coordonnées des points 3D sont obtenues dans un autre repère que celui des caméras (en l’occurrence, le repère du scanner), nous appliquons une transformation rigide connue $F(r_x, r_y, r_z, t_x, t_y, t_z)$ aux points N_i et les points résultant de cette transformation sont notés M_i . Cette transformation F est choisie aléatoirement de la façon suivante: la rotation est paramétrée par un vecteur rotation dont la norme (resp. la direction) est choisie uniformément dans le segment $] - 2\pi; 2\pi]$ (resp. dans la sphère unité), et les coordonnées de la translation sont choisies uniformément dans $] - 750mm; 750mm]$. Ce choix aléatoire est représentatif de la réalité puisque les caméras sont positionnées sur le côté du scanner, ce qui signifie que la translation entre les deux repères n’excède pas 1 mètre.

Ensuite, nous modélisons l’erreur de localisation des marqueurs par un bruit gaussien centré d’écart-type 2 affectant les coordonnées $[x, y, z]^T$ des M_i (en mm pour les images scanner 3D) et les coordonnées $[u, v]^T$ des m_i (en pixel pour les images vidéo). Pour respecter les notations établies dans le paragraphe 4.1.1, les points bruités sont notés \tilde{M}_i et \tilde{m}_i . La minimisation des critères aboutit donc sur l’estimation \hat{T} de la transformation $T = F^{-1}$. Les minimisations sur les paramètres de transformation ou de coordonnées 3D (dans le cas de l’EPPC) sont réalisées avec l’algorithme de Powell [Pow64, PFTV92] avec une tolérance de minimisation valant 10^{-4} mm.

4.2.2 Définition des mesures réalisées

Nous évaluons l’erreur de recalage à l’aide de N points de contrôle C_i définis dans le repère de référence R_f , en calculant l’erreur RMS qui sépare les $\hat{T} \star T^{-1} \star C_i$ des C_i :

$$\text{Erreur RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N \| \hat{T} \star C_i - T \star C_i \|^2}$$

On notera que cette erreur RMS correspond en fait à une erreur de recalage de la cible (*Target Registration Error* (TRE)) communément utilisée dans les applications médicales lorsqu'on cherche à estimer les caractéristiques d'un recalage, et non pas à une erreur de recalage de marqueurs (*Fiducial Registration Error* (FRE)), qui refléterait seulement le bruit infectant les données (cf. article de Fitzpatrick [Fit99]).

Chaque tableau présenté dans ce paragraphe indique les valeurs comparatives de temps de calcul et de précision des 3 critères par rapport à un paramètre, les autres étant fixés aux valeurs par défaut détaillées dans le tableau 4.1.

Pour chaque valeur du paramètre qui varie, nous indiquons la moyenne du temps de calcul et l'erreur RMS sur 10000 recalages. Pour caractériser la variabilité de la précision, nous ajoutons l'écart-type des mesures RMS ainsi que sa valeur maximale sur l'ensemble des recalages. A titre indicatif nous précisons aussi dans la ligne *Pourcentage de succès*, le pourcentage du nombre de cas dans lesquels l'EPPC offre une meilleure précision que les deux autres critères.

Nombre de points	Ecart-type du bruit 3D/2D	Tolérance de la minimisation	Angle entre les caméras
15	2 mm/ 2 pixels	10^{-4} mm	45°

TAB. 4.1 – Valeur par défaut des conditions initiales des expériences.

Pour comparer visuellement les critères, nous représentons graphiquement l'erreur relative et le rapport des temps de calcul entre chaque couple de critère. La mesure d'erreur relative n'étant pas symétrique, nous calculons en fait l'exponentielle de la moyenne des logarithmes des rapports de RMS pour obtenir une mesure symétrique de la forme :

$$\text{Erreur relative}(A/B) = [\text{Erreur relative}(B/A)]^{-1} = \exp \left(\frac{1}{10000} \cdot \sum_1^{10000} \log \left(\frac{RMS_{\text{méthode A}}}{RMS_{\text{méthode B}}} \right) \right)$$

Ainsi, une valeur au-dessus de 1 indique que la méthode B est relativement plus précise que la méthode A. Pour alléger la lecture des tableaux, nous utiliserons les abréviations IS, OS, E pour ISPPC, OSPPC et EPPC.

4.3 Temps de calcul et précision

Ce paragraphe étant uniquement consacré à la précision dans le cas où la minimisation a correctement convergé, les algorithmes de recalage sont tous initialisés avec une transformation proche de la solution. Ces expériences ont été réalisées sur un PC 1.1 GHz.

4.3.1 Choix du seuil d'arrêt d_{max} pour la minimisation alternée de l'EPPC

Le tableau 4.2 évalue l'influence des seuils d'arrêt d_{max}^r et d_{max}^t (définis dans la section 4.1.3) sur la précision de l'EPPC. Dans cette section, nous définissons d_{max} de sorte que :

$$d_{max}^r = d_{max} \cdot \Pi \quad \text{et} \quad d_{max}^t = d_{max} \cdot dim$$

où dim correspond aux dimensions de l'objet à recalage. Nous remarquons que les erreurs RMS et relatives ne changent pas significativement pour une valeur de d_{max} en dessous de 10^{-4} , mais deviennent légèrement plus mauvaises pour des valeurs au dessus. D'un point de vue temps de

calcul, la durée de l'optimisation est très sensible à d_{max} : elle passe de 0.050 seconde à 1.811 seconde pour un résultat en précision qui ne varie quasiment pas. Le choix $d_{max} = 10^{-4}$ est donc un très bon compromis entre la précision et le temps de calcul nécessaire à l'optimisation de l'EPPC.

Valeur d'arrêt d_{max} de l'EPPC	10^{-1}	10^{-3}	10^{-4}	10^{-5}	10^{-7}	10^{-10}
Temps de calcul moyen	0.020s	0.029s	0.050s	0.132s	0.569s	1.811s
Erreur RMS moyenne en mm	2.07 ± 0.79	2.00 ± 0.77	1.96 ± 0.74	1.95 ± 0.74	1.94 ± 0.73	1.97 ± 0.75
Erreur RMS maximale en mm	6.21	5.26	5.56	5.68	5.62	5.77

TAB. 4.2 – Temps de calcul et précision en fonction de la valeur d'arrêt d_{max} de l'EPPC. Nous choisissons pour d_{max} la valeur à partir de laquelle la précision reste stable i.e. 10^{-4} .

4.3.2 Influence de la tolérance de minimisation de Powell

Les résultats du tableau 4.3 nous montrent que les temps de calcul et la précision des trois critères ne dépendent pas vraiment de la tolérance de minimisation pour des valeurs inférieures à 10^{-2} . Pour des valeurs plus élevées, les temps de calcul sont plus faible mais la précision se dégrade légèrement pour l'EPPC et violemment pour l'ISPPC et l'OSPPC. La précision restant stable à partir de 10^{-4} , nous avons donc choisi cette valeur pour mener nos expériences de précision. Néanmoins, nous verrons lors de l'étude sur la robustesse (section 4.3.3) que cette valeur ne permet pas de converger souvent vers le minimum global en l'absence d'une initialisation assez proche de la solution. Pour la recherche d'une transformation sans initialisation correct, il sera nécessaire de diminuer ce paramètre afin d'éviter une convergence trop fréquente vers un minimum local.

Afin d'alléger la lecture de ce chapitre, nous avons reporté toutes les mesures comparatives dans l'annexe A.1.

4.3.3 Etude de la robustesse

Cette partie contient les résultats d'un grand nombre d'expérience destinées à cerner la différence des bassins de convergence des trois critères en fonction des données initiales. Nous nous intéressons particulièrement à l'influence de la transformation recherchée et de son initialisation.

Les conditions d'expérimentation Les conditions expérimentales sont globalement les mêmes que dans la section 4.2.1, mais cette-fois-ci, nous n'initialisons plus l'algorithme avec des paramètres de transformation proches de la solution.

Nous calculons les même mesures de performances que la section précédente et nous rajoutons le pourcentage du nombre de cas où chaque critère ne converge pas. Les valeurs de RMS correspondant à une mauvaise convergence ne sont évidemment pas comptabilisées pour le calcul de la moyenne de précision et de temps de calcul. Afin de décider si l'algorithme a correctement convergé pour chaque recalage, nous comparons l'erreur RMS trouvée avec l'erreur RMS maximale calculée dans les expériences de la section précédente qui présentent les

4.3. Temps de calcul et précision

Tolérance de Powell		10^{-1}	10^{-3}	10^{-4}	10^{-8}	10^{-10}	10^{-14}
Temps de calcul moyen	IS	0.003s	0.004s	0.004s	0.008s	0.011s	0.014s
	OS	0.011s	0.017s	0.018s	0.044s	0.049s	0.057s
	E	0.041s	0.045s	0.050s	0.090s	0.112s	0.150s
Erreur RMS moyenne en mm \pm écart-type	IS	4.12 \pm 2.05	2.20 \pm 0.85	2.18 \pm 0.82	2.17 \pm 0.85	2.16 \pm 0.81	2.17 \pm 0.84
	OS	4.05 \pm 1.96	2.19 \pm 0.82	2.17 \pm 0.84	2.18 \pm 0.85	2.16 \pm 0.87	2.17 \pm 0.86
	E	2.01 \pm 0.85	1.96 \pm 0.78	1.95 \pm 0.75	1.95 \pm 0.71	1.94 \pm 0.73	1.95 \pm 0.73
Erreur RMS maximale en mm	IS	13.42	8.58	8.91	8.56	8.51	7.94
	OS	12.56	8.85	8.97	7.67	8.79	6.99
	E	8.31	7.82	5.82	4.59	5.11	4.54
Pourcentage de succès	E vs. IS	87.2	70.8	66.2	68.6	69.1	65.9
	E vs. OS	86.5	69.3	66.5	67.5	69.2	64.6

TAB. 4.3 – Temps de calcul et précision en fonction de la tolérance de minimisation de Powell. Les temps de calcul deviennent plus importants lorsque la tolérance est plus faible. Comme la précision reste stable à partir de 10^{-4} , nous choisissons cette valeur pour les expériences.

mêmes conditions initiales. Il est en effet fort improbable que l'algorithme converge au minimum global et que l'erreur RMS obtenue soit supérieure à l'erreur RMS maximale calculée précédemment étant donné que les conditions initiales sont identiques à celles de la section précédente (exceptée la transformation recherchée T).

4.3.4 Nouveau choix pour la tolérance de Powell

Tolérance de Powell		10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-10}	10^{-14}
Temps de calcul moyen	IS	0.0070s	0.0080s	0.0090s	0.0100s	0.0110s	0.0110s	0.0140s
	OS	0.029s	0.032s	0.037s	0.040s	0.044s	0.049s	0.057s
	E	0.373s	0.398s	0.283s	0.374s	0.302s	0.524s	1.071s
Pourcentage convergences incorrects	IS	41.9%	31.1%	21.7%	18.9%	17.4%	11.9%	8.5%
	OS	35.6%	29.1%	20.4%	17.7%	11.7%	6.5%	4.3%
	E	2.1%	1.6%	1.2%	0.9%	0.8%	0.9%	0.8%

TAB. 4.4 – Influence de la tolérance de minimisation sur la robustesse. Nous choisissons une valeur de 10^{-7} pour l'EPCC puisque le pourcentage n'évalue plus pour des valeurs plus faibles. Pour l'ISPPC et l'OSPPC nous choisissons la valeur qui fournit les meilleurs résultats de robustesse (i.e. 10^{-14}) et qui correspond à la limite de la précision numérique.

Dans la section précédente nous utilisons une tolérance de Powell de 10^{-4} qui était suffisante pour assurer une bonne convergence pour les 3 algorithmes. Nous nous sommes rendu compte que cette valeur n'était pas assez faible pour fournir régulièrement une bonne convergence quand la transformation et/ou l'initialisation étaient choisies au hasard. Le tableau 4.4 présente l'influence de ce paramètre sur la robustesse des 3 critères. Une brève considération des résultats de l'ISPPC et de l'OSPPC montrent clairement que nous devons utiliser une valeur de 10^{-14} si nous voulons limiter au mieux le nombre de convergences erronées. Nous soulignons qu'une valeur encore plus faible n'apporterait rien vu que nous atteignons la limite de la précision numérique (10^{-15}).

Considérant maintenant l'EPPC, nous remarquons immédiatement qu'il est moins sensible à la valeur de tolérance que les 2 autres critères. Deuxièmement, il ne semble pas vraiment utile de choisir une valeur en dessous de 10^{-7} pour atteindre l'efficacité totale de l'EPPC. En conséquence, nous avons décidé de choisir pour les expériences de robustesse une valeur de 10^{-14} (resp. 10^{-7}) pour l'ISPPC/L'OSPPC (resp. l'EPPC).

4.3.5 Les résultats de robustesse

Maintenant que nous avons trouvé une valeur de tolérance équitable pour comparer les 3 critères, nous présentons dans l'annexe A.2 trois types d'expériences pour évaluer leur robustesse. Les 3 premiers tableaux correspondent au cas où T est choisi aléatoirement tandis que l'initialisation est fixée à l'identité. Pour la plupart des problèmes de recalages, cela représente la plus vraisemblable des expériences : quand nous ne disposons d'aucune information préalable sur la transformation, il est raisonnable d'initialiser l'algorithme avec une rotation et une translation nulle. Ensuite, échangeant le rôle de la transformation recherchée et initiale, nous observons que le problème de minimisation ne se comporte pas symétriquement. Cela souligne le fait que la méthode de minimisation semble être biaisée vers l'identité. Le dernier tableau offre des résultats intéressants sur le cas où la transformation recherchée et l'initialisation sont choisies aléatoirement. Plus précisément, cela correspond à l'occurrence d'une très mauvaise initialisation.

4.3.6 Synthèse des résultats

Sur le plan de la précision, l'EPPC fournit toujours une meilleure erreur RMS que les deux autres critères, mais au coût d'un temps de calcul 10 à 30 fois supérieur. Il est très important de noter que le gain en précision est essentiellement dépendant de l'angle entre les caméras et du rapport signal sur bruit entre les données 2D et 3D. Nous voyons dans l'annexe A.1.3 que si l'angle est supérieur à 80° et l'écart-type du bruit 3D est équivalent à celui sur les données 2D, alors l'EPPC n'apporte qu'un faible gain en précision (erreur relative $\simeq 1.03$). En revanche, lorsque l'angle est faible ($< 30^\circ$) ou le rapport σ_{3D}/σ_{2D} proche de 2 (cf. annexe A.1.1), l'erreur relative devient significative ($\simeq 1.15$). Dans un cas où ces deux conditions sont retrouvées simultanément, l'erreur relative dépasse même 1.30. Pour un cas intermédiaire avec moins de contraintes (angle $\simeq 45^\circ$ and $\sigma_{3D}/\sigma_{2D} \simeq 1$), nous pouvons compter sur une erreur relative de 1.10.

Au niveau de la robustesse, l'étude montre que l'EPPC est plus robuste que les deux critères dans quasiment tous les cas. Le pourcentage de convergence incorrect vaut 1% contre 4% pour l'OSPPC et 9% pour l'ISPPC lorsque T est choisie aléatoirement et avec une initialisation fixée à l'identité (cf. annexe A.2.1). Par ailleurs, lorsque T et l'initialisation sont choisis aléatoirement (cf. A.2.3), le pourcentage de convergence incorrect de l'EPPC reste faible (5%) alors que ceux de l'ISPPC et de l'OSPPC deviennent considérable (42% et 15%). Une interprétation alternative de ce résultat est qu'en cas d'initialisation erronée, la convergence de l'EPPC convergera vers la bonne solution dans 95% des cas. L'EPPC n'apporte évidemment pas d'avantage en robustesse lorsqu'une bonne initialisation est disponible.

Ainsi, le choix du critère dépend principalement des contraintes de l'application. Dans un contexte large pour lequel les contraintes temporelles ne sont pas importantes, nous recommandons l'EPPC puisqu'il apporte systématiquement une meilleure erreur RMS. De plus il autorise une initialisation plus large sans compromettre la qualité du recalage. En vue d'un système temps réel comme le notre, nous pouvons espérer un temps d'initialisation de l'ordre

de 3 sec. et un taux de rafraîchissement entre 10 et 25 Hz si les mouvements sont suffisamment lents par rapport au flux vidéo. Il serait éventuellement possible d'augmenter la vitesse de l'algorithme en rafraîchissant l'estimation des "coordonnées 3D exactes" seulement de temps en temps ou à l'aide d'un filtre de Kalman.

4.4 Validation sur des données réelles

Ce paragraphe est consacré à des expériences sur des données réelles afin de comparer les 3 critères exposés précédemment. L'expérience n'étant réellement intéressante que sur des données correspondant à celles que nous aurions en salle d'opération, nous nous servons d'un mannequin en plastique (Body Form®, Limbs & Things Ltd, Bristol, UK) contenant un foie synthétique et sur la peau duquel nous avons collée 25 marqueurs radio-opaques (Multi-Modality Markers type MM3002, IZI Medical Products, Baltimore, MD, USA) couramment utilisés en condition clinique.

Idéalement, la précision de chacun des critères devraient être évaluées en comparant chaque recalage estimé à une mesure étalon. Une telle mesure étalon peut être obtenue en pré-calculant la transformation reliant le repère scanner au repère caméra à l'aide d'un système externe extrêmement précis. Cependant, si un tel système était disponible, nous n'aurions pas développé un algorithme de recalage adapté à nos conditions. Nous devons donc contourner le problème que pose l'absence de mesure étalon avec un autre moyen. En conséquence nous présenterons tout d'abord notre méthodologie inspirée par [PGT98, RPMA01, PBH⁺01] qui permet d'évaluer la *Target Registration Error* (TRE) sans mesure de référence, en calculant l'*erreur résiduelle* de boucles de recalages. Les caméras induisant des distorsions non linéaires sur la position des points 2D dans les images, nous expliquons les modifications apportées aux critères afin de pouvoir en tenir compte. Cette partie sera suivie par les résultats de précision obtenues en condition réelles pour chacun des critères.

4.4.1 Méthodologie de la mesure d'erreur sans mesure étalon

Comme énoncé ci-dessus l'absence d'une mesure étalon nous oblige à nous servir d'une astuce pour évaluer la précision des recalages sur des données réelles. Le principe est de réaliser plusieurs acquisitions scanner du mannequin qui sont recalées entre elles par la méthode décrite ci-après (et qui conduit à une estimation très précise des transformations inter-CT appelé *bronze standard*). Ensuite, plusieurs couples d'images vidéo sont acquises avec des caméras calibrées simultanément afin de pouvoir comparer des recalages 3D/2D *indépendants* du même objet (mais avec des images 2D et 3D qui varient). Une boucle typique est schématisée sur la figure 4.3 : un point de contrôle localisé dans le foie du CT1 est transformé dans le repère caméra CAM1 en utilisant un premier recalage 3D/2D, puis dans le repère scanner CT2 en utilisant un deuxième recalage 3D/2D (les repères CAM1 et CAM2 sont identiques puisque les caméras sont toutes calibrées dans le même repère), et finalement transformé vers CT1 grâce au *bronze standard*. Le fait de se servir de 2 acquisitions CT différentes permet de décorréler les 2 recalages 3D/2D. En effet, si nous recalons les points 2D extraits des images vidéo sur le même ensemble de points 3D (extrait d'une unique acquisition CT), les deux transformations calculées seront affectées de manière similaire par l'erreur de mesure sur les points 3D. Ce qui signifie que la variabilité sur l'extraction des points 3D (ainsi qu'un éventuel biais) sera cachée.

Si toutes les transformations calculées étaient exactes, nous obtiendrions la même position pour notre point de contrôle. Evidemment, comme les transformations ne sont pas parfaites,

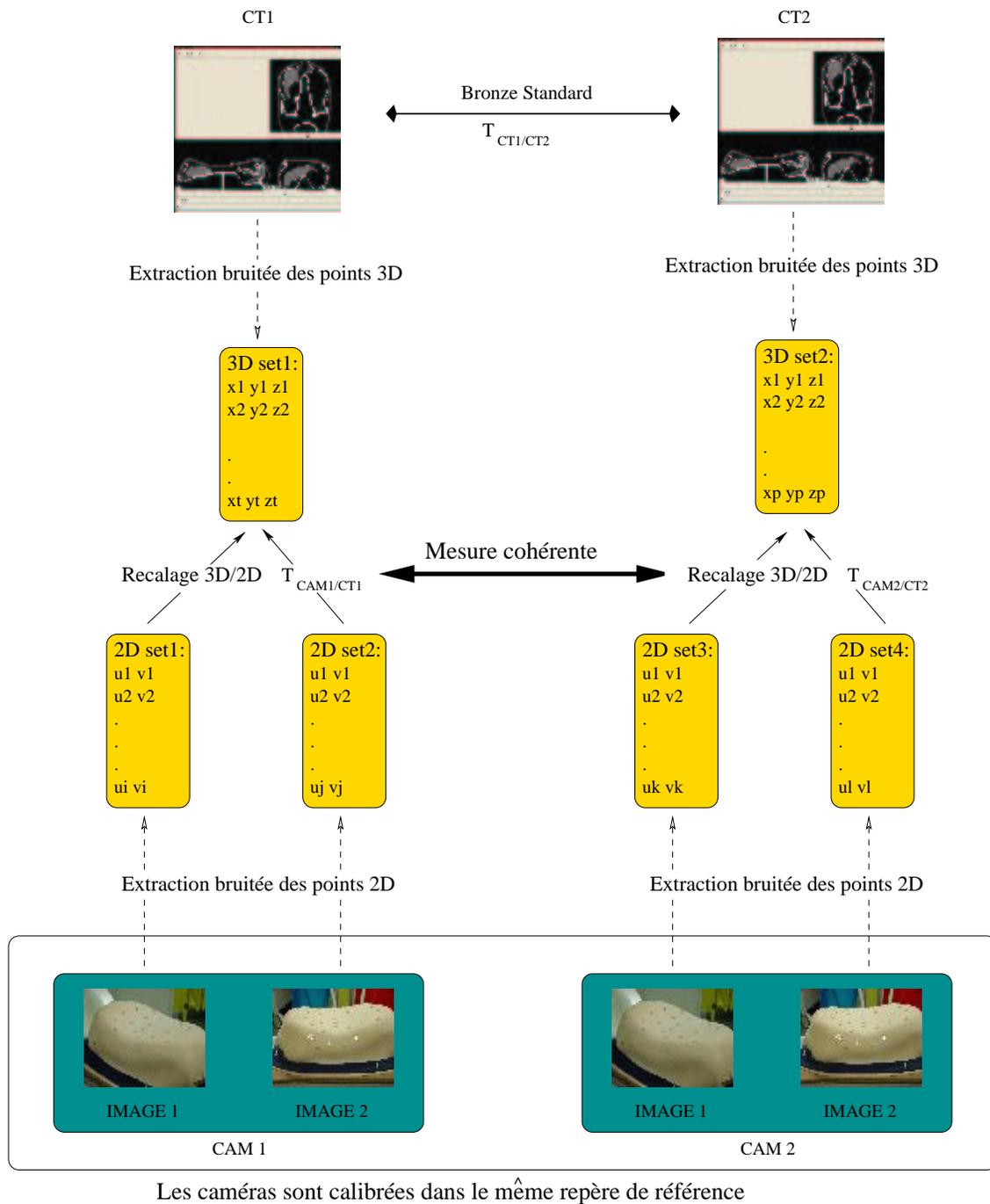


FIG. 4.3 – Boucles de recalage utilisée pour estimer l'erreur résiduelle de recalage : un point de contrôle C défini dans le repère $CT1$ et choisi à l'intérieur du foie du mannequin est transformé dans le repère $CAM1$ en utilisant un premier recalage 3D/2D \hat{T}_1 , puis dans le repère $CT2$ en appliquant le second recalage 3D/2D \hat{T}_2 fourni par l'autre paire de caméras (les repères $CAM1$ et $CAM2$ sont identiques puisque les caméras sont calibrées dans le même repère). Finalement C est transformé dans son repère d'origine $CT1$ grâce au bronze standard calculé entre $CT1$ et $CT2$. Si toutes les transformations étaient parfaites, nous obtiendrions la même position pour C . Evidemment, les transformations n'étant pas parfaites nous mesurons une erreur dont la variance $\sigma^2_{loop} = 2\sigma^2_{CAM/CT} + \sigma^2_{CT/CT}$ correspond à une TRE (Target Registration Error).

nous mesurons une TRE dont la variance vaut $\sigma^2_{loop} = 2\sigma^2_{CAM/CT} + \sigma^2_{CT/CT}$. Cette expérience ne nous fournissant qu'une seule mesure d'erreur, nous devons la répéter avec des données différentes pour obtenir des valeurs statistiquement significatives. Afin de tenir compte des erreurs possibles de calibrage et/ou de biais, il est nécessaire de re-itérer ces expériences avec différents calibrages et positions des caméras, et non pas seulement de bouger le mannequin dans l'espace. Nous pourrions donc calculer σ^2_{loop} et donc estimer la variabilité due aux recalages 3D/2D $\sigma^2_{CAM/CT}$, à condition de connaître $\sigma^2_{CT/CT}$. Dans ce but nous mettons en place la procédure expérimentale suivante basée sur de multiples recalages entre acquisitions CT qui, non seulement fournit une très bonne estimation des transformations entre chaque acquisition, mais qui évalue aussi la précision des recalages fournis.

Création d'un recalage étalon entre les images CT

Notre but est de calculer les $n - 1$ transformations $\bar{T}_{i,i+1}$ qui mettent en relation les n acquisitions CT successives. Il est possible d'obtenir facilement différentes estimations de ces transformations en calculant toutes les transformations possibles $T_{i,j}$ entre les acquisitions CT par différentes méthodes ([RPMA01]). Ensuite, les transformations $\bar{T}_{i,i+1}$ qui satisfont au mieux les mesures sont calculées en minimisant la somme des distances au carré entre les transformations $T_{i,j}$ observées et la combinaison correspondante de la transformation recherchée $\bar{T}_{i,i+1} \circ \bar{T}_{i+1,i+2} \dots \bar{T}_{j-1,j}$. La distance entre les transformations est une estimation robuste de la distance invariante à gauche entre transformations rigides développée dans [PGT98].

L'estimation $\bar{T}_{i,i+1}$ de la transformation parfaite $T_{i,i+1}$ est appelé *bronze standard* puisque le résultat converge vers la transformation exacte $T_{i,i+1}$ quand m et n augmentent. En effet, étant donné une méthode de recalage, la variabilité due au bruit sur les données va décroître si le nombre d'images n augmente, et le recalage calculé va converger vers le recalage parfait à un biais intrinsèque près (s'il y en a un) introduit par la méthode. Maintenant, si l'on utilise plusieurs types de recalage basés sur des méthodologies différentes, le biais intrinsèque va devenir à son tour une variable aléatoire, qui devrait être centré, proche de 0, et moyenné durant la procédure de minimisation. Ainsi, les différents biais des méthodes sont intégrés dans la variabilité de la transformation calculée. Pour atteindre ce but il est donc nécessaire de se servir d'autant de méthodes de recalage indépendantes que possible.

Dans notre protocole, nous avons 5 acquisitions scanner du mannequin dans différentes positions, et 5 méthodes de recalage différentes utilisant des caractéristiques géométriques ou iconiques différentes. Trois de ces méthodes sont basées sur l'intensité: l'algorithme *aladin* [ORPA00] fonctionne sur une stratégie en appariement de région, où les appariements sont déterminés grâce au coefficient de corrélation et la transformation est estimée de manière robuste grâce à une minimisation LTS; l'algorithme *yasmina* [RPMA01] optimise (avec l'algorithme de Powell) le critère SSD ou une estimation robuste de la mesure de corrélation entre les images. Pour les méthodes géométriques, nous avons utilisé le recalage de **lignes de crêtes** décrit dans [PAT00], et l'algorithme **EM-ICP** multi-échelle de [GP02] sur les passages par zéro du laplacien (les images ont été sous-échantillonnées par un facteur 2 pour limiter le nombre de points sur la surface à environ 1.5 million...). Etant donné qu'aucune de ces méthodes n'utilise les coordonnées 3D des marqueurs comme données, nous assurons l'indépendance par rapport au bruit de localisation sur les marqueurs qui corrompt le recalage 3D/2D.

En supplément, nous pouvons utiliser 4 des méthodes citées pour déterminer le *bronze standard*, et s'en servir pour calculer la précision de la cinquième méthode (une espèce de méthode "leave one out"). Cette incertitude est ensuite propagée dans le recalage final du

	σ_{rot} (deg)	σ_{trans} (mm)
Aladin	0.09	0.56
Yasmina SSD	0.02	0.41
Yasmina CR	0.06	0.41
Crest lines	0.04	0.27
EM-ICP	0.08	0.68
Bronze standard	0.01	0.07
Marqueurs	0.15	0.85

TAB. 4.5 – *Evaluation de l’erreur de chacune des méthodes par rapport à un bronze standard obtenu en moyennant les autres méthodes. L’erreur finale obtenue pour le bronze standard incluant toutes les méthodes est très faible (en dessous du dixième de millimètre pour la translation). La variance obtenue sur la rotation est exceptionnellement précise grâce aux structures très rigides dont est constitué le mannequin (coque en plastique dur, socle en bois).*

bronze standard (qui inclut toutes les méthodes) pour estimer sa précision. Dans le tableau 4.5, nous indiquons l’écart-type calculé sur les composantes rotation et translation pour chacune des méthodes, l’incertitude du recalage étalon et l’incertitude du recalage ponctuel 3D des marqueurs obtenu en utilisant une minimisation aux moindres carrés (par rapport au recalage étalon).

Le recalage basé sur les **lignes de crêtes** est le meilleur, suivi par **yasmina**. L’EM-ICP ne donne pas de très bon résultat à cause du sous-échantillonnage des images. La précision du recalage étalon final est très bonne (cela correspond à 0.08 mm TRE sur les points de contrôle). Finalement, nous soulignons que le manque de précision du recalage 3D des marqueurs par rapport aux autres méthodes est principalement due au fait que les marqueurs sont collés sur la “peau” du mannequin, qui est élastique et bougea de 1 à 2 mm entre les acquisitions (cf. figure 4.4), alors que les autres méthodes se sont concentrées sur les structures rigides à l’intérieur du mannequin.

4.4.2 Prise en compte de la distorsion dans les critères

Les distorsions auxquelles sont sujettes nos optiques n’étant pas négligeables, nous devons en tenir compte et les intégrer à notre approche. Pour cela, lors de notre étape de calibrage, nous ajoutons une phase d’estimation des paramètres de distorsions suivant le modèle détaillé dans la section 5.2.3. En possession de ces paramètres, nous sommes à même de calculer la position après distorsion de la projection dans l’espace image d’un point 3D de l’espace. Nous appelons P_{Dist} la nouvelle fonction projective réalisant cette opération.

Afin d’implémenter ce nouveau modèle dans nos critères, nous remplaçons simplement P dans la formule (4.2) et (4.6) par P_{Dist} qui contient en fait les paramètres d’un modèle plus réaliste des caméras. Le tableau 4.6 contient l’erreur moyenne de recalage obtenue pour l’EPPC dans les cas où nous incluons ou pas la distorsion des caméras (Sony®DXC-390P) dans notre critère. Les conditions expérimentales correspondent à celles de la section suivante dans laquelle nous comparons les précisions de l’EPPC, l’ISPPC et de l’OSPPC.

Comme attendu, la précision est meilleure lorsque le modèle choisi est plus complexe. En l’occurrence nous pouvons compter sur un gain en précision de l’ordre de 8% si nous tenons compte de la distorsion.

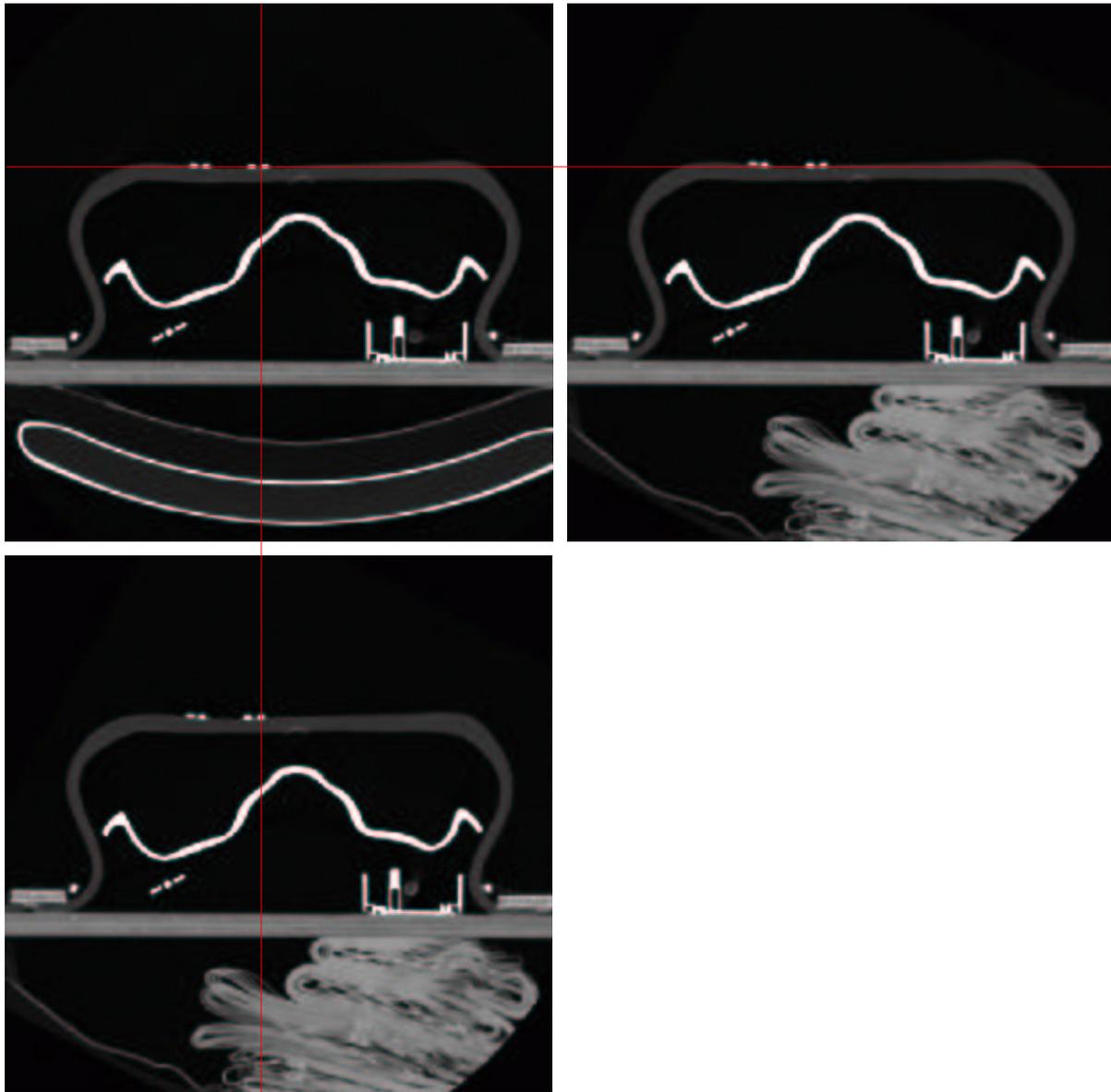


FIG. 4.4 – Mise en évidence du mouvement de la peau du mannequin entre deux acquisitions (dimension des voxels $0.68 \times 0.68 \times 1 \text{ mm}^3$). Nous comparons la position d'un marqueur de l'acquisition 1 (en haut à gauche) avec celle du même marqueur dans l'acquisition 2 (image en haut à droite et en bas à gauche). L'image 2 a été préalablement recalée dans le repère de l'image 1 avec le bronze standard. Les deux droites orthogonales rouges permettent de constater le mouvement du marqueur entre les 2 acquisitions dans un plan de coupe (O, X, Y) de l'image scanner : 2 voxels en X et 1 voxel en Y soit $\sqrt{(2 \cdot 0.68)^2 + 0.68^2} = 1.52 \text{ mm}$ environ.

Angle entre les caméras		60°	40°	10°	
Erreur moyenne de recalage de l'EPPC	avec P	1.84	1.98	2.47	
	avec P_{Dist}	1.69	1.82	2.17	
Erreur relative		P vs. P_{Dist}	1.087	1.086	1.13

TAB. 4.6 – Evaluation de la précision du recalage de l'EPPC en fonction de la prise en compte ou non de la distorsion des deux caméras.

Angle entre les caméras		60°	40°	10°
Temps de calcul moyen	ISPPC	0.023 s	0.025 s	0.027 s
	OSPPC	0.090 s	0.092 s	0.089 s
	EPPC	0.21 s	0.21 s	0.25 s
TRE moyenne en mm	ISPPC	1.80	1.99	2.30
	OSPPC	1.81	2.02	2.29
	EPPC	1.69	1.82	2.17
Erreur relative	IS vs. E	1.065	1.092	1.072
	OS vs. E	1.067	1.098	1.070

TAB. 4.7 – Temps de calcul et TRE moyenne obtenus sur des données réelles provenant de notre mannequin. L'EPPC apporte un gain allant jusqu'à 9.8% pour des temps de calcul de 2 à 10 fois supérieurs. L'erreur de recalage obtenue pour un point cible dans le foie est légèrement inférieure à 2 mm pour un angle supérieur à 40°.

4.4.3 Résultats quantitatifs et qualitatifs sur des données réelles

Après la création du recalage étalon entre nos 5 acquisitions CT du mannequin, nous avons pris 4 images du mannequin dans des positions différentes avec quatre caméras conjointement calibrées. Une fois la série d'images acquise, nous avons modifié les paramètres des caméras (position et focale), les avons recalibré dans un même référentiel et avons pris à nouveau des images du mannequin dans 4 positions différentes. Au total, cette procédure fut réalisée 3 fois, et nous avons donc obtenu 2×12 paires d'images vidéo 2D à recalcer sur 5 acquisitions TDM. Pour rendre l'expérience plus robuste, nous avons réalisé toutes les paires possibles de caméras, ce qui nous mène à 180 boucles de recalages. Environ 25 marqueurs furent localisés dans chaque image ($\sigma_{3D} = 1.5$ mm, $\sigma_{2D} = 2.0$ pixel). Finalement, nous avons obtenu l'évaluation quantitative présentée dans le tableau 4.7.

En définitive, l'EPPC est toujours plus précis que l'ISPPC et que l'OSPPC mais l'erreur relative augmente moins que prévu quand l'angle est petit (10°). Cela pourrait être expliqué par les conditions variées dans lesquelles les mesures ont été réalisées. Nous n'avons pas conservé pour chaque valeur d'angle les mêmes paramètres des caméras (distances focales différentes, nombre de marqueurs observés).

Une autre explication est évidemment le mouvement cohérent mais non-rigide des marqueurs sur la peau grâce auxquels le recalage a été réalisé. Néanmoins, ces mouvements légers simulent de manière réaliste le repositionnement imparfait de la peau et des organes qui est induit par le protocole qui permet de suivre l'évolution de la respiration. Dans ce contexte, l'hypothèse (sur laquelle est basée notre critère) de bruit indépendant sur chacun des points 3D n'est pas forcément satisfaite. Malgré ces différences par rapport aux résultats attendus, nous soulignons que la précision finale obtenue est proche de 2 mm, ce qui est bien en dessous



FIG. 4.5 – Validation visuelle de la précision du recalage sur notre mannequin : le recalage a été réalisé avec 2 caméras. L'image en haut à gauche montre le mannequin avec les marqueurs radio-opaques. L'image en haut à droite est un exemple de vue augmentée dont pourrait bénéficier le praticien, c'est à dire la superposition sur l'image vidéo des reconstructions 3D des marqueurs et des parties internes (structure en plastique et foie synthétique) après recalage. Pour vérifier visuellement la qualité du recalage sur le foie, nous avons enlevé la peau et superposé la reconstruction des marqueurs sur le foie. Les images de gauche au milieu et en bas montrent le mannequin sans sa peau selon les 2 points de vue des caméras utilisées. Sur les 2 images correspondantes à leur droite, nous pouvons observer la qualité de la superposition des marqueurs virtuels sur le foie.

des 5 mm nécessaire à notre application. La figure 4.5 permet d'évaluer visuellement la qualité du recalage sur un cas particulier.

4.5 Conclusion

L'ensemble du travail réalisé dans ce chapitre permet d'aboutir à deux conclusions générales.

4.5.1 L'EPPC, un critère plus précis et plus robuste aux multiples applications

Nous avons développé un critère de recalage optimal en présence de bruit sur les données mesurées 2D et 3D qui généralise les critères de recalages 3D/2D existants (l'ISPPC et l'OSPPC) ainsi que la méthode basée sur une reconstruction stéréoscopique suivie d'un recalage 3D/3D. Notre évaluation des performances sur des données synthétiques montre que l'EPPC est plus précis et plus robuste que l'ISPPC et l'OSPPC, et que le gain obtenu dépend essentiellement de l'angle entre les caméras et du rapport des RSB sur les données 2D et 3D. L'EPPC est nettement supérieur en terme de précision lorsque l'angle entre les caméras est faible (gain supérieur à 20% pour un angle inférieur à 20°) et quand le RSB_{3D} est équivalent ou supérieur au RSB_{2D} . Nous avons ensuite mis au point une méthode originale afin de comparer les différents critères sur des données réelles en l'absence de mesure étalon. Les résultats de précision obtenus, validant la supériorité de l'EPPC, ont montré que nous pouvions compter sur un gain de l'ordre de 10% dans des conditions proches de notre application.

Nous tenons à faire remarquer que les conditions dans lesquelles l'EPPC apporte un bénéfice de précision significatif ne sont pas rares en pratique. Par exemple, les appareillages portables vidéo (présentés dans la section 2.2.2) fonctionnent avec deux caméras disposées à moins de 10 cm l'une de l'autre. Dans le cas d'un recalage réalisé avec l'information provenant de ces deux caméras, la différence entre leur angle de visée serait d'environ 10° pour un champ d'action au niveau des mains de l'utilisateur : le gain pourrait alors dépasser 25%. Un autre exemple est l'ensemble des systèmes de repérages existants (de type Optotrak©, Flashpoint©, Polaris©...) dont les caméras sont distantes d'un maximum de 1 mètre et qui détectent très précisément ($\sigma < 0.1$ pixel) la position de points d'intérêt utilisés ensuite dans de nombreuses applications nécessitant un recalage. Il suffit que le champ d'action soit situé à 2 mètres du dispositif pour que l'angle entre les caméras soit inférieur à 30°. Les configurations pour lesquelles notre critère apporte un réel bénéfice se retrouvent donc dans de multiples applications déjà existantes.

Parallèlement au travail réalisé sur l'EPPC, nous avons effectué pour la première fois (à notre connaissance) une étude comparative de l'ISPPC et de l'OSPPC qui montre sur l'ensemble des tests que leurs précisions sont équivalentes¹⁵, et que l'OSPPC se révèle plus robuste (le pourcentage de mauvaises convergences est à peu près 2 fois inférieur à celui de l'ISPPC).

4.5.2 L'EPPC, un critère adapté à notre application médicale

Le critère que nous avons conçu est adapté aux hypothèses statistiques de notre application puisqu'il ne considère pas implicitement que les données extraites des images scanner sont parfaites, contrairement aux critères déjà existants. Nous avons d'ailleurs pu montrer, sur des données réelles proches de celles que nous devrions retrouver au bloc opératoire, que nous

15. Dans l'article présentant l'OSPPC, Lu [LHM00] réalisa une évaluation des performances de l'OSPPC dans le but de comparer trois méthodes d'optimisation différentes.

pouvions compter sur un gain en précision de l'ordre de 10%. Une interprétation alternative de ce gain est qu'avec 20 marqueurs, l'EPPC atteint la même précision que l'ISPPC avec 24 marqueurs. Etant donné que des marqueurs peuvent être occultés par la main du praticien et que nous pouvons subir des contraintes lors du positionnement des marqueurs sur la peau, ce gain ne doit pas être pris à la légère. Par ailleurs, comme les conditions en salle d'opération n'autorisent pas un positionnement libre des caméras, nous pouvons rencontrer des situations pour lesquelles l'angle entre les caméras est en dessous de 20° , ce qui signifie alors un gain supérieur à 18%.

Cette étude comparative sur des données réelles nous a aussi permis de prouver que nous atteignons en moyenne une erreur de recalage d'environ 2 mm pour un point cible situé dans le foie du mannequin. L'amplitude de cette erreur, additionnée à l'erreur de repositionnement du foie due à la respiration ($\simeq 1$ mm), permet déjà de supposer que nous parviendrons à respecter notre objectif initial de précision (5 mm) étant donné que nous pensons obtenir une erreur de repérage de l'aiguille inférieure à 2 mm.

Chapitre 5

Calibrage des caméras et de l'électrode

Ce chapitre est principalement dévolu aux deux étapes de calibrage nécessaires au fonctionnement du système : le calibrage de nos caméras et celui de l'aiguille manipulée par le praticien. En effet, la définition d'un repère dans lequel acquérir les données du patient nécessite un *calibrage* des caméras par rapport à une mire de référence. Par ailleurs, pour que le chirurgien puisse continuer à manœuvrer l'aiguille vers la cible une fois qu'elle a disparu après introduction sous la peau, nous devons superposer sa position virtuelle dans l'image vidéo. Pour cela, nous devons la lier rigidement à un marqueur repéré dans le référentiel des caméras et *calibrer* sa position par rapport à ce marqueur.

Notre système devant être à la fois précis et utilisable en salle d'opération, nous réalisons préalablement une étude de la précision en fonction du nombre de caméras employées. Nous montrons ainsi que l'emploi de deux caméras représente un bon compromis entre la précision et l'encombrement au bloc opératoire. Ensuite, nous menons une étude bibliographique des nombreuses méthodes de calibrage existantes, parmi lesquelles nous mettons de côté les méthodes directes au profit des méthodes itératives plus précises puisque basées sur des modélisations physiques plus complexes des caméras. Nous expliquons alors pourquoi la méthode multi-vues de Zhang [Zha99] est à même de satisfaire nos exigences de précision et de flexibilité.

Dans la troisième partie de ce chapitre, nous expliquons comment nous calibrons et repérons l'aiguille fixée à un marqueur en nous servant d'une librairie graphique disponible sur internet (ARToolkit). Le fonctionnement de l'ARToolkit étant basé sur des méthodes peu précises, nous en proposons une version améliorée et nous clôturons cette partie par une phase d'évaluation de la précision avec laquelle nous sommes capables de repositionner l'aiguille virtuelle sur l'aiguille réelle. Les résultats montrent que notre version est environ 1.5 fois plus précise que l'ARToolkit et que l'erreur moyenne de recalage du modèle virtuelle est de l'ordre de 0.95 mm pour la pointe et de 0.7° pour l'orientation, ce qui est compatible avec notre contrainte globale de précision.

5.1 Deux caméras : notre consensus précision/encombrement

Dans cette partie, nous testons l'évolution de la précision du système en fonction du nombre de caméra et du bruit sur les acquisitions. N'ayant à notre disposition qu'un nombre limité de caméras, nous avons réalisé des expériences avec des données synthétiques proches de celles que

nous aurons en condition clinique.

5.1.1 Les conditions expérimentales

Les conditions expérimentales concernant la création de la transformation recherchée, des données observées et des points de contrôle sont les mêmes que dans la section 4.2. En revanche, nous ne nous intéressons ici qu'à l'influence du nombre de caméras, de la différence entre leur angle de visée et du rapport signal sur bruit *dans l'image vidéo*. En effet, il n'est plus intéressant de faire varier σ_{3D} car nous ne nous plaçons plus dans un contexte d'évaluation générale de performance. Nous avons donc choisi une valeur proche de nos observations réelles (en l'occurrence $\sigma_{3D} = 2$ mm). L'objet synthétique à recaler possède une forme et des dimensions semblables à celles d'un abdomen humain, et l'évolution de la position des caméras en fonction de l'angle choisi est représentée sur la figure 5.1.

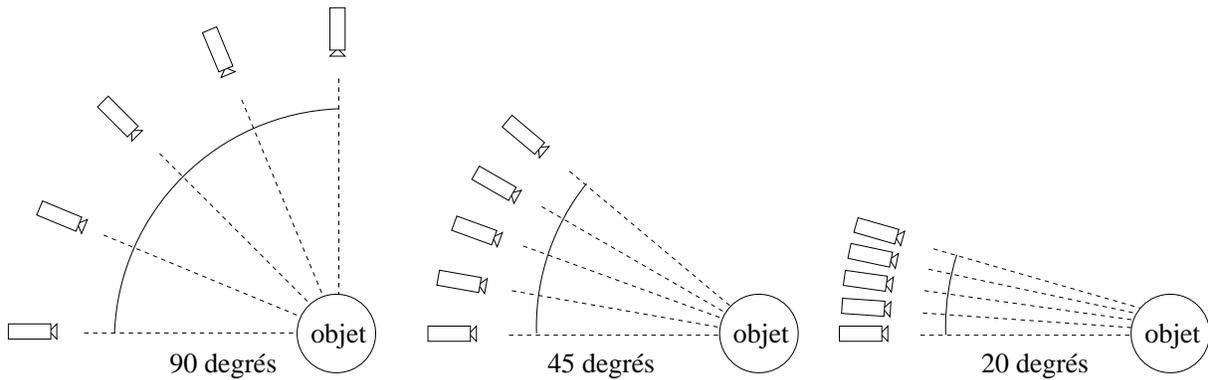


FIG. 5.1 – Exemple pour 5 caméras de l'évolution de leur position respective en fonction de la valeur du paramètre angle entre les caméras. Pour 3 caméras et plus, l'angle entre les caméras représente la différence d'angle de visée entre les deux caméras les plus éloignées.

Remarquons que la configuration choisie pour le positionnement des caméras n'est pas optimale puisque les caméras sont coplanaires. Nous avons fait ce choix car une salle d'opération a des dimensions limitées, il n'est donc pas réaliste d'y apporter plus de trois caméras sur trépied et de les positionner dans une configuration optimale prenant beaucoup d'espace. En revanche, il est raisonnable d'apporter un ensemble de trois caméras ou plus fixées rigidement sur un support rectiligne d'une dimension inférieure à 1m50. Nous avons donc envisagé la configuration la moins encombrante et la plus réaliste possible pour nos expériences.

Notons aussi qu'une configuration non-coplanaire de caméras fixées au plafond pourra être envisagée lorsque notre système de guidage sera validé. Cette étude partielle devra alors être complétée.

5.1.2 Etude de la précision en fonction du nombre de caméras

La courbe à gauche de la figure 5.2 caractérise l'évolution de la précision en fonction du nombre de caméras et de la différence entre leur angle de visée. Nous remarquons tout d'abord que l'erreur RMS diminue quand l'angle augmente, quel que soit le nombre de caméras (sauf pour une caméra évidemment). Ce résultat était attendu puisqu'en augmentant l'angle de visée, nous faisons croître l'information tridimensionnelle recueillie. Par ailleurs, l'évolution des

courbes suggère que le gain en précision devient minime à partir d'un angle de 50° (1.6 mm d'erreur à 90° contre 1.8 mm à 50° pour 2 caméras). On peut aussi noter que pour diviser par 2 la précision obtenue avec 2 caméras, il faut 5 fois plus de caméras. Mais le résultat remarquable est la très forte imprécision obtenue avec 1 caméra (6 mm d'erreur contre 3 mm avec 2 caméras à 20°), le gain est donc très significatif d'une caméra à 2 et beaucoup moins de 2 caméras à 10.

La courbe à droite de la figure 5.2 indique l'erreur RMS en fonction du bruit affectant les données 2D (pour un angle de 45° entre les caméras). Il apparaît à nouveau que la précision obtenue avec une caméra est mauvaise par rapport à celle obtenue avec 2 caméras et plus (il y a environ un rapport 3 entre l'erreur obtenue avec 1 et 2 caméras quel que soit l'amplitude du bruit). En outre nous retrouvons à peu près la même valeur du rapport de précision entre 2 caméras et 10 que dans l'expérience précédente, c'est à dire qu'il faut 5 fois plus de caméras pour diminuer l'erreur d'un facteur 2. A titre indicatif, nous avons ajouté sur les courbes de la figure 5.3 les erreurs RMS maximales obtenues sur l'ensemble des 1000 recalages réalisés. Ces courbes permettent de voir qu'avec une caméra, il est possible de se retrouver avec une erreur de repositionnement spatiale de 2.5 cm, alors qu'avec 2 caméras à 50° , nous risquons 5 mm d'erreur dans le pire des cas, ce qui reste encore acceptable pour notre application.

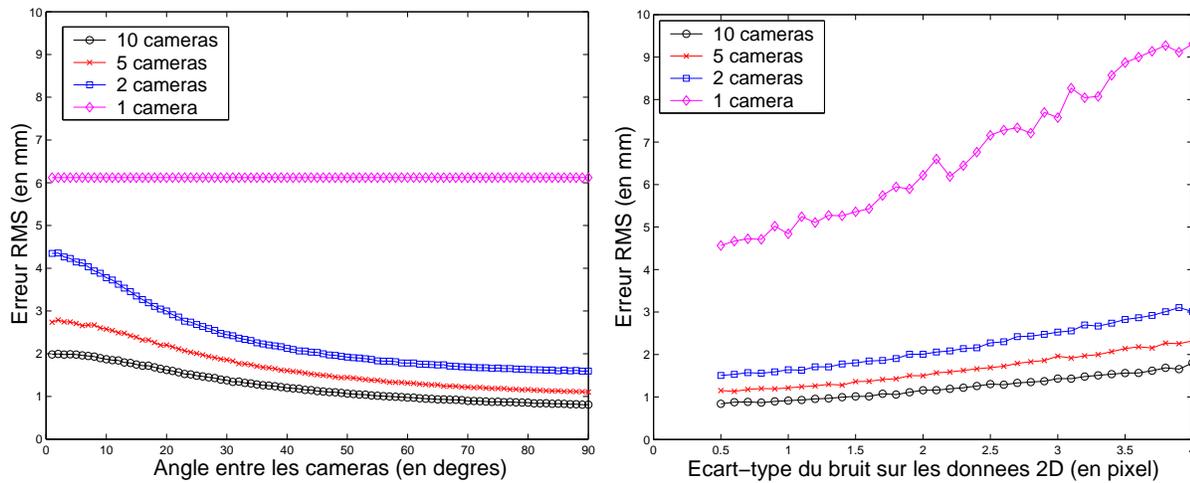


FIG. 5.2 – Évolution de l'erreur RMS en fonction du nombre de caméras et de l'angle (à gauche) ou du RSB sur les données 2D (à droite). Sur la courbe de gauche, l'erreur RMS obtenue avec une caméra n'a pas été recalculée pour chaque valeur d'angle, la courbe est donc constante. A partir de 45° , l'erreur avec deux caméras est 3 fois moins élevée qu'avec une caméra (2 mm contre 6.2 mm). Pour obtenir une précision 2 fois plus élevée qu'avec deux caméras, il faut utiliser 5 fois plus de caméras.

5.1.3 Choix du nombre de caméras

Evidemment, un système composé de 10 caméras serait plus fiable et plus précis qu'un système stéréoscopique. Cependant, nous devrions alors confectionner un support sur lequel fixer les caméras pour que l'encombrement en salle d'opération soit acceptable. De plus, il faudrait résoudre le problème de la gestion synchrone temps réel de 10 caméras, ce qui nécessite des moyens informatiques considérables (multiples cartes d'acquisition et plusieurs PC). Etant donné que nous avons des contraintes matérielles et que l'emploi de deux caméras représente

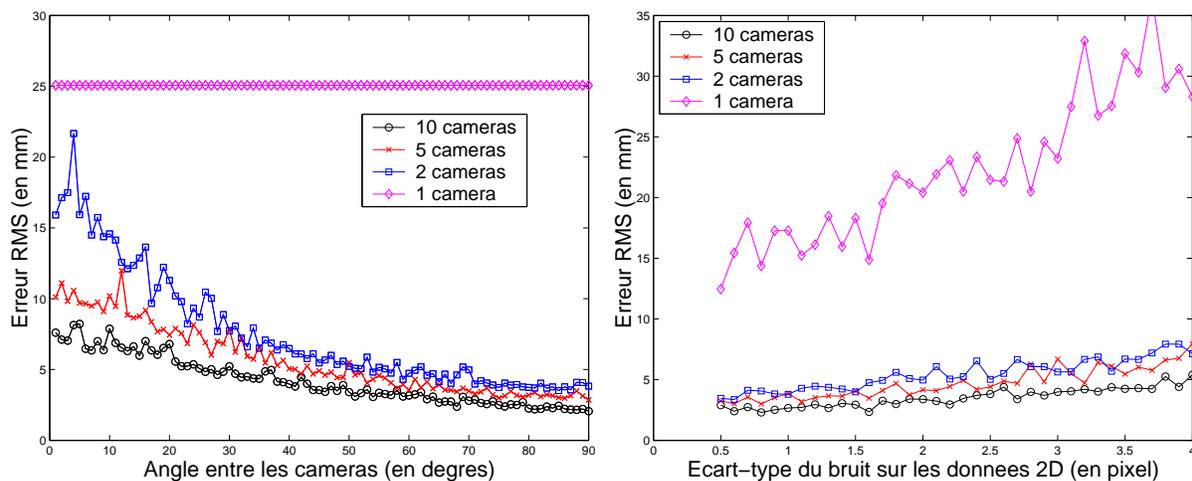


FIG. 5.3 – Évolution de l'erreur RMS maximale en fonction du nombre de caméras et de l'angle (à gauche) ou du RSB sur les données 2D (à droite). Sur la courbe de gauche, l'erreur RMS obtenue avec une caméra n'a pas été recalculée pour chaque valeur d'angle, la courbe est donc constante. On remarque que l'erreur maximale pour une caméra est 5 fois plus élevée qu'avec deux caméras à 50° (2.5 cm contre 5 mm).

un bon compromis entre la précision et l'encombrement, nous avons décidé de ne pas en utiliser plus.

5.2 Choix du calibrage des caméras

Dans cette section, nous abordons le problème du calibrage des caméras, plus connu sous le nom de *problème d'orientation externe* en photogrammétrie. Cette étape est fondamentale pour pouvoir situer le patient par rapport aux caméras visionnant la scène. Ce sujet étant traité depuis plus de 50 ans¹⁶, nous avons à notre disposition un grand nombre de méthodes qui diffèrent par le modèle physique utilisé pour représenter la caméra, le matériel nécessaire (notamment la mire étalon) et les techniques algorithmiques employées afin de calculer les paramètres du modèle physique choisi. Nous souhaitons trouver la méthode qui offre le meilleur compromis entre la précision, la simplicité d'utilisation et la rapidité d'implémentation sur un PC.

Nous montrons tout d'abord que la méthode introduite par Zhang [Zha99] satisfait nos exigences. Nous décrivons ensuite les paramètres physiques du modèle choisi et l'expression résultante qui relie un point de l'espace à son correspondant dans l'image vidéo. Nous terminons la section en montrant comment tirer profit d'un calibrage stéréoscopique pour optimiser la rapport volume de données/paramètre à estimer.

5.2.1 Les méthodes couramment utilisés

Les méthodes de calibrage sont généralement classées dans 2 catégories : les méthodes directes et itératives. Le choix du type de méthode dépend du modèle utilisé afin de paramétrer

16. Szczepanski [Szc58] a présenté en 1958 un état de l'art regroupant déjà plus de 80 solutions à ce problème.

l'optique des caméras employées. Les méthodes directes ne permettent d'obtenir en général que les paramètres du *modèle sténopé* qui ne prend pas en compte les distorsions induites par les imperfections des optiques. La détermination des paramètres se fait généralement par une résolution directe d'un système linéaire [AAK71, FT86, Gan84, HTMS82], par décomposition de la matrice de projection si la résolution est implicite [WM93] ou alors par une considération géométrique du problème [FB81]. Les méthodes itératives sont utilisées lorsqu'on souhaite augmenter la complexité du modèle choisi pour les optiques. La détermination des paramètres se fait alors par une optimisation itérative sur les paramètres. Il faut donc fournir préalablement une initialisation correcte afin d'éviter une mauvaise convergence. En général cette initialisation est obtenue grâce à une méthode directe. Dans ce cadre nous retrouvons diverses variantes différant par le nombre de paramètres utilisées pour la distorsion et la procédure aboutissant aux calculs des paramètres. Tsai [Tsa87] propose une détermination directe et algorithmique d'un certain nombre de paramètres à partir de la contrainte radiale d'alignement qui est suivie d'une optimisation itérative sur les paramètres restants. Weng [WCH92], Heikkila [HS97], Faugeras [Fau93] et Zhang [Zha99] proposent, en libérant la contrainte orthonormale sur la matrice de rotation, 4 méthodes linéaires différentes afin de calculer une initialisation de tous les paramètres (distorsion non comprise). Cette étape est suivie d'une optimisation itérative sur *tous* les paramètres.

5.2.2 Le calibrage de Zhang : précis et fiable

Les méthodes directes sont extrêmement rapides par rapport aux méthodes itératives, néanmoins, il a été clairement montré que la précision qu'il est possible d'atteindre avec un modèle sans distorsion est bien inférieure à celle obtenue avec un modèle considérant les distorsions. Salvi [SAB02] constate que le gain en précision des méthodes itératives sur les méthodes directes est de l'ordre de 200%. Le protocole expérimentale est néanmoins trop peu détaillé. Weng [WCH92] propose une étude plus explicite montrant que le gain obtenu par une estimation itérative est compris entre 50% et 71%. L'écart significatif de gain entre les 2 études citées provient du fait que la localisation des points utilisés dans les images pour calibrer avec la méthode directe n'est pas la même d'une étude à l'autre. Weng choisit ces points spécifiquement au centre de l'image, ce qui a pour effet de limiter les erreurs induites par la distorsion, alors que Salvi n'effectue aucun tri particulier sur la position des données extraites des images.

On peut se demander ensuite quelles sont les différences de précision entre les différentes méthodes itératives. Il semblerait que pour un jeu de données étalon identique, il n'y ait quasiment aucune différence de précision entre elles [SAB02]. Dans cette étude, les méthodes de Tsai optimisé, Weng et Faugeras optimisé ne présente qu'une différence de l'ordre de 1% en terme de précision, malgré une méthodologie différente et un modèle physique parfois moins compliqué pour estimer les paramètres du calibrage (Tsai considère uniquement la distorsion radiale). En revanche, Weng [WCH92] montre qu'une modélisation radiale et tangentielle de la distorsion apporte un gain en précision d'environ 25% sur un modèle ne tenant compte que de la distorsion radiale. A nouveau, la différence de conclusion s'explique vraisemblablement par la différence du protocole de mesure d'erreur. Nous ne pouvons donc pas conclure à une différence significative entre les précisions obtenues par ces 2 modèles de distorsion. Étant donné que notre étape de calibrage n'est pas soumise à des contraintes de temps particulières (les 2 caméras en salle d'opération fixées au plafond ou posées sur un trépied peuvent être calibrées avant l'intervention), nous estimons que nous avons tout à gagner à choisir un modèle précis

quitte à ce que l'étape d'optimisation prenne plusieurs secondes. Nous nous orientons donc vers une méthode de calibrage itérative.

Parmi la multitude de méthodes non-linéaires, un très petit nombre est disponible via internet ou sur demande auprès des auteurs. Notre recherche nous a permis de trouver le code (en C ou en Matlab) des 3 méthodes suivantes : Faugeras (linéaire et non-linéaire), Tsai et Zhang. Des 3 méthodes trouvées, nous avons mis de côté l'implémentation proposé par Tsai puisque le code ne contient pas les outils nécessaires à l'extraction des points d'intérêt dans les images, ainsi qu'au raffinement sub-pixelique de leur détection. Les méthodes de Faugeras et de Zhang diffèrent principalement par la méthode d'acquisition de données étalon définissant le repère dans lequel la caméra va être positionnée. La méthode de Faugeras est basée sur l'acquisition vidéo d'une mire non-coplanaire (cf. figure 5.4). L'obtention d'une telle mire est difficile et nécessite soit un usinage coûteux, soit les services d'une entreprise de métrologie (ce qui est coûteux aussi) afin de connaître très précisément la position relative des points d'intérêts. Zhang propose une méthode utilisant pour mire un damier aux dimensions connues, obtenue à partir d'une imprimante traditionnelle. Cette méthode est disponible sur internet sous un environnement en C ou Matlab, et permet de calibrer une caméra sans objet étalon difficile à se procurer. Nous avons donc choisi de nous servir de cette méthode afin de calibrer nos caméras.

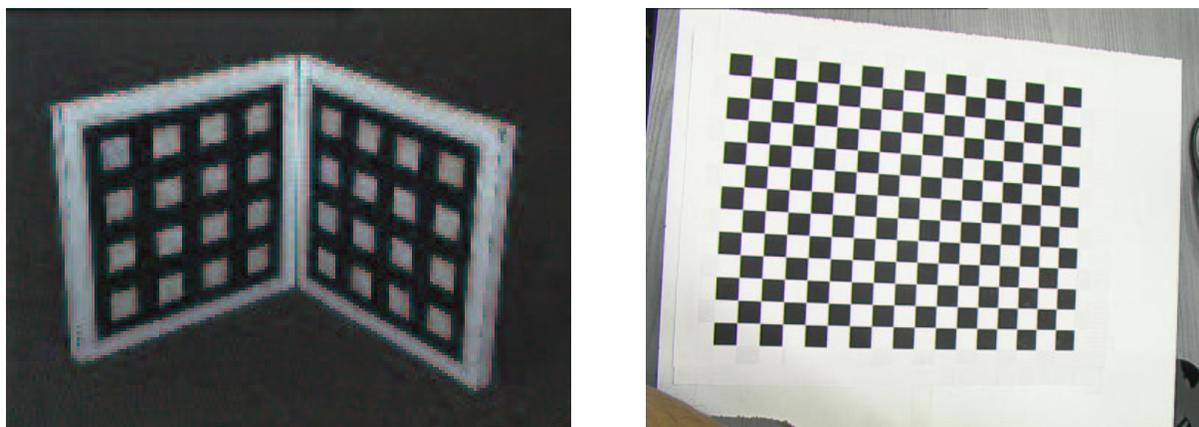


FIG. 5.4 – Exemple de mires de calibrage utilisées par la méthode de Faugeras (à gauche) et par la méthode de Zhang (à droite).

5.2.3 Le modèle physique choisi

La mise en place d'un modèle géométrique pour calculer la projection d'un point de l'espace sur la matrice CCD d'une caméra correspond à déterminer la fonction F qui permet de calculer, à partir des coordonnées 3D du point défini dans le repère étalon des caméras, son emplacement 2D en pixel dans l'image vidéo. Nous détaillons ci-dessous la succession des étapes menant à la projection d'un point de l'espace et clarifions un par un les paramètres introduits. Cette explication est illustrée par la figure 5.5.

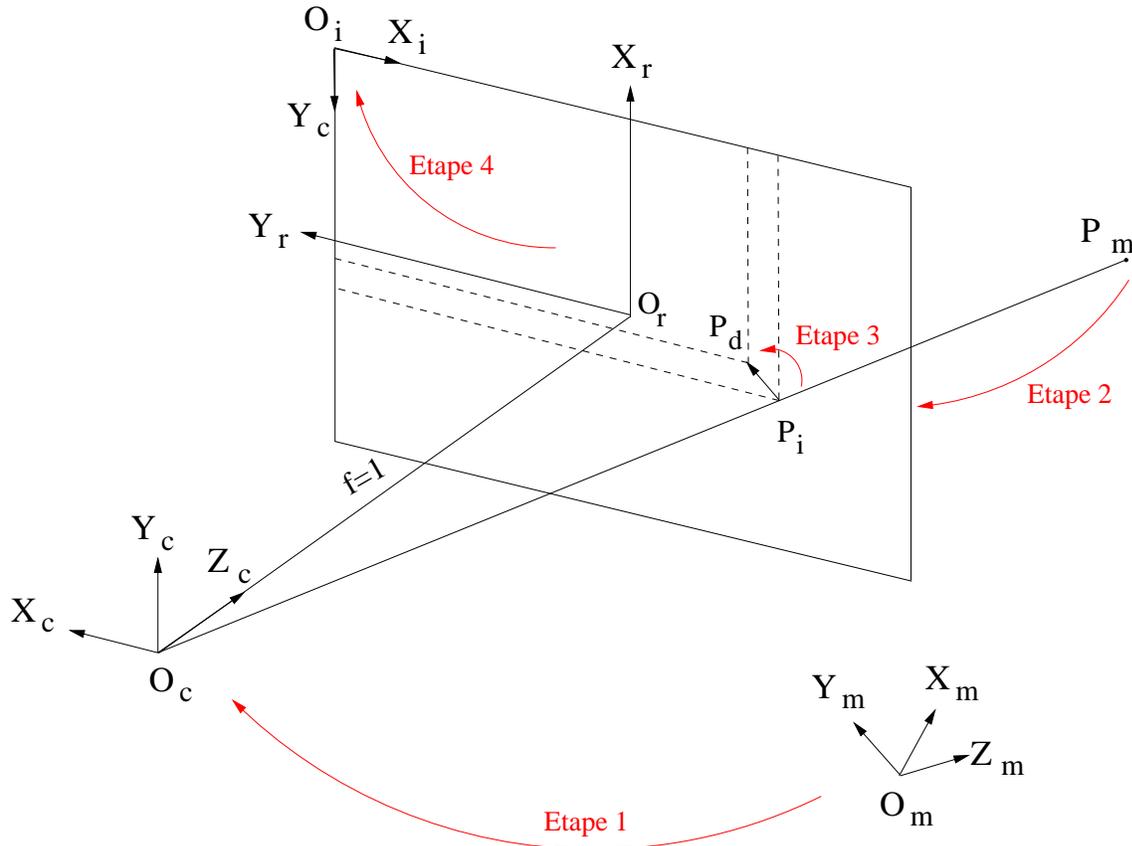


FIG. 5.5 – Illustration des étapes menant à la projection d'un point de l'espace sur le capteur d'une caméra. Le point P_m défini dans le repère étalon $R_m(O_m, X_m, Y_m, Z_m)$ est d'abord passé dans le repère $R_c(O_c, X_c, Y_c, Z_c)$ lié au centre optique de la caméra (étape 1), puis projeté sur le plan focal image en P_i (étape 2). Les effets de distorsion étant appliqués (étape 3) à P_i , le point résultant P_d subit un dernier changement de repère correspondant au passage du repère réel au repère défini en pixel du capteur.

Détermination de la position de la caméra Soit un point P_m défini dans le repère monde $R_m(O_m, X_m, Y_m, Z_m)$ (qui correspond concrètement au repère défini par notre mire de calibrage). La première étape consiste à savoir où se situe le centre optique et l'axe de visée de la caméra par rapport à R_m . Cela se traduit concrètement par la connaissance de la matrice de passage entre le repère R_m et un repère $R_c(O_c, X_c, Y_c, Z_c)$ lié au centre optique (classiquement l'axe des Z correspond à l'axe de visée et les axes des X et des Y sont parallèles aux côtés du capteur de la caméra). Les 6 paramètres (3 pour la rotation et 3 pour la translation) impliqués sont appelés communément paramètres extrinsèques de la caméra car ils ne dépendent pas de la structure interne de la caméra.

Projection sur le plan focal image La seconde étape correspond à la projection du point $P[X_P, Y_P, Z_P]^T$ défini dans R_C sur le plan focal image de la caméra. Considérant sans perte

de généralité que la distance focale f est égale à un¹⁷, on obtient dans le plan focal image la projection P_i de P_m :

$$P_i : \left[\frac{X_P}{Z_P}, \frac{Y_P}{Z_P}, 1 \right]^\top$$

Application de la distorsion La troisième étape correspond à l'application des effets de distorsion. En pratique, les lentilles subissent principalement 2 types de distorsion : radiale et tangentielle. La distorsion radiale est due à la courbure de la lentille et va induire une distorsion symétrique par rapport à son axe optique qui va dilater ou contracter l'image idéale autour de l'axe (cf. figure 5.6 gauche). La distorsion tangentielle est générée par l'alignement imparfait des centres optiques des lentilles constituant l'objectif et va déplacer l'image tangentiellement à l'axe optique (cf. figure 5.6 droite). En général, la distorsion radiale prédomine sur la tangentielle (au moins d'un facteur 10).

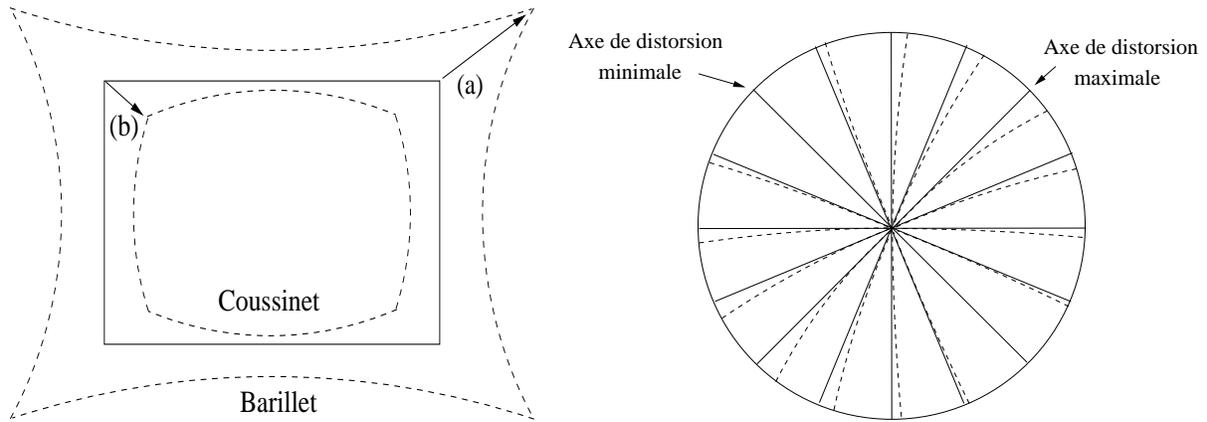


FIG. 5.6 – A gauche : effets dus à la distorsion radiale. L'image originale (trait plein) subit une contraction (déformation de type coussinet (a)) ou une dilatation (déformation de type barillet(b)) symétrique autour de l'axe optique. A droite : effets dus à la distorsion tangentielle. En trait plein : image originale sans distorsion. En pointillé : avec distorsion.

Dans le cadre du modèle sténopé, ces effets sont totalement négligés. Les autres modèles de distorsion [Bro71, Fai75, oP80] proposent des paramétrisations différentes des mêmes phénomènes physiques. Nous avons choisi le modèle employé par Heikkilä ([HS97]) qui modélise la distorsion radiale (au second ordre) et la tangentielle (au second ordre aussi). Notant $P_i = [x_i, y_i, 1]^\top$ et $r^2 = x_i^2 + y_i^2$, P_i devient $P_d = [x_d, y_d, 1]^\top$ après distorsion :

$$x_d = \underbrace{(1 + k_1 \cdot r^2 + k_2 \cdot r^4)}_{\text{distorsion radiale}} \cdot x_i + \underbrace{2 \cdot t_1 \cdot x_i \cdot y_i + t_2 \cdot (r^2 + 2 \cdot x_i^2)}_{\text{distorsion tangentielle}}$$

$$y_d = \underbrace{(1 + k_1 \cdot r^2 + k_2 \cdot r^4)}_{\text{distorsion radiale}} \cdot y_i + \underbrace{2 \cdot t_2 \cdot x_i \cdot y_i + t_1 \cdot (r^2 + 2 \cdot y_i^2)}_{\text{distorsion tangentielle}},$$

17. La focale est un paramètre de similitude qui peut être pris en compte dans les paramètres intrinsèques [Fau93, WCH92].

où k_1 et k_2 (resp. t_1 et t_2) sont les paramètres modélisant la distorsion radiale (resp. tangentielle).

Nous pourrions considérer les distorsions à des ordres supérieurs, néanmoins, l'expérience montre que cela n'apporte pas de gain en précision (l'erreur d'extraction sur les données étant largement supérieure aux corrections apportées au troisième ordre). En plus cela peut conduire à une convergence erronée lors de l'optimisation itérative si le nombre de données n'est pas suffisant. Nous avons en effet constaté sur un faible jeu de données un calibrage imputant principalement la distorsion sur la composante tangentielle et non radiale, ce qui n'est pas réaliste.

Passage du repère réel au repère pixel La dernière étape consiste à passer du monde réel au monde pixel. Il faut passer de la métrique réelle à la métrique en pixel du capteur. Les capteurs n'étant pas parfaitement carrés, 2 paramètres multiplicatif α_u et α_v sont introduits. Par ailleurs, l'axe optique de l'objectif n'est pas forcément aligné avec le centre du capteur (en fait ce n'est jamais le cas). Deux paramètres sont ainsi ajoutés afin de modéliser la translation de l'origine du repère. Mathématiquement, cette étape se concrétise par la multiplication du point $P_d[x_d, y_d, 1]^T$ avec la matrice de paramètres intrinsèques $I(\alpha_u, \alpha_v, u_0, v_0)$ et aboutit aux coordonnées en pixel du point P_{pix} dans l'image vidéo :

$$\underbrace{\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}}_{P_{pix}} = \underbrace{\begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_I \cdot \underbrace{\begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}}_{P_d}$$

Remarquons au passage que nous avons supposé que les capteurs étaient parfaitement rectangulaires. Cette hypothèse, autrefois peu réaliste, est désormais justifiée par le degré de précision atteint lors de la conception des capteurs. Au final nous avons un total de 6 paramètres extrinsèques et 8 intrinsèques à estimer.

5.2.4 Notre calibrage stéréoscopique

Dans cette section, nous commençons par présenter les principes du calibrage de Zhang appliqué à une caméra. Ensuite, nous expliquons comment nous adaptons cette méthode à notre système stéréoscopique afin d'optimiser le rapport volume de données/nombre de paramètres à estimer

La méthode de Zhang : une méthode multi-vues

L'idée intéressante consiste en la prise de vues successive d'une même mire dans des positions et orientations différentes, sans connaissance a priori sur les orientations relatives entre chaque prise de vue. Soient M le nombre d'images acquises de la mire et N le nombre de points d'intérêt qu'elle contient. Pour chaque prise de vue il y a 6 paramètres extrinsèques à estimer, en revanche, les paramètres intrinsèques ne changeant pas d'une prise à l'autre, nous disposons de $M \cdot N$ données pour estimer $6 \cdot M + 8$ paramètres, ce qui donne un ratio paramètres/données de $\frac{6 \cdot M + 8}{M \cdot N} = 6/N + \frac{8}{M \cdot N}$ paramètres. En conséquence, plus le nombre de vues M sera grand, plus le ratio baissera. Plus simplement, cette méthode permet de baisser virtuellement le nombre de paramètres, pour un minimum limite de 6. Par exemple, considérons la méthode mono-vue de Faugeras avec la mire non-coplanaire (cf. figure 5.4 à gauche). Il y a 128 points pour 14

paramètres, ce qui donne un ratio de $14/128 = 0.109$. Avec la mire plane que nous utilisons (cf. figure 5.4 à droite) et avec 10 prises de vue, nous obtenons un ratio de $6/192 + \frac{8}{10 \cdot 192} = 0.0354$, ce qui signifie qu'avec une méthode mono-vue nous avons 9.17 données pour estimer un paramètre et 28.2 avec une méthode multi-vues.

Résumé de la méthode d'estimation des paramètres :

En premier lieu, une première estimation des paramètres est nécessaire afin d'initialiser l'optimisation itérative. Cette initialisation est calculée tout d'abord par une résolution implicite des paramètres (calcul d'une homographie) suivie d'une résolution explicite par décomposition de la matrice d'homographie. En deuxième lieu, on va minimiser sur les paramètres intrinsèques et extrinsèques la somme des distances aux carrés entre les points du damier observés dans l'image vidéo \tilde{m}_i et la reprojection 2D des points du damier M_i (dont les positions 3D sont connues), i.e. :

$$\sum_{l=1}^{l=M} \sum_{i=1}^{i=N} \| \tilde{m}_i^l - m_i^l(I, k_1, k_2, t_1, t_2, R^l, t^l, M_i) \|^2$$

où I correspond à la matrice des paramètres intrinsèques, M au nombre de prises de vue et N au nombre de points de la mire.

Ce critère, dérivé d'un maximum de vraisemblance, est optimale à la condition que l'extraction des données vidéo 2D soit perturbée par un bruit gaussien centré, de même amplitude sur tous les points d'intérêt et indépendant d'un point à l'autre. En pratique, même si ce n'est pas tout à fait le cas (l'extraction est légèrement biaisée puisqu'elle est réalisée sur des images sujettes aux effets de distorsion), l'approximation reste très bonne.

Application de la méthode de Zhang à notre système stéréoscopique

Pour que nos deux caméras soient calibrées dans le même référentiel, il suffit que, sur la dizaine d'images de la mire acquises par chaque caméra, une seule corresponde à une position identique de la mire par rapport aux deux caméras. Cependant, pour optimiser le rapport volume de données/nombre de paramètres à estimer, il est préférable de réaliser une acquisition simultanée avec les deux caméras pour chaque position de la mire. Appelant $(R_{C_1 \rightarrow C_2}, t_{C_1 \rightarrow C_2})$ les paramètres de la transformation rigide relative entre les deux caméras, nous pouvons estimer les paramètres des caméras en minimisant la somme S_{reproj} des erreurs de reprojection pour chaque caméra :

$$\begin{aligned} S_{reproj} &= \sum_{l=1}^{l=M} \sum_{i=1}^{i=N} \| \tilde{m}_{i,l}^{C_1} - m_{i,l}^{C_1}(I^{C_1}, k_1^{C_1}, k_2^{C_1}, t_1^{C_1}, t_2^{C_1}, R^l, t^l, M_i) \|^2 \\ &+ \sum_{l=1}^{l=M} \sum_{i=1}^{i=N} \| \tilde{m}_{i,l}^{C_2} - m_{i,l}^{C_2}(I^{C_2}, k_1^{C_2}, k_2^{C_2}, t_1^{C_2}, t_2^{C_2}, R^l \star R_{C_1 \rightarrow C_2}, t^l + t_{C_1 \rightarrow C_2}, M_i) \|^2 \end{aligned}$$

où $\tilde{m}_{i,l}^{C_1}$ (resp. $\tilde{m}_{i,l}^{C_2}$) correspond à l'observation du point i de la mire dans l'image l vue par la caméra C_1 (resp. C_2). Notons que ce critère est semblable au critère ISPPC (cf. section 4.1.2) à la différence que la minimisation de S_{reproj} permet d'estimer plusieurs transformations rigides (et non pas une seule) en plus des paramètres intrinsèques des caméras.

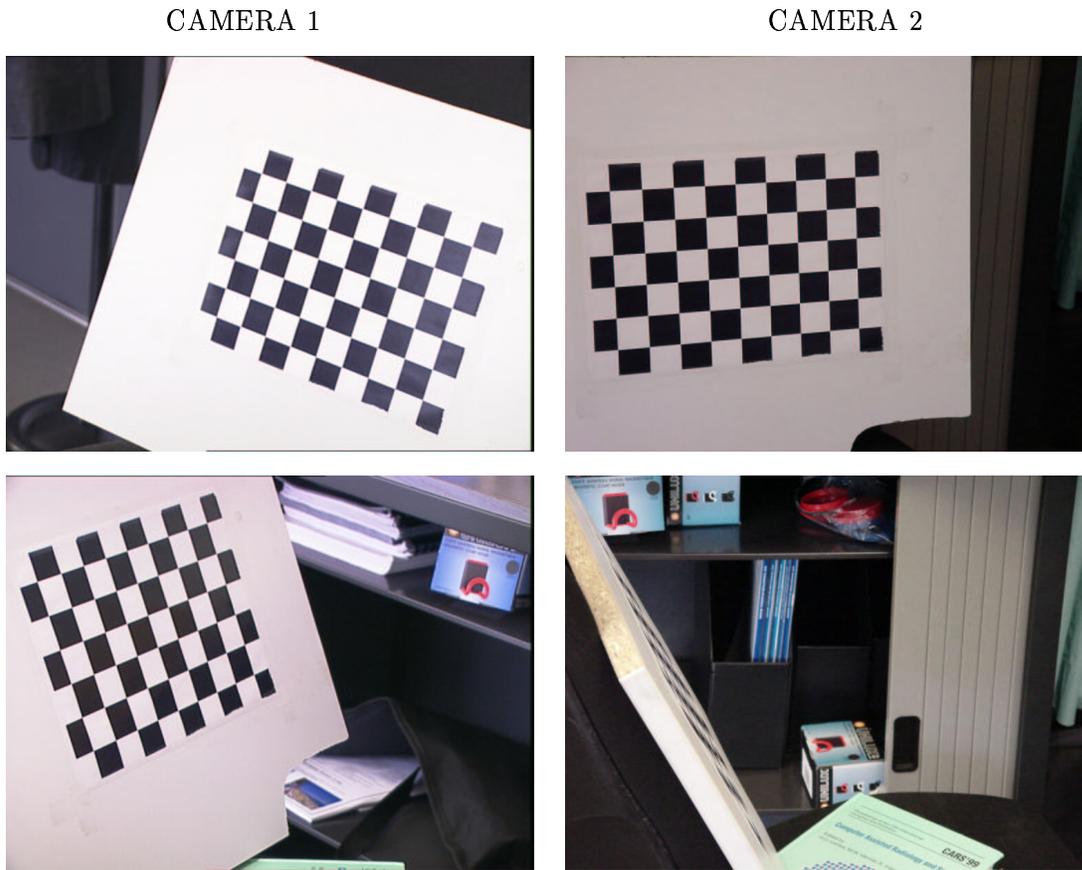


FIG. 5.7 – A gauche : deux images de la mire vue par la caméra C_1 . A droite : les deux vues correspondantes pour la caméra C_2 . Dans les images du bas, nous avons changé l'orientation de la mire d'environ 70° par rapport à sa position dans les images en haut. Nous voyons alors que la mire n'est plus visible dans la caméra 2 : l'angle entre les caméras étant de 40° , les variations d'orientation de la mire sont limitées dans le cas d'un calibrage stéréoscopique. Les paramètres intrinsèques ne seront pas estimés de manière optimale.

Dans ce cas de figure, pour M images et N points sur la mire, le rapport paramètres/données vaut $\frac{6 \cdot M + 8 + 8 + 6}{2 \cdot M \cdot N} = \frac{3 \cdot M + 11}{M \cdot N}$ au lieu de $\frac{6 \cdot M + 8}{M \cdot N}$ dans le cas du calibrage indépendant d'une caméra. Ainsi, dès que le nombre d'images M acquises simultanément dépasse 2, le calibrage stéréoscopique fournit un meilleur rapport.

Néanmoins, nous remarquons que dans notre application, nos deux caméras seront disposées de telle sorte que la différence entre leur angle de visée soit au dessus de 40° . Il ne sera donc pas aisé d'acquérir simultanément des images de la mire dans des conditions d'orientation vraiment différentes (cf. figure 5.7). Or la qualité de l'estimation des paramètres intrinsèques (distorsion comprise) dépend de l'importance des variations d'orientation entre chaque position de la mire (cf. [Zha99]). Un calibrage stéréoscopique n'est peut-être pas le meilleur choix du fait de cette considération.

Pour limiter ce défaut, nous réalisons notre calibrage en 2 étapes :

- nous calibrons tout d'abord indépendamment nos caméras en optimisant le critère de

la section précédente. L'orientation de la mire n'étant plus contrainte, le calibrage des paramètres intrinsèques est optimal.

- nous réalisons ensuite un calibrage stéréoscopique en optimisant le critère S_{Reproj} , les paramètres intrinsèques étant fixés aux valeurs estimées à l'étape précédente.

Finalement, notre implémentation nous permet d'obtenir une estimation précise des paramètres de calibrage en moins de 3 minutes.

5.3 Repérage de l'aiguille dans le repère des caméras

Dans le chapitre précédent nous avons montré que nous étions capable de recalibrer les structures internes de notre mannequin avec une précision moyenne de l'ordre de 2 mm. Étant donné que l'aiguille n'est plus visible une fois que le praticien l'introduit sous la peau, nous devons aussi superposer son modèle virtuel dans l'image vidéo. Pour cela, il est nécessaire de repérer sa position dans le repère des caméras. Ce problème de repérage est très fréquent (tant dans le domaine médical [Ra01, Sa02, SLG⁺96, LLB02, SLY⁺03] que le domaine industriel [BNR03, FAMS03, WS02, EW03, BK02]) et de nombreux systèmes industriels (Optotrak©, Polaris©, Flashpoint© et cie.) le résolvent avec une grande précision. Néanmoins, l'emploi d'un tel système de repérage est encombrant puisqu'il faut ajouter ces dispositifs d'acquisition spécifiques dans la salle d'opération. De plus, nous devrions mettre en place une méthode afin de calibrer rapidement et précisément le système de repérage par rapport au système stéréoscopique. Ce calibrage induirait une erreur supplémentaire qui nécessiterait une phase d'évaluation et de validation.

Nous pourrions alternativement ne pas nous servir de caméras vidéo et repositionner le modèle scanner dans le repère du système de repérage en localisant les marqueurs sur la peau à l'aide d'une sonde et en effectuant un recalage 3D/3D. Cependant, l'absence d'acquisition vidéo¹⁸ par le système à deux conséquences. D'une part nous perdrons la possibilité de fournir au praticien une image réelle du patient sur laquelle est superposé le modèle virtuel, d'autre part nous devrions systématiquement localiser manuellement les marqueurs dans le repère du système alors qu'un traitement des images vidéo du patient permet d'extraire automatiquement la position des marqueurs radio-opaques (cf. chapitre 6). De fait, nous pensons qu'il est plus commode et plus judicieux de réaliser nous-mêmes un traitement d'image sur nos acquisitions vidéo afin de localiser l'aiguille directement dans le repère de référence de nos caméras.

La méthodologie classique est de lier rigidement l'aiguille à un marqueur dont la géométrie spatiale est connue. Le marqueur doit pouvoir être détecté de telle sorte que nous puissions déterminer sa position en 3D dans le repère de référence. Ainsi, si la position de l'aiguille par rapport au marqueur est connue, il devient alors possible de la connaître dans le repère de référence et d'afficher sa position virtuelle dans les images vidéo. Par suite, nous devons tout d'abord choisir un marqueur aisément détectable dans une image vidéo que nous fixerons à l'aiguille.

Dans cette section, nous commençons par présenter concisément la librairie ARToolkit (Augmented Reality Toolkit), qui est une librairie graphique dont le but est la localisation de marqueur de forme carré dans une image vidéo. Cette librairie n'étant pas conçue pour des applications dans lesquelles la précision est importante, nous expliquons ensuite les améliorations que nous lui avons apporté afin d'augmenter la précision. Dans la partie suivante,

18. Les systèmes ne sont pas tous basés sur une analyse d'image acquise par des caméras, et lorsque c'est la cas, le flux vidéo n'est pas rendu accessible par les constructeurs.

nous détaillerons la méthode de calibrage permettant de déterminer la position de l'aiguille par rapport au marqueur, puis nous terminerons avec une étude de précision afin d'évaluer l'amélioration apportée ainsi que l'erreur de recalage entre l'aiguille réelle et l'aiguille virtuelle que nous superposons dans l'image vidéo.

5.3.1 Amélioration de la librairie standard ARToolkit

La librairie Augmented Reality Toolkit disponible sur internet [Hirdu] permet de détecter en temps réel dans une image vidéo les coins de plusieurs marqueurs carrés différenciés par leur motif central (cf. figure 5.8). Ce motif non-isotrope sert à orienter le carré, définissant ainsi un repère orienté sur ses coins. La géométrie spatiale du carré étant connu, l'ARToolkit fournit une estimation de la transformation rigide liant le repère du marqueur au repère de référence en minimisant le critère aux moindres carrés classique (ISPPC) :

$$\sum_{i=1}^{i=4} \| P(T \star M_i) - m_i \|^2$$

où m_i correspond aux coins détectés dans l'image vidéo, et M_i à leurs coordonnées 3D définis dans le repère du marqueur.



FIG. 5.8 – Exemple de marqueurs détectés dans des images vidéo par l'ARToolkit. Les points d'intérêts sont les 4 coins du carré et le motif central sert à les identifier quand plusieurs marqueurs sont présents dans la même image.

Connaissant l'orientation et la position de la pointe de l'aiguille dans le repère du marqueur, la connaissance de T nous permet donc de savoir où se situe l'aiguille dans le repère de référence. Cependant, cette librairie ayant été créée dans un but plutôt ludique [BRF01, RS01, PT02], la précision fournie pour le calcul de T est loin d'être optimale [ZFN03]. Étant donné que notre contrainte de précision est forte, nous avons jugé important de modifier quelques fonctionnalités.

Détection sub-pixellique des coins

L'ARToolkit détecte les coins dans une image sur laquelle a été appliqué le filtre de Sobel suivi d'un seuillage. L'image résultante étant binaire, cette détection est de l'ordre du pixel. De plus, elle est extrêmement sensible à la valeur du seuil choisi. Nous avons décidé de réaliser une détection sub-pixellique des coins par optimisation d'un critère dépendant du gradient en intensité de l'image. La méthode proposée provient de la librairie de traitement d'image OpenCV d'Intel©.

Soit q le coin recherché dans l'image. Nous supposons que nous avons une première estimation grossière q_I de sa position et nous effectuons notre recherche sur une fenêtre f centrée autour. Nous pouvons remarquer que pour tout p_i situé dans f , $\left(\overrightarrow{(p_i - q)} \mid \overrightarrow{\nabla I_{p_i}}\right) = 0$, où $(\cdot \mid \cdot)$ représente le produit scalaire. En effet, si p_i est un contour alors le vecteur $\overrightarrow{(p_i - q)}$ est perpendiculaire à la direction du gradient en p_i , et lorsque p_i n'est pas un contour, il est dans une zone de couleur homogène, $\overrightarrow{\nabla I_{p_i}}$ est donc nul (cf. figure 5.9). Nous cherchons donc q tel que :

$$\epsilon^2 = \sum_{i \in f} \left(\overrightarrow{(p_i - q)} \mid \overrightarrow{\nabla I_{p_i}}\right)^2 = 0 \quad (5.1)$$

l'unicité de la solution étant théoriquement assurée puisque pour tout point différent de q , $\left(\overrightarrow{(p_i - q)} \mid \overrightarrow{\nabla I_{p_i}}\right) \neq 0$. Comme les observations sont soumises à du bruit, $\left(\overrightarrow{(p_i - q)} \mid \overrightarrow{\nabla I_{p_i}}\right)$ n'est jamais nul. Nous cherchons donc q qui minimise la fonction d'erreur 5.1. L'estimation de q est réalisée par une minimisation itérative de ϵ^2 .

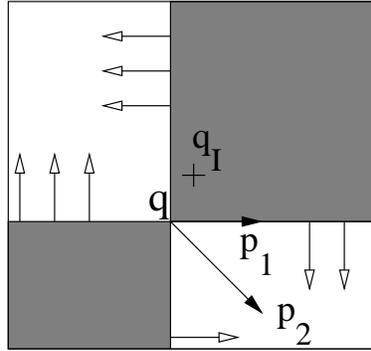


FIG. 5.9 – Nous cherchons à trouver le coin q à partir de l'initialisation q_I . Les flèches vides indiquent la direction du gradient dans l'image. Pour tout point p_i de la fenêtre, le produit scalaire $\epsilon_i = \left(\overrightarrow{(p_i - q)} \mid \overrightarrow{\nabla I_{p_i}}\right)$ est nul. Ce fait est évident pour un point sur un contour comme p_1 puisque $(p_1 - q)$ est orthonormale à la direction du gradient en p_1 . Pour un autre point comme p_2 , la zone étant uniforme le gradient ∇I_{p_2} vaut 0. ϵ_i étant non nul pour tout point différent de q , nous cherchons q tel que $\sum_{i \in f} \epsilon_i^2$ soit minimum.

Un critère de recalage incluant la distorsion

Le critère proposé par l'ARToolkit ne prend pas du tout en compte la distorsion potentielle des caméras. Nous avons donc étendu le critère en intégrant dans la fonction de projection les paramètres de distorsion disponibles grâce à la phase de calibrage préalable.

Validation par reprojection dans l'image

Afin de vérifier le gain de précision que nos modifications apportent, nous avons acquis une vingtaine de paires d'images du marqueur dans le repère des caméras et avons recalé le modèle

3D du marqueur en optimisant :

$$\sum_{i=1}^{i=4} \| P_1(T \star M_i) - m_i^1 \|^2 + \sum_{i=1}^{i=4} \| P_2(T \star M_i) - m_i^2 \|^2$$

Dans le premier cas, nous nous sommes servi de la détection des coins fournis par l'ARToolkit et n'avons pas inclus les paramètres de distorsion dans P_1 et P_2 , et dans le second cas, nous avons utilisé nos nouvelles fonctionnalités. Avec l'ARToolkit, l'erreur RMS de reprojection des coins du marqueur est de 1.2 pixel et avec notre version, nous obtenons 0.17 pixel.

Validation par reconstruction stéréoscopique

A partir des images de l'expérience précédente, nous avons reconstruit par stéréoscopie les coins détectés en minimisant sur M_i :

$$\sum_{i=1}^{i=4} \| P_1(M_i) - m_i^1 \|^2 + \sum_{i=1}^{i=4} \| P_2(M_i) - m_i^2 \|^2$$

où P_1 et P_2 contiennent les paramètres de calibrage *avec* distorsion. Nous avons ensuite réalisé un recalage aux moindres carrés 3D/3D entre le marqueur modèle et celui reconstruit dans le référentiel des caméras. L'erreur RMS de recalage obtenu avec la version ARToolkit est de l'ordre de 1.4 mm alors que notre détection sub-pixellique la réduit à 0.3 mm.

Ces 2 expériences démontrent donc que les améliorations que nous avons apportées à la librairie ARToolkit entraînent un gain en précision significatif.

5.3.2 Le calibrage de l'aiguille

Nous présentons maintenant la méthode que nous avons utilisé afin d'obtenir la position 3D de la pointe de l'aiguille et de son orientation dans le repère lié au marqueur.

Calibrage de la pointe : la méthode du pivot

Pour calibrer la pointe de l'aiguille, nous utilisons la méthode du pivot. La pointe de l'aiguille restant immobile (positionnée par exemple dans une encoche métallique), nous réalisons des acquisitions stéréoscopiques du marqueur dans de multiples positions différentes (typiquement une cinquantaine). Le marqueur étant en liaison rotule autour de la pointe, le barycentre des 4 coins du marqueur est censé décrire une sphère (cf. figure 5.10). En reconstruisant pour chaque paire d'image la position du barycentre dans le référentiel des caméras, nous disposons alors d'un jeu de 50 points appartenant à cette sphère, aux erreurs de calibrage et de reconstruction près. La position de la pointe de l'aiguille correspond au centre de cette sphère. Nous allons donc estimer les caractéristiques de la sphère passant par ces 50 points. En l'absence de bruit, chaque barycentre calculé $[x_i, y_i, z_i]^T$ satisfait l'équation :

$$(x_i - a)^2 + (y_i - b)^2 + (z_i - c)^2 - R^2 = 0$$

où (a, b, c) correspond au centre de la sphère et R à son rayon. Une méthode d'estimation des paramètres de la sphère consiste donc à rechercher (a, b, c, R) tel que la fonction $f(a, b, c, R)$ suivante soit minimum :

$$f(a,b,c,R) = \sum_{i=1}^{50} \| (x_i - a)^2 + (y_i - b)^2 + (z_i - c)^2 - R^2 \|^2$$

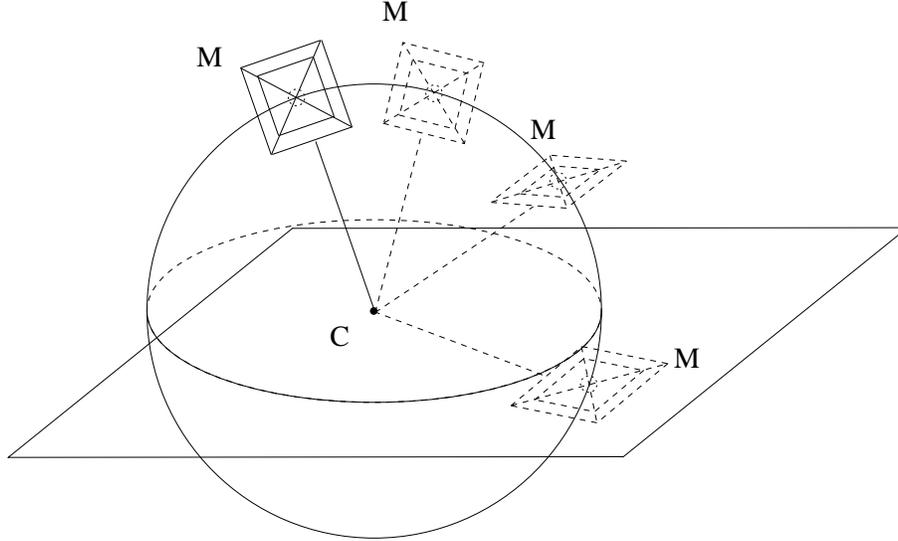


FIG. 5.10 – Le marqueur est déplacé, la pointe de l'aiguille restant immobile au point C . Le point M du marqueur effectue donc une trajectoire appartenant à la sphère de rayon MC et de centre C , i.e. la pointe de l'aiguille.

Considérant que $\hat{R}_i^2 = (x_i - a)^2 + (y_i - b)^2 + (z_i - c)^2$ est un estimateur du rayon R de la sphère, optimiser $f(a,b,c,R)$ revient à rechercher la sphère qui minimise la variance de l'estimateur \hat{R} .

Il existe deux autres méthodes (à notre connaissance) permettant d'estimer la position du centre de la sphère (elles sont décrites dans l'annexe B). Comme il n'existe aucune étude comparant la précision de ces méthodes, nous avons choisi la seule des trois ne nécessitant pas d'optimisation itérative. En écrivant l'équation de la sphère sous la forme :

$$x^2 + y^2 + z^2 + d \cdot x + e \cdot y + f \cdot z + g = 0$$

la solution minimisant f est :

$$\begin{bmatrix} d \\ e \\ f \\ g \end{bmatrix} = - \left(\sum_{i=1}^{50} [x_i, y_i, z_i]^T \cdot [x_i, y_i, z_i] \right)^{-1} \cdot \left(\sum_{i=1}^{50} (x_i^2 + y_i^2 + z_i^2) \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \right)$$

à condition que la matrice $\sum_{i=1}^{50} [x_i, y_i, z_i]^T \cdot [x_i, y_i, z_i]$ soit inversible, ce qui est faux uniquement si tous les points sont alignés (auquel cas il n'y a pas de sphère à trouver...).

Calibrage de l'orientation

Pour estimer l'orientation de l'aiguille, il suffit de la faire tourner autour de son axe et de rechercher l'axe fixe autour duquel tourne le marqueur. Dans notre cas de figure, la détermination de cet axe peut se faire en calculant les paramètres du cercle par lequel passe chaque coin du marqueur (cf. figure 5.11). Obtenant 4 cercles, l'axe fixe est alors calculé en moyennant les 4 vecteurs qui lui sont orthogonaux.

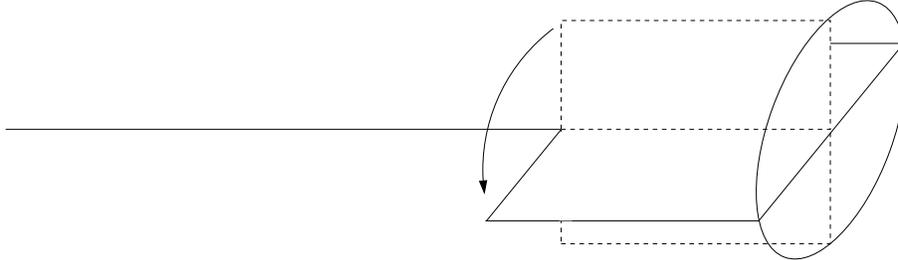


FIG. 5.11 – Lorsque l'aiguille tourne sur elle-même, la trajectoire décrite par chaque coin du marqueur est un cercle perpendiculaire à l'axe de rotation et dont le centre lui appartient.

Le problème de cette méthode réside dans la difficulté de maintenir l'orientation de l'aiguille pendant qu'on la tourne. Le meilleur moyen est de faire usiner un tube cylindrique fixé sur un support et dont le diamètre intérieur est très légèrement supérieur à celui de l'aiguille (au dixième de millimètre près). Comme les aiguilles utilisées par les praticiens ont des diamètres différents suivant l'intervention à réaliser, il faudrait faire concevoir un dispositif qui s'adapte aux dimensions de l'aiguille. Cette solution est onéreuse et peu pratique puisqu'elle est dépendante de la présence d'un instrument externe.

Il nous semble plus simple et plus judicieux de déterminer l'orientation de l'aiguille dans l'espace à partir de son orientation dans les 2 images vidéo. En effet, si nous parvenons à calculer automatiquement les équations des 2 droites par lesquelles passe l'aiguille dans les 2 images vidéo, alors nous sommes capables de calculer 2 plans de l'espace dans le repère de référence auquel appartient l'aiguille. Ces 2 plans étant forcément non-coplanaires du fait de la position relative de nos deux caméras, leur intersection nous fournit une droite qui correspond à l'orientation de l'aiguille (cf. figure 5.12).

Pour calculer automatiquement l'équation de la droite à laquelle correspond l'aiguille dans l'image vidéo, nous recherchons les droites qui passent par la pointe de l'aiguille en utilisant une transformée de Hough. L'étape précédente nous ayant fourni les coordonnées 3D de la position de la pointe dans le repère de référence, nous sommes à même de calculer automatiquement sa position en pixel dans l'image vidéo par reprojection. En fait, nous extrayons plusieurs droites que nous moyennons (cf. figure 5.13).

Finalement, l'étape de calibrage de l'aiguille nous fournit la position de la pointe de l'aiguille $P_{Marqueur}^{Aiguille}$ et de son orientation $\vec{V}_{Marqueur}^{Aiguille}$ dans le repère du marqueur.

5.3.3 Étude de la précision du positionnement de l'aiguille virtuelle

Dans cette section nous cherchons à évaluer la précision dans l'espace avec laquelle nous sommes capable de positionner l'aiguille virtuelle par rapport à l'aiguille réelle. Nous décrivons

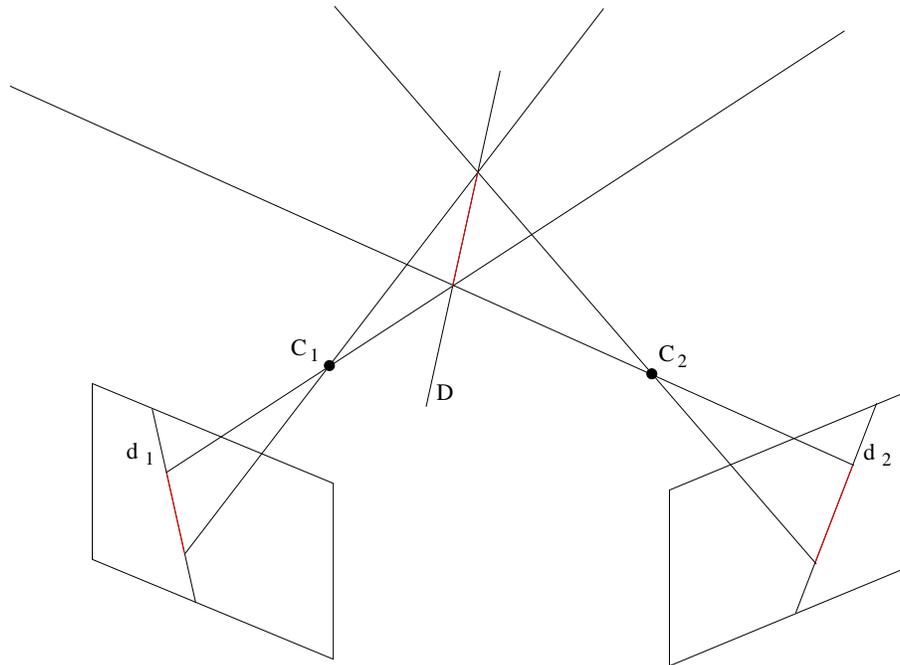


FIG. 5.12 – La connaissance de 2 droites d_1 et d_2 auxquelles appartient l'aiguille (en rouge) dans les images vidéos permet de déterminer la droite D par laquelle passe l'aiguille dans le monde réel. D est définie par l'intersection des plans (C_1, d_1) et (C_2, d_2) .

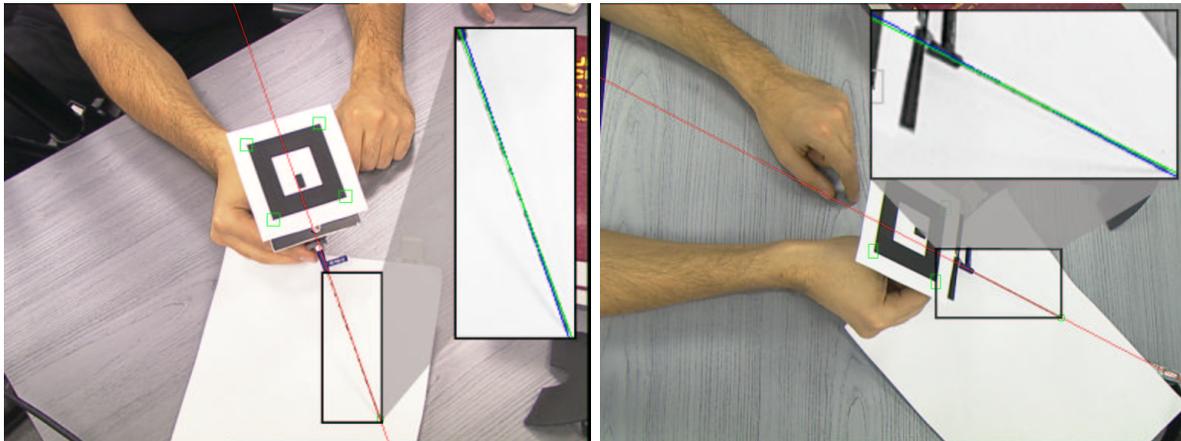


FIG. 5.13 – Détection automatique dans les images vidéo de la droite à laquelle appartient l'aiguille. Le résultat de la détection de droites par la transformée de Hough est grossi dans le rectangle noir. On peut distinguer qu'il y a en fait deux droites. Nous les moyennons et obtenons les droites tracées en rouge dans les deux images vidéo

tout d'abord le protocole expérimental utilisé pour réaliser nos mesures, puis nous présentons les résultats obtenus.

Le protocole expérimental

Une fois le calibrage de l'aiguille effectué nous re-calibrons les caméras afin d'éviter toute corrélation entre les mesures utilisées pour calibrer l'aiguille et celles que nous allons réaliser afin d'évaluer la précision. A la fin de la procédure de calibrage, la mire servant de référence aux caméras est immobilisée : les intersections du damier peuvent nous servir de mesures étalon puisque nous connaissons leur coordonnées. Nous avons réalisé 50 acquisitions stéréoscopiques dans lesquelles nous avons pointé une intersection du damier avec notre aiguille (cf. figure 5.14). Nous calculons alors pour chaque paire d'images la transformation rigide $T_{\text{Marqueur} \rightarrow \text{Mire}}$ reliant le repère du marqueur à celui de la mire de référence en minimisant le critère de recalage 3D/2D :

$$\sum_{i=1}^{i=4} \| P_1(T_{\text{Marqueur} \rightarrow \text{Mire}} \star M_i) - m_i^1 \|^2 + \sum_{i=1}^{i=4} \| P_2(T_{\text{Marqueur} \rightarrow \text{Mire}} \star M_i) - m_i^2 \|^2$$

et nous appliquons $T_{\text{Marqueur} \rightarrow \text{Mire}}$ à $P_{\text{Marqueur}}^{\text{Aiguille}}$. Nous obtenons ainsi les coordonnées de la pointe de l'aiguille virtuelle dans le repère de référence. Connaissant la position réelle de l'intersection pointée avec l'aiguille, nous obtenons une évaluation de l'erreur 3D de repositionnement de la pointe de l'aiguille virtuelle. Comme cette mesure dépend de la qualité de l'orientation du marqueur dans les images vidéo et de l'interprétation de l'opérateur, nous faisons varier la position du marqueur et de la pointe pour chaque paire d'images acquise.

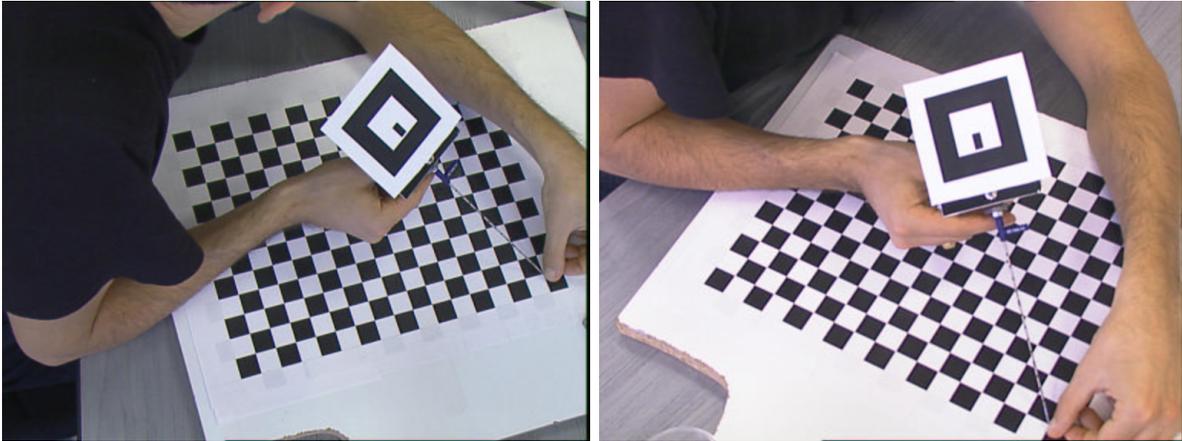


FIG. 5.14 – Évaluation de l'erreur de repositionnement de la pointe après calibrage de l'aiguille : à la fin du calibrage des caméras, la mire de référence est immobilisée et nous utilisons les intersections du damier comme mesures étalon.

Signalons au passage que nous aurions pu reconstruire les quatre coins du carré par stéréoscopie, et calculer T en minimisant le critère LS classique 3D : $C = \sum_{i=1}^{i=4} \| T \star M_i - M_i^! \|^2$. Nous n'en avons rien fait puisque ce critère n'est pas optimal au sens statistique. En effet, il faudrait préalablement calculer l'incertitude Σ_{3D}^i sur les reconstructions stéréoscopiques (qui sont anisotropes) et minimiser plutôt le critère : $\sum_{i=1}^{i=4} (T \star M_i - M_i^!) \cdot \Sigma_{3D}^i{}^{-1} \cdot (T \star M_i - M_i^!)^\top$ qui n'a pas de solution directe. De fait nous préférons utiliser le critère 3D/2D qui est optimal puisque l'écart-type du bruit sur l'extraction des données vidéo est constant et indépendant d'un point à l'autre.

Afin d'évaluer l'erreur d'orientation de l'aiguille virtuelle, nous avons acquis 50 paires d'images de l'aiguille dans lesquelles nous l'avons aligné le long des droites du damier dont nous connaissons l'orientation (cf. Figure 5.15). Après application de $T_{\text{Marqueur} \rightarrow \text{Mire}}$ à $\vec{V}_{\text{Marqueur}}^{\text{Aiguille}}$, nous pouvons alors calculer l'erreur d'orientation de l'aiguille virtuelle par rapport à l'aiguille réelle.

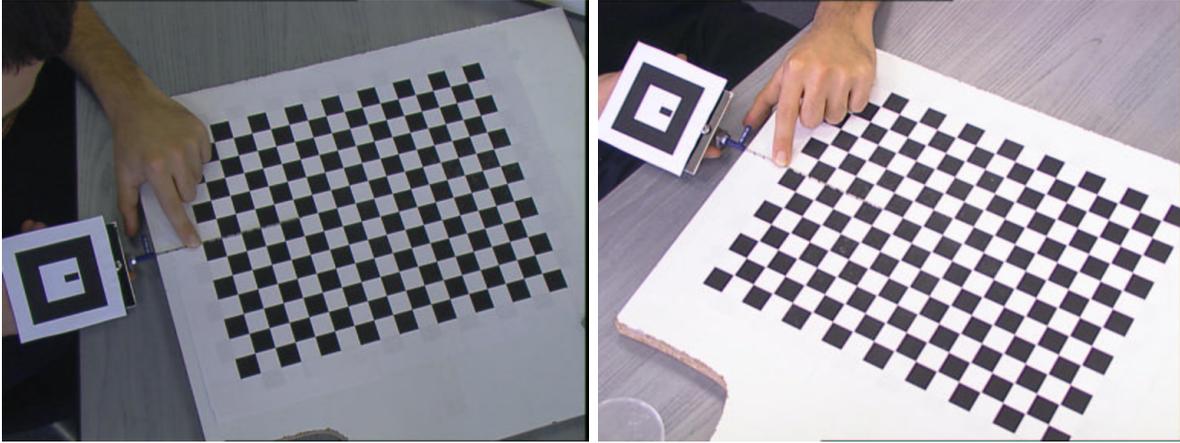


FIG. 5.15 – Les droites du damier nous servent d'orientation étalon pour évaluer l'erreur d'orientation de l'aiguille virtuelle sur l'aiguille réelle.

Résultats

Le tableau 5.1 contient les résultats des calibrages de l'aiguille obtenus avec l'ARToolkit et avec notre version. En premier lieu, nous constatons qu'il y a environ 1 mm de différence sur l'estimation de $P_{\text{Marqueur}}^{\text{Aiguille}}$. A priori, il n'est pas possible de savoir laquelle de ces deux estimations est la plus proche de la réalité, néanmoins nous constatons aussi que l'écart-type sur les 50 estimations réalisées de $P_{\text{Marqueur}}^{\text{Aiguille}}$ est beaucoup plus faible avec notre version qu'avec l'ARToolkit. Cela signifie que la précision de notre version sur l'estimation sera de meilleure qualité s'il n'y a pas de biais sur l'extraction des coins du carré. Cette conjecture se voit confirmée par les résultats du tableau 5.2 qui contient les erreurs de repositionnement de la pointe ainsi que l'erreur d'orientation de l'aiguille virtuelle. Nous y lisons effectivement que notre version fournit un gain de 52% en précision sur le positionnement de la pointe. Au niveau de l'orientation nous notons une amélioration d'un facteur 1.5. Notre version améliore donc considérablement les performances initiales de l'ARToolkit. Le résultat d'un de nos calibrages est illustré sur la figure 5.16.

Pour confirmer la qualité du recalage obtenu avec notre version améliorée de l'ARToolkit, nous avons aussi mesuré la répétabilité de notre méthode de calibrage. Nous avons réalisé 20 calibrages successifs de l'aiguille et avons calculé les écart-types sur les coordonnées des 20 estimations obtenues de $P_{\text{Marqueur}}^{\text{Aiguille}}$ et de $\vec{V}_{\text{Marqueur}}^{\text{Aiguille}}$ (les résultats sont résumés dans le tableau 5.3). Nous constatons à nouveau la supériorité de notre version puisque les écart-types obtenues sont approximativement 2 fois plus faibles avec notre version. Par ailleurs, nous soulignons la qualité de la répétabilité de notre méthode puisque, d'un calibrage à un autre, la distance entre deux estimations différentes de la position de la pointe est inférieure à 0.9 mm dans 95% des cas.

5.3. Repérage de l'aiguille dans le repère des caméras

Coordonnées en mm	$P_{\text{Marqueur}}^{\text{Aiguille}}$			$\vec{V}_{\text{Marqueur}}^{\text{Aiguille}}$		
	X	Y	Z	X	Y	Z
ARToolkit	-0.9±1.9	-239.8±2.3	-16.4±1.5	-1.7e-3 ±1e-2	-0.999±1.4e-2	-1.4e-2±9e-2
Notre version	-0.65±1.1	-239.9±0.9	-15.6±0.8	2.6e-3±8e-3	-0.999±9e-3	6e-3±6e-2

TAB. 5.1 – Résultats avec l'ARToolkit et avec notre version de l'estimation de la position de la pointe et de l'orientation de l'aiguille par rapport au marqueur. Les deux estimations calculées sont distantes d'environ 1 mm et l'écart-type sur l'estimation de notre version est environ 1.9 fois plus faible.

	Moyenne de l'erreur de repositionnement de la pointe (\pm écart-type)	Moyenne de l'erreur d'orientation de l'aiguille virtuelle (\pm écart-type)
ARToolkit	1.45±1.3 mm	1.11°±0.66
Notre version	0.95±0.5 mm	0.72°±0.15

TAB. 5.2 – Résultats de l'étude de précision. Le positionnement de la pointe est 1.5 fois plus précis et l'orientation de l'aiguille est 1.6 fois plus proche de l'orientation étalon avec notre version qu'avec l'ARToolkit.

Ecart-type en mm	$P_{\text{Marqueur}}^{\text{Aiguille}}$				$\vec{V}_{\text{Marqueur}}^{\text{Aiguille}}$			
	X	Y	Z	$\sqrt{X^2 + Y^2 + Z^2}$	X	Y	Z	$\sqrt{X^2 + Y^2 + Z^2}$
ARToolkit	0.58	0.62	0.41	0.94	4.1e-3	5.2e-3	2.2e-2	2.5e-2
Notre version	0.33	0.32	0.23	0.51	2.3e-3	3.8e-3	1.4e-2	1.7e-2

TAB. 5.3 – Résultats de l'étude de répétabilité sur le calibrage de l'aiguille. Les colonnes X, Y et Z contiennent les écart-types sur les coordonnées isolées alors que les colonnes $\sqrt{X^2 + Y^2 + Z^2}$ contiennent les écart-types en 3D. La répétabilité est bien meilleure avec notre version, l'écart-type en 3D étant presque deux fois plus faible.

Bien que la précision finalement atteinte avec notre version soit relativement satisfaisante, les travaux antérieurs concernant le calibrage de la pointe d'une aiguille rigide [CL98, FDP⁺02] nous avait laissé penser que nous atteindrions une erreur de repositionnement de l'ordre de 0.5 mm. Bien que l'interprétation humaine a un impact (la mesure étalon dépend de la qualité du positionnement par l'opérateur de l'aiguille sur la mire), nous estimons que le nombre de mesures réalisées réduit considérablement son effet néfaste sur l'évaluation de la précision. Nous pensons plutôt, à l'instar de [BTFC03, BFTC03], que la flexibilité de l'aiguille (dont le diamètre vaut seulement 2 mm) au moment de son calibrage est principalement responsable des erreurs. En effet, nous avons observé comme eux qu'une légère pression suffit à modifier de 1 mm la position de la pointe.



FIG. 5.16 – *Validation visuelle de la précision du recalage de l'aiguille virtuelle sur l'aiguille réelle. La croix dans les images correspond à la position virtuelle de la pointe de l'aiguille, et la droite à son orientation. Les zones encadrées ont été agrandies pour une meilleure vérification visuelle.*

5.4 Conclusion

Dans ce chapitre, nous nous sommes d'abord penchés sur le problème du nombre de caméra. La précision du système dépendant de leur nombre, nous avons mené une étude montrant qu'à partir de 2 caméras, la précision obtenue satisfaisait nos contraintes. Comme l'encombrement en salle d'opération est limité, nous avons choisi de ne pas en utiliser plus.

Dans une seconde partie, nous avons traité le problème du choix de la méthode de calibrage de nos caméras. La solution idéale consiste à utiliser du matériel métrologique (banc de calibrage, mire de calibrage usinée). D'une part nous ne disposons pas de matériel de ce type, d'autre part il ne serait pas pratique de devoir emmener un tel matériel au bloc opératoire. Nous avons alors montré que, parmi les techniques de calibrage existantes, la méthode multi-vues de Zhang est très précise et ne nécessite qu'une mire coplanaire : un objet plan et un damier imprimé collé dessus suffisent donc pour obtenir un calibrage de qualité.

Dans la dernière partie, nous expliquons comment nous repérons l'aiguille fixée à un marqueur, et comment nous la calibrons en utilisant uniquement un traitement des images acquises par nos caméras. Nous démontrons ensuite en menant une étude expérimentale que l'erreur de repositionnement de la pointe de l'aiguille virtuelle vaut 0.95 mm en moyenne, et que l'erreur

moyenne d'orientation est de 0.7° . La précision obtenue avec notre procédé est donc compatible avec nos objectifs puisqu'en ajoutant les trois sources d'erreurs de notre système (repositionnement du foie, erreur de recalage du modèle virtuelle du patient, erreur de recalage de l'aiguille virtuelle), nous obtenons : $\sqrt{1^2 + 2^2 + 0.95^2} = 2.42$ mm. Nous restons donc en dessous des 5 mm imposés par nos contraintes initiales, avec une certaine marge d'erreur.

Finalement, ce chapitre contient, à notre connaissance, la première évaluation de la précision qu'il est possible d'atteindre en utilisant une version améliorée d'une librairie gratuite sur internet (l'ARToolkit). Il existe des articles comparant la précision de la détection de marqueurs *dans les images vidéo* [ZFN03], mais aucun n'a étudié l'erreur de recalage *dans l'espace* lors de l'utilisation de marqueurs carrés fixés sur un outil. Ainsi, nous avons montré qu'il est possible d'avoir à moindre frais un système de repérage basé sur une acquisition stéréoscopique dont la précision est sub-millimétrique.

Chapitre 6

Automatisation de la localisation et de l'appariement des marqueurs

Dans le **chapitre 4**, nous avons défini un nouveau critère de minimisation adapté aux hypothèses de notre problème afin de déterminer la transformation liant le repère caméra au repère scanner. Lorsque nous avons formulé l'EPPC, nous supposons implicitement que nous connaissons la position des marqueurs radio-opaques dans les images vidéo et scanner ainsi que leurs appariements respectifs. En pratique, cette procédure pourrait être réalisée interactivement par le praticien qui pointerait les marqueurs dans les images et les mettrait en correspondance ultérieurement. Néanmoins, cette phase de pré-traitement interactive est longue et fastidieuse (environ 10 minutes sont nécessaires pour localiser et appairer une vingtaine de marqueurs). Comme nous souhaitons réduire la durée des interventions de radiologie interventionnelle, la mise au point d'algorithmes d'extraction et d'appariement automatiques représenterait un bénéfice majeur pour le praticien lors de l'emploi de notre système.

De ce fait, ce chapitre est consacré à la description et l'évaluation de nos algorithmes d'extraction et d'appariement des marqueurs radio-opaques. Le temps nécessaire à l'extraction des marqueurs dans les images TDM est de l'ordre de 2 minutes sur un PC à 1 GigaHz alors que le suivi des marqueurs fonctionne en temps réel après une première itération des autres procédures (extraction dans les images vidéo et appariement) qui prend moins de 10 secondes. Etant donné que la localisation des marqueurs dans l'image scanner n'est réalisée qu'une seule fois, le temps de calcul de notre algorithme automatique n'est pas prohibitif pour notre application. Nous pensons exécuter cette tâche en moins de 30 secondes dans des conditions plus favorables (PC de dernière génération, optimisation du code). L'évaluation de l'ensemble des algorithmes sur des données réelles provenant de notre mannequin montre que nous extrayons et appariions correctement plus de 90% des marqueurs présents dans les images sources (sans aucun faux positif). Bien que d'autres améliorations soient encore envisageables, les algorithmes développés sont robustes dans le cadre de notre application et satisfont nos objectifs en terme de vitesse d'exécution.

6.1 Extraction des marqueurs radio-opaques

Cette section contient une description de nos deux algorithmes d'extraction. Les marqueurs radio-opaques sont caractérisés principalement par une étude de l'histogramme des couleurs dans les images vidéo, et par l'application de l'opérateur morphologique chapeau haut de forme

dans les images tridimensionnelles TDM.

Pour les marqueurs des images vidéo, une comparaison statistique du résultat de notre extraction automatique avec un ensemble de repérage interactif montre que notre algorithme est aussi précis qu'une localisation manuelle en moyenne (et plus centré). Une évaluation du même type est en cours pour les images scanners et les résultats préliminaires semblent confirmer des résultats équivalents.

6.1.1 Extraction dans les images vidéos

Méthodologie

Nous supposons initialement que la peau du patient occupe au moins 50% de l'image. Cette hypothèse semble vraisemblable puisque nous choisissons nous-même la focale des caméras orientées vers le patient. Dans le cas où le champ opératoire (de couleurs uniforme verte) était prépondérant dans l'image, son identification serait aisée en raison de sa localisation en bordure d'image.

La première étape consiste à identifier la peau du patient dans l'image. Nous passons tout d'abord du modèle de représentation des couleurs RGB au modèle HSV afin de caractériser les couleurs dans un système à une dimension. En effet, nos études préliminaires nous ont montré que la considération d'une des composantes RGB (nous avons immédiatement pensé au canal rouge) ne permettait pas de différencier efficacement les marqueurs de la peau. Nous recherchons ensuite dans l'histogramme H_{chrom} de la chrominance de l'image, la couleur moyenne de la peau CM_{peau} . Nous définissons cette valeur comme étant la position du mode dont l'intégrale (au sens mathématique) est la plus importante et non pas la position de celui dont la hauteur est la plus élevée. En effet, une petite région de couleur plus homogène que la peau peut correspondre à un mode d'écart-type bien plus faible mais de valeur maximale plus élevée. Considérant que la distribution de chrominance autour du mode détecté est approximativement gaussienne, un seuillage est alors réalisé. Un choix classique pour les valeurs de seuil T_{min} et T_{max} est :

$$T_{min} = CM_{peau} - 3 \cdot \sigma_{peau} \quad \text{et} \quad T_{max} = CM_{peau} + 3 \cdot \sigma_{peau}$$

où σ_{peau} est l'écart-type du mode correspondant à la peau. Afin d'évaluer σ_{peau} , nous pourrions calculer la largeur à mi-hauteur de la distribution de chrominance de la peau, qui est théoriquement égale à $2.3548 \cdot \sigma_{peau}$ sous l'hypothèse gaussienne. Cependant, comme cette distribution n'est pas parfaitement gaussienne, il est préférable d'évaluer σ_{peau} à partir d'une version lissée de l'histogramme H_{chrom} . Cet histogramme optimal est défini comme l'histogramme original lissé itérativement avec un filtre gaussien jusqu'à ce que le nombre de modes soit stable (cf. figure 6.1). Etant donné qu'un lissage nivellement la distribution mais ne modifie pas son intégrale, σ_{peau} s'exprime en fonction de σ_{opt} de la manière suivante :

$$\sigma_{peau} \cdot Max_{peau} = \sigma_{opt} \cdot Max_{opt}$$

où σ_{opt} est l'écart-type de la version lissée et Max_{opt} (resp. Max_{peau}) sa valeur maximale (resp. la valeur maximale de H_{chrom}).

Après cette étape, comme nous supposons que la peau représente la plus grande surface, nous extrayons la plus grande composante connexe. Les marqueurs ayant une couleur très différente de la peau (ils sont verts), nous pouvons considérer que toutes les composantes connexes incluses

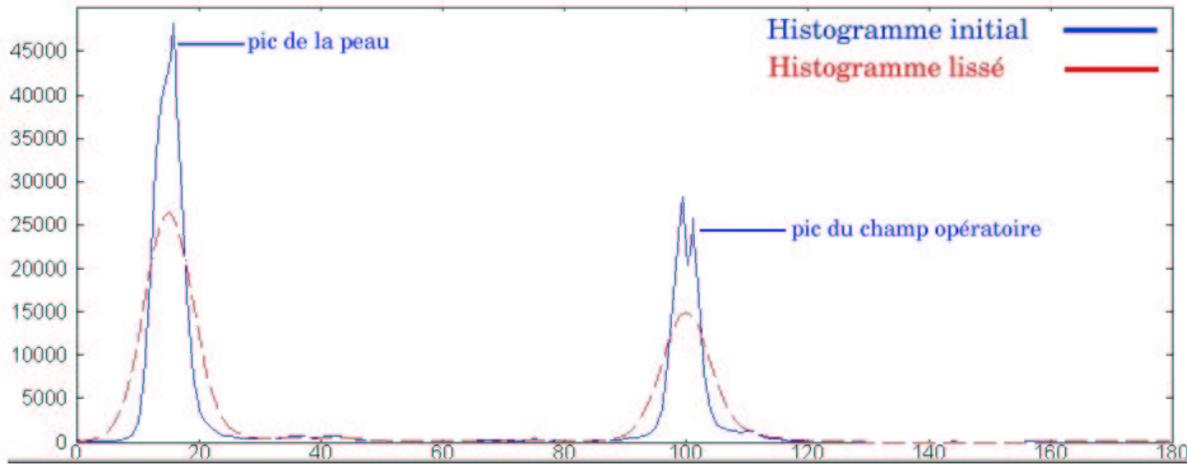


FIG. 6.1 – Exemple d’histogramme de chrominance obtenu sur des données réelles. La courbe en bleu correspond à l’histogramme original. Bien que la forme de la distribution de couleur relative à la peau (pic de gauche) ressemble à une distribution gaussienne, son pic n’est pas symétrique et semble trop élevé par rapport à une distribution classique. Une estimation de σ_{peau} à partir de la largeur à mi-hauteur risque d’être sous-évaluée. Il est préférable de lisser l’histogramme H_{peau} et d’estimer σ_{peau} à partir de cette version lissée H_{opt} (histogramme en pointillé rouge).

dans la peau détectée sont des marqueurs potentiels¹⁹. La caractérisation des vrais marqueurs parmi toutes ces composantes est réalisée en plusieurs étapes. En premier lieu, nous éliminons celles dont la surface est trop petite (typiquement moins de 10 pixels de surface). Ensuite, nous calculons la médiane des surfaces des composantes S_{median} et nous rejetons celles dont la surface n’est pas dans l’intervalle $[\alpha \cdot S_{\text{median}}; 1/\alpha \cdot S_{\text{median}}]$. Le choix de α est réalisé en fonction de l’inclinaison maximale des marqueurs par rapport à l’axe optique des caméras (cf. figure 6.2). Pour notre application, nous considérons une inclinaison maximale θ d’environ 60° et choisissons en conséquence $\alpha = 0.5$. Finalement, nous éliminons les composantes dont le moment d’inertie est trop petit (caractérisant les composantes trop plates). La position des marqueurs est estimée en calculant le centre de gravité des composantes restantes. Un résumé de la méthode d’extraction est illustré sur la figure 6.3.

Evaluation

Afin d’évaluer notre algorithme, nous avons collé sur 9 personnes un nombre de marqueurs compris entre 10 et 20, et réalisé 2 acquisitions vidéo stéréoscopiques pour chacun des sujets.

19. En pratique, il n’y a pas d’autre objet dans l’image dont la chrominance soit suffisamment proche de celle de la peau pour être conservé après seuillage. Si jamais cela arrive, la distribution relative à la chrominance de la peau sera bruitée par celle de la composante dont la chrominance est proche de la peau. Il y aura donc un mode dans l’histogramme qui regroupera ces deux objets différents. Cela peut nous mener à un seuillage caractérisant incorrectement les marqueurs. Pour supprimer les effets de ce bruit, nous pouvons ajouter quelques traitements supplémentaires à notre algorithme. Par exemple, une fois que la peau et que l’objet indésirable ont été extraits par seuillage, nous identifions la peau comme la composante principale et nous calculons un nouvel histogramme des chrominances de l’image vidéo masquée par la composante connexe de la peau préalablement remplie. La recherche de CM_{peau} et des seuils T_{min} et T_{max} est alors réalisée sur cet histogramme exempt de bruit.

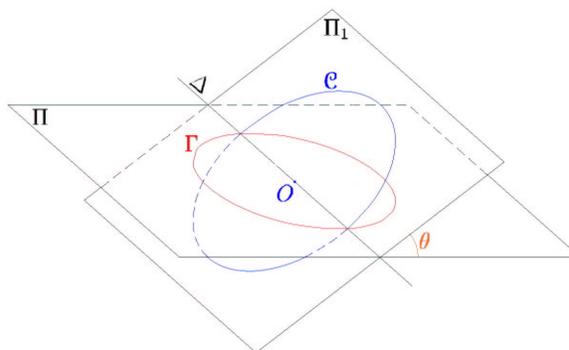


FIG. 6.2 – La surface d'un marqueur dans l'image vidéo dépend de son inclinaison par rapport à l'axe optique de la caméra. Les marqueurs étant suffisamment éloignés de la caméra, nous pouvons approximer la transformation projective par une projection orthogonale sur le plan image de la caméra. La projection d'un disque C de rayon R résulte donc en une ellipse Γ de demi-axes R et $R \cdot \cos\theta$. Le rapport entre les surfaces de C et Γ vaut donc $\cos\theta$. Dans nos conditions θ n'excédera pas 60° , un choix conservateur pour α et $1/\alpha$ est donc $\cos 60^\circ = 0.5$ et $\frac{1}{\cos 60^\circ} = 2.0$.

A chaque nouvelle acquisition les caméras étaient déplacées et leur focale modifiée afin d'augmenter la diversité des conditions expérimentales. Pour certaine acquisition, un léger réglage du diaphragme était nécessaire afin de détecter correctement les marqueurs. En effet, une luminosité trop faible ou trop forte réduit le contraste des couleurs et ne permet plus de dissocier le pic de la peau de celui des marqueurs.

Ayant acquis récemment un autre type de marqueur (Multi-Modality Marker TM 4001 IZI Product) que nous souhaitions comparer aux marqueurs verts classiques (cf. figure 6.4), la procédure d'acquisition présentée précédemment fut réalisée une seconde fois. Nous avons constaté avec ces marqueurs que la qualité de l'extraction était beaucoup plus dépendante de la luminosité : un réglage fin du diaphragme était nécessaire afin de tous les détecter dans l'image. Leur couleur rouge claire plus proche de la teinte de la peau explique vraisemblablement ce phénomène. Finalement, un total de 72 images a été obtenu sur lesquelles nous avons relevé le nombre de faux positifs et négatifs après exécution de notre algorithme d'extraction. Un exemple d'extraction est présenté sur la figure 6.5. Les résultats totaux (indiqués dans le tableau 6.1) sont très satisfaisants puisque 2% seulement des marqueurs n'ont pas été détectés, il n'y a que 1% de faux positifs et les temps de calcul sont très honorables (valeur maximale en dessous d'une seconde). Par ailleurs, une étude des images dans lesquelles nous avons constaté les faux positifs indique que leur présence est due à un vêtement de teinte semblable à celle de la peau. Nous pouvons donc raisonnablement penser que le risque d'occurrence d'un faux positif est quasiment nul durant une intervention.

Comparaison statistique avec des relevés interactifs

Les marqueurs autocollants radio-opaques n'étant pas conçus de manière à ce qu'une mesure étalon de leur position puisse être établie, il n'est pas possible de quantifier la précision de notre algorithme d'extraction. Néanmoins, nous pouvons la comparer qualitativement avec une extraction interactive afin de montrer que la version automatique est équivalente à un relevé

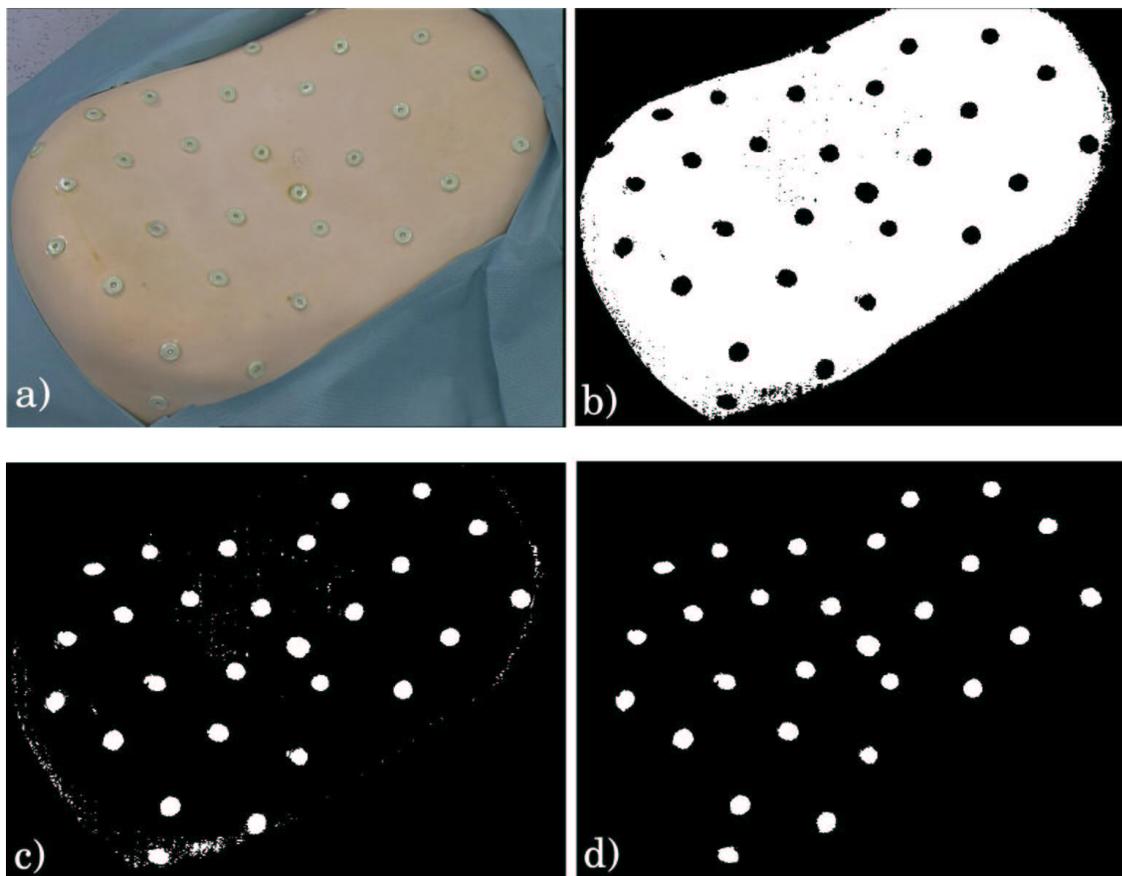


FIG. 6.3 – (a) Image originale du mannequin. (b) Application du seuillage et extraction de la plus grande composante connexe. (c) Les marqueurs potentiels sont toutes les composantes connexes en blancs. (d) Extraction finale des marqueurs dans la peau détectée. Les marqueurs réels sont caractérisés parmi ces composantes en fonction de leur surface et de leur forme.



FIG. 6.4 – Marqueurs utilisés pour nos expériences. A gauche : Multi-Modality Marker MM 3005 IZI Product. A droite : Multi-Modality Marker TM 4001 IZI Product

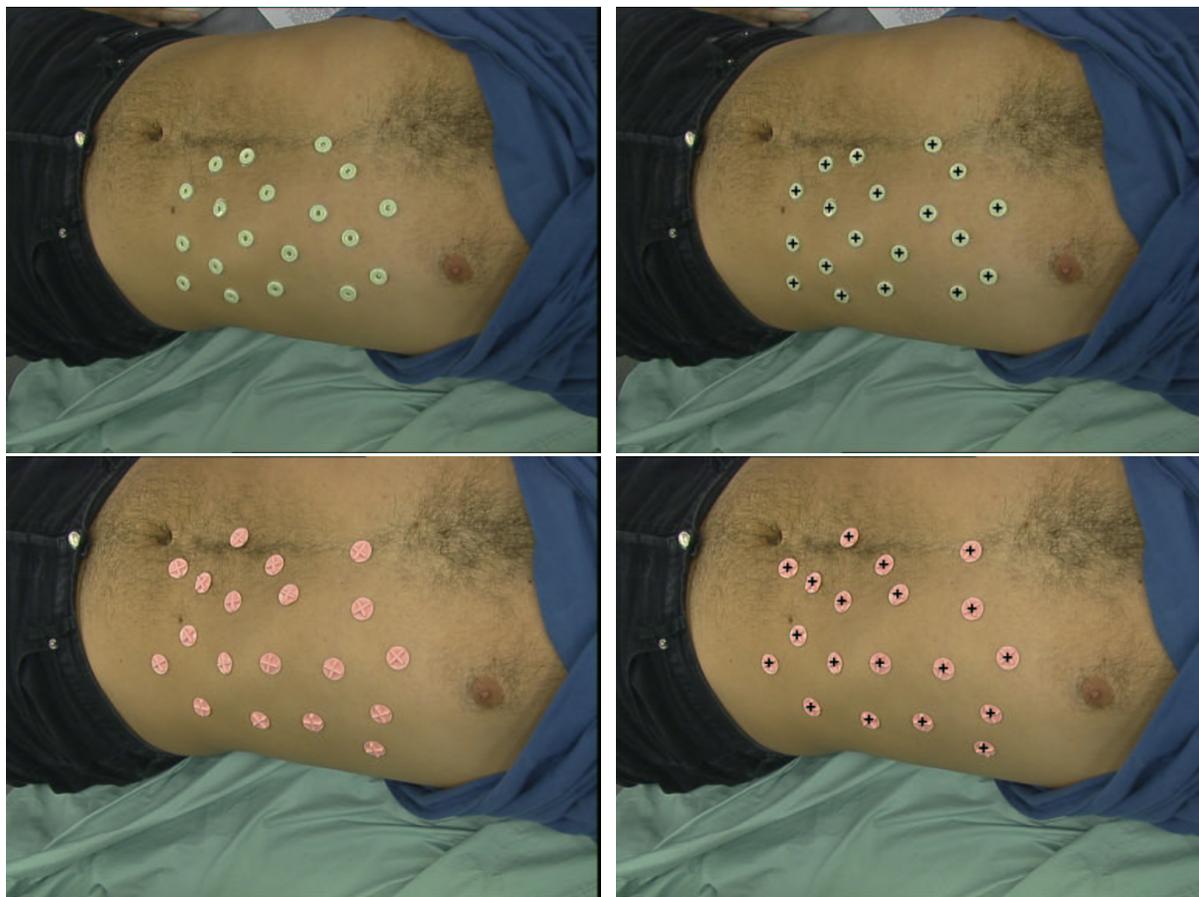


FIG. 6.5 – Exemple d'extraction automatique de marqueurs sur des torses humains. Les croix noires dans les images de droites indiquent le résultat de la localisation automatique. Les deux types de marqueurs ont été correctement détectés.

	Nombre de marqueurs	Faux positifs	Faux négatifs	Temps de calcul maximal
Marqueurs verts	545	0.6%	2.0%	0.95 sec.
Marqueurs roses	539	1.0%	2.0%	0.99 sec.
Total	1084	0.6%	2.0%	0.97 sec.

TAB. 6.1 – Pourcentage de faux positifs et faux négatifs obtenus après application de notre algorithme d'extraction sur les 72 images vidéo. Ces résultats et le temps d'exécution maximal sur l'ensemble des images sont donnés en fonction du type de marqueur.

manuel. Dans ce but, nous avons demandé à 6 personnes de relever dans les images vidéo la position des marqueurs. Nous disposons ainsi, pour chaque marqueur, d'une détection automatique et de 6 observations humaines.

Biais indépendant pour chaque marqueur : dans ce premier test, nous souhaitons savoir si, pour chaque marqueur m_i^l de l'image l , l'extraction automatique A_i^l est significativement différente de la moyenne \bar{P}_i^l des 6 relevés interactifs. Le test de Hotelling [LKH96] nous aurait permis de tester l'hypothèse nulle: "la moyenne de la population dont sont issus les P_i^l vaut A_i " sans nécessairement connaître précisément la covariance de la distribution des P_i^l (elle est estimée empiriquement à partir des données). En pratique, nous n'avons pas pu nous servir de ce test à cause du faible nombre de données par marqueur (6 mesures) et du bruit de quantification qui corrompt les relevés interactifs (dû à la nature discrète de l'image). En effet, les matrices de covariance empiriques obtenues n'étaient pas toujours inversibles, elles étaient donc très peu représentatives de la réalité.

Pour contourner ce problème nous avons choisi d'estimer la matrice de covariance relative aux extractions manuelles à partir de toutes les données disponibles :

$$\Sigma_{mm} = \frac{1}{N \cdot L \cdot 6 - 1} \sum_{l=1}^N \sum_{i=1}^{L^l} \sum_{k=1}^6 \| P_{i,k}^l - \bar{P}_i^l \|^2$$

où N est le nombre total d'image et L^l le nombre de marqueurs de l'image l . Pour différencier les deux types de marqueur, nous avons calculé séparément les matrices de covariance et avons obtenu :

$$\Sigma_{mm}^{vert} = \begin{bmatrix} 0.85 & 0.12 \\ 0.12 & 1.3 \end{bmatrix} \quad \Sigma_{mm}^{rose} = \begin{bmatrix} 0.95 & 0.21 \\ 0.21 & 1.2 \end{bmatrix}$$

Le nombre de données étant très important (plus de 3000 relevés interactifs par type de marqueur), nous considérons que la matrice de covariance est connue, auquel cas nous changeons de test et utilisons un test du χ^2 . Pour ce test, notre hypothèse nulle est: "La moyenne des relevés interactif est égal à A_i^l et sa covariance vaut $\Sigma_{mm}^{couleur}$ ". Choisisant une valeur de rejet (p-value) à 99%, nous obtenons les résultats décrits dans le tableau 6.2.

	Nombre total de marqueurs	Rejet de l'hypothèse nulle
Marqueurs verts	545	49
Marqueurs roses	539	214

TAB. 6.2 – Nombre de marqueurs pour lesquels l'hypothèse nulle a été rejetée. Dans le cas des marqueurs verts, plus de 90% des cas ont passé le test avec succès, alors que pour les roses, plus de 40% des marqueurs ont conduit à un rejet de l'hypothèse nulle.

Il semble donc que l'extraction automatique soit bien plus fiable pour les marqueurs verts que pour les roses puisque 9% seulement des relevés automatiques pour les marqueurs verts ont été considérés comme significativement éloignés des relevés manuels (contre 35% pour les roses). Une différence aussi importante provient du fait que les utilisateurs ont cliqué sur la croix moulée dans les marqueurs roses alors que notre algorithme extrait le centre de la forme du marqueur reprojecté dans l'image. Ainsi, dès qu'un marqueur était légèrement incliné par rapport à l'axe optique de la caméra, les relevés automatiques étaient biaisés (cf. figure 6.6).

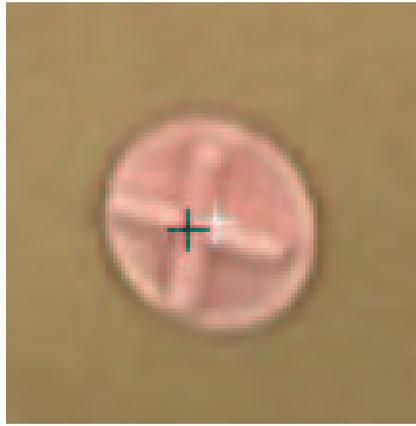


FIG. 6.6 – Mise en évidence sur un exemple du biais observé avec les marqueurs roses. En noir : moyenne des relevés interactif. En blanc : résultat de l'extraction automatique. Les opérateurs ont localisé le centre du marqueur au niveau du centre de la croix en relief sur le marqueur (qui correspond bien au barycentre du marqueur). L'extraction automatique en est éloigné de plusieurs pixels (au moins 4), puisqu'elle correspond au centre de la projection du marqueur rose dans l'image vidéo.

Bien que 90% de succès pour les marqueurs verts semble être un bon résultat, nous nous attendions à encore mieux. Nous avons donc décidé de considérer visuellement l'extraction automatique et le barycentre des relevés manuels pour les marqueurs dont le test a été rejeté, afin de voir si l'un des deux relevés semblait réellement éloigné de la réalité. Nous avons tout d'abord remarqué que les 49 marqueurs en question étaient fortement inclinés par rapport à l'axe optique des caméras (au moins 45°). Nous nous sommes ensuite rendu compte que les opérateurs avaient systématiquement cliqué au centre de l'ellipse supérieure du marqueur alors que nous leur avons demandé de relever la projection du barycentre du marqueur.

Le relevé interactif semble donc avoir été biaisé à chaque fois en fonction de l'inclinaison du marqueur. Une vérification attentive a permis de déceler 27 marqueurs (parmi les 50) pour lesquels le relevé interactif moyen était sensiblement éloigné de la projection du barycentre du marqueur d'au moins 2 pixels (alors que l'extraction automatique semblait en être plus proche). Pour les 23 autres marqueurs, les précisions obtenues ont semblé équivalentes.

Nous pouvons donc conclure que, dans le cas des marqueurs verts, notre algorithme d'extraction automatique fournit un résultat globalement équivalent en précision à celui d'un relevé interactif, et que, dans le cas de marqueur incliné par rapport à l'axe optique de la caméra, il se révèle peut être même plus précis. En revanche, pour les marqueurs roses, la détection est bien moins efficace à cause de la forme particulière des marqueurs. Nous verrons dans le paragraphe suivant qu'une faible inclinaison des marqueurs peut provoquer un biais d'une amplitude d'environ 1 pixel sur l'ensemble de la détection des marqueurs dans l'image.

Biais dépendant pour les marqueurs d'une image : nos marqueurs n'étant pas parfaitement plats, leur projection dans une image vidéo ne correspond pas vraiment à une ellipse²⁰, notamment quand ils sont inclinés par rapport à l'axe optique de la caméra. Si une majorité

20. En réalité, la projection d'un cercle dans une image vidéo n'est pas non plus une ellipse. Néanmoins lorsque l'objet est suffisamment éloigné de la caméra, cette approximation est très proche de la réalité.

des marqueurs n'est pas face à la caméra, il est alors possible qu'un biais systématique de même direction (mais d'amplitude différente) affecte l'extraction automatique. Le test établi précédemment permet de jauger la différence d'extraction entre une version automatique et une version interactive au niveau d'un marqueur. Cependant, s'il existe un biais systématique selon une orientation pour une image donnée entre les deux types de détection, ce test ne le décelera pas puisqu'il ne tient pas compte de l'information relative à l'extraction des autres marqueurs de la même image.

Nous allons nous servir comme précédemment d'un test du χ^2 , mais avec des données différentes. Etant donné une image l , nous resituons pour chacun des marqueur m_i^l les 6 relevés interactifs P_i^l dans le repère centré en A_i^l . Nous obtenons alors un échantillon Q_l d'une taille égale à 6 fois le nombre de marqueurs présents dans l'image l . Nous testons alors l'hypothèse nulle: "la population dont est issu Q_l a pour moyenne 0", et nous employons à nouveau les covariances estimées préalablement Σ_{mm}^{vert} et Σ_{mm}^{rose} .

	Nombre total d'images	Rejet de l'hypothèse nulle
Marqueurs verts	36	30
Marqueurs roses	36	36

TAB. 6.3 – Nombre d'images pour lesquelles l'hypothèse nulle a été rejetée. Dans le cas des marqueurs verts (resp. roses), un biais systématique de même direction est détecté dans 84% (100%) des images testées.

Les résultats obtenus (dans le tableau 6.3) sont éloquentes puisque l'hypothèse nulle est presque systématiquement rejetée quel que soit le type de marqueur considéré. Cela signifie qu'il y a un biais détecté presque à chaque fois avec une confiance à 99%. Ce résultat, qui paraît étonnant, s'explique par la puissance élevée de notre test statistique²¹ due à une taille d'échantillon importante (entre 60 et 100 par image). Le test que nous menons est capable de déceler un biais d'une valeur inférieure au pixel. En l'occurrence, il suffit qu'une faible proportion de marqueurs dans l'image vidéo soit inclinée selon une direction à peu près équivalente pour qu'un biais d'une amplitude d'un pixel soit détecté et que l'hypothèse nulle soit rejetée. Nous avons effectivement constaté que l'hypothèse nulle était acceptée lorsque les marqueurs trop inclinés étaient enlevés de l'échantillon de départ. La remarque établie dans le paragraphe précédent, soulignant l'erreur d'interprétation des opérateurs dans le cas de marqueur incliné, suggère donc que l'extraction automatique fournit un résultat plus précis pour les marqueurs verts.

Nous pourrions tenter d'affiner la qualité de la détection en utilisant un modèle géométrique de nos marqueurs. Cependant, leur modélisation semble difficile puisque les marqueurs d'un même type ne sont pas parfaitement identiques. En effet, la radio-opacité des marqueurs verts est due à la présence d'une mousse verte qui n'est pas répartie de la même manière d'un marqueur à l'autre. Dans le cas des marqueurs roses, la position de la croix par rapport au socle n'est pas la même non plus d'un marqueur à l'autre.

6.1.2 Extraction dans les images TDM

Nous décrivons dans cette section la méthode employée afin d'extraire la position des marqueurs dans les images TDM en niveaux de gris. Nous utilisons une technique issue de la

21. Un test est dit puissant s'il est capable de détecter la présence d'un biais faible par rapport aux valeurs prises par l'échantillon.

morphologie mathématique qui consiste, en résumé, à appliquer un filtre de type “chapeau haut de forme” à l'image tridimensionnelle. Notre méthodologie, bien qu'équivalente à celle de Wang [WMFM96], ne suppose pas, comme lui, que les marqueurs sont éloignés de la surface de la peau. En effet, nos marqueurs sont collés sur la surface de la peau et ne sont pas sur l'extrémité d'un bâtonnet vissé sur le crâne. Par ailleurs, notre caractérisation des marqueurs potentiels ne nécessite pas forcément qu'ils soient très opaques aux rayons X.

Méthodologie

La première étape consiste à seuiller l'image TDM afin d'extraire le patient et les marqueurs. Par ailleurs, l'exécution de nos traitements sur des images binaires réduit les temps de calculs qui peuvent être considérables pour des images aussi volumineuses ($512 \times 512 \times 200$ voxels). Les scanners étant calibrés (en unités Hounsfield), une valeur de seuil indépendante du scanner peut-être choisie (cette valeur est typiquement au dessus de la densité de l'air : nous avons choisi -800 HU pour le mannequin et -300 HU pour un être humain). La figure A.1 montre le résultat du seuillage à partir de l'image initiale en niveau de gris.

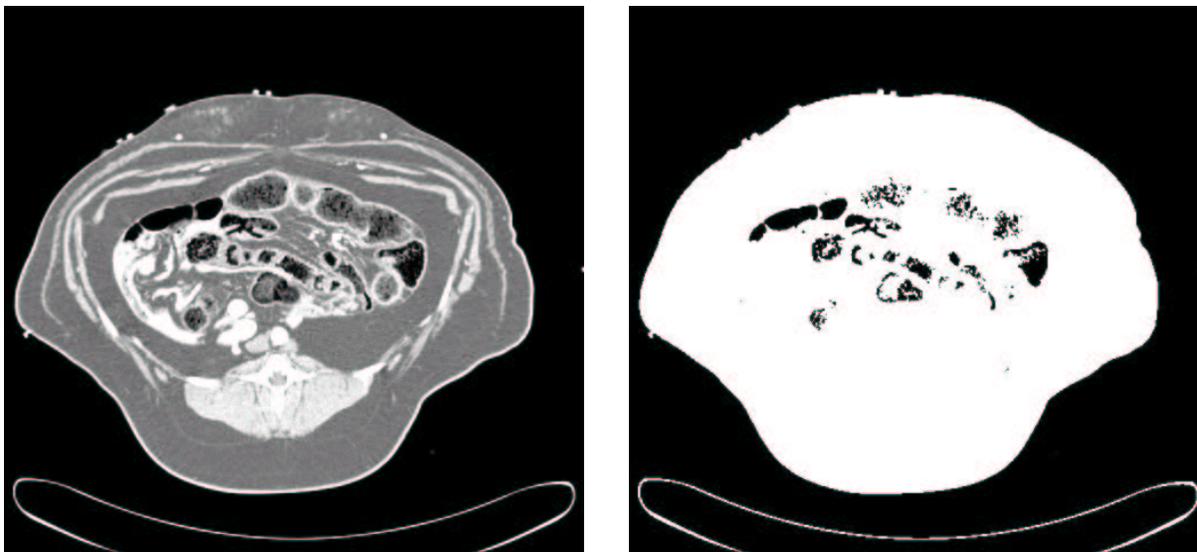


FIG. 6.7 – *Effet du seuillage sur une image originale d'un porc en niveaux de gris. Le seuil a été fixé à -300 HU.*

L'image binaire contient la représentation du corps du patient, et, bien souvent, de la table qui le soutient. Bien que le manipulateur radio puisse fenêtrer l'image du scanner autour du patient et effacer ainsi la présence de la table, certains artefacts peuvent apparaître (tubes, draps). Aussi, nous extrayons la composante connexe principale de l'image binaire afin de n'extraire que le corps du patient (cf. figure 6.8).

L'opération suivante consiste à “remplir” le patient (cf. figure 6.8). La technique classique est d'effectuer une recherche de la plus grande composante connexe sur les voxels noirs, caractérisant le vide autour du corps, et de fixer en blanc le reste de l'image (donc les zones gazeuses à l'intérieur du corps). Finalement, nous identifions les marqueurs potentiels en les effaçant par application d'une ouverture morphologique (cf. figure 6.8). En effet, il suffit alors de soustraire l'image précédente avec la nouvelle image sans marqueur pour les mettre en évidence. Le choix

de l'élément structurant est effectué en fonction de la taille des marqueurs. Typiquement, un élément structurant en forme de boule de rayon 3 fois supérieur au diamètre des marqueurs permet de les effacer de la surface de la peau. En général, cette opération fait aussi apparaître des artefacts, notamment une pelure de peau (cf. figure 6.8). Nous la faisons disparaître en appliquant une ouverture avec un élément structurant de très petite taille puis en triant les composantes connexes restantes en fonction de leur volume et de leur forme comme dans la section précédente. Finalement, nous obtenons les coordonnées du centre de chaque marqueur en calculant le barycentre pondéré \vec{G} de chaque composante connexe extraite afin de tenir compte des effets de volume partiel :

$$\vec{G} = \frac{\sum_{P \in Comp} I(\vec{P}) \cdot \vec{P}}{\sum_{P \in Comp} I(\vec{P})}$$

où \vec{P} est un point de la composante et $I()$ la fonction associant à un point de l'espace son intensité lumineuse.

Bien que cette méthode fonctionne avec nos deux types de marqueurs, nous avons remarqué que les marqueurs rouges sont bien plus radio-opaques. Il est alors possible de les segmenter à l'aide d'un seuillage d'une valeur supérieure à la peau et les os, suivi d'une caractérisation en fonction de leur dimension.

Nous mentionnons que nous aurions pu raffiner le calcul du centre en nous servant d'une méthode de type "template matching". Bien que ce traitement supplémentaire implique un temps de calcul additionnel, le fait que nous pouvons disposer d'une bonne initialisation du centre (grâce à la version actuelle de notre méthode) et de l'orientation du marqueur (en calculant sa matrice d'inertie) permettrait de l'exécuter en des temps raisonnables relativement courts par rapport au temps demandé par l'ensemble du traitement. En pratique, ce raffinement s'avère superflu puisque le gain en précision serait d'un ordre sub-voxellique, alors que la variabilité à prendre en compte est le mouvement corrélé des marqueurs, qui malgré la ventilation contrôlée des poumons du patient, est de l'ordre du millimètre (et donc d'environ 2 voxels).

Evaluation

Nous disposons de 3 images scanners du porc et 6 du mannequin sur lesquels ont été collés des marqueurs radio-opaques. Le tableau 6.4 indique les pourcentages de faux positifs et faux négatifs obtenus par notre algorithme d'extraction, ainsi que les temps de calcul. Le temps d'exécution est de 2 minutes pour les marqueurs verts et de 30 secondes pour les rouges. La différence est due au traitement simple effectué pour localiser les rouges (essentiellement un seuillage). Dans les deux cas, ces valeurs ne sont pas prohibitives pour notre application puisque cette procédure n'est réalisée qu'une seule fois au début de l'intervention. Nous pensons que nous réduirons facilement sa durée en optimisant les traitements. Le nombre de faux positifs est très faible et correspond généralement à la présence des tétons ou de tuyaux le long du corps. Une caractérisation supplémentaire en fonction des niveaux de gris permettrait éventuellement d'éliminer des objets artefacts, mais pas les tétons qui ont la même radio-opacité que les marqueurs verts. Néanmoins, la présence de quelques faux positifs ne nuit pas à l'ensemble du traitement global rendu robuste lors de la phase d'appariement.

Le pourcentage de faux négatifs est sensiblement différent d'un marqueur à l'autre. Cette différence provient du fait que les marqueurs verts contiennent une moisissure (radio-opaque) qui, une fois à la température ambiante, s'assèche progressivement (en 2 semaines). Ils deviennent donc beaucoup plus fins dans l'image scanner et sont donc éliminés par l'algorithme

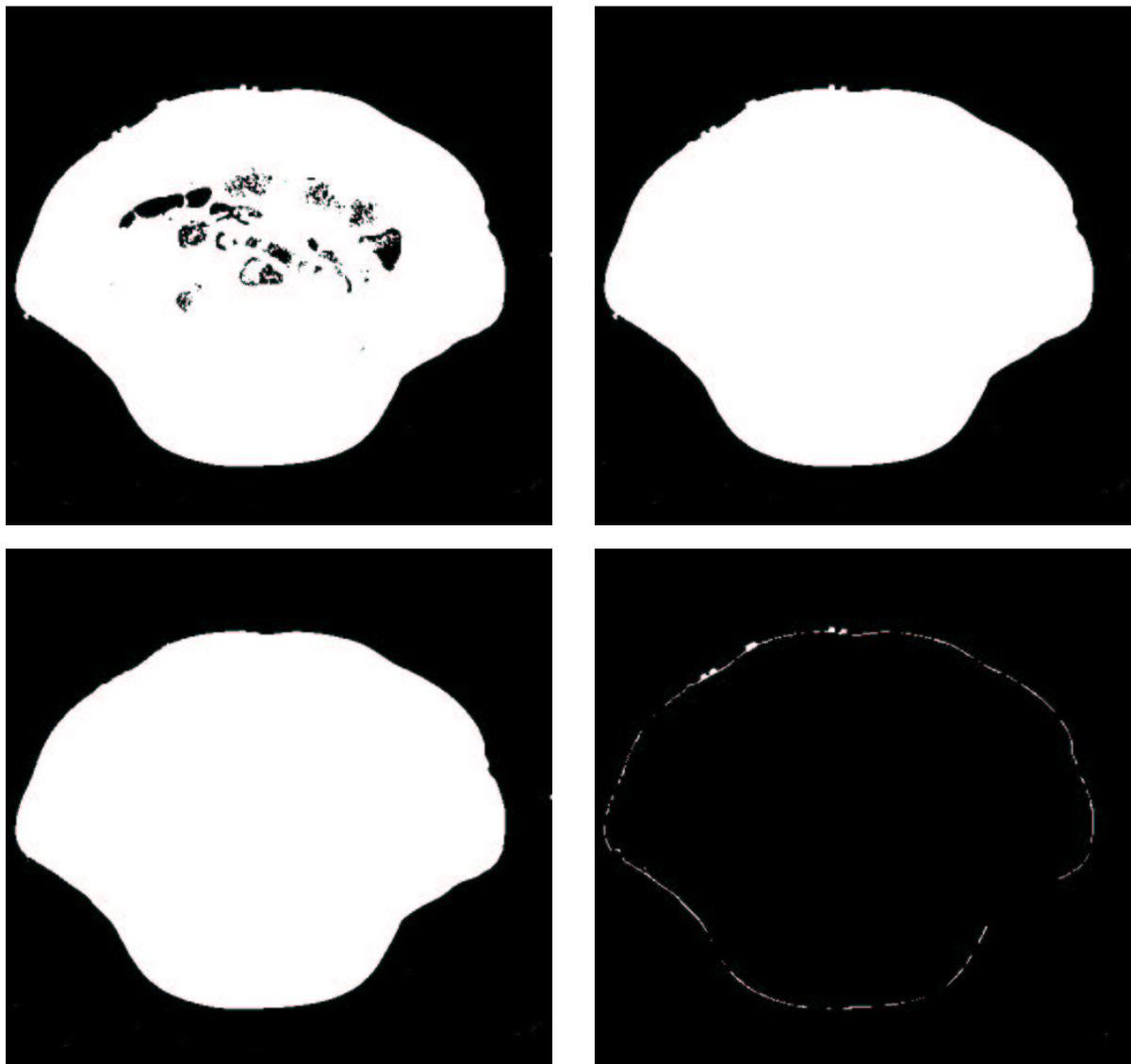


FIG. 6.8 – *En haut à gauche* : extraction de la composante principale de l'image, la table du scanner a disparu. *En haut à droite* : remplissage en blanc de la composante principale. *En bas à gauche* : application d'une ouverture avec un élément structurant en forme de boule de rayon 3 fois supérieurs à la taille des marqueurs. Les marqueurs ont été effacés. *En bas à droite* : soustraction de l'image en haut à droite avec l'image en bas à gauche. Les marqueurs sont mis en évidence. Une pelure de peau est aussi extraite. Elle est traitée par une ouverture suivie d'une caractérisation des composantes connexes en fonction de leur volume et de leur forme.

6.2. Mise en correspondances des marqueurs extraits

	Nombre de marqueurs	Faux positifs	Faux négatifs	Temps de calcul moyen
Marqueurs verts	260	2.3%	4.6%	117 sec.
Marqueurs roses	80	1.25%	1.25%	38 sec.
Total	340	2.05%	3.8%	98.4 sec.

TAB. 6.4 – *Pourcentage de faux positifs et faux négatifs obtenus après application de notre algorithme d'extraction sur les images scanner. Ces résultats et le temps d'exécution maximal sur l'ensemble des images sont donnés en fonction du type de marqueur.*

d'extraction lors de la caractérisation par le volume des composantes connexes. En pratique, les marqueurs sont conservés dans un réfrigérateur jusqu'au moment de l'opération, ce phénomène ne se produira donc pas en conditions cliniques. Les 2 types de marqueurs devraient donc fournir des performances équivalentes quant au nombre de faux positifs et faux négatifs. En revanche, d'un point de vue temps de calcul, les marqueurs roses sont bien plus avantageux puisque la procédure est 3 fois plus rapide. Les premiers essais en conditions cliniques permettront de choisir définitivement lesquelles des marqueurs choisir.

Une validation statistique de notre algorithme d'extraction dans les images TDM (de même type que celle réalisée pour l'algorithme d'extraction vidéo) est en cours.

6.2 Mise en correspondances des marqueurs extraits

Dans cette section nous nous attaquons au problème de l'appariement des marqueurs TDM 3D et vidéo 2D. Une technique serait d'apparier, avec un algorithme de prédiction/vérification [AF86] par exemple, les marqueurs extraits de l'image TDM avec ceux d'une image vidéo sans se servir de l'information potentielle contenue dans la seconde image. La mise en correspondance de points 3D avec des points 2D d'une image vidéo est une tâche bien souvent difficile et très coûteuse en temps de calcul en l'absence d'information a priori. En effet, la transformation projective ne conserve ni les distances, ni les angles, ni leur rapport²². Nos deux caméras, étant calibrées dans un même repère, il semble bien plus raisonnable d'apparier préalablement les marqueurs dans les deux images vidéo en nous servant de la contrainte épipolaire, de les reconstruire dans le repère des caméras et de procéder à un appariement de type prédiction/vérification entre deux ensembles de points 3D en utilisant des contraintes métriques et/ou angulaires.

Certains marqueurs ne seront pas visibles dans les deux images vidéo et d'autres ne seront pas appariés malgré la contrainte épipolaire pour des raisons d'ambiguïté. Bien qu'ils ne puissent être reconstruits, ces points participeront à l'étape suivante d'appariement 3D/3D, non pas durant la phase de prédiction, mais durant celle de vérification. Ces points, écartés à la fin de la phase d'appariement 2D/2D, seront donc reconsidérés et finalement appariés lorsque la prédiction sera correcte.

Nous verrons lors des résultats expérimentaux que cette approche est très robuste et qu'elle réalise la mise en correspondance de tous les appariements potentiels. Cette caractéristique est fondamentale pour notre application puisque la précision du recalage obtenu après optimisation de l'EPPC dépend du nombre d'appariement correct. Par ailleurs, l'étape de prédiction

²². La géométrie projective conserve le birapport. Néanmoins, cette propriété nous est inutile puisqu'elle nécessite l'alignement spatial de 4 marqueurs, ce qui ne sera pratiquement jamais notre cas.

fournit une première estimation grossière de la transformation caméra/scanner qui assure lors de l'optimisation de l'EPPC une convergence rapide vers le minimum global.

La présente section se divise donc logiquement en deux parties traitant séparément de deux problèmes d'appariements différents. Nous présentons en premier lieu la méthode basée sur la géométrie épipolaire permettant de mettre en correspondance deux points provenant de deux images vidéo et correspondant au même point de l'espace. En second lieu, nous décrivons notre algorithme de prédiction/vérification qui réalise la mise en correspondance entre les points 3D issus de la reconstruction stéréoscopique dans le repère des caméras et ceux provenant du modèle scanner. Les résultats obtenus sur des données réelles provenant du mannequin montrent que nos deux algorithmes sont suffisamment rapides (procédure totale en dessous de 5 secondes) et que la liste définitive des appariements est exempt d'erreur et contient tous les appariements qui pouvaient être réalisés.

6.2.1 Mise en correspondances des données vidéos 2D

La contrainte épipolaire

Le problème de la mise en correspondance de données ponctuelles dans des images vidéo provenant de caméras calibrées dans le même repère est largement traité dans la littérature [Fau93, HM95, Aya89, Aya91]. Sa résolution est basée sur une considération géométrique, communément appelée géométrie épipolaire, de la position des caméras, de leur centre optique et de leur plan focal.

Considérons deux caméras fixant un même point M dans l'espace. Appelons C et C' leurs deux centres optiques respectifs, et m la projection de M sur le plan image P associé à la caméra de centre optique C . Nous souhaitons déterminer quel marqueur m' du plan image P' est apparié à m . Nous savons que M appartient à la demi-droite $[mC)$, et notre connaissance a priori de la position relative de la caméra et du patient nous permet de fixer une valeur minimale M_{min} et maximale M_{max} à sa profondeur. La projection du lieu d'existence de M , i.e. le segment $[M_{min}M_{max}]$, sur le plan image P' permet de définir un segment $[m'_{min}m'_{max}]$ sur lequel doit se situer le correspondant m' de m (cf. figure 6.9).

Il est donc possible d'associer à chaque marqueur de l'image P un segment épipolaire dans l'image P' sur lequel se trouve son correspondant. Le calibrage n'étant pas parfait et l'extraction des marqueurs étant bruitée, le correspondant n'appartient pas réellement au segment épipolaire. Nous sélectionnons alors les points les plus proches dont la distance au segment est inférieure à un seuil d_{seuil} . La valeur de d_{seuil} à partir de laquelle un point est considéré comme proche du segment épipolaire est déterminée en fonction du bruit sur l'extraction 2D σ_{2D} , des distances caméra/objet et des distances focales f_i . En effet, pour σ_{2D} et f_i fixés, plus la première caméra est éloignée du patient, plus d_{seuil} doit être faible lors de la recherche d'un correspondant dans l'image de la seconde caméra. Dans notre cas de figure, nous avons estimé que l'écart-type *dans le plan image* de notre extraction automatique valait environ 1.5 pixels, ce qui correspond à une erreur maximale de l'ordre de $3 \times 1.5 = 4.5$ pixels (règle des trois sigmas). Comme les caméras jouent un rôle approximativement symétrique dans notre application, une valeur conservatrice pour d_{seuil} est $2 \times 4.5 = 9$ pixels. S'il n'y a qu'un seul point proche de la droite épipolaire, l'appariement est considéré comme fiable, sinon il est ambigu (cette configuration, illustrée sur la figure 6.10, apparaît quand les deux marqueurs radio-opaques forment une droite parallèle à (CC')). Nous essayons alors de lever l'ambiguïté à l'aide de la contraintes d'ordre (illustrée sur la figure 6.11). Dans le cas où l'ambiguïté persiste (par exemple, dans le cas où 2 marqueurs dans la première image ont chacun 3 mêmes correspondants potentiels dans

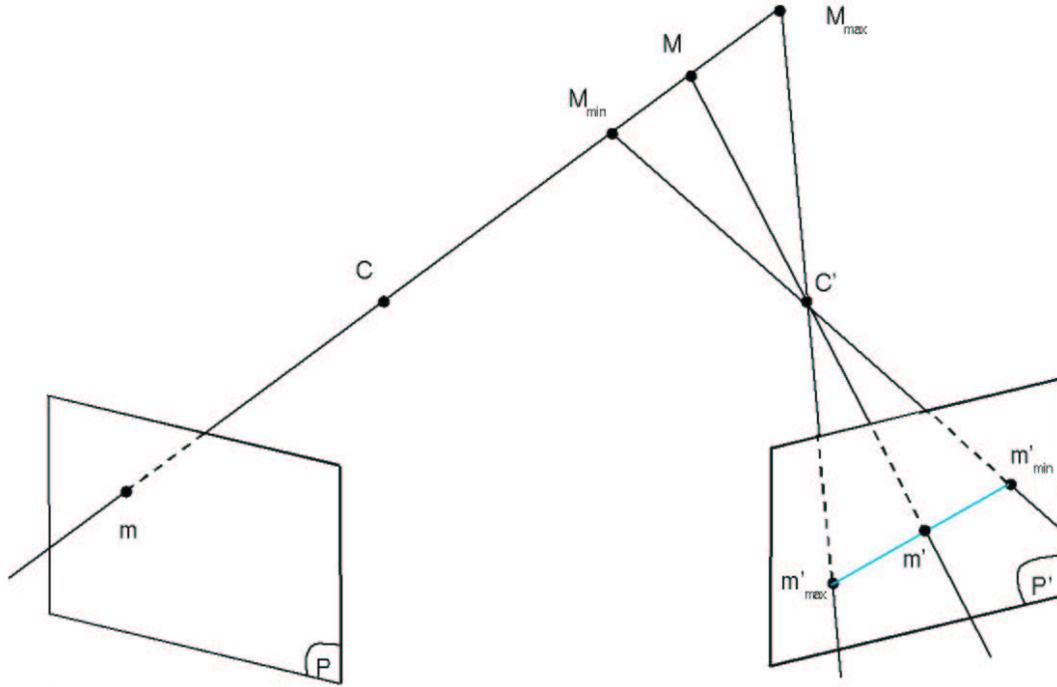


FIG. 6.9 – Illustration de la contrainte épipolaire. Etant donné un marqueur 2D m dans l'image vidéo P , nous recherchons son correspondant m' dans l'image vidéo P' . Le marqueur 3D M , dont la projection sur P est m , se situe sur la demi-droite $[mC)$. Connaissant l'intervalle de profondeur dans lequel se trouve le patient, nous pouvons définir un segment $[M_{min}M_{max}]$ dans $[mC)$ contenant M . La projection de ce segment sur l'image vidéo P' définit le segment $[m'_{min}m'_{max}]$ sur lequel se situe m' .

l'autre image), les marqueurs ne sont pas appariés.

Les points appariés sont finalement reconstruits en 3D par optimisation sur les coordonnées de M du critère classique :

$$C_{rec} = \| P_1(M) - m_1 \|^2 + \| P_2(M) - m_2 \|^2$$

où M est initialisé par un calcul direct ne tenant pas compte de la distorsion. Comme indiqué dans l'introduction de la section, les points non-appariés seront reconsidérés lors de l'étape d'appariement suivante.

Résultats Nous avons réalisé 240 paires d'images de notre mannequin avec deux caméras conjointement calibrées. Les caméras ont été déplacées et recalibrées 60 fois. Pour chaque position fixe des caméras, 4 acquisitions du mannequin dans des positions différentes ont été réalisées. Les marqueurs ont été relevés automatiquement et appariés interactivement quand cela était possible afin de disposer d'une mesure étalon à laquelle comparer le résultat de notre algorithme. Les résultats indiquent que notre algorithme est très efficace puisque, sur une moyenne de 24.8 appariements possibles, 21.3 ont été correctement retrouvés, 2.5 ont été considérés comme ambigus et seulement 1 était erroné. Pour que l'étape d'appariement 3D/3D conduise à une prédiction incorrecte, il faut que moins de 3 appariements soient justes. Nous

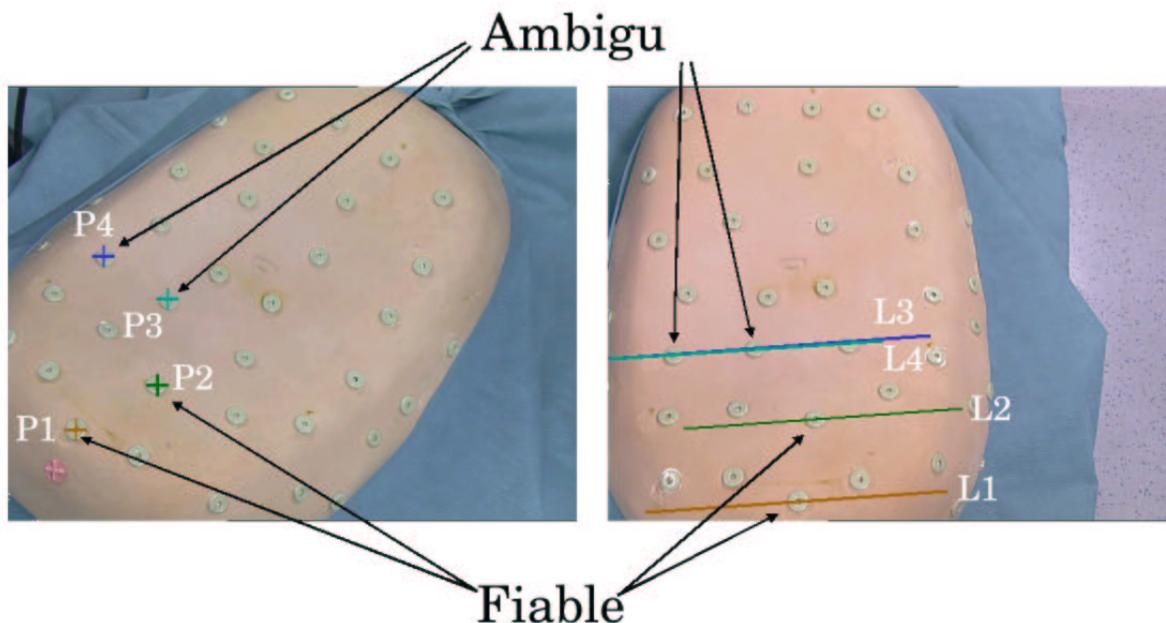


FIG. 6.10 – Exemple d'appariements épipolaires sur une paire d'image vidéo du mannequin. Les segments épipolaires L_i tracés dans l'image droite correspondent aux marqueurs P_i de l'image gauche. On peut remarquer qu'il n'y a qu'un seul marqueur sur les segments L_1 et L_2 (en bas). Dans ce cas, l'appariement est considéré comme fiable. Les segments L_3 et L_4 étant très proches de 3 marqueurs différents, ce cas est ambigu. Si la contrainte d'ordre ne permet pas de lever l'ambiguïté, ces marqueurs ne seront pas appariés.

avons donc vérifié expérimentalement que notre algorithme d'appariement vidéo fournissait à chaque fois un nombre d'appariements corrects suffisant. Cette procédure de mise en correspondance nécessite en moyenne moins de 1 ms.

6.2.2 Mise en correspondances des données 3D et 2D

Méthode Nous disposons d'un ensemble de marqueurs reconstruits dans le repère des caméras, que nous souhaitons appairer avec les marqueurs correspondants extraits de l'image scanner. Nous avons décidé d'utiliser un algorithme de prédiction/vérification [AF86] contraints par les distances entre les points. Cet algorithme nous fournit non seulement une estimation des appariements mais aussi une estimation de la transformation rigide liant le repère des caméras et scanner. Le principe est de choisir un triplet dans l'ensemble des marqueurs reconstruits et de chercher un triplet dans l'ensemble des marqueurs extraits de l'image scanner qui possède les mêmes dimensions (les côtés des triangles ont approximativement la même longueur). Si la recherche ne fournit aucun résultat, un nouveau triplet est choisi dans l'ensemble des points reconstruits. Lorsqu'un triplet de dimensions équivalentes a été trouvé, nous calculons la transformation rigide T_{estim} qui les recalcule en minimisant le critère classique LSQ 3D/3D [AHB87] (cf. figure 6.12).

Afin de vérifier cette prédiction, nous appliquons T_{estim} à tous les marqueurs extraits de l'image scanner et les projetons dans les 2 images vidéo. Comme l'algorithme d'extraction dans l'image TDM (resp. dans les images vidéo) fournit, dans le pire des cas, 2 (resp. 1) faux négatifs

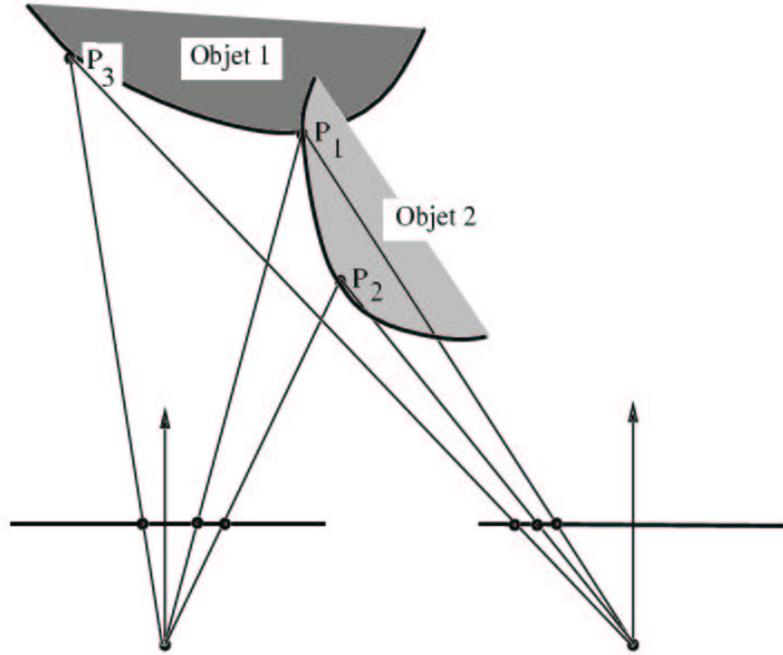


FIG. 6.11 – Lorsque deux points se trouvent sur la surface d'un objet opaque peu incliné par rapport aux axes optiques des deux caméras, l'ordre de leurs projections est le même dans les deux images vidéo (cas de l'objet 1). Dans le cas contraire (cas de l'objet 2) les projections ne sont pas vues simultanément dans les deux images.

si 20 marqueurs ont été collés, le nombre minimum de marqueurs vidéo extrait proches de leur correspondant reprojecté vaut $20 - (2 + 1) = 17$. Nous acceptons donc la transformation si, pour chaque image vidéo, plus de 85% des marqueurs vidéo extraits sont à une distance d'un point projeté inférieure à $d_{seuil3D}$ (ce cas est illustré en haut de la figure 6.13). Pour notre application, nous avons choisi de fixer $d_{seuil3D}$ à la moitié de la distance moyenne d'un point à son plus proche voisin dans l'image vidéo :

$$d_{seuil3D} = \frac{1}{2 \cdot N} \cdot \sum_{i=1}^N \min_{j \neq i} (d(m_i, m_j))$$

où $d(m_i, m_j)$ correspond à la distance entre les marqueurs i et j extraits d'une image vidéo. Dans le cas contraire (illustré en bas de la figure 6.13), les 2 triplets ne sont pas homologues malgré la correspondance de leur dimension : nous continuons notre recherche et choisissons un nouveau triplet dans l'ensemble des marqueurs issus de l'image scanner. En fait, afin de limiter les temps de calcul, il est plus intéressant de trier préalablement les triplets dans l'ordre décroissant de leur dimension. Ainsi, les premiers triplets choisis sont ceux dont le nombre potentiel de triplet correspondant est le plus faible.

Lorsqu'une transformation correcte est trouvée, toutes les correspondances 3D/2D établies durant la phase de vérification sont conservées et sont ensuite utilisées pour raffiner T_{estim} lors de l'optimisation de l'EPPC. Les marqueurs qui faisaient partie des appariements vidéo ambigus sont donc implicitement reconsidérés grâce à la vérification.

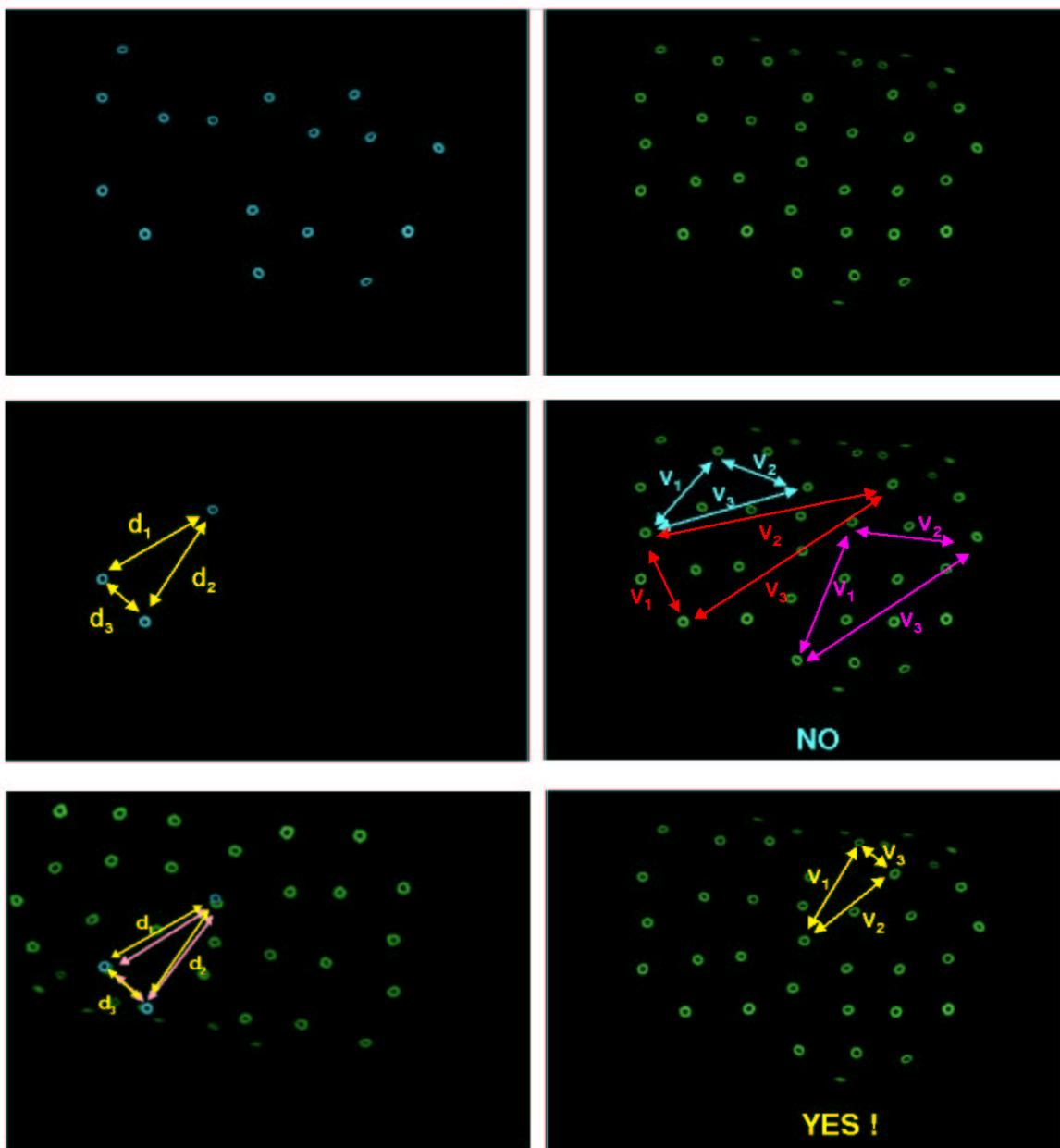


FIG. 6.12 – Illustration de la recherche de deux triplets de même dimension. En haut à gauche : ensemble des marqueurs reconstruits dans le repère des caméras après l'appariement vidéo. En haut à droite : ensemble des marqueurs extraits de l'image scanner. Au milieu à gauche : un triplet de marqueur est choisi parmi les points reconstruits. Au milieu à droite : recherche d'un triplet de dimension équivalente. Les exemples proposés sont tous rejetés. En bas à droite : finalement un triplet de dimension similaire a été trouvé. En bas à gauche : la transformation rigide 3D/3D reliant ces 2 triplets est calculée, puis appliquée aux autres marqueurs extraits de l'image scanner pour l'étape de vérification.

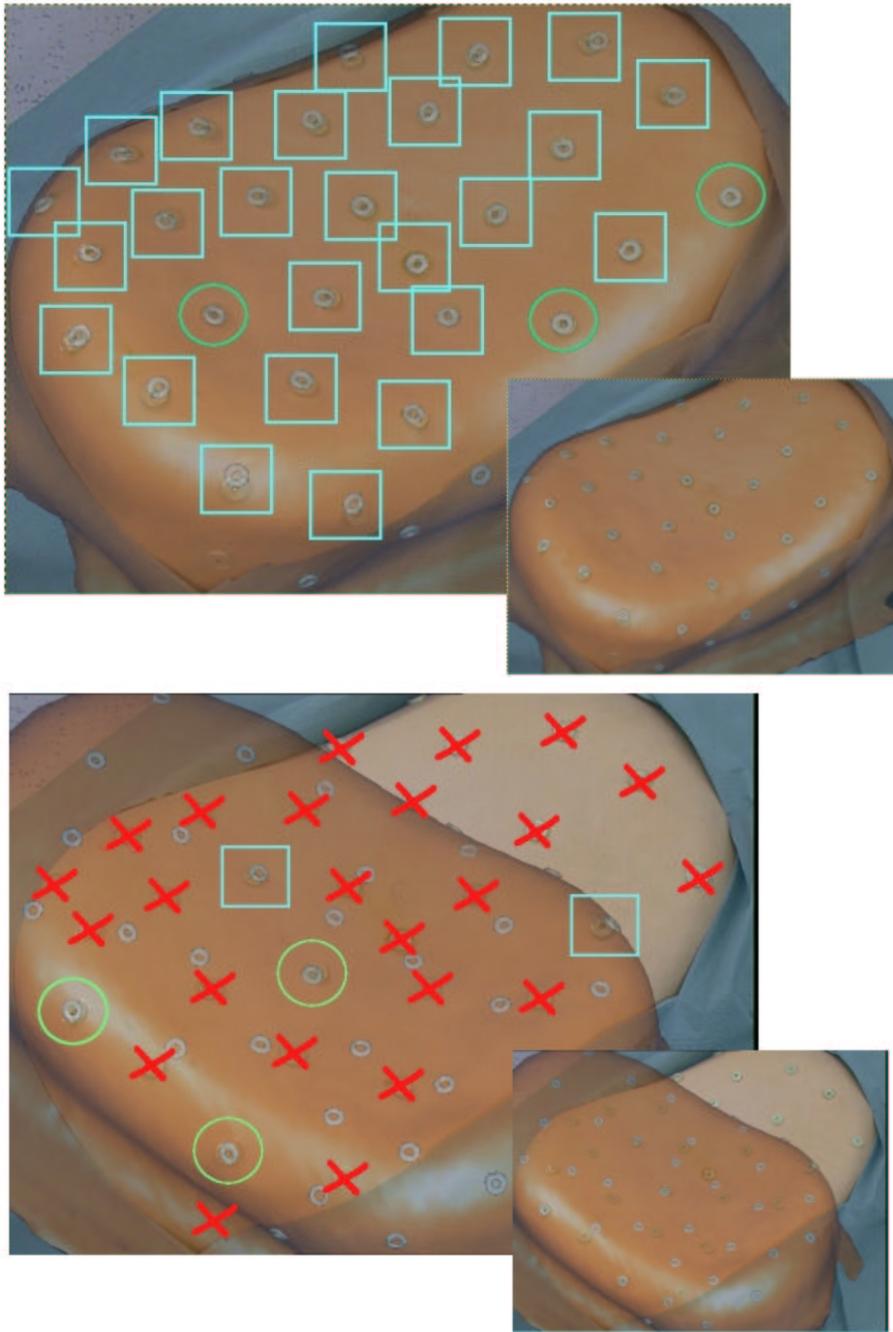


FIG. 6.13 – *En haut*: illustration de l'étape de vérification dans le cas où les 2 triplets sont homologues. Médaille en bas à droite: le modèle virtuel issu du scanner est reprojété en semi-transparente dans l'image vidéo. Image centrale: les marqueurs du triplet apparié sont entourés d'un cercle vert. Les marqueurs 3D de l'image TDM reprojétés dans l'image vidéo sont entourés d'un carré bleu lorsqu'ils sont appariés à un marqueur 2D extrait de l'image vidéo. Plus de 85% des marqueurs extraits des images vidéo ont été appariés, la transformation estimée est acceptée. *En bas*: illustration de l'étape de vérification dans le cas où les 2 triplets ne sont pas homologues. Les marqueurs 2D de l'image vidéo sont entourés d'un carré bleu lorsqu'ils sont appariés à un marqueur 3D extrait de l'image TDM et barré avec une croix rouge s'il n'y a pas de correspondant. Seulement 2 marqueurs ont été appariés en plus de ceux du triplet. Plus de la moitié des marqueurs de l'image TDM sont sans correspondant. La transformation estimée est rejetée.

Evaluation de l'ensemble de la boucle automatique

Nous avons extrait et apparié automatiquement les marqueurs visibles dans les 240 paires d'images du mannequin de la section précédente. L'algorithme de prédiction/vérification a été évalué sur les reconstructions stéréoscopiques des points vidéo appariés et les marqueurs extraits automatiquement des différentes acquisitions scanner du mannequin. Pour chaque expérience, notre algorithme fournit une transformation globalement correcte et retrouve la totalité des appariements corrects possibles. Le temps de calcul nécessaire est en moyenne inférieur à 2 sec. pour une valeur maximale de l'ordre de 3 sec. Les appariements réalisés serviront à l'étape suivante dans laquelle nous raffinerons la transformation avec le critère EPPC développé dans le chapitre 4.

6.3 Suivi des marqueurs en temps réel

Le patient restant immobile durant l'intervention, les marqueurs sur son torse sont quasiment immobiles dans les images vidéo (aux mouvements respiratoires près). Une recherche exhaustive dans l'image n'est donc plus nécessaire. Il suffit d'effectuer la recherche des marqueurs dans des imagerie entourant les marqueurs localisés lors de l'itération précédente, et dont la taille est choisie en fonction de la dimension des marqueurs dans l'image vidéo (carré de côté 3 fois supérieur à leur diamètre). Les appariements ayant été déterminés lors de la première itération grâce à l'algorithme de prédiction/vérification, il n'est pas nécessaire d'effectuer une nouvelle mise en correspondance. En effet, nous supposons évidemment que le marqueur dont nous cherchons la position dans l'imagerie correspond à celui de départ. Par conséquent, la seule étape nécessaire correspond à l'extraction des marqueurs dans les imagerie dont le temps d'exécution est très faible (inférieur à 20 ms) puisque la somme des surfaces des imagerie représente moins de 30% de l'image entière et la détermination du pic de la peau et de ses seuils a déjà été réalisée. Un suivi en temps réel est donc possible.

6.4 Conclusion

La motivation de ce chapitre était axée sur le temps nécessaire au traitement des données utilisées pour le calcul de la transformation rigide caméra/scanner. Dans le domaine médical, les pointages manuels sont réalisés en routine clinique. Or, une réponse interactive au problème de l'extraction et de l'appariement des marqueurs radio-opaques est opérateur-dépendant et nécessite une durée proportionnelle à leur nombre. Par exemple, [Mou03] et [LYT⁺02] localisent manuellement 10 et 5 de nos marqueurs verts radio-opaques dans une image TDM en 2 et 1 minutes. Dans [Mou03], la localisation dans l'espace est réalisée avec un bras-robot (DaVinci©) dont les articulations sont libérées, alors que dans [LYT⁺02], les marqueurs sont localisés à l'aide d'une sonde, qui est pointée successivement au centre de chaque marqueur par le praticien. Dans notre application, l'extraction et l'appariement d'une vingtaine de marqueurs nécessite environ 10 minutes à un opérateur expérimenté.

L'intérêt de notre système réside non seulement dans l'apport de précision, mais aussi dans la réduction potentielle de la durée d'une intervention de radiologie interventionnelle, nous avons présenté nos algorithmes d'extraction et d'appariement automatique pour chacun desquels une évaluation sur des données réelles a été réalisée. Nous avons ainsi pu prouver leur efficacité et leur robustesse par rapport à un traitement manuel. Bien qu'une localisation interactive soit exempt de faux négatif et positif (à moins d'une erreur d'interprétation humaine), nous

montrons que nos algorithmes réalisent plus de 95% des appariements potentiels en des délais bien plus bref (au moins 4 fois plus rapide), et permettent de réaliser un suivi des marqueurs en temps réel dans les images vidéo une fois l'initialisation réalisée (pour des mouvements de faible amplitude). En cas de mouvement du patient, une relocalisation n'est donc pas nécessaire. De même, si les 2 caméras sont rigidement liées, il est alors possible de déplacer les caméras afin de changer de point de vue sans traitement supplémentaire.

Par ailleurs, nos algorithmes résolvent automatiquement un problème d'extraction et d'appariement d'un type de marqueurs qui est très couramment employé dans des domaines médicaux adjacents [VG99, HE98, WNvdBS⁺01, DAP⁺99, PPR⁺03, NAM⁺98, SKR⁺97, BKSW93]. Ils pourraient donc être utilisés par de nombreuses personnes qui travaillent dans des conditions expérimentales similaires aux nôtres, mais qui continuent de localiser et d'apparier manuellement les marqueurs qu'ils utilisent. Le travail réalisé dans ce chapitre constitue donc une contribution pratique pour le domaine clinique.

Chapitre 7

Assurer la précision du système

A la fin du chapitre 6, nous proposons un système de guidage par réalité augmentée dont l'extraction et le traitement des données sont réalisés de manière automatique. La précision du recalage des organes internes a été évalué en moyenne à 2 mm dans des conditions proches de la réalité en intervention clinique. Nous avons aussi évalué à 0.95 mm l'erreur moyenne de repositionnement de l'aiguille virtuelle sur l'aiguille réelle. En général, l'étape suivante pour un système de guidage qui a été évalué et validé *in vitro* est son introduction et sa validation en bloc opératoire [SFB94, GKO⁺00, Mou03]. Nous mettons l'accent sur le fait que les études de précision sont parfois menées dans des conditions qui ne sont pas forcément reproductibles en salle d'opération. Typiquement, il existe de nombreuses études de précision qui sont dépendantes des dimensions de l'aiguille utilisée, du système de repérage employé, de la forme de l'objet repéré [HED⁺99, DAP⁺99, CL98]. La précision de ces systèmes n'est donc assurée que dans des conditions d'utilisation plutôt réduites. En l'occurrence, la précision de notre système n'est pas constante et dépend de nombreux paramètres pouvant varier d'une intervention à l'autre (angle entre les caméras, nombre de marqueurs, focales des caméras, courbure de la peau du patient, position relative de la cible par rapport aux marqueurs radio-opaques, dimension de l'aiguille). Par exemple, nous avons vu dans le chapitre 4 sur l'évaluation de l'EPPC que la précision peut diminuer d'un facteur 2 quand l'angle entre les caméras passe de 60° à 20°.

N'étant pas assuré de pouvoir systématiquement avoir des conditions optimales en salle d'opération, nous devrions joindre au système une étude statistique qui donnerait l'erreur de recalage de la cible (TRE) et l'erreur de repositionnement de l'aiguille en fonction des nombreux paramètres possibles. Il s'agirait de fournir l'équivalent de la posologie et des effets secondaires obligatoires pour tous les médicaments, et des tables de précision et de fiabilité pour les outils robotiques. Ces informations décrivent un éventail de conditions expérimentales pour lesquelles une caractéristique particulière (par exemple la précision) a été évaluée statistiquement. Sachant que la dimension des aiguilles peut varier d'une intervention à l'autre et que la configuration des marqueurs par rapport à la cible change aussi, il est impossible de fournir une table de précision en fonction de tous les paramètres. De plus, comme nous souhaitons un système fonctionnant en temps réel, les conditions peuvent changer durant l'intervention (occultation des marqueurs par les bras du chirurgien). L'évaluation de la précision doit donc être constamment mise à jour pour éviter un geste potentiellement dangereux durant l'intervention.

Afin de pouvoir quantifier la précision moyenne du système, nous proposons une analyse des perturbations sur la transformation estimée après optimisation des critères employés (ISPPC et EPPC). La connaissance de la covariance sur la transformation estimée permettrait après

propagation sur le point cible (ou sur la pointe de l'aiguille), de déduire la TRE moyenne et l'erreur de repositionnement de la pointe de l'aiguille.

Ce chapitre, consacré à l'étude de la prédiction de l'erreur, est divisé en 2 parties. La première partie est consacrée aux calculs conduisant à une formulation *analytique* de la covariance du point cible en fonction de la covariance sur les données 2D et 3D observées. Introduisant un formalisme matriciel pratique pour mener ce genre de calcul [Pen96], nous calculons, pour l'ISPPC puis pour l'EPCC la matrice de covariance après optimisation et obtenons un résultat analytique facilement implémentable.

En second lieu, nous nous attelons à la tâche primordiale de la validation de nos formules théoriques. En effet, cette phase de validation, bien souvent oubliée ou négligée, est absolument nécessaire puisque nos calculs théoriques sont menés sous des hypothèses qui ne seront pas toujours vérifiées dans la réalité (non-linéarité du critère, calibrage parfait, bruit gaussien non-biaisé sur l'extraction des données 3D et 2D). Afin de vérifier pas à pas que nos hypothèses ne sont pas abusives, nous réalisons une validation incrémentale sur des données synthétiques, avec des caméras réelles filmant un objet précisément défini, et finalement avec le mannequin et l'aiguille. A l'issue de la validation, nous prouvons que notre prédiction de la TRE est fiable dans des conditions expérimentales variées et réaliste, mais qu'il faudrait pousser la modélisation plus loin et/ou disposer d'un calibrage plus précis pour prédire correctement l'erreur de repositionnement de la pointe de l'aiguille.

7.1 Étude théorique de l'incertitude

Dans cette section, nous commençons par rappeler la théorie générale de la propagation d'incertitude au travers d'un critère d'optimisation. Ensuite, nous conformant au cadre méthodologique introduit par [PT97, PGT98], nous présentons pour l'ISPPC et l'EPCC les calculs aboutissant aux formulations analytiques de la matrice de covariance sur la transformation après optimisation du critère. Nous terminons par un résumé algorithmique permettant au lecteur d'implémenter rapidement les résultats obtenus.

7.1.1 Rappel sur la propagation de l'incertitude appliquée aux critères de maximisation

Nous rappelons ici brièvement les étapes de calcul qui conduisent à la matrice de covariance sur la transformation T à partir de la covariance sur les données et d'un critère faisant intervenir les données et T (pour plus de détail voir [Pen96]).

Supposons que le critère $C(\chi, T)$ soit une fonction de classe C^2 de $\mathbb{R}^m \times \mathbb{R}^p$ vers \mathbb{R}^+ , où χ est le vecteur m -dimensionnel de données, et T le vecteur p -dimensionnel des paramètres recherchés (dans notre cas la transformation). \hat{T} est la transformation optimale minimisant le critère : $\hat{T} = \arg \min_T (C(\chi, T))$.

Un minimum local est atteint et défini si et seulement si :

$$\Phi(\chi, T) = \left(\frac{\partial C}{\partial T}(\chi, T) \right)^\top = 0 \quad \text{et} \quad H = \frac{\partial^2 C}{\partial T^2}(\chi, T) \quad \text{défini positif.} \quad (7.1)$$

La fonction Φ définit alors de manière implicite \hat{T} comme une fonction dépendante des données χ . Une extension de Taylor au premier ordre de Φ nous donne :

$$\Phi(\chi + \delta\chi, T + \delta T) = \Phi(\chi, T) + \frac{\partial \Phi}{\partial \chi} \cdot \delta\chi + \frac{\partial \Phi}{\partial T} \cdot \delta T + O(\delta\chi^2, \delta T^2)$$

ce qui signifie qu'autour de la solution optimale \hat{T} nous avons :

$$\hat{T}(\chi + \delta\chi) = \hat{T}(\chi) - \left(\frac{\partial\Phi}{\partial T} \right)^{(-1)} \cdot \frac{\partial\Phi}{\partial\chi} \cdot \delta\chi + O(\delta\chi^2)$$

En conséquence, si χ est une variable aléatoire de moyenne $\bar{\chi}$ et de covariance $\Sigma_{\chi\chi}$, la transformation optimale \hat{T} est (au second ordre près), une variable aléatoire de moyenne et de covariance :

$$\bar{T} = \arg \min_T (C(\bar{\chi}, T)) \quad \text{et} \quad \Sigma_{TT} = H^{-1} \cdot \left(\frac{\partial\Phi}{\partial\chi} \right) \cdot \Sigma_{\chi\chi} \cdot \left(\frac{\partial\Phi}{\partial\chi} \right)^\top \cdot H^{-\top}$$

où H est le hessien $\frac{\partial^2 C(\chi, T)}{\partial T^2}$ de C et J_{Φ_x} son jacobien $\frac{\partial^2 C(\chi, T)}{\partial\chi\partial T}$. Il s'ensuit que pour propager la matrice de covariance des données aux paramètres de la transformation minimisant le critère, nous devons calculer H et J_{Φ_x} . Pour obtenir la matrice de covariance finale sur un point cible C_i après recalage, nous devons simplement propager l'incertitude à travers l'action de la transformation T :

$$\Sigma_{T \star C_i} = \frac{\partial(T \star C_i)}{\partial T} \cdot \Sigma_{TT} \cdot \frac{\partial(T \star C_i)}{\partial T}^\top \quad (7.2)$$

Afin d'alléger la lecture de ce chapitre, le détail des calculs est donné dans l'annexe C.1 pour le critère ISPPC et dans l'annexe C.2 pour l'EPPC.

Dans le cas de l'ISPPC, les étapes de calculs présentées ci-dessus permettent d'aboutir directement à la covariance sur la transformation rigide. En revanche, la procédure est légèrement différente pour l'EPPC. En effet, le vecteur des paramètres recherchés contient alors non seulement les paramètres de la transformation rigide, mais aussi les variables auxiliaires du critère i.e. les coordonnées 3D exactes inconnues (cf. section 4.1.3). Il est alors nécessaire d'extraire la variabilité sur la transformation de la variabilité sur l'ensemble des paramètres recherchés.

7.1.2 Résumé algorithmique

Nous obtenons une expression assez simple de Σ_{TT} pour les deux critères, en fonction d'un nombre de variables intermédiaires plutôt réduit. Nous résumons dans cette section les étapes nécessaires pour implémenter le calcul de Σ_{TT} .

Variables communes

$$\begin{aligned} D_i &= \frac{\partial(T \star M_i)}{\partial T} \\ m_i^{(l)} &= P^{(l)}(T \star M_i) = \frac{Q^{(l)} \cdot (T \star M_i) + b^{(l)}}{1 + C_{(l)}^\top \cdot (T \star M_i)} \\ L_i &= \sum_{l=1}^S \xi_i^l \cdot \frac{(Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)^\top \cdot (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)^2} \\ K_i &= L_i - \sum_{l=1}^S \xi_i^l \cdot \frac{C_{(l)} \cdot (m_i^{(l)} - \bar{m}_i^{(l)})^\top \cdot (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top) + (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)^\top \cdot (m_i^{(l)} - \bar{m}_i^{(l)}) \cdot C_{(l)}^\top}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)^2} \end{aligned}$$

Variabes intermédiaires pour le critère ISPPC Pour ce critère, nous rappelons que les données observées 3D sont considérées exacts, de telle sorte que $M_i = \tilde{M}_i$ est retrouvée dans les variables ci-dessous.

$$\begin{aligned}\Gamma &= \sum_{i=1}^N D_i^\top (\sigma_{3D}^2 \cdot K_i \cdot K_i + L_i) \cdot D_i \\ H &= \sum_{i=1}^N D_i^\top \cdot K_i \cdot D_i\end{aligned}$$

Variabes intermédiaires pour le critère EPPC Pour ce critère, les positions des points 3D exacts M_i et la transformation T sont estimées de manière alternée par la minimisation du critère. Les paramètres recherchés étant plus nombreux, nous obtenons une formule plus complexe pour Γ et H :

$$\begin{aligned}\Gamma &= \sum_{i=1}^N D_i^\top (\sigma_{3D}^2 \cdot Id + K_i^{-1})^{-1} \cdot (\sigma_{3D}^2 \cdot Id + K_i^{-1} \cdot L_i \cdot K_i^{-1}) \cdot (\sigma_{3D}^2 \cdot Id + K_i^{-1})^{-1} \cdot D_i \\ H &= \sum_{i=1}^N D_i^\top \cdot (\sigma_{3D}^2 \cdot Id + K_i^{-1})^{-1} \cdot D_i\end{aligned}$$

Estimation finale de l'incertitude Pour les deux critères, la matrice de covariance sur la transformation est finalement calculée comme suit :

$$\Sigma_{TT} = H^{-1} \cdot \Gamma \cdot H^{-1}$$

Équivalence de l'incertitude des deux critères pour des données 3D exacts Dans le cas limite où $\sigma_{3D} = 0$, nous savons que les 2 critères sont équivalents. Nous pouvons vérifier que leur incertitude est aussi égale. En effet, nous avons pour les deux critères les simplifications suivantes :

$$\begin{aligned}\Gamma(\sigma_{3D} = 0) &= \sum_{i=1}^N D_i^\top \cdot K_i \cdot (K_i^{-1} \cdot L_i \cdot K_i^{-1}) \cdot K_i \cdot D_i = \sum_{i=1}^N D_i^\top \cdot L_i \cdot D_i \\ H(\sigma_{3D} = 0) &= \sum_{i=1}^N D_i^\top \cdot K_i \cdot D_i\end{aligned}$$

7.2 Validation de la prédiction théorique

A l'aide des formules analytiques précédemment établies, nous sommes à même de prédire la covariance de la transformation après optimisation des critères ISPPC et EPPC. Néanmoins nous nous sommes servis de certaines approximations classiques (linéarisation du critère à l'optimum, calibrage supposé parfait sans distorsion) afin de dériver nos formules théoriques. Il est toutefois possible que ces approximations ne soient pas valables dans nos conditions expérimentales et que nous produisions systématiquement une surestimation ou une sous-estimation de la covariance réelle. Nous n'avons donc réalisé que la moitié du travail : nous devons maintenant valider les hypothèses utilisées (faibles non-linéarité du critère, calibrage parfait sans distorsion, bruit gaussien non-biaisé sur l'extraction des données). Dans cette section, nous vérifions incrémentalement que ces hypothèses sont justifiées dans nos conditions applicatives. Nous commençons notre validation avec des données synthétiques (pour les non-linéarité du

critère), puis avec des images vidéo réelles d'un objet 3D défini précisément (pour le calibrage des caméras et les distorsions) et finalement avec des données réelles provenant du mannequin et de l'aiguille (pour les hypothèses sur le bruit d'extraction des données).

7.2.1 Validation de la linéarisation

Le but de ce paragraphe est de vérifier que les non-linéarités du critère induisent des erreurs négligeables sur notre prédiction théorique de la TRE moyenne. Pour cela nous réalisons des expériences sur des données synthétiques construites comme cela est expliqué dans la section 4.2.1. Deux caméras synthétiques sont orientées vers N points 3D. La position relative de ces points est connue dans un repère différent de celui des caméras. Nous recherchons la transformation rigide recalant la projection 2D des N points 3D sur ceux visibles dans les deux images vidéo. Afin de simuler l'erreur de mesure sur les données 3D et 2D, nous les corrompons avec un bruit gaussien centré uniforme d'écart-type σ_{3D} et σ_{2D} . Les données synthétiques et les paramètres des caméras sont alors aléatoirement choisis afin de balayer toutes les valeurs envisageables dans la réalité :

- Le nombre de point est choisi uniformément entre 4 et 25.
- La position de chaque point est choisi parmi 40 possibles qui sont représentatives de la forme cylindrique de l'abdomen.
- l'angle entre les caméras est une variable uniforme avec une valeur minimale de 5° et maximale de 120° .
- Le rapport distance focale/distance caméra-objet varie entre 20 et 50.
- σ_{2D} et σ_{3D} sont choisis uniformément entre 0.5 et 4 (ce qui correspond à un rapport signal sur bruit (RSB) allant de 60 dB à 90 dB²³).

Notre prédiction d'erreur n'étant valide que sous l'hypothèse que l'optimisation a convergé vers le minimum global, nous initialisons avec une transformation suffisamment proche de la solution

Étant donné que chaque expérience est différente, nous devons calculer l'erreur relative de la TRE prédite (PTRE) par rapport à la TRE expérimentale (ETRE) pour mesurer quantitativement la qualité de la prédiction d'incertitude. Toutefois, comme le rapport $ETRE^2/PTRE^2$ suppose implicitement une distribution d'erreur isotrope, nous préférons nous servir de l'index de validation introduit dans [PT97]. En effet, ayant observé des anisotropies importantes, cet index permet de les prendre en compte puisqu'il pondère le vecteur d'erreur observé avec l'inverse de sa matrice de covariance prédite (nous obtenons ainsi une distance de Mahalanobis μ^2). Si l'erreur réelle, après recalage sur les points test, suit une loi de distribution gaussienne de moyenne nulle et de matrice de covariance celle que nous avons prédite, alors l'index de validation suit une loi du χ_3^2 . Disposant, à l'issue des expériences, de nombreuses valeurs indépendantes dans des configurations diverses, nous pouvons vérifier que l'index est vraiment distribué comme un χ_3^2 . Le test de Kolmogorov-Smirnov [PFTV91] (ou KS-test) est adapté pour faire cela, mais comme il ne fournit qu'une réponse binaire, nous nous servons aussi du fait que la valeur moyenne d'une distribution en χ_3^2 vaut 3 et sa variance 6.

Nous appelons index de validation la moyenne estimée des μ_i^2 :

$$I = \bar{\mu}^2 = \frac{1}{N} \sum_{i=0}^N \mu_i^2 \quad \text{et sa variance est calculée avec :} \quad \sigma_I^2 = \frac{1}{N-1} \sum_{i=0}^N (\mu_i^2 - \bar{\mu}^2)^2$$

23. $RSB_{dB} = 10 \log_{10}(\frac{\sigma_s}{\sigma_n})$ où σ_s (resp. σ_n) est la variance du signal (resp. bruit).

Cet index peut aussi être interprété comme un indicateur de surestimation ($I > 3$) ou de sous-estimation ($I < 3$) de l'erreur prédite sur le point cible. C'est une sorte d'erreur relative sur notre estimation de l'erreur.

Finalement nous calculons pour chaque recalage l'erreur prédite sur la transformation et la propageons sur 9 points test C_i . Cependant, les points test étant situés dans la même région, leur matrice de covariance sont hautement corrélées. Donc pour satisfaire la condition d'indépendance du K-S test, nous devons choisir 1 point seulement (parmi les 9) par expérience.

Les résultats

Les figures 7.1 à 7.5 montrent l'index de validation et le résultat du KS-test pour 5 paramètres variant indépendamment. Chaque valeur de l'index et du KS-test correspondant est calculée à l'issue de 1000 recalages. Elles montrent clairement que la prédiction de l'incertitude est toujours fiable. En effet, l'index de validation est toujours très proche de 3.0 et la valeur du KS-test ne descend jamais en dessous de 1%, ce qui signifie que la probabilité que le bruit sur la transformation ne se comporte pas comme la distribution que nous prédisons est négligeable. Nous pouvons juste faire remarquer que pour un nombre de point inférieur à 7, l'index est proche de 3.5, ce qui signifie que nous sous-estimons légèrement le bruit résultant sur la transformation.

Pour conclure cette étude préliminaire, le tableau 7.1 résume les statistiques obtenues avec 20000 recalages provenant d'une expérience dans laquelle tous les 5 paramètres varient aléatoirement. Les valeurs obtenues pour l'index de validation et le KS-test valident pleinement la fiabilité de la prédiction d'incertitude. En conclusion, nous avons démontré que l'influence des non-linéarités du critère sur notre prédiction est vraiment négligeable, et qu'en conséquence nous sommes à même de prédire correctement l'incertitude pourvu que les erreurs de calibrage soit faible et que l'erreur d'extraction sur les données suivent une loi gaussienne connue.

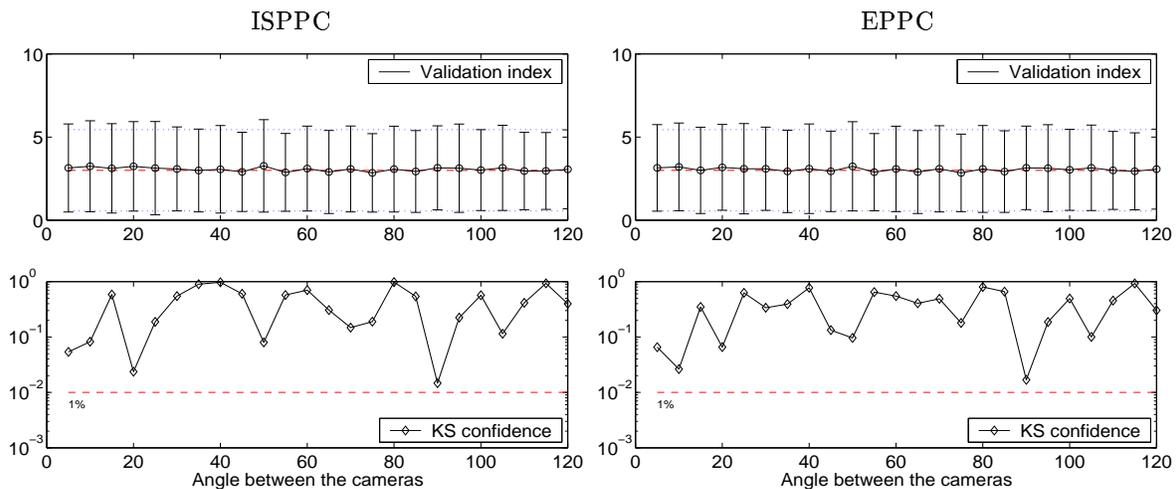


FIG. 7.1 – Validation de la prédiction de l'incertitude par rapport à l'angle entre les caméras. En haut : moyenne et variance de l'index de validation. En bas : valeur du KS-test. La prédiction est validée dans tous les cas avec une certitude à 1%.

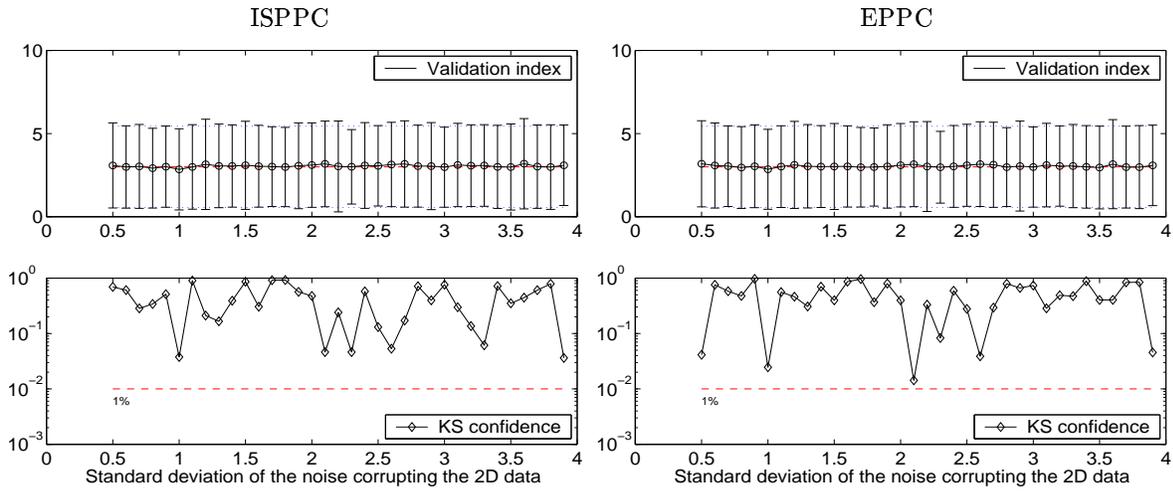


FIG. 7.2 – Validation de la prédiction de l'incertitude par rapport au bruit corrompant les données 2D (σ_{3D} fixé à 2 mm). En haut : moyenne et variance de l'index de validation. En bas : valeur du KS-test. La prédiction est validée dans tous les cas avec une certitude à 1%.

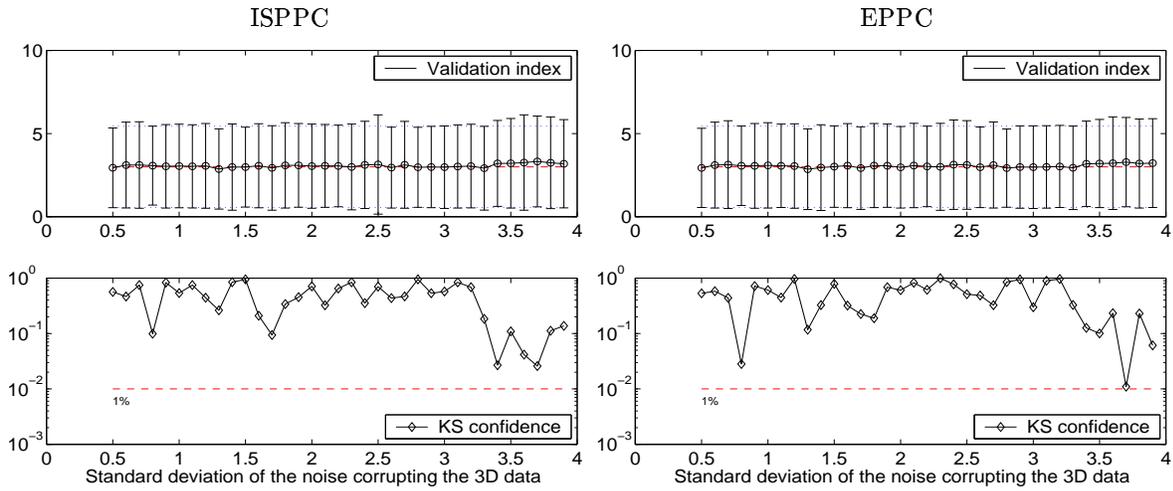


FIG. 7.3 – Validation de la prédiction de l'incertitude par rapport au bruit corrompant les données 3D (σ_{2D} fixé à 2 pixels). En haut : moyenne et variance de l'index de validation. En bas : valeur du KS-test. La prédiction est validée dans tous les cas avec une certitude à 1%.

	Index de validation μ^2 (3.0)	Ecart-type (2.449)	KS-confiance (0.01 < x < 1.0)
ISPPC	3.020	2.506	0.353
EPPC	3.016	2.486	0.647

TAB. 7.1 – Validation sur 20000 recalages de la prédiction d'incertitude lorsque tous les paramètres varient aléatoirement.

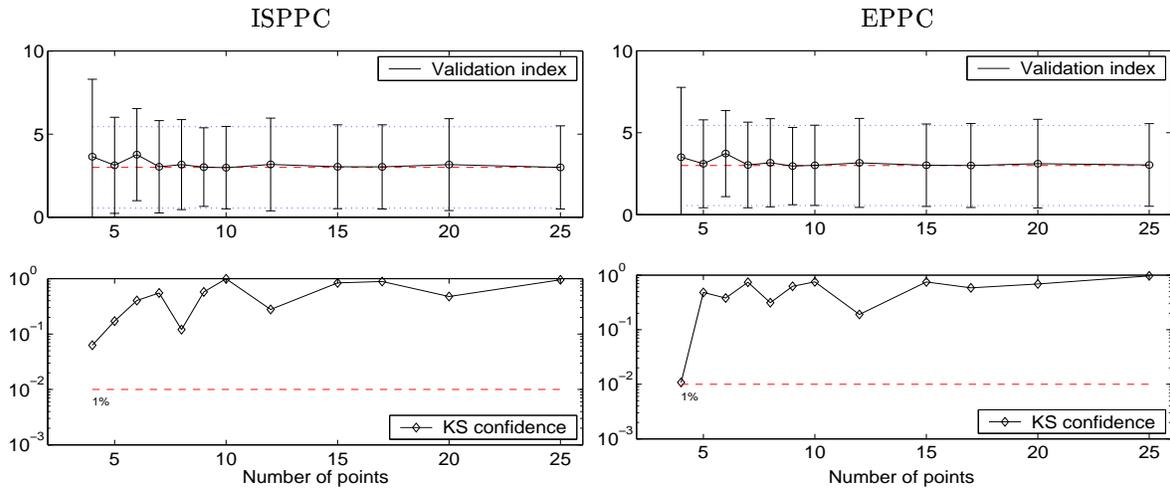


FIG. 7.4 – Validation de la prédiction de l'incertitude par rapport au nombre de points utilisés pour le recalage. En haut : moyenne et variance de l'index de validation. En bas : valeur du KS-test. La prédiction est validée dans tous les cas avec une certitude à 1%.

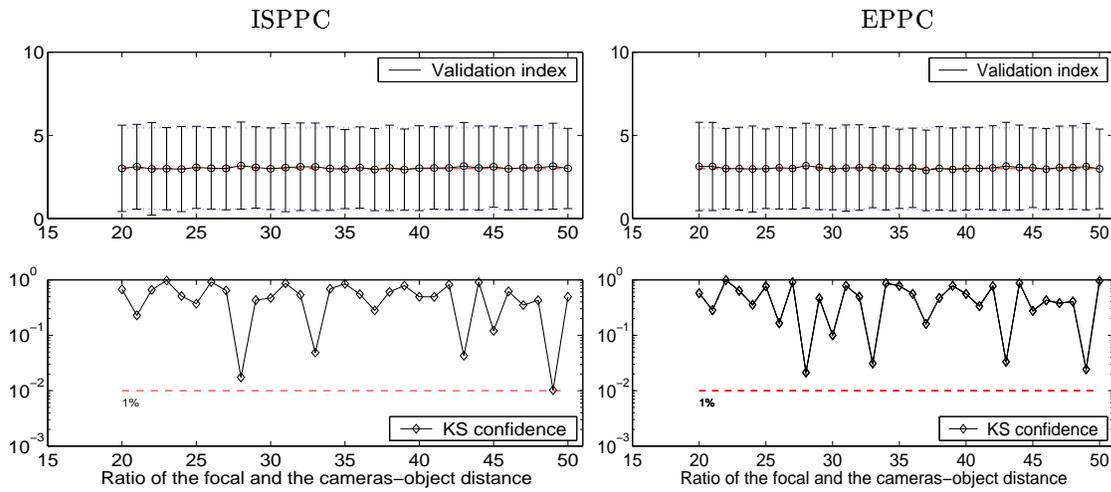


FIG. 7.5 – Validation de la prédiction de l'incertitude par rapport au rapport distance focale/distance caméras-objet. En haut : moyenne et variance de l'index de validation. En bas : valeur du KS-test. La prédiction est validée dans tous les cas avec une certitude à 1%.

7.2.2 Influence du calibrage

La validation de notre prédiction sur données synthétiques ne tient pas compte des erreurs de calibrage et exclut les distorsions vraisemblables du modèle sténopé utilisé pour dériver les formules théoriques. Le but des expériences présentées dans cette section est d'évaluer l'influence des erreurs de calibrage (et de sa sous modélisation) sur la qualité de notre prédiction d'erreur. Nous nous servons d'un damier imprimé (1200x1200 dpi) contenant 54 points d'intérêts (cf. figure 7.6) que nous pouvons localiser avec une grande précision ($\sigma_{3D} < 0.1$ mm, $\sigma_{2D} < 0.2$ pixel). Une telle précision est évidemment largement en dessous de celle obtenue pour les marqueurs radio-opaques, autant dans l'image CT ($\sigma_{3D} \simeq 1$ mm) que dans les images vidéo ($\sigma_{2D} \simeq 2$ pixels). En conséquence, pour simuler des conditions réalistes proches de notre application, nous rajoutons nous-mêmes un bruit gaussien centré sur les mesures 3D et 2D relevées. Cette astuce nous permet donc de valider notre prédiction en conditions réelles vis à vis des conditions d'acquisition, mais en restant sous l'hypothèse que le bruit d'extraction sur les données est gaussien, centré et que nous en avons une estimation réaliste.

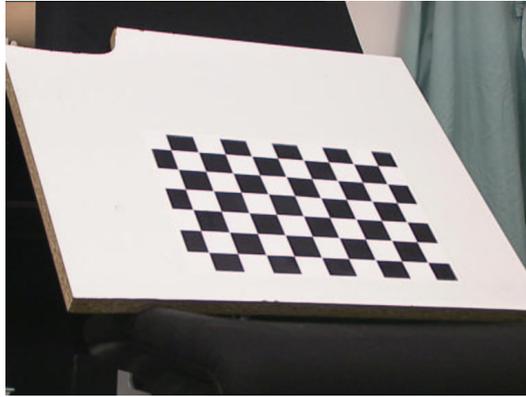


FIG. 7.6 – Damier utilisé pour le recalage 3D/2D.

Afin de nous servir de l'index de validation comme dans la section précédente, nous devons avoir à notre disposition une transformation étalon obtenue à l'aide d'un système hautement précis. Nous faisons alors face au même problème que celui présenté dans la section 4.4 : l'absence de recalage étalon. Pour contourner cette difficulté, nous allons nous servir d'une version modifiée du protocole présenté en section 4.4.1, basé sur des boucles de recalages et permettant de mesurer une erreur vectorielle TRE pour un ensemble de points test.

Le principe est le même sauf que nous substituons la grille de calibrage au mannequin. Nous n'avons plus besoin d'élaborer un bronze standard puisque nous rajoutons nous-mêmes synthétiquement sur les données 3D deux tirages de bruit aléatoire. Une boucle type, schématisée sur la figure 7.7, décrit le processus permettant d'obtenir une valeur de μ^2 . Plus précisément, nous savons calculer la matrice de covariance des 2 transformations \hat{T}_1 et \hat{T}_2 de la boucle, nous pouvons alors propager le bruit et calculer la matrice de covariance de C après la boucle. Nous obtenons finalement une matrice 3×3 $\Sigma_{loop} = \Sigma_{\hat{T}_1, \hat{T}_2^{-1} \star C}$ qui contient la prédiction de l'incertitude sur C après application des 2 recalages. Réalisant plusieurs fois cette expérience, nous pouvons calculer l'index :

$$\mu^2 = \frac{1}{N} \sum_{i=1}^N (C - E \star C)^\top \cdot \Sigma_{loop}^{-1} \cdot (C - E \star C)$$

et vérifier que sa distribution suit bien une loi du chi-2 à 3 degrés de liberté²⁴.

Conditions expérimentales et résultats

Pour une configuration donnée de nos 4 caméras et de notre objet étalon, nous pouvons calculer $T_1 \star T_2^{-1} \star C$ en changeant le nombre de points et/ou l'amplitude du bruit sur les données. En faisant varier la position du damier et la configuration et le calibrage des 4 caméras, nous obtenons finalement 144 valeur de μ^2 pour chaque valeur de nombre de points et d'amplitude de bruit. Pratiquement, l'angle entre les caméras varie de 10° à 60° , et la distance objet-caméra est 25 à 30 fois supérieure à la distance focale.

La figure 7.8 présente la moyenne, la variance et la valeur du KS-test de l'index de validation par rapport aux nombres de points utilisés (choisi aléatoirement parmi les 54 possibles). Il en résulte que notre prédiction est valide jusqu'à 40 points (ce qui reste conforme aux conditions de notre application pour laquelle nous n'envisageons pas l'utilisation de plus de 25 marqueurs). Cette valeur critique s'explique par la réduction progressive de l'erreur de recalage qui rejoint finalement l'erreur de calibrage ignorée dans notre modèle (et qui semble être comprise entre 0.5 et 1 mm). De même, nous observons sur la figure 7.9 un comportement similaire lorsque le bruit sur les données devient trop faible (σ_{3D} et σ_{2D} en dessous de 0.7).

7.2.3 Validation sur des données entièrement réelles

Nous abordons maintenant l'étape finale de la validation. Après avoir démontré que les effets de la linéarisation étaient négligeables et quantifié l'importance des erreurs induites par le calibrage, il nous reste à vérifier sur des données réelles que le bruit d'extraction sur les données 3D et 2D est suffisamment proche d'une loi gaussienne non-biaisée. La précision de notre système étant basée sur les recalages 3D/2D des structures internes reconstruites et de l'aiguille virtuelle, nous réalisons notre validation finale sur des données réelles provenant du mannequin et de l'aiguille.

Prédiction de l'erreur de recalage du mannequin

Les conditions expérimentales sont les même que celles utilisées pour le protocole de mesure de la section 4.4. Cependant, la validation statistique est conduite différemment par rapport à la section précédente. Les points cibles C_i sont maintenant choisis à l'intérieur du foie du mannequin et non plus à une distance et une position aléatoire par rapport au damier, et les marqueurs visibles dans le CT et dans les images vidéo sont détectés avec nos algorithmes d'extraction. En outre comme il y a un recalage CT/CT intermédiaire, nous devons insérer sa matrice de covariance dans la boucle totale.

24. Nous indiquons au passage que nous pouvons aussi réaliser la même validation sur l'incertitude prédite sur $T_1 \star T_2^{-1}$. Dans ce cas, $\mu^2 = \frac{1}{N} \sum_{i=1}^N T_1^\top \cdot \Sigma_{T_{loop}}^{-1} \cdot T_2^{-1}$ et nous devons vérifier que μ^2 suit une loi du χ_6^2 . L'interprétation de μ^2 étant difficile dans ce cas-là, nous avons mené notre validation sur l'incertitude d'un point cible dans l'espace 3D.

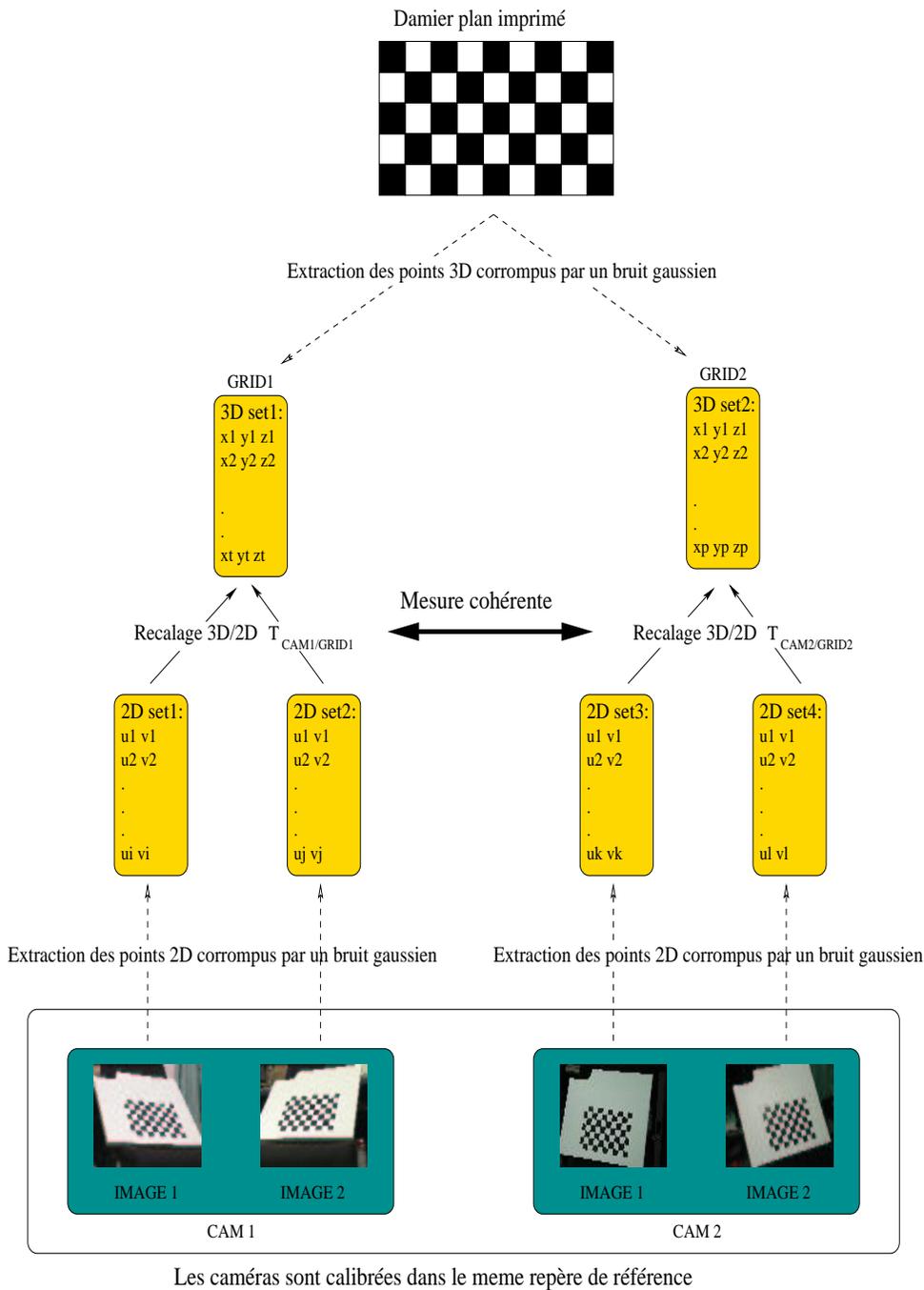


FIG. 7.7 – Boucle de recalage utilisée pour obtenir une valeur de *ETRE* vectorielle : un point test C choisi à une certaine distance de la grille imprimée (typiquement 20 cm) est recalé dans le repère de référence $CAM1$ à l'aide d'un premier recalage 3D/2D \hat{T}_1 , puis remis dans le repère de la mire après application d'un second recalage 3D/2D \hat{T}_2 fourni par l'autre couple de caméra (les repères de référence $CAM1$ et $CAM2$ sont identiques puisque les caméras sont calibrées dans le même repère). Si toutes les transformations étaient exactes, nous obtiendrions la même position pour le point test. Évidemment, puisque les transformations ne sont pas parfaites, nous mesurons un vecteur de variance 3D $\sigma_{loop}^2 = 2\sigma_{CAM/GRID}^2$ qui correspond physiquement à une *TRE*. Finalement, une valeur de μ^2 est donnée par la distance de Mahalanobis entre C et $\hat{T}_1 \cdot \hat{T}_2^{-1} \star C$.

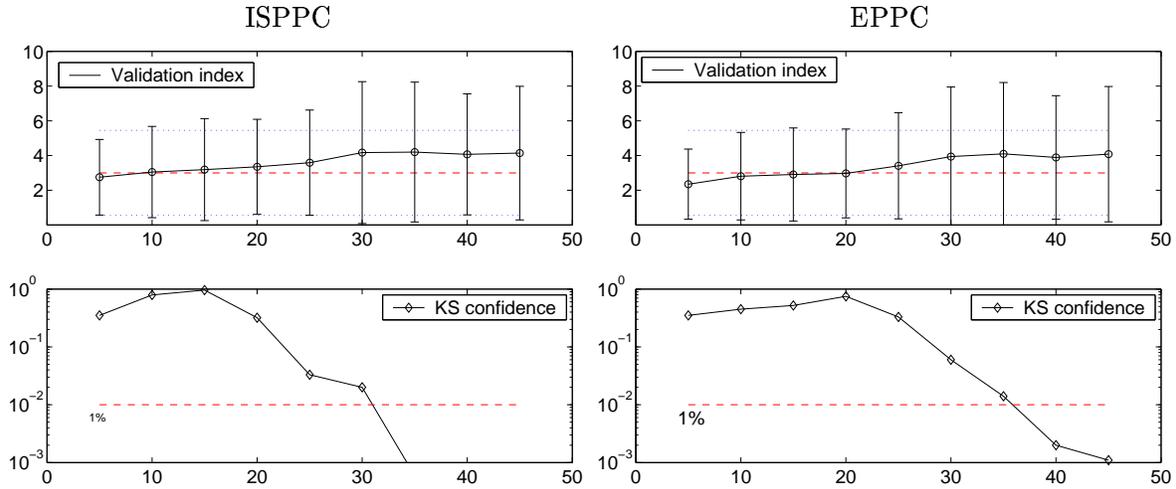


FIG. 7.8 – Validation de la prédiction d'incertitude avec la grille par rapport aux nombres de points utilisés pour réaliser le recalage. En haut : moyenne et variance de l'index de validation. En bas : valeur du KS-test. La prédiction est validée jusqu'à 40 points avec une certitude à 1%.

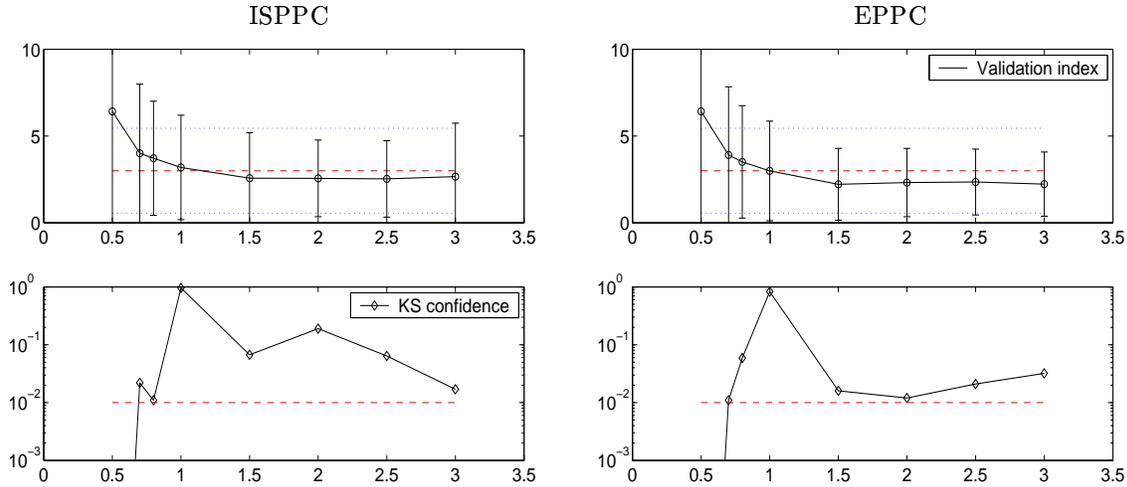


FIG. 7.9 – Validation de la prédiction d'incertitude avec la grille par rapport à l'écart-type du bruit corrompant les données 2D et 3D. En haut : moyenne et variance de l'index de validation. En bas : valeur du KS-test. La prédiction est validée jusqu'à une valeur minimale pour σ_{2D} et σ_{3D} de 0.7 avec une certitude à 1%.

Les marqueurs utilisés sont choisis aléatoirement parmi les 40 possibles et nous obtenons 80 valeurs de μ^2 pour chaque expérience. Comme nous avons observé expérimentalement qu'il y avait un mouvement non-rigide mais cohérent de la peau souple du mannequin (environ 1 mm, cf. section 4.4.1.0), nous avons choisi $\sigma_{3D} \simeq 2.0$ mm (à la place de 1 mm) pour tenir compte de cette incertitude additionnelle. La figure 7.10 présente la moyenne et la variance de μ^2 par rapport au nombre de points utilisés. Nous remarquons tout d'abord que l'index augmente lentement avec le nombre de points. Cela peut être expliqué par le biais introduit par l'erreur de calibrage et le mouvement corrélé des marqueurs avec la peau. En effet, la précision mesurée ne converge pas vers 0 mm avec un nombre de points important mais plutôt vers 1 mm, ce qui correspond aux mouvements de la peau.

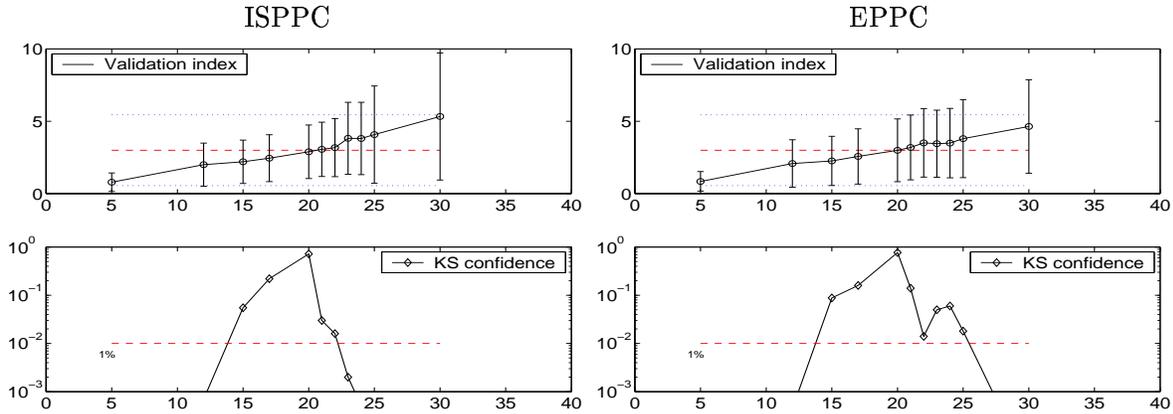


FIG. 7.10 – Validation de la prédiction d'incertitude sur le mannequin par rapport au nombre de points utilisés. En haut: moyenne et variance de l'index de validation. En bas: valeur du KS-test. La prédiction est validée entre 15 et 25 points avec une certitude à 1%.

Néanmoins, il apparaît que la prédiction est correctement validée pour un nombre de points variant entre 15 et 25. Comme $\bar{\mu}^2$ peut être interprété comme une erreur relative sur la prédiction de l'erreur (cf. section 7.2.1), nous surestimons la TRE moyenne d'un facteur 1.7 pour un faible nombre de points ($\bar{\mu}^2 \simeq 1$), et nous la sous-estimons d'un facteur 1.3 pour plus de 25 points ($\bar{\mu}^2 \simeq 5$). Pour notre application, le nombre de points visibles n'excédera pas 25, cela veut dire que nous prédirons correctement l'amplitude de l'erreur. Dans le pire des cas nous la surestimerons, ce qui peut être considéré comme une bonne mesure de sécurité²⁵

25. *Utilisation des intervalles de confiance*: nous pouvons aussi nous servir de l'index de validation afin d'établir des intervalles de confiance permettant notamment d'évaluer le risque d'occurrence d'un mauvais recalage. Dans notre cas nous avons obtenu une erreur moyenne de l'ordre de 2 mm en moyenne. Souhaitant connaître les valeurs maximales envisageables dans 90% des cas, nous consultons une table de loi de χ_2^3 et nous regardons le chiffre correspondant ($p < 0.1$): 6.251. Nous aurons alors 90% de chance d'avoir une erreur de recalage en dessous de $\sqrt{6.251} * 1.8 = 4.5mm$. En général, une majeure partie des 10% restant correspondent aux occurrences se situant en queue de gaussienne. En pratique, la détection des marqueurs suit une loi de gaussienne tronquée: nous n'avons jamais une erreur supérieure à 4 pixels dans les images vidéo et supérieure à 2 mm dans les images CT. La majorité des cas représentant ces 10% ne sont donc pas réalistes dans nos conditions. Nous avons donc bien plus de 90% de chance d'être en dessous de 4.25 mm.

Prédiction de l'erreur de recalage de l'aiguille virtuelle

La validation est menée suivant un protocole identique à celui utilisée avec le mannequin. Cette fois-ci, les coins du carré sont les marqueurs et la cible visée est représentée par la pointe de l'aiguille. Afin de simuler le fait que le marqueur est fixé à l'aiguille d'une manière qui n'est pas forcément reproductible, la position de la pointe est choisie aléatoirement dans une zone sphérique de 20 cm de diamètre dont le centre est situé à 20 cm du marqueur sur le plan formé par ses 4 coins (cf. figure 7.11). Nous avons observé expérimentalement que $\sigma_{2D} \simeq 0.1mm$ et, bien que l'imprimante offre une précision de 1200 dpi (ce qui correspond à $\sigma_{3D} \simeq 10\mu m$), nous choisissons $\sigma_{3D} = 0.05$ afin de tenir compte des erreurs de calibrage et des distorsions de la tête d'impression. Malgré un bruit sur les données 2D sensiblement supérieure à celui sur les données 3D ($RSB_{3D} \simeq 50 dB$ et $RSB_{2D} \simeq 35 dB$) nous décorrélons les données 3D en créant 4 marqueurs à l'aide de 4 imprimantes différentes.

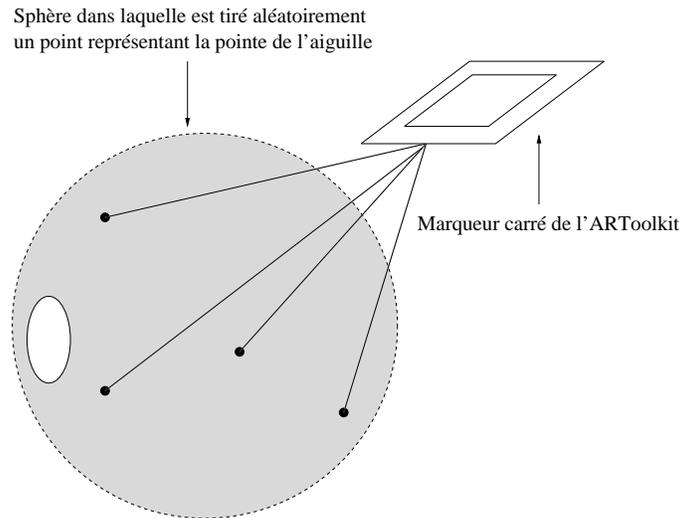


FIG. 7.11 – Nous choisissons aléatoirement un point virtuel représentant la pointe de l'aiguille dans la sphère grisée dont le centre est situé à 20 cm du marqueur et dans le plan formé par ses 4 coins.

Nous déplaçons le marqueur dans les images vidéos de manière à décrire l'ensemble des positions réalistes possibles. Les résultats obtenus figurent dans le tableau 7.2. Tout d'abord, nous voyons que le KS-test n'a pas été validé, ce qui signifie que l'erreur réelle ne suit pas la loi gaussienne prédite. La valeur élevée de l'index de validation ($I=12$) indique que nous sous-estimons l'erreur d'un facteur 2. Ces résultats, pouvant paraître surprenants, étaient en fait attendus à cause de la conclusion obtenue avec les données artificiellement bruitées provenant du damier (cf. section précédente). En effet, on peut lire sur la figure 7.9 que pour 15 points nous surestimons significativement l'erreur dès que le bruit descend en dessous du seuil 0.7 pix. et 0.7 mm. En l'occurrence, les bruits observés sur les données extraites du marqueur lié à l'aiguille sont bien en-dessous de ces valeurs. La précision du recalage de l'aiguille virtuelle est donc du même ordre de grandeur que l'erreur de calibrage.

Pour réussir à valider notre prédiction de l'erreur en présence de bruits 2D et 3D aussi faibles, nous devrions soit disposer d'un calibrage plus précis, soit améliorer notre modèle pour tenir compte d'erreurs supposées jusqu'à présent négligeables (typiquement les paramètres de

	Index de validation μ^2 (3.0)	Ecart-type (2.449)	KS-confiance ($0.01 < x < 1.0$)
ISPPC	12.1	21.7	4e-11
EPPC	11.6	18.8	5e-10

TAB. 7.2 – *Index de validation et test de KS sur des données réelles provenant du marqueur sur lequel l’aiguille est fixée. La valeur de μ^2 proche de 12 indique une sous-estimation de l’erreur de recalage d’un facteur 2.*

calibrage des caméras). Pour l’instant, afin de fournir une prédiction conservatrice de l’erreur globale du système, nous fixons l’erreur de recalage de l’aiguille à la valeur maximale observée lors des expériences de précision dans le chapitre 5 (section 5.3.3), c’est à dire 2.1 mm.

7.3 Conclusion

Dans ce chapitre nous nous sommes intéressés à la sécurité de notre système en terme de précision. Nous avons tout d’abord mis l’accent sur le fait qu’une évaluation de la précision dans des conditions expérimentales spécifiques était généralement insuffisantes à cause de la variabilité des paramètres retrouvés lors d’une intervention clinique. En effet, même si nous pouvions assurer une géométrie des caméras fixe, la dimension de l’aiguille employée par le chirurgien peut changer (de 10 jusqu’à 20 mm), et la position de la cible par rapport aux marqueurs change d’un patient à l’autre.

Il est donc indispensable de fournir aux chirurgiens un moyen de connaître avant l’intervention la précision moyenne du recalage calculé par notre système. Dans ce but, nous avons mené une étude de la propagation des perturbations au travers du critère d’optimisation utilisé. Nous obtenons ainsi une formulation analytique de la covariance sur la transformation, et pouvons donc fournir une estimation de la TRE et de l’erreur de repositionnement de l’aiguille virtuelle. Comme la validité de nos calculs théoriques est liée à certaines hypothèses, il est nécessaire de procéder à une validation rigoureuse sur des données réelles proches des conditions cliniques envisageables.

Afin de vérifier la validité des approximations réalisées, nous menons une étude incrémentale qui nous permet de quantifier l’impact de chacune des hypothèses. Nous vérifions d’abord que les non-linéarités dans le critère et que les erreurs de calibrage sont négligeables. Notre système dépendant de 2 recalages 3D/2D, nous procédons ensuite à deux validations sur données réelles provenant d’abord du mannequin en plastique de l’abdomen puis de l’aiguille lié à son marqueur. Cette étude confirme que notre prédiction de la TRE dans les conditions de notre application (i.e. entre 15 et 25 marqueurs) est correcte, avec une légère surestimation si trop de marqueurs sont occultés, ce qui est une bonne mesure de sécurité. En revanche, les résultats obtenus avec les données de l’aiguille mettent en évidence les limites de notre système actuel. Notre prédiction de l’erreur de repositionnement de l’aiguille virtuelle, deux fois inférieures à l’erreur réelle, vient du fait que l’erreur de recalage est de l’ordre de l’erreur introduite par le calibrage, et qui n’est pas modélisée dans notre critère pour mener les calculs théoriques.

Pour améliorer la qualité de notre prédiction dans un avenir proche, nous envisageons d’inclure un modèle de bruit sur les paramètres extrinsèques des caméras.

Chapitre 8

Du prototype à l'application clinique

Nous avons développé dans les chapitres précédents toutes les briques du système que nous nous étions proposés de concevoir et les avons évaluées de manière indépendante. Nous avons premièrement démontré que l'erreur moyenne de recalage du modèle scanner du patient dans les images vidéo était de 2 mm. Ensuite nous avons évalué l'erreur de repositionnement de la pointe de l'aiguille virtuelle à 0.95 mm. Nous avons alors confirmé l'intérêt de nos algorithmes automatiques d'extraction et d'appariement, puisqu'ils exécutent une tâche usuellement interactive en 4 fois moins de temps et avec une précision équivalente. Finalement, nous avons validé notre prédiction de l'erreur de recalage du modèle scanner. L'objectif de ce chapitre est de valider le système complet, toutes les briques étant assemblées. Nous commençons par montrer comment nous évaluons les performances de notre système en utilisant des marqueurs radio-opaques pour cibles dans un mannequin de l'abdomen. Cette expérimentation mettant en évidence les faiblesses du système, nous expliquons ensuite les améliorations au niveau de l'ergonomie que nous avons imaginées et développées avec les chirurgiens. Nous évaluons ensuite les conditions nécessaires à une première expérience sur l'animal. D'une part, les démarches administratives sont excessivement longues pour obtenir un créneau horaire au scanner avec un cochon (six mois au minimum), d'autre part, la durée de mise à disposition du scanner est très réduite (90 minutes). Pour contourner ce problème, nous montrons comment évaluer notre système sur un patient grâce à un protocole sans danger qui permet de comparer la précision de notre système à celle du praticien. Finalement, nous présentons les premiers résultats obtenus sur 3 patients à l'aide de cette méthode.

8.1 Validation du système sur un mannequin

Le but de notre expérience est de mesurer la précision de ponctions percutanées²⁶ guidées par notre système sur des cibles situées dans le mannequin.

8.1.1 Protocole expérimental

Cette expérience a été réalisée dans une salle d'opération à l'IRCAD (cf. figure 8.1). Nous avons collé une trentaine de marqueurs radio-opaques sur la peau du mannequin et une dizaine à la surface de son foie. Une acquisition scanner du mannequin a été réalisée, et nous avons

²⁶. Une ponction percutanée est un acte médical qui consiste à introduire une aiguille dans un patient pour analyser ou pour détruire une zone pathologique.

reconstruit les modèles de sa peau, de son foie et de tous les marqueurs radio-opaques (cf. figure 8.2). Les centres des marqueurs sur la peau ont été repérés automatiquement par notre algorithme d'extraction et ceux collés sur le foie ont été localisés semi-interactivement dans l'image scanner : ils représentaient les cibles à atteindre.

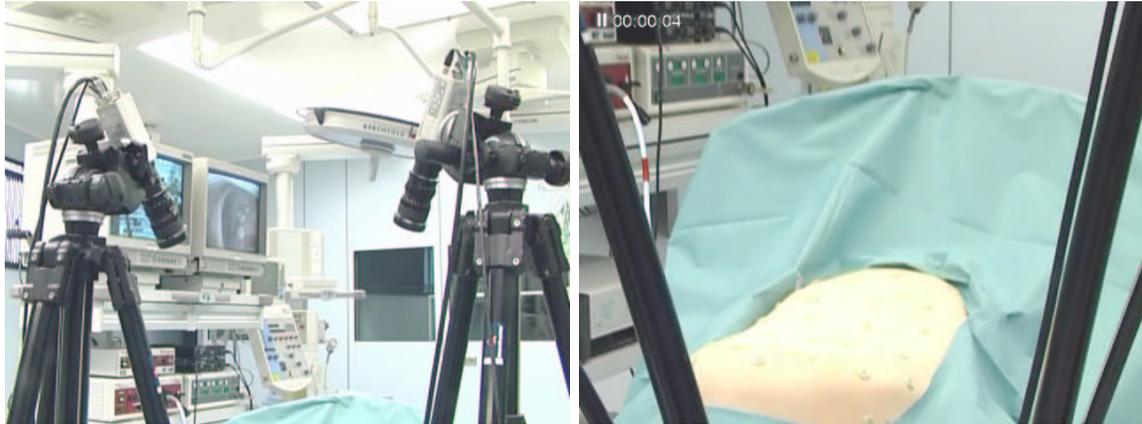


FIG. 8.1 – A gauche : vue d'ensemble de la salle d'opération et des deux caméras. A droite : le mannequin est disposé sous les caméras.

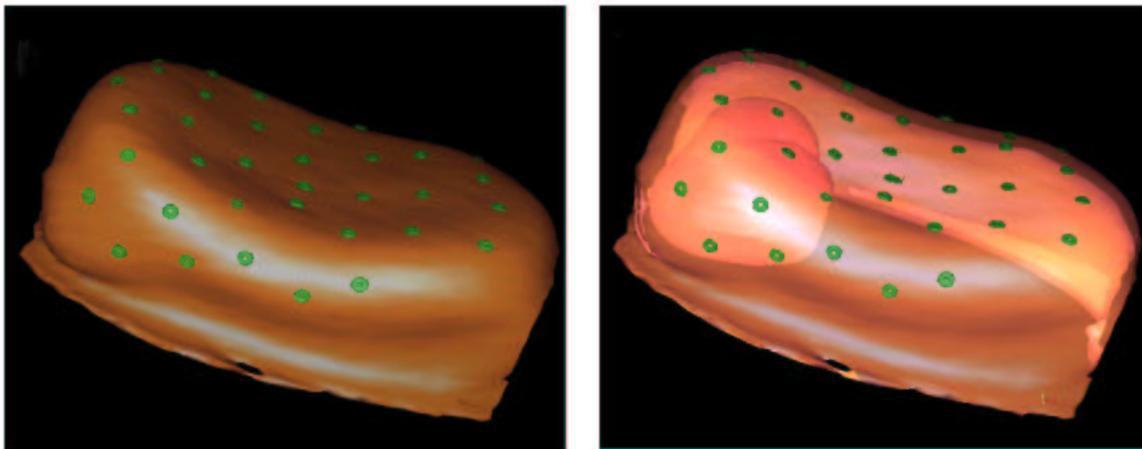


FIG. 8.2 – A gauche : reconstruction surfacique du mannequin à partir d'une segmentation réalisée dans l'image scanner. A droite : vue en transparence du modèle 3D. Le foie devient visible sous la peau.

Nos deux caméras ont été disposées au dessus du mannequin, orientées vers la zone contenant le foie. Leurs focales ont été choisies de sorte que le marqueur fixé à l'aiguille soit visible dans le champ de vue des deux caméras durant sa manipulation. Les caméras ont été ensuite calibrées dans un même repère de référence en suivant la méthode décrite dans la section 5.2.4. La mire utilisée est un damier imprimé sur une feuille de format A3 contenant 17×13 points de référence (les carreaux ont une dimension de 20×20 mm). Le calibrage de l'imprimante n'étant pas parfait, il peut arriver que les dimensions de la mire imprimée soient sensiblement diffé-

rentes de celles rentrées en paramètre au moment de l'impression (nous avons parfois observé que la longueur de la mire imprimée atteignait 341 mm alors que nous avons paramétré la longueur à 340 mm). Lorsque cela arrive, nous modifions les paramètres de format d'impression en appliquant une règle de trois, de sorte que la mire imprimée ait la longueur requise. La position de la pointe de l'aiguille et son orientation ont été calibrées par rapport au marqueur carré la veille de l'expérience. Afin d'obtenir un calibrage de l'aiguille aussi précis que possible (cf. section 5.3.2 pour la méthodologie du calibrage stéréoscopique), les focales des deux caméras ont été choisies de telle sorte que la taille du marqueur carré soit la plus importante possible dans les deux images vidéo (le rapport signal sur bruit de la détection des coins du marqueur carré est alors plus élevé).

Durant l'expérience, nous n'avons pas utilisé une détection en temps réel des marqueurs dans les images vidéo, puisque la peau du mannequin est très élastique et n'est pas retenue en des endroits proches des points d'insertion de l'aiguille. En effet, lorsque nous introduisons l'aiguille, la peau s'affaissait de 4-5 cm et déplaçait considérablement quelques marqueurs. Tenir compte de ces marqueurs reviendrait à introduire des données corrompues par un bruit corrélé, ce qui ne correspond pas à nos hypothèses sur le bruit d'extraction dans les images vidéo (bruit centré gaussien). L'extraction automatique des marqueurs n'a donc été réalisée qu'une seule fois avant le début de l'expérimentation. Dans le cas d'un patient réel, nous pourrions employer une détection en temps réel puisque l'introduction de l'aiguille ne déplace la peau que localement et légèrement. Dans le pire des cas, si un marqueur est situé trop près de la zone d'insertion, nous le décollerons.

Nous avons ensuite exécuté les deux procédures d'appariement qui ont mis en correspondance 100% des marqueurs détectés dans les deux images vidéo. La transformation reliant le repère scanner au repère des caméras a été alors calculée par optimisation du critère EPPC. Après application de la transformation trouvée sur le modèle extrait du scanner, nous l'avons reprojété dans les deux images vidéo²⁷. Nous avons pu alors vérifier visuellement que la re-projection des marqueurs extraits de l'image scanner recouvraient correctement les marqueurs visibles dans les images vidéo (cf. figure 8.3). Finalement, une caméra endoscopique fut installée à l'intérieur du mannequin et orientée vers le foie pour permettre à un opérateur indépendant de relever la précision de chaque ponction (cf. figure 8.4).

8.1.2 Guidage de ponctions percutanées sur des cibles à l'intérieur du mannequin

Un chirurgien et moi-même avons réalisé chacun un total de cinquante ponctions percutanées sur une des cibles situées sur le foie du mannequin. Nous aurions souhaité utiliser plusieurs cibles différentes, cependant le manque d'espace dans le mannequin a limité l'angle de vue de la caméra endoscopique de sorte que nous ne pouvions juger visuellement la précision que pour une seule cible. Les images vidéo contenaient le foie et la cible à atteindre, et la position virtuelle de l'aiguille était superposée en bleu foncé sur l'aiguille réelle (cf. figure 8.5).

En plus de la position virtuelle de l'aiguille et de la cible pour se guider, nous indiquions en temps réel les trois valeurs décrites sur la figure 8.6 (la distance séparant le centre de la cible de la pointe de l'aiguille, la distance entre le centre de la cible et l'axe de l'aiguille, et l'erreur d'orientation de l'aiguille). Leur lecture et interprétation simultanée étant difficile, l'orientation était aussi guidée visuellement en fonction de la couleur de la cible (blanc : $\beta > 30^\circ$, rouge : $10^\circ < \beta < 30^\circ$, jaune : $2^\circ < \beta < 10^\circ$, vert : $\beta < 2^\circ$).

27. Le mixage des modèles virtuels dans les images réelles est réalisé avec la librairie graphique OpenGL.



FIG. 8.3 – Vérification visuelle de la reprojection des marqueurs extraits de l'image scanner après recalage du modèle virtuel dans le repère de référence des caméras. En haut : images vidéo sans reprojection des marqueurs. En bas : avec reprojection des marqueurs. Le résultat visuel est de bonne qualité : les marqueurs réels sont tous recouverts par leur homologues virtuels. Les procédures d'extraction et d'appariement ont bien fonctionné. Dans l'image en bas à droite, trois marqueurs rouges sont superposés sur le champ opératoire et semblent être mal recalés. En fait, les trois marqueurs réels leur correspondant sont cachés sous le champ opératoire.



FIG. 8.4 – Vérification par vision endoscopique de la qualité du positionnement de l'aiguille dans la cible. Le déplacement de la caméra dans le mannequin étant extrêmement limité: un seul marqueur était suffisamment visible pour permettre une vérification pertinente.

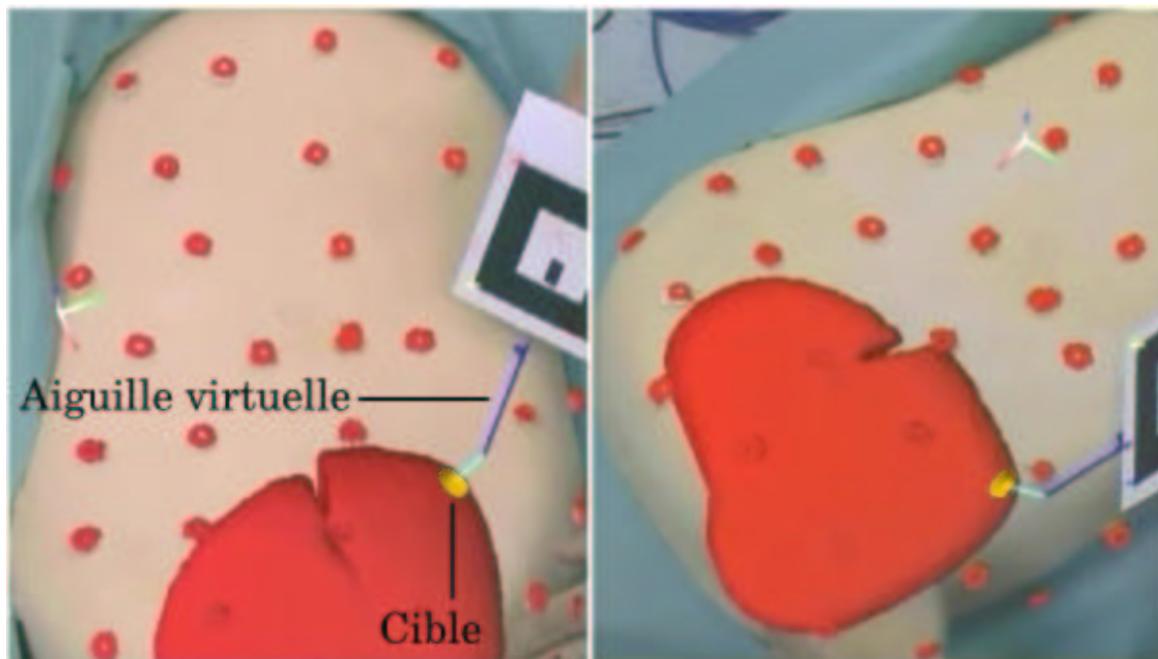


FIG. 8.5 – Exemple de vision augmentée dont nous disposons pour guider l'aiguille au centre de la cible (en jaune). Le foie virtuel est superposé en rouge et l'aiguille virtuelle en bleu foncé. Le trait bleu clair correspond au trajet séparant la cible de la pointe de l'aiguille virtuelle.

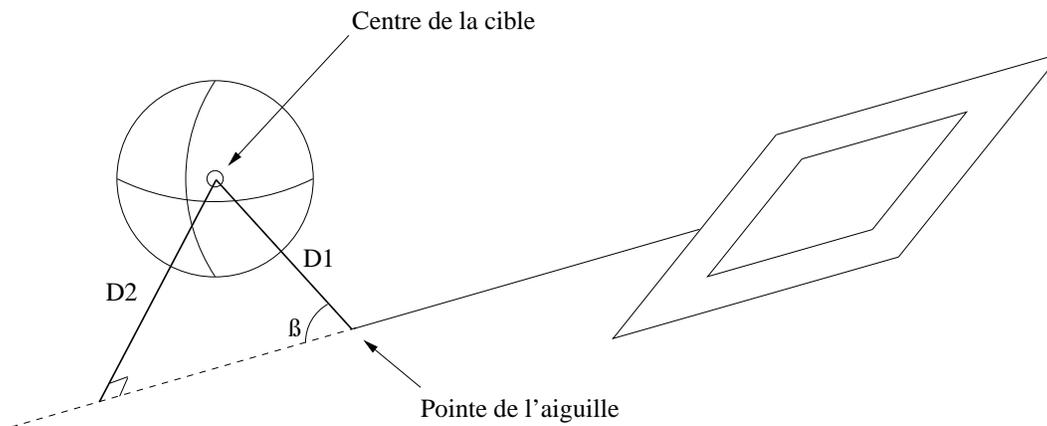


FIG. 8.6 – Description des trois valeurs indiquées à l'utilisateur pour guider l'aiguille. $D1$ correspond à la distance séparant le centre de la cible de la pointe de l'aiguille. $D2$ correspond à la distance entre le centre de la cible et l'axe de l'aiguille. β traduit l'erreur d'orientation de l'aiguille.

Au début de chaque tentative, l'aiguille était à l'extérieur du mannequin. L'opérateur décidait quand il commençait son geste et un chronomètre était alors déclenché. Celui-ci était arrêté lorsque l'opérateur estimait qu'il avait suffisamment bien placé la pointe de l'aiguille. La vérification de la précision du positionnement était alors faite par une autre personne grâce à la vue endoscopique. A la fin de chaque tentative, nous avons relevé la durée du positionnement, l'erreur réelle observée grâce à la caméra endoscopique et l'erreur virtuelle $D1$ indiquée par notre logiciel. La caméra endoscopique étant enregistrée, en la visionnant nous avons pu précisément décrire la distribution spatiale de toutes les ponctions réalisées sur le marqueur.

8.1.3 Résultats et discussion

Le tableau 8.2 résume toutes les mesures que nous avons effectuées, et la figure 8.7 regroupe les distributions spatiales des ciblage réalisés par le chirurgien et par moi-même. La précision globale obtenue est extrêmement satisfaisante puisqu'en moyenne, elle est de 2 mm inférieure à la limite de 5 mm fixée dans le cahier des charges. De même, la durée moyenne nécessaire à l'insertion et au positionnement de l'aiguille étant de l'ordre de la minute, est bien inférieure à celle que l'on retrouve dans les conditions classiques (10 minutes). Nous remarquons aussi une différence significative entre les précision obtenues par le chirurgien et moi-même. Le chirurgien a été plus précis et plus concentré dans ses pointages. Ce constat est très positif puisqu'il est courant que le concepteur d'un système de guidage médical soit plus habile que le praticien lors de son utilisation alors que le système lui est initialement destiné. Dans notre cas, les types d'information de guidage ont été mis au point en concertation avec le chirurgien pour éviter de favoriser l'expertise du système développé au détriment de l'expertise médicale.

Nous pouvons remarquer que la distribution est centrée autour d'un point décalé du centre d'environ 2 mm. Ce phénomène peut provenir du fait que, le mannequin n'ayant pas été déplacé durant toute l'expérience, un seul recalage 3D/2D a été exécuté : l'erreur due au recalage est restée statique. Dans ce cas, la cible virtuelle était vraisemblablement décalée de 2-3 mm vers le haut du marqueur.

	Total	Informaticien	Chirurgien
Erreur RMS réelle (mm) (\pm écart-type)	3.1 ± 1.4	3.74 ± 1.03	2.3 ± 1.27
Erreur RMS virtuelle D1 (mm) (\pm écart-type)	2.5 ± 1.2	2.5 ± 0.88	2.0 ± 0.71
Durée moyenne (sec.) (\pm écart-type)	46.5 ± 24.6	38.5 ± 21.8	54.6 ± 24.9

TAB. 8.1 – Résultats obtenus lors de notre expérience de guidage avec le mannequin. La colonne “Total” regroupe les résultats indépendants obtenus par l’informaticien et le chirurgien. L’erreur RMS (Root Mean Square) finalement observée est très satisfaisante puisqu’en moyenne, elle est inférieure de 2 mm à notre objectif en précision. La durée moyenne est également très honorable puisqu’elle vaut moins d’une minute alors que le temps standard de ce type d’intervention est de l’ordre de 10 minutes. L’erreur virtuelle D1 est représentative de la difficulté rencontrée pour maintenir l’aiguille dans une position statique.

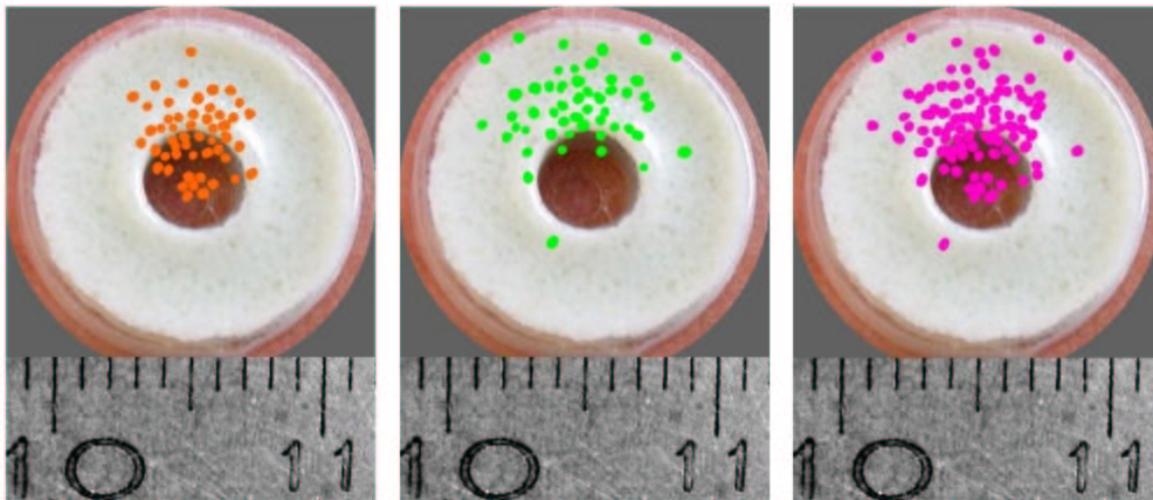


FIG. 8.7 – Distributions spatiales des ponctions réalisées sur le marqueur cible. A gauche : distribution obtenue par le chirurgien. Au milieu : par l’informaticien. A droite : somme des deux distributions.

L'erreur virtuelle D1 représentant la distance séparant la pointe de l'aiguille virtuelle du centre de la cible virtuelle, nous nous attendions à obtenir une moyenne beaucoup plus proche de 0. Néanmoins, le mannequin étant vide entre le foie et la peau, il n'était pas possible de maintenir l'aiguille dans une position statique à cause des tremblements de la main. L'erreur virtuelle oscillait quand nous étions proches de la cible et nous décidions d'arrêter notre geste lorsque nous stabilisons l'erreur entre 0 et 3 mm. En condition clinique, cette erreur serait sans doute plus faible étant donné que l'aiguille serait stable une fois introduite dans le corps du patient. Par ailleurs, nous mentionnons que notre prédiction sur l'erreur de recalage du modèle virtuel était égale à 2.8 mm pour un choix $\sigma_{2D} = 2$ pixels et $\sigma_{3D} = 2$ mm. L'erreur de repérage de l'aiguille ayant été évaluée à 0.95 mm, nous devons nous attendre à une erreur *RMS* finale de $\sqrt{2.8^2 + 0.95^2} = 2.95$ mm. Notre prédiction était donc très proche de la réalité finalement observée : 3.1 mm.

Nous devons toutefois faire attention au fait que l'erreur observée peut être une sous-évaluation de l'erreur réelle. En l'occurrence, l'erreur observée ne représente pas toute l'erreur possible. En effet, le foie étant en plastique trop dense pour pouvoir y insérer l'aiguille, nous avons été contraint de placer les cibles à la surface du foie. La mesure de l'erreur de positionnement a donc été réalisée dans le plan contenant le marqueur : la vérification endoscopique ne permettait de considérer que la position de l'aiguille sur le marqueur, et non pas d'évaluer l'erreur de positionnement en profondeur. Nous avons donc sous-estimé l'erreur moyenne de positionnement. Cependant, l'enregistrement vidéo a permis de vérifier que la pointe de l'aiguille était systématiquement dans le marqueur ou bien dans son orifice central. Le marqueur ayant une épaisseur de 2 mm, une évaluation conservatrice de l'erreur peut donc être donnée par la somme *RMS* de l'erreur observée et de la demi-épaisseur du marqueur : $\sqrt{3.1^2 + 1.0^2} = 3.26$ mm, qui reste un résultat très satisfaisant en regard des 5 mm d'erreur requis.

8.2 Un système ergonomique et sécurisé

Bien que les résultats obtenus aient été satisfaisants, nous avons mis en évidence pendant l'expérience certaines imperfections de l'interface que le chirurgien a souhaité que nous modifions. Par ailleurs, il a souligné que nous n'avions pas intégré dans notre système des solutions de repli dans certains cas de figure extrêmes (nombre insuffisant de marqueurs détectés ou d'appariement correct). Dans cette section, nous présentons en premier lieu les améliorations imaginées avec le chirurgien que nous avons apporté à notre système afin de le guider efficacement tout au long de l'intervention. En second lieu, nous détaillons les situations critiques relevées par le chirurgien et leurs alternatives appropriées.

8.2.1 Un guidage optimal de la pose des marqueurs au positionnement de l'aiguille

Guidage du positionnement des marqueurs

Au début de l'expérience, nous avons observé au moment de la phase d'appariement 2D que seulement 4 marqueurs sur 26 avaient été correctement appariés. Nous nous sommes rendu compte après vérification que les marqueurs avaient été disposés maladroitement le long des droites épipolaires, ce qui augmentait considérablement les chances d'avoir des appariements ambigus. Malgré le fait que notre algorithme de prédiction/vérification ait permis de retrouver tous les appariements possibles, il est évidemment préférable d'éviter ce genre de configuration

qui pourrait en cas de malchance empêcher une mise en correspondance 3D/2D. Pour guider le praticien durant la pose des marqueurs, nous faisons actuellement apparaître sur les images vidéo du patient une zone à éviter autour des segments épipolaires relatifs aux marqueurs, au fur et à mesure qu'il les colle. Il peut ainsi éviter visuellement les configurations ambiguës.

Estimation de la précision avant le passage au scanner du patient

Dans le système actuel, le praticien colle au début d'une intervention un certain nombre de marqueurs radio-opaques sur la peau du patient et le système évalue ensuite la précision du système une fois que le scanner du patient est réalisé et que les marqueurs ont été extraits de l'image résultante. Dans le cas où la précision prédite est insuffisante, il est probable que le praticien juge indispensable de coller des marqueurs supplémentaires sur le patient et de réaliser une nouvelle acquisition scanner.

Pour éviter une telle situation, une possibilité pour le praticien est de coller plus de marqueurs que cela ne lui semble nécessaire. Cependant, cette surestimation ne lui assurera pas pour autant une précision suffisante étant donné que l'angle entre les caméras et leur distance focale ont un impact majeur sur la précision qu'il ne saura pas estimer. Pour éviter des irradiations inutiles au patient, il est primordial de pouvoir fournir au praticien un moyen de prédire l'erreur de recalage *avant* le passage du patient au scanner.

D'un point de vue théorique, notre prédiction de l'erreur sur le recalage ne dépend pas de la valeur de la transformation recherchée reliant le référentiel scanner à celui des caméras. En effet, elle ne dépend que de la position relative des marqueurs, des paramètres des caméras et des caractéristiques du bruit corrompant les données. Étant donné que les caméras sont calibrées et que les bruits ont déjà été caractérisés, si nous sommes capables de connaître à l'avance la position relative des marqueurs dans l'espace, cela signifie que nous pouvons calculer l'erreur de recalage avant le passage du patient au scanner. Évidemment, il est impossible de connaître la position des marqueurs dans le référentiel scanner avant d'avoir scanné le patient. En revanche, nous pouvons calculer une estimation de leur position relative par reconstruction stéréoscopique. Comme tous les marqueurs ne sont pas forcément visibles et appariés, cela signifie que nous pourrions fournir avant le scanner du patient une sous-estimation de l'erreur de recalage, ce qui correspond à une bonne mesure de sécurité. Ce nouveau module a été implémenté mais n'a pas encore été testé.

Un guidage ergonomique du positionnement de l'aiguille

Pendant notre expérience, nous parvenions assez rapidement à nous approcher de la cible à moins de 5 mm en nous servant du code de couleur indiquant la qualité de l'orientation vers la cible. En revanche, le positionnement précis de l'aiguille s'est avéré être une tâche ardue puisque la connaissance de la distance séparant la cible de la pointe ne nous informait pas sur leur position relative. Nous étions alors obligé de procéder par tâtonnement afin de nous rapprocher à moins de 3 mm de la cible. Par ailleurs, le chirurgien aurait parfois souhaité disposer d'un angle de vue différent de celui fourni par les caméras afin de mieux comprendre spatialement la position de l'aiguille par rapport aux modèles virtuels.

Pour permettre au praticien de savoir où se situe la cible lorsque la pointe de l'aiguille est à moins de 4 mm, nous lui affichons actuellement une sphère dont le centre correspond à la pointe et contenant un vecteur orienté vers la cible (cf. figure 8.8).

Pour que le praticien puisse changer d'angle de vue, nous avons décidé de remplacer la vue stéréoscopique des caméras par une vue virtuelle des organes et de la position de l'aiguille, dont

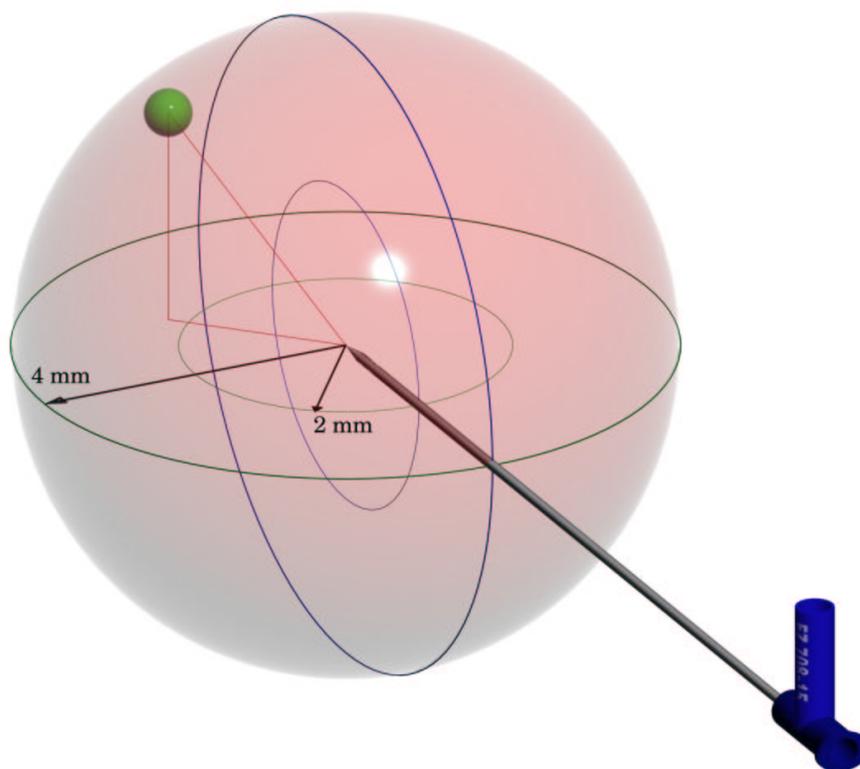


FIG. 8.8 – Interface de guidage pour le positionnement fin de l'aiguille au centre de la tumeur. Nous affichons la position relative de l'aiguille par rapport à la tumeur (en vert) dans une sphère de rayon 4 mm. Le triangle rectangle rouge, dont l'hypoténuse relie la pointe de l'aiguille au centre de la tumeur, fournit l'information de profondeur nécessaire au positionnement de l'aiguille.

on peut modifier l'angle de vue et la focale. Néanmoins, cela nécessite l'aide d'un agent externe qui manipule le point de vue de l'image afin de fournir au praticien la vue virtuelle qu'il désire. Pour limiter le besoin d'interaction que cela implique entre le praticien et l'agent, nous avons ajouté une caméra virtuelle au bout de l'aiguille dirigée selon son axe (cf. figure 8.9). Cet outil supplémentaire, très pratique pour orienter l'aiguille vers la cible, permet au praticien de poser l'aiguille sur le point d'entrée de la peau et de visualiser, en jouant sur la transparence des modèles 3D reconstruits, les zones qu'il va traverser avant d'atteindre la cible.

Finalement, notre interface de guidage se compose d'un module pour le positionnement grossier (vue externe + caméra embarqué dans l'aiguille) et d'un autre pour le positionnement précis (sphère 3D + distance numérique séparant la cible de la pointe de l'aiguille virtuelle).

Afin de démontrer que notre nouvelle interface est plus ergonomique que la version initiale, nous avons mené la même expérience que celle décrite dans la section 8.1.1. Cette fois-ci, plusieurs marqueurs ont pu servir de cibles, et le mannequin était légèrement déplacé tous les 10 pointages réalisés. Les résultats quantitatifs sont indiqués dans le tableau 8.2.

Nous constatons effectivement que cette version de l'interface permet d'obtenir de meilleurs résultats. Si la précision n'a été que légèrement améliorée (1.8 mm contre 2.3 mm), le temps nécessaire au positionnement de l'aiguille a été réduit d'un facteur 2 (30 secondes contre 54 secondes auparavant).

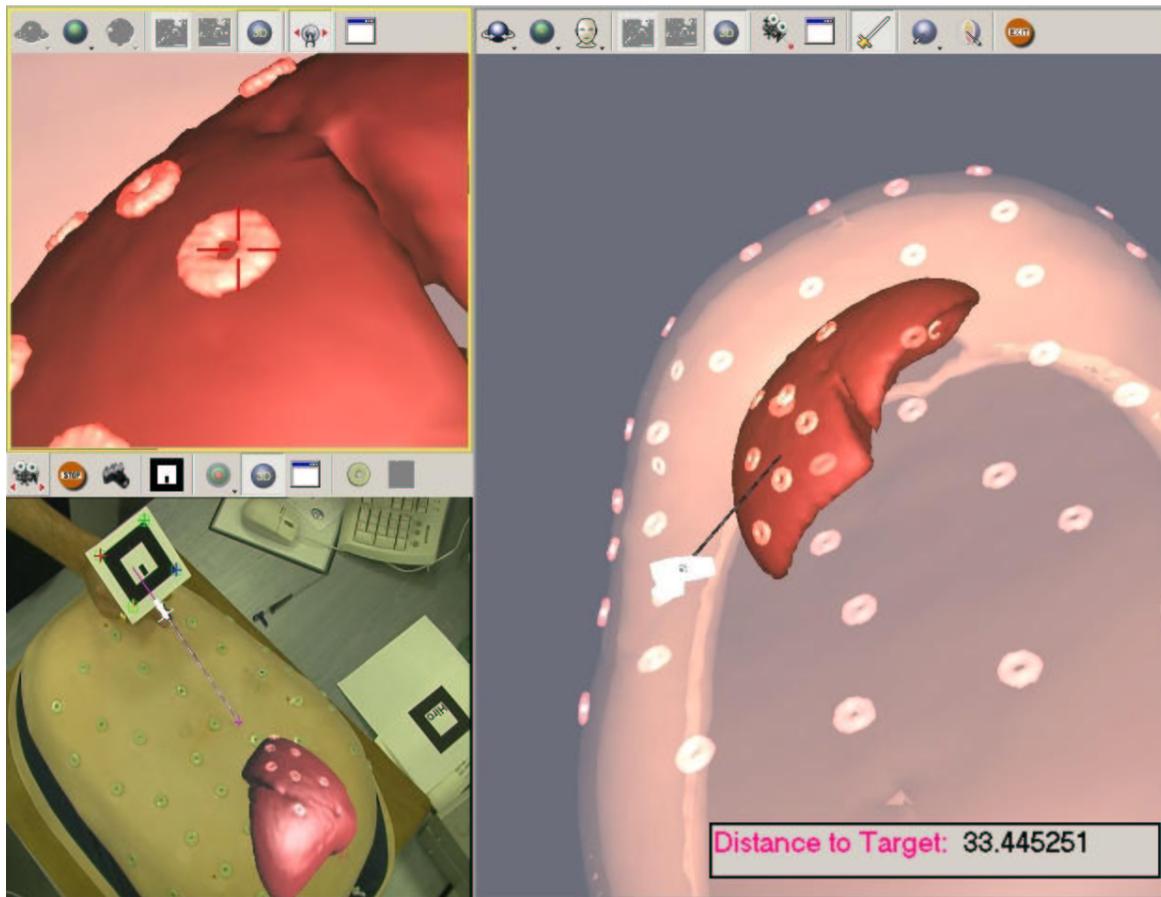


FIG. 8.9 – Nouvelle interface de guidage pour le positionnement de l'aiguille au centre de la tumeur. Nous affichons la position relative de l'aiguille par rapport aux modèles virtuels (image de droite). La vue de type cible est positionnée en haut à gauche. Nous affichons finalement en bas à gauche la vue réelle sur laquelle sont superposés les modèles virtuels.

	Erreur RMS moyenne (mm) \pm écart-type	Erreur minimale	Erreur maximale	Durée moyenne (sec.) \pm écart-type
Chirurgien 1	2.2 ± 0.57	1	3	32 ± 12.2
Chirurgien 2	2.9 ± 1.25	0	5	22 ± 3.1
Chirurgien 3	1.3 ± 1.16	0	3	32 ± 3.7
Chirurgien 4	0.84 ± 0.48	0	1	32 ± 6.4
Total	1.8 ± 0.7	-	-	29.5 ± 7.3

TAB. 8.2 – Résultats obtenus lors de l'expérience de guidage avec la nouvelle interface. L'erreur moyenne de positionnement de l'aiguille a été légèrement améliorée par rapport à l'ancienne version. La durée, quant à elle, a été divisée par 2. Cette version de l'interface semble donc bien plus adaptée au guidage que la version précédente. Notons que les erreurs minimales et maximales relevées sont des valeurs entières à cause de la quantification de l'opérateur vérifiant les pointages. En effet, la vue endoscopique permettait seulement une évaluation millimétrique de l'erreur de pointage.

Malgré les résultats très encourageants obtenus, l'interface de notre système est encore en cours d'optimisation. En effet, nous attendons les premières expériences *in vivo* pour récupérer l'avis des praticiens, et pour savoir quelles sont les améliorations supplémentaires dont ils aimeraient bénéficier.

8.2.2 Un système sécurisé à chaque étape

Dans le cas de figure improbable où un faux positif entraîne une erreur d'appariement, le praticien ne s'en rendra pas forcément compte même en reprojétant le modèle virtuel dans l'image vidéo. Il est donc indispensable de pouvoir vérifier rapidement à chaque étape que nos algorithmes ont fonctionné correctement. Par ailleurs, si trop peu de marqueurs ont été extraits, il faut donner la possibilité au praticien de compléter interactivement les données. Nous détaillons dans cette section tous les utilitaires qui doivent encore être développés afin de sécuriser l'ensemble des procédures de notre système.

Vérification de l'extraction automatique

Si un simple coup d'oeil suffit pour vérifier que les marqueurs radio-opaques ont été correctement extraits dans les images vidéo, il semble moins aisé de vérifier l'extraction des marqueurs 3D. La solution la plus simple, consistant à rechercher les marqueurs extraits en parcourant les coupes scanner, est une tâche relativement fastidieuse. En fait, comme nous créons, durant l'extraction des marqueurs de l'image scanner, les masques binaires des marqueurs extraits et de la peau (sans les marqueurs), une reconstruction surfacique rapide (cf. article de Delinquette [Del99]) en deux couleurs différentes et une visualisation 3D permettra de constater très rapidement la présence de faux positifs ou de faux négatifs.

Dans le cas peu probable où trop peu de marqueurs auront été détectés, notre système permettra au praticien d'extraire manuellement ceux qui sont absents (que ce soit dans l'image scanner ou vidéo). Dans une telle situation, le gain de temps que le système procure usuellement sera amoindri, néanmoins, le patient conservera le bénéfice de subir moins d'irradiation que si le protocole de guidage avait été radiologique.

Vérification des appariements

Une vérification visuelle rapide des appariements peut être réalisée après le recalage et le superpositionnement des marqueurs virtuels dans les deux images vidéo. Pour les mettre en évidence, nous entourerons d'un rectangle les marqueurs vidéo et scanner mis en correspondance (cf. figure 8.10). Pour les rares cas où la procédure d'appariement échouerait, nous développerons une application supplémentaire permettant de mettre en correspondance rapidement trois marqueurs extraits dans l'image scanner avec trois marqueurs issus des images vidéo.

Vérification permanente du calibrage

Dans la version actuelle de notre système, si une des deux caméras est très légèrement déplacée (quelques millimètres suffisent) ou bien si la focale est modifiée, alors le calibrage et le recalage calculé deviennent erronés. Le bruit sur l'extraction des marqueurs vidéo ayant des caractéristiques relativement stables, l'erreur moyenne de reprojektion des marqueurs dans les images devrait être constante durant toute l'intervention. Ainsi, en inspectant régulièrement cette valeur, nous pourrions prévenir le praticien d'un tel incident si jamais l'erreur augmente

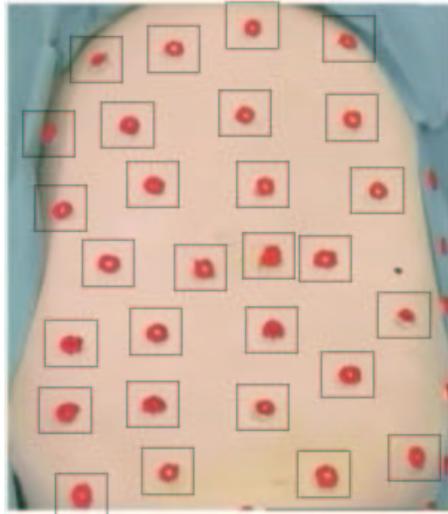


FIG. 8.10 – Vérification visuelle des appariements réalisés entre les marqueurs extraits de l'image TDM et ceux extraits des images vidéo. Des rectangles englobant les marqueurs appariés sont tracés pour permettre au praticien de vérifier rapidement leur validité.

subitement et se stabilise. En effet, une modification soudaine des paramètres de la caméra se traduit par l'ajout d'un bruit de forme particulière (translation dans le cas d'un déplacement de la caméra, homothétie si la focale change) sur l'extraction des marqueurs qui entraînera automatiquement une forte augmentation de l'erreur de reprojection.

Ce constat suggère qu'il serait préférable de lier rigidement des caméras à focale fixe, voire même de les calibrer à l'avance. Cependant, il existe peu de matériaux (l'acier n'en fait pas partie) dont le pouvoir de dilatation soit suffisamment faible pour permettre un calibrage unique durant une longue période. Par ailleurs, si un lien rigide entre les caméras offre une sécurité et un confort (puisque l'on peut alors déplacer les caméras durant l'intervention), la configuration des caméras est limitée aux dimensions de la barre les reliant, ce qui peut alors réduire significativement la précision potentielle du système (si l'écart entre les caméras reste faible). Notre procédure de calibrage ne nécessitant pas plus de trois minutes nous pensons pour l'instant que rigidifier la position relative des caméras n'est pas un choix qui s'impose.

8.3 Evaluation du système sur des patients

Notre première expérience sur un mannequin de l'abdomen s'est avérée être un succès. Cependant, bien que la forme du mannequin et la position du foie soient représentatives de la réalité, le mannequin n'est pas rempli comme un corps humain (il y a du vide entre la peau et le foie), et il n'y a aucune simulation des effets de la respiration. Nous pourrions éventuellement tenter de le perfectionner afin d'augmenter son réalisme (par exemple en le remplissant avec de la mousse et/ou en construisant un mécanisme déplaçant régulièrement le foie). Néanmoins, nous pensons que les résultats que nous obtiendrions avec ce nouveau modèle de l'abdomen seront très proches de ceux de notre première validation. De plus, même si nous obtenons des résultats satisfaisants avec une version améliorée de notre mannequin, l'étape suivante réellement significative est l'expérimentation *in vivo*. Il semble donc plus judicieux de tenter

directement une expérience sur l'animal.

Cependant, la réservation d'un créneau horaire au scanner de l'hôpital de Strasbourg nécessite une longue attente (il faut six mois au minimum pour obtenir les autorisations et pour réunir toutes les personnes nécessaires à une intervention de ce type). Nous avons donc pensé réaliser des expériences directement sur des patients (il y a au moins une opération de radiologie interventionnelle par semaine à l'hôpital civil de Strasbourg). Néanmoins, il n'est pas envisageable de demander au praticien de se servir de notre système pour se guider étant donné que la moindre erreur dans nos évaluations d'erreur pourrait mettre en danger la vie du patient. Il faudrait donc parvenir à tester notre système sans que cela représente un quelconque risque pour le patient. Dans cette section, nous décrivons préalablement le protocole expérimental que nous avons mis au point qui permet d'évaluer la précision de notre système par rapport au praticien sans que cela ait une incidence sur le patient pendant l'intervention. Nous présentons ensuite les résultats que nous avons obtenus sur 3 cas cliniques.

8.3.1 Un protocole de validation sans danger pour le patient

Évaluation de la précision de notre système par rapport au praticien

Le but de ce protocole d'évaluation est de mesurer l'écart entre la position atteinte par le praticien en se servant du protocole radiologique standard et la position que nous lui aurions suggéré d'atteindre s'il s'était servi de notre système. L'opération se déroule normalement comme décrit dans la section 2.1.3. Pour établir nos mesures, nous devons demander au praticien de fixer un marqueur carré sur l'aiguille, de coller des marqueurs radio-opaques sur la peau du patient et de positionner nos deux caméras dans la salle d'opération. Conformément au protocole usuel, l'abdomen du patient est scanné une première fois. Sur cette acquisition SCAN1 que nous récupérons, nous demandons au praticien d'indiquer les coordonnées du point C^{scan} qu'il va tenter d'atteindre avec la pointe de l'aiguille. Nous le laissons alors réaliser son opération comme il le fait d'habitude en se guidant à l'aide de multiples acquisitions scanner jusqu'à ce qu'il estime avoir atteint la cible qu'il s'était fixée. Nos deux caméras étant préalablement calibrées²⁸, nous réalisons alors une acquisition vidéo contenant l'abdomen du patient (avec les marqueurs radio-opaques) et le marqueur carré lié à l'aiguille. L'aiguille ayant été calibrée, nous pouvons déterminer la position de la pointe P^{cam} dans le référentiel des caméras (aux erreurs de repérage près). Les coordonnées de P^{cam} ainsi calculées correspondent donc à la position de l'aiguille que le praticien estime correct.

Les marqueurs radio-opaques étant localisés dans le scanner et les images vidéo, nous pouvons recalculer l'image scanner SCAN1 du patient dans le repère des caméras et donc calculer dans ce repère les coordonnées du point cible C^{cam} initialement choisi par le praticien dans le SCAN1. Le point cible recalculé correspond à la cible vers laquelle nous aurions guidé le praticien s'il s'était servi de notre système. Considérant que P^{cam} est une mesure étalon fournie par le praticien, nous pouvons alors mesurer la distance qui sépare P^{cam} de C^{cam} et évaluer ainsi l'erreur ϵ commise par notre système.

Évidemment, le praticien ne parvient jamais à positionner parfaitement l'aiguille là où il souhaiterait la placer. Les chirurgiens à l'IRCAD estiment qu'ils ont une précision de l'ordre de 3 mm. En conséquence, si l'erreur ϵ que nous mesurons vaut en moyenne 3 mm, cela signifierait que la précision de notre système et celle du praticien seraient approximativement équivalentes.

²⁸ Les deux caméras peuvent aussi être calibrées après l'opération. Cela n'a pas d'importance puisque nous pouvons effectuer tous nos traitements après l'intervention.

Nous démontrerions alors l'intérêt de notre système puisqu'il permettrait, à précision équivalente, de réduire le temps de l'intervention et de se passer d'un protocole radiologique irradiant le patient.

Évaluation de l'erreur du praticien

Bien que les praticiens aient une idée quantitative de leur précision, il serait préférable de l'évaluer à l'aide d'un protocole expérimental. Dans le cadre du traitement d'une tumeur par radio-fréquence, un scanner SCAN2 du foie est généralement acquis après thermo-ablation de la cible afin d'examiner la quantité de cellules nécrosées. En général, l'aiguille n'est pas extraite du patient pour ne pas avoir à la réintroduire si jamais la thermo-ablation n'est pas suffisante, elle est donc présente dans l'image scanner SCAN2. Ainsi, si nous parvenons à segmenter les foies dans les images SCAN1 et SCAN2 et à les recalculer, nous pourrions alors déterminer à quel distance de la cible initiale le praticien a placé la pointe de l'aiguille. Nous obtiendrions alors une estimation de l'erreur de positionnement du praticien à laquelle comparer l'erreur ϵ de notre système calculée par rapport aux praticiens.

Contrainte du protocole

Le protocole que nous proposons est indépendant de l'acte médical réalisé par le praticien pendant l'intervention. Cependant, nous devons demander au praticien de se servir d'une aiguille sur laquelle est fixé le marqueur carré qui nous sert à la repérer. Les exigences hygiéniques de cette intervention étant la stérilisation du matériel utilisé (et donc de l'aiguille), nous avons dû préalablement concevoir un marqueur susceptible de supporter l'étape de stérilisation (température à 134° C sous une pression de 2 bars.).

Afin de réduire au minimum le dérangement occasionné par la présence du marqueur sur l'aiguille, nous l'avons fait confectionner en un plastique certifié conforme à la stérilisation médicale. Notre marqueur est donc léger (son poids est inférieur à 100 grammes) et n'occasionne quasiment aucun artefact dans les images scanner. L'artisan n'ayant pu respecter les contraintes de planéité demandée (l'erreur de planéité observée sur le marqueur est supérieure à 1 millimètre), nous avons dû réaliser un outil permettant de calibrer précisément la position relative des 4 coins du marqueur carré (cf. Annexe D).

8.3.2 Expériences sur trois cas cliniques

Pour réaliser notre première évaluation du système en conditions opératoires, nous avons choisi les trois premiers patients qui allaient subir une thermo-ablation par radio-fréquence à l'hôpital civil de Strasbourg. Le premier patient souffrait d'une tumeur située dans un os du bassin, et les deux patients suivants de tumeurs hépatiques. Tout d'abord, nous décrivons les conditions expérimentales de chacune des expériences. Ensuite nous présentons les résultats obtenus et discutons des différences de précision observées entre l'expérience avec le mannequin et celles avec les patients.

Patient 1 : radio-fréquence osseuse

Dans cette expérience, la pathologie visée était une tumeur osseuse située au niveau du bassin. Le patient était sous anesthésie générale et sa respiration était contrôlée. Nous avons collé 18 marqueurs radio-opaques sur la peau du patient, autour de la zone d'entrée de l'aiguille.

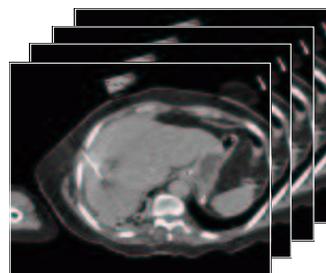
Pose des marqueurs radio-opaques sur le patient.
Le marqueur carré est fixé sur l'aiguille.
Les deux caméras sont installées.



Acquisition scanner SCAN1.
Le praticien définit la cible qu'il
va essayer d'atteindre.



Le praticien positionne l'aiguille
à l'aide du protocole radiologique
habituel.



Acquisition vidéo du patient et
de l'aiguille.
Calcul des coordonnées de la pointe
de l'aiguille dans le repère des caméras.



Recalage de l'image SCAN1 dans
le repère des caméras grâce aux
marqueurs radio-opaques.
Nous pouvons alors calculer les
coordonnées du point cible dans le
repère des caméras



Recalage 3D/2D



Calcul de la distance 3D entre la pointe
de l'aiguille et la cible vers laquelle nous
aurions guidé le praticien avec notre
système.



FIG. 8.11 – Schéma descriptif du protocole permettant de comparer la précision de notre système à celle du praticien.

Après une installation et un ajustement rapide des paramètres de nos deux caméras, nous les avons calibrées juste avant la première acquisition scanner.

Le radiologue a estimé que les mouvements respiratoires n'avaient pas d'influence sur la zone traitée, la respiration du patient a donc été régulière pendant toute l'opération et n'a jamais été interrompue (même pendant l'acquisition scanner). Après avoir récupéré cette première acquisition, nous avons demandé au radiologue d'indiquer dans l'image TDM les coordonnées 3D du voxel qu'il désirait atteindre avec la pointe de l'aiguille, conformément au protocole décrit dans la section 8.3.1. Ensuite, nous avons extrait automatiquement les marqueurs et réalisé le recalage reliant le repère des caméras à celui du scanner. Nous avons alors pu vérifier visuellement que le recalage était de bonne qualité (cf. Figure 8.13).

Le radiologue débuta alors l'intervention. Avant d'introduire l'aiguille, il fixa notre marqueur carré dessus. Il réalisa son positionnement dans les conditions habituelles, c'est à dire en se dirigeant à l'aide de multiples acquisitions scanner de contrôle (cf. section 2.1.3). Les risques d'infection étant bien plus importants pendant le traitement de tumeurs osseuses que lors de tumeurs hépatiques, le radiologue souhaite ne pas élargir le champ opératoire, comme nous le lui avons demandé, pour des raisons hygiéniques. En conséquence, nous ne pouvions plus réaliser le recalage entre le repère des caméras et celui du scanner en temps réel. Pour contourner ce problème, nous avons demandé à la manipulatrice radio de repositionner la table du scanner à la même abscisse chaque fois qu'elle ressortait le patient du tunnel. Le recalage n'a donc été effectué qu'une seule fois avant la pose du champ opératoire.

Lorsque le radiologue termina le positionnement de l'aiguille, nous avons mesuré la distance entre la pointe de l'aiguille virtuelle et la cible vers laquelle nous aurions guidé le radiologue. En fait, les effets induits par la respiration n'étaient pas aussi négligeables que le pensait le radiologue. Certains des marqueurs subissaient un déplacement de l'ordre de 2 mm. Par ailleurs nous avons visuellement constaté sur l'acquisition TDM de la patiente que la respiration était importante (cf. figure 8.12). L'aiguille étant légèrement déplacée au cours de la respiration, nous avons pris plusieurs paires d'images vidéo pendant plusieurs cycles respiratoires et avons moyenné les résultats. Un exemple de la position relative de l'aiguille par rapport à la cible initiale fixée par le radiologue est visible sur la figure 8.14.

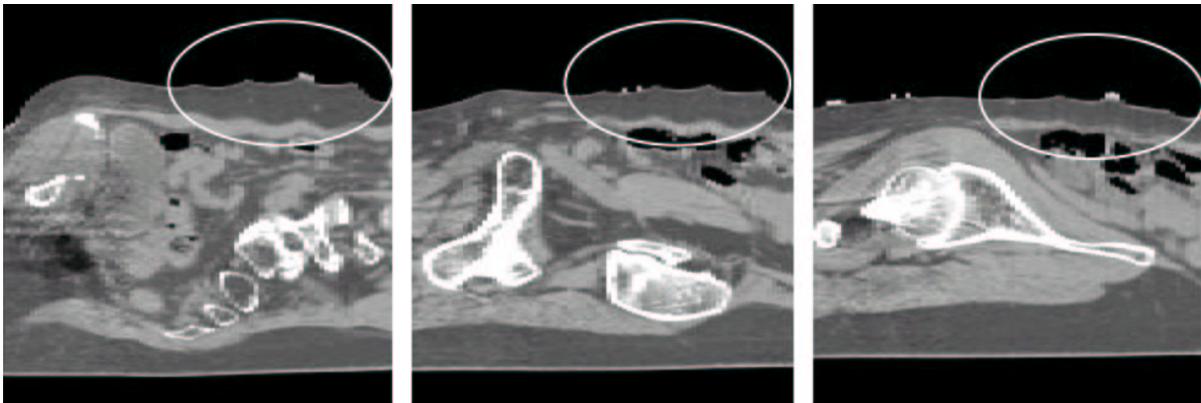


FIG. 8.12 – Mise en évidence de l'impact de la respiration dans l'image TDM. Les 3 images correspondent à 3 coupes sagittales du bassin de la patiente, distantes d'environ 10 cm. Nous voyons dans les ellipses blanches de petites vagues sur la peau qui sont significatives d'une respiration importante pendant l'acquisition TDM.

Patient 2 : radio-fréquence hépatique

Pour cette expérience, la pathologie visée était une tumeur hépatique. Le patient étant régulièrement hospitalisé et anesthésié, il était souhaitable d'éviter une anesthésie générale. Comme la tumeur était située dans une zone peu vascularisée du foie (et donc peu douloureuse), l'intervention s'est déroulée sous anesthésie locale. Le patient était donc éveillé et pouvait respirer librement. Afin de limiter les mouvements du foie, le radiologue a demandé au patient de retenir sa respiration pendant l'acquisition scanner et à chaque fois qu'il manipulerait l'aiguille. Nous avons effectivement pu vérifier sur l'image TDM pré-opératoire l'absence d'artefact signifiant un mouvement respiratoire important.

Par ailleurs, la tumeur étant peu visible dans les images TDM, le radiologue s'est aidé d'une sonde échographique pour vérifier la qualité du positionnement de l'aiguille. La manipulation d'une telle sonde nécessitant de la place sur la peau près de l'aiguille, nous avons dû coller nos marqueurs plus loin que prévu de la zone d'entrée de l'aiguille. De plus, le radiologue n'a pas respecté notre demande concernant la dimension du champ opératoire. En effet, on voit sur l'image 8.15 que seulement 7 marqueurs sur 18 sont visibles (et pas entièrement).

Afin d'effectuer un recalage tenant compte de la totalité des marqueurs visibles dans l'image TDM, nous avons demandé à la manipulatrice radio de repositionner la table à la même abscisse à chaque fois que le patient était sorti de l'anneau du scanner (comme pour le patient 1). Malheureusement, le patient a bougé sur la table durant l'intervention, de sorte que la transformation que nous avons calculée avant la pose du champ opératoire n'était plus valable. Pour réaliser notre mesure d'erreur, nous avons donc effectué un recalage post-opératoire en détectant manuellement les 7 marqueurs visibles dans les images vidéo. Sur la figure 8.15 (en bas) nous montrons la qualité du recalage du modèle reconstruit de la peau, et sur les images de la figure 8.16 nous présentons les résultats de réalité augmentée obtenus avant et après la pose du champ opératoire. Finalement, nous montrons sur la figure 8.17 la position de l'aiguille virtuelle par rapport à la cible initialement visée par le radiologue.

Patient 3 : radio-fréquence hépatique avec deux tumeurs

Pour ce dernier cas clinique, le foie du patient comptait deux tumeurs à traiter. Pour des raisons équivalentes à celles présentées pour le patient 2, ce patient ne fut pas sous anesthésie générale pendant l'intervention. Afin d'obtenir un champ opératoire suffisamment large tout en satisfaisant des conditions hygiéniques optimales, nous avons convenu avec le praticien de coller des "steristrip" (bandes autocollantes transparentes et stériles) autour du point d'entrée de l'aiguille et sur les marqueurs (cf. image en haut à gauche de la figure 8.18). Nous avons ainsi pu repérer durant toute l'intervention la position des marqueurs radio-opaques. Il ne fut donc pas nécessaire (comme ce fut le cas pour les patient 1 et 2) de repérer la position de la table du scanner et de la replacer précisément à la même abscisse.

Les figures 8.18 et 8.19 montrent les résultats de réalité augmentée pour les deux tumeurs traitées. Ces images permettent de voir que 5 marqueurs sur les 14 collés n'étaient pas visibles dans l'acquisition TDM. Nous avons donc dû réaliser le recalage avec les 9 marqueurs les plus éloignés du foie. Par ailleurs, le patient a été légèrement déplacé après la destruction de la première tumeur. L'image TDM, réalisée au début de l'intervention, était donc moins représentative de la réalité lors de la destruction de la seconde tumeur : sa position par rapport aux marqueurs radio-opaques était donc sensiblement différente. Ce phénomène est visible dans l'image en haut de la figure 8.19 : on constate nettement le mauvais alignement de la reconstruction virtuelle sur le patient (dans l'ellipse noire). Nous verrons d'ailleurs dans la

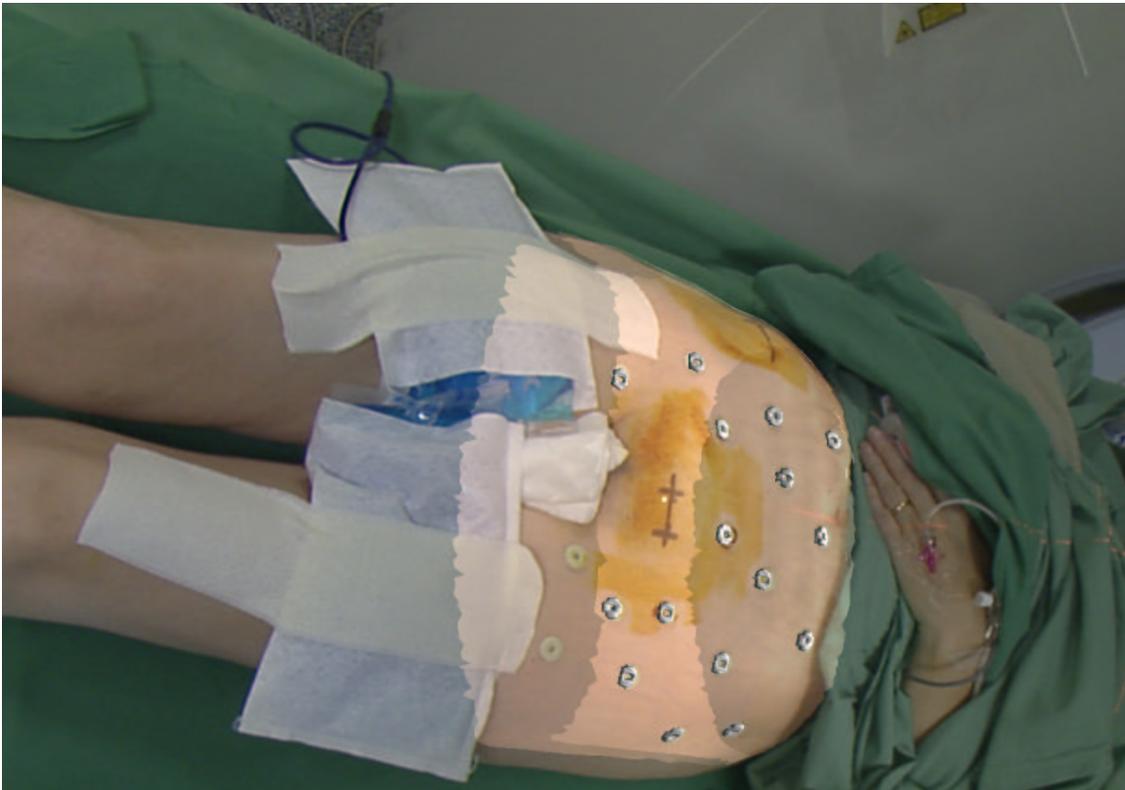
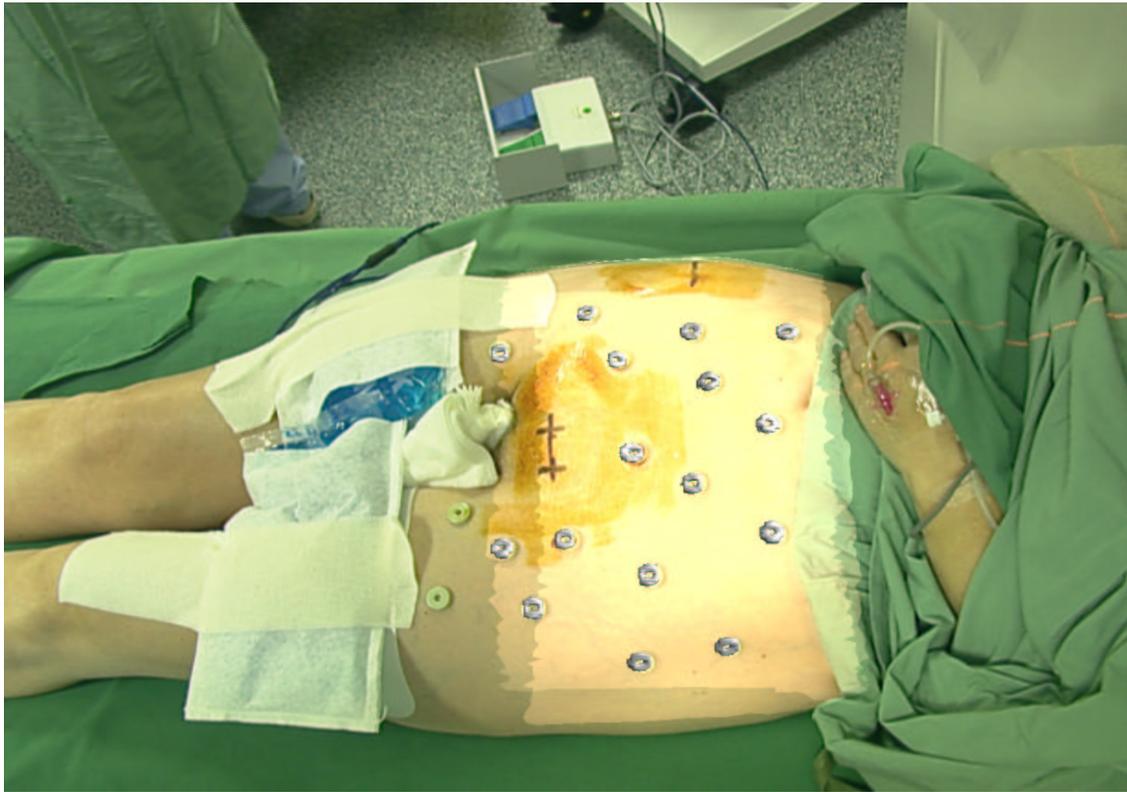


FIG. 8.13 – Vérification visuelle de la qualité du recalage. Les marqueurs du modèle reconstruit sont bien superposés sur ceux visibles dans les images vidéo. On peut remarquer que deux des marqueurs ne sont pas apparus sur l'acquisition scanner. D'un point de vue qualitatif, tous les marqueurs présents dans les images vidéo et TDM ont été extraits, et le nombre d'appariement retrouvé a été optimal (tous les marqueurs extraits de l'image TDM ont été appariés correctement).

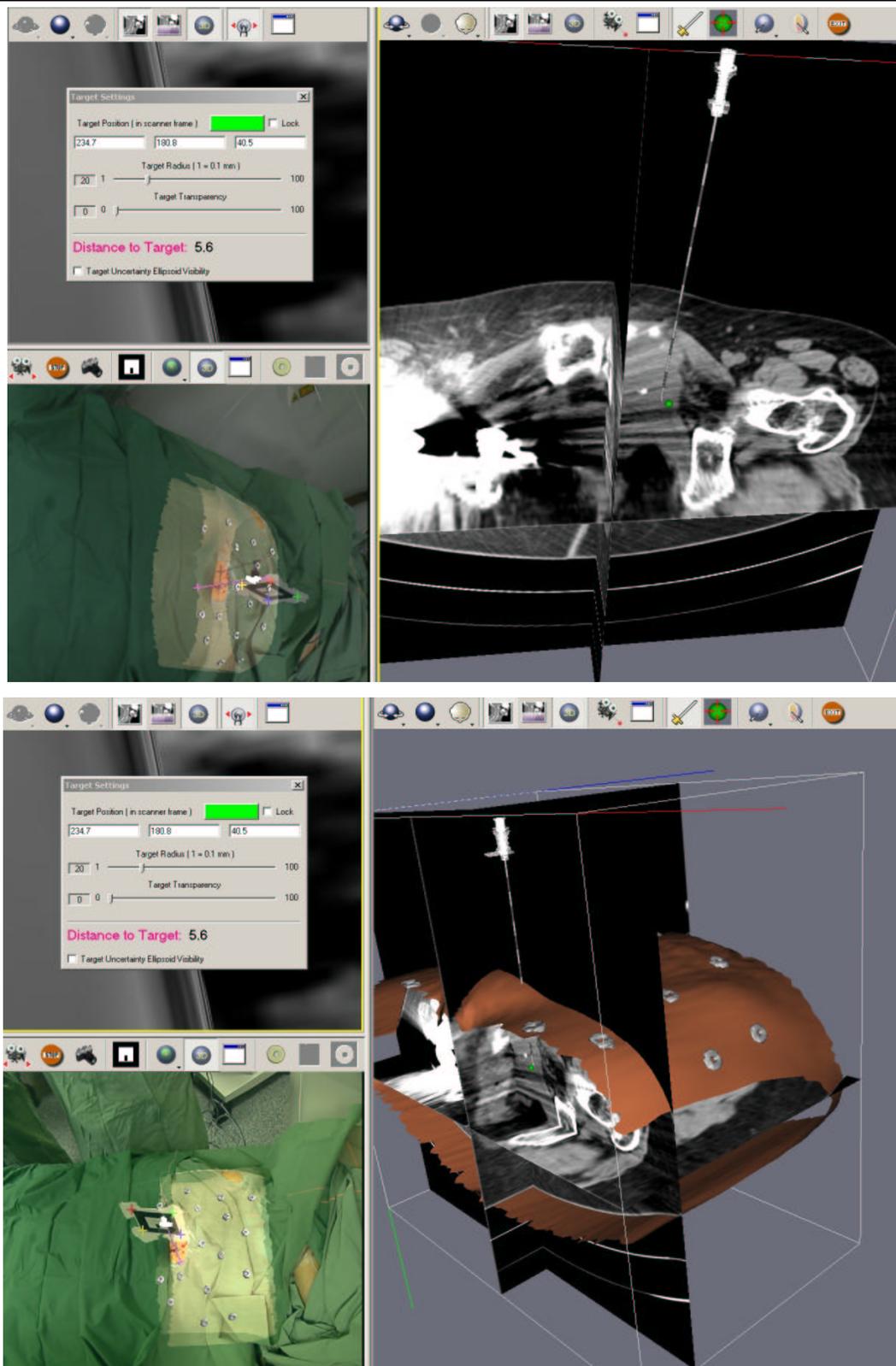


FIG. 8.14 – Vérification visuelle du positionnement de l'aiguille par rapport à la cible initialement prévue par le radiologue (sphère verte). Les deux images n'ont pas été prises au même moment du cycle respiratoire. L'aiguille étant à une position légèrement différente d'une image à l'autre, les distances mesurées ne sont pas identiques. Dans l'image du haut, la distance séparant la pointe de l'aiguille de la sphère vaut 5.8 mm, et dans l'image du bas 6.8 mm.

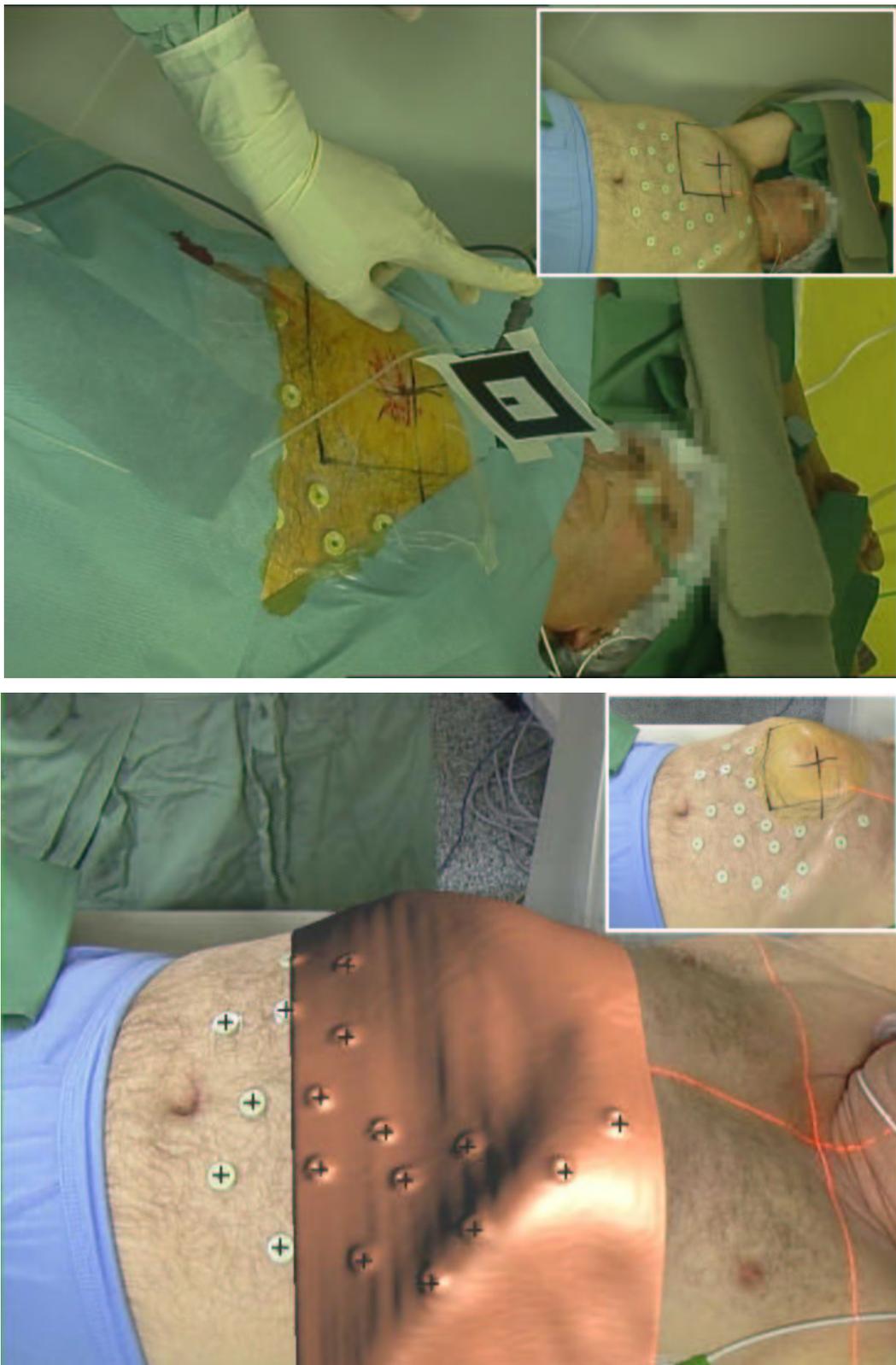


FIG. 8.15 – *En haut*: image du patient après la pose du champ opératoire. Sur un total de 18 marqueurs collés, seulement 7 étaient décelables dans nos images vidéo. *En bas*: illustration de la qualité du recalage du modèle reconstruit dans l'image vidéo. Les croix noires correspondent au résultat de la détection automatique des marqueurs radio-opaques dans l'image vidéo. On peut voir que les marqueurs visibles sur le modèle reconstruit sont correctement superposés sur les croix noirs.

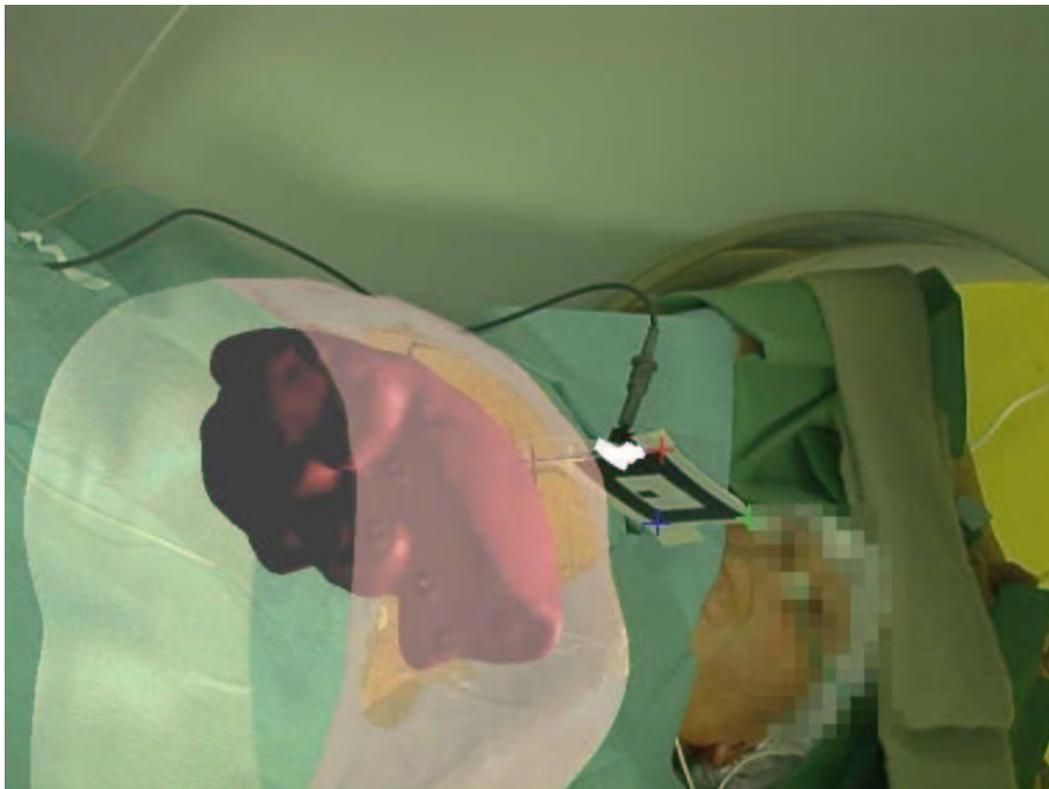
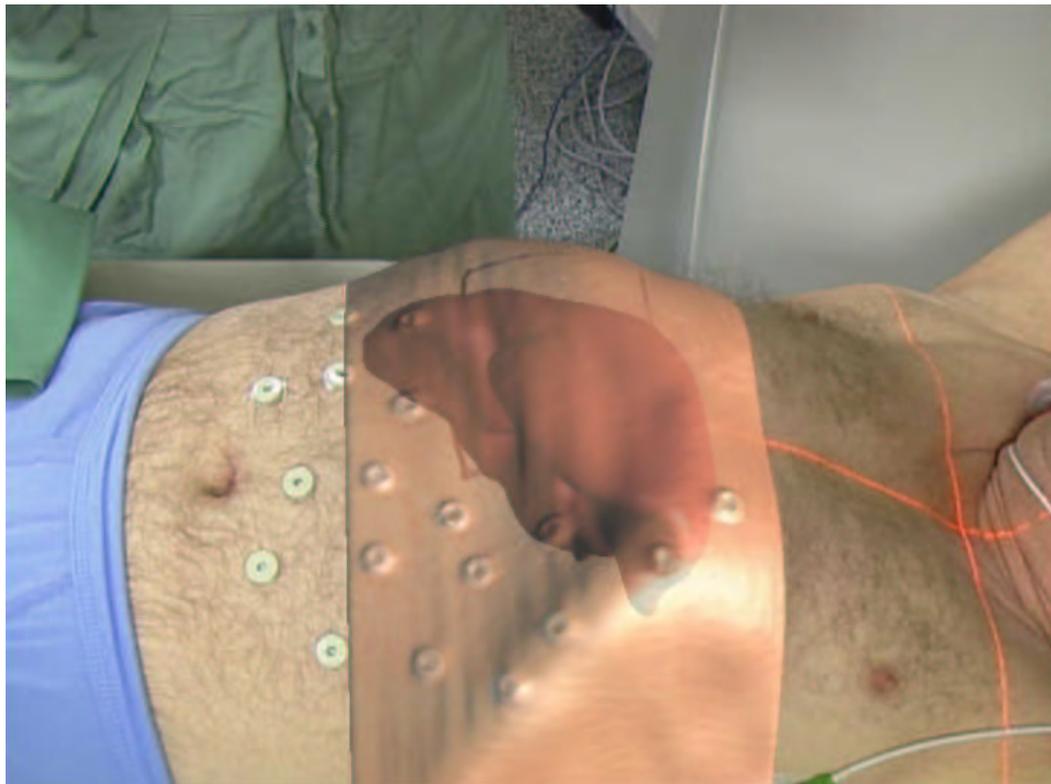


FIG. 8.16 – *En haut*: illustration de la vue de réalité augmentée. La peau transparente permet de voir la position virtuelle du foie du patient. Dans cet exemple, le recalage a été réalisé avec 14 marqueurs. *En bas*: illustration de la vue de réalité augmentée pendant la manipulation de l'aiguille. Le champ opératoire masquant plusieurs marqueurs, le recalage n'a été réalisé qu'avec 7 marqueurs, interactivement localisés.

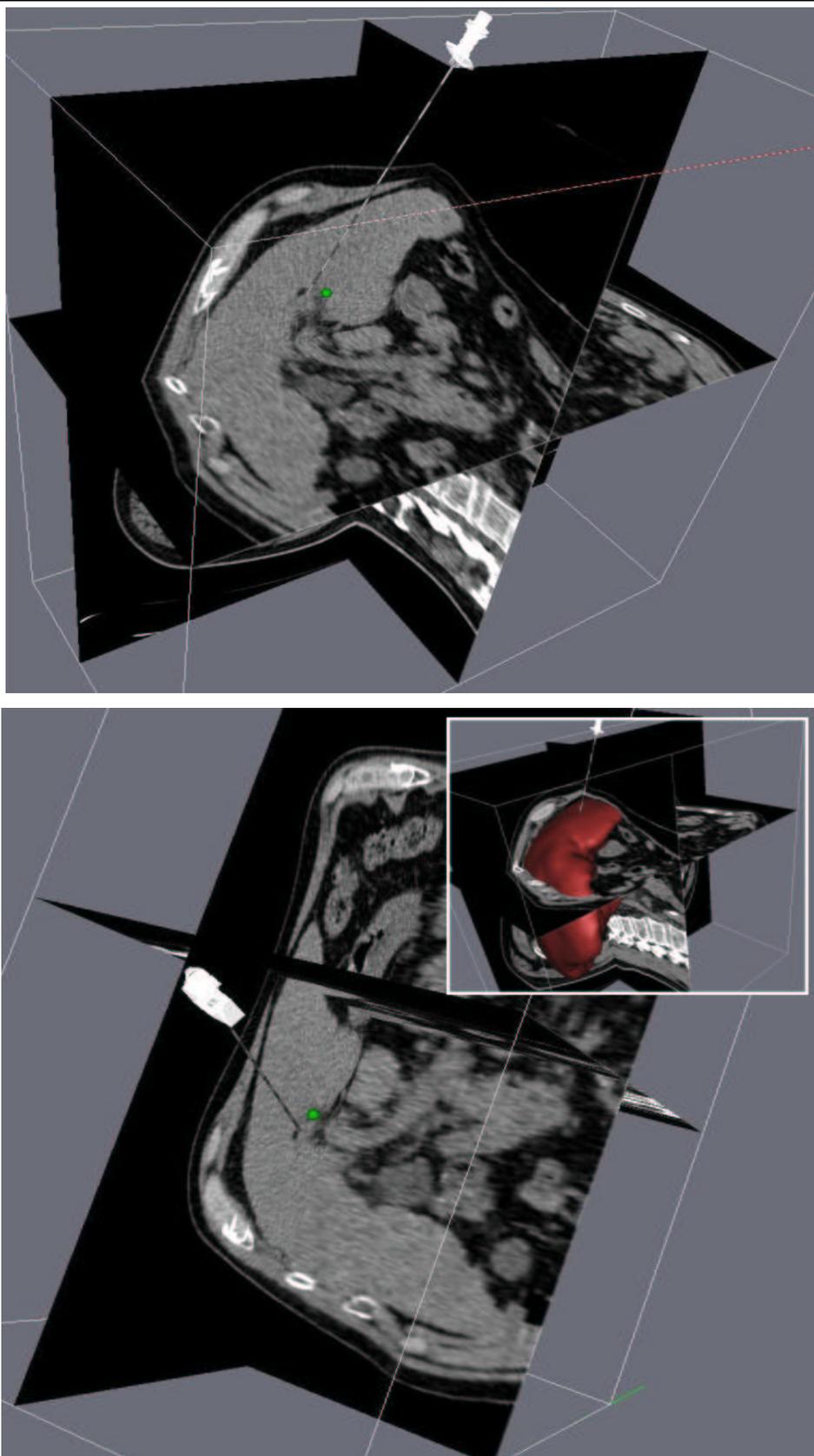


FIG. 8.17 – Vérification visuelle, selon deux points de vue différents, du positionnement de l'aiguille virtuelle par rapport à la cible initialement prévue par le radiologue (sphère verte). Nous avons affiché la reconstruction du foie dans l'encadré de l'image du bas.

sous-section suivante que la précision du système a été moins bonne pour la seconde tumeur.

Des résultats encourageants dans des conditions difficiles

Nous résumons dans le tableau 8.3 les précision obtenues par rapport au praticien pour chacun des patients. Nous remarquons que la précision moyenne est bien supérieure à celle que nous avons obtenue avec le mannequin (9.5 mm contre 2.3 mm). Cette différence provient de deux faits. Le premier est le nombre de marqueurs utilisés pour le recalage. Dans le cas du mannequin, 25 marqueurs étaient employés, alors que pour nos trois patients, nous n'avons pu nous servir que de 9.7 marqueurs en moyenne sur les 15-16 collés sur la peau. De plus, dans le cas des patients 2 et 3, les marqueurs utilisés étaient très éloignés du foie, ce qui a augmenté l'imprécision du recalage de la tumeur. L'influence du facteur "nombre et position des marqueurs" sur l'imprécision a pu être évaluée grâce au calcul de l'erreur théorique de recalage de la cible. Nous avons effectivement observés que dans les conditions propres aux patients 2 et 3, la précision théorique du recalage de la tumeur valait en moyenne 4.4 mm contre 2.8 mm auparavant.

	Anesthésie générale	Mises en apnée	Nombre de marqueurs utilisés	Précision en mm	Précision théorique (mm)
Patient 1	oui	non	14	6.7	2.8
Patient 2	non	non	7	12.0	4.9
Patient 3 : tumeur 1	non	non	9	8.0	4.2
Patient 3 : tumeur 2	non	non	9	11.3	4.7
Moyenne	-	-	9.75	9.5	4.15

TAB. 8.3 – Résultats de l'évaluation du système sur trois patients. La précision correspond à la distance mesurée entre le positionnement de l'aiguille par le praticien et la cible vers laquelle nous l'aurions guidée. La précision théorique correspond à l'erreur de recalage prédite pour la cible préalablement défini par le praticien dans le scanner pré-opératoire du patient. La faible précision obtenue (9.5 mm contre 2.3 mm avec le mannequin) provient du nombre réduit de marqueurs utilisés pour le recalage et du fait que les patients n'ont pas été mis - ou n'ont pas pu être mis - en apnée durant les interventions.

Le second facteur expliquant la faible précision obtenue est le fait que les patients n'aient pas été mis en apnée répétitivement dans les mêmes conditions volumétriques, contrairement à ce que nous souhaitions. Dans le cas du patient 1, c'est une erreur d'appréciation du praticien qui en est la cause; pour les deux autres patients, ce sont des raisons de sécurité cliniques.

Ces résultats montrent donc qu'une précision moyenne en dessous de 1 cm reste difficile à obtenir pour le traitement de tumeurs hépatiques lorsque le patient n'est pas sous anesthésie générale et que des mises en apnée à volume d'air constant ne sont pas effectuées. En revanche, dans le cas de tumeurs osseuses éloignées des poumons, le résultat obtenu (6.7 mm) suggère qu'une précision suffisante est possible sans mise en apnée pour traiter des tumeurs d'une taille supérieure à 3 cm.

Par ailleurs, nous avons pu constater qu'un patient sous anesthésie locale ne semble pas parvenir à conserver une position statique par rapport à la table du scanner. En effet, une vérification sur les images vidéo acquises nous a révélé que le patient 2 s'était déplacé d'environ 1.5 cm le long de la table pendant l'intervention, sans que nous nous en soyons rendus compte. Ce constat souligne la nécessité d'un repérage en temps réel des marqueurs collés sur le patient. Les

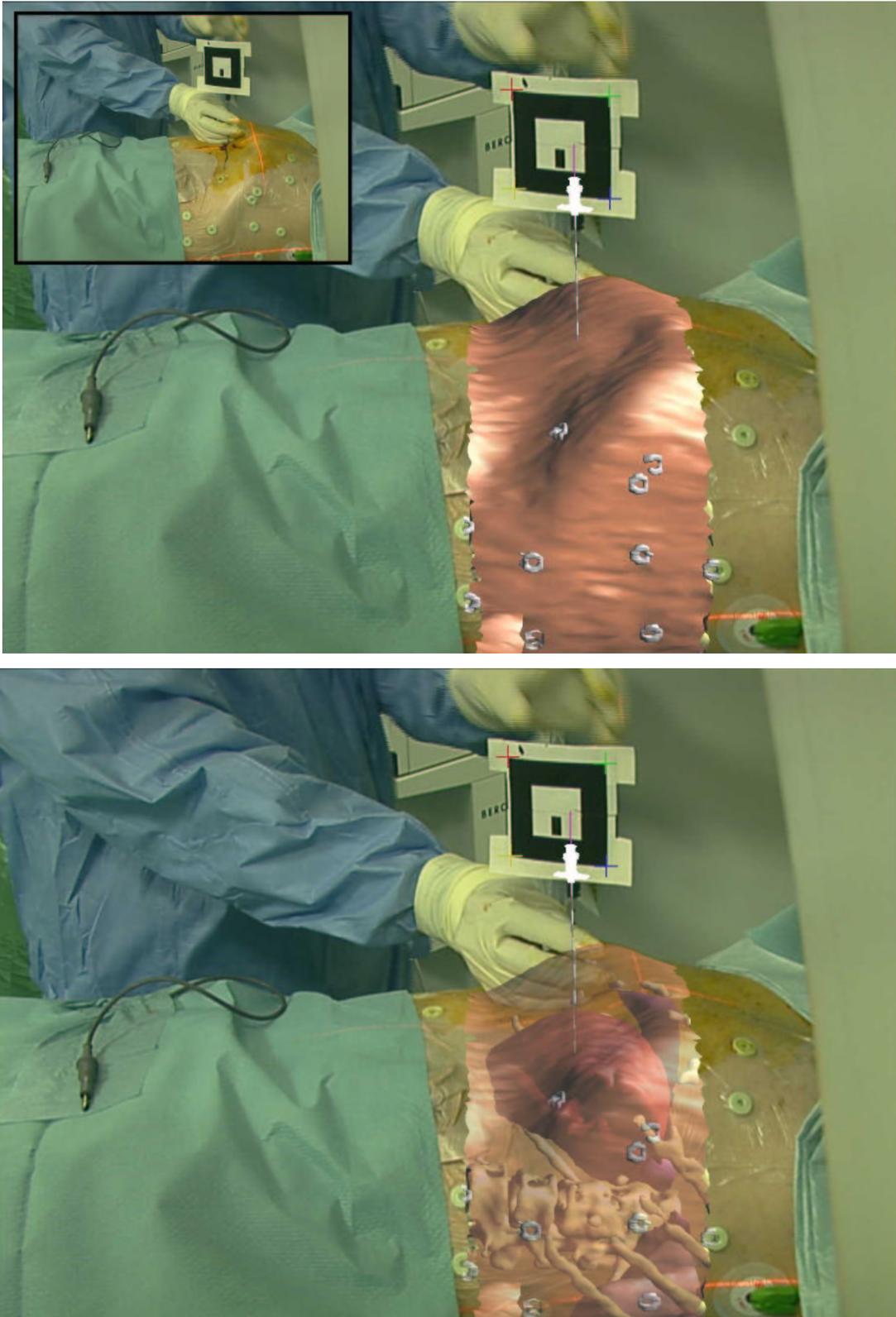


FIG. 8.18 – Illustration de la vue de réalité augmentée lors du traitement de la première tumeur. En haut : la peau est opaque. En bas : la peau semi-transparente permet de voir les organes sous la peau du patient. Le recalage a été réalisé avec 9 marqueurs automatiquement localisés.

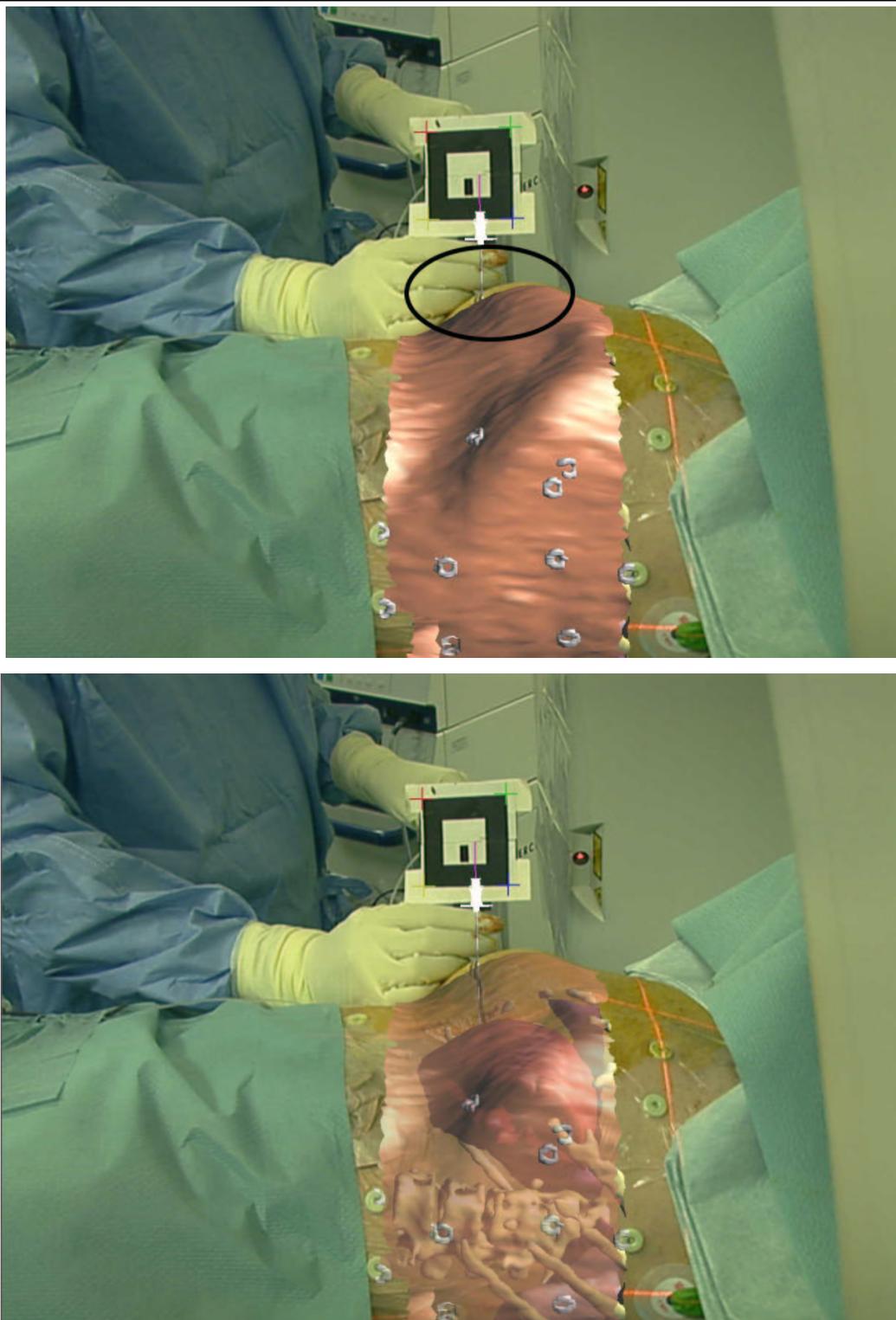


FIG. 8.19 – Illustration de la vue de réalité augmentée lors du traitement de la seconde tumeur. En haut : la peau est opaque et nous avons surligné dans l'ellipse noire le mauvais alignement de la reconstruction virtuelle sur le patient. Ce phénomène provient du fait que le patient a été légèrement déplacé après le traitement de la première tumeur et que l'image TDM n'a pas été réactualisée. En bas : la peau semi-transparente permet de voir les organes sous la peau du patient. Le recalage a été réalisé avec 9 marqueurs automatiquement localisés.

systèmes de guidage nécessitant un pointage manuel des marqueurs sur le patient, risqueraient donc de subir un biais significatif, ce qui rendrait leur emploi dangereux pour le patient.

Finalement, cette évaluation nous a permis de vérifier le bon fonctionnement de tous nos algorithmes en conditions opératoires. Bien que les traitements aient été menés avec succès, nous planifions une nouvelle validation de la détection des marqueurs sur la peau lorsque les bandes de “steristrip” sont collées dessus. En effet, il semble important de s’assurer que les réflexions spéculaires induites par ces bandes n’affectent pas la précision de notre algorithme d’extraction.

Une validation future sur des cas sélectionnés

Les expériences menées sur les patients nous ont montré la difficulté d’une validation sur un patient. En effet, notre étude ne porte pas sur un cadavre ou sur un animal, auquel cas les conditions sont très facilement contrôlables puisque le temps ne fait pas défaut et qu’il n’y a pas d’enjeu vital. Bien que notre protocole expérimental ne soit pas théoriquement dangereux pour le patient, nous demandons au praticien de nous accepter dans la salle avec notre matériel, et de faire un effort supplémentaire durant l’intervention (coller les marqueurs, manipuler l’aiguille avec un marqueur attaché dessus). Il essaie de mener son intervention comme d’habitude en tenant compte de paramètres inhabituels dus à notre présence. Il ne faut donc pas être étonné par le fait, par exemple, qu’il n’ait pas agrandi suffisamment le champ opératoire sur le patient 2 : il est principalement préoccupé par son geste opératoire, qu’il doit mener en un temps limité (d’autres patients attendent que le scanner soit disponible).

Afin de mener une validation dans les meilleures conditions possibles, nous devons limiter au maximum le stress du praticien provoqué par notre intrusion dans la salle d’opération. Actuellement, nous introduisons un PC de bureau, deux volumineux trépieds, deux caméras et de nombreux câbles gênants laissés au sol. Nous projetons de limiter notre système à un PC portable, un trépied et deux petites caméras haute définition USB2 rigidement fixées sur une barre (afin de pouvoir procéder au calibrage en dehors de la salle d’opération). De même nous essaierons de réduire au maximum le poids du porte aiguille, qui, selon le praticien, est encore un peu trop lourd.

L’étape suivante consistera en une validation du système sur des patients préalablement sélectionnés. Les conditions requises seront l’anesthésie générale du patient et les mises en apnée dans des conditions volumétriques d’air identiques dans les poumons. En effet, notre évaluation a confirmé le fait que nous ne pourrions pas obtenir une précision en dessous de 5 mm sans le respect de ces conditions (qui sont respectées dans la plupart des thermo-ablations par radio-fréquence). Les trois patients ayant servis à notre première expérience n’auraient donc pas été choisis pour la validation. Concernant le positionnement des marqueurs sur le patient, nous devons coopérer étroitement avec la manipulatrice radio et le praticien afin de prédéfinir la zone qui sera scannée, et éviter ainsi de devoir effectuer un recalage avec moins de marqueurs que prévu du fait qu’ils ne sont pas visibles dans l’image TDM.

Finalement, notons que l’IRCAD a fait l’acquisition d’un micro-scanner pour petits animaux qui sera livré en septembre 2004. Il nous sera alors possible d’effectuer des expériences in vivo afin d’étudier les moyens de nous passer des mises en apnée et de l’anesthésie générale du patient, indispensable actuellement pour obtenir une précision compatible avec les exigences cliniques.

8.4 Conclusion

Dans ce chapitre, nous souhaitons montrer que notre système fonctionnait correctement lorsque tous les différents modules étaient assemblés, et que nous pouvions envisager une première expérience sur des patients. Nous avons donc tout d'abord réalisé une validation de l'ensemble de notre système sur le mannequin en utilisant des marqueurs radio-opaques pour cibles et une caméra endoscopique pour la vérification. Les résultats ont été très satisfaisants puisque la précision globale a été de 3.1 mm, ce qui est en dessous de la précision minimale définie dans notre cahier des charges. De même, la durée moyenne nécessaire au positionnement de l'aiguille, proche de 30 secondes, est bien inférieure au temps nécessaire au praticien dans les conditions classiques (environ 10 minutes). Par ailleurs, la comparaison entre les ciblage réalisés par le concepteur du système et le chirurgien nous a permis de vérifier que notre système favorise l'expertise du geste et non pas celle du système développé. Cette expérience ayant souligné certaines lacunes de notre système, nous avons alors présenté les solutions ergonomiques conçues avec le chirurgien qui permettront de guider le praticien tout au long de l'intervention et de lui assurer le contrôle du système en cas de défaillances d'un de nos algorithmes.

Dans une dernière partie, nous avons montré que les contraintes liées à une expérience sur l'animal sont trop lourdes. Afin de nous en affranchir, nous avons présenté une méthode d'évaluation sans danger pour le patient qui nous a permis de comparer la précision de notre système à celle d'un praticien sur 3 cas cliniques.

Ces expériences nous ont fourni l'occasion de tester notre système dans les conditions classiques de radiologie interventionnelle. Nous avons ainsi pu valider le fonctionnement des modules de traitement automatique (extraction et appariement des marqueurs radio-opaques) avec des données provenant de cas cliniques. De plus, nous avons pu vérifier que le temps d'installation de notre système est compatible avec l'application clinique. Cette évaluation préliminaire démontre la faisabilité d'un guidage par réalité augmentée en conditions réelles. En effet, la précision moyenne obtenue, bien qu'insuffisante (entre 6 et 10 mm), est très encourageante puisque certaines des conditions requises pour la validation future n'étaient pas respectées. Nous sommes donc confiant quant au succès de la validation du système. Finalement, nous mettons en avant le fait que, à notre connaissance, nous fournissons pour la première fois une évaluation sur *des patients humains* d'un système de guidage de ponctions percutanées *dans la zone abdominale*.

Chapitre 9

Conclusion

Nous nous étions donné comme objectif de concevoir un système de réalité augmentée pour les opérations du foie en radiologie interventionnelle. Afin que notre système apporte un bénéfice substantiel, nous devons surmonter des contraintes de précision, de rapidité et de fiabilité. Pour satisfaire toutes ces exigences et fournir un système adapté à la demande des praticiens, nous avons examiné les conditions cliniques présentes en radiologie interventionnelle. Nous avons tout d'abord analysé les effets de la respiration lorsque le patient est intubé, ce qui nous a permis de choisir une technique de recalage appropriée pour obtenir la meilleure précision possible. Pour rendre le système rapide, le traitement des données nécessaires au système a été automatisé. Finalement, une étude de la propagation des erreurs a été menée pour prédire l'erreur moyenne de recalage.

Dans le détail, nous pouvons séparer l'ensemble de notre travail en trois parties.

9.1 Conclusion théorique

Un nouveau critère de recalage ponctuel 3D/2D

Les conditions en radiologie interventionnelle nous ont conduits à dériver un nouveau critère de recalage ponctuel 3D/2D qui tient compte, grâce à l'introduction de variables cachées, non seulement du bruit 2D, mais aussi du bruit 3D. Nous avons ensuite montré sur un large éventail de données synthétiques que notre critère (baptisé EPPC) surpassait en terme de précision et de robustesse les critères de recalage standard. Cette évaluation des performances nous a d'ailleurs permis de mettre en évidence que l'EPPC apporte un gain considérable (supérieur à 30%) lorsque l'angle entre les caméras est faible (inférieur à 30°) et que le ratio des rapports signal sur bruit sur l'extraction des données 2D et 3D (RSB_{2D}/RSB_{3D}) est supérieur à 2.

Pour confirmer les conclusions de l'étude synthétique, nous avons souhaité utiliser des données réelles provenant d'un mannequin de l'abdomen. Comme nous n'avions pas de mesure étalon, nous avons mis au point un protocole d'évaluation original, ne nécessitant aucun étalon, basé sur de multiples boucles de recalage. Ainsi, nous avons pu évaluer l'erreur moyenne de recalage de notre système à 2 mm et montrer que nous pouvions compter sur un gain en précision atteignant 9%.

Nous insistons sur la versatilité de notre critère puisque les conditions dans lesquelles il apporte un gain majeur ne sont pas inhabituelles. En effet, les applications utilisant un système stéréoscopique avec des caméras très proches sont assez nombreuses (robot mobile, système de repérage 3D de type PolarisTM, appareillage portable (Head Mounted Display) vidéo...). Etant

donné que les corps rigides repérés dans les images vidéo (dans notre cas le marqueur carré de l'ARToolkit) sont fabriqués avec une précision atteignant le centième de millimètre, on pourrait penser que le RSB_{3D} est trop faible par rapport au RSB_{2D} pour que l'EPPC apporte un réel avantage. Cet argument, valide lorsque la détection de points d'intérêt dans une image vidéo était de l'ordre du pixel, n'est plus d'actualité puisque nous obtenons déjà une précision au dixième de pixel avec des images 768×576 . De plus, la dimension des images utilisées dans les systèmes de vision par ordinateur ne cesse d'augmenter : le RSB_{2D} ne peut donc pas être considéré comme négligeable par rapport au RSB_{3D} .

Validation de la prédiction sur l'erreur de recalage

Nous avons propagé les incertitudes au travers du critère standard ISPPC (défini dans la section 3.2.3) et de notre nouveau critère EPPC (cf. section 4), et avons obtenu pour la première fois une formulation *analytique* de la covariance sur la transformation. Nous avons ensuite réalisé un nombre considérable d'expériences afin de valider toutes les hypothèses sur lesquelles nous nous sommes appuyés pour réaliser nos calculs théoriques. Nous avons tout d'abord montré sur des données synthétiques que les non-linéarités des critères étaient négligeables. Ensuite, nous avons mis en évidence que les erreurs de calibrage étaient négligeables tant que les écarts-types des bruits d'extraction 2D et 3D étaient supérieurs à 0.7 pixel et 0.7 mm. Cette étape de validation a d'ailleurs permis d'évaluer à 1 mm environ les erreurs qu'induit le calibrage imparfait sur les transformations que nous calculons avec notre système et dans les conditions de notre expérience.

Nous avons finalement vérifié sur des données réelles provenant de notre mannequin de l'abdomen que notre prédiction était valide pour le nombre de marqueurs employés (entre 15 et 25) et surévalué pour un nombre plus faible.

En dehors de l'intérêt que ce travail représente pour la fiabilité de notre système, les résultats analytiques obtenus pourront être très utiles pour l'évaluation des erreurs commises par les systèmes de repérage de type OptotrakTM, PolarisTM, FlashPointTM. En effet, la méthodologie actuelle pour estimer l'erreur moyenne de repérage de la pointe d'un instrument, est de réaliser de multiples tests de précision à l'aide de mesures étalon. Avec notre prédiction, il nous suffit de connaître la position relative des caméras et la matrice de covariance du bruit d'extraction des points d'intérêts dans les images, pour prédire l'erreur de repérage de la pointe de l'instrument. D'une part, nous pourrions fournir *en temps réel* les erreurs de repérage, d'autre part, cela éviterait de nombreux tests fastidieux, long et coûteux.

9.2 Conclusion pratique

Un système de repérage précis au millimètre à moindre coût

Pour permettre au praticien de suivre l'évolution de l'aiguille dans le corps du patient, nous avons réalisé un système de repérage stéréoscopique, à l'aide de bibliothèques LGPL²⁹ et de quelques traitements d'images supplémentaires. Une expérience d'évaluation a montré que nous obtenions une précision de 1 mm dans notre volume d'intérêt, ce qui est tout à fait honorable par rapport aux systèmes de repérage commercial [WTF04]. De plus, le matériel requis est minimal : un PC, une imprimante et deux caméras sont nécessaires. Ce travail aura permis de prouver qu'un système de repérage précis est accessible à moindre frais.

29. Lesser General Public License

Extraction et appariement automatique des marqueurs

Nous avons développé tous les outils nécessaires à l'extraction et l'appariement automatique des données nécessaires à notre système. Il est donc opérateur-indépendant et permet d'éviter les tâches interactives classiques présentes en routine clinique. De plus, nous avons validé tous nos algorithmes sur de nombreuses images réelles, et avons même validé la précision de notre extraction automatique dans les images vidéo en la comparant à des relevés manuels.

Comme les marqueurs que nous avons utilisés sont couramment employés, ce travail représente une contribution médicale extrêmement pratique, puisque notre mécanisme d'extraction automatique peut être réutilisé dans toutes les applications cliniques dont les conditions sont proches des nôtres.

9.3 Conclusion médicale

Pour répondre aux exigences de précision de notre système thérapeutique, nous avons mené une étude de l'influence de la respiration sur la position du foie, dans les conditions spécifiques à la radiologie interventionnelle. Notre étude bibliographique, appuyée par notre propre expérimentation menée sur deux patients, indiquent que l'erreur de repositionnement du foie est de l'ordre du millimètre. Ces résultats auront permis de justifier l'objectif de précision que nous nous étions proposé d'atteindre avant la conception du système.

D'un point de vue plus général, nous avons conçu un système thérapeutique pour le foie en radiologie interventionnelle et nous avons testé le bon fonctionnement de tous les modules qui le compose une fois qu'ils étaient assemblés. Pour démontrer l'efficacité générale de notre système, nous avons réalisé une expérience dans laquelle un chirurgien cherchait à atteindre, avec une aiguille, des cibles situées à l'intérieur du mannequin, en se guidant grâce à notre système. Les résultats obtenus ont été très satisfaisants puisque, en moyenne, l'erreur RMS de positionnement commise par les chirurgiens était de 2.3 mm, et le temps moyen nécessaire au déroulement de l'acte était en dessous d'une minute. Nous avons ainsi démontré que notre système atteignait les objectifs initialement fixés (précision et durée du positionnement inférieure à 5 mm et 10 minutes) sur un mannequin de l'abdomen. Nous avons complété cette validation par trois expériences en conditions réelles sur des patients. Nous avons préalablement montré comment comparer la précision de notre système à celle du praticien à l'aide d'un protocole *sans danger pour le patient*. Les résultats de précision obtenus, bien qu'encore insuffisants (entre 6 et 10 mm d'erreur) sont très encourageants du fait que les conditions pendant ces expériences n'étaient pas assez favorables pour obtenir une précision optimale: le patient respirait librement et le nombre de marqueurs radio-opaques utilisés pour le recalage était deux fois inférieur à celui des expériences sur le mannequin. De plus, ces premiers résultats sont novateurs, puisque, à notre connaissance, nous avons réalisé la première évaluation sur un patient d'un système de guidage par réalité augmentée sur la zone abdominale.

9.4 Perspectives théoriques

Nous avons développé un nouveau critère de recalage 3D/2D qui tenait compte de la présence de bruit sur les données 3D et 2D. Cependant, nous avons supposé implicitement que les erreurs de calibrage étaient négligeables. Cette hypothèse a entraîné, pendant notre travail de validation sur la prédiction d'incertitude, le constat suivant: lorsque l'amplitude du bruit corrompant les données est en dessous d'un certain seuil, notre prédiction n'est plus fiable.

En l'occurrence, l'erreur prédite est plus faible que l'erreur observée à cause des erreurs de calibrage.

Pour réussir à prédire correctement l'erreur de recalage dans les cas où le bruit est de très faible amplitude, il faudrait soit réaliser un calibrage plus précis, ce qui nécessiterait des moyens matériels très lourds, soit modéliser les caractéristiques des erreurs affectant les paramètres des caméras lors de leur calibrage.

Afin de s'affranchir d'instrument de calibrage industriel, il serait intéressant de mener un travail de modélisation et de validation sur la possibilité d'intégrer, sous forme statistique, la présence des erreurs de calibrage dans le critère EPPC. L'intérêt de ce travail est double : d'une part, il pourrait mener à un critère englobant et encore plus précis que l'EPPC ; d'autre part, il permettrait de prédire les erreurs de recalage, même quand les bruits corrompant les données seraient très faibles.

Un autre aspect théorique sur l'EPPC est à considérer. Dans la version actuelle du système, les caractéristiques des bruits corrompant les données 2D et 3D ont été évaluées préalablement, puisque nous supposons que la dimension des images utilisées ne variaient pas pendant nos expériences. Les systèmes technologiques étant régulièrement améliorés, la résolution des images scanner et vidéo va augmenter et les caractéristiques des bruits d'extraction ne seront plus les mêmes. Il serait donc intéressant de modifier notre critère afin de pouvoir estimer, en plus de la transformation, les caractéristiques des bruits 2D et 3D. Ce travail est en cours et nécessitera un effort de validation considérable.

9.5 Perspectives médicales

Nous profiterons des prochaines expériences sur des patients afin de compléter notre étude sur les déplacements du foie et nous essaierons aussi de l'étendre à d'autres organes (reins, prostate, vessie, poumons...). Si nous réussissons à montrer que l'erreur de repositionnement est également de l'ordre de 1 mm pour ces organes, alors nous pourrions envisager de nouvelles applications de notre système sur ces organes

Par ailleurs, nous espérons pouvoir évaluer l'erreur de positionnement que commet le praticien lorsqu'il essaie de placer la pointe d'une aiguille en un point préalablement défini de l'image TDM. Pour l'instant nous sommes limités à leur avis subjectifs. Il serait intéressant de savoir quantitativement à partir de quelle précision un système de guidage apporte réellement un bénéfice au praticien.

La précision de notre système est basée sur l'hypothèse classique que l'aiguille ne se tord pas lorsqu'elle est introduite dans le patient. Bien que cette supposition soit raisonnable, il semble néanmoins important d'en vérifier la validité. Si nous parvenons à obtenir des images TDM dans lesquelles apparaît l'aiguille (cas de la radio-fréquence), nous pourrions essayer de la segmenter afin d'évaluer sa flexion lorsqu'elle est introduite dans le patient.

Dans le cadre des opérations en radiologie interventionnelle, nous pourrions essayer de fournir un modèle déformé du foie en fonction du cycle respiratoire du patient. Pour réaliser cela, une solution serait de réaliser plusieurs acquisitions TDM (ou IRM...) du patient à des instants différents de son cycle respiratoire et d'effectuer une analyse en composantes principales (ACP) des mouvements du foie en fonction de ceux des marqueurs radio-opaques. Il serait alors possible d'interpoler sa position et sa forme quel que soit le volume de gaz dans les poumons.

Finalement, nous envisageons d'appliquer le travail réalisé dans cette thèse en laparoscopie pour les opérations d'organe à mobilité réduite (par exemple la glande surrénalienne). Cette fois-ci nous chercherons à superposer dans l'image vidéo fournie par l'endoscope, les reconstructions

3D pour guider le praticien. Nous devons évaluer alors les erreurs dues au calibrage de la caméra endoscopique et à son repérage dans l'espace. Par ailleurs, nous pourrions envisager dans ce contexte un recalage déformable du modèle pré-opératoire du foie sur une reconstruction surfacique du foie obtenue par projection de lumière structurée sur sa surface.

9.6 Perspectives pratiques

La librairie dont nous nous sommes servie pour repérer l'aiguille dans l'espace est peu précise et peu évolutive. La précision obtenue est au dessus d'environ 1 mm et les caractéristiques du carré servant au repérage ne peuvent être modifiées. Nul doute qu'avec une forme géométrique plus adaptée (par exemple un damier avec une douzaine de points d'intérêt), la précision obtenue serait meilleure et plus fiable. Il serait donc très utile, pour la communauté des chercheurs en vision par ordinateur, de créer une librairie améliorée de repérage, qui tiendrait compte de la distorsion, du nombre de caméra et qui laisserait la possibilité à l'utilisateur de concevoir lui-même son marqueur.

L'évaluation de la précision du repérage de l'aiguille ayant été réalisée dans des conditions restreintes, correspondant à celles de notre application (le volume dans lequel la précision a été mesurée est d'environ 25 dm³), nous souhaiterions mettre en place un protocole d'évaluation plus complet avec du matériel plus précis (banc de calibrage) que celui utilisé jusqu'à présent. En effet, il serait souhaitable de fournir aux utilisateurs une idée de la précision envisageable dans des conditions moins favorables que les nôtres (par exemple lorsque le marqueur est très éloigné des caméras). Nous pourrions, de surcroît, en profiter pour comparer la précision de notre système de repérage avec ceux existants.

9.7 Conclusion générale

D'un point de vue global, nous avons appliqué à un domaine médical des outils de vision par ordinateur afin de créer un système dédié à des opérations sur le foie. Ce travail s'inscrit donc dans le cadre de la recherche en informatique appliquée à la médecine. Les résultats obtenus sont d'autant plus précieux que le système final a été confronté aux avis d'experts médicaux et qu'il a été introduit en salle d'opération. Nous avons ainsi démontré la réalité de son emploi et de son bénéfice dans un futur proche. Bien que nous ayons focalisé notre travail sur le foie, nous pensons que ce système pourra être utilisé sur d'autres organes abdominaux. Nous espérons ainsi avoir apporté notre pierre à l'édifice que représente la chirurgie de l'abdomen guidé par l'image et l'informatique.

Annexe A

Evaluation des performances de l'EPPC

Nous proposons dans cette annexe une étude comparative de trois critères de recalage ponctuel 3D/2D décrit dans le **chapitre 4** (ISPPC, OSPPC et EPPC). Le but est d'évaluer les performances de chacun des critères en temps de calcul, en précision et en robustesse.

Les expériences sont réalisées avec deux caméras synthétiques observant la même zone avec une différence d'angle de visée de 45° , et calibrées dans le même repère de référence R_f . Nous utilisons 15 points N_i dont les coordonnées 3D sont connues dans R_f , et dont les projections m_i dans les images vidéos sont précisément pré-calculées (en unité pixel). Le détail des conditions expérimentale est donnée dans la section 4.2.1. Nous évaluons l'erreur de recalage à l'aide de N points de contrôle C_i définis dans le repère de référence R_f .

A.1 Temps de calcul et précision

Chaque tableau présenté dans ce paragraphe indique les valeurs comparatives de temps de calcul et de précision des 3 critères par rapport à un paramètre, les autres étant fixés (cf. section 4.2.1 pour la définition des paramètres). Pour chaque valeur du paramètre qui varie, nous indiquons la moyenne du temps de calcul et l'erreur RMS sur 10000 recalages. Pour caractériser la variabilité de la précision, nous ajoutons l'écart-type des mesures RMS ainsi que sa valeur maximale sur l'ensemble des recalages. A titre indicatif nous précisons aussi dans la ligne *Pourcentage de succès*, le pourcentage du nombre de cas dans lesquels l'EPPC offre une meilleure précision que les deux autres critères.

Pour comparer visuellement les critères, nous représentons graphiquement l'erreur relative et le rapport des temps de calcul entre chaque couple de critère. La mesure d'erreur relative n'étant pas symétrique, nous calculons en fait l'exponentielle de la moyenne des logarithmes des rapports de RMS pour obtenir une mesure symétrique de la forme :

$$\text{Erreur relative}(A/B) = [\text{Erreur relative}(B/A)]^{-1} = \exp \left(\frac{1}{10000} \cdot \sum_1^{10000} \log \left(\frac{RMS_{\text{méthode A}}}{RMS_{\text{méthode B}}} \right) \right)$$

Ainsi, une valeur au-dessus de 1 indique que la méthode B est relativement plus précise que la méthode A. Pour alléger la lecture des tableaux, nous utiliserons les abréviations IS, OS, E pour ISPPC, OSPPC et EPPC.

A.1.1 Influence de l'amplitude du bruit

Le tableau A.1 montre que l'amplitude du bruit a une influence relativement linéaire sur la précision. En effet, l'erreur RMS des 3 critères est grossièrement égale (dans ce cas) à l'écart-type des bruit 2D et 3D, même si l'EPPC reste plus précis que les autres (gain de 10%).

Amplitude du bruit (2D/3D)		0.2 / 0.2	1 / 1	2 / 2	3 / 3	4 / 4	10 / 10
Temps de calcul moyen	IS	0.0050s	0.0040s	0.0040s	0.0040s	0.0040s	0.0040s
	OS	0.020s	0.018s	0.017s	0.017s	0.017s	0.016s
	E	0.031s	0.035s	0.050s	0.060s	0.067s	0.091s
Erreur RMS moyenne en mm ± écart-type	IS	0.22±0.08	1.08±0.42	2.18±0.84	3.25±1.26	4.32±1.67	10.85±4.16
	OS	0.22±0.08	1.08±0.41	2.17±0.84	3.25±1.26	4.32±1.66	10.82±4.16
	E	0.20±0.07	0.98±0.37	1.96±0.75	2.93±1.12	3.90±1.48	9.73±3.69
Erreur RMS maximale en mm	IS	0.59	2.98	7.99	9.34	11.64	32.29
	OS	0.60	3.00	7.87	9.26	11.94	32.44
	E	0.55	2.90	7.16	8.16	10.46	26.19
Pourcentage de succès	E vs. IS	67.8	67.3	66.2	65.7	64.7	65.2
	E vs. OS	67.6	67.2	66.5	65.7	64.0	65.0

TAB. A.1 – Temps de calcul et précision en fonction du bruit corrompant les données. La précision se dégrade lorsque le bruit sur les données augmente. Par ailleurs, il est intéressant de constater que les précisions de l'ISPPC et de l'OSPPC sont quasiment équivalentes quelle que soit l'amplitude des bruits.

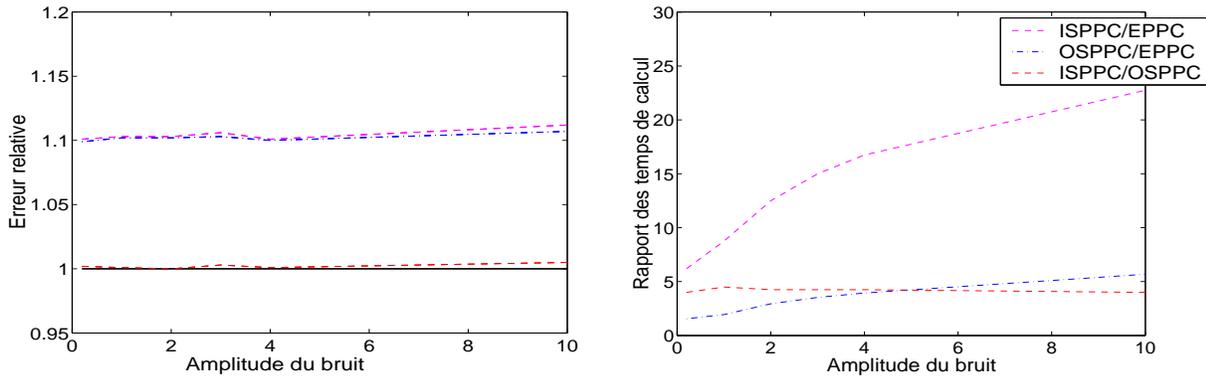


FIG. A.1 – Erreur relative et rapport des temps de calcul entre les 3 critères en fonction de l'amplitude du bruit sur les données. Les erreurs relatives ISPPC/EPPC et OSPPC/EPPC restent constantes (à 1.1) si le rapport $\frac{\sigma_{3D}}{\sigma_{2D}}$ est stable quelle que soit l'amplitude de σ_{2D} et σ_{3D} .

Si l'on considère maintenant une augmentation du bruit 3D par rapport au bruit 2D, la figure A.2 nous montre que l'erreur relative augmente aussi (erreur relative de 1.10 pour un rapport $\sigma_{2D}/\sigma_{3D} \simeq 1$ contre 1.18 pour un rapport de 2). Cela signifie que l'EPPC est relativement de plus en plus précis, ce qui était attendu puisque ce critère capture mieux la nature de l'information bruitée sur les données 3D. Ce gain en précision est cependant compensé par un temps de calcul bien plus élevé. (0.05 sec. contre 0.23 sec. par exemple). Nous remarquons aussi avec surprise qu'il y a une limite à ce comportement : quand le bruit 3D devient trop important (pour un rapport σ_{3D}/σ_{2D} égal à 8), l'erreur relative diminue au lieu d'augmenter.

Rapport des écart-type: $\frac{\sigma_{3D}}{\sigma_{2D}}$		8	4	2	1	0.5	0.25
Temps de calcul moyen	IS	0.004s	0.004 s	0.004 s	0.004	0.004 s	0.004 s
	OS	0.017 s	0.017 s	0.017 s	0.017 s	0.018 s	0.018s
	E	0.23 s	0.18 s	0.11 s	0.05 s	0.03 s	0.015s
Pourcentage de succès	E vs. IS	83.6	77.5	72.5	66.2	58.5	53.5
	E vs. OS	72.2	75.7	72.2	66.5	58.8	54.5

TAB. A.2 – Temps de calcul et pourcentage de succès de l'EPPC en fonction du rapport des valeurs des écart-types des bruits utilisés dans l'algorithme EPPC. Comme l'EPPC tend à se comporter comme l'ISPPC en l'absence de bruit 3D, les temps de calcul de l'EPPC diminuent considérablement lorsque le rapport $\frac{\sigma_{3D}}{\sigma_{2D}}$ baisse. En revanche, lorsque ce rapport augmente, nous constatons logiquement que le pourcentage de succès de l'EPPC augmente puisque l'EPPC capture mieux l'information 3D que les deux autres critères.

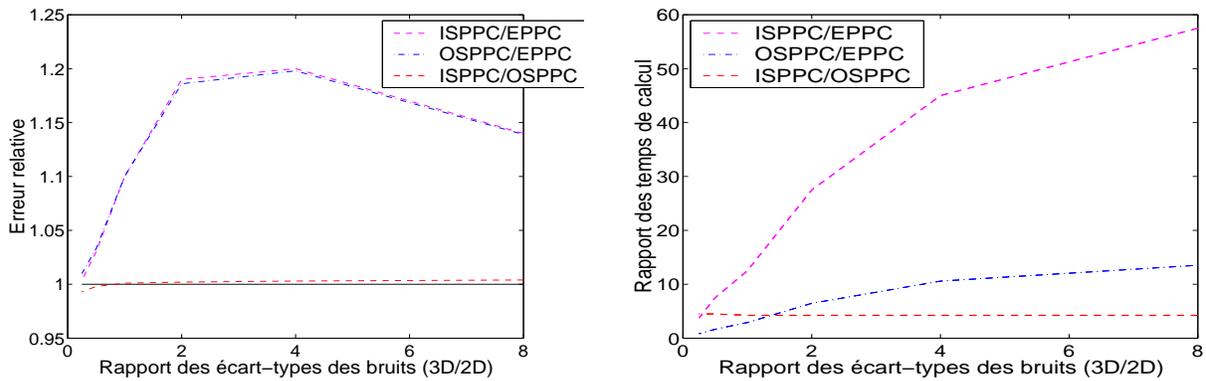


FIG. A.2 – Erreur relative et rapport des temps de calcul entre les 3 critères en fonction du rapport des bruits sur les données. L'erreur relative augmente considérablement en fonction du rapport $\frac{\sigma_{3D}}{\sigma_{2D}}$ (1.22 pour un rapport 4). L'EPPC nécessite alors des temps de calcul conséquents par rapport aux autres critères (10 fois plus que l'OSPPC et jusqu'à 50 fois plus que l'ISPPC).

A.1.2 Influence de la distance objet/caméras

Les résultats de la figure A.3 indiquent clairement que l'EPPC devient relativement plus précis que les deux autres critères au fur et à mesure que les caméras se rapprochent de l'objet. Ce phénomène provient du fait que rapprocher les caméras correspond à diminuer l'impact du bruit sur les données 2D, et donc à augmenter artificiellement le rapport σ_{3D}/σ_{2D} . En effet, les caméras se rapprochant, la taille de l'objet dans l'image va augmenter pour une même valeur de bruit infectant les données 2D et le rapport signal sur bruit (RSB) va donc baisser. Nous constatons de même que lorsque σ_{3D}/σ_{2D} augmente, que le temps de calcul de l'EPPC augmente lorsque la distance caméra/objet diminue.

Distance objet/caméras (mm)		300	500	700	1000	2000	3000
Temps de calcul moyen	IS	0.0040s	0.0040s	0.0040s	0.0040s	0.0040s	0.0040s
	OS	0.018s	0.018s	0.018s	0.018s	0.017s	0.017s
	E	0.247s	0.105s	0.071s	0.050s	0.027s	0.021s
Erreur RMS moyenne en mm ± écart-type	IS	2.17±0.82	2.08±0.79	2.08±0.80	2.17±0.84	2.74±1.05	3.50±1.33
	OS	2.05±0.78	2.06±0.78	2.07±0.80	2.17±0.83	2.74±1.05	3.50±1.33
	E	1.66±0.62	1.70±0.64	1.78±0.67	1.96±0.74	2.66±1.02	3.47±1.32
Erreur RMS maximale en mm	IS	6.34	5.58	5.93	6.27	8.26	9.73
	OS	5.74	5.33	5.98	6.16	8.21	9.82
	E	4.61	4.78	4.94	5.09	7.95	9.84
Pourcentage de succès	E vs. IS	82.4	76.0	70.5	65.6	58.8	54.2
	E vs. OS	74.4	74.6	70.7	65.6	58.7	54.4

TAB. A.3 – Temps de calcul et précision en fonction de la distance objet/caméras. Lorsque la distance diminue le RSB des données 2D augmente, la précision s'améliore donc pour les trois critères.

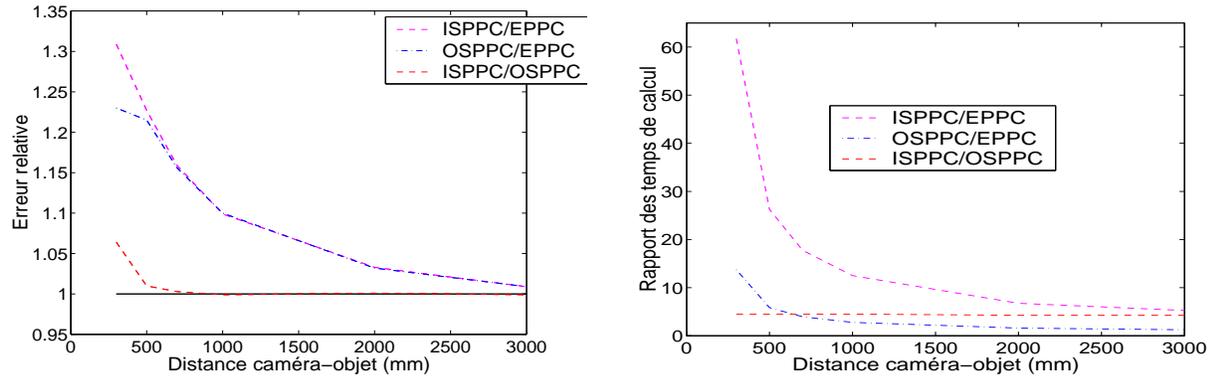


FIG. A.3 – Erreur relative et rapport des temps de calcul entre les 3 critères en fonction de la distance objet/caméras. Le rapport σ_{3D}/σ_{2D} augmentant lorsque la distance diminue, nous observons une amélioration de l'erreur relative de l'EPPC sur les autres critères (dépassant 1.20 pour une distance inférieure à 500 mm). Les temps de calcul relatifs de l'EPPC deviennent alors plus importants.

A.1.3 Influence de l'angle entre les caméras

Le tableau A.4 présente des résultats en fonction de l'angle entre les caméras qui varie de 10 à 90 degrés. Nous remarquons tout d'abord que l'erreur RMS diminue pour un angle se rapprochant de 90°. Cela était prévisible puisque la position des points est d'autant plus contrainte que l'angle entre les caméras augmente.

Ensuite, l'EPPC est relativement d'autant plus précis, que l'angle diminue, jusqu'à obtenir 22% de gain en précision pour un angle de 10° (cf. figure A.4). Cette évolution est cohérente étant donné que la quantité d'information 3D dépend de l'angle entre les caméras. Lorsque cette information est maximale, l'EPPC n'apporte donc que peu d'avantage (erreur relative = 1.03).

Angle entre les caméras		90°	60°	45°	30°	10°
Temps de calcul moyen	IS	0.0040s	0.0040s	0.0040s	0.0040s	0.0040s
	OS	0.017s	0.018s	0.017s	0.017s	0.016s
	E	0.032s	0.044s	0.050s	0.056s	0.087s
Erreur RMS moyenne en mm ± écart-type	IS	1.78±0.67	1.97±0.75	2.18±0.84	2.51±0.95	4.09±1.80
	OS	1.78±0.67	1.96±0.75	2.17±0.84	2.52±0.96	4.10±1.81
	E	1.75±0.66	1.84±0.70	1.96±0.75	2.19±0.82	3.25±1.30
Erreur RMS maximale en mm	IS	5.10	5.44	7.99	6.49	14.24
	OS	5.12	5.35	7.87	6.37	12.96
	E	5.09	5.16	7.16	5.88	11.52
Pourcentage de succès	E vs. IS	58.1	62.9	66.2	68.4	69.4
	E vs. OS	57.0	62.1	66.5	68.2	68.8

TAB. A.4 – Temps de calcul et précision en fonction de l'angle entre les caméras. L'information de profondeur diminue lorsque l'angle entre les caméras s'affaiblit, nous observons alors une dégradation de la précision.

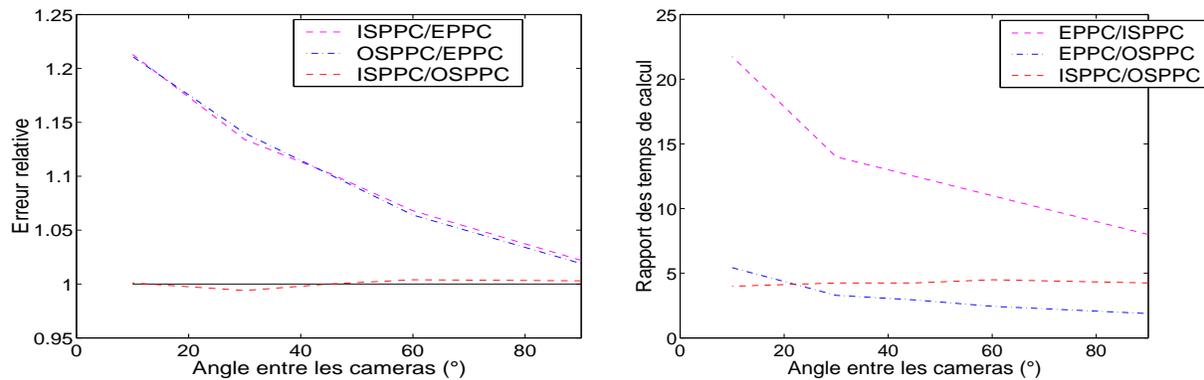


FIG. A.4 – Erreur relative et rapport des temps de calcul entre les 3 critères en fonction de l'angle entre les caméras. La quantité d'information 3D diminuant avec l'angle entre les caméras, les erreurs relatives de l'ISPPC et de l'OSPPC par rapport à l'EPPC augmentent jusqu'à 1.22 pour un angle de 10°.

A.1.4 Influence du nombre de points

Nous clôturons cette étude avec une évaluation de la précision du recalage en fonction du nombre de points utilisés (cf. tableau A.5 et figure A.5). Nous pouvons premièrement remarquer que les temps de calcul sont presque proportionnels au nombre de points. Cela était prévisible puisque la complexité numérique est linéaire par rapport au nombre de données. La deuxième observation est que l'erreur RMS semble inversement proportionnel à la racine carré de la variation du nombre de point (multiplier par 2 le nombre de points diminue l'erreur RMS d'un facteur $\sqrt{2}$), ce qui est un comportement standard en statistique étant donné que les points sont bien distribués uniformément dans l'espace.

Nombre de points		30	15	8	4
Temps de calcul moyen	IS	0.0070s	0.0040s	0.0030s	0.0010s
	OS	0.031s	0.018s	0.012s	0.008s
	E	0.137s	0.050s	0.022s	0.009s
Erreur RMS moyenne en mm \pm écart-type	IS	1.56 \pm 0.60	2.17 \pm 0.84	2.78 \pm 1.05	5.27 \pm 2.22
	OS	1.56 \pm 0.60	2.17 \pm 0.84	2.77 \pm 1.04	5.26 \pm 2.22
	E	1.41 \pm 0.54	1.95 \pm 0.74	2.52 \pm 0.94	4.98 \pm 2.13
Erreur RMS maximale en mm	IS	4.42	6.40	7.79	15.53
	OS	4.37	6.21	8.16	15.26
	E	3.99	5.32	7.21	14.75
Pourcentage de succès	E vs. IS	66.1	67.0	64.6	61.9
	E vs. OS	65.7	66.9	65.5	61.6

TAB. A.5 – Temps de calcul et précision en fonction du nombre de points. Comme attendu, la précision s'améliore et les temps de calcul augmentent lorsque le nombre de points augmente.

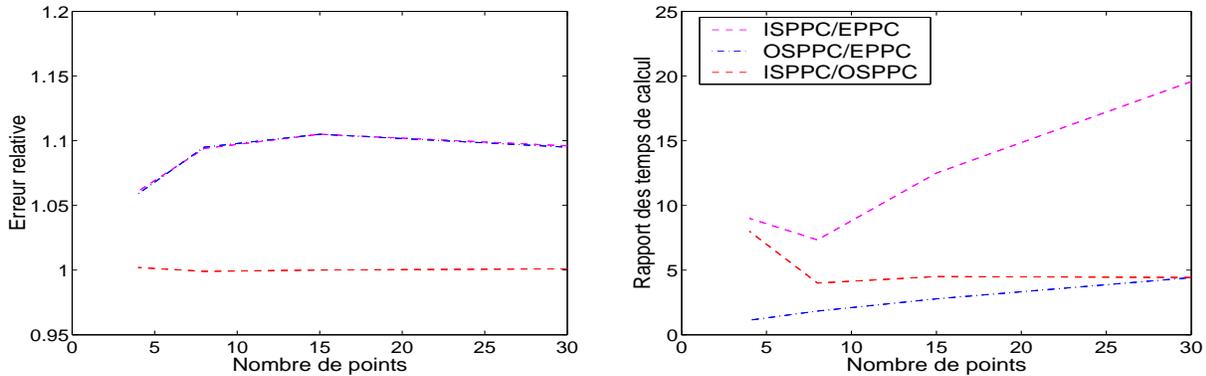


FIG. A.5 – Temps de calcul et précision en fonction du nombre de points. L'erreur relative reste stable à partir de 7 points. L'EPPC est relativement plus lent que les autres critères lorsque le nombre de points augmente.

A.2 Etude de la robustesse

Dans cette section, nous cherchons à évaluer la différence des bassins de convergence des trois critères en fonction des données initiales. Nous nous intéressons particulièrement à l'influence de la transformation recherchée et de son initialisation. Les conditions expérimentales et les mesures de performances sont globalement les mêmes que dans la section précédente, mais cette-fois-ci, nous n'initialisons plus l'algorithme avec des paramètres de transformation proches de la solution. De plus nous indiquons le pourcentage du nombre de cas où chaque critère ne converge pas (cf. 4.3.3 pour plus de détails).

A.2.1 T aléatoire avec une initialisation à l'identité

Le tableau A.6 montre l'incidence de l'amplitude du bruit sur la robustesse des critères. Les temps de calcul ne semblent pas affectés, alors que le rapport de mauvaises convergences descend de 9.1% à 7.1% pour l'ISPPC et de 4.6% à 2.9% pour l'OSPPC. En revanche, les valeurs restent stables pour l'EPPC (1%).

Amplitude du bruit (2D/3D)		0.2 / 0.2	1 / 1	2 / 2	3 / 3	4 / 4	10 / 10
Temps de calcul moyen	IS	0.0150s	0.0140s	0.0140s	0.0140s	0.0140s	0.0140s
	OS	0.064s	0.062s	0.061s	0.061s	0.061s	0.060s
	E	0.357s	0.425s	0.385s	0.386s	0.402s	0.368s
Pourcentage de mauvaises convergences	IS	9.1%	9.4%	8.6%	8.6%	8.6%	7.1%
	OS	4.6%	4.5%	3.8%	4.1%	4.2%	2.9%
	E	0.9%	1.0%	1.0%	1.0%	1.0%	1.0%

TAB. A.6 – Evaluation de la robustesse en fonction de l'amplitude du bruit. La robustesse des critères ne dépend pas des variations d'amplitude de bruit.

Le tableau A.7 examine l'influence du nombre de points utilisés pour le recalage. Le rapport de convergences incorrects reste faible pour l'EPPC (1%) et à peu près constant pour l'ISPPC et l'OSPPC (respectivement 8% et 4%). En revanche, pour 4 points, l'ISPPC et l'OSPPC sont très peu performants car dans les deux cas plus de 35% des recalages ont mal convergés. Ce phénomène vient probablement du fait que la configuration à 4 points admet deux solutions stables. Dans le cas de l'ISPPC et de l'OSPPC, si l'initialisation est plus proche de la mauvaise solution, l'algorithme convergera dans sa direction. En revanche, l'EPPC sera beaucoup moins sensible à cette singularité car la seconde solution n'optimise pas simultanément les deux termes de ce critère.

Nombre de points		30	15	8	4
Temps de calcul moyen	IS	0.0260s	0.0140s	0.0090s	0.0060s
	OS	0.104s	0.059s	0.037s	0.028s
	E	0.881s	0.423s	0.135s	0.089s
Pourcentage de mauvaises convergences	IS	10.7%	9.5%	7.6%	40.6%
	OS	4.5%	4.0%	3.5%	35.2%
	E	0.6%	1.1%	1.0%	0.6%

TAB. A.7 – Evaluation de la robustesse en fonction du nombre de points. Les configurations à 4 points admettant 2 solutions stables, les robustesses de l’ISPPC et de l’OSPPC se dégradent très violemment. L’EPPC n’est pas sensible à ce phénomène étant donné que la mauvaise solution n’optimise pas simultanément les 2 termes du critères.

Angle entre les caméras		90°	60°	45°	30°	10°
Temps de calcul moyen	IS	0.0140s	0.0140s	0.0140s	0.0140s	0.0140s
	OS	0.065s	0.059s	0.059s	0.062s	0.065s
	E	0.146s	0.410s	0.347s	0.621s	0.585s
Pourcentage de mauvaises convergences	IS	7.6%	8.7%	8.3%	7.8%	6.9%
	OS	4.3%	4.0%	4.0%	4.7%	10.2%
	E	0.02%	0.8%	0.8%	1.8%	2.2%

TAB. A.8 – Evaluation de la robustesse en fonction de l’angle entre les caméras. L’EPPC est plus robuste lorsque l’angle est important et devient quasiment nulle pour 90°.

Finalement, le tableau A.8 présente les mêmes conclusions pour l’influence de l’angle entre les caméras. En effet, le pourcentage de mauvaises convergences pour l’EPPC est toujours faible (proche de 1%) et 8 fois plus élevé (resp. 4 fois plus élevé) pour l’ISPPC (resp. pour l’OSPPC).

En résumé, l’ISPPC (resp. l’OSPPC) converge vers une mauvaise solution dans 9% (resp. 4.5%) des cas, alors que l’EPPC ne converge mal que dans 1% des cas, ce qui en fait un critère plus robuste pour notre application avec un bassin de convergence beaucoup plus large. Cependant, nous devons rester prudent puisque les résultats de robustesse dépendent aussi de la méthode d’optimisation. Une autre méthode d’optimisation pourrait donner des résultats légèrement différents.

A.2.2 Initialisation aléatoire avec T fixée à l’identité

Nous présentons brièvement dans ce paragraphe une version symétrique de l’expérience précédente: la transformation recherchée T est fixée à l’identité, tandis que la transformation initiale est déterminée aléatoirement. En principe, nous devrions obtenir à peu près les mêmes pourcentages de convergences incorrects et de temps de calcul.

En fait, le tableau A.9 décrit un comportement peu attendu: les temps de calcul sont plus faibles en général et les rapport de convergences incorrects ont baissé (il y a deux fois moins d’erreur de convergence pour l’ISPPC et exactement 0 pour l’OSPPC). L’EPPC ne semble pas être affecté par cette différence de condition initiale. Ce phénomène est probablement dû à la méthode d’optimisation, ou à la représentation particulière de l’espace des rotations. Il serait intéressant de réaliser quelques expériences avec une méthode d’optimisation différente pour clarifier ce comportement.

Temps de calcul moyen	IS	0.010s
	OS	0.007s
	E	0.230s
Pourcentage de mauvaises convergences	IS	4.7%
	OS	0.0%
	E	1.5%

TAB. A.9 – *Evaluation de la robustesse avec T fixée à l'identité et une initialisation aléatoire. Les résultats ne sont pas identiques à ceux obtenus pour les conditions symétriques (transformation choisie aléatoirement et initialisation fixée à l'identité). Ce phénomène est vraisemblablement dû à la méthode d'optimisation.*

A.2.3 Initialisation aléatoire avec T aléatoire

Le tableau A.10 présente les résultats quand T et l'initialisation sont choisis aléatoirement. Si nous comparons ces résultats avec ceux obtenus quand l'initialisation vaut l'identité, nous remarquons que les temps de calculs changent peu pour l'ISPPC et l'OSPPC, mais que le pourcentage de convergences incorrectes augmente considérablement de 9% à 42.9% pour l'ISPPC et de 4% à 15% pour l'OSPPC. Bien que l'EPPC soit aussi très sensible à ce changement de conditions initiales, le pourcentage d'erreur de convergence reste faible (proche de 5%) malgré des temps de calcul élevé (plus de 2 sec.). Cela signifie que même si l'initialisation est complètement fautive, la convergence sera assurée dans 95% des cas.

Temps de calcul moyen	IS	0.012s
	OS	0.041s
	E	2.411s
Pourcentage de mauvaises convergences	IS	42.9%
	OS	15.4%
	E	5.3%

TAB. A.10 – *Evaluation de la robustesse avec T et l'initialisation choisies aléatoirement. L'EPPC est bien plus robuste que l'ISPPC et l'OSPPC, mais au prix d'un temps de calcul 50 à 200 fois supérieur.*

A.3 Conclusion

Les résultats comparatifs présentés dans cette annexe démontrent la supériorité de l'EPPC sur l'ISPPC et l'OSPPC. Sur le plan de la précision, l'EPPC fournit toujours une meilleure erreur RMS que les deux autres critères, mais au coût d'un temps de calcul 10 à 30 fois supérieur. De plus, l'étude montre que l'EPPC est plus robuste que les deux critères dans quasiment tous les cas. Le pourcentage de convergence incorrecte vaut 1% contre 4% pour l'OSPPC et 9% pour l'ISPPC lorsque T est choisie aléatoirement et avec une initialisation fixée à l'identité. En revanche, l'EPPC n'apporte pas d'avantage en robustesse lorsqu'une bonne initialisation est disponible.

Annexe B

Méthodes d'estimation de la pointe d'une aiguille

Nous présentons dans cette annexe deux méthodes, différentes de celle proposée dans la section 5.3.2, permettant d'estimer la position de la pointe d'une aiguille par rapport à un corps rigide connu sur lequel l'aiguille est attachée. Nous supposons que nous disposons, dans un même repère de référence R , de N acquisitions de l'aiguille réalisées dans des positions différentes, la pointe restant immobile. Nous supposons aussi que nous sommes capables de déterminer dans R la position du corps rigide pour chacune des acquisitions réalisées. Dans la suite, nous appellerons M un repère lié au corps rigide.

B.1 Estimation simultanée du point fixe dans les repères de référence et du corps rigide

Cette méthode nécessite, pour chaque acquisition, le calcul de la transformation T_k reliant le repère du corps rigide M au repère de référence R . Dans le cas où 3 points du corps rigide ont pu être calculés dans le repère R (par reconstruction stéréoscopique par exemple), T_k est donnée par la transformation rigide recalant ces 3 points connus dans M sur leurs correspondants calculés dans R . Nous obtenons alors en l'absence de bruit les N relations :

$$\begin{aligned} T^1 \star P_{fixe}^M &= P_{fixe}^R \\ &\vdots \\ T^k \star P_{fixe}^M &= P_{fixe}^R \\ &\vdots \\ T^N \star P_{fixe}^M &= P_{fixe}^R \end{aligned}$$

où P_{fixe}^M (resp. P_{fixe}^R) correspond à la position exacte de la pointe de l'aiguille dans le repère M (resp. dans le repère R). Cet ensemble d'équations à 6 inconnues (3 pour P_{fixe}^M et 3 pour P_{fixe}^R) peut se résoudre par une minimisation aux moindres carré de la fonction d'erreur ϵ :

$$\epsilon = \sum_{k=1}^N \| T^k \star P_{fixe}^M - P_{fixe}^R \|^2$$

B.2 Estimation du point fixe à partir des diamètres

Cette méthode nécessite, pour chaque acquisition, la reconstruction 3D d'un point appartenant au corps rigide lié à l'aiguille. Nous utilisons alors la propriété géométrique suivante : étant donné 3 points D, E, F sur une sphère S , le centre de S appartient à la droite perpendiculaire au triangle DEF et passant par le centre du cercle circonscrit à DEF . Ainsi, chaque triplet de points choisis parmi les N acquisitions permet de calculer une droite Δ_i sur laquelle se trouve la pointe de l'aiguille (une illustration avec 6 points est donnée sur la figure B.1).

En présence de bruit, toutes les droites ne s'intersectent pas et nous devons recourir à une optimisation d'un critère physique. Une possibilité est de chercher le point C qui minimise la distance spatiale aux droites Δ_i :

$$Err = \sum_{i=1}^{i=N} \| d(C, \Delta_i) \|^2$$

où C correspond à une estimation de P_{fixe}^R . Finalement, pour obtenir P_{fixe}^M , nous déterminons préalablement pour chaque acquisition la transformation rigide T^k reliant le repère M au repère R et nous moyennons les $T_k^{-1} \star P_{fixe}^R$.

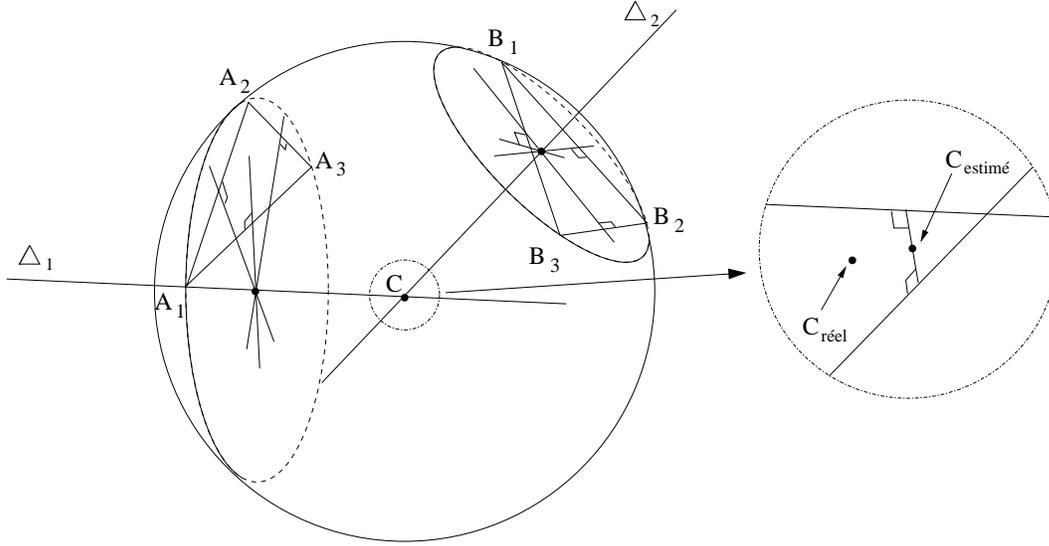


FIG. B.1 – $A_1A_2A_3$ (resp. $B_1B_2B_3$) correspondent à 3 positions différentes d'un des coins du marqueur carré. Nous savons que C se situe sur la droite Δ_1 perpendiculaire au triangle $A_1A_2A_3$ et passant par le centre de son cercle circonscrit. La détermination de Δ_2 (grâce à $B_1B_2B_3$) permet de fournir une estimation de la position de C . Evidemment les coordonnées recueillies de A et B sont bruitées, il faut donc choisir un critère afin d'obtenir une estimation de C en fonction des droites Δ_i obtenues.

À notre connaissance, ces deux méthodes et celle détaillée dans la section 5.3.2 n'ont pas été comparées en terme de précision. Nous avons donc choisi la seule des trois ne nécessitant pas d'optimisation itérative.

Annexe C

Propagation de la covariance sur la transformation estimée à partir de l'ISPPC et de l'EPCC

Nous présentons dans cette annexe le détail des calculs permettant d'obtenir une formule analytique de la covariance sur la transformation estimée à partir des critères ISPPC et EPPC (cf. section 7.1).

C.1 Cas du critère ISPPC

Le critère ISPPC présenté dans la section 4.1.2 correspond à une estimation de la transformation au sens du maximum de vraisemblance avec des points 3D exacts M_i (nous considérons dans cette partie que $M_i = \tilde{M}_i$) et des points 2D bruités $\tilde{m}_i^{(l)}$. Cependant, pour tenir compte de toutes les sources possibles d'incertitude, nous devons voir comment la solution du critère varie s'il y a du bruit sur les données 2D, mais aussi sur les données 3D. Nous devons donc inclure les M_i dans le vecteur de données et supposer une matrice de covariance Σ_{M_i} sur chacune des mesures 3D effectuées. Le critère à minimiser est donné par l'équation 4.2 :

$$C(T) = \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \frac{\|P^{(l)}(T \star M_i) - \tilde{m}_i^{(l)}\|^2}{2\sigma_{2D}^2}$$

Préliminaires Pour réduire la complexité des calculs, nous décomposons la matrice de projection 3×4 comme indiqué ci-dessous :

$$P^{(l)} = \left[\begin{array}{c|c} Q_{2 \times 3}^{(l)} & b_{2 \times 1}^{(l)} \\ \hline C_{(l)}^\top & 1 \end{array} \right]$$

Nous pouvons maintenant écrire analytiquement l'effet de la fonction de projection $P^{(l)}$ sur un point 3D :

$$m_i^{(l)} = P^{(l)}(T \star M_i) = \frac{Q^{(l)} \cdot (T \star M_i) + b^{(l)}}{1 + C_{(l)}^\top \cdot (T \star M_i)},$$

que nous pouvons maintenant dériver par rapport aux paramètres de la transformation :

$$\frac{\partial m_i^{(l)}}{\partial T} = \left(\frac{Q^{(l)}}{1 + \langle C_{(l)}^\top | T \star M_i \rangle} - \frac{(Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top) \cdot C_{(l)}^\top}{(1 + \langle C_{(l)}^\top | T \star M_i \rangle)^2} \right) \cdot \frac{\partial (T \star M_i)}{\partial T} = \frac{Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top}{1 + \langle C_{(l)}^\top | T \star M_i \rangle} \cdot D_i \quad (C.1)$$

où $D_i = \frac{\partial (T \star M_i)}{\partial T}$ correspond à la dérivée de l'action de la transformation T par rapport aux paramètres de la transformation (voir [Pen96] pour son expression en fonction des vecteurs rotation et translation). De même, nous avons besoin de la dérivée de $m_i^{(l)}$ par rapport au point 3D M_i :

$$\frac{\partial m_i^{(l)}}{\partial M_i} = \left(\frac{Q^{(l)}}{1 + \langle C_{(l)}^\top | T \star M_i \rangle} - \frac{(Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top) \cdot C_{(l)}^\top}{(1 + \langle C_{(l)}^\top | T \star M_i \rangle)^2} \right) \cdot \frac{\partial (T \star M_i)}{\partial M_i} = \frac{Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top}{1 + \langle C_{(l)}^\top | T \star M_i \rangle} \cdot R \quad (C.2)$$

où R est la matrice de rotation 3×3 de la transformation T tel que $T \star M = R \cdot M + t$. pour clarifier quelques expressions très lourdes à venir, nous définissons dès maintenant les matrices

suivantes que nous rencontrerons fréquemment durant nos calculs :

$$\begin{aligned}
 G_i^{(l)} &= (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)^\top \cdot (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top) - C_{(l)} \cdot (m_i^{(l)} - \tilde{m}_i^{(l)})^\top \cdot (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top) \\
 &\quad - (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)^\top \cdot (m_i^{(l)} - \tilde{m}_i^{(l)}) \cdot C_{(l)}^\top \\
 L_i^{(l)} &= \frac{(Q - m_i^{(l)} \cdot C_{(l)}^\top)^\top \cdot (Q - m_i^{(l)} \cdot C_{(l)}^\top)}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)^2} \quad \text{and} \quad L_i = \sum_{l=1}^S \xi_i^l \cdot L_i^{(l)} \\
 K_i^{(l)} &= \frac{G_i^{(l)}}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)^2} \quad \text{and} \quad K_i = \sum_{l=1}^S \xi_i^l \cdot K_i^{(l)}
 \end{aligned} \tag{C.3}$$

Calcul du jacobien et du hessien Selon le paragraphe 7.1.1, nous devons tout d'abord calculer la fonction implicite $\Phi = \frac{\partial C}{\partial T}$:

$$\frac{\partial C}{\partial T} = \frac{\partial}{\partial T} \left(\sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \frac{\| P^{(l)}(T \star M_i) - \tilde{m}_i^{(l)} \|^2}{2\sigma_{2D}^2} \right) = \frac{1}{\sigma_{2D}^2} \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot (m_i^{(l)} - \tilde{m}_i^{(l)})^\top \cdot \frac{\partial m_i^{(l)}}{\partial T}$$

Maintenant, à l'aide des dérivées de $m_i^{(l)}$ données en C.1, nous obtenons :

$$\Phi = \left(\frac{\partial C}{\partial T} \right)^\top = \frac{1}{\sigma_{2D}^2} \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot D_i^\top \frac{\left(Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top \right)^\top \cdot (m_i^{(l)} - \tilde{m}_i^{(l)})}{1 + \langle C_{(l)} | T \star M_i \rangle}$$

Dérivant une fois encore par rapport à T , nous obtenons (après simplification) la matrice hessienne :

$$H = \frac{\partial^2 C}{\partial T^2} = \frac{\partial \Phi}{\partial T} = \frac{1}{\sigma_{2D}^2} \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \frac{D_i^\top \cdot G_i^{(l)} \cdot D_i}{(1 + \langle C_{(l)} | T \star M_i \rangle)^2} = \sum_{i=1}^N D_i^\top \cdot K_i \cdot D_i \tag{C.4}$$

Pour calculer le jacobien J_Φ , nous devons dériver Φ par rapport aux M_i et aux $\tilde{m}_i^{(l)}$:

$$\frac{\partial \Phi}{\partial \tilde{m}_i^{(l)}} = -\xi_i^l \cdot \frac{1}{\sigma_{2D}^2} \frac{D_i^\top \cdot \left(Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top \right)^\top}{1 + \langle C_{(l)} | T \star M_i \rangle} \tag{C.5}$$

$$\frac{\partial \Phi}{\partial M_i} = \sum_{l=1}^S \xi_i^l \cdot \frac{1}{\sigma_{2D}^2} \frac{D_i^\top \cdot G_i^{(l)} \cdot R}{(1 + \langle C_{(l)} | T \star M_i \rangle)^2} = D_i^\top \cdot K_i \cdot R \tag{C.6}$$

Nous pouvons maintenant décrire le jacobien de Φ :

$$J_\Phi = \frac{\partial \Phi}{\partial \chi} = \left[\frac{\partial \Phi}{\partial M_1}, \dots, \frac{\partial \Phi}{\partial M_N}, \frac{\partial \Phi}{\partial \tilde{m}_1^{(1)}}, \dots, \frac{\partial \Phi}{\partial \tilde{m}_1^{(N)}}, \dots, \frac{\partial \Phi}{\partial \tilde{m}_1^M}, \dots, \frac{\partial \Phi}{\partial \tilde{m}_N^M} \right]$$

où χ est un vecteur contenant toutes nos données :

$$\chi = \begin{pmatrix} M_1 \\ \vdots \\ M_N \\ \tilde{m}_1^1 \\ \vdots \\ \tilde{m}_N^1 \\ \vdots \\ \tilde{m}_1^M \\ \vdots \\ \tilde{m}_N^M \end{pmatrix} \quad \text{avec} \quad \Sigma_{\chi\chi} = \begin{bmatrix} \Sigma_{M_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & & & & & & & \vdots \\ \vdots & \ddots & \Sigma_{M_N} & \ddots & & & & & & & \vdots \\ \vdots & & \ddots & \Sigma_{\tilde{m}_1^1} & \ddots & & & & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & & & & \vdots \\ \vdots & & & & \ddots & \Sigma_{\tilde{m}_1^N} & \ddots & & & & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & & & & \ddots & \Sigma_{\tilde{m}_1^M} & \ddots & & \vdots \\ \vdots & & & & & & & \ddots & \Sigma_{\tilde{m}_1^M} & \ddots & \vdots \\ \vdots & & & & & & & & \ddots & \Sigma_{\tilde{m}_1^M} & \vdots \\ 0 & \cdots & 0 & \Sigma_{\tilde{m}_N^M} \end{bmatrix}$$

Calcul de Σ_{TT} Selon la section 7.1.1, la matrice de covariance des paramètres de la transformation vaut $\Sigma_{TT} = H^{-1} \cdot \frac{\partial \Phi}{\partial \chi} \cdot \Sigma_{\chi\chi} \cdot \frac{\partial \Phi}{\partial \chi}^\top \cdot H^{-\top}$. Pour simplifier les calculs, nous réduisons en premier lieu $\Gamma = \frac{\partial \Phi}{\partial \chi} \cdot \Sigma_{\chi\chi} \cdot \frac{\partial \Phi}{\partial \chi}^\top$.

Supposant un bruit isotrope et d'amplitude égale sur tous les points 3D et sur tous les points 2D (i.e. $\Sigma_{M_i} = \sigma_{3D}^2 \cdot \text{Id}_3$ et $\Sigma_{\tilde{m}_i^{(l)}} = \sigma_{2D}^2 \cdot \text{Id}_2$), nous avons:

$$\begin{aligned} \Gamma &= \sum_{i=1}^N \frac{\partial \Phi}{\partial M_i} \cdot \Sigma_{M_i} \cdot \frac{\partial \Phi}{\partial M_i}^\top + \sum_{l=1}^S \sum_{i=1}^N \frac{\partial \Phi}{\partial \tilde{m}_i^{(l)}} \cdot \Sigma_{\tilde{m}_i^{(l)}} \cdot \frac{\partial \Phi}{\partial \tilde{m}_i^{(l)}}^\top \\ &= \underbrace{\sigma_{3D}^2 \sum_{i=1}^N \frac{\partial \Phi}{\partial M_i} \cdot \frac{\partial \Phi}{\partial M_i}^\top}_{\Gamma_{MM}} + \underbrace{\sigma_{2D}^2 \sum_{l=1}^S \sum_{i=1}^N \frac{\partial \Phi}{\partial \tilde{m}_i^{(l)}} \cdot \frac{\partial \Phi}{\partial \tilde{m}_i^{(l)}}^\top}_{\Gamma_{mm}} \end{aligned}$$

Finalement, nous aboutissons à la formulation suivante de la matrice de covariance de T :

$$\Sigma_{TT} = H^{-1} \cdot (\Gamma_{MM} + \Gamma_{mm}) \cdot H^{-\top} \quad (\text{C.7})$$

avec les simplifications suivantes dans les expressions :

$$\begin{aligned} \Gamma_{mm} &= \frac{1}{\sigma_{2D}^2} \sum_{i=1}^N \sum_{l=1}^S \xi_i^l \cdot \frac{D_i^\top \cdot (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)^\top \cdot (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top) \cdot D_i}{(1 + \langle C_{(l)} | T * M_i \rangle)^2} = \sum_{i=1}^N D_i^\top \cdot L_i \cdot D_i \\ \Gamma_{MM} &= \sigma_{3D}^2 \sum_{i=1}^N D_i^\top \cdot K_i \cdot K_i \cdot D_i \\ H &= \sum_{i=1}^N D_i^\top \cdot K_i \cdot D_i \end{aligned}$$

C.2 Cas du critère EPPC

Nous nous tournons maintenant vers la prédiction de l'incertitude sur T après optimisation de l'EPPC défini à l'équation 4.6. Ignorant les termes constants, le critère s'écrit :

$$C(T, M_1, \dots, M_N) = \sum_{i=1}^N \frac{\|\tilde{M}_i - M_i\|^2}{2 \cdot \sigma_{3D}^2} + \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot \frac{\|\tilde{m}_i^{(l)} - m_i^{(l)}\|^2}{2 \cdot \sigma_{2D}^2}$$

où $m_i^{(l)} = P^{(l)}(T \star M_i)$ est défini comme dans la section précédente. Ses dérivées ont donc encore la forme données aux équations C.1 et C.2.

Équations implicites à l'optimum Cette fois, la minimisation est réalisée non seulement sur les paramètres de la transformation T mais aussi sur les coordonnées 3D parfaites M_i . Nous avons donc à l'optimum :

$$\Phi_T = \left(\frac{\partial C}{\partial T} \right)^\top = 0 \quad \text{et} \quad \Phi_{M_i} = \left(\frac{\partial C}{\partial M_i} \right)^\top = 0$$

Étant donné que le premier terme du critère ne dépend pas directement de la transformation, nous pouvons réutiliser directement les résultats de la dernière section pour obtenir Φ_T :

$$\Phi_T = \frac{1}{\sigma_{2D}^2} \sum_{l=1}^S \sum_{i=1}^N \xi_i^l \cdot D_i^\top \frac{(Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)^\top (m_i^{(l)} - \tilde{m}_i^{(l)})}{1 + \langle C_{(l)} | T \star M_i \rangle} \quad (\text{C.8})$$

Puis une courte étape clarifie Φ_{M_i} :

$$\begin{aligned} \Phi_{M_i} &= \left(\frac{(M_i - \tilde{M}_i)^\top}{\sigma_{3D}^2} + \sum_{l=1}^S \xi_i^l \cdot \frac{(m_i^{(l)} - \tilde{m}_i^{(l)})^\top}{\sigma_{2D}^2} \cdot \frac{\partial m_i^{(l)}}{\partial M_i} \right)^\top \\ &= \frac{(M_i - \tilde{M}_i)}{\sigma_{3D}^2} + \sum_{l=1}^S \xi_i^l \cdot \frac{R^\top \cdot (Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top)^\top \cdot (m_i^{(l)} - \tilde{m}_i^{(l)})}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)} \end{aligned} \quad (\text{C.9})$$

Extraction de la variabilité sur la transformation optimale Dans notre application particulière, les points 3D exacts M_i sont seulement des variables auxiliaires et non pas des paramètres d'intérêts : nous recherchons uniquement la covariance $\Sigma_{TT} = E((T - \bar{T})(T - \bar{T})^\top) = E(\delta T \cdot \delta T^\top)$ de la transformation, et non pas la covariance conjointe de tous les paramètres. Regroupons les coordonnées de tous les points 3D exacts dans M , de telle sorte que le vecteur de paramètres et la fonction implicite définissant l'optimum deviennent $p = \begin{bmatrix} T \\ M \end{bmatrix}$ et

$\Phi(p, \chi) = \begin{bmatrix} \Phi_T \\ \Phi_M \end{bmatrix} = 0$. Dans les paragraphes suivant, le vecteur de données χ contient à la suite les coordonnées des points 3D et 2D observés \tilde{M}_i et $\tilde{m}_i^{(l)}$. Quand les données varient de χ à $\chi + \delta\chi$, les paramètres optimaux varient de p (où $\Phi(p, \chi) = 0$) à $p + \delta p$ de telle sorte qu'au second ordre près, nous ayons :

$$\Phi(p + \delta p, \chi + \delta\chi) = \frac{\partial \Phi}{\partial p} \cdot \delta p + \frac{\partial \Phi}{\partial \chi} \cdot \delta\chi = 0$$

Cela signifie que :

$$\delta p = \begin{bmatrix} \delta T \\ \delta M \end{bmatrix} = - \left(\frac{\partial \Phi}{\partial p} \right)^{-1} \cdot \frac{\partial \Phi}{\partial \chi} \cdot \delta \chi = - \begin{bmatrix} \frac{\partial \Phi_T}{\partial T} & \frac{\partial \Phi_T}{\partial M} \\ \frac{\partial \Phi_M}{\partial T} & \frac{\partial \Phi_M}{\partial M} \end{bmatrix}^{(-1)} \cdot \begin{bmatrix} \frac{\partial \Phi_T}{\partial \chi} \\ \frac{\partial \Phi_M}{\partial \chi} \end{bmatrix} \cdot \delta \chi$$

Arrivés à ce point, nous pourrions calculer la forme analytique de $\frac{\partial \Phi}{\partial p}$ et $\frac{\partial \Phi}{\partial \chi}$, utiliser une méthode numérique pour inverser la matrice $\frac{\partial \Phi}{\partial p}$, calculer $\Sigma_{pp} = E(\delta p \cdot \delta p^\top) = \begin{bmatrix} \Sigma_{TT} & \Sigma_{TM} \\ \Sigma_{MT} & \Sigma_{MM} \end{bmatrix}$ et finalement extraire Σ_{TT} de cette grande matrice. Néanmoins, il est plus judicieux de pousser les calculs un pas plus loin pour résoudre analytiquement δT puisque la formule finale sera plus aisée à comprendre et à implémenter. De plus, nous éviterons le calcul difficile et inutile de δM parce que nous nous focalisons seulement sur la première partie de l'incertitude sur p , i.e. δT . Pour exprimer analytiquement $\frac{\partial \Phi}{\partial p}^{-1}$, nous réalisons une inversion matricielle par bloc avec $A = \frac{\partial \Phi_T}{\partial T}$, $B = \frac{\partial \Phi_T}{\partial M}$ et $C = \frac{\partial \Phi_M}{\partial M}$:

$$\begin{bmatrix} A & B \\ B^\top & C \end{bmatrix} \cdot \begin{bmatrix} D & E \\ E^\top & F \end{bmatrix} = Id \quad \iff \quad \begin{cases} D = (A - B \cdot C^{-1} \cdot B^\top)^{-1} \\ F = (C - B^\top \cdot A^{-1} \cdot B)^{-1} \\ E = -A^{-1} \cdot B(C - B^\top \cdot A^{-1} \cdot B)^{-1} \\ \quad = -(A - B \cdot C^{-1} \cdot B^\top)^{-1} \cdot B \cdot C^{-1} \end{cases}$$

Nous obtenons donc l'expression analytique de δT :

$$\delta T = \left(D \cdot \frac{\partial \Phi_T}{\partial \chi} + E \cdot \frac{\partial \Phi_M}{\partial \chi} \right) \cdot \delta \chi = (A - B \cdot C^{-1} \cdot B^\top)^{-1} \left[\frac{\partial \Phi_T}{\partial \chi} - B \cdot C^{-1} \cdot \frac{\partial \Phi_M}{\partial \chi} \right] \cdot \delta \chi \quad (\text{C.10})$$

Dérivées partielles de Φ Puisque $\frac{\partial \Phi_{M_i}}{\partial M_j} = 0$ pour $i \neq j$, nous devons encore calculer

$$A = H_T = \frac{\partial \Phi_T}{\partial T}, \quad B = \frac{\partial \Phi_T}{\partial M} = \left[\frac{\partial \Phi_T}{\partial M_1}, \dots, \frac{\partial \Phi_T}{\partial M_N} \right] \quad \text{et} \quad C = H_M = \begin{bmatrix} \frac{\partial \Phi_{M_1}}{\partial M_1} & & 0 \\ & \ddots & \\ 0 & & \frac{\partial \Phi_{M_N}}{\partial M_N} \end{bmatrix}$$

en plus de :

$$\frac{\partial \Phi_T}{\partial \chi} = \left[\dots \frac{\partial \Phi_T}{\partial \tilde{M}_i} \dots \frac{\partial \Phi_T}{\partial \tilde{m}_i^{(l)}} \dots \right] \quad \text{et} \quad \frac{\partial \Phi_{M_j}}{\partial \chi} = \left[\dots \frac{\partial \Phi_{M_j}}{\partial \tilde{M}_i} \dots \frac{\partial \Phi_{M_j}}{\partial \tilde{m}_i^{(l)}} \dots \right]$$

A l'aide des résultats de la section précédente (notamment les formules C.4, C.5 et C.6), nous savons déjà que :

$$\begin{aligned} H_T = \frac{\partial \Phi_T}{\partial T} &= \sum_{i=1}^N D_i^\top \cdot K_i \cdot D_i \\ \frac{\partial \Phi_T}{\partial \tilde{m}_i^{(l)}} &= -\xi_i^l \cdot \frac{1}{\sigma_{2D}^2} \frac{D_i^\top \cdot \left(Q^{(l)} - m_i^{(l)} \cdot C_{(l)}^\top \right)^\top}{1 + \langle C_{(l)} | T \star M_i \rangle} \\ B_i = \frac{\partial \Phi_T}{\partial M_i} &= D_i^\top \cdot K_i \cdot R \end{aligned}$$

Les autres dérivées sont calculées similairement :

$$\begin{aligned}
 \frac{\partial \Phi_T}{\partial \tilde{M}_i} &= 0 \\
 H_{M_i} &= \frac{\partial \Phi_{M_i}}{\partial M_i} = \frac{Id_3}{\sigma_{3D}^2} + R^t \cdot K_i \cdot R \\
 \frac{\partial \Phi_{M_i}}{\partial \tilde{M}_j} &= -\delta_{ij} \cdot \frac{Id_3}{\sigma_{3D}^2} \\
 \frac{\partial \Phi_{M_i}}{\partial \tilde{m}_j^{(l)}} &= -\delta_{ij} \cdot \xi_j^l \cdot \frac{R^\top \cdot (Q - m_j^{(l)}) \cdot C_{(l)}^\top}{\sigma_{2D}^2 (1 + \langle C_{(l)} | T \star M_i \rangle)}
 \end{aligned}$$

Covariance de la transformation Écrivons :

$$\begin{aligned}
 H &= A - B \cdot C^{-1} \cdot B^\top = H_T - B \cdot H_M^{-1} \cdot B^\top \\
 \Gamma &= \left[\frac{\partial \Phi_T}{\partial \chi} - B \cdot H_M^{-1} \cdot \frac{\partial \Phi_M}{\partial \chi} \right] \Sigma_{\chi\chi} \left[\frac{\partial \Phi_T}{\partial \chi} - B \cdot H_M^{-1} \cdot \frac{\partial \Phi_M}{\partial \chi} \right]^\top
 \end{aligned}$$

A l'aide de C.10, la covariance de la transformation peut être écrite :

$$\Sigma_{TT} = E \left(\delta T \cdot \delta T^\top \right) = H^{(-1)} \cdot \Gamma \cdot H^{(-1)} \quad (C.11)$$

Pour calculer Γ , considérons tout d'abord la composante relative aux \tilde{M}_i et aux $\tilde{m}_i^{(l)}$ de χ dans la matrice $\left[\frac{\partial \Phi_T}{\partial \chi} - B \cdot C^{-1} \cdot \frac{\partial \Phi_M}{\partial \chi} \right]$:

$$\left[\frac{\partial \Phi_T}{\partial \chi} - B \cdot C^{-1} \cdot \frac{\partial \Phi_M}{\partial \chi} \right]_{\tilde{M}_i} = \underbrace{\frac{\partial \Phi_T}{\partial \tilde{M}_i}}_0 - \sum_{k=1}^N B_k \cdot H_{M_k}^{-1} \cdot \frac{\partial \Phi_{M_k}}{\partial \tilde{M}_i} = \frac{B_i \cdot H_{M_i}^{-1}}{\sigma_{3D}^2} \quad (C.12)$$

$$\begin{aligned}
 \left[\frac{\partial \Phi_T}{\partial \chi} - B \cdot C^{-1} \cdot \frac{\partial \Phi_M}{\partial \chi} \right]_{\tilde{m}_i^{(l)}} &= \frac{\partial \Phi_T}{\partial \tilde{m}_i^{(l)}} - \sum_{k=1}^N B_k \cdot H_{M_k}^{-1} \cdot \frac{\partial \Phi_{M_k}}{\partial \tilde{m}_i^{(l)}} \\
 &= \xi_i^l \cdot \left(-\frac{D_i^\top \cdot (Q - m_i^{(l)}) \cdot C_{(l)}^\top}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)} + \frac{B_i \cdot H_{M_i}^{-1} \cdot R^\top \cdot (Q - m_i^{(l)}) \cdot C_{(l)}^\top}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)} \right) \\
 &= \xi_i^l \cdot (B_i \cdot H_{M_i}^{-1} \cdot R^\top - D_i^\top) \frac{(Q - m_i^{(l)}) \cdot C_{(l)}^\top}{\sigma_{2D}^2 \cdot (1 + \langle C_{(l)} | T \star M_i \rangle)} \quad (C.13)
 \end{aligned}$$

Les points 3D \tilde{M}_i et les points 2D $\tilde{m}_i^{(l)}$ étant indépendants, $\Sigma_{\chi\chi}$ est une matrice diagonale par bloc. De plus, nous supposons un bruit isotrope et d'amplitude égale sur tous les points 3D et sur tous les points 2D (i.e. $\Sigma_{M_i} = \sigma_{3D}^2 \cdot Id_3$ et $\Sigma_{\tilde{m}_i^{(l)}} = \sigma_{2D}^2 \cdot Id_2$). A l'aide de (C.12) and (C.13) nous pouvons alors écrire Γ :

$$\Gamma = \sum_{i=1}^N \frac{B_i \cdot H_{M_i}^{-1} H_{M_i}^{-1} \cdot B_i^\top}{\sigma_{3D}^2} + \sum_{i=1}^N \sum_{l=1}^S \xi_i^l \cdot (B_i \cdot H_{M_i}^{-1} \cdot R^\top - D_i^\top) \cdot L_i^{(l)} \cdot (B_i \cdot H_{M_i}^{-1} \cdot R^\top - D_i^\top)^\top$$

Pour simplifier encore plus l'expression, remarquons que

$$B_i \cdot H_{M_i}^{-1} \cdot R^\top = D_i^\top \cdot \left(\frac{K_i^{-1}}{\sigma_{3D}^2} + Id_3 \right)^{-1}$$

Cette expression sera utilisée pour simplifier le premier terme de Γ . Pour le second terme, nous utilisons le lemme d'inversion $(A + B)^{-1} = A^{-1} - A^{-1}(B^{-1} + A^{-1})^{-1}A^{-1}$ pour obtenir :

$$B_i \cdot H_{M_i}^{-1} \cdot R^\top = D_i^\top \cdot \left(Id_3 - \left(\sigma_{3D}^2 Id_3 + K_i^{-1} \right)^{-1} \cdot K_i^{-1} \right)$$

Cela nous conduit à :

$$\Gamma = \sum_{i=1}^N \frac{D_i^\top}{\sigma_{3D}^2} \cdot \left(\frac{K_i^{-1}}{\sigma_{3D}^2} + Id \right)^{-2} \cdot D_i + \sum_{i=1}^N D_i^\top (\sigma_{3D}^2 Id + K_i^{-1})^{-1} \cdot K_i^{-1} \cdot L_i \cdot K_i^{-1} \cdot (\sigma_{3D}^2 Id + K_i^{-1})^{-1} \cdot D_i$$

et finalement, après quelques simplifications supplémentaires à :

$$\Gamma = \sum_{i=1}^N D_i^\top (\sigma_{3D}^2 Id + K_i^{-1})^{-1} (\sigma_{3D}^2 Id + K_i^{-1} \cdot L_i \cdot K_i^{-1}) (\sigma_{3D}^2 Id + K_i^{-1})^{-1} \cdot D_i \quad (C.14)$$

Il ne nous reste plus qu'à simplifier H . Grâce aux dérivées précédemment calculées, nous arrivons rapidement à :

$$H = H_T - \sum_{i=1}^N B_i \cdot H_{M_i}^{-1} \cdot B_i^\top = \sum_{i=1}^N D_i^\top \cdot K_i \cdot D_i - \sum_{i=1}^N D_i^\top \cdot \left[K_i \cdot \left(\frac{Id}{\sigma_{3D}^2} + K_i \right)^{-1} \cdot K_i \right] \cdot D_i$$

Utilisant encore une fois le lemme d'inversion, nous aboutissons finalement à :

$$H = \sum_{i=1}^N D_i^\top \cdot \left(\sigma_{3D}^2 \cdot Id + K_i^{-1} \right)^{-1} \cdot D_i \quad (C.15)$$

En introduisant Γ et H dans l'équation C.11, nous obtenons la formulation analytique de Σ_{TT} pour l'EPCC.

Annexe D

Estimation de la position 3D relative de n points

D.1 Estimation à partir de mesures issues d'images stéréoscopiques

Nous présentons dans cette annexe la méthode que nous avons utilisé pour estimer la position relative des 4 coins du marqueur carré qui est fixé sur l'aiguille du praticien. La méthode proposée, présentée pour 4 points est évidemment généralisable au cas de n points.

Nous disposons de N paires d'images dans lesquelles est visible le marqueur carré de l'AR-Toolkit. Nous désignons par $\tilde{M}_1^i, \tilde{M}_2^i, \tilde{M}_3^i, \tilde{M}_4^i$ les reconstructions stéréoscopiques des 4 coins obtenues à partir de la i^{ieme} paire d'images. $\tilde{m}_j^{1,i}$ (resp. $\tilde{m}_j^{2,i}$) correspond à l'extraction du j^{ieme} point (j allant de 1 à 4) dans la i^{ieme} acquisition vidéo de la première (resp. seconde) caméra. P_1 et P_2 correspondent aux fonctions projectives des deux caméras. Nous appelons $M_1 \dots M_4$ les positions 3D exactes des 4 coins dans un repère quelconque inconnu appelé U . L'introduction de ce repère traduit le fait que nous recherchons une forme qui est indépendante du repère dans lequel elle est définie. $T_1 \dots T_i \dots T_N$ correspondent aux transformations rigides exactes qui recalent dans le repère U le marqueur vu dans la i^{ieme} paire d'images. Nous minimisons alors le critère suivant, dérivé d'un maximum de vraisemblance (les hypothèses sont équivalentes à celles utilisées pour dériver l'ISPPC dans la section 4.1.2):

$$C(T_1 \dots T_N, M_1 \dots M_4) = \sum_{j=1}^4 \sum_{i=1}^N \left(\frac{\| P_1(T_i \star M_j) - \tilde{m}_j^{1,i} \|^2}{2 \cdot \sigma_{2D}^2} + \frac{\| P_2(T_i \star M_j) - \tilde{m}_j^{2,i} \|^2}{2 \cdot \sigma_{2D}^2} \right)$$

Ce critère est symétrique puisqu'il ne suppose pas implicitement que nous cherchons une solution $\hat{M}_1 \dots \hat{M}_4$ dans un des repère liés à une acquisition vidéo particulière. En fait, il y a une infinité de solution puisque la forme recherchée peut être définie dans un repère quelconque. Cela se concrétise mathématiquement par le fait que le changement de variable: $\hat{T}'_i = \hat{T}'_i \star T^{-1}$ et $\hat{M}'_j = T \star \hat{M}'_j$ fournit une autre solution (les \hat{T}'_i et les \hat{M}'_j , quel que soit T).

Méthode d'optimisation

Comme la forme ne dépend pas du repère, nous pouvons choisir arbitrairement le repère dans lequel nous allons chercher une solution. En l'occurrence, nous avons décidé de la chercher dans le repère lié à la première acquisition vidéo. Nous initialisons donc les $M_1 \dots M_4$ par $\tilde{M}_1^1 \dots \tilde{M}_4^1$, et nous optimisons sur les $T_1 \dots T_i \dots T_N$. L'initialisation de T_i est donnée par la transformation rigide qui recalc les $\tilde{M}_1^i \dots \tilde{M}_4^i$ sur les $\tilde{M}_1^1 \dots \tilde{M}_4^1$ (nous employons un recalage 3D aux moindres carrés classique [AHB87]). A la fin de cette première phase d'optimisation, nous effectuons le changement de variable suivant afin d'estimer, à l'étape suivante, la position de $M_1 \dots M_4$ dans le repère choisi (en l'occurrence celui lié à la première acquisition):

$$T'_i = T_i \star T_1^{-1} \quad \text{et} \quad M'_j = T_1 \star M_j$$

Nous optimisons alors les $M_1 \dots M_4$, et alternons ensuite les optimisations en nous replaçant à chaque fois dans le repère choisi après la phase d'estimation des T_i . Dans nos conditions, nous estimons que la convergence est atteinte lorsque la distance entre deux estimations successives s et $s + 1$ des points $M_1 \dots M_4$ sont très proches:

$$1/4 \cdot \sum_{k=1}^4 d(\hat{M}_{k,s+1}, \hat{M}_{k,s}) < 10^{-4} \text{ mm}$$

Bibliographie

- [AAK71] Y. I. Abdel-Aziz and H. M. Karara. Direct linear transformation into object space coordinates in close-range photogrammetry. In *Proceedings of the Symposium on close-Range Photogrammetry*, pages 1–18, Urbana, 1971.
- [ABK⁺03] René Anxionnat, Marie-Odile Berger, Erwan Kerrien, Serge Bracard, and Luc Picard. Intra- and inter-observer variability in the angiographic delineation of brain arterio-venous malformations (avms). In *17th International Congress and Exhibition on Computer Assisted Radiology and Surgery - CARS'2003, Londres, Royaume-Uni*, pages 1297–1298. Elsevier, Jun 2003.
- [AD03] A. Ansar and K. Daniilidis. Linear pose estimation from points or lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):578–589, 2003.
- [AF86] N Ayache and O. Faugeras. Hyper: A new approach for the recognition and positioning of 2d objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):44–54, January 1986.
- [AHB87] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3d point sets. *IEEE Transactions on Pattern analysis and machine intelligence*, 9(5):698–700, september 1987.
- [Aya89] N. Ayache. *Vision Stéréoscopique et Perception Multisensorielle: Application à la robotique mobile*. Inter-Editions (MASSON), 1989. 314 pages, in French.
- [Aya91] N. Ayache. *Artificial Vision for Mobile robots - Stereo-vision and Multisensor Perception*. MIT-Press, 1991. 342 pages.
- [Aya03] N. Ayache. Epidaure: A research project in medical image analysis, simulation, and robotics at inria. *IEEE Transactions on Medical Imaging*, 22(10):1185–1201, October 2003.
- [Azu97] R. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, August 1997.
- [Ba02] F. Banovac and al. Liver tumor biopsy in a respiring phantom with the assistance of a novel electromagnetic navigation device. In *Sixth International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'03)*, pages 200–207, 2002.
- [BCMDV02] Olivier Bantiche, Ève Coste-Manière, Frédéric Devernay, and Thierry Viéville. A video-projector-mono-camera stereoscope for surgical planning transfer. In *Surgetica CAS*, pages 158–163, Grenoble, France, September 2002.
- [BFTC03] EM Boctor, G. Fichtinger, RH. Taylor, and M. Choti. Tracked 3d ultrasound in radio-frequency liver ablation. In *SPIE Annual Conference on Medical Imaging 2003*, volume 5035, pages 174–182, 2003.

- [BGK⁺04] J.M. Blackall, Penney G.P., A.P. King, A. Adam, and D.J. Hawkes. Tracking alignment of sparse ultrasound with preoperative images of the liver and an interventional plan using models of respiratory motion and deformation. In *In Proceedings SPIE Medical Imaging 2004.*, 2004.
- [BK02] M. Billinghurst and H. Kato. Collaborative augmented reality. *Communications of the ACM*, 45(7):64–70, July 2002.
- [BKG⁺01] J.M. Blackall, A.P. King, Penney G.P., A. Adam, and D.J. Hawkes. A statistical model of respiratory motion and deformation of the liver. In *Fourth International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'01)*, volume volume 2208 LNCS, pages 1338–1340. Springer-Verlag, October 2001.
- [BKS⁺93] G. Barnett, D. Kormos, C. Steiner, and J. Weisenberger. Use of frameless, armless stereotactic wand for brain tumor localisation with two-dimensional and three-dimensional neuroimaging. *Neurosurgery*, 33(4):674–678, 1993.
- [BLM⁺98] J.M. Balter, K.L. Lam, C.J. McGinn, T.S. Lawrence, and R.K. Ten Haken. Improvement of CT-based treatment-planning models of abdominals targets using static exhale imaging. *Int. J. Radiation Oncology Biol. Phys.*, 41(4):939–943, 1998.
- [BM92] P.J. Besl and N.D. McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [BNR03] G. Baratoff, A. Neubeck, and H. Regenbrecht. Interactive multi-marker calibration of augmented reality applications. In IEEE Computer Society, editor, *IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pages 107–116, 2003.
- [BRF01] D. Bandyopadhyay, R. Raskar, and H. Fuchs. Dynamic shader lamps: Painting on movable objects. In *IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, pages 207–216, New-York, October 2001. IEEE and ACM.
- [Bro71] D. C. Brown. Decentering distortion of lenses. *Photometric Engineering Remote Sensing*, 37(8):855–866, 1971.
- [BTFC03] EM Boctor, RH. Taylor, G. Fichtinger, and M. Choti. Robotically assisted intraoperative ultrasound with application to ablative therapy of liver cancer. In *SPIE Annual Conference on Medical Imaging 2003*, volume 5029, pages 281–291, 2003.
- [BTHL⁺96] J. Balter, R. Ten Haken, T. Lawrence, K. Lam, and J. Robertson. Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing. *Int. J. Radiation Oncology Biol. Phys.*, 36(1):167–174, 1996.
- [BWDPS99] M.-O. Berger, B. Wrobel-Dautcourt, S. Petitjean, and G. Simon. Mixing synthetic and video images of an outdoor urban environment. *Machine Vision and Applications*, 11(3):145–159, 1999.
- [CBLC02] M. Clifford, F. Banovac, E. Levy, and K. Cleary. Assessment of hepatic motion secondary to respiration for computer assisted interventions. *Computer Aided Surgery*, 7:291–299, 2002.
- [CL98] F. Chassat and S. Lavallée. Experimental protocol for accuracy evaluation of 6-d localizers for computer-integrated surgery: Application to four optical

-
- localizers. In *First International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'98)*, pages 277–284, 1998.
- [CMAMB04] Ève Coste-Manière, Louaï Adhami, Fabien Mourgues, and Olivier Bantiche. Optimal planning of robotically assisted heart surgery: Transfer precision in the operating room. *International Journal of Robotics Research*, 23(4):539–548, May 2004.
- [Cot97] S. Cotin. *Modèles anatomiques déformables en temps réel: Application à la simulation de chirurgie avec retour d'effort*. Thèse de sciences, Université de Nice Sophia-Antipolis, November 1997.
- [CRHH99] M.J. Clarkson, D. Rueckert, D. Hill, and D.J. Hawkes. Registration of multiple video images to pre-operative CT for image guided surgery. In *Proceedings of SPIE, Medical Imaging*, pages 14–23, February 1999.
- [CRHH00] M. Clarkson, D. Rueckert, D. Hill, and D. Hawkes. A multiple 2d video - 3d medical image registration algorithm. In *Medical Imaging 2000*, pages 342–352, 2000.
- [DAP⁺99] NL Dorward, O Alberti, JD Palmer, ND Kitchen, and DG Thomas. Accuracy of true frameless stereotaxy: in vivo measurement and laboratory phantom studies. *Journal of Neurosurgery*, 90(1):160–168, 1999.
- [DBC⁺01] S. De Buck, J. V. Cleynebreugel, I. Geys, T. Koninckx, P. Koninck, and P. Suetens. A system to support laparoscopic surgery by augmented reality visualization. In W. Niessen and M. Viergever, editors, *Fourth International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'01)*, pages 691–698, Utrecht, October 2001.
- [DBCM01] Frédéric Devernay, Olivier Bantiche, and Ève Coste-Manière. Structured light on dynamic scenes using standard stereoscopy algorithms. Research report, ChIR Medical Robotics group. INRIA Sophia Antipolis, March 2001.
- [DD95] D. DeMenthon and L. Davis. Model-based object pose in 25 lines of code. *International Journal of Computer Vision*, 15:123–141, 1995.
- [Del99] H. Delingette. General object reconstruction based on simplex meshes. *International Journal of Computer Vision*, 32(2):111–146, September 1999.
- [Dev97] F. Devernay. *Vision stéréoscopique et propriétés différentielles des surfaces*. PhD thesis, Ecole Polytechnique, Février 1997.
- [DHH⁺94] S. Davies, A. Hill, R. Holmes, M. Halliwell, and P. Jackson. Ultrasound quantitation of respiratory organ motion in the upper abdomen. *British Journal Radiology*, 67:1096–1102, 1994.
- [Dra93] D. Drascic. Stereoscopic video and augmented reality. *Scientific Computing and Automation*, 9(7):31–34, June 1993.
- [DRL89] M. Dhome, M. Richetin, and G. Lapreste, J. and Rives. Determination of the attitude of 3d objects from a single perspective view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12):1265–1278, December 1989.
- [Dub01] E. Dubois. *Chirurgie augmentée, un cas de réalité augmentée; Conception et réalisation centrées sur l'utilisateur*. Thèse de doctorat, Université Joseph Fourier, Grenoble, 2001.
- [EKM⁺00] P.J. Edwards, A.P. King, C.R. Maurer, D.A. De Cunha, D.J. Hawkes, D.L. Hill, R.P. Gaston, M.R. Fenlon, A. Jusczyzck, A.J. Strong, C.L. Chandler, and

- M.J. Gleeson. Design and evaluation of a system for microscope-assisted guided interventions (MAGI). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1082–1093, November 2000.
- [EW03] M. Billinghamurst E. Woods, P. Mason. Magicmouse: an inexpensive 6-degree-of-freedom mouse. In *Proceedings of Graphite*, Melbourne, 2003.
- [Fai75] W. Faig. Calibration of close-range photogrammetric systems: Mathematical formulation. *Photometric Engineering Remote Sensing*, 41(12):1479–1486, 1975.
- [FAMS03] M. Fiorentino, R. Amicis, G. Monno, and A. Stork. Spacedesign: A mixed reality workspace for aesthetic industrial design. In IEEE Computer Society, editor, *IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pages 86–94, 2003.
- [Fau93] O. D. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.
- [FB81] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, June 1981.
- [FDP⁺02] G. Fichtinger, T. DeWeese, A. Patriciu, A. Tanacs, D. Mazilu, J. Anderson, K. Masamune, R. Taylor, and D. Stoianovici. System for robotically assisted prostate biopsy and therapy with intraoperative CT guidance. *Journal of Academic Radiology*, 9(1):60–74, January 2002.
- [Fel95] J. Feldmar. *Recalage rigide, non rigide et projectif d'images médicales tridimensionnelles*. PhD thesis, Ecole Polytechnique, décembre 1995.
- [FH83] O. Faugeras and M. Hebert. A 3d recognition and positioning algorithm using geometrical matching between primitive surfaces. In *International Joint Conference on Artificial Intelligence*, pages 996–1002, Karlsruhe, August 1983.
- [Fit99] J. Fitzpatrick. Predicting error in rigid-body point-based registration. *IEEE Trans. on Medical Imaging*, 17(5):694–702, 1999.
- [For03] C. Forest. *Simulation de chirurgie par coelioscopie: contributions à l'étude de la découpe volumique, au retour d'effort et à la modélisation des vaisseaux sanguins*. PhD thesis, École Polytechnique, March 2003.
- [FSJ01] F. Fleig, O. and Devernay, J.-M. Scarabin, and P. Jannin. Surface reconstruction of the surgical field from stereoscopic microscope views in neurosurgery. In *CARS'2001 (Computer Assisted Radiology and Surgery)*, Berlin, 2001.
- [FT86] O.D. Faugeras and G. Toscani. Calibration problem for stereo. In *Proc. Int. Conf. Computer Vision and Patten Recognition*, pages 15–20, Miami, 1986.
- [G.00] Bianchi G. Réalité augmentée en chirurgie hépatique. Technical report, Université Blaise Pascal Clermont-Ferrand II, Septembre 2000. Rapport de stage de DEA.
- [Gan84] S. Ganapathy. Decomposition of transformation matrices for robot vision. *Pattern Recognition Letters*, 2(6):401–412, December 1984.
- [GBDH95] P. Graebing, C. Boucher, C. Daul, and E. Hirsch. 3d sculptured surface analysis using a structured-light approach. In *Proceedings of SPIE Volume: 2598*, pages 128–139, 1995.
- [GEW⁺96] W. Grimson, G. Ettinger, S. White, W. Wells T. Lozano-Perez, and R. Kikinis. An automatic registration method for frameless stereotaxy, image-guided

-
- surgery and enhanced reality visualization. *IEEE Trans. Medical Imaging*, 15(2):129–140, April 1996.
- [GKO⁺00] A. Gronningsaeter, A. Kleven, S. Ommedal, T.E. Aarseth, T. Lie, F. Lindseth, T. Lango, and G. Unsgard. Sonowand, an ultrasound-based neuronavigation system. *Neurosurgery*, 47(6):1373–1380, 2000.
- [GNK⁺99] D. Gering, A. Nabavi, R. Kikinis, W. Grimson, N. Hata, P. Everett, F. Jolesz, and W. Wells. An integrated visualization system for surgical planning and guidance using image fusion and interventional imaging. In *Second International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'99)*, pages 808–819, 1999.
- [GP02] S. Granger and X. Pennec. Multi-scale EM-ICP: A fast and robust approach for surface registration. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *European Conference on Computer Vision (ECCV 2002)*, volume 2353 of *LNCS*, pages 418–432, Copenhagen, Denmark, 2002. Springer.
- [Gra03] S. Granger. *Une approche statistique multi-échelle au recalage rigide de surfaces : Application à l'implantologie dentaire*. Thèse de sciences, Ecole des Mines de Paris, April 2003.
- [GRS⁺02] Y. Genc, S. Riedel, C. Souvannavong, C. Akinlar, and N. Navab. Marker-less tracking for ar: A learning-based approach. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pages 295–304, October 2002.
- [GTN02] Y. Genc, M. Tuceryan, and N. Navab. Practical solutions for calibration of optical see-through devices. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pages 169–175, October 2002.
- [GW03] N. Glossop and Z. Wang. Laser projection augmented reality system for computer-assisted surgery. In *Computer Assisted Radiology and Surgery (CARS2003)*, volume 1256, pages 65–71, June 2003.
- [GWL⁺99] H. Gumprecht, D. Widenka, C. Lumenta, R. Maciunas, G. Barnett, P. Kelly, and D. Roberts. Brainlab vectorvision neuronavigation system: Technology and clinical experiences in 131 cases. *Neurosurgery*, 44(1):97–105, 1999.
- [HB79] G. Harauz and M. Bronskill. Comparison of the liver's respiratory motion in the suspine and upright positions. *Journal of Nuclear Medicine*, 20:733–735, 1979.
- [HBE86] T.S. Huang, S.D. Blostein, and Margerum E.A. Least-squares estimation of motion parameters from 3d point correspondences. In *IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, June 24-26 1986.
- [HE98] P. A. Helm and T. S. Eckel. Accuracy of registration methods in frameless stereotaxis. *Computer Aided Surgery*, 3:51–56, 1998.
- [HED⁺99] A. Hartov, S. D. Eisner, David, W. Roberts, K. Paulsen, L. A. Platenik, and M. I. Miga. Error analysis for a free-hand three-dimensional ultrasound system for neuronavigation. *Neurosurgery Focus*, 6(3):Article 5, 1999.
- [Her98] G. Hermosillo. Recalage par maximisation de l'information mutuelle. Technical report, Université de Nice, Septembre 1998. Rapport de stage de DEA.
- [HHS⁺00] A. Herline, J. Herring, J. Stefansic, W. Chapman, R. Galloway, and B. Dawa. Surface registration for use in interactive, image-guided liver surgery. In *Computer Aided Surgery*, volume 5, pages 11–17, 2000.

- [Hirdu] Hiro. Human interface technology laboratory, <http://www.hitl.washington.edu/>.
- [HJL⁺89] R. Haralick, H. Joo, CN. Lee, X. Zhuang, V. Vaidya, and MB. Kim. Pose estimation from corresponding point data. *IEEE Trans. on Systems, Man, and Cybernetics*, 19(06):1426–1446, December 1989.
- [HM95] R. Horaud and O. Monga. *Vision par ordinateur, outils fondamentaux*. Edition Hermès, 1995.
- [HS97] J. Heikkilä and O. Silvén. A four-step camera calibration procedure with implicit image correction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pages 1106–1112, San Juan, Puerto Rico, 1997.
- [HSD99] A. Herline, J. Stefansic, and J. et al. Debelak. Image guided surgery: Preliminary feasibility studies of frameless stereotactic liver surgery. *Archives of Surgery*, 134:644–650, June 1999.
- [HTMS82] E. L. Hall, M. B. K. Tio, C. A. McPherson, and F. A. Sadjadi. Curved surface measurement and recognition for robot vision. In *Conf. Rec. IEEE Workshop on Industrial Applications of Machine*, May 1982.
- [HYH85] Y. Hung, PS. Yeh, and D. Harwood. Passive ranging to known planar point sets. *IEEE Int. Conf. on Robotics and Automation*, pages 80–85, March 1985.
- [JFH⁺02] P. Jannin, J.M. Fitzpatrick, D.J. Hawkes, X. Pennec, R. Shahidi, and M. Vannier. Validation of medical image processing in image-guided therapy. *IEEE Transactions on Medical Imaging*, 21(12):1445–1449, December 2002.
- [JFS⁺00] P. Jannin, O.J. Fleig, E. Seigneuret, C. Grova, X. Morandi, and J.M. Scarabin. A data fusion environment for multimodal and multi-informational neuro-navigation. *Journal of Computer Aided Surgery*, 5(1):1–10, 2000.
- [JMF⁺02] P. Jannin, X. Morandi, O.J. Fleig, E. Le Rumeur, P. Toulouse, B. Gibaud, and J.M. Scarabin. Integration of sulcal and functional information for multimodal neuronavigation. *Journal of Neurosurgery*, 96:713–723, April 2002.
- [KER⁺92] H. Korin, R. Ehman, S. Riederer, J. Felmlee, and R. Grimm. Respiratory kinematics of the upper abdominal organs: a quantitative study. *Magnetic Resonance Medicine*, 23:172–178, 1992.
- [KH96] H. D. Kubo and B. C. Hill. Respiration-gated radiotherapy treatment: A technical study. *Phys Med Biol*, 41:83–91, 1996.
- [Kum89] R. Kumar. Determination of camera location and orientation. In *DARPA Image Understanding Workshop*, pages 870–881, Palo Alto, Calif., 1989.
- [Lev97] M. Leventon. A registration, tracking, and visualization system for image guided surgery. Master's thesis, MIT, May 1997.
- [LHF90] Y. Liu, T. Huang, and O. Faugeras. Determination of camera location from 2d to 3d line and point correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(01):28–37, January 1990.
- [LHM00] C. Lu, G.D Hager, and E. Mjolsness. Fast and globally convergent pose estimation from video images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(6):610–622, June 2000.
- [LKH96] X. Liu, T. Kanungo, and R. Haralick. Statistical validation of computer vision software, 1996.

-
- [LLB02] F. Lindseth, T. Lango, and J. Bang. Accuracy evaluation of a 3d ultrasound-based neuronavigation system. *Computer Aided Surgery*, 7:197–222, 2002.
- [LLS02] H. Lamecker, T. Lange, and M. Seebass. A statistical shape model for the liver. In *Sixth International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'03)*, pages 412–427, 2002.
- [Low91] D. Lowe. Fitting parameterized three-dimensional models to images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):441–450, May 1991.
- [LVD98] J.-M. Lavest, M. Viala, and M. Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In *Proceedings of the 5th European Conference on Computer Vision-Volume I (ECCV'98)*, pages 158–174. Springer-Verlag, 1998.
- [LWG97] M. Leventon, W. Wells, and W. Grimson. Multiple view 2d-3d mutual information registration. In *Image Understanding Workshop*, pages 625–630, 1997.
- [LYT⁺02] T. Langø, B. Ystgaard, G. Tangen, T. Hernes, and R. Mårvik. Feasibility of 3d navigation in laparoscopic surgery., Oral presentation at the SMIT (Society for Medical Innovation and Technology) Conference. September 2002. Oslo. Norway.
- [MBHN00] M Mitschke, A. Bani-Hashemi, and N. Navab. Interventions under video-augmented x-ray guidance: Application to needle placement. In *Third International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'00)*, pages 858–868, Pittsburgh, Pennsylvania USA, October 11-14 2000.
- [MDI01] J.F. McGahan and G.D. Dodd III. Radiofrequency ablation of the liver: Current status. *American Journal of Roentgenology*, 176(1):3–16, 2001.
- [MFW⁺97] C. Maurer, J. Fitzpatrick, M. Wang, R. Galloway, R. Macinuas, and G. Allen. Registration of head volume images using implantable fiducial markers. *IEEE Trans. Med. Imaging*, 16:447–462, 1997.
- [MLTP00] P. Merloz, F. Lavalée, J. Tonneti, and L. Pittet. Image-guided spinal surgery: technology, operative technique and clinical practice. *Operative Techniques in Orthopaedics*, 10(1):56–63, 2000.
- [MMN⁺00] K. Masamune, Y. Masutani, N. Nakajima, I. Sakuma, T. Dohi, H. Isek, and K. Takakura. Three-dimensional slice image overlay system with accurate depth perception for surgery. In *Third International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'00)*, Pittsburgh, Pennsylvania USA, October 11-14 2000.
- [Mou03] F. Mourgues. *Guidage par réalité augmentée: application à la chirurgie cardiaque robotisée*. PhD thesis, Université de Nice - Sophia Antipolis, September 2003.
- [MVFCM03] Fabien Mourgues, Thierry Vieville, Volkmar Falk, and Ève Coste-Manière. Interactive guidance by image overlay in robot assisted coronary artery bypass. In *Medical Image Computing and Computer Assisted Intervention (MICCAI'03)*, LNCS, Montréal, Canada, November 2003. Springer.
- [MZD93] P. Milgram, S. Zhai, and D. Drascic. Applications of augmented reality for human-robot communication. In *IROS93*, pages 1467–1476, Yokohama Japan, July 1993.

- [NAM⁺98] S. Nakajima, H. Atsumi, D. C. Metcalf, T. Yoshimine, F. A. Jolesz, P. Black, and R. Kikinis. A simple method of scalp localization using multiplanar reconstruction of mr images. *Surgical Neurology*, 50:597–599, 1998.
- [Nav03] N. Navab. Industrial augmented reality(iar): Challenges in design and commercialization of killer apps. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR'03)*, Tokyo, Japan, October 2003.
- [Nav04] N. Navab. Developing killer apps for industrial augmented reality. *IEEE Computer Graphics and Applications*, pages 16–20, May/June 2004.
- [NBHM99] N. Navab, A. Bani-Hashemi, and M. Mitschke. Merging visible and invisible: two camera-augmented mobile c-arm (camc) applications. In *2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*, pages 131–141, October 1999.
- [NGP⁺03a] S. Nicolau, A. Garcia, X. Pennec, L. Soler, and N. Ayache. Augmented reality guided radio-frequency tumor ablation. In *Proceedings of Augmented and Virtual Reality Workshop (AVIR03)*, pages 34–35, Geneve, 2003.
- [NGP⁺03b] S. Nicolau, A. Garcia, X. Pennec, L. Soler, and N. Ayache. Guidage de ponctions percutanées à l'aide d'un système de vision basé sur une méthode de recalage 3d/2d. In *Actes du colloque Imagerie pour les sciences du vivant et de l'ingénieur (IMVIE03)*, Strasbourg, 2003.
- [NGP⁺04] S. Nicolau, A. Garcia, X. Pennec, L. Soler, and N. Ayache. An augmented reality system to guide radio-frequency tumor ablation. In *the Journal of Computer Animation and Virtual World*, 2004. in press.
- [NPSA03a] S. Nicolau, X. Pennec, L. Soler, and N. Ayache. Evaluation of a new 3D/2D registration criterion for liver radio-frequencies guided by augmented reality. In N. Ayache and H. Delingette, editors, *International Symposium on Surgery Simulation and Soft Tissue Modeling (IS4TM'03)*, volume 2673 of *LNCS*, pages 270–283, Juan-les-Pins, France, 2003. INRIA Sophia Antipolis, Springer-Verlag.
- [NPSA03b] S. Nicolau, X. Pennec, L. Soler, and N. Ayache. Validation of a new 3d/2d registration criterion including error prediction. application to image guided radio-frequency ablation of the liver tumors. Research report 4993, INRIA, 2003.
- [NPSA03c] S. Nicolau, X. Pennec, L. Soler, and N. Ayache. Vision augmentée de structures anatomiques abdominales recalées par stéréoscopie. In *Actes du quatrième colloque francophone. Méthodes et Techniques Optiques pour l'Industrie*, volume 1, pages 413–418, Belfort, 2003.
- [NPSA04] S. Nicolau, X. Pennec, L. Soler, and N. Ayache. An accuracy certified augmented reality system for therapy guidance. In *European Conference on Computer Vision (ECCV'04)*, LNCS 3023, pages 79–91. Springer-Verlag, 2004.
- [OLWK98] Siu-Hang Or, W. S. Luk, K. H. Wong, and Irwin King. An efficient iterative pose estimation algorithm. In *ACCV*, pages 559–566, 1998.
- [OOA⁺89] K. Ohara, T. Okumura, M. Akisada, T. Inada, T. Mori, H. Yokota, and M.J. Calaguas. Irradiation synchronised with respiration gate. *Int. J. Radiation Oncology Biol. Phys.*, 17:853–857, 1989.
- [oP80] American Society of Photogrammetry, editor. *Manual of Photogrammetry*. 1980.

-
- [ORPA00] S. Ourselin, A. Roche, S. Prima, and N. Ayache. Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images. In *Third International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'00)*, pages 557–566, Pittsburgh, Pennsylvania USA, October 11-14 2000.
- [PARC01] X. Pennec, N. Ayache, A. Roche, and P. Cachier. Non-rigid MR/US registration for tracking brain deformations. In IEEE Computer Society Press, editor, *Proc of Int. Workshop on Medical Imaging and Augmented Reality (MIAR 2001), 10-12 June 2001, Shatin, Hong Kong*, pages 79–86, June 2001.
- [PAT00] X. Pennec, N. Ayache, and J.-P. Thirion. Landmark-based registration using features identified through differential geometry. In I. Bankman, editor, *Handbook of Medical Imaging*, chapter 31, pages 499–513. Academic Press, September 2000.
- [PBH⁺01] G. Penney, J. Blackall, D. Hayashi, T. Sabharwal, A. Adam, and D. Hawkes. Overview of an ultrasound to CT or MR registration system for use in thermal ablation of liver metastases. In *Medical Image Understandings and Analysis*, July 2001.
- [PCA03] X. Pennec, P. Cachier, and N. Ayache. Tracking brain deformations in time-sequences of 3D US images. *Pattern Recognition Letters - Special Issue on Ultrasonic Image Processing and Analysis*, 24(4-5):801–813, February 2003.
- [Pen96] X. Pennec. *L'incertitude dans les problèmes de reconnaissance et de recalage. Application en imagerie médicale et biologie moléculaire*. PhD thesis, École Polytechnique, 1996.
- [PFTV91] W.H. Press, B.P. Flannery, S.A Teukolsky, and W.T. Vetterling. *Numerical Recipes in C*. Cambridge Univ. Press, 1991.
- [PFTV92] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1992.
- [PGT98] X. Pennec, C.R.G. Guttman, and J.-P. Thirion. Feature-based registration of medical images: Estimation and validation of the pose accuracy. In *Proc. of First Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'98)*, volume 1496 of *LNCS*, pages 1107–1114, Cambridge, USA, October 1998. Springer Verlag.
- [PHT95] T. Phong, R. Horaud, and P. Tao. Object pose from 2-d to 3-d points and line correspondences. *International Journal of Computer Vision*, 15:225–243, 1995.
- [Pic01] G. Picinbono. *Modèles géométriques et physiques pour la simulation d'interventions chirurgicales*. Thèse de sciences, université de Nice Sophia-Antipolis, February 2001.
- [Pow64] M. J. D. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal*, 7(2):155–162, 1964.
- [PPR⁺03] S. Poggi, S. Pallotta, S. Russo, P. Gallina, A. Torresin, and M. Bucciolini. Neuronavigation accuracy dependence on CT and MR imaging parameters: a phantom-based study. *Physics in Medicine and Biology*, 48:2199–2216, 2003.
- [PT97] X. Pennec and J.-P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, 25(3):203–229, 1997.

- [PT02] W. Piekarski and B. Thomas. Arquake: The outdoor augmented reality gaming system. *Communications of the ACM*, 45(1):36–38, January 2002.
- [PTE95] B. Peuchot, A. Tanguy, and M. Eude. Virtual reality as an operative tool during scoliosis surgery. In *First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, pages 550–554, Nice, France, April 1995.
- [QL99] L. Quan and Z. Lan. Linear n-point camera pose determination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(07):774–780, July 1999.
- [Ra01] Rosenthal and al. Augmented reality guidance for needle biopsies: A randomized, controlled trial in phantoms. In *Fourth International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'01)*, pages 240–248, 2001.
- [RF00] J. Rolland and H. Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoperators and Virtual Environments*, 9(3):287–309, June 2000.
- [RMOZ01] T. Rohlfing, C. Maurer, W. O'Dell, and J. Zhong. Modeling liver motion and deformation during the respiratory cycle using intensity-based free-form registration of gated mr images. In *Medical Imaging*, volume 4319 of *Proceedings of SPIE*, pages 337–348, 2001.
- [RPMA01] A. Roche, X. Pennec, G. Malandain, and N. Ayache. Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging*, 20(10):1038–1049, October 2001.
- [RS01] G. Reitmayr and D. Schmalstieg. Mobile collaborative augmented reality. In *IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, pages 114–123, New-York, October 2001. IEEE and ACM.
- [RSL⁺02] M. Rosenthal, A. State, J. Lee, G. Hirota, J. Ackerman, K. Keller, E. Pisano, M. Jiroutek, M. Keith, and H. Fuchs. Augmented reality guidance for needle biopsies; an initial randomized, controlled trial in phantoms. *Medical Image Analysis*, 6:313–320, 2002.
- [Sa02] R. Shahidi and al. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Transactions on Medical Imaging*, 21(12):1524–1535, December 2002.
- [SAB02] J. Salvi, X. Armangué, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617–1635, 2002.
- [SAN⁺04] L. Soler, N. Ayache, S. Nicolau, X. Pennec, C. Forest, H. Delingette, and J. Marescaux. Traitement d'images médicales pour la planification, la simulation et l'aide intra-opératoire des actes chirurgicaux. *La Revue de l'Electricité et de l'Electronique*, pages 64–71, janvier 2004.
- [SB98] G. Simon and M.-O. Berger. A two stage robust statistical method for temporal registration from features of various type. In *In Proceedings of 6th International Conference on Computer Vision (ICCV'98)*, pages 261–266, Bombay, India, January 1998.

-
- [SB02a] G. Simon and M.-O. Berger. Pose estimation for planar structures. *IEEE Computer Graphics and Applications*, 22(6):46–53, Nov 2002.
- [SB02b] G. Simon and M.-O. Berger. Reconstructing while registering : a novel approach for markerless augmented reality. In *International Symposium on Mixed and Augmented Reality - ISMAR'02, Darmstadt, Germany*, Sep 2002.
- [SC01] G. Stetten and V. Chib. Magnified real-time tomographic reflection. In W. Niessen and M. Viergever, editors, *Fourth International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'01)*, pages 683–690, Utrecht, October 2001.
- [SCHB01] G. Stetten, V. Chib, D. Hildebrand, and J. Bursee. Real time tomographic reflection: Phantoms for calibration and biopsy. In *IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, pages 11–19, New-York, October 2001.
- [SDM⁺01] L. Soler, H. Delingette, G. Malandain, N. Ayache, and J. Marescaux. Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surgery*, 6(3), August 2001.
- [Ser03] M. Sermesant. *Modèle électromécanique du coeur pour l'analyse d'image et la simulation (Electromechanical Model of the Heart for Image Analysis and Simulation)*. PhD thesis, Université de Nice Sophia Antipolis, May 2003.
- [SFB94] K. R. Smith, K. J. Frank, and R. D. Bucholz. The neurostationTM - a highly accurate, minimally invasive solution to frameless stereotactic neurosurgery. *Computerized Medical Imaging and Graphics*, 18(4):247–256, 1994.
- [SKB⁺01] F. Sauer, A. Khamen, B. Bascle, L. Schimmang, F. Wenzel, and S. Vogt. Augmented reality visualization of ultrasound images: System description, calibration, and features. In *IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, pages 30–39, New-York, October 2001.
- [SKR⁺97] U. Spetzger, G. Krombach, M. Reinges, J. Gilsbach, and T. Schmidt. Navigational microneurosurgery: Experience with the easyguide neuro. *Medicamundi*, 41(1):28–35, March 1997.
- [SLG⁺96] A. State, M. Livingston, W.F. Garrett, G. Hirota, M. Whitton, E. Pisano, and H. Fuchs. Realizing ultrasound-guided needle biopsies. In *In Computer Graphics Proceedings*, pages pp 439–446. ACM SIGGRAPH, August 1996.
- [SLY⁺03] E. Sjolie, T. Lango, B. Ystgaard, G. Tangen, A. Nagelhus Hernes, and R. Marvik. 3d ultrasound-based navigation for radiofrequency thermal ablation in the treatment of liver malignancies. *Surgical Endoscopy*, 17:933–938, 2003.
- [SND01] O. Skrinjar, A. Nabavi, and J. Duncan. A stereo-guided biomechanical model for volumetric deformation analysis. In *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 95–104, December 2001.
- [SPM84] I. Suramo, M. Paivansalo, and V. Myllyla. Cranio-caudal movements of the liver, pancreas and kidneys in respiration. *Acta Radiologica Diagnosis*, 25:129–131, 1984.
- [SSA⁺00] S. Schimizu, H. Shirato, H. Aoyama, S. Hashimoto, T. Nishioka, A. Yamasaki, K. Kagei, and K. Miyasaka. High-speed magnetic resonance imaging for four-dimensional treatment planning of conformal radiotherapy of moving body tumors. *Int. J. Radiation Oncology Biol. Phys.*, 48(2):471–474, 2000.

- [SSC⁺02] G. Stetten, D. Shelton, W. Chang, V. Chib, R. Tamburo, D. Hildebrand, L. Lobes, and J. Sumkin. Towards a clinically useful sonic flashlight. In *IEEE International Symposium on Biomedical Imaging*, pages 417–420, 2002.
- [SSX⁺99] S. Schimizu, H. Shirato, B. Xo, K. Kagei, T. Nishioka, S. Hashimoto, K. Tsuchiya, H. Aoyama, and K. Miyasaka. Four-dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor. *Radiotherapy Oncology*, 50:367–370, 1999.
- [Szc58] W. Szczepanski. *Die Lösungsvorschläge für den räumlichen Rückwärtseinschnitt*, Deutsche Geodätische Kommission, pages 1–144. Reihe C: Dissertationen-Heft Nr. 29 1958.
- [TGN02] M. Tuceryan, Y. Genc, and N. Navab. Single point active alignment method (spaam) for optical see-through hmd calibration for augmented reality. *Presence: Teleoperators and Virtual Environments*, 11(3):259–276, June 2002.
- [THBM97] R.K. Ten Haken, J.M. Belter, and L.H. Marsh. Potential benefits of eliminating planning target volume expansions for patient breathing in the treatment of liver tumors. *Int. J. Radiation Oncology Biol. Phys.*, 38(3):613–617, 1997.
- [Tsa87] 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*, 3(4):323–344, 1987.
- [VG99] H. Villalobos and I. Germano. Clinical evaluation of multimodality registration in frameless stereotaxy. *Computer Aided Surgery*, 4:45–49, 1999.
- [Vio95] P.A. Viola. Alignment by maximization of mutual information. Technical Report 1548, MIT, June 1995.
- [VW97] P. Viola and W.M. Wells. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.
- [WCH92] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, 1992.
- [WH01] H. Worn and H. Hoppe. Augmented reality in the operating theatre of the future. In W. Niessen and M. Viergever, editors, *Fourth International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'01)*, pages 1195–1196, Utrecht, October 2001.
- [WM93] G. Q. Wei and S. D. Ma. A complete two-planes camera calibration method and experimental comparisons. In *Proceedings of the Fourth International Conference on Computer Vision (ICCV 93)*, pages 439–446, Berlin, 1993.
- [WMFM96] M. Y. Wang, C. R. Maurer, J. M. Fitzpatrick, and R. J. Maciunas. An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head. *IEEE Transactions on Biomedical Engineering*, 43(6):627–637, June 1996.
- [WNvdBS⁺01] P.W.A. Willems, H.J. Noordmans, J.W. van der Berkelbach Sprenkel, M.A. Viergever, and C.A.F. Tulleken. An MKM-mounted instrument holder for frameless point-stereotactic procedures: a phantom-based accuracy evaluation. *Journal of neurosurgery*, 95:1067–1074, 2001.
- [WS02] S. Wiedenmaier and A. Stadler. Augmented reality applications for effective manufacturing and service. In *Proceedings of the 6th International Scientific Conference on Work With Display Units*, pages 393–395, 2002.

-
- [WSJ⁺99] J. Wong, M. Sharpe, D. Jaffray, V. Kini, J. Robertson, J. Stromberg, and A. Martinez. The use of active breathing control (abc) to reduce margin for breathing motion. *Int. J. Radiation Oncology Biol. Phys.*, 44(4):911–919, 1999.
- [WTF04] A. Wiles, D. Thompson, and D. Frantz. Accuracy assessment and interpretation for optical tracking systems. In SPIE, editor, *Proceedings of Medical Imaging: Visualization, Image-Guided Procedures, and Display*, volume 5367, February 2004.
- [Yua89] J. Yuan. A general photogrammetric method for determining object position and orientation. *IEEE Transactions on Robotics and Automation*, 5(2):129–142, April 1989.
- [ZFN03] X. Zhang, S. Fronz, and N. Navab. Visual marker detection and decoding in AR systems: A comparative study. In IEEE Computer Society, editor, *IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'02)*, pages 97–106, 2003.
- [Zha94] Z. Zhang. Iterative point matching for registration of free-form curves and surface. *International Journal of Computer Vision*, pages 119–152, 1994.
- [Zha99] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Proceedings of the Ninth International Conference on Computer Vision (ICCV 99)*, pages 666–673, 1999.