

HABILITATION À DIRIGER DES RECHERCHES

présentée devant

L'Université de Rennes 1 **Spécialité : Traitement du signal**

par

Alexandre Krupa

Contributions à l'asservissement visuel échographique

soutenue le 21 décembre 2012 devant le jury composé de

Mme.	Jocelyne	Troccaz	Rapporteur
M.	Philippe	Poignet	Rapporteur
M.	Seth	Hutchinson	Rapporteur
M.	Michel	de Mathelin	Examineur
M.	Nicolas	Andreff	Examineur
M.	Christian	Barillot	Examineur
M.	François	Chaumette	Examineur

Table des matières

Avant-propos et remerciements	7
Notations utilisées	9
1 Introduction	11
1.1 Contributions	14
1.2 Applications médicales	16
1.3 Organisation du document	17
2 L'asservissement visuel échographique	19
2.1 Particularités du capteur et de l'image échographique	20
2.1.1 Différences de modèle entre caméra et sonde échographique	20
2.1.2 Formation de l'image échographique	22
2.1.3 Différents types de sonde	26
2.2 Techniques d'asservissement visuel	29
2.2.1 Une classification selon trois critères	29
2.2.2 Formalisme de la fonction de tâche	30
2.2.3 Matrice d'interaction	31
2.2.4 Commande cinématique	31
2.2.5 Tâches hybrides : le formalisme de la redondance	34
2.2.6 Choix des grandeurs asservies	35
2.2.7 Cas de l'imagerie échographique	37
2.3 Asservissements visuels utilisant la modalité échographique : état de l'art	38
2.3.1 Pour le guidage de la sonde	38

2.3.2	Pour le guidage d'un outil	40
2.4	Positionnement de l'étude	44
2.5	Conclusion	45
3	Approches utilisant des informations géométriques	47
3.1	Approche pour l'information de type « point »	48
3.1.1	Modélisation de l'interaction de la coupe avec une droite 3D	48
3.1.2	Application à une procédure d'étalonnage de la sonde	49
3.2	Approche utilisant le contour de la section observée	52
3.3	Approche basée sur les moments 2D de la section observée	54
3.3.1	Modélisation de la matrice d'interaction	55
3.3.2	Interaction avec un objet de forme ellipsoïdale	56
3.3.3	Interaction avec un objet de forme quelconque	57
3.4	Approche tri-plans utilisant les moments 2D	62
3.4.1	Choix des 6 primitives visuelles et modélisation de l'interaction	62
3.4.2	Application au recalage par asservissement virtuel	65
3.5	Contour actif basé sur les descripteurs de Fourier	68
3.5.1	Formulation générique d'un contour actif paramétrique	69
3.5.2	Formulation basée sur les descripteurs de Fourier	70
3.5.3	Initialisation du contour actif et changement de topologie	72
3.6	Application à la télé-échographie robotisée	75
3.6.1	Tâche de récupération de section	76
3.6.2	Maintien automatique de la visibilité lors de la télé-opération	79
3.7	Conclusion	84
4	Approches utilisant l'information dense	85
4.1	Approche basée sur la corrélation spatiale du « speckle » échographique	86
4.1.1	Estimation du mouvement	87
4.1.2	Asservissement visuel 2D-3D pour une tâche de suivi	91
4.2	Approche utilisant directement l'intensité des pixels	93
4.2.1	Modélisation de l'information visuelle et de son interaction	93
4.2.2	Détermination du gradient 3D	95
4.2.3	Application au positionnement de sondes 2D, 3D et bi-plans	96
4.2.4	Application à la stabilisation de l'image échographique	101
4.2.5	Application au suivi d'un outil	105
4.3	Approche 3D avec prise en compte des déformations des tissus	107

4.4	Conclusion	109
5	Conclusion et perspectives	111
5.1	Contributions : primitives visuelles géométriques	112
5.2	Contributions : information dense	113
5.3	Vers une prise en compte des déformations	114
5.4	Perspectives	114
	Bibliographie	129

Avant-propos et remerciements

Ce document présente un aperçu des travaux de recherche que j'ai menés et encadrés dans l'équipe Lagadic depuis mon recrutement en 2004 en tant que chargé de recherche à l'Inria jusqu'en 2012. Il présente également certains travaux que j'ai conduits lors de ma mise à disposition à The Johns Hopkins University durant l'année 2006.

Je tiens tout d'abord à remercier les membres de mon jury qui ont évalué, critiqué et commenté le travail accompli. J'adresse mes remerciements tout particulièrement à mes rapporteurs Jocelyne Troccaz (directeur de recherche CNRS, TIMC-IMAG, Grenoble), Seth Hutchinson (professeur à University of Illinois, Urbana Champaign), Philippe Poignet (professeur à l'université de Montpellier 2, LIRMM) qui ont pris le temps de lire en profondeur ce manuscrit. Je remercie également Michel de Mathelin (professeur à l'université de Strasbourg, ICUBE) et Nicolas Andreff (professeur à l'université de Franche-Comté, FEMTO-ST) d'avoir participé à mon jury de soutenance. Je tiens également à remercier Christian Barillot (directeur de recherche CNRS, équipe Visages de Inria/Irisa, Rennes) qui a présidé mon jury.

Je remercie sincèrement François Chaumette de m'avoir accueilli en 2004 dans son équipe pour démarrer une nouvelle activité de recherche sur la robotique médicale guidée par l'imagerie échographique. Je lui suis reconnaissant de m'avoir confié la conduite de cette thématique si passionnante et si innovante.

Je remercie également très amicalement, l'ensemble de mes collègues de travail et plus particulièrement Fabien Spindler qui m'a démontré expérimentalement qu'il était possible de manger une excellente choucroute Alsacienne en Bretagne. Un grand merci également à Eric Marchand pour les discussions enrichissantes que nous avons eues continuellement depuis ces 8 dernières années lors du café du matin ou des déjeuners le midi. Je souhaite également remercier Marie Babel et Céline Gharsalli pour leur bonne humeur et les moments de rires que nous partageons au quotidien. Merci également à Maud Marchal et Christophe Collewet pour les collaborations menées sur Rennes et à mes collègues hors Bretagne des projets ANR USComp et Prosit.

Bien évidemment, les résultats des travaux de recherche réalisés ne pourraient être obtenus

sans l'implication des 3 thésards que j'ai encadrés durant cette période. Je remercie donc très chaleureusement Rafik Mebarki, Caroline Nadeau et Tao Li pour les travaux qu'ils ont accomplis. Je remercie également Deukhee Lee d'avoir choisi d'effectuer un séjour post-doctoral à mes côtés. Je n'oublie pas les différents stagiaires que j'ai encadrés lors de cette période et qui ont débroussaillé un grand nombre de sujets.

Enfin, je termine ces remerciements en exprimant toute ma reconnaissance à ma famille et dédie cette habilitation à mes enfants et mon épouse.

Notations utilisées

Notations générales

- a, λ : scalaire noté en minuscule.
- \mathbf{v} : vecteur noté en minuscule gras, v_i désigne la i -ème composante du vecteur \mathbf{v} .
- \mathbf{M} : matrice notée en majuscule gras, M_{ij} désigne l'élément situé à l'intersection de la i -ème ligne et de la j -ème colonne de la matrice \mathbf{M} .
- $[\mathbf{v}]_{\times}$ matrice de préproduit vectoriel associée au vecteur \mathbf{v} .
- \mathbf{M}^+ pseudo-inverse de la matrice \mathbf{M} .
- L'estimation d'une grandeur (scalaire ou matricielle) est indiquée grâce au symbole $\hat{\cdot}$, par exemple \hat{a} indique une estimation de la grandeur a .

Changements de repère – transformations – coordonnées

- ${}^a\mathbf{M}_b$: matrice homogène définie par

$${}^a\mathbf{M}_b = \begin{pmatrix} {}^a\mathbf{R}_b & {}^a\mathbf{t}_b \\ \mathbf{0}_{1 \times 3} & 1 \end{pmatrix}$$

où ${}^a\mathbf{R}_b$ et ${}^a\mathbf{t}_b$ sont respectivement la matrice de rotation et le vecteur de translation exprimant la transformation entre les repères \mathcal{R}_a et \mathcal{R}_b . ${}^a\mathbf{M}_b$ décrit la localisation du repère \mathcal{R}_b exprimé dans le repère \mathcal{R}_a .

- ${}^s\mathbf{M}_o$: cas particulier de la matrice de pose ${}^s\mathbf{M}_o$ qui décrit la localisation du repère \mathcal{R}_o d'un objet dans le repère de la sonde \mathcal{R}_s .
- ${}^b\mathbf{x}_A = ({}^bx_A, {}^by_A, {}^bz_A)$: coordonnées 3D du point nommé A exprimées dans le repère \mathcal{R}_b .

Asservissement visuel

- \mathbf{v} : vitesse de la sonde exprimée dans son repère, $\mathbf{v} = (\mathbf{v}, \boldsymbol{\omega})$ où $\mathbf{v} = (v_x, v_y, v_z)$ est la vitesse de translation $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ est la vitesse de rotation.
- \mathbf{s} : vecteur des informations visuelles impliquées dans la loi de commande.

- \mathbf{L}_s : matrice d'interaction associée à s .
- Les grandeurs calculées en position désirée sont notées à l'aide du symbole $*$, par exemple : s^* ou \mathbf{L}_s^* .

CHAPITRE 1

Introduction

Apparue au milieu des années 1980, la robotique médicale et chirurgicale n'a cessé depuis de progresser. En effet, grâce à l'introduction de nouvelles technologies de l'information et de la communication dans la médecine, une nouvelle discipline est apparue, laquelle s'intéresse aux Gestes Médicaux et Chirurgicaux Assistés par Ordinateur (GMCAO). Ainsi, les techniques d'imagerie médicale informatisées permettent aujourd'hui d'assister le praticien de façon quantitative et fiable. Historiquement, les premiers systèmes robotiques ont été introduits en neurochirurgie [Benabid 1987]. L'opération consiste à insérer une aiguille de biopsie pour atteindre une zone profonde du cerveau ou encore à introduire un guide afin de déposer des électrodes de stimulation dans le but de traiter des pathologies telle la maladie de Parkinson. Les robots sont également apparus en chirurgie orthopédique pour réaliser, de manière autonome à partir d'un planning pré-opératoire, l'usinage de la cavité osseuse nécessaire au placement de prothèses de hanche [Taylor 1994].

Par la suite, l'utilisation de systèmes robotiques a permis d'effectuer des opérations chirurgicales peu invasives pour le patient, voire même des interventions ne pouvant pas être réalisées par une technique manuelle. A leur début, ces systèmes ont principalement été télé-opérés par le praticien. Nous pouvons mentionner par exemple les systèmes utilisés en chirurgie laparoscopique [Guthart 2000] et microchirurgie [Mitsubishi 2012] où le chirurgien commande à distance, par l'intermédiaire d'une console, un manipulateur esclave motorisé qui actionne les déplacements des instruments chirurgicaux. Ces systèmes permettent d'augmenter la dextérité du geste chirurgical par l'ajout de mobilités supplémentaires aux outils et par la démultiplication et/ou le filtrage des mouvements des mains du chirurgien. La télé-opération est également utilisée pour réaliser des examens à distance. Il existe par exemple des robots légers télé-opérés porteurs de sonde échographique qui se posent sur l'abdomen du patient [Vilchis 2003, Delgorgue 2005]. Ces systèmes permettent d'examiner un patient par un expert d'une pathologie non présent sur le site du patient.

Depuis une dizaine d'années, l'émergence des techniques poussées d'imagerie a permis d'envisager de nouvelles applications robotiques d'assistance aux gestes médicaux. Il est à présent possible d'utiliser l'information provenant d'une ou de plusieurs modalités d'imagerie telles l'endoscopie, l'échographie, l'imagerie à rayonnement X, ou l'imagerie à résonance magnétique IRM pour guider des robots médicaux.

Actuellement, une grande majorité des interventions réalisées sous imagerie médicale repose sur l'utilisation d'un système de navigation pour localiser à l'aide d'un système optique ou magnétique la position d'un instrument chirurgical dans le référentiel de l'imageur. Le médecin peut alors s'aider de cette information de pose pour guider l'instrument vers une cible définie dans l'image. La cible peut être soit directement visible dans l'image per-opératoire ou projetée dans celle-ci après avoir été définie par le médecin dans une image pré-opératoire. Dans ce dernier cas, une étape de recalage est nécessaire pour convertir les coordonnées de la cible, initialement fixées dans l'image pré-opératoire, dans le référentiel associé à l'image per-opératoire. Généralement, ce recalage est réalisé au moyen d'un algorithme d'optimisation qui met en correspondance des marqueurs artificiels ou anatomiques observés dans les deux images (pré et per-opératoires) [Maintz 1998, Pluim 2003, Wyawahare 2009]. La précision du geste est donc directement fonction d'une part de la précision du système utilisé pour localiser l'outil et d'autre part de la précision de la procédure de recalage. L'erreur de positionnement est également plus grande lorsque la cible réelle à atteindre est mobile, ce qui est généralement le cas en raison des perturbations induites par les mouvements physiologiques du patient. Il est à noter que la compensation des mouvements physiologiques reste actuellement un défi important à relever. Ces systèmes de navigation sont couramment utilisés en radiologie interventionnelle pour assister le médecin lors de gestes d'insertion d'aiguille dans un organe afin de prélever des tissus ou traiter localement des tumeurs par radiothérapie ou cryoablation. Cette dernière technique consiste à injecter un gaz réfrigérant à l'endroit où se trouve la tumeur afin de tuer les tissus cancéreux par l'application d'une température extrêmement basse. L'insertion d'aiguille est pour l'instant réalisée manuellement par le chirurgien à l'aide d'un système de navigation mais plusieurs robots prototypes ont déjà été proposés par des équipes de recherche pour réaliser la tâche avec plus de précision [Maurin 2005, Bricault 2008, Hungr 2009].

La commande par vision ou asservissement visuel est une approche alternative aux techniques de recalage classiques. Cette approche consiste à commander, en boucle fermée, les mouvements d'un système dynamique, généralement un robot, à partir d'informations visuelles extraites de l'image fournie par un capteur embarqué ou observant le système. Plusieurs équipes de recherche ont proposé son utilisation dans des applications médicales. Elle permet de commander en boucle fermée le déplacement de l'imageur ou celui d'un instrument directement à partir d'informations visuelles extraites de l'image médicale per-opératoire. Le principe de cette commande est d'adapter le déplacement du robot de manière à minimiser une erreur visuelle qui est directement mesurée à chaque instant dans l'image per-opératoire. Cette erreur peut par exemple correspondre à la distance entre un instrument et une cible à atteindre qui sont tous deux détectés par un algorithme de traitement d'images. Néanmoins pour sa mise en œuvre, il est nécessaire que les images per-opératoires soient rafraîchies et traitées à une cadence élevée afin de garantir la contrainte du temps réel nécessaire à la commande d'un robot. Le respect de cette contrainte est primordial et constitue une difficulté supplémentaire au développement d'un asservissement visuel. Pour cette raison, principalement deux types d'imagerie ont été considérés dans la littérature : l'imagerie optique utilisant des capteurs vidéo (caméras)

comme par exemple l'endoscopie, la fibroscopie, la microscopie et l'imagerie échographique. En effet, toutes ces modalités ont le principal avantage de fournir des images à la cadence vidéo de 25 images/seconde ou plus.

Historiquement, les premiers travaux mettant en œuvre des asservissements visuels dans le contexte de la robotique médicale ont considéré la modalité optique. Ils ont porté sur le positionnement automatique de robots porteurs d'endoscope afin de suivre et maintenir dans le champ visuel de la caméra un ou plusieurs instruments de chirurgie laparoscopique lors d'une intervention chirurgicale [Uecker 1995, Casal 1996, Wei 1997]. Ils ont également concerné le positionnement automatique d'un instrument de chirurgie laparoscopique comme cela a été le cas dans mes travaux de thèse [Krupa 2003b]. Par la suite, la compensation active du mouvement cardiaque par asservissement visuel est devenue un challenge que plusieurs équipes de recherche se sont proposées de relever afin de permettre à un chirurgien d'opérer sur un cœur battant virtuellement stabilisé [Nakamura 2001, Ginhoux 2004]. Pour avoir une vue plus approfondie sur ces travaux en robotique médicale utilisant la vision optique directement dans la commande du système, j'invite le lecteur à lire l'article d'état de l'art [Krupa 2009a] que j'ai présenté à l'édition 2009 des Journées Nationales de la Recherche en Robotique (JNRR'09).

Parmi les autres modalités d'imagerie médicale, l'imagerie ultrasonore 2D possède trois principaux atouts. Elle permet une imagerie anatomique non invasive, en temps réel et à faible coût. Néanmoins, le rapport signal sur bruit des images échographiques est faible et rend difficile leur interprétation et traitement. Le premier asservissement visuel utilisant cette modalité a été proposé aux alentours de l'an 2000 pour contrôler certains mouvements de la sonde lors d'une application de diagnostic de l'artère carotide [Abolmaesumi 2001]. Dans cette étude pionnière, la commande visuelle est utilisée pour centrer automatiquement la section de l'artère dans l'image échographique 2D, permettant ainsi de maintenir l'artère visible tout au long de l'examen. Néanmoins, l'approche proposée permet uniquement de contrôler les mouvements de la sonde dans le plan d'observation du capteur (deux translations et une rotation) et aucune solution n'a été fournie pour contrôler également les déplacements en dehors du plan d'observation. La raison de cette limitation est liée à une modélisation incomplète du capteur échographique qui agit de manière totalement différente en comparaison avec une caméra. Cette dernière fournit une projection de l'environnement 3D perçu vers une image 2D, alors qu'une sonde échographique 2D a la particularité de fournir une information complète dans le plan d'observation du capteur, mais de ne fournir aucune information en dehors de ce plan. Il est par conséquent impossible d'appliquer des méthodes d'asservissement visuel reposant sur la modélisation de l'interaction d'une caméra avec l'environnement observé pour contrôler les 6 degrés de liberté (ddl) d'une sonde échographique. C'est pour cette raison que l'asservissement visuel échographique a été dans un premier temps limité à des applications médicales où uniquement les mouvements dans le plan d'observation sont contrôlés par la vision, que ce soit pour la commande de la sonde [Abolmaesumi 2002] ou d'un instrument chirurgical telle une aiguille de biopsie [Hong 2004, Neubach 2010].

1.1 Contributions

Suite à mes travaux de thèse qui ont porté sur la mise en œuvre d'un asservissement visuel utilisant des images endoscopiques [Krupa 2003a] et à mon recrutement à l'INRIA en septembre 2004, j'ai démarré une nouvelle thématique de recherche, inexistante alors dans l'équipe, portant sur l'étude et l'élaboration de méthodes d'asservissement visuel utilisant les images échographiques. A cette époque, très peu de travaux dans la littérature avaient été menés dans ce domaine et les méthodes existantes se limitaient, comme je l'ai mentionné précédemment, à contrôler uniquement les mouvements de la sonde dans le plan de coupe. Par conséquent, j'ai orienté mon activité de recherche en vue de fournir des solutions capables de contrôler le mouvement complet d'une sonde échographique embarquée sur un robot médical.

A cette fin, un certain nombre de verrous scientifiques majeurs doivent être levés afin de garantir la généralité, la robustesse et la fiabilité d'un asservissement visuel échographique. D'une part, il s'agit de déterminer les informations visuelles robustes qui sont pertinentes à la réalisation d'une tâche de positionnement automatique de la sonde. D'autre part, une formulation complète de l'interaction entre la sonde et son environnement doit être entreprise pour gérer la totalité des degrés de liberté, c'est-à-dire également les déplacements effectués en dehors du plan de coupe. Une problématique concerne également l'extraction de l'information visuelle utile à la commande du système à partir de l'observation des tissus organiques dans les images échographiques qui sont par nature fortement bruitées et faiblement structurées. Actuellement, la segmentation du contenu de ces images est toujours une thématique de recherche à part entière. La difficulté est encore plus prononcée lorsqu'il s'agit d'extraire des informations pertinentes en temps réel. Contrairement à une application de robotique manufacturière, s'ajoute également la complexité non maîtrisée de l'environnement perçu par la sonde, comme les déformations internes induites par le mouvement physiologique du patient.

Pour relever ces verrous et répondre à nos objectifs nous avons proposé diverses approches pour contrôler le mouvement complet de la sonde directement à partir d'informations issues de l'image échographique. Ces approches se distinguent par la nature des informations visuelles considérées, et forment ainsi deux catégories de solutions. La première catégorie concerne les informations visuelles géométriques tels des points, des contours et autres représentations géométriques de sections pouvant être extraites au moyen d'une segmentation de l'image. La seconde catégorie considère quant à elle directement l'information dense contenue dans l'image.

Primitives visuelles géométriques. Les approches utilisant des primitives visuelles géométriques sont basées sur la modélisation de l'interaction entre l'objet observé et le plan de coupe de la sonde 2D. Cette modélisation consiste alors à déterminer la relation analytique qui relie, d'une part, la variation d'un ensemble d'informations visuelles géométriques 2D qui sont pertinentes à la réalisation d'une tâche robotique donnée, aux déplacements de la sonde échographique d'autre part. Afin de tenir compte de la variation des informations visuelles introduite par les mouvements effectués en dehors du plan de coupe, nous avons défini un modèle simplifié ou estimé en ligne de la surface de l'objet observé.

Cette modélisation de l'interaction permet ensuite de déterminer la loi de commande à appliquer au robot porteur de sonde afin d'atteindre la coupe échographique présentant l'information visuelle désirée, ou pour réaliser la poursuite d'une coupe de référence. Cette étude a été

conduite pour différents types d'information visuelle :

- Information visuelle de type « point ». J'ai considéré l'information visuelle de type point correspondant à l'intersection d'une droite 3D traversant le plan de coupe [Krupa 2006b]. Du point de vue applicatif, ce travail a permis d'automatiser le positionnement d'une sonde par rapport à un objet en forme de croix lors d'une procédure d'étalonnage. Cette dernière a été mise en œuvre pour estimer les paramètres intrinsèques et extrinsèques du capteur échographique, ceci en vue de permettre l'imagerie échographique 3D par un système constitué d'un robot médical actionnant une sonde échographique 2D [Krupa 2006a].
- Information visuelle de type « contour ». Afin de positionner la sonde par rapport à un élément anatomique, l'information visuelle correspondant au contour 2D de la section de l'élément observé a été considérée [Bachta 2006]. L'asservissement visuel consiste alors à commander la sonde de manière à atteindre le contour 2D désiré d'un organe d'intérêt. Dans cette étude, nous avons proposé une méthode d'asservissement visuel basée sur une formulation implicite du contour 2D d'un élément anatomique qui est observé dans la coupe et dont la géométrie 3D est a priori connue au moyen d'une imagerie pré-opératoire.
- Information visuelle de type « moments ». Dans le cadre de la thèse de Rafik Mebarki [Mebarki 2010a], nous avons amélioré les performances de cette tâche de positionnement de la sonde en utilisant les moments de la section observée en tant qu'informations visuelles. Ce choix a permis d'augmenter la robustesse de la commande vis-à-vis des erreurs de mesure dans l'image et d'agrandir le domaine de stabilité. Cette méthode a été étendue pour s'affranchir de la connaissance d'un modèle pré-opératoire de l'organe considéré [Mebarki 2010b]. Son principe repose sur une estimation du vecteur normal à la surface de l'organe pour chaque point du contour de la section observée.

Les approches décrites ci-dessus reposent ainsi toutes sur l'utilisation d'informations visuelles géométriques tels des points, des contours ou les moments d'une section d'un organe d'intérêt qui sont extraites de l'image par des techniques de segmentation. Une méthode d'extraction de contour temps réel, basée sur une approche paramétrique intégrant des descripteurs de Fourier [Li 2011a], a ainsi été développée et mise en œuvre lors des travaux de thèse de Tao Li [Li 2012a] pour extraire ces informations de l'image.

Néanmoins, les approches utilisant des primitives visuelles géométriques sont essentiellement appropriées à des applications où les contours des objets d'intérêt, tels des organes, des artères, des tumeurs ou des kystes, sont décelables dans l'image échographique. Par conséquent, afin de considérer des images très faiblement structurées, j'ai également travaillé à l'élaboration de méthodes utilisant une seconde catégorie d'information en entrée de la commande.

Information dense. L'information dense de l'image échographique peut être directement utilisée. Ce type d'approche permet en outre de s'affranchir de l'étape de segmentation coûteuse en temps de calcul.

- Lors de ma mise à disposition à “The Johns Hopkins University” à Baltimore, j'ai démarré une étude qui utilise directement l'information de corrélation de la texture de type « speckle » présente dans l'image échographique pour réaliser la poursuite d'un organe en mouvement par une sonde échographique robotisée selon les 6 degrés de liberté [Krupa 2009b]. La méthode consiste à minimiser la position relative entre une coupe de référence apprise à un instant donné et se déplaçant avec l'objet d'intérêt, et la coupe actuellement observée par la sonde. Une technique de décorrélation du « speckle » est utilisée pour obtenir une estimation de cette position relative qui est immédiatement minimisée par un asservissement visuel hybride 2D/3D.
- Dans le cadre de la thèse de Caroline Nadeau [Nadeau 2011a], un asservissement visuel 2D, essentiellement basé sur l'intensité des pixels, a été proposé. Les informations visuelles choisies correspondent dans ce cas directement à la valeur d'intensité d'un ensemble de pixels de l'image. La variation de ces informations en fonction du mouvement de la sonde a été modélisée afin de mettre en œuvre des asservissements visuels capables de réaliser des tâches de positionnement ou de suivi de coupes anatomiques [Nadeau 2011d]. Cette méthode a été adaptée pour différents types de capteurs ultrasonores (2D, 3D, bi-plans).
- Nous avons également proposé une méthode qui permet d'estimer en temps réel les déformations des tissus mous dans une séquence de volumes échographiques denses [Lee 2011]. L'approche proposée consiste à estimer, à partir des différences d'intensité des voxels mesurées entre volumes successifs, les déformations d'une région d'intérêt par une grille 3D déformable. La compensation active du mouvement est ensuite réalisée par un asservissement visuel 3D dont l'objectif est de déplacer le capteur échographique 3D de manière à minimiser le mouvement rigide estimé.

1.2 Applications médicales

Les applications qui découlent de mes travaux se situent principalement dans le domaine de l'assistance à l'examen échographique. Elles concernent d'une part le positionnement automatique, par asservissement visuel échographique, de la sonde sur une section anatomique désirée afin de permettre au médecin de naviguer automatiquement parmi un ensemble de coupes échographiques apprises au préalable. Cette fonctionnalité de récupération de section ainsi qu'une assistance de maintien de visibilité [Li 2012b] par asservissement visuel échographique ont été intégrées, dans le cadre de la thèse de Tao Li, sur le robot de télé-échographie du projet ANR PROSIT 2008-2012. Ce projet collaboratif porté par le PRISME inclut l'INRIA dont je suis le responsable scientifique et les partenaires Robosoft, Pprime et le CHU de Tours.

D'autre part, les applications que nous avons traitées portent sur la stabilisation active de l'image échographique à des fins de diagnostic. En effet, dans le cadre du projet ANR USComp 2008-2012 dont je suis le porteur, et qui regroupe des collègues du LIRMM et du LSIT, nous

avons proposé différentes approches permettant de compenser le mouvement de tissus mous par asservissement visuel échographique en synchronisant efficacement les déplacements du mouvement de la sonde.

1.3 Organisation du document

Les méthodes d'asservissement visuel échographique que je propose se classent selon deux catégories d'informations visuelles : géométrique ou dense. Les contributions décrites ci-après sont donc organisées selon ces deux types d'informations.

Ce mémoire se structure en quatre chapitres principaux :

- Le chapitre 2 présente les particularités du capteur et de l'image échographique et rappelle les concepts fondamentaux de l'asservissement visuel. Il expose ensuite un état de l'art complet sur les travaux réalisés en asservissement visuel échographique.
- Le chapitre 3 présente mes contributions à l'asservissement visuel échographique utilisant l'information visuelle géométrique.
- Les approches utilisant directement l'information dense pour commander les mouvements de la sonde et permettant ainsi de s'affranchir de la segmentation de l'image échographique sont présentées dans le chapitre 4.

Ce mémoire s'achève par un chapitre 5 de conclusion dans lequel sont également évoquées mes perspectives de recherche.

Notons que les références qui correspondent à mes travaux, et à ceux des personnes que j'ai encadrées, apparaissent en italique dans le texte, contrairement aux autres références.

L'asservissement visuel échographique

La plupart des travaux portant sur l'asservissement visuel ont été conduits dans le cadre de l'imagerie optique. L'image est généralement fournie par une caméra vidéo monoculaire qui perçoit les réflexions de la lumière sur les objets constituant la scène. Elle correspond dans ce cas à une projection perspective de l'environnement 3D sur un plan image matérialisé par le capteur électronique de la caméra. A leur origine, les concepts de l'asservissement visuel ont été développés pour l'usage d'une caméra monoculaire classique. Ils ont été étendus par la suite pour d'autres types de capteurs optiques telles les caméras stéréoscopiques fournissant deux vues différentes de la scène ou les caméras omnidirectionnelles fournissant une projection sphérique de l'environnement sur le plan du capteur. Plus récemment, les caméras RGB-D, comme la Kinect de Microsoft, font l'objet d'études en vue d'être intégrées à des systèmes utilisant l'asservissement visuel.

L'imagerie échographique a été pour sa part encore peu étudiée à des fins de contrôle de robots directement par la vision. Dans le but de mettre en œuvre des systèmes robotiques commandés par la vision échographique, il s'agit donc de vérifier dans quelle mesure les concepts de l'asservissement visuel restent valides dans le cas de l'imagerie échographique.

Dans ce chapitre, nous présenterons dans un premier temps les particularités du capteur échographique en soulignant ses différences par rapport à une caméra et introduirons brièvement les principes physiques de la formation de l'image échographique (section 2.1). Nous rappellerons ensuite les notions de base de l'asservissement visuel en détaillant le formalisme de la fonction de tâche qui est actuellement très largement utilisé dans la communauté robotique pour la mise en œuvre de la commande de systèmes par la vision (section 2.2). Nous présenterons ensuite l'état de l'art en asservissement visuel échographique en mettant en évidence les différents défis à relever (section 2.3) et positionnerons notre étude dans ce contexte (section 2.4).

2.1 Particularités du capteur et de l'image échographique

Comme nous l'avons évoqué dans l'introduction de ce document, nos travaux portent sur l'étude de méthodes d'asservissement visuel utilisant la modalité échographique. Le capteur utilisé est dans ce cas une sonde ultrasonore et non plus une caméra. Nous nous attachons ici à décrire les spécificités du capteur échographique et de l'image résultante.

2.1.1 Différences de modèle entre caméra et sonde échographique

Caméra et sonde échographique ne fournissent pas une même représentation de l'environnement. En effet, une caméra apporte une information visuelle correspondant à la projection perspective de la scène observée sur le plan image, comme illustrée par la figure 2.1. A la différence, une sonde échographique 2D fournit une image représentant une coupe de l'environnement sans procurer d'information en dehors de son plan d'observation (figure 2.2). Il est par conséquent inapproprié d'appliquer directement les techniques d'asservissement visuel fondées sur l'utilisation d'une caméra représentée par le modèle du sténopé qui n'est plus valide dans le cas d'une sonde.

Les différences se situent principalement sur la modélisation de l'interaction du capteur avec son environnement.

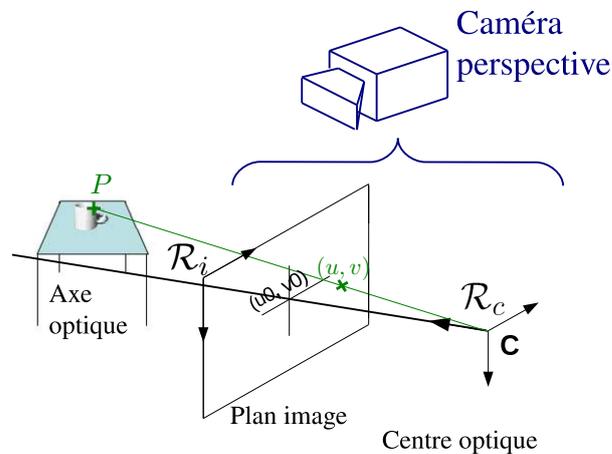


FIGURE 2.1 – Formation de l'image par projection perspective avec une caméra classique.

Afin de souligner les spécificités géométriques d'une sonde échographique, il est nécessaire de rappeler tout d'abord le modèle géométrique associé à la projection perspective qui fournit les coordonnées en pixel (u, v) de l'image d'un point physique P de coordonnées 3D, ${}^c\mathbf{x} = ({}^c x, {}^c y, {}^c z)$, exprimées dans le repère de la caméra \mathcal{R}_c . Soient f la distance focale de la caméra, (u_0, v_0) les coordonnées en pixel de la projection du centre optique de la caméra sur

son plan image et (k_u, k_v) les facteurs d'échelle de l'image. Le modèle s'écrit

$$\begin{cases} u = u_0 + f k_u \frac{c_x}{c_z}, \\ v = v_0 + f k_v \frac{c_y}{c_z}. \end{cases} \quad (2.1)$$

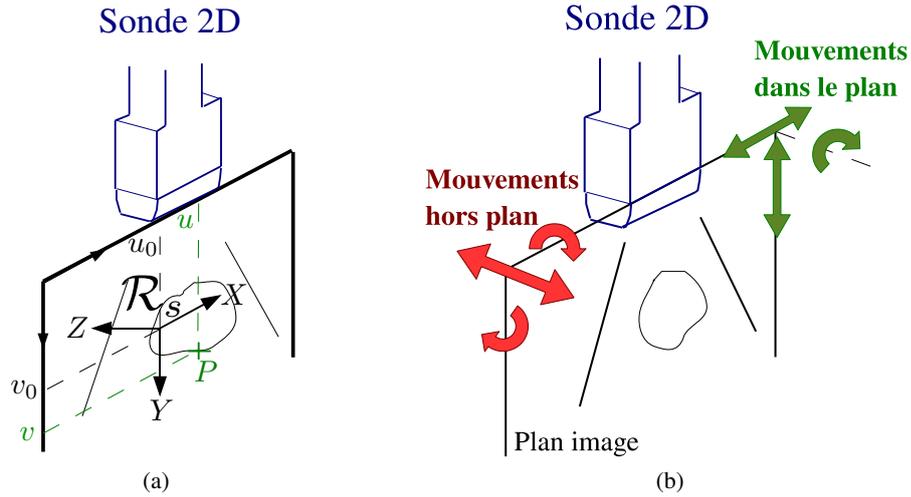


FIGURE 2.2 – (a) Plan de coupe observé par une sonde échographique et repère cartésien \mathcal{R}_s associé à la sonde. (b) Mouvements dans le plan (2 translations, 1 rotation) et mouvements hors du plan (1 translation, 2 rotations) d'observation d'une sonde 2D.

Si nous considérons à présent le cas d'une sonde échographique 2D, seuls les points physiques de la scène appartenant au plan de coupe sont visibles dans l'image. Le modèle géométrique, reliant les coordonnées pixelliques aux coordonnées 3D, ${}^s\mathbf{x} = ({}^s x, {}^s y, {}^s z)$, d'un point P de la coupe exprimées dans le repère de la sonde \mathcal{R}_s avec ${}^s z = 0$, s'écrit alors

$$\begin{cases} u = u_0 + k_u {}^s x, \\ v = v_0 + k_v {}^s y, \end{cases} \quad (2.2)$$

où (u_0, v_0) sont les coordonnées pixelliques du centre de l'image sur lequel nous avons fixé l'origine du repère de la sonde (voir figure 2.2a), et (k_u, k_v) sont les facteurs d'échelle de l'image échographique nécessaires à la conversion des coordonnées métriques en coordonnées pixelliques. Réciproquement on introduit les facteurs d'échelle (l_x, l_y) qui convertissent les coordonnées pixelliques en coordonnées métriques. Ces facteurs représentent les paramètres intrinsèques de la sonde. Ils permettent de déterminer, pour tout point de l'image échographique, les coordonnées 3D dans le repère de la sonde du point physique correspondant. Afin d'exprimer ces coordonnées 3D dans un autre repère de référence, il faut également connaître les paramètres extrinsèques de la sonde qui caractérisent la pose du repère de la sonde par rapport au repère de référence. Une connaissance correcte de ces paramètres est donc essentielle pour réaliser une imagerie 3D à l'aide d'une sonde 2D. Ces paramètres peuvent être estimés à l'aide d'une procédure de calibration. Nous y reviendrons dans la section 3.1 où une solution

d'automatisation d'une telle procédure au moyen d'un asservissement visuel échographique sera présentée.

De ce fait, l'image échographique a la particularité de fournir une information complète dans le plan de coupe. Elle donne un accès direct aux coordonnées 3D des points physiques situés dans le plan de la sonde, et notamment l'information de profondeur, contrairement à l'image obtenue par une caméra. Par contre, comme nous l'avons indiqué, elle ne fournit aucune information en dehors de ce plan. Cette particularité permet de mettre en évidence une différence importante entre une sonde et une caméra. Quel que soit le mouvement appliqué à une caméra, grâce à la projection perspective il est toujours possible d'observer les mêmes points physiques de l'objet considéré sous deux angles de vue légèrement différents dans la mesure où l'objet d'intérêt ne sort pas du champ visuel. Au contraire, les points observés dans une image échographique 2D ne correspondent plus aux mêmes points 3D physiques de la scène si un déplacement de la sonde est réalisé hors de son plan de coupe. Nous pouvons donc faire la distinction entre deux types de mouvements (figure 2.2b) :

- les mouvements dans le plan qui correspondent aux deux translations selon les axes X et Y du repère de la sonde et à la rotation autour de l'axe Z ;
- les mouvements hors du plan qui se définissent par la translation selon la direction d'élévation de la sonde (portée par l'axe Z) et les deux rotations restantes autour des axes X et Y .

Ainsi, si l'objectif de l'asservissement visuel échographique est de contrôler exclusivement les mouvements dans le plan, il est possible d'adapter sans trop de difficultés les approches classiques de l'asservissement visuel 2D qui sont basées sur l'utilisation de primitives visuelles simples de type points ou orientation principale d'une section. En revanche, si l'objectif est de considérer également le contrôle d'un ou de plusieurs mouvements hors plan, voire de la totalité des degrés de liberté, l'absence d'information en dehors du plan d'observation de la sonde fait du contrôle des mouvements hors plan un défi majeur de l'asservissement visuel basé sur les images échographiques. C'est justement ce défi que nous avons proposé de relever dans nos travaux de recherche.

2.1.2 Formation de l'image échographique

Le principe de l'imagerie échographique consiste à émettre une onde acoustique dans le spectre des ultrasons (20 KHz à 500 Mhz) et réceptionner l'écho renvoyé par la matière qu'elle traverse. Historiquement, l'existence des ondes ultrasonores (US) a été mise en évidence en 1794 par Spallanzi lors de l'observation du vol des chauves-souris. Elles ont été produites artificiellement pour la première fois en 1880, lorsque Pierre Curie découvrit le principe de la piézo-électricité. Néanmoins, leur première application en médecine ne remonte qu'aux années 1940 où Dussik les utilisa dans le cadre d'un examen médical.

Un élément principal d'un dispositif d'imagerie échographique est la sonde ultrasonore. Cette dernière est constituée d'éléments céramiques piézo-électriques (PZT) qui sont excités par des impulsions électriques de fréquence proche de la fréquence de résonance des céramiques. Il en résulte une vibration mécanique des éléments qui produit une onde ultrasonore. Ces éléments, également appelés transducteurs, génèrent l'onde ultrasonore d'émission puis sont également utilisés en tant que récepteurs en convertissant l'onde réfléchie par la matière

TABLE 2.1 – Caractéristiques de quelques milieux du corps humain traversés par les ondes ultrasonores

Tissus	Impédance acoustique (kg/m ² /s)	Vitesse de propagation (m/s)
Sang	1.6610 ⁶	1566
Foie	1.6610 ⁶	1566
Rein	1.6210 ⁶	1567
Graisse	1.3310 ⁶	1446
Os	3.75 – 7.3810 ⁶	2070 - 5350
Air	0.410 ³	333

en signal électrique.

Interaction des ultrasons avec la matière. La vitesse de propagation c de l'onde US dépend de la masse volumique ρ et de l'impédance acoustique Z du milieu traversé, avec

$$c = \frac{Z}{\rho}. \quad (2.3)$$

A titre d'exemple, le tableau 2.1 présente les caractéristiques de quelques milieux organiques en termes d'impédance acoustique et de vitesse de propagation de l'onde ultrasonore. Lors de la propagation de l'onde, son intensité I est soumise à un phénomène d'atténuation qui est dû à l'absorption d'une partie de son énergie par le milieu. Dans le cas d'un milieu homogène, cette atténuation est fonction d'un coefficient d'absorption noté μ et de la distance d du chemin parcouru par l'onde depuis la position où elle avait une intensité initiale I_0 , à savoir

$$I(d) = I_0 e^{-\mu d}. \quad (2.4)$$

Ce coefficient d'absorption $\mu = K f^2$ dépend d'une constante K propre au milieu et est proportionnel au carré de la fréquence f d'émission de l'onde. Le réglage par l'utilisateur d'une fréquence peu élevée permet donc d'imager des structures anatomiques profondes qui sont présentes par exemple dans l'abdomen du patient. Cependant, la résolution de l'image échographique dépend de la longueur d'onde et une fréquence basse s'accompagne alors d'une dégradation de la qualité de l'image.

Lorsque l'onde incidente d'amplitude I_i rencontre, en formant un angle droit, une interface entre deux milieux ayant respectivement des impédances acoustiques Z_1 et Z_2 , un écho de l'onde d'intensité initiale I_r est renvoyé vers le capteur selon

$$I_r = R I_i, \text{ avec } R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2. \quad (2.5)$$

Si le coefficient de réflexion R est inférieur à 10%, l'intensité de l'onde transmise est alors suffisante pour imager les structures plus profondes. Par contre, si la valeur du coefficient de réflexion atteint 40% comme dans le cas d'une interface tissus mous/os ou 99% pour une interface tissus mous/air, il n'est alors plus possible d'explorer les structures situées au-delà de cette interface.

Dans le cas où l'onde incidente n'est pas perpendiculaire à l'interface, elle est partiellement transmise et réfléchié selon les lois de l'optique géométrique avec un angle de réflexion θ_r égal à l'opposé de l'angle d'incidence θ_i (figure 2.3). L'angle du faisceau transmis θ_t est alors

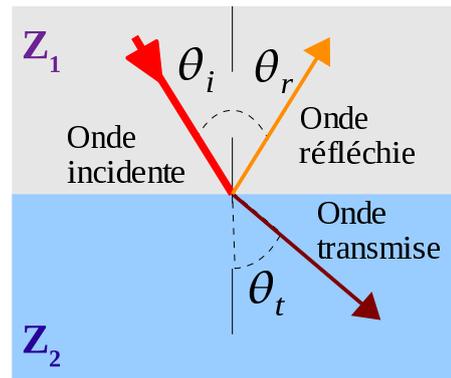


FIGURE 2.3 – Transmission et réflexion du faisceau ultrasonore à l'interface entre deux milieux d'impédance acoustique différente.

fonction de l'angle incident θ_i et des vitesses de propagation c_1 et c_2 de l'onde ultrasonore dans les milieux formant l'interface, suivant

$$\frac{\sin\theta_t}{\sin\theta_i} = \frac{c_2}{c_1}. \quad (2.6)$$

Un autre phénomène physique observé est la diffusion de l'énergie de l'onde ultrasonore lorsque l'interface rencontrée est de petite taille par rapport à la longueur d'onde. Dans ce cas, l'onde incidente est renvoyée dans toutes les directions. En pratique, cette diffusion est principalement rencontrée au niveau des microstructures constituant les tissus organiques.

Au final, les échos réceptionnés par les éléments piézo-électriques proviennent à la fois de la réflexion et de la diffusion des ondes émises à travers les différentes interfaces rencontrées. Les ondes diffusées apparaissent alors sur l'image échographique sous la forme d'un aspect granulaire appelé « speckle » en anglais ou encore « chatoiement » ou « tavelure » en français. Ce speckle fournit une texture aux organes mais dégrade la visibilité de leur contour. Il a souvent été considéré comme un bruit. Plusieurs travaux de recherche dans le contexte de l'imagerie médicale ont eu pour objectif de le supprimer. Cependant, il traduit une mesure réelle puisqu'il est généré par des inhomogénéités de structures, appelées diffuseurs. Ces derniers sont répartis dans les tissus mous et réfléchissent chacun une partie du signal initial. Ces diffuseurs peuvent par exemple correspondre à des cellules ou ensembles de cellules, à des fibres musculaires ou à d'autres éléments microscopiques constituant les tissus organiques. Nous verrons dans le chapitre 4 que cette information de type speckle peut se révéler très utile pour l'asservissement visuel échographique.

Modes d'acquisition. Il existe plusieurs modes d'acquisition en imagerie échographique. Nous décrivons brièvement ici les modes les plus répandus.

Le mode A (Amplitude) représente l'amplitude des échos renvoyés par les interfaces traversées par l'onde ultrasonore selon une simple ligne de tir. Ce mode ne fournit par une image 2D mais un signal monodimensionnel qui est fonction de la profondeur. Cette technique, autrefois employée en neurochirurgie (figure 2.4) et en ophtalmologie pour mesurer la taille du

globe oculaire, est à présent très peu utilisée.

Le mode B (Brillance) est actuellement le plus utilisé et correspond à celui que nous avons retenu dans nos travaux car il fournit une image 2D. Il dérive du mode A auquel est associé un balayage électronique permettant d'émettre l'onde ultrasonore et de réceptionner les échos selon différentes lignes de tir (figure 2.5). En pratique ce balayage est obtenu par une excitation séquentielle d'éléments piézo-électriques qui sont disposés de manière rectiligne ou convexe sur l'extrémité de la sonde. La formation de l'image est ensuite réalisée par différents traitements parmi lesquels se succèdent une phase d'amplification, de démodulation et de redressement des signaux réceptionnés, puis une numérisation et interpolation afin de former une image 2D en niveau de gris dont l'intensité est fonction de l'amplitude de l'écho. Une compensation de l'atténuation en fonction de la profondeur et une adaptation de la dynamique par une compression logarithme sont également appliquées avant l'affichage de l'image sur l'écran vidéo de la station échographique.

Le mode Doppler permet de mesurer la vitesse du sang dans les artères ou les veines. Comme son nom l'indique, il est basé sur l'effet Doppler. Il consiste à mesurer la différence de fréquence entre l'émission du signal ultrasonore et la réception de l'écho. Si la fréquence de l'onde écho est plus élevée que celle de l'onde émise, cela signifie que l'interface rencontrée se rapproche du capteur et inversement si sa fréquence est plus faible. En utilisant l'information fournie par plusieurs éléments piézo-électriques, il est alors possible de reconstruire la vitesse et la direction des globules rouges. Il existe deux modes Doppler. Un mode dit continu où l'onde ultrasonore est émise en continu par la sonde, et un mode dit pulsé où les éléments piezo-électriques fonctionnent de manière séquentielle en émetteur et récepteur. Ce dernier mode permet de régler la profondeur de la zone à analyser mais est incapable de fournir la mesure d'une vitesse élevée. Cette mesure ne peut se faire qu'avec le mode Doppler continu. Les dispositifs d'échographie actuels permettent de combiner l'acquisition en Doppler avec celle du mode B. L'image affichée est une image en mode B sur laquelle sont superposées des zones de couleur représentant les flux sanguins. L'amplitude de la vitesse ainsi que la direction de ces flux dans un sens ou l'autre sont représentées par un code de couleur.

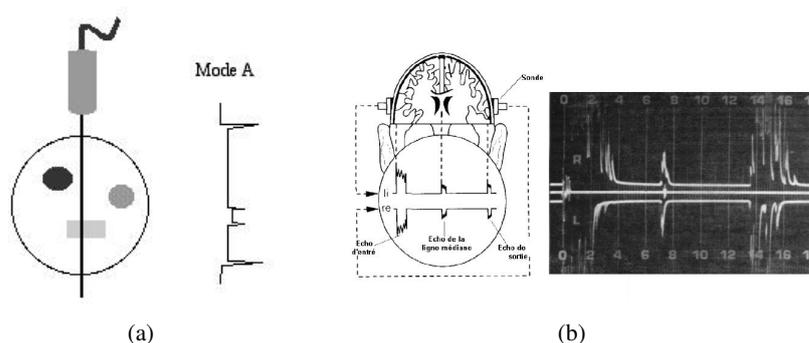


FIGURE 2.4 – (a) Le mode A correspond à une seule ligne de tir de l'onde ultrasonore. (b) Utilisation du mode A en neurologie pour détecter un épanchement intracrânien ou une masse par mesure du déplacement de l'écho de la ligne médiane [Dagon 2004].

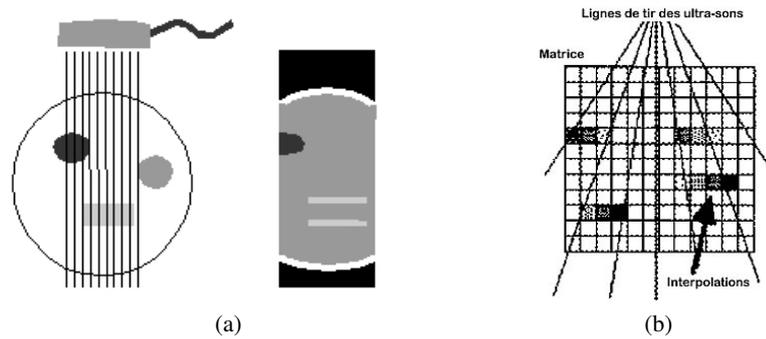


FIGURE 2.5 – (a) Le mode B correspond à une représentation spatiale du signal obtenue après interpolation (b) des lignes de tir.

2.1.3 Différents types de sonde

En pratique, il existe plusieurs types de sonde ultrasonore qui permettent l'emploi de l'imagerie échographique dans de nombreuses disciplines médicales.

Les capteurs 2D. Les sondes 2D sont actuellement les capteurs échographiques les plus répandus dans les centres hospitaliers. Elles se déclinent sous plusieurs formes, en fonction de la cible anatomique considérée. La figure 2.6 regroupe un ensemble de trois capteurs 2D classiquement utilisés dans les examens médicaux. Les capteurs linéaires offrent une bonne résolution d'image notamment en surface mais ont une zone d'observation restreinte. Les sondes convexes ont un cône d'observation plus large et sont particulièrement utilisées pour les examens abdominaux et gynéco-obstétricaux. Enfin, des sondes plus allongées, d'accès transrectal ou transvaginal, permettent d'imager par exemple la prostate ou la vessie.

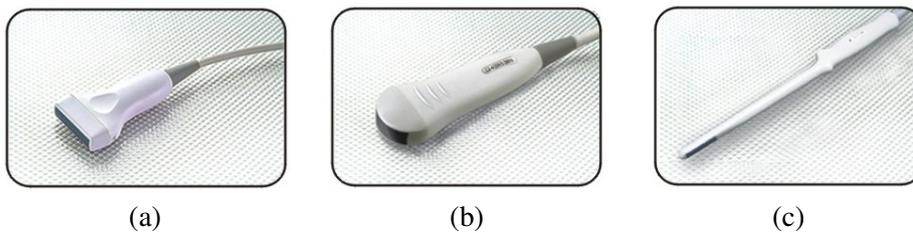


FIGURE 2.6 – Différents modèles de sondes 2D, linéaire (a), convexe (b) et transrectale (c), commercialisées par Meditech.

Aux sondes classiques 2D sont venus s'ajouter récemment quelques modèles de sondes présentant une géométrie plus complexe. Ces sondes permettent de visualiser simultanément plusieurs coupes, voire un volume complet, de l'organe ciblé.

Les capteurs bi-plans. Actuellement les sondes bi-plans sont généralement des sondes transrectales ou transœsophagiennes, constituées d'une barrette droite (capteur linéaire) et d'un transducteur convexe situé à son extrémité. Une telle sonde transrectale a été développée par

la société Vermon pour permettre une visualisation optimale de la prostate [Felix 2005]. Cette sonde présentée en figure 2.7(a) comprend un capteur convexe de 96 éléments PZT qui fournit un plan de visualisation transverse (c'est-à-dire orthogonal à l'axe de la sonde) et deux capteurs linéaires positionnés de part et d'autre du plan transverse comprenant également 96 éléments chacun. La sonde qui permet ainsi la visualisation simultanée des plans transverses et longitudinaux est bien adaptée dans le contexte d'une curiethérapie de la prostate où elle permet à la fois la détection de l'aiguille dans le plan longitudinal et la superposition de la grille d'implantation des noyaux radioactifs sur la vue transverse [Felix 2005].

Tout récemment, une nouvelle sonde bi-plan a également été développée par société Vermon pour des explorations abdominales ou cardiaques. Cette sonde XYPA2.5 (figure 2.7(b)) est constituée de deux capteurs convexes orthogonaux de 2.5 MHz et fournit les deux images à une cadence de 25Hz. Elle présente par ailleurs une petite empreinte, qui la rend adaptée à l'imagerie intercostale.

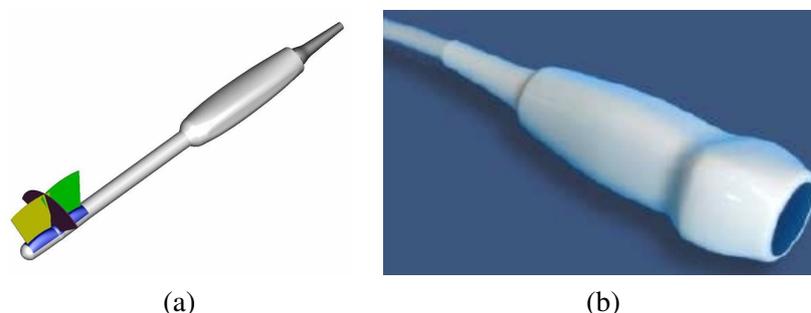


FIGURE 2.7 – Deux modèles de sondes bi-plans, transrectale [Felix 2005] (a) et abdominale (b), commercialisées par Vermon.

Les capteurs 3D. La première solution de capteur 3D dérive naturellement de la géométrie de la sonde 2D classique, en modifiant l'empreinte de la sonde. Les cellules PZT ne sont plus disposées selon une ligne mais selon une matrice rectangulaire, d'où le nom de sondes matricielles donné à cette catégorie de capteurs. Ce type de sonde permet la capture quasi-instantanée d'un volume 3D pyramidal grâce aux données fournies par plusieurs milliers de cellules PZT (figure 2.8). C'est à l'heure actuelle la seule méthode permettant la visualisation 3D temps réel d'un cœur battant. Cependant, le développement de telles sondes matricielles est freiné par de nombreuses difficultés techniques liées au nombre considérable de transducteurs PZT nécessaires pour obtenir une image 3D de bonne qualité (interférences entre les transducteurs et électronique complexe).

La seconde solution, qui constitue la majorité des sondes échographiques 3D, consiste à créer une image volumétrique à partir d'une séquence d'images 2D obtenue par le balayage mécanique motorisé d'un capteur 2D. Ce balayage est généralement réalisé en éventail pour les sondes abdominales ou cardiaques et par rotation pour les sondes trans-vaginales (figure 2.9). Les coupes 2D ainsi générées fournissent un échantillonnage du volume 3D qui est finalement reconstruit par une méthode d'interpolation.

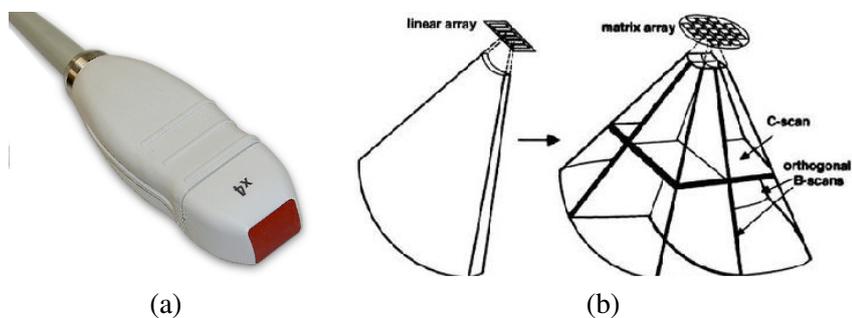


FIGURE 2.8 – Sonde matricielle 3D. (a) Un modèle de sonde 3D matricielle commercialisé par Philips. (b) Une matrice de transducteurs permet de reconstruire un volume pyramidal de l'objet observé.

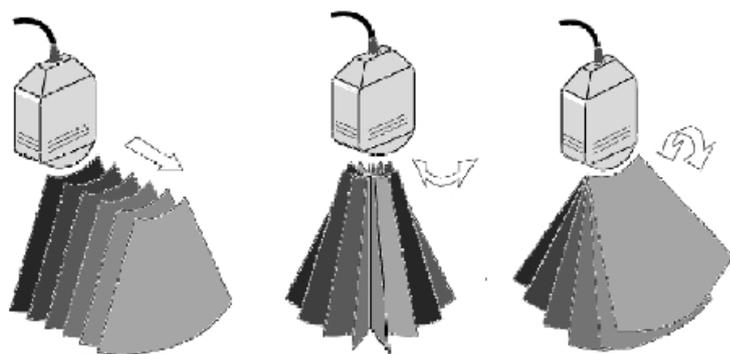


FIGURE 2.9 – Reconstruction du volume 3D par balayage d'un transducteur 2D [Rohling 1998].

2.2 Techniques d'asservissement visuel

Les notions de base de l'asservissement visuel sont rappelées dans cette section. Afin d'obtenir une explication plus approfondie, nous invitons le lecteur à se reporter aux articles tutoriels [Chaumette 2006] et [Chaumette 2007].

Par définition, la commande par vision ou asservissement visuel consiste à contrôler les mouvements d'un système dynamique en intégrant les informations fournies par un capteur de vision dans un système de commande en boucle fermée. Initialement, les techniques d'asservissement visuel ont été développées pour contrôler des robots de type manipulateurs à partir de l'image fournie par une caméra monoculaire. La figure 2.10 expose un schéma de commande classique incluant les différents maillons de la chaîne de traitement.

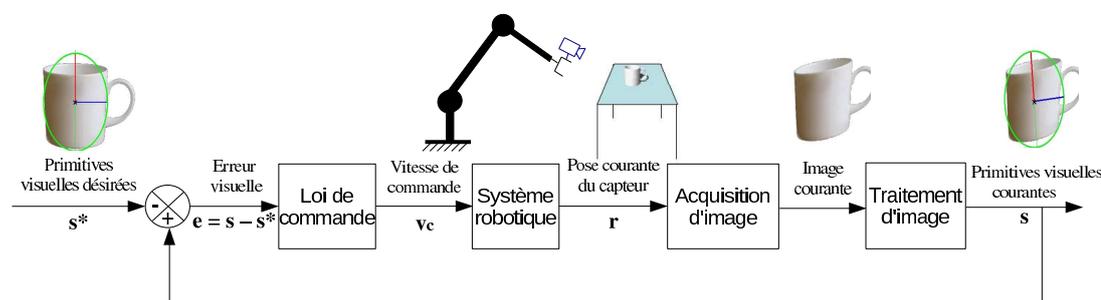


FIGURE 2.10 – Schéma de commande par asservissement visuel d'un système robotique.

2.2.1 Une classification selon trois critères

Une classification à présent bien connue permet de qualifier les différents types d'asservissement visuel selon trois critères suivants : la configuration spatiale du capteur visuel dans le système robotique, le type de grandeur à asservir et le type d'architecture de commande utilisé.

La position du capteur visuel. Dans la situation où le capteur visuel est monté sur l'organe terminal du robot, la configuration du capteur est alors dite embarquée. L'objectif de l'asservissement est généralement de déplacer le capteur afin d'observer un objet d'intérêt selon une vue désirée pour l'atteindre ou le suivre s'il est mobile. L'autre possibilité est la configuration dite déportée où le capteur visuel n'est pas monté sur l'effecteur du robot asservi. Le but est alors de faire évoluer le robot par rapport à son environnement de travail, en percevant à l'aide du capteur visuel l'image du robot et des objets avec lesquels il interagit.

Les grandeurs à asservir. Différents types de grandeurs peuvent être considérés en entrée du schéma de commande de l'asservissement visuel. Elles peuvent être de nature tridimensionnelle comme la mesure de l'attitude d'un objet exprimée dans le repère de la caméra ; ou de type 2D lorsqu'elles représentent directement des primitives visuelles géométriques extraites

du plan image. Nous verrons plus en détail dans la section 2.2.6 différents choix possibles d'informations visuelles dans le cadre de l'utilisation d'une caméra.

L'architecture de commande du robot. Il existe principalement deux approches de commande permettant de contrôler un robot avec un retour visuel bouclé :

- **La commande cinématique** est la plus répandue. Elle consiste à envoyer périodiquement des consignes de vitesses articulaires au contrôleur bas niveau du robot. La cadence de rafraîchissement des consignes est imposée entre autres par la fréquence d'acquisition des images, la durée des traitements nécessaires à l'extraction de l'information visuelle et le calcul de la loi de commande. Ce type de commande est généralement utilisé pour des fréquences d'acquisition des images allant de 1 à environ 100 Hz. Les effets dynamiques n'étant pas modélisés, la fonction de transfert du robot est considérée comme étant un intégrateur pur. Cette commande a l'avantage d'être facilement mise en œuvre si le contrôleur bas niveau du robot permet de générer et d'appliquer une trajectoire de positions articulaires à partir de consignes de vitesses articulaires. Néanmoins, elle n'est pas adéquate pour des applications nécessitant une réactivité élevée du robot du fait de la non modélisation des effets dynamiques.
- **La commande dynamique** tient compte d'une modélisation plus réaliste du comportement du robot. Cependant une identification des paramètres dynamiques (masses, inerties, frottements,...) du robot est nécessaire afin de mettre en œuvre une commande en vitesse ou en couple qui linéarise le comportement dynamique du système. Dans ce cas, le système de vision commande directement les variateurs des actionneurs du robot en se substituant au contrôleur bas niveau du robot. Cette architecture permet de tirer pleinement partie des possibilités dynamiques offertes par le robot pour réaliser des asservissements visuels rapides. Elle nécessite cependant une cadence élevée du système de vision (plus de 100 images/s) afin de garantir la stabilité et la convergence de la commande. Nous invitons le lecteur à se référer aux travaux [Gangloff 1999, Gangloff 2003] s'il souhaite une description plus approfondie dans le cas d'une commande dynamique.

Dans la suite de ce document nous nous positionnerons dans le cadre des asservissements visuels basés sur une commande cinématique.

2.2.2 Formalisme de la fonction de tâche

Un formalisme introduit dans [Samson 1991] permet d'exprimer une tâche robotique à effectuer comme étant la régulation vers zéro sur un horizon temporel d'une fonction $\mathbf{e}(\mathbf{q}, t)$ appelée fonction de tâche. Cette fonction de dimension n appartient à la classe C^2 et est dépendante du vecteur \mathbf{q} contenant les n coordonnées articulaires du robot. Elle exprime l'erreur entre la configuration courante et la configuration souhaitée du capteur par rapport à la scène. Dans le cas de l'asservissement visuel, la fonction de tâche, également appelée tâche référencée vision, se définit par l'expression [Espiau 1992]

$$\mathbf{e}(\mathbf{q}, t) = \mathbf{C} (\mathbf{s}(\mathbf{r}(\mathbf{q}, t), t) - \mathbf{s}^*(t)), \quad (2.7)$$

où :

- \mathbf{s} est le vecteur de mesures de dimension k obtenu à partir des informations visuelles courantes qui doivent être choisies convenablement. Il est fonction de la situation courante $\mathbf{r}(\mathbf{q}, t)$ du capteur vis-à-vis de la scène ;
- \mathbf{s}^* est la consigne sur les informations visuelles de la tâche. Elle peut être constante ou variable en fonction du temps ;
- $(\mathbf{s} - \mathbf{s}^*)$ est donc l'erreur visuelle que la tâche a pour objectif de minimiser ;
- $\mathbf{r}(\mathbf{q}, t)$ est la pose du capteur. Elle dépend de la position articulaire \mathbf{q} du robot et constitue un élément de l'ensemble $\mathbf{SE}(3)$ permettant de décrire la transformation entre corps rigides par une rotation dans $\mathbf{SO}(3)$ et une translation dans \mathbb{R}^3 ;
- \mathbf{C} est une matrice de combinaison de dimensions $n \times k$, avec n correspondant au nombre de degrés de liberté du robot à commander. Elle permet de prendre en compte par combinaison plus d'informations que de degrés de liberté, de manière à gérer une éventuelle redondance de l'information.

2.2.3 Matrice d'interaction

Afin de réguler la fonction de tâche, il est nécessaire de modéliser la variation des informations visuelles induite par le mouvement du capteur visuel (configuration embarquée) ou de l'organe terminal du robot interagissant avec la scène (configuration déportée). Cette variation est obtenue en dérivant $\mathbf{s}(\mathbf{r}, t)$ par rapport au temps selon

$$\dot{\mathbf{s}} = \frac{\partial \mathbf{s}}{\partial \mathbf{r}} \dot{\mathbf{r}} + \frac{\partial \mathbf{s}}{\partial t} = \mathbf{L}_s \mathbf{v} + \frac{\partial \mathbf{s}}{\partial t}, \quad (2.8)$$

où \mathbf{L}_s est la matrice d'interaction [Espiau 1992] de dimension $k \times 6$ associée au vecteur \mathbf{s} qui lie la variation des informations visuelles au torseur cinématique $\mathbf{v} = (\mathbf{v}, \boldsymbol{\omega})$ du capteur. Ce torseur, exprimé dans le repère du capteur, se compose des vitesses de translation $\mathbf{v} = (v_x, v_y, v_z)$ et de rotation $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ du capteur qui sont relatives à son environnement. Le terme $\frac{\partial \mathbf{s}}{\partial t}$ introduit la variation propre des informations visuelles \mathbf{s} qui est nulle si la scène est immobile ou, dans le cas contraire, rarement connue à l'avance.

2.2.4 Commande cinématique

Généralement, la régulation à zéro de la fonction de tâche (2.7) est réalisée par une commande cinématique garantissant une décroissance exponentielle de $\|\mathbf{e}\|$. Cette décroissance est obtenue pour chaque composante de \mathbf{e} en vérifiant l'équation différentielle du premier ordre suivante :

$$\dot{\mathbf{e}} = -\lambda \mathbf{e}, \quad (2.9)$$

où λ est un gain positif dont l'inverse correspond directement à la constante de temps de la décroissance exponentielle de chaque composante de \mathbf{e} .

Nous présentons ici la commande cinématique dans le cas où la matrice de combinaison \mathbf{C} est constante. La dérivée temporelle de la fonction de tâche est alors égale à

$$\dot{\mathbf{e}} = \mathbf{C} (\dot{\mathbf{s}} - \dot{\mathbf{s}}^*). \quad (2.10)$$

En tenant compte de l'équation (2.8), cette expression peut s'écrire directement en fonction du torseur cinématique \mathbf{v} du capteur visuel selon

$$\dot{\mathbf{e}} = \mathbf{CL}_s \mathbf{v} + \mathbf{C} \frac{\partial \mathbf{s}}{\partial t} - \mathbf{C} \dot{\mathbf{s}}^*, \quad (2.11)$$

ou encore

$$\dot{\mathbf{e}} = \mathbf{CL}_s \mathbf{v} + \frac{\partial \mathbf{e}}{\partial t}. \quad (2.12)$$

Le terme $\frac{\partial \mathbf{e}}{\partial t}$ représente les variations de \mathbf{e} dues à la variation propre de \mathbf{s} et à la variation de la consigne \mathbf{s}^* . En remplaçant le terme de gauche par l'équation (2.9), nous obtenons la commande cinématique caractérisant la décroissance exponentielle découplée de la fonction de tâche, à savoir

$$\mathbf{v} = -(\mathbf{CL}_s)^{-1} \left(\lambda \mathbf{e} + \frac{\partial \mathbf{e}}{\partial t} \right). \quad (2.13)$$

La mise en œuvre d'une telle commande est d'ordre théorique, car elle suppose que :

- les dynamiques du système soient négligeables et n'interviennent pas,
- la matrice d'interaction \mathbf{L}_s soit exactement connue, ce qui implique la connaissance des paramètres du capteur visuel, d'un modèle exact de son interaction avec l'environnement, voire même de la géométrie de l'environnement,
- la variation propre $\frac{\partial \mathbf{s}}{\partial t}$ des informations visuelles doit être connue à l'avance, ce qui est rarement le cas,
- la matrice \mathbf{CL}_s soit inversible, carrée, et être telle que \mathbf{e} soit de dimension au moins égale au nombre de degrés de liberté n du robot. En effet, si $\dim(\mathbf{e}) > n$, le système est sous-actionné et il n'est pas possible en général de faire converger \mathbf{e} vers $\mathbf{0}$. Si $\dim(\mathbf{e}) < n$, le système comporte des degrés de liberté redondants et il est possible de réaliser la tâche de vision tout en effectuant une tâche secondaire comme nous le détaillerons dans la section 2.2.5.

En pratique, la matrice d'interaction \mathbf{L}_s est souvent construite sur des hypothèses fortes qui l'éloignent de son exactitude. Dans la majorité des cas, elle dépend des paramètres intrinsèques du capteur visuel dont les valeurs sont généralement estimées expérimentalement. La loi de commande réellement appliquée est donc

$$\mathbf{v} = -(\widehat{\mathbf{CL}}_s)^{-1} \left(\lambda \mathbf{e} + \frac{\partial \widehat{\mathbf{e}}}{\partial t} \right), \quad (2.14)$$

où $\widehat{\mathbf{L}}_s$ est une valeur approchée ou estimée de la matrice d'interaction et $\frac{\partial \widehat{\mathbf{e}}}{\partial t}$ est une prédiction de la variation propre de \mathbf{e} . Si la scène considérée est immobile et la consigne visuelle est constante, la loi de commande devient

$$\mathbf{v} = -\lambda (\widehat{\mathbf{CL}}_s)^{-1} \mathbf{e}. \quad (2.15)$$

Le choix de la matrice de combinaison \mathbf{C} dépend du nombre d'informations visuelles k considéré.

- Si $k = n$, la matrice \mathbf{C} peut être choisie égale à la matrice identité. Dans ce cas, la fonction de tâche correspond directement à l'erreur visuelle et un comportement de décroissance exponentielle pour tous les éléments du vecteur \mathbf{s} est obtenu avec la loi de commande :

$$\mathbf{v} = -\lambda \widehat{\mathbf{L}}_{\mathbf{s}}^{-1} (\mathbf{s} - \mathbf{s}^*) \quad (2.16)$$

- Si $k > n$, la matrice \mathbf{C} est de dimension $n \times k$ et doit être de rang n . Elle peut être choisie comme étant la pseudo-inverse d'une estimation de la matrice d'interaction à la position désirée [Espiau 1992], soit $\mathbf{C} = \widehat{\mathbf{L}}_{\mathbf{s}=\mathbf{s}^*}^+$. Une autre solution consiste à choisir \mathbf{C} égale à la matrice identité de dimension k et à remplacer dans (2.15) l'inverse de la matrice d'interaction par sa pseudo-inverse de Moore-Penrose $\mathbf{L}_{\mathbf{s}}^+ = (\mathbf{L}_{\mathbf{s}}^{\top} \mathbf{L}_{\mathbf{s}})^{-1} \mathbf{L}_{\mathbf{s}}^{\top}$. Pour ce dernier choix nous obtenons alors la loi de commande suivante :

$$\mathbf{v} = -\lambda \widehat{\mathbf{L}}_{\mathbf{s}}^+ (\mathbf{s} - \mathbf{s}^*) \quad (2.17)$$

Analyse de stabilité

La méthode de Lyapunov peut être utilisée pour étudier la stabilité du système en boucle fermée. La fonction candidate choisie est la norme au carré de la fonction de tâche

$$\mathcal{L} = \frac{1}{2} \|\mathbf{e}(t)\|^2. \quad (2.18)$$

L'analyse consiste alors à démontrer les conditions permettant de garantir la décroissance de \mathcal{L} en fonction du temps. En dérivant (2.18), il en résulte l'expression

$$\dot{\mathcal{L}} = \mathbf{e}^{\top} \dot{\mathbf{e}}, \quad (2.19)$$

où la variation $\dot{\mathbf{e}}$ de la fonction de tâche du système bouclé est obtenue en combinant (2.12) et (2.14), soit

$$\dot{\mathbf{e}} = -\lambda \mathbf{C} \mathbf{L}_{\mathbf{s}} (\mathbf{C} \widehat{\mathbf{L}}_{\mathbf{s}})^{-1} \mathbf{e} - \mathbf{C} \mathbf{L}_{\mathbf{s}} (\mathbf{C} \widehat{\mathbf{L}}_{\mathbf{s}})^{-1} \frac{\partial \widehat{\mathbf{e}}}{\partial t} + \frac{\partial \mathbf{e}}{\partial t}. \quad (2.20)$$

En supposant que $\frac{\partial \mathbf{e}}{\partial t} = \frac{\partial \widehat{\mathbf{e}}}{\partial t} = 0$, ce qui signifie que \mathbf{s} ne présente pas de variation propre (scène immobile) et que la consigne à atteindre \mathbf{s}^* est constante, nous obtenons

$$\dot{\mathcal{L}} = -\lambda \mathbf{e}^{\top} \mathbf{C} \mathbf{L}_{\mathbf{s}} (\mathbf{C} \widehat{\mathbf{L}}_{\mathbf{s}})^{-1} \mathbf{e}, \quad (2.21)$$

qui conduit à la condition de stabilité suivante :

$$\mathbf{Q} = \mathbf{C} \mathbf{L}_{\mathbf{s}} (\mathbf{C} \widehat{\mathbf{L}}_{\mathbf{s}})^{-1} > \mathbf{0}. \quad (2.22)$$

Comme souligné dans [Chaumette 2006], la positivité de la matrice \mathbf{Q} est suffisante pour garantir la stabilité asymptotique globale dans le cas où $k = n = 6$ et les matrices $\mathbf{L}_{\mathbf{s}}$ et $\widehat{\mathbf{L}}_{\mathbf{s}}^{-1}$ de dimension 6×6 sont de rang plein. La convergence de $\|\mathbf{e}\|$ vers zéro est alors assurée puisque \mathcal{L} est décroissante. Par contre, si le nombre d'informations visuelles est supérieur au nombre de degrés de liberté $k > n$, la matrice $\mathbf{L}_{\mathbf{s}}$ de dimension $k \times n$ a un rang qui au plus est égal à n . Par conséquent, le noyau associé n'est pas nul et il existe des configurations du vecteur \mathbf{e} qui correspondent à des minima locaux. Dans ce cas, la condition (2.22) permet uniquement de garantir la stabilité asymptotique locale dans un voisinage proche de la configuration désirée. Bien que cette relation soit simple, l'étude de la stabilité du système en tenant compte des erreurs de mesure et d'estimation reste un problème complexe qui est toujours d'actualité.

2.2.5 Tâches hybrides : le formalisme de la redondance

Pour une application donnée, l'objectif de l'asservissement n'est pas forcément de réguler une tâche visuelle unique qui contraint les n degrés de liberté du système dans le but d'obtenir une liaison virtuelle rigide entre le capteur et son environnement. En effet, il peut être intéressant de combiner une tâche visuelle qui commande $m < n$ degrés de liberté avec une tâche secondaire dont les mesures proviennent soit directement de l'image, soit de la reconstruction d'informations à partir de l'image ou d'autres types de capteurs (capteur de proximité, position, force...).

Le formalisme de la redondance qui complète celui des fonctions de tâche permet de poser de manière élégante le problème [Samson 1991]. Dans cette approche, une fonction de tâche principale, notée e_1 , est considérée prioritaire devant une tâche secondaire e_2 qui consiste à minimiser une fonction de coût sous la contrainte que la tâche principale soit réalisée. Dans le contexte de l'asservissement visuel, la tâche visuelle est généralement choisie comme étant la tâche principale dont la dimension m est inférieure aux n degrés de liberté du système selon

$$e_1 = C(s - s^*), \quad (2.23)$$

avec une matrice de combinaison C de dimension $m \times k$ qui doit toujours être définie de rang plein. La fonction de tâche globale, minimisant la tâche secondaire e_2 sous la contrainte que la tâche de vision principale soit réalisée, soit $e_1 = 0$, s'écrit [Samson 1991, Espiau 1992]

$$e = W^+ e_1 + \alpha(I_n - W^+ W) e_2, \quad (2.24)$$

où :

- W est une matrice $m \times n$ de rang plein m telle que son noyau soit identique au noyau de la matrice d'interaction L_s , soit

$$\ker(W) = \ker(L_s). \quad (2.25)$$

En pratique, W est construite à partir de \widehat{L}_s car nous avons uniquement accès à une valeur approchée ou estimée de la matrice d'interaction.

- $(I_n - W^+ W)$ est un opérateur qui réalise la projection orthogonale sur le noyau de \widehat{L}_s . Cette projection permet de limiter les mouvements induits par la tâche secondaire e_2 afin qu'ils ne perturbent pas la tâche prioritaire e_1 . Notons que cette limitation des mouvements ne garantit pas que la tâche secondaire soit atteinte au final. De plus, étant donné que W est définie à partir d'une estimation \widehat{L}_s , la tâche secondaire peut introduire de légères perturbations sur la tâche principale si la qualité de l'estimation de L_s est mauvaise.
- $\alpha > 0$ est un scalaire permettant de régler la pondération de la tâche principale par rapport à la tâche secondaire.

Généralement, la tâche secondaire est définie par le gradient d'une fonction de coût h_s à minimiser selon

$$e_2^T = \frac{\partial h_s}{\partial \mathbf{r}}. \quad (2.26)$$

La régulation de la fonction de tâche globale e est à présent obtenue par loi de commande [Chaumette 2002] exprimée sous la forme

$$\mathbf{v} = -\lambda e - \mathbf{W}^+ \widehat{\frac{\partial e_1}{\partial t}} - \alpha (\mathbf{I}_n - \mathbf{W}^+ \mathbf{W}) \frac{\partial e_2}{\partial t}, \quad (2.27)$$

où $\widehat{\frac{\partial e_1}{\partial t}}$ correspond à une prédiction d'une éventuelle variation propre de la tâche visuelle. La prise en compte de ce terme est souhaitable dans le cas d'une application de poursuite d'objet mobile car il permet d'annuler l'erreur de trainage.

La condition de stabilité de cette commande a été démontrée dans [Samson 1991] et reprise dans [Espiau 1992, Chaumette 2002] pour une adaptation du formalisme aux asservissements visuels. Nous ne la détaillons pas ici mais invitons le lecteur à s'y référer.

2.2.6 Choix des grandeurs asservies

Comme nous l'avons indiqué, un aspect important des techniques d'asservissement visuel est le choix du type d'informations visuelles s considérées en entrée de la commande du système.

Asservissement visuel 3D. Dans un asservissement visuel 3D (position-based visual servoing), la grandeur à asservir s est une mesure tridimensionnelle. Elle peut représenter des coordonnées de points 3D de l'objet observé [Martinet 1996] ou directement caractériser la pose relative de l'effecteur du robot par rapport à un repère cible à atteindre [Wilson 1996]. Dans tous les cas il est nécessaire de reconstruire cette pose à partir d'informations visuelles extraites de l'image et d'un modèle géométrique de l'objet d'intérêt.

Une caractérisation classique consiste à définir le vecteur d'entrée s comme étant la pose entre le repère courant de la caméra \mathcal{R}_c et son repère désiré \mathcal{R}_{c^*} , soit $s = ({}^{c^*}\mathbf{t}_c, \theta \mathbf{u})$. Les vecteurs ${}^{c^*}\mathbf{t}_c$ et $\theta \mathbf{u}$ décrivent respectivement la translation et la rotation à réaliser. Cette dernière est décrite par un angle θ autour d'un vecteur unitaire \mathbf{u} . Dans ce cas, la matrice d'interaction est diagonale par blocs et un découplage complet entre les mouvements de translation et de rotation est obtenu. Ce choix de s conduit à une trajectoire rectiligne de l'effecteur du robot et la stabilité asymptotique globale a été prouvée dans le cas où la pose est parfaitement estimée. Par contre, la visibilité de l'objet d'intérêt dans l'image n'est pas assurée si la caméra est embarquée car sa trajectoire en ligne droite peut amener l'objet à sortir du champ visuel et il n'est alors plus possible d'estimer sa pose. Une solution est alors de choisir $s = ({}^c\mathbf{t}_o, \theta \mathbf{u})$ où ${}^c\mathbf{t}_o$ représente la translation entre le repère courant de la caméra et le repère \mathcal{R}_o de l'objet. Cette approche permet d'obtenir une trajectoire rectiligne dans l'image de la projection de l'origine du repère de l'objet garantissant ainsi sa visibilité. Cependant, son principal défaut est de générer une trajectoire complexe dans l'espace.

Un asservissement visuel 3D nécessite une étape de reconstruction permettant de fournir une mesure d'attitude de l'objet. Les techniques de reconstruction de pose utilisent pour la plupart des primitives géométriques extraites de l'image ainsi qu'une connaissance du modèle

3D de l'objet observé. Les plus simples utilisent des marqueurs artificiels afin de mettre en correspondance des points 3D de l'objet d'intérêt avec les coordonnées dans l'image de la projection de ces points [DeMenthon 1995, Liu 1990]. D'autres approches également basées sur un modèle CAO de l'objet s'affranchissent de l'utilisation de marqueurs en utilisant des primitives géométriques de type "segments" qui correspondent aux projections des arêtes de l'objet observées dans l'image [Lowe 1991, Drummond 2002, Comport 2006].

Néanmoins, quelle que soit la méthode de reconstruction de pose utilisée, l'estimation de l'attitude est très sensible aux bruits de mesure et aux erreurs de modélisation et d'étalonnage de la caméra. Il en résulte, après convergence de l'asservissement lorsque le régime permanent est atteint, un biais entre la position du robot et l'objet d'intérêt à atteindre. Ce biais constitue le défaut majeur de l'asservissement visuel 3D.

Asservissement visuel 2D. Dans un asservissement visuel 2D (image-based visual servoing), la grandeur à asservir est exprimée sous la forme de primitives visuelles dans l'image. Contrairement à un asservissement visuel 3D, le contrôle n'est pas réalisé dans l'espace cartésien, mais directement dans l'image. On parle alors d'asservissement dans le plan image ou référencé image. Les primitives extraites de l'image peuvent être des formes géométriques élémentaires qui proviennent de la projection de l'objet d'intérêt dans le plan image tels les coordonnées d'un point, les paramètres d'une droite, d'un segment, d'une ellipse [Chaumette 1990]. Afin de considérer des objets complexes, les primitives peuvent également représenter les paramètres d'un contour [Collewet 2000] ou les moments 2D d'une surface [Chaumette 2004] ou les moments d'un nuage de points [Tahri 2005]. Dans tous les cas, il est possible de définir l'entrée de commande s comme étant une combinaison de différents types de primitives 2D extraites de l'image. L'avantage majeur de l'asservissement visuel 2D est sa robustesse vis-à-vis des erreurs de modélisation du système ou de calibration de la caméra. L'erreur à minimiser est en effet directement exprimée dans l'image sans passer par une étape de reconstruction. De plus, la visibilité des primitives est assurée durant le déplacement car leur convergence est obtenue en ligne droite dans l'image. Par contre, le prix à payer est d'une part, une trajectoire dans l'espace qui peut se révéler très complexe, voire impossible à réaliser dans l'espace de travail du robot, et d'autre part, la possibilité de converger vers des minima locaux si les primitives visuelles ne sont pas correctement choisies. La problématique principale est donc de définir les primitives visuelles qui permettent d'obtenir un bon comportement du système.

- Dans le cas où s est choisi de dimension 6, il est nécessaire que les primitives visuelles retenues assurent un isomorphisme avec l'espace des configuration $\mathbf{SE}(3)$ atteignable. Dans le cas contraire, plusieurs minima globaux existeraient car différentes poses de l'effecteur conduiraient au même état des primitives visuelles voire à la même image. Le choix idéal serait de retenir des primitives visuelles en vue d'obtenir une matrice d'interaction diagonale conduisant ainsi à un système linéaire. La détermination de telles primitives constitue toujours l'objectif de nombreux travaux de recherche dans la communauté.
- Dans le cas où s est de dimension supérieure à 6, comme nous l'avons déjà indiqué, la stabilité globale n'est pas assurée et les primitives visuelles peuvent converger vers des minima locaux. Néanmoins, ce choix peut s'avérer utile dans un voisinage proche de la configuration désirée car il permet de considérer des informations visuelles redondantes

et d'augmenter ainsi la robustesse vis-à-vis d'éventuelles occultations de primitives ou d'erreurs d'extraction. Ce choix permet également de garantir l'isomorphisme entre s et l'espace des configuration $SE(3)$.

Asservissement visuel $d2D/dt$. Pour des scènes complexes qui ne contiennent aucune primitive géométrique simple à extraire, l'information du mouvement dans l'image peut être choisie comme primitive visuelle [Papanikolopoulos 1993, Crétual 2001]. Dans ce cas, une mesure du flot optique représentant le mouvement 2D apparent dans l'image est régulée à un champ de vitesses désiré.

Asservissement visuel 2D 1/2. L'asservissement visuel 2D 1/2 tire partie des avantages des approches 3D et 2D sans en prendre leurs inconvénients. Un vecteur s de dimension 6 est établi à la fois à partir de grandeurs exprimées dans le plan image et dans le repère de la caméra. L'approche la plus connue [Malis 1999] permet d'obtenir un découplage partiel entre les mouvements de translation et de rotation de la caméra. Elle génère une trajectoire correcte dans l'espace tout en garantissant le contrôle de la projection du centre de l'objet dans l'image afin de maintenir sa visibilité. De plus, cette approche ne requiert pas de calcul de pose et la condition de stabilité asymptotique globale du système a été démontrée [Malis 2002].

2.2.7 Cas de l'imagerie échographique

Les techniques existantes d'asservissement visuel ont été fondées pour la plupart sur l'utilisation d'une caméra dont le modèle d'interaction avec l'environnement diffère totalement de celui d'une sonde échographique. Par conséquent certaines d'entre elles ne peuvent pas être appliquées à l'imagerie échographique. Cette incompatibilité ne se situe pas au niveau des lois de commande issues des formalismes génériques de la fonction de tâche et de la redondance, mais porte sur les informations visuelles retenues et leur modèle d'interaction avec l'environnement.

Les approches d'asservissement visuel 3D existantes ne posent pas de problème et peuvent être appliquées en l'état si une mesure de pose peut être reconstruite à partir d'informations visuelles extraites à partir de l'image échographique. Néanmoins, pour obtenir une mesure de pose, il est nécessaire d'avoir un modèle géométrique de l'objet observé ou d'utiliser une sonde échographique 3D fournissant un volume dans lequel peut être extrait l'information de pose.

En revanche, les approches d'asservissement visuel 2D fondées sur l'utilisation de l'imagerie optique ne sont pas applicables en imagerie échographique ou alors uniquement dans un mode dégradé qui ne permet pas de contrôler les mouvements hors du plan d'observation de la sonde, comme nous le verrons à la section 2.3. La raison principale est une modélisation inappropriée de la matrice d'interaction qui ne prend pas en compte les particularités du capteur échographique. De plus, les différents choix d'informations visuelles proposés dans le cas d'une caméra ne sont pas forcément les plus adaptés pour contrôler les 6 ddl du système échographique.

La section suivante s'attache donc à exposer un état de l'art dédié à l'asservissement visuel échographique.

2.3 Asservissements visuels utilisant la modalité échographique : état de l'art

Nos travaux sur l'asservissement visuel échographique ont démarré en fin d'année 2004. A cette époque, très peu d'activités de recherche avaient été menées sur cette thématique qui a pris place essentiellement dans le contexte de la robotique médicale. Nous sommes donc en mesure de présenter dans cette section un état de l'art quasiment complet depuis l'apparition des travaux pionniers aux alentours de l'an 2001 jusqu'à la date de rédaction de ce document. Les critères de classification que nous proposons de retenir sont la configuration du capteur échographique dans le système robotique et le type d'asservissement visuel considéré : 2D ou 3D. A cet effet, nous distinguerons ci-après les travaux qui portent sur l'asservissement visuel échographique où le déplacement de la sonde est contrôlé (configuration embarquée), de ceux qui adressent le contrôle du mouvement d'un instrument médical (configuration déportée). Nous différencierons également ces travaux en fonction de la nature du mouvement contrôlé : déplacement uniquement dans le plan d'observation ou incluant également des composantes hors du plan.

2.3.1 Pour le guidage de la sonde

Asservissement visuel 2D : déplacement uniquement dans le plan d'observation

Des chercheurs de Vancouver ont été les premiers à mettre en œuvre un asservissement visuel utilisant l'image échographique pour contrôler un robot porteur de sonde [Abolmaesumi 2001]. Ce robot avait été initialement développé pour constituer un système de télé-échographie robotisée [Salcudean 1999]. Parmi les six degrés de liberté du robot actionnant la sonde, seuls les trois degrés de liberté correspondant aux mouvements dans le plan de coupe sont commandés par la vision échographique et les 3 ddl générant le mouvement hors du plan sont laissés au contrôle manuel de l'utilisateur via la télé-opération. L'application envisagée est l'imagerie échographique 3D de l'artère carotide. Elle consiste à acquérir une série de coupes transversales de l'artère qui sont ensuite positionnées dans un référentiel fixe au moyen des mesures odométriques du robot. Durant l'acquisition des coupes, l'asservissement visuel a pour fonction de centrer automatiquement la section de l'artère dans l'image pendant que l'utilisateur déplace la sonde le long du cou du patient par télé-opération.

L'information visuelle retenue est de type point et plus précisément les coordonnées dans l'image du centre de la section de l'artère. Cinq méthodes permettant de détecter le centre de l'artère ont été comparées dans [Abolmaesumi 2002] : mesures de similarité de motifs, avec des algorithmes de corrélation croisée (CC) ou de détection par similarité séquentielle (SSD), ou segmentation de contours par un algorithme Star, Star-Kalman ou contour actif de type Snake. Une matrice d'interaction L_s associée à la primitive visuelle de type point est donc considérée dans la commande du robot. Cependant, les auteurs ne

l'ont pas modélisée de manière complète car ils n'ont pas tenu compte de l'influence des mouvements hors du plan sur la variation des coordonnées dans l'image du centre de l'artère. Néanmoins, cet oubli ou ce choix intentionnel conduit à une valeur approchée \widehat{L}_s qui assure tout de même la réalisation de la tâche de centrage dans le plan image dans la mesure où les mouvements hors du plan ne sont pas contrôlés par la vision. Etant donné que l'utilisation comme informations visuelles des deux coordonnées ($k = 2$) dans l'image d'un point unique ne permet pas de contraindre les trois ddl ($n = 3$) dans le plan, les auteurs ont considéré deux points dans l'image correspondant aux centres de deux artères. Toutefois, lors du mouvement hors plan de la sonde qui est généré par téléopération, aucune solution n'a été proposée pour définir la position désirée à atteindre dans l'image par les deux artères. Cet asservissement visuel 2D qui se limite au mouvement dans le plan a été validé sur un fantôme composé de trois tubes plongés dans une bassine d'eau.

Asservissement visuel 3D : déplacement dans et hors du plan d'observation

L'application visée dans [Lee 2007] est quant à elle une assistance robotisée à la procédure de lithotripsie ultrasonique qui consiste à dissoudre un calcul rénal par focalisation d'une forte énergie ultrasonore. Le dispositif présenté est constitué d'un robot cartésien à trois translations dont l'effecteur embarque un transducteur d'ondes de forte intensité (HIFU). Deux sondes échographiques 2D sont également fixées sur l'effecteur du robot de manière à fournir deux coupes orthogonales dont l'intersection contient le point focal des ondes HIFU. Un asservissement visuel échographique est mis en œuvre pour contrôler les trois translations de l'effecteur afin de suivre automatiquement les mouvements du calcul rénal dûs aux mouvements physiologiques du patient tout au long de la procédure de lithotripsie. L'intérêt majeur est de permettre ainsi le traitement d'un calcul en mouvement tout en épargnant les tissus voisins.

L'information visuelle retenue correspond aux coordonnées 3D du centre du calcul rénal exprimées dans un repère attaché à l'effecteur. Ces coordonnées sont reconstruites à partir des coordonnées 2D du calcul extraites dans les deux coupes orthogonales. L'erreur de position 3D entre le point focal et le calcul est ensuite directement appliquée en entrée du contrôleur bas-niveau du robot pour être compensée. Pour chaque image, la position de la section du calcul est extraite à l'aide d'une technique de recherche de motif (pattern matching) où la corrélation normalisée est choisie comme mesure de similarité. Le motif de référence du calcul est initialement défini dans l'image par l'utilisateur au démarrage de la procédure. Cette approche correspond donc à un asservissement visuel échographique 3D où la reconstruction de la position de la cible et du point focal de l'émetteur HIFU nécessite une calibration précise du système qui limite la précision de la tâche de suivi. Une validation en condition in-vitro de ce système a été présentée où une bille plongée dans une bassine d'eau est actionnée par un piston afin de simuler un calcul rénal en mouvement.

Une approche récente [Nakadate 2011] a pour objet la compensation du mouvement de la carotide en maintenant la coupe échographique centrée sur le plan médian de la section longitudinale de l'artère. L'intérêt de cette application est de rendre possible l'analyse

de l'écoulement sanguin sur une durée de plusieurs minutes en vue de diagnostiquer un éventuel problème vasculaire. Seul 1 ddl de la sonde qui correspond au mouvement de translation hors du plan est contrôlé. Une procédure initiale consiste à acquérir une série d'images parallèles ainsi que leur position respective autour de l'image désirée de l'artère. Ensuite lors de l'asservissement, une mise en correspondance de l'image courante avec les différentes images pré-enregistrées est réalisée. Elle repose sur le calcul d'une mesure de similarité (SSD) basée sur les intensités des pixels des images considérées. L'image pré-enregistrée présentant la meilleure similarité avec l'image courante de la sonde est alors sélectionnée comme étant la plus proche de l'image courante : sa position est alors prise comme position courante de la sonde. La vitesse appliquée au robot est définie proportionnellement à l'erreur de position entre cette position estimée de la sonde et la position de l'image cible qui a été définie au milieu de la série des images pré-enregistrées. La méthode a été validée en condition *in vivo* pour suivre la section longitudinale de la carotide selon 1 ddl. Nous pouvons considérer cette approche comme étant basée sur un asservissement visuel 3D. En effet, la mesure de position de la sonde, selon le degré de liberté porté par la translation hors du plan, est reconstruite à partir d'un volume échographique obtenu par balayage au démarrage de la procédure.

2.3.2 Pour le guidage d'un outil

Asservissement visuel 2D : déplacement uniquement dans le plan d'observation

Dans [Hong 2004], un robot médical est commandé par un asservissement visuel échographique 2D pour actionner automatiquement une aiguille rigide lors d'une procédure de cholécystotomie où l'aiguille est introduite dans la vésicule biliaire afin de réaliser un drainage. Le système est constitué d'un manipulateur passif à 5 ddl sur lequel est monté un robot à deux ddl. Ce dernier est rigidement fixé par sa base à une sonde échographique 2D et a été spécialement conçu pour orienter l'aiguille autour d'un point d'entrée et effectuer son insertion tout en contraignant mécaniquement l'axe de l'aiguille à rester dans le plan de coupe de la sonde pour garantir sa visibilité. Au début de la procédure, le manipulateur passif est déplacé manuellement par le médecin pour positionner l'extrémité de l'aiguille sur la peau du patient au niveau du point d'entrée. Un asservissement visuel échographique 2D permet ensuite de réaliser l'insertion automatique de l'aiguille pour atteindre la vésicule tout en compensant d'éventuels mouvements physiologiques ou involontaires du patient.

Les informations visuelles retenues sont les coordonnées dans l'image de l'extrémité de l'aiguille qui est détectée à l'aide d'une transformée de Hough. La position désirée à atteindre correspond au centre de la section de la vésicule qui est segmentée dans l'image par un algorithme basé sur un contour actif. Cependant cet asservissement visuel ne permet que de contrôler deux ddl (une translation le long de l'axe de l'aiguille et une rotation autour du point d'insertion) qui correspondent uniquement à des mouvements dans le plan de la sonde.

Récemment, une approche utilisant un asservissement visuel échographique a également été définie en vue de réaliser l'insertion d'une aiguille flexible dans le plan d'observa-

tion fourni par une sonde échographique 2D déportée [Neubach 2010]. Cette dernière est initialement positionnée de manière à observer le point d'entrée de l'aiguille et la cible à atteindre tout au long de la procédure. L'extrémité externe de l'aiguille est fixée à l'effecteur d'un robot parallèle à 6 ddl. L'objectif de la tâche est d'amener le bout de l'aiguille sur une cible définie par l'utilisateur dans l'image tout en contournant des obstacles en modifiant la courbure de l'aiguille.

La mesure visuelle considérée est la position 2D du bout de l'aiguille qui est détectée lorsque l'aiguille est en mouvement par un simple calcul de différence entre l'image courante et l'image précédente. Cette position est ensuite utilisée en entrée d'un générateur de trajectoire de l'aiguille qui va appliquer un déplacement au robot de manière à minimiser l'erreur de position avec la cible. La méthode employée pour générer la trajectoire de l'aiguille a été présentée dans une étude préliminaire [Glozman 2004] où les auteurs ont décrit un modèle direct capable de prédire la trajectoire planaire du bout de l'aiguille en fonction du déplacement de la base de l'aiguille (selon 1 rotation, et 2 translations dans le plan) ainsi que des raideurs mécaniques des tissus rencontrés. Ces derniers sont modélisés par des ressorts qui appliquent des forces latérales sur le corps de l'aiguille. Un modèle inverse permet ensuite de déterminer la position de la base de l'aiguille en fonction d'une trajectoire donnée. Cet asservissement visuel a été validé sur un fantôme simulant deux couches de tissus mous caractérisées par des raideurs différentes. Ces raideurs sont estimées en ligne à partir d'une détection des déformations des tissus dans l'image dues à l'insertion de l'aiguille.

Asservissement visuel 2D : déplacement dans et hors du plan d'observation

Dans le contexte de la chirurgie laparoscopique, des travaux français ont été menés en vue d'assister la chirurgie cardiaque sur cœur battant et plus particulièrement le geste de réparation de la valve mitrale. Un asservissement visuel échographique basé image utilisant une sonde échographique 2D est proposé dans [Vitrani 2005] pour contrôler les 4 ddl d'un instrument de chirurgie laparoscopique muni à son extrémité d'un forceps et actionné par un robot manipulateur.

Les informations visuelles retenues sont de type point et correspondent à l'intersection des 2 mâchoires du forceps avec le plan de coupe observé par la sonde déportée. Les échos dans l'image correspondant à ces deux sections ont la particularité de présenter une intensité élevée. Ils sont détectés après filtrage et seuillage de l'image par un algorithme qui consiste dans un premier temps à appliquer des opérations morphologiques (ouverture et fermeture), puis à discriminer les échos restants par une méthode heuristique basée sur un critère de taille de région et de vitesse de déplacement de celles-ci [Ortmaier 2005]. Les coordonnées des deux points sont ensuite fournies par un calcul de barycentre des deux régions retenues. Les auteurs ont modélisé de manière complète la matrice d'interaction qui lie les variations des coordonnées dans l'image des deux points aux déplacements de l'outil dans et en dehors du plan de coupe de la sonde. Le modèle qui lie la variation aux mouvements hors plan est basé sur le fait que chaque point d'intersection est contraint à se déplacer le long d'une droite matérialisée par une mâchoire du forceps. Cette modélisation est fondée sur l'hypothèse que l'angle d'ouverture des forceps est connu et reste constant au cours du temps. L'objectif de l'asservissement vi-

suel est alors de positionner automatiquement ces deux points vers des positions définies par l'utilisateur dans l'image.

Par la suite, de nouvelles primitives visuelles ont été retenues afin d'assurer une meilleure robustesse de la commande [Vitrani 2007]. Elles correspondent aux coordonnées du milieu du segment défini par les deux points dans l'image, sa longueur et son orientation par rapport à l'axe horizontal de l'image. Etant donné que la commande est appliquée au robot actionnant l'outil et non pas à la sonde, une connaissance de la pose relative entre le repère de la sonde et le repère du robot est nécessaire pour calculer la vitesse de consigne à appliquer à l'effecteur de ce dernier. Une méthode de calibration hors ligne est proposée dans [Vitrani 2008] pour fournir une estimation de cette pose ainsi que les facteurs d'échelles de l'image échographique. Cette méthode consiste à mesurer la position dans l'image d'un point d'intersection de l'outil pour différentes poses de l'instrument. Il s'agit d'estimer à l'aide d'un algorithme d'optimisation les valeurs de ces paramètres recherchés qui minimisent au mieux la distance entre la position du point observé et sa position reconstruite dans l'image à partir des paramètres estimés. Cet asservissement visuel 2D a été validé en conditions *in vivo* lors d'une expérimentation où l'outil est inséré dans le ventricule gauche d'un cœur de cochon dont l'accès a été ménagé par une sternotomie (ouverture du thorax) et où la sonde échographique déportée est maintenue sur la surface du cœur.

Dans le même contexte applicatif, une approche différente a été proposée pour guider un instrument chirurgical également à partir de deux mesures visuelles de type point qui correspondent à l'intersection d'un forceps avec le plan d'observation d'une sonde échographique 2D [Sauvée 2008]. La différence principale par rapport à l'approche présentée précédemment porte sur la commande mise en œuvre. Cette dernière n'est pas une commande issue du formalisme de la fonction de tâche mais une commande prédictive non linéaire (CPNL) qui considère le contrôle comme un problème d'optimisation sous contraintes [Allgower 1999].

Cette approche n'utilise pas une modélisation de la matrice d'interaction mais un modèle direct fournissant une prédiction de la position des points d'intersection dans l'image en fonction des positions articulaires du robot. Ce modèle est construit sur plusieurs hypothèses. D'une part, les deux points correspondant à l'intersection du forceps avec le plan de coupe sont considérés comme rigidement liés à l'outil et à une distance imposée constante du point de jonction des mâchoires du forceps. En réalité, ces points glissent le long des 2 droites matérialisées par les mâchoires lorsqu'un mouvement hors plan est appliqué à l'outil. D'autre part, l'orientation de l'axe principal de l'instrument est considéré comme étant orthogonal avec le plan de coupe. Les coordonnées 3D de ces points sont exprimées dans le repère de l'effecteur du robot à partir d'un modèle géométrique de l'instrument qui est fonction de l'angle d'ouverture du forceps et de la longueur de tige de l'instrument. L'utilisation du modèle géométrique direct du robot et d'une estimation de la pose relative entre le repère de la sonde et le repère de la base du robot permet ensuite d'exprimer les coordonnées 3D, ${}^s\mathbf{x} = ({}^sx, {}^sy, {}^sz)$ des deux points considérés dans le repère de la sonde. Ces coordonnées qui ne représentent pas forcément des points situés dans le plan d'observation (${}^sz \neq 0$) sont projetées orthogonalement dans le plan de coupe puis exprimées en pixel à l'aide des facteurs d'échelle de l'image afin de fournir une prédiction approchée de la position des points d'intersection du forceps dans

l'image. Le calcul de la vitesse articulaire de commande \mathbf{u}_k du robot consiste alors à minimiser à chaque itération de commande k une fonction de coût de forme quadratique définie sur un horizon temporel fini de taille N_p . Elle s'exprime par

$$\mathcal{C} = \frac{1}{2} \sum_{i=0}^{N_p-1} \left[\epsilon_{i|k}^T \mathbf{Q} \epsilon_{i|k} + \mathbf{u}_{i|k}^T \mathbf{R} \mathbf{u}_{i|k} \right], \quad (2.28)$$

où $\epsilon_{i|k}$ est la séquence d'erreurs de prédiction entre la position désirée des points dans l'image et leur position prédite sur l'horizon N_p et $\mathbf{u}_{i|k}$ correspond à la séquence de commande optimale sur l'horizon N_p dont uniquement le premier élément est appliqué au robot $\mathbf{u}_k = \mathbf{u}_{1|k}$. Les matrices positives et symétriques \mathbf{Q} et \mathbf{R} permettent de pondérer respectivement l'erreur de prédiction et l'énergie mise en jeu par la commande. Dans cette approche la mesure de la position des points dans l'image est utilisée pour réaliser un réajustement en sortie du modèle de prédiction. Ce réajustement consiste à ajouter aux coordonnées prédites la différence entre les coordonnées mesurées et les coordonnées prédites afin de compenser les erreurs de modèle. Lors de la minimisation de la fonction de coût, les auteurs ont également proposé d'ajouter des contraintes qui assurent le maintien de la visibilité des points dans l'image et leur coordonnée en Z dans l'épaisseur du plan de coupe. Ils ont aussi proposé d'ajouter des contraintes d'évitement des butées articulaires et de saturation des vitesses. L'optimisation sous contraintes est réalisée au moyen de la librairie CFSQP (a C library of the Feasible Sequential Quadratic Programming algorithm) [Lawrence 1997] et l'approche a été validée avec un robot manipulateur à 6 ddl équipé d'un instrument plongé dans une bassine d'eau et observé par une sonde immobile déportée. Bien que cette approche soit intéressante car elle permet de prendre en compte différentes contraintes, les points ne suivent pas de lignes droites dans l'image à la différence d'un asservissement visuel 2D où l'erreur entre la consigne et la mesure visuelle est directement minimisée dans l'image.

Asservissement visuel 3D : déplacement dans et hors du plan d'observation

Dans [Novotny 2007], l'utilisation d'une sonde échographique matricielle 3D a été proposée pour contrôler les 4 degrés de liberté d'un instrument robotisé de chirurgie laparoscopique actionné par un robot manipulateur. Dans cette étude réalisée en condition *in vitro* où l'outil et la sonde sont plongés dans une bassine d'eau, une reconstruction complète du volume observé par le capteur permet d'estimer la position de l'instrument dans un repère cartésien associé à la sonde déportée. Cette localisation 3D de l'instrument est ensuite utilisée pour le guider automatiquement vers une cible constituée par un marqueur artificiel en forme de croix plongé dans d'eau. L'extraction de la position et de l'orientation de l'instrument à partir du volume échographique est réalisée en deux étapes [Stoll 2006]. Un algorithme basé sur une version modifiée de la transformée de Radon est proposé pour extraire la droite 3D représentant l'axe principal de l'instrument. La rotation propre de l'instrument autour de son axe ainsi que la position de l'extrémité de l'outil sont ensuite estimées grâce à l'utilisation d'un marqueur passif matérialisé par un manchon cylindrique qui est fixé sur l'extrémité de l'outil. Ce manchon est composé de trois anneaux orientés avec des angles différents dont les valeurs sont connues. Une

coupe 2D obtenue par interpolation le long de l'axe de l'instrument fournit alors une section du manchon où 3 pics correspondant aux sections des anneaux sont détectés au moyen d'un algorithme de détection de motifs (template matching). Les distances relatives des pics sont ensuite utilisées pour déterminer l'orientation propre de l'outil et la position de son extrémité. L'extraction de la cible en forme de croix est également obtenue à l'aide de la transformée de Radon. Ces données de position sont ensuite exprimées dans un repère du robot et un correcteur proportionnel dérivé est appliqué pour minimiser l'erreur de position relative entre la cible et l'outil.

2.4 Positionnement de l'étude

L'utilisation d'une sonde échographique 3D ou d'un dispositif fournissant simultanément plusieurs coupes 2D permet de mettre en œuvre un asservissement visuel 3D contrôlant l'ensemble des degrés de liberté d'un outil ou de la sonde à partir d'une estimation de pose. Néanmoins, comme nous l'avons déjà indiqué, les approches d'asservissement visuel 3D utilisant la pose sont sensibles aux erreurs de reconstruction dues aux bruits de mesure et surtout aux erreurs de calibration du système. Il en résulte une faible précision de positionnement puisque le contrôle est réalisé sur des mesures biaisées. Ce défaut est encore plus prononcé pour une configuration déportée du capteur puisqu'il est alors nécessaire d'avoir une connaissance précise de la pose relative entre le repère du capteur dans lequel sont exprimées les mesures et le repère associé au robot où est réalisé le contrôle.

Au contraire, les asservissements visuels 2D offrent une meilleure robustesse à d'éventuelles erreurs de calibration, et permettent d'atteindre ainsi de meilleures précisions de positionnement puisque la tâche à réaliser est directement définie dans l'image. Cependant la difficulté de mettre en œuvre cette approche réside dans la détermination de la variation des informations visuelles induite par la translation et les deux rotations en dehors du plan de coupe. Pour cette raison la majorité des travaux existants s'est limitée à contrôler exclusivement le mouvement de la sonde ou d'une aiguille dans le plan échographique à l'exception de deux travaux basés sur une modélisation géométrique de l'interaction entre un instrument de forme 3D connue et le plan de coupe. En revanche, cette approche n'a jamais été considérée dans une configuration embarquée pour contrôler le mouvement complet de la sonde. En effet, la variation de l'information visuelle due au mouvement en dehors du plan d'observation dépend de la forme 3D de l'objet observé, ce qui rend difficile la détermination de la matrice d'interaction uniquement à partir d'informations visuelles extraites de l'image. De plus, dans la configuration embarquée, l'information visuelle observée est principalement de nature anatomique et contrairement aux applications de manipulation d'un instrument, il n'est généralement pas possible d'utiliser une connaissance a priori sur la géométrie de l'élément anatomique observé.

Le contrôle de la totalité des mouvements d'une sonde échographique par asservissement visuel échographique 2D est donc un défi majeur. C'est justement ce challenge que nous avons proposé de relever via des méthodes génériques capables de contrôler les mouvements d'une sonde dans et hors de son plan d'observation.

2.5 Conclusion

Ce chapitre nous a permis de mettre en évidence la problématique associée à l'utilisation de l'image fournie par un capteur échographique. En effet, il est important de noter qu'une sonde échographique 2D a la particularité de fournir une information complète dans le plan d'observation du capteur sans toutefois fournir d'information en dehors de ce plan. Une caméra agit d'une manière totalement différente. Elle fournit une projection de l'environnement perçu vers une image 2D. Par conséquent, les méthodes classiques d'asservissement visuel 2D fondées pour la plupart sur l'utilisation d'un modèle du sténopé de la caméra ne peuvent pas être appliquées directement dans le cas de l'imagerie échographique.

En outre, une problématique importante concerne l'extraction en temps réel des informations visuelles, nécessaires à la commande du système. En effet, les images échographiques sont par nature très fortement bruitées en raison du phénomène de diffusion de l'onde ultrasonore dans les tissus qui a pour effet de générer une texture granulaire (speckle) dans l'image.

Il s'agit donc d'élaborer des solutions robustes permettant de contrôler l'ensemble des degrés de liberté de la sonde en vue de positionner cette dernière par rapport à l'environnement. Les deux chapitres suivants traiteront de solutions d'asservissement visuel en considérant respectivement des informations visuelles géométriques et des approches fondées sur l'information dense de l'image.

Asservissement visuel échographique : approches utilisant des informations géométriques

Dans le chapitre précédent, nous avons fait la distinction entre les mouvements dans le plan et les mouvements hors du plan de la sonde. Le contrôle de l'ensemble de ces mouvements peut être réalisé sans trop de difficultés en considérant un asservissement visuel 3D, comme cela a été présenté dans l'état de l'art. Dans ce cas, une reconstruction de pose peut être effectuée à partir de mesures extraites d'un volume échographique 3D [Novotny 2007] ou de plusieurs coupes 2D [Lee 2007]. Néanmoins, une telle approche ne permet pas de définir l'erreur à minimiser directement dans l'image et il en résulte alors à la convergence un biais de positionnement dû à la fois aux erreurs d'étalonnage du système et au bruit de mesures. En revanche, une approche basée sur un asservissement visuel 2D permet d'atteindre des précisions de l'ordre du pixel puisque l'erreur de positionnement est directement exprimée dans l'image.

Ainsi, nous nous sommes orientés vers l'étude d'approches d'asservissement visuel 2D utilisant des informations géométriques en vue de contrôler les 6 ddl d'une sonde échographique embarquée sur un manipulateur robotique.

En outre, l'obtention d'un asservissement visuel robuste est conditionnée par la qualité d'extraction des primitives géométriques de l'image. A cet effet, il est nécessaire de porter une attention particulière sur les méthodes et algorithmes de traitement d'image inclus dans la boucle de contrôle, tout en respectant les contraintes temps réel du système.

Ce chapitre présente différentes solutions basées sur l'information de type point (section 3.1), contour (section 3.2) et moments géométriques (section 3.3) pour le positionnement d'une sonde 2D conventionnelle. Une approche tri-plans est proposée à la section 3.4 pour la commande d'un capteur échographique 3D par asservissement visuel 2D. La section 3.5 s'intéresse quant à elle à la problématique d'une segmentation robuste et faiblement complexe en présentant une solution adaptée à l'imagerie échographique pour extraire des informations visuelles géométriques.

Enfin, chaque solution est validée par une application démontrant les performances des approches proposées. En particulier, des fonctionnalités d'assistance au diagnostic ont été intégrées sur un système de télé-échographie (section 3.6).

3.1 Approche pour l'information de type « point »

Les informations géométriques naturellement les plus simples sont directement les points d'intersection de deux droites remarquables d'un objet d'intérêt. Cette section s'appuie donc sur cette information pour élaborer un asservissement visuel 2D en vue de positionner la sonde par rapport à cet objet.

3.1.1 Modélisation de l'interaction de la coupe avec une droite 3D

Dans un premier temps, il s'agit d'automatiser une procédure d'étalonnage des paramètres intrinsèques et extrinsèques de la sonde. La solution consiste à considérer l'information visuelle de type point correspondant à l'intersection d'une droite 3D traversant le plan de coupe d'une sonde échographique 2D [Krupa 2005]. Du point de vue applicatif, l'idée est d'automatiser au moyen d'un asservissement visuel 2D le positionnement de la sonde par rapport à un marqueur artificiel immergé dans un bac d'eau. Ce marqueur est constitué de deux droites \mathcal{D}_1 et \mathcal{D}_2 matérialisées par des fils de nylon tendus qui se croisent en un point fixe P^* de l'espace et dont les points P_1 et P_2 correspondent aux points d'intersection avec le plan de coupe, comme illustré sur la figure 3.1. A partir de ce modèle, la matrice d'interaction, qui fournit la vitesse des coor-

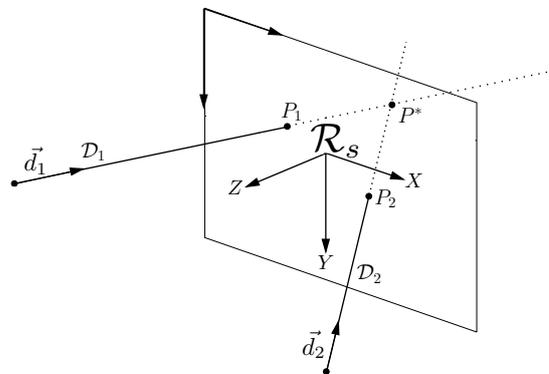


FIGURE 3.1 – Intersection de droites avec le plan de coupe

données dans l'image du point d'intersection d'une droite en fonction du torseur cinématique v de la sonde échographique, s'exprime selon

$$\mathbf{L}_p = \begin{bmatrix} -1 & 0 & \frac{d_x}{d_z} & \frac{d_x}{d_z}y & -\frac{d_x}{d_z}x & y \\ 0 & -1 & \frac{d_y}{d_z} & \frac{d_y}{d_z}y & -\frac{d_y}{d_z}x & -x \end{bmatrix}. \quad (3.1)$$

Cette matrice dépend donc des coordonnées (x, y) dans l'image du point d'intersection et également des composantes du vecteur unitaire ${}^s\mathbf{d} = (d_x, d_y, d_z)$ de la droite exprimées dans le

repère de la sonde \mathcal{R}_s (figure 3.1). Une connaissance de l'orientation de la droite est donc nécessaire pour lier la variation du point dans l'image aux mouvements hors plan de la sonde. A partir de cette modélisation de l'interaction, nous avons mis en œuvre un asservissement visuel échographique 2D en vue de positionner la sonde de manière à amener les points d'intersection des deux droites sur des consignes définies dans l'image [Krupa 2006b]. Dans notre approche, le vecteur unitaire de chaque droite est estimé en ligne par une méthode des moindres carrés. Celle-ci utilise en entrée les coordonnées 3D des points des droites qui sont reconstruites, lors du déplacement de la sonde en dehors de son plan, à partir des mesures 2D dans l'image et de l'odométrie du robot. Cet asservissement a permis d'automatiser la procédure d'étalonnage présentée ci-après.

3.1.2 Application à une procédure d'étalonnage de la sonde

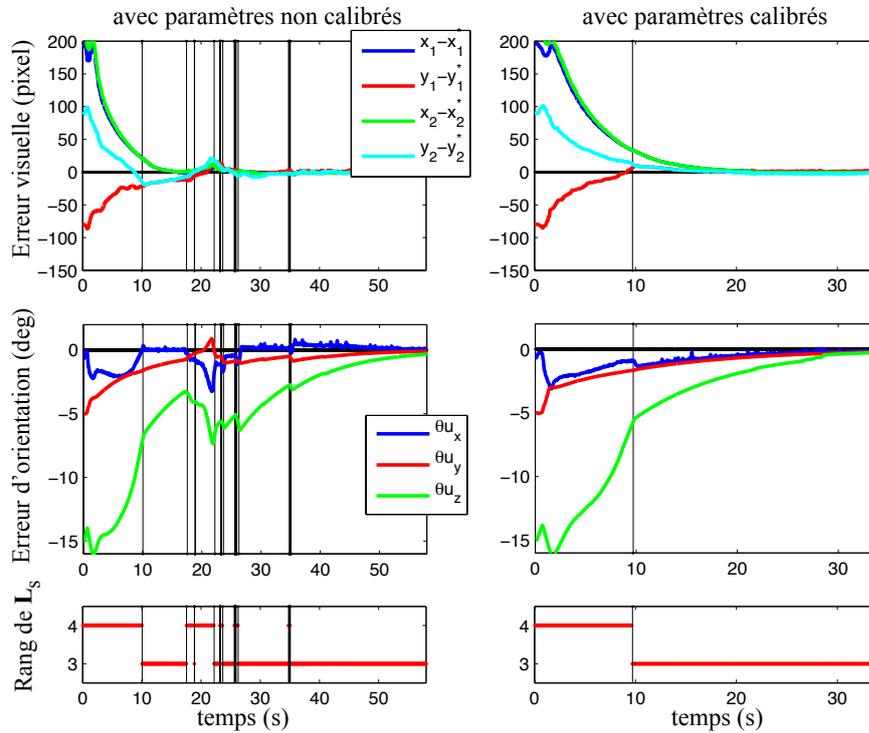
Comme nous l'avons signalé dans la section 2.1.1, la connaissance des paramètres intrinsèques et extrinsèques du capteur échographique est indispensable pour la réalisation d'une imagerie échographique 3D avec une sonde 2D. Cette technique d'imagerie, également appelée échographique 3D de type « main libre » [Nelson 1998], est une méthode alternative pour reconstruire l'information de volume. Elle présente l'avantage de ne pas utiliser de sonde échographique 3D dont les principales limitations sont la faible taille des volumes acquis et leur faible résolution spatiale. Elle consiste à déplacer une sonde échographique 2D classique munie d'un système de localisation sur la zone d'intérêt du patient à examiner. Les mesures de pose de chaque coupe échographique fournies par le système de localisation permettent ensuite de recalculer les coupes échographiques enregistrées dans un repère de référence 3D, afin de reconstruire l'information tridimensionnelle. Les systèmes de localisation utilisés sont soit magnétiques [Pagoulatos 1999], soit optiques [Pagoulatos 2000] ou mécaniques et ont pour contrainte de fournir en continu la position et l'orientation de la sonde. Un robot porteur de sonde peut également être utilisé pour fournir grâce à ses mesures odométriques la pose de la sonde par rapport à un repère fixe. Le robot Hippocrate [Pierrot 1999], utilisé en cardiologie vasculaire pour effectuer l'imagerie 3D d'artères dans le but de détecter ou suivre l'évolution de dépôts graisseux, s'appuie sur un tel système.

Dans tous les cas, quel que soit le système de localisation utilisé pour mesurer la pose de la sonde, il est indispensable d'étalonner les paramètres intrinsèques et extrinsèques de la sonde. Les paramètres intrinsèques correspondent aux 2 facteurs d'échelle de l'image nécessaires à la conversion des coordonnées pixéliques en coordonnées métriques. Les paramètres extrinsèques correspondent quant à eux aux 3 translations et aux 3 angles de rotation qui caractérisent la transformation rigide entre le repère de l'image et le repère du système de localisation attaché à la sonde (effecteur dans le cas d'un robot porteur de sonde). Plusieurs méthodes d'étalonnage ont été proposées dans la littérature. Un état de l'art récent est présenté dans [Hsu 2009]. Toutes ces méthodes reposent sur un même principe. Il consiste à observer un objet géométrique de forme connue avec différentes positions et orientations de la sonde. Une fonction de coût est ensuite définie à partir de l'erreur entre les mesures visuelles de points caractéristiques observés dans l'ensemble des images et la prédiction de ces mêmes points fournis par un modèle géométrique. Cette prédiction dépend d'une connaissance de l'objet et met en jeu les paramètres d'étalonnage de la sonde. Un algorithme d'optimisation calcule alors les paramètres qui minimisent au mieux cette fonction de coût non linéaire. Les objets considérés sont

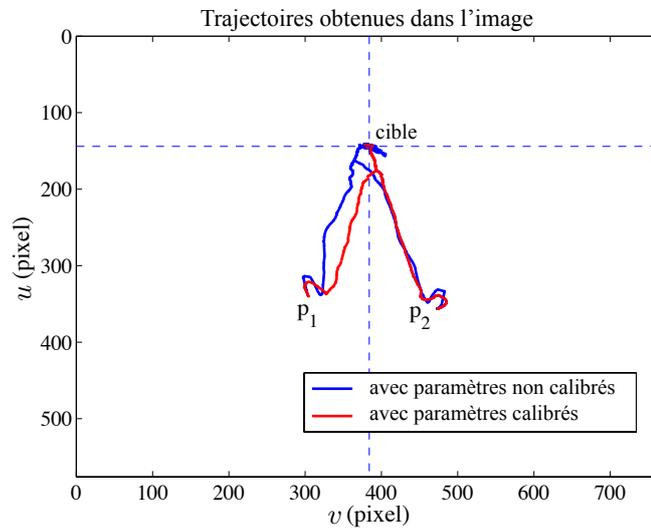
généralement de forme simple pour faciliter d'une part le traitement d'image nécessaire à leur détection dans l'image échographique, et d'autre part, pour éviter d'introduire dans l'estimation des paramètres des erreurs dues à une mauvaise modélisation de l'objet. Par exemple, dans [Rousseau 2005], l'objet considéré est un plan matérialisé par une plaque de plexiglas et la mesure visuelle correspond aux paramètres de la droite 2D résultante de l'intersection de ce plan avec le plan de coupe de la sonde. La méthode que nous avons retenue repose sur l'utilisation d'un simple point 3D matérialisé par l'intersection de deux fils de nylon tendus et immobiles dans une bassine d'eau [Detmer 1994]. L'observation de ce point avec différentes positions et orientations de la sonde pose toutefois une difficulté à l'utilisateur car un positionnement précis de la sonde doit être effectué pour maintenir le point 3D dans le plan de coupe. Cette méthode d'estimation, bien que simple en théorie, se révèle donc très fastidieuse à mettre en œuvre en pratique lors de la collecte des mesures si la sonde est positionnée manuellement par un utilisateur. En revanche, dans le cas où la sonde est actionnée par un robot manipulateur, ce problème de positionnement devient une tâche qu'un asservissement visuel peut réaliser adroitement. C'est justement dans ce contexte applicatif que nous avons proposé un asservissement visuel échographique 2D utilisant des primitives visuelles de type point, en s'appuyant sur la définition de la matrice d'interaction donnée par l'équation (3.1).

Positionner le plan de coupe de la sonde sur le point d'intersection P^* des fils modélisés par deux droites revient à amener les points d'intersection P_1 et P_2 vers un même point cible dans l'image (figure 3.1). Par conséquent, les coordonnées dans l'image $\mathbf{s} = (x_1, y_1, x_2, y_2)$ des deux points d'intersection avec le plan de coupe ont été retenues comme primitives visuelles de commande, et la consigne a été fixée à une position donnée $\mathbf{s} = (x^*, y^*, x^*, y^*)$. La matrice d'interaction de dimension 4×6 associée à \mathbf{s} est obtenue dans ce cas en empilant simplement les matrices d'interaction (3.1) associées à chaque point. Etant donné que cette tâche visuelle de dimension 4 ne contraint pas les 6 degrés de liberté de la sonde, une infinité de positions de la sonde peuvent être obtenues à la convergence. En outre, à la position désirée dans l'image les points étant confondus, le rang de la matrice d'interaction initialement à 4 chute à 3 : de ce fait, seuls 3 ddl restent contraints par la tâche visuelle lorsque celle-ci a convergé. Avec les 3 ddl restants, il est donc possible d'effectuer une tâche secondaire en appliquant le formalisme de la redondance introduit à la section 2.2.5. Afin d'automatiser la procédure de collecte des mesures, nous avons ainsi mis en œuvre une tâche secondaire dont l'objectif est d'amener la sonde à une orientation donnée dans l'espace tout en garantissant la réalisation de la tâche principale de vision [Krupa 2006a]. Cette tâche secondaire est définie à partir d'une fonction de coût $h = \frac{1}{2} \theta \mathbf{u}^T \theta \mathbf{u}$, où $\theta \mathbf{u}$ représente l'orientation relative entre le repère courant de la sonde et son orientation désirée à atteindre. L'application d'une séquence de consignes dans l'image avec plusieurs orientations de la sonde a ainsi permis d'automatiser la campagne de mesures nécessaires à l'estimation des paramètres de la sonde.

La figure 3.2a présente l'évolution de l'erreur visuelle, l'erreur d'orientation de la sonde ainsi que la valeur du rang de la matrice d'interaction lors d'un positionnement de la sonde par la méthode proposée avec ou sans les paramètres calibrés de la sonde. Les trajectoires des points obtenues dans l'image sont présentées à la figure 3.2b. Comme prévu, l'asservissement visuel 2D a l'avantage d'être robuste aux erreurs d'étalonnage, et réalise la tâche visuelle avec des valeurs approximatives des paramètres intrinsèques et extrinsèques de la sonde. En effet, dans les deux cas (sonde calibrée ou non calibrée) les points convergent vers la cible et l'orientation désirée de la sonde est atteinte.



(a)



(b)

FIGURE 3.2 – (a) Evolution de l'erreur visuelle $s - s^*$ (ligne du haut), erreur d'orientation de la sonde θu (ligne du milieu) et rang de \hat{L}_s (ligne du bas) sans (colonne de gauche) ou avec (colonne de droite) étalonnage des paramètres de la sonde ; (b) Trajectoires des points dans l'image avec ou sans paramètres calibrés de la sonde

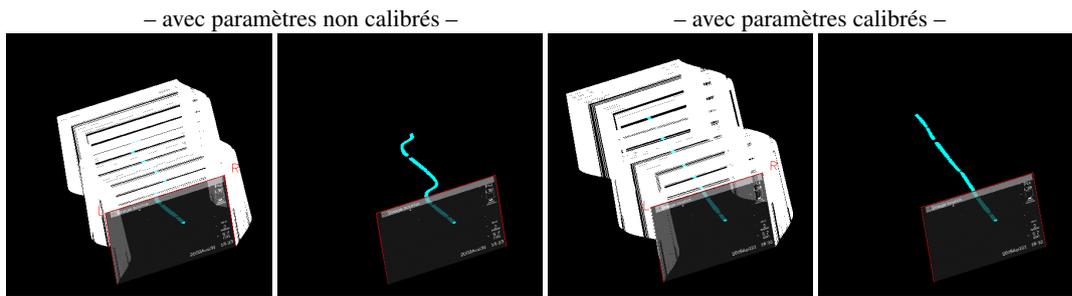


FIGURE 3.3 – Reconstruction 3D d'un fil de nylon tendu (visualisation avec le logiciel Stradx [Prager 1999])

Deux exemples de reconstructions 3D, obtenues à partir d'une même série de coupes observant un fil de nylon tendu sont présentés à la figure 3.3 en utilisant des paramètres de la sonde non-calibrés ou estimés par la méthode de Detmer [Detmer 1994], pour laquelle la campagne de mesures a été automatisée par asservissement visuel échographique. Ces résultats confirment l'efficacité de la procédure d'étalonnage car la forme du fil de nylon, reconstruite à partir des paramètres estimés de la sonde, représente réellement une droite dans l'espace contrairement à celle obtenue en considérant des paramètres approchés.

3.2 Approche utilisant le contour de la section observée

L'approche précédente a démontré sa validité pour le positionnement de la sonde vis-à-vis d'un objet constitué de droites. Toutefois, pour des objets plus complexes, il est nécessaire de s'appuyer sur des informations géométriques plus élaborées. A cet effet, le contour d'une section a d'abord été considéré.

Ainsi dans une étude préliminaire, menée dans le cadre du stage de Master Recherche de Wael Bachta, nous avons proposé une méthode d'asservissement visuel échographique 2D pour réaliser le positionnement automatique de la sonde, selon 6 ddl, sur une section désirée d'un objet immobile dont la géométrie 3D est a priori connue au moyen d'une imagerie pré-opératoire [Bachta 2006]. L'étude a porté sur un objet ovoïde, car il représente géométriquement un grand nombre de kystes ou tumeurs rencontrés en pratique. Une formulation implicite de la surface de l'ovoïde est donnée par l'expression

$$\frac{{}^o x^2}{a^2} + \frac{{}^o y^2}{b^2} \left(1 + \frac{d^2 + 2 d {}^o x}{a^2}\right) + \frac{{}^o z^2}{c^2} \left(1 + \frac{d^2 + 2 d {}^o x}{a^2}\right) = 1, \quad (3.2)$$

où $({}^o x, {}^o y, {}^o z)$ sont les coordonnées 3D des points de la surface exprimées dans un repère \mathcal{R}_o associé à l'objet, et $\Theta = (a, b, c, d)$ sont les paramètres géométriques constants supposés connus de l'ovoïde. A partir de ce modèle de surface, nous avons déterminé l'expression mathématique du contour de la section observée dans le plan de coupe de la sonde. Ainsi,

$$x^3 + \sum_{i,j \geq 0, i+j \leq 3, i < 3} a_{ij} x^i y^j = 0, \quad (3.3)$$

où (x, y) représentent les coordonnées 2D des points du contour dans l'image et a_{ij} des coefficients de pondération. Cette expression est un polynôme d'ordre 3, dont les 9 coefficients a_{ij} sont dépendants des paramètres Θ de l'ovoïde et également de la pose relative courante du repère de l'objet \mathcal{R}_o exprimée dans le repère de la sonde \mathcal{R}_s . Cette pose relative est décrite par la matrice homogène ${}^s\mathbf{M}_o$. Nous avons proposé dans cette étude théorique de considérer les 9 coefficients a_{ij} en tant que mesures visuelles \mathbf{s} et de déterminer analytiquement l'expression de la matrice d'interaction \mathbf{L}_s associée à ce choix. De ce fait, on a

$$\dot{\mathbf{s}} = \begin{bmatrix} \dot{a}_{ij} \end{bmatrix}_{i,j \geq 0, i+j \leq 3, i < 3} = \mathbf{L}_s(\Theta, {}^s\mathbf{M}_o)\mathbf{v}, \quad (3.4)$$

qui est obtenue en dérivant \mathbf{s} par rapport au temps. Bien entendu, \mathbf{L}_s est également dépendante des paramètres de l'ovoïde et de la pose relative entre l'objet et la sonde. Bien que les paramètres constants Θ soient supposés connus au moyen d'une imagerie 3D préalable de l'objet, il est toutefois nécessaire d'obtenir une mesure en ligne de ${}^s\mathbf{M}_o$ pour déterminer la valeur courante de la matrice d'interaction. En formulant la pose relative par ${}^s\mathbf{M}_o = {}^s\mathbf{M}_b {}^b\mathbf{M}_o$, où ${}^s\mathbf{M}_b$ fournie par l'odométrie du robot décrit la pose du repère de la base du robot \mathcal{R}_b dans le repère de la sonde, le problème peut alors se simplifier à l'estimation d'une transformation constante ${}^b\mathbf{M}_o$ entre le repère de l'objet et le repère de la base. La méthode, que nous avons mise en oeuvre pour estimer ${}^b\mathbf{M}_o$, est basée sur le même principe que celui sur lequel sont fondées les différentes méthodes d'étalonnage de sonde. Elle consiste à déterminer, lors d'une procédure hors-ligne, à partir d'un ensemble de coupes différentes de l'objet, la pose ${}^b\mathbf{M}_o$ qui minimise au mieux les erreurs observées entre les mesures visuelles \mathbf{s} et leurs prédictions fournies par le modèle du contour exprimé par l'expression (3.3).

Concernant la mise en oeuvre de la commande, étant donné que le nombre d'informations visuelles $k = 9$ est supérieur aux 6 dll du robot, nous avons considéré une matrice de combinaison $\mathbf{C} = \widehat{\mathbf{L}}_{\mathbf{s}|\mathbf{s}=\mathbf{s}^*}^+$ dans la définition de la fonction de tâche et appliqué la commande cinématique (2.15) où une estimation de la valeur courante de la matrice d'interaction est considérée. La consigne \mathbf{s}^* a été définie à partir du contour observé à la position désirée à atteindre. Les résultats de simulation de cette tâche de positionnement sont présentés à la figure 3.4. Nous pouvons observer une décroissance exponentielle de l'erreur visuelle vers zéro qui confirme le succès de la tâche.

Cette étude nous a permis de vérifier la faisabilité théorique d'un asservissement visuel contrôlant les 6 dll d'une sonde échographique 2D en vue de la positionner sur une coupe désirée d'un objet. Bien que les informations visuelles aient convergé vers la consigne désirée, le choix des coefficients du polynôme décrivant le contour de la section ne s'est pas révélé optimal dans la mesure où il ne représente pas des informations géométriques au sens physique de la section observée tels son barycentre, son orientation principale, sa surface... En effet, comme la minimisation de l'erreur visuelle est réalisée au niveau des coefficients a_{ij} , le déplacement du contour dans l'image ne suit pas une ligne droite, et il en résulte également une trajectoire complexe réalisée par la sonde dans l'espace. Il se pose donc le problème du maintien de la visibilité : la section observée peut en effet sortir de la zone utile de l'image lors de la phase de positionnement.

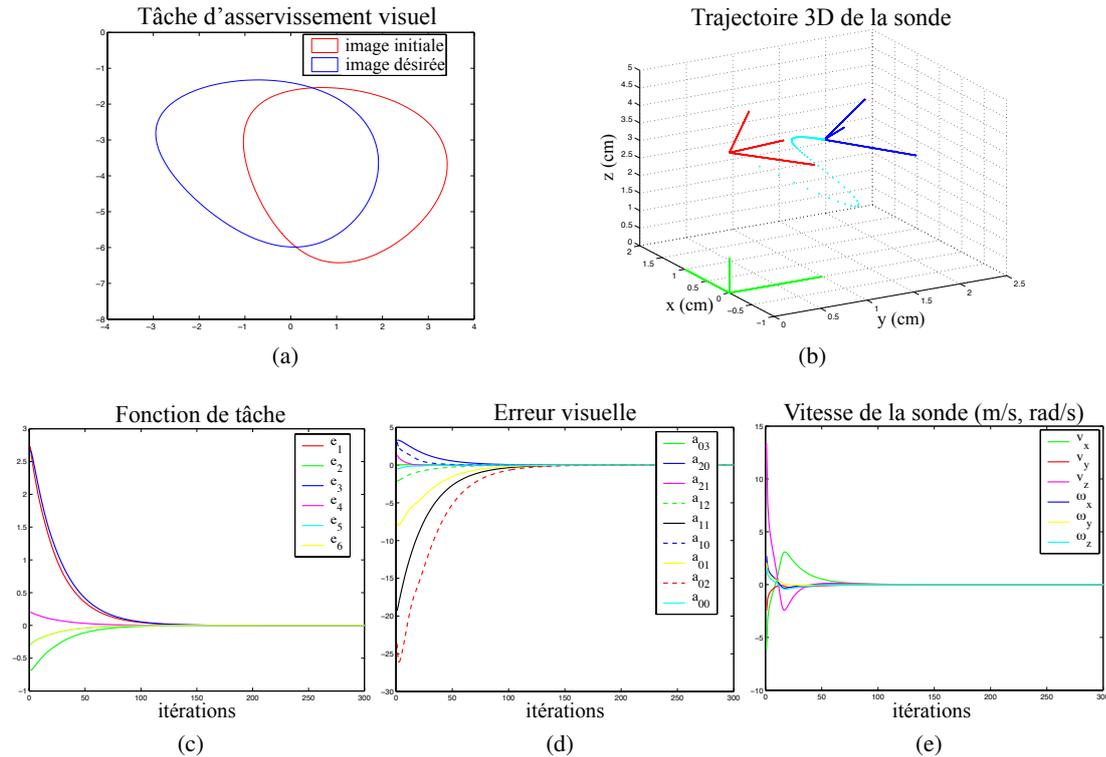


FIGURE 3.4 – Simulation de la tâche de positionnement. (a) section initiale (rouge) et section désirée et atteinte à la convergence (bleue) ; (b) trajectoire 3D de la sonde réalisée lors de la tâche ; (c) évolution de la fonction de tâche en fonction des itérations de la commande ; (d) évolution de l'erreur visuelle $s - s^*$; (e) vitesse appliquée à la sonde.

Nous avons par conséquent, dans le cadre des travaux de thèse de Rafik Mebarki [Mebarki 2010a], orienté le choix des informations visuelles sur des primitives visuelles construites à partir des moments géométriques 2D de la section, qui sont plus à même de caractériser la position et la forme de la section dans l'image.

3.3 Approche basée sur les moments 2D de la section observée

Au vu des limites de la méthode basée sur les coefficients du polynôme décrivant le contour, l'utilisation dans la commande de primitives visuelles définies à partir des moments 2D de la section observée de l'objet devrait apporter une réponse quant à la problématique de l'évolution de la section dans l'image. En effet, les moments de l'image sont des informations qui ont l'avantage d'être génériques quant à la caractérisation de formes 2D très variées. Les moments de faible ordre ont un sens géométrique intuitif puisqu'ils représentent directement l'aire, le barycentre et l'orientation principale dans l'image de l'objet d'intérêt. Etant obtenus par intégration et plus particulièrement par une somme discrétisée sur la section observée dans l'image, ils ont l'avantage d'être peu sensibles aux bruits de l'image et semblent donc appropriés pour

une utilisation avec des images échographiques.

Les moments de l'image ont été largement utilisés dans le domaine de la vision par ordinateur notamment pour des applications de reconnaissance de formes [Mukundan 1998, Prokop 1992], où des fonctions de moments invariantes aux mouvements de translation et rotation dans l'image ont été élaborées [Hu 1962]. Ils ont également été introduits en asservissement visuel, comme cela a été évoqué à la section 2.2.6, où 6 primitives visuelles construites à partir d'une combinaison des moments 2D ont été proposées dans la commande du système pour réaliser la tâche de positionnement d'une caméra [Chaumette 2004, Tahri 2005]. Dans ce travail, des moments invariants ont été proposés pour obtenir une matrice d'interaction de forme quasi triangulaire conduisant ainsi à un découplage important de la commande du système. Néanmoins, la modélisation de la matrice d'interaction associée aux moments ainsi que la combinaison des moments retenue comme primitives visuelles s'ont été déterminées pour le seul cas de l'utilisation d'une caméra basée sur le modèle projectif.

Par conséquent, afin de rendre possible l'utilisation des moments 2D dans un asservissement visuel échographique, il s'agit dans un premier temps, de modéliser la matrice d'interaction permettant d'exprimer leur variation en fonction du torseur cinématique de la sonde [Mebarki 2008a, Mebarki 2009].

3.3.1 Modélisation de la matrice d'interaction

Supposons à présent que nous disposons d'une image binaire représentant la section fermée \mathcal{S} d'un objet observé dans le plan de coupe de la sonde qui a été extraite par un algorithme de traitement d'images. Nous détaillerons plus tard dans la section 3.5 la méthode de segmentation utilisée en pratique. Par définition, les moments 2D d'ordre $i + j$ de cette section binaire sont donnés par l'expression

$$m_{ij} = \iint_{\mathcal{S}} x^i y^j dx dy, \quad (3.5)$$

où (x, y) sont les coordonnées dans l'image des points appartenant à la section \mathcal{S} . En appliquant le théorème de Green [Stewart 1991], qui établit une relation entre une intégrale curviligne le long d'un contour fermé C et une intégrale double sur la section \mathcal{S} délimitée par ce contour C , nous pouvons reformuler l'expression des moments 2D à partir d'une intégrale curviligne sur le contour selon

$$m_{ij} = \frac{-1}{j+1} \oint_C x^i y^{j+1} dx = \frac{1}{i+1} \oint_C x^{i+1} y^j dy. \quad (3.6)$$

La détermination de la matrice d'interaction $\mathbf{L}_{\mathbf{m}_{ij}}$ associée au moment d'ordre $i + j$ consiste alors à calculer la dérivée temporelle de m_{ij} (équation 3.5). D'après le théorème de Green, on a alors

$$\dot{m}_{ij} = - \oint_C x^i y^j \dot{y} dx + \oint_C x^i y^j \dot{x} dy. \quad (3.7)$$

Cette dérivée étant fonction de la variation temporelle (\dot{x}, \dot{y}) des coordonnées des points du contour C , la vitesse exprimée dans le plan de coupe d'un point du contour en fonction du torseur cinématique de la sonde s'écrit

$$\begin{cases} \dot{x} = -v_x - K_x v_z - y K_x \omega_x + x K_x \omega_y + y \omega_z \\ \dot{y} = -v_y - K_y v_z - y K_y \omega_x + x K_y \omega_y - x \omega_z, \end{cases} \quad (3.8)$$

où les scalaires K_x et K_y se déduisent du vecteur normal à la surface de l'objet $\nabla\mathbf{F} = (f_x, f_y, f_z)$ exprimé dans le repère de la sonde en chaque point du contour, tels que

$$\begin{cases} K_x &= f_x f_z / (f_x^2 + f_y^2) \\ K_y &= f_y f_z / (f_x^2 + f_y^2). \end{cases} \quad (3.9)$$

Finalement, en substituant (3.8) dans (3.7), après identification des différents termes, l'expression analytique de la matrice d'interaction de dimension 1×6 associée au moment m_{ij} est donnée par

$$\mathbf{L}_{m_{i,j}} = [m_{v_x} \ m_{v_y} \ m_{v_z} \ m_{\omega_x} \ m_{\omega_y} \ m_{\omega_z}], \quad (3.10)$$

avec

$$\begin{cases} m_{v_x} &= -i.m_{i-1,j} \\ m_{v_y} &= -j.m_{i,j-1} \\ m_{v_z} &= {}^x m_{i,j} - {}^y m_{i,j} \\ m_{\omega_x} &= {}^x m_{i,j+1} - {}^y m_{i,j+1} \\ m_{\omega_y} &= -{}^x m_{i+1,j} + {}^y m_{i+1,j} \\ m_{\omega_z} &= i.m_{i-1,j+1} - j.m_{i+1,j-1}. \end{cases} \quad (3.11)$$

Les composantes $(m_{v_x}, m_{v_y}, m_{\omega_z})$ relatives aux mouvements dans le plan de la sonde dépendent donc uniquement des moments d'ordre $i + j - 1$ et $i + j$ qui peuvent être directement mesurés dans l'image. En revanche, les autres composantes $(m_{v_z}, m_{\omega_x}, m_{\omega_y})$, relatives aux mouvements hors du plan de la sonde, sont dépendantes de moments modifiés ${}^x m_{i,j}$ et ${}^y m_{i,j}$, définis par

$$\begin{cases} {}^x m_{ij} &= \oint_C x^i y^j K_y dx \\ {}^y m_{ij} &= \oint_C x^i y^j K_x dy. \end{cases} \quad (3.12)$$

Ces derniers font intervenir le vecteur normal à la surface de l'objet pour chaque point du contour. Une connaissance 3D de l'objet et plus particulièrement du vecteur normal $\nabla\mathbf{F}$ à sa surface est donc indispensable pour la mise en œuvre d'un asservissement visuel contrôlant également les mouvements hors plan de la sonde.

3.3.2 Interaction avec un objet de forme ellipsoïdale

Afin de valider le modèle précédent, nous avons tout d'abord étudié le cas où une formulation analytique de la forme de l'objet est disponible [Mebarki 2008a], en considérant un ellipsoïde dont la surface est représentée par la fonction

$$F({}^o x, {}^o y, {}^o z) = \left(\frac{{}^o x}{a_1}\right)^2 + \left(\frac{{}^o y}{a_2}\right)^2 + \left(\frac{{}^o z}{a_3}\right)^2 - 1 = 0, \quad (3.13)$$

où ${}^o \mathbf{x} = ({}^o x, {}^o y, {}^o z)$ sont les coordonnées 3D des points de la surface exprimées dans le repère de l'objet \mathcal{R}_o et a_1, a_2, a_3 représentent les demi-longueurs des axes principaux qui sont supposées connues au moyen d'une imagerie 3D préalable. Le vecteur normal à la surface est dans ce cas directement donné par le gradient normalisé de la surface au point considéré, soit ${}^o \nabla\mathbf{F} = \frac{\nabla F}{\|\nabla F\|}$. Il est alors possible de déterminer son expression ${}^s \nabla\mathbf{F}$ pour chaque point du contour ${}^s \mathbf{x} = (x, y, 0)$ observé dans le repère de la sonde \mathcal{R}_s à partir de la transformation homogène ${}^s \mathbf{M}_o$ décrivant la pose du repère de l'objet dans le repère de la sonde. Bien entendu,

une mesure en ligne de cette pose relative est nécessaire, et peut s'obtenir avec le même principe que celui décrit dans la section 3.2.

Etant donné que le contour de la section résultante de l'intersection d'un objet ellipsoïdal avec le plan de coupe correspond dans tous les cas à une ellipse, uniquement 5 primitives visuelles peuvent être utilisées pour caractériser cette dernière. Nous avons ainsi proposé le vecteur de primitives visuelles suivant en entrée de l'asservissement visuel

$$\mathbf{s} = (x_g, y_g, \alpha, l_1, a), \quad (3.14)$$

où les coordonnées 2D du barycentre (x_g, y_g) , l'angle principal de la section dans l'image α , la longueur de son axe principal l_1 ainsi que son aire a sont des primitives géométriques définies à partir des moments d'ordre 0 à 2 extraits de la section de l'image. La matrice d'interaction \mathbf{L}_s associée à ces primitives se dérive alors directement à partir de celles des moments $\mathbf{L}_{m_{i,j}}$ (3.10). Nous pouvons dès à présent noter que les coordonnées du barycentre et l'orientation principale de la section sont fortement appropriées pour contrôler les déplacements dans le plan de la sonde. En effet, leurs variations sont très sensibles à ces mouvements tout en étant très peu couplées entre elles. Réciproquement, l'aire et la longueur de l'axe principal de la section sont des primitives adaptées pour le contrôle des déplacements hors du plan car elles sont invariantes aux mouvements dans le plan.

La figure 3.5 présente les résultats expérimentaux obtenus en appliquant la loi de commande (2.17) pour une tâche de positionnement de la sonde sur une section désirée d'un cœur de lapin plongé dans une bassine d'eau [Mebarki 2008b]. Bien que la forme du cœur soit différente du modèle de l'ellipsoïde considéré pour la détermination de la matrice d'interaction utilisée dans la commande, les primitives visuelles ont convergé vers leurs valeurs de consigne, ce qui démontre la robustesse de l'asservissement visuel 2D aux erreurs de modélisation. En revanche, la position à laquelle a été apprise la section désirée n'est pas atteinte à la convergence (figure 3.5f) car les 5 primitives visuelles ne permettent pas de contraindre les 6 dll du robot ($k < n$). En effet, il existe une infinité de positions de la sonde qui conduisent à observer dans l'image une section caractérisée par les mêmes paramètres désirés de l'ellipse.

Par conséquent, si l'objectif de la tâche d'asservissement visuel est de réaliser un positionnement mettant en jeu les 6 dll de la sonde afin d'obtenir à la convergence une liaison virtuelle rigide entre la sonde et l'objet d'intérêt et permettre ainsi une compensation automatique du mouvement de l'objet, 6 primitives visuelles non dépendantes entre elles sont à définir. Bien entendu, dans ce cas, un objet de forme simple tel un ellipsoïde n'est pas approprié. Nous avons donc étendu notre approche utilisant les moments de la section pour le cas d'objets de forme plus complexe.

3.3.3 Interaction avec un objet de forme quelconque

Dans l'approche précédente, une formulation analytique des composantes du vecteur normal ∇F a été fournie pour calculer les moments modifiés (3.12) intervenant dans l'expression des composantes relatives aux mouvements hors du plan de la matrice d'interaction $\mathbf{L}_{m_{i,j}}$. Cependant, cette formulation analytique peut se dériver exclusivement à partir d'un modèle géométrique de l'objet et de la pose de ce dernier par rapport à la sonde. Pour s'affranchir de ces connaissances géométriques et considérer ainsi des objets de forme quelconque, nous

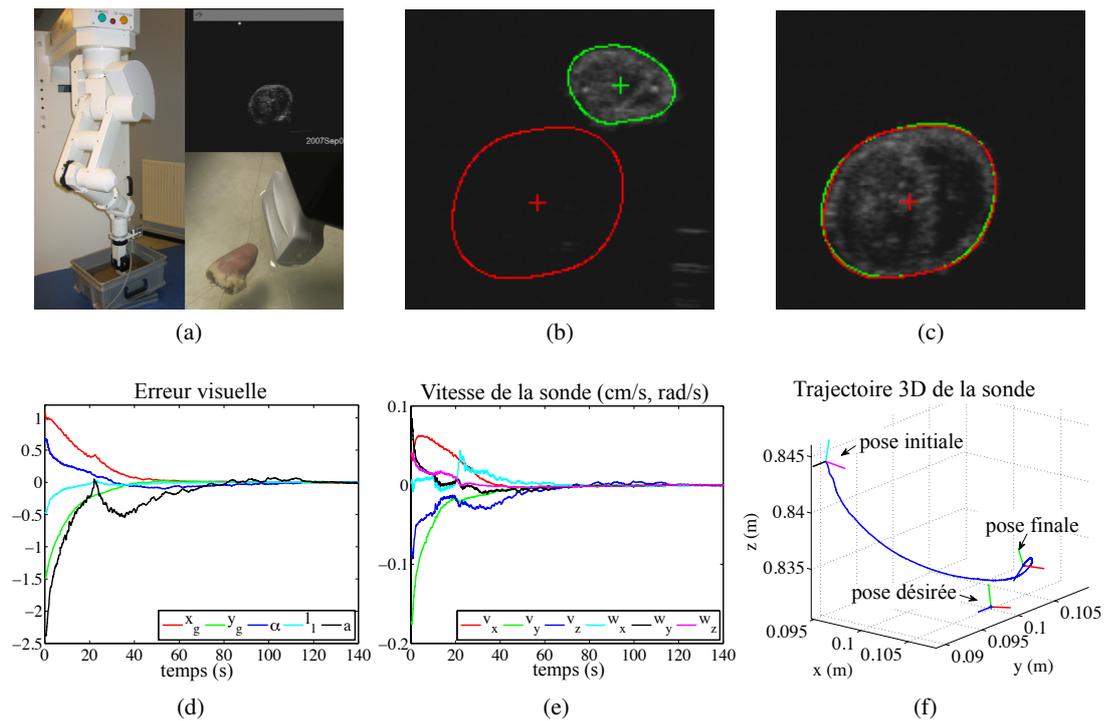


FIGURE 3.5 – Expérimentation *in vitro* de la tâche de positionnement : (a) plate-forme expérimentale constituée d'un robot Hippocrate 6 axes portant une sonde linéaire 5-10 Mhz interagissant avec un cœur de lapin immergé dans de l'eau ; (b) section désirée (rouge) et initiale (verte) ; (c) section désirée (rouge) et courante (verte) obtenue à la convergence de l'asservissement visuel ; (d) évolution de l'erreur visuelle (cm, cm, rad, cm, cm²) ; (e) vitesse appliquée à la sonde ; (f) trajectoire 3D (m, m, m) de la sonde réalisée lors de l'asservissement visuel.

avons proposé une méthode d'estimation en ligne des composantes du vecteur normal $\nabla\mathbf{F}$ à la surface de l'objet pour chaque point du contour de la section observée. Cette approche permet ainsi d'obtenir une estimation de la matrice d'interaction courante durant l'asservissement visuel pour des objets de forme inconnue [Mebarki 2009, Mebarki 2010b].

Estimation du vecteur normal à la surface de l'objet

Le vecteur normal $\nabla\mathbf{F}$ à la surface de l'objet en un point P du contour C peut se définir par le produit vectoriel du vecteur tangent \mathbf{d}_i au contour en P dans le plan image avec un second vecteur tangent à la surface de l'objet \mathbf{d}_t , non colinéaire à \mathbf{d}_i (voir figure 3.6). Le vecteur \mathbf{d}_i est calculé directement à partir du contour extrait de l'image courante et \mathbf{d}_t est estimé à partir d'un ensemble de coupes acquises à chaque intervalle de temps dt lors d'un mouvement hors plan de la sonde. Pour le calcul de ce second vecteur \mathbf{d}_t , le principe consiste à estimer les paramètres d'une courbe 3D, représentée ici par un modèle du second ordre. Cette courbe \mathcal{K} épouse localement la surface de l'objet en passant au mieux par les points des contours successifs de l'objet ayant les mêmes coordonnées polaires γ dans l'image (voir figure 3.6). Les paramètres de la courbe sont alors estimés en ligne à l'aide d'un algorithme des moindres

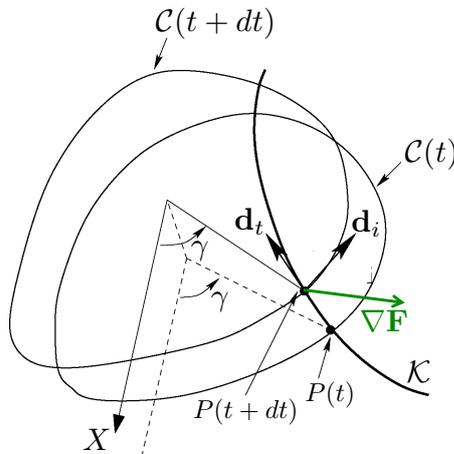


FIGURE 3.6 – Estimation de la normale aux points de contour de l'objet. Les points de régression de la courbe κ sont définis par la même orientation polaire γ .

carrés récursif qui utilise en entrée les coordonnées 3D de ces points successifs exprimées dans le repère de la sonde. Les coordonnées 3D des points sont obtenues par reconstruction à partir de leurs coordonnées 2D dans l'image et des poses relatives entre les différentes coupes fournies par l'odométrie du robot. Afin de permettre une variation locale de la courbe lors du déplacement de la sonde, un facteur d'oubli est mis en œuvre dans l'algorithme des moindres carrés récursif pour prendre en compte uniquement les mesures les plus récentes. Le vecteur \mathbf{d}_t est ensuite obtenu en calculant le vecteur tangent à la courbe estimée.

Choix des 6 primitives visuelles et modélisation de l'interaction

Comme nous l'avons déjà souligné, il est nécessaire de retenir au minimum 6 primitives visuelles pour contraindre les 6 ddl du robot et contrôler ainsi tous les mouvements de la sonde dans l'espace. De plus, afin d'obtenir un découplage total de la commande, l'idéal serait de définir 6 primitives indépendantes entre elles dont la variation de chacune serait uniquement liée à une seule composante du torseur cinématique de la sonde. Il en résulterait une matrice d'interaction diagonale qui permettrait d'obtenir un système linéaire. Cette quête est bien entendu toujours d'actualité et de nombreux travaux sur l'asservissement visuel par vision optique ont été orientés dans ce sens. Nous n'allons donc pas exposer ici la solution capable d'atteindre ce découplage idéal dans le contexte de l'asservissement visuel échographique 2D, mais plutôt présenter le choix retenu pour obtenir un découplage « convenable ». Nous avons proposé dans [Mebarki 2010b] le vecteur de primitives visuelles suivant

$$\mathbf{s} = (x_g, y_g, \alpha, \sqrt{a}, \phi_1, \phi_2). \quad (3.15)$$

Comme pour le cas de l'objet ellipsoïdal (section 3.3.2), les coordonnées du barycentre de la section (x_g, y_g) ainsi que l'angle α caractérisant l'orientation principale de la section dans l'image ont été retenus : ces primitives sont en effet indépendantes entre elles et chacune est fortement liée à un mouvement dans le plan de la sonde (respectivement v_x, v_y et ω_z). Concernant le choix des primitives visuelles pour le contrôle des mouvements hors du plan de la sonde, nous avons retenu trois primitives invariantes aux mouvements dans le plan. La première est définie par l'aire de la section de l'objet et plus précisément par sa racine carrée \sqrt{a} afin de conserver la même unité de mesure que x_g et y_g . Les deux dernières primitives (ϕ_1, ϕ_2) sont construites à partir de moments invariants à l'échelle de l'image ainsi qu'aux deux translations et à la rotation dans l'image [Tahri 2005]. Elles sont donc invariantes aux mouvements dans le plan et découplées de l'aire de la section observée. Par ailleurs, elles sont également indépendantes entre elles car ϕ_1 est obtenue à partir d'une combinaison des moments d'ordre 2 et ϕ_2 est définie à partir des moments d'ordre 3. Les expressions analytiques de ces 6 primitives visuelles en fonction des moments sont détaillées dans [Mebarki 2010a, Mebarki 2010b] ainsi que le calcul à partir de $\mathbf{L}_{m_{i,j}}$ de la matrice d'interaction \mathbf{L}_s associée à ce choix. La forme finale de \mathbf{L}_s s'exprime par

$$\mathbf{L}_s = \begin{bmatrix} -1 & 0 & x_{gvz} & x_{g\omega_x} & x_{g\omega_y} & y_g \\ 0 & -1 & y_{gvz} & y_{g\omega_x} & y_{g\omega_y} & -x_g \\ 0 & 0 & \alpha_{vz} & \alpha_{\omega_x} & \alpha_{\omega_y} & -1 \\ 0 & 0 & \frac{a_{vz}}{2\sqrt{a}} & \frac{a_{\omega_x}}{2\sqrt{a}} & \frac{a_{\omega_y}}{2\sqrt{a}} & 0 \\ 0 & 0 & \phi_{1vz} & \phi_{1\omega_x} & \phi_{1\omega_y} & 0 \\ 0 & 0 & \phi_{2vz} & \phi_{2\omega_x} & \phi_{2\omega_y} & 0 \end{bmatrix}. \quad (3.16)$$

La figure 3.7 présente les résultats expérimentaux obtenus pour une tâche de positionnement de la sonde par rapport à un objet de forme complexe en appliquant la loi de commande (2.16) avec les 6 primitives visuelles proposées. Cette expérimentation a été réalisée en condition in vitro en modélisant un organe d'intérêt par un objet confectionné en gélatine et immergé dans un bac d'eau. L'estimation du vecteur normal à la surface de l'objet pour chaque point de contour est effectuée en ligne par la méthode introduite précédemment. Elle permet de considérer une estimation de la matrice d'interaction courante (3.16) dans la commande. Lors de

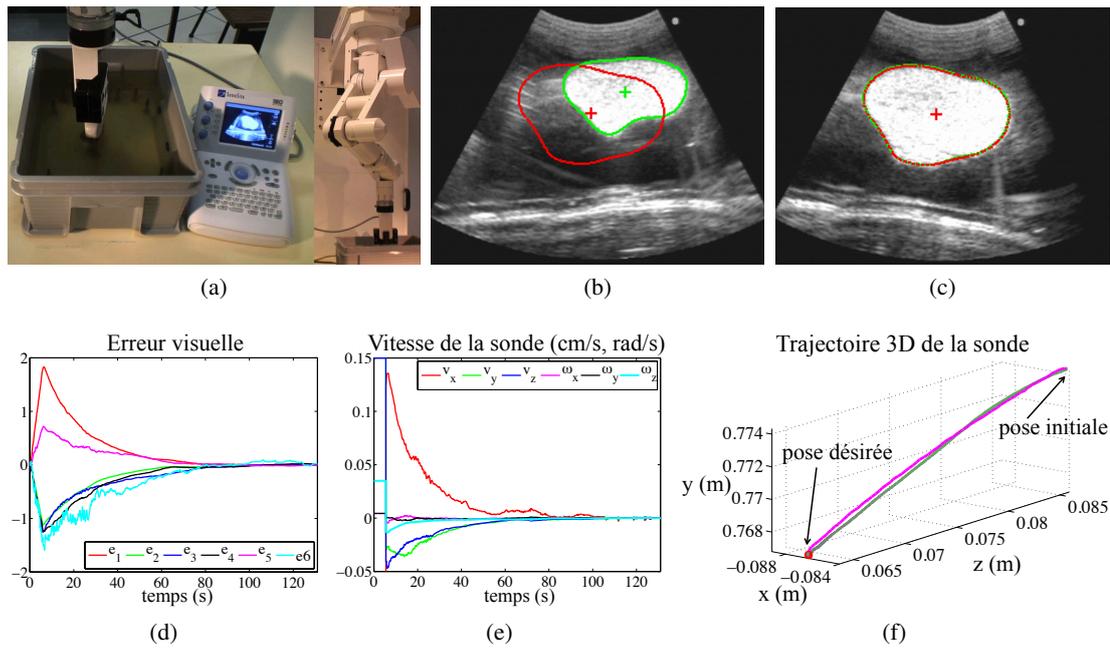


FIGURE 3.7 – Expérimentation *in vitro* de la tâche de positionnement : (a) plate-forme expérimentale constituée d'un robot Hippocrate 6 axes portant une sonde convexe 2-5 Mhz interagissant avec un objet en gélatine de forme complexe immergé dans de l'eau ; (b) section désirée (rouge) et initiale au démarrage de l'asservissement visuel (verte) ; (c) section désirée (rouge) et courante (verte) obtenue à la convergence de l'asservissement visuel ; (d) évolution de l'erreur visuelle (cm, cm, deg/10, cm, unité, unité) et (e) vitesse appliquée à la sonde durant la phase ($t=0$ à 6s) de positionnement en boucle ouverte permettant d'éloigner la sonde de la position où a été apprise la section désirée et la phase ($t>6s$) d'asservissement visuel ; (f) trajectoire 3D (m, m, m) de la sonde lors du déplacement en boucle ouverte (magenta) et trajectoire 3D (verte) réalisée durant l'asservissement visuel.

cette expérience, la valeur désirée des 6 primitives visuelles a été apprise à une position initiale de la sonde. Une vitesse constante a ensuite été appliquée en boucle ouverte afin d'amener la sonde vers une autre position. A partir de celle-ci, l'asservissement visuel a été démarré à l'instant $t=6s$ dans le but de ramener automatiquement la sonde vers la section initiale désirée. Comme prévu et souligné à la section 2.2.4, une décroissance exponentielle de l'erreur visuelle est obtenue. D'autre part, la trajectoire de la sonde a cette fois-ci convergé vers la pose désirée où a été apprise la section désirée ; en effet, les 6 dll du système sont ici totalement contraints par la tâche visuelle.

Depuis une position initiale peu éloignée de la position désirée, le choix de ces 6 primitives extraites d'un seul plan de coupe assure donc la convergence de la sonde en termes d'erreur visuelle et de position. Néanmoins, l'information contenue dans un seul plan de coupe n'est pas toujours suffisamment discriminante de la position relative de la sonde par rapport à l'objet, auquel cas la minimisation de l'erreur visuelle ne garantit pas la convergence de l'asservissement visuel en termes de positionnement. Cette limitation est particulièrement flagrante dans le cas d'un objet grossièrement symétrique où plusieurs positions différentes de la sonde échographique peuvent générer une section de caractéristiques géométriques semblables. Il existe alors plusieurs minima globaux qui conduisent au même état des primitives visuelles voire à la même image.

3.4 Approche tri-plans utilisant les moments 2D

Pour assurer une meilleure convergence de l'asservissement visuel basé sur les moments en termes de positionnement dans l'espace, lors des travaux de thèse de Caroline Nadeau [Nadeau 2011a], nous avons proposé de considérer plusieurs coupes échographiques simultanément. Par rapport à l'approche mono-plan, l'information hors plan apportée par l'ajout de nouvelles coupes permet ainsi d'obtenir une meilleure caractérisation de la pose de la sonde par rapport à des objets de nature fortement symétriques tels des ovoïdes ou ellipsoïdes pouvant ressembler par exemple à des kystes ou tumeurs. C'est dans ce sens qu'une approche basée sur l'extraction de primitives géométriques à partir de trois plans orthogonaux associés au capteur échographique a été développée dans [Nadeau 2010]. En pratique, ces trois plans peuvent être extraits en temps réel par interpolation à partir d'un volume échographique fourni soit par une sonde 3D matricielle, soit par une sonde à balayage motorisé dont les principes de fonctionnement ont été introduits à la section 2.1.3.

3.4.1 Choix des 6 primitives visuelles et modélisation de l'interaction

L'idée générale de cette approche tri-plans est de sélectionner, pour chaque plan orthogonal, des primitives géométriques qui sont fortement couplées aux mouvements dans le plan considéré et très peu sensibles aux mouvements hors de ce plan. A cet effet, les coordonnées dans l'image du centre de gravité de la section de l'objet ainsi que son orientation principale sont sélectionnées. Les coordonnées du barycentre sont particulièrement représentatives des mouvements de translation dans le plan de l'image et l'orientation principale présente un fort couplage avec la rotation dans le plan tout en étant invariante aux translations dans l'image. La figure 3.8 représente la configuration tri-plans proposée. Dans ce cas, le vecteur s des 6

primitives visuelles utilisées dans la loi de commande pour contrôler les 6 ddl du capteur échographique s'exprime selon

$$\mathbf{s} = (x_{g_0}, y_{g_0}, x_{g_1}, \alpha_1, \alpha_2, \alpha_0), \quad (3.17)$$

où (x_{g_0}, y_{g_0}) et α_0 correspondent aux coordonnées du barycentre et l'orientation principale de la section observée dans la coupe frontale nommée US_0 , x_{g_1} et α_1 représentent respectivement l'abscisse du barycentre et l'orientation de la section observée dans la coupe sagittale US_1 , et α_2 décrit l'orientation de la section observée dans la coupe transverse US_2 .

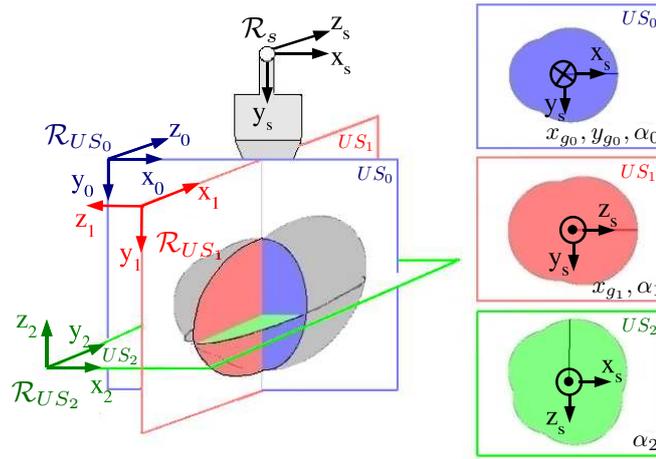


FIGURE 3.8 – Approche tri-plans. Par souci de clarté, les différents repères \mathcal{R}_{US_i} , avec $i \in \{0, 1, 2\}$ associés aux différents plans image de la sonde, sont indiqués aux coins supérieurs gauches des plans images. En pratique, les origines de ces repères sont confondues et positionnées au milieu des trois images. A droite les trois images orthogonales fournies à un instant donné sont représentées et le repère \mathcal{R}_s de commande de la sonde est indiqué sur chaque plan.

La configuration choisie permet ainsi d'assimiler chacun des 6 ddl de la sonde à un mouvement du plan dans l'un des trois plans image définis. Par rapport à la définition du repère de commande de la sonde \mathcal{R}_s (voir figure 3.8), les trois mouvements de la sonde dans le plan (v_x, v_y, ω_z) correspondent aux mouvements dans le plan image US_0 $(v_{x_0}, v_{y_0}, \omega_{z_0})$, les composantes hors plan (v_z, ω_x) correspondent aux deux mouvements dans le plan image US_1 $(v_{x_1}, -\omega_{z_1})$ et enfin la rotation ω_y de la sonde correspond à la rotation dans le plan image US_2 $(-\omega_{z_2})$. Par conséquent, tous les mouvements de la sonde peuvent être contrôlés à l'aide de primitives géométriques associées aux mouvements du plan d'où elles sont extraites.

Etant donné que les repères \mathcal{R}_{US_i} sont rigidement liés au repère de la sonde \mathcal{R}_s et que leurs origines sont confondues, nous pouvons exprimer leur vitesse instantanée \mathbf{v}_{US_i} en fonction du torseur cinématique de la sonde \mathbf{v} par

$$\forall i \in \{0, 1, 2\}, \quad \mathbf{v}_{US_i} = {}^{US_i}\mathbf{W}_s \mathbf{v}, \quad (3.18)$$

avec

$${}^{US_i}\mathbf{W}_s = \begin{bmatrix} {}^{US_i}\mathbf{R}_s & \mathbf{0}_3 \\ \mathbf{0}_3 & {}^{US_i}\mathbf{R}_s \end{bmatrix}, \quad (3.19)$$

où ${}^{US_i}\mathbf{R}_s$ est la matrice de rotation caractérisant la transformation rigide entre les repères \mathcal{R}_{US_i} et \mathcal{R}_s . Par la suite, nous considérons que le repère \mathcal{R}_{US_0} est aligné avec celui de la sonde, soit ${}^{US_0}\mathbf{W}_s = \mathbf{I}_6$. La matrice d'interaction \mathbf{L}_s qui lie la variation de s au torseur cinématique de la sonde \mathbf{v} s'obtient alors directement à partir des matrices d'interaction élémentaires exprimant chacune la variation d'une primitive par rapport à la vitesse du plan où elle est extraite. Ainsi, on obtient

$$\mathbf{L}_s = \begin{bmatrix} \mathbf{L}_{xg_0} \\ \mathbf{L}_{yg_0} \\ \mathbf{L}_{xg_1} \quad {}^{US_1}\mathbf{W}_s \\ \mathbf{L}_{\alpha_1} \quad {}^{US_1}\mathbf{W}_s \\ \mathbf{L}_{\alpha_2} \quad {}^{US_2}\mathbf{W}_s \\ \mathbf{L}_{\alpha_0} \end{bmatrix}. \quad (3.20)$$

Les matrices d'interaction élémentaires \mathbf{L}_{x_g} , \mathbf{L}_{y_g} , \mathbf{L}_{α} (de dimension 1×6) sont respectivement associées à l'abscisse, à l'ordonnée et à l'orientation principale de la section dans l'image qui sont données par les 3 premières lignes de la matrice d'interaction présentée à l'équation (3.16). En substituant dans (3.20) les expressions de ces matrices, nous obtenons la forme suivante

$$\mathbf{L}_s = \begin{bmatrix} -1 & 0 & x_{g_0v_z} & x_{g_0\omega_x} & x_{g_0\omega_y} & y_{g_0} \\ 0 & -1 & y_{g_0v_z} & y_{g_0\omega_x} & y_{g_0\omega_y} & -x_{g_0} \\ -x_{g_1v_z} & 0 & -1 & -y_{g_1} & x_{g_1\omega_y} & x_{g_1\omega_x} \\ -\alpha_{1v_z} & 0 & 0 & 1 & \alpha_{1\omega_y} & \alpha_{1\omega_x} \\ 0 & -\alpha_{2v_z} & 0 & \alpha_{2\omega_x} & 1 & \alpha_{2\omega_y} \\ 0 & 0 & \alpha_{0v_z} & \alpha_{0\omega_x} & \alpha_{0\omega_y} & -1 \end{bmatrix}, \quad (3.21)$$

où l'indice 0, 1 ou 2 associé à chaque élément de cette matrice indique le plan image dont l'information visuelle est extraite, et l'indice $(v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)$ fait référence au mouvement de ce même plan image.

Comme établi précédemment, les six primitives géométriques considérées dans le vecteur d'informations visuelles sont très fortement couplées à un mouvement particulier dans le plan où elles sont extraites. C'est pourquoi au final, nous avons proposé d'utiliser l'estimation suivante de la matrice d'interaction dans la loi de commande

$$\widehat{\mathbf{L}}_s = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & y_{g_0} \\ 0 & -1 & 0 & 0 & 0 & -x_{g_0} \\ 0 & 0 & -1 & -y_{g_1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \quad (3.22)$$

Cette matrice décrit un comportement approché du système dans la mesure où l'influence des mouvements hors plan sur la variation des informations visuelles est négligée par rapport à l'influence des mouvements dans le plan. Par comparaison avec la matrice complète (3.21), cette nouvelle matrice présente de très bonnes propriétés de découplage et ne dépend que de données mesurées directement dans l'image. En particulier, elle n'implique pas d'estimation du vecteur normal à la surface de l'objet et garantit une stabilité asymptotique locale de la commande si la condition de positivité de $\mathbf{L}_s(\widehat{\mathbf{L}}_s)^{-1}$ est respectée.

En pratique, cette solution nécessite cependant l'usage d'une sonde 3D fournissant des volumes en temps réel dans lesquelles les 3 coupes échographiques peuvent être extraites. Ces sondes de technologie matricielle sont actuellement interfacées à des systèmes fermés qui ne peuvent être ouverts sans l'accord des fabricants. Néanmoins, cette approche tri-plans peut être appliquée dans le cadre d'un asservissement visuel virtuel en vue d'effectuer le recalage de coupes échographiques avec un volume préalablement acquis comme nous le verrons ci-après.

3.4.2 Application au recalage par asservissement virtuel

L'asservissement visuel virtuel a été proposé dans la littérature pour estimer la pose d'objets observés par une caméra à des fins d'applications de réalité augmentée [Marchand 2002]. Le principe consiste à contrôler, par un asservissement visuel 2D, les déplacements d'une caméra virtuelle interagissant avec un objet virtuel. L'objectif est d'atteindre une image simulée correspondant à la vue du même objet fournie par une vraie caméra. Une estimation de la pose relative entre l'objet et la caméra réels est ainsi disponible à la convergence de l'asservissement visuel virtuel.

Dans le même ordre d'idées, nous avons proposé de valider l'approche tri-plans par la mise en œuvre d'un asservissement visuel virtuel dont l'objectif est de réaliser un recalage rigide automatique entre un volume échographique pré-opératoire et 3 coupes orthogonales per-opératoires qui ont été acquises séquentiellement à l'aide d'une sonde 2D positionnée par un robot [Nadeau 2010]. Le principe consiste à commander par asservissement visuel une sonde virtuelle tri-plans interagissant avec le volume pré-opératoire de manière à atteindre à la convergence les 3 coupes per-opératoires qui ont été observées par la sonde réelle. Les images fournies par la sonde virtuelle sont directement générées à partir du volume per-opératoire grâce à une technique d'interpolation, puis segmentées par une méthode basée sur un contour actif. Les 6 primitives visuelles désirées ont été extraites des 3 images échographiques réelles après détection du contour de la section observée dans chaque image par le même algorithme de segmentation. L'asservissement est réalisé en appliquant à la sonde virtuelle la commande cinématique (2.16) utilisant l'estimation de la matrice d'interaction (3.22). Les expériences de recalage ont été réalisées avec un fantôme contenant un objet en forme d'œuf. La figure 3.9 présente les résultats obtenus pour un recalage monomodal où nous avons considéré un volume échographique pré-opératoire du fantôme. Comme prévu, nous pouvons observer la convergence de l'asservissement visuel en termes d'erreur visuelle et de position. La transformation rigide recherchée entre les images per-opératoires et le volume pré-opératoire se caractérise alors directement par la pose de la sonde virtuelle obtenue à la convergence de l'asservissement visuel. Cette approche peut également servir à réaliser un recalage multimodal comme le démontrent les résultats présentés à la figure 3.10 où un volume pré-opératoire tomodensitométrique (TDM) du fantôme est considéré.

A la différence des approches utilisant des informations visuelles extraites d'une seule coupe, l'approche tri-plans a l'avantage de s'affranchir de l'estimation d'informations 3D de l'objet considéré, tels l'orientation des droites pour l'information de type point ou le vecteur normal à la surface de l'objet pour l'approche mono-plan basée sur les moments. En effet, les composantes de la matrice d'interaction $\widehat{\mathbf{L}}_s$ utilisée dans la loi de commande sont uniquement dépendantes d'informations géométriques extraites des trois plans de coupe. De plus, le do-

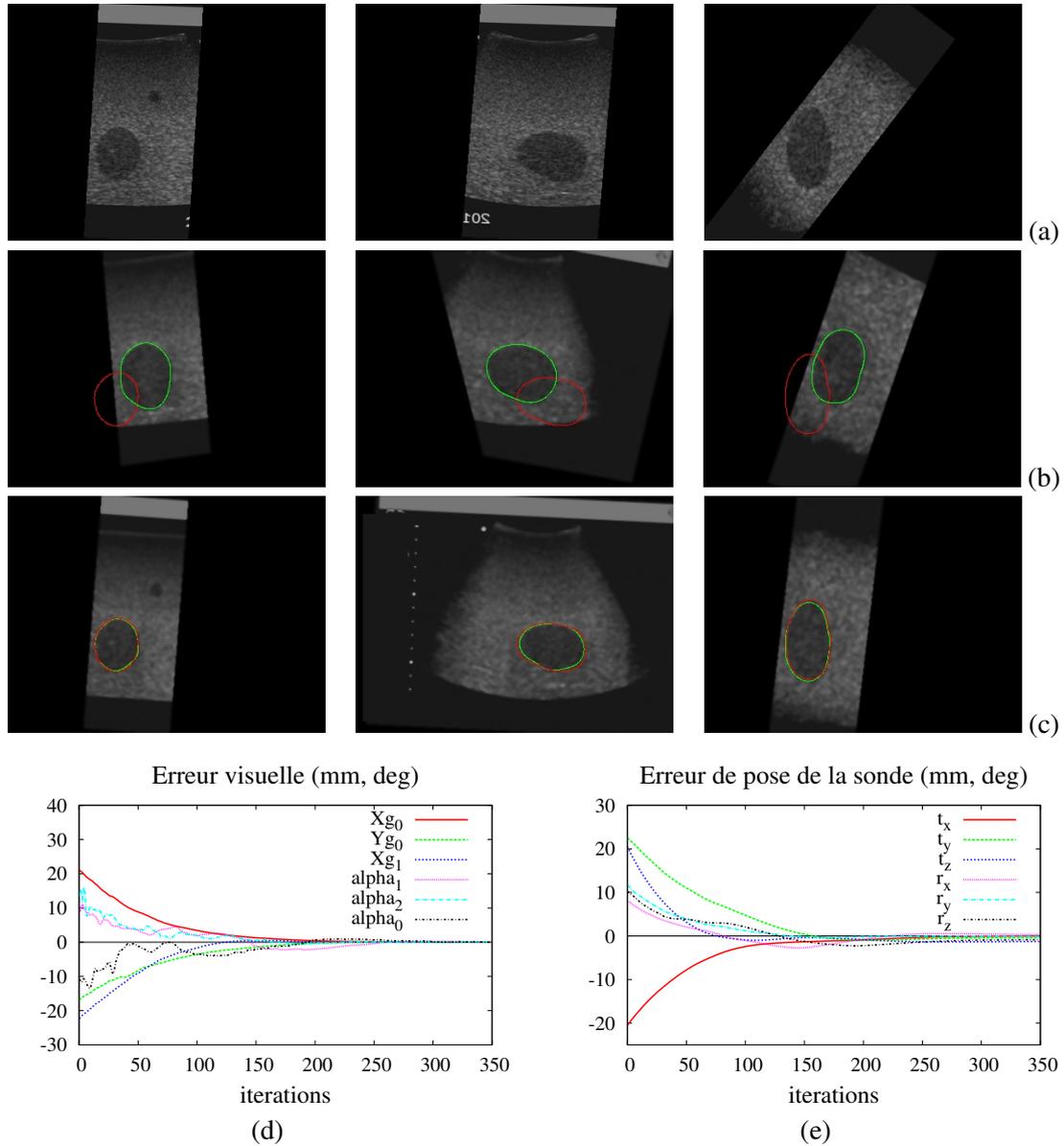
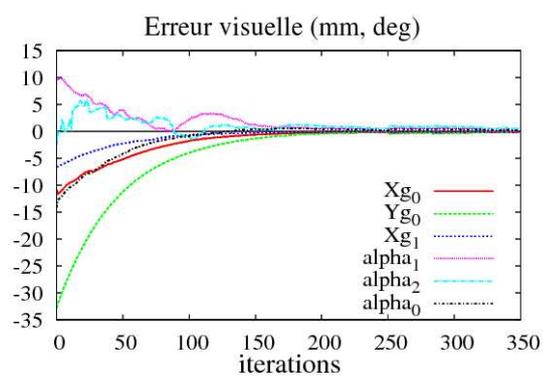
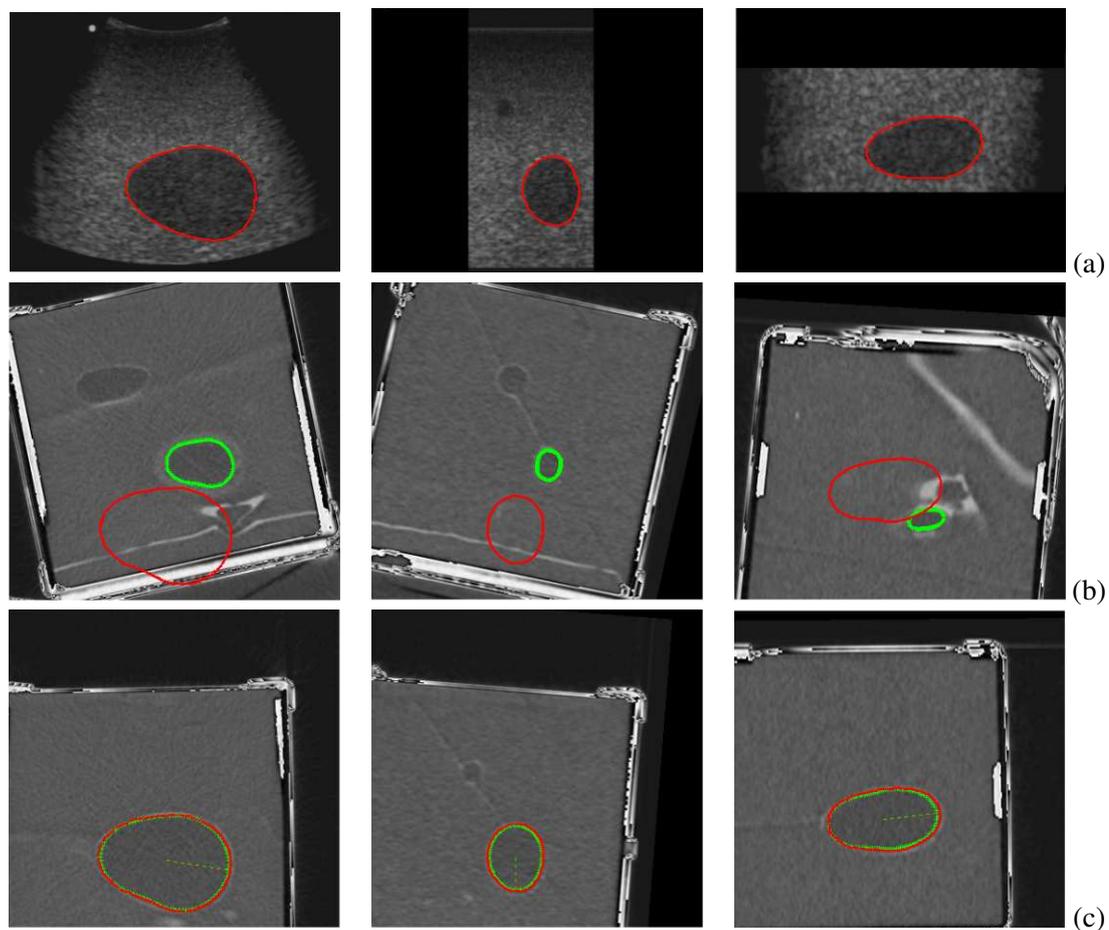


FIGURE 3.9 – Recalage US/US : (a) Vues échographiques per-opératoires et extraction du contour désiré de l'organe. (b) Coupes échographiques pré-opératoires associées à la position initiale de la sonde virtuelle tri-plans. (c) Vues acquises à la position finale de la sonde virtuelle. (d), (e) Convergence de l'algorithme en termes d'erreur visuelle et de positionnement.



(d)

FIGURE 3.10 – Recalage US/TDM : (a) Vues échographiques per-opératoires et extraction du contour désiré de l'organe. (b) Coupes tomodensitométriques pré-opératoires associées à la position initiale de la sonde virtuelle tri-plans. (c) Vues acquises à la position finale de la sonde virtuelle. (d) Convergence de l'algorithme en termes d'erreur visuelle.

maine de convergence de cette approche tri-plans est largement supérieur à celui des approches mono-plan en raison de la prise en compte de vues complémentaires de l'objet qui contraignent d'avantage l'asservissement pour atteindre le minimum global en terme de pose.

Néanmoins, quelque soit la méthode utilisée, il est nécessaire d'extraire de l'image échographique les informations géométriques considérées en entrée de l'asservissement visuel. La faible qualité des images échographiques rend cette extraction difficile et il est donc nécessaire de proposer des solutions adaptées à cette modalité d'image qui doivent être robustes, tout en restant faiblement complexes pour respecter la contrainte du temps réel. Pour répondre à cette attente, la section suivante fournit une solution générique permettant d'extraire les informations géométriques de l'image échographique.

3.5 Contour actif basé sur les descripteurs de Fourier

Pour mettre en œuvre les méthodes d'asservissement visuel échographique proposées, il nous est nécessaire de concevoir des algorithmes de traitement d'images échographiques capables de détecter et de suivre à la cadence vidéo des informations géométriques relatives à la section observée d'un élément anatomique d'intérêt. La segmentation d'images est une opération qui permet de répondre à ce besoin. Dans la littérature, il existe deux principales catégories de méthodes de segmentation : les méthodes de segmentation basées région [Chan 2001, Subasic 2003, Slabaugh 2006, Belaid 2009] et les méthodes de segmentation basées contour [Brigger 2000, Precioso 2005, Collewet 2009].

Les méthodes basées sur les régions permettent de séparer les pixels appartenant à l'objet d'intérêt de la scène selon un critère défini par exemple à partir de l'homogénéité d'une couleur, de la texture ou du mouvement d'un sous-ensemble des pixels de l'image. Néanmoins les temps de calcul de ces méthodes sont élevés et ne permettent donc pas leur application en vue de la commande d'un système robotique.

Les méthodes basées sur les contours, également dénommées contours actifs, peuvent être exécutées beaucoup plus rapidement. Ces dernières se décomposent en deux classes selon la description du contour utilisée. La première classe regroupe les méthodes de type contour actif géométrique où le contour est décrit par une courbe de niveau [Malladi 1995] et la seconde définit les méthodes de type contour actif paramétrique pour lesquelles le contour est décrit par une fonction paramétrique [Brigger 2000].

Etant donné que notre contrainte principale est de réaliser la segmentation de la section observée à une cadence proche de 25 images/seconde, nous nous sommes orientés vers une méthode basée sur un contour actif paramétrique. Le contour actif paramétrique a également l'avantage d'introduire une contrainte géométrique sur la forme du contour recherché. Cette contrainte est utile pour augmenter la robustesse de la segmentation vis-à-vis du bruit contenu dans l'image qui est fortement présent dans les images échographiques.

3.5.1 Formulation générique d'un contour actif paramétrique

Le contour actif paramétrique, également appelé « snake » en anglais [Kass 1987], est un contour déformable. Il est décrit par une courbe dont la forme minimise l'énergie issue de forces externes et internes. La force interne traduit les caractéristiques mécaniques internes du contour telle sa raideur. La force externe inclut en particulier la force issue de l'image qui attire le contour actif sur le contour réel de l'objet observé. Dans la littérature, il existe différents modèles de paramétrisation d'un contour comme par exemple les B-splines [Brigger 2000], les courbes de Bézier, la signature polaire [Lie 1986] ou les descripteurs de Fourier [Impe-dovo 1978]. Nous présentons ici une formulation générique du contour et les différentes étapes de calculs permettant d'obtenir son équation d'évolution [Collewet 2009].

Un modèle générique de contour actif paramétrique $C_{\mathbf{q}}$ est défini par une fonction vectorielle de dimension 2, dépendante d'une variable de parcours $u \in [0, 2\pi]$ et d'un vecteur de paramètres \mathbf{q} de dimension n , telle que

$$C_{\mathbf{q}}(u) : \mathbf{x}(u) = \mathbf{x}_c + \sum_{l=0}^n q_l \Phi_l(u), \quad (3.23)$$

où \mathbf{x}_c sont les coordonnées 2D d'un point défini à l'intérieur de la section délimitée par le contour et $\Phi_l(u)$ ($l = 0, \dots, n$) sont des fonctions vectorielles de dimension 2 définies à partir des paramètres du modèle retenu pour le contour.

Cette courbe, plongée dans un milieu visqueux, se déforme sous l'influence de forces internes \mathbf{f}_i et de forces externes \mathbf{f}_e . L'équation d'évolution du contour traduisant sa déformation peut être obtenue grâce au formalisme de Lagrange. En considérant \mathbf{q} le vecteur des coordonnées généralisées du système, pour chaque composante q_i de \mathbf{q} , l'équation d'Euler-Lagrange conduit à l'expression

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial U}{\partial q_i} = Q_i, \quad (3.24)$$

où T et U correspondent respectivement à l'énergie cinétique et l'énergie potentielle du contour. Q_i décrit donc la force généralisée associée à q_i , soit

$$Q_i = \int_{C_{\mathbf{q}}} \mathbf{f}^\top(u) \frac{\partial C_{\mathbf{q}}(u)}{\partial q_i} du, \quad (3.25)$$

avec $\mathbf{f} = \mathbf{f}_{vis} + \mathbf{f}_{im}$ désignant le vecteur des forces dissipatives où \mathbf{f}_{vis} représente la force due au frottement visqueux et \mathbf{f}_{im} la force engendrée par l'image. L'énergie cinétique T est définie par

$$T = \int_{C_{\mathbf{q}}} \frac{1}{2} \mu \| C_t \|^2 du, \quad (3.26)$$

avec $C_t = \frac{\partial C_{\mathbf{q}}}{\partial t}$ et où μ représente la masse linéique du contour $C_{\mathbf{q}}$. Le plus souvent, l'énergie potentielle U est décrite par le biais de termes de régularisation de Tikhonov tels que

$$U = \int_{C_{\mathbf{q}}} \frac{1}{2} k_1 \| C_u \|^2 du + \int_{C_{\mathbf{q}}} \frac{1}{2} k_2 \| C_{uu} \|^2 du, \quad (3.27)$$

avec $C_u = \frac{\partial C_{\mathbf{q}}}{\partial u}$ et $C_{uu} = \frac{\partial^2 C_{\mathbf{q}}}{\partial u^2}$. Les scalaires k_1 et k_2 permettent de limiter respectivement l'extension et la courbure du contour $C_{\mathbf{q}}$.

Intéressons-nous à présent à la force généralisée Q_i et plus précisément aux forces dissipatives \mathbf{f} impliquées dans (3.25). La force due au frottement est généralement définie par $\mathbf{f}_{vis} = -\gamma C_t$ où γ est la viscosité du milieu. Sa contribution dans (3.25) est donnée par

$$\int_{C_q} \mathbf{f}_{vis}^\top(u) \frac{\partial C_q(u)}{\partial q_i} du = - \int_{C_q} \gamma \left(\sum_l \dot{q}_l \Phi_l \right)^\top \Phi_i du. \quad (3.28)$$

D'autre part, la force engendrée par l'image est définie par $\mathbf{f}_{im} = -\nabla E_{im}$ où E_{im} correspond à l'énergie potentielle de l'image I . E_{im} est caractérisée par $E_{im} = -\|\nabla I\|$, avec $\|\nabla I\|$ correspondant à la norme du gradient ∇I de l'image I .

Finalement (3.24) peut s'écrire sous la forme vectorielle suivante

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{Q}_{im}(\mathbf{q}) = - \int_{C_q} \nabla E_{im}^\top(u) \frac{\partial C_q(u)}{\partial \mathbf{q}} du, \quad (3.29)$$

où $\mathbf{M} = [M_{mn}]$, $\mathbf{C} = [C_{mn}]$ et $\mathbf{K} = [K_{mn}] = \mathbf{K}_1 + \mathbf{K}_2$, et avec

$$\begin{cases} M_{mn} = \mu \int_{C_q} \Phi_m^\top \Phi_n du \\ C_{mn} = \gamma \int_{C_q} \Phi_m^\top \Phi_n du \\ K_{1mn} = k_1 \int_{C_q} \Phi_m'^\top \Phi_n' du \\ K_{2mn} = k_2 \int_{C_q} \Phi_m''^\top \Phi_n'' du \end{cases} \quad \text{avec} \quad \Phi_l' = \frac{\partial \Phi_l}{\partial u} \quad \text{et} \quad \Phi_l'' = \frac{\partial^2 \Phi_l}{\partial u^2} \quad (3.30)$$

En fixant la masse μ à zéro afin d'annuler le comportement inertiel du contour actif et d'augmenter ainsi sa stabilité dynamique, le système d'équations différentielles (3.29) peut finalement être résolu en temps discret par la méthode d'Euler explicite qui conduit à la formulation récursive du vecteur des paramètres du contour selon

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \Delta t \mathbf{C}^{-1} (\mathbf{Q}_{im}(\mathbf{q}_k) - \mathbf{K}\mathbf{q}_k). \quad (3.31)$$

En pratique, la méthode consiste dans un premier temps à initialiser manuellement le contour actif à une distance proche du contour réel à détecter. L'évolution automatique du contour permettant d'atteindre le contour réel est ensuite effectuée en itérant le calcul de l'équation (3.31) jusqu'à l'obtention de la convergence des paramètres \mathbf{q} du contour. Notons que les matrices impliquées dans (3.29) sont diagonales si les fonctions vectorielles $\Phi_l(u)$ ainsi que leur première et seconde dérivées sont orthogonales. Il est donc judicieux de retenir une paramétrisation du contour qui permet d'obtenir ces conditions afin de simplifier grandement l'équation d'évolution du contour actif. Il en résulte ainsi un temps de calcul très rapide de la convergence du contour qui garantit le respect de la contrainte du temps réel.

3.5.2 Formulation basée sur les descripteurs de Fourier

Une paramétrisation du contour basée sur la signature polaire a été proposée dans la littérature pour extraire le contour d'un objet d'intérêt par rapport au fond de la scène à la cadence

du flux vidéo [Collewet 2009]. Cette description a été considérée dans le cadre des travaux de thèse de Rafik Mebarki pour la mise en œuvre d'asservissements visuels échographiques utilisant les moments de la section de l'objet segmenté [Mebarki 2008b]. L'équation du contour (3.23) devient dans ce cas

$$C_{\mathbf{q}}(u) : \mathbf{x}(u) = \mathbf{x}_c + \rho_q(u) \begin{pmatrix} \cos u \\ \sin u \end{pmatrix}, \quad (3.32)$$

où $\rho_q(u)$ est le rayon polaire qui est décomposé en série de Fourier avec un nombre fini h d'harmoniques tel que

$$\rho_q(u) = q_0 + \sum_{k=1}^h q_k \cos(ku) + q_{k+h} \sin(ku). \quad (3.33)$$

Les expressions des différentes matrices mises en jeu dans l'équation différentielle d'évolution (3.29) des paramètres \mathbf{q} de la signature polaire sont détaillées dans [Collewet 2009].

Cependant, la signature polaire est, comme son nom l'indique, une description basée sur les coordonnées polaires, ce qui réduit son domaine d'application à la représentation de contours dont les formes sont convexes. Par conséquent, nous avons proposé de considérer une description cartésienne du contour afin de permettre la détection et le suivi dans des séquences d'images échographiques d'objet dont le contour peut être de forme concave ou convexe. Etant donné que les courbes de Bézier ou les B-splines ne sont pas basées sur des fonctions orthogonales, nous avons retenu les descripteurs de Fourier comme paramétrisation du contour. En remplaçant le terme $\Phi_l(u)$ par l'expression des descripteurs de Fourier [Impedovo 1978] dans la formulation générique du contour (3.23), nous obtenons l'équation du contour formulée par

$$C_{\mathbf{q}}(u) : \mathbf{x}(u) = \begin{pmatrix} a_0 \\ c_0 \end{pmatrix} + \sum_{k=1}^h \begin{pmatrix} a_k & b_k \\ c_k & d_k \end{pmatrix} \begin{pmatrix} \cos ku \\ \sin ku \end{pmatrix}, \quad (3.34)$$

où h est le nombre d'harmoniques permettant de complexifier plus ou moins la forme du contour.

Afin de mettre en œuvre un contour actif basé sur cette description du contour, nous avons proposé dans [Li 2011a] de définir le vecteur des coordonnées généralisées du système par

$$\mathbf{q} = (a_0, a_1 \cdots a_h, b_1 \cdots b_h, c_0, c_1 \cdots c_h, d_1 \cdots d_h). \quad (3.35)$$

Cette expression permet ainsi de reformuler l'expression du contour (3.34) par

$$C_{\mathbf{q}}(u) : \mathbf{x}(u) = \sum_{l=0}^{4h+1} q_l \Phi_l(u), \quad (3.36)$$

avec les fonctions vectorielles $\Phi_l(u)$ définies selon

$$\left\{ \begin{array}{l} 0 \leq l < h : \Phi_l(u) = \begin{pmatrix} \cos lu \\ 0 \end{pmatrix}, \\ h+1 \leq l < 2h : \Phi_l(u) = \begin{pmatrix} \sin(l-h)u \\ 0 \end{pmatrix}, \\ 2h+1 \leq l < 3h+1 : \Phi_l(u) = \begin{pmatrix} 0 \\ \cos(l-2h-1)u \end{pmatrix}, \\ 3h+2 \leq l < 4h+2 : \Phi_l(u) = \begin{pmatrix} 0 \\ \sin(l-3h-1)u \end{pmatrix}. \end{array} \right. \quad (3.37)$$

En substituant (3.37) dans (3.30) et en fixant la masse μ à zéro pour annuler le comportement inertiel du contour actif, les matrices \mathbf{C} , $\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$, impliquées dans le système d'équations différentielles décrivant l'évolution du contour, sont données par

$$\left\{ \begin{array}{l} \mathbf{C} = \gamma \text{diag}(2\pi, \pi \dots \pi, \pi \dots \pi, \\ \quad \quad \quad 2\pi, \pi \dots \pi, \pi \dots \pi), \\ \mathbf{K}_1 = k_1 \text{diag}(0, \pi \dots \pi h^2, \pi \dots \pi h^2, \\ \quad \quad \quad 0, \pi \dots \pi h^2, \pi \dots \pi h^2), \\ \mathbf{K}_2 = k_2 \text{diag}(0, \pi \dots \pi h^4, \pi \dots \pi h^4, \\ \quad \quad \quad 0, \pi \dots \pi h^4, \pi \dots \pi h^4). \end{array} \right. \quad (3.38)$$

Etant donné que ces matrices sont diagonales et constantes, le système d'équations différentielles (3.29) devient optimal. Il en résulte une réduction du temps de calcul de l'expression récursive (3.31) permettant d'actualiser la valeur des paramètres du contour.

La figure 3.11 présente les résultats obtenus en appliquant le contour actif basé sur les descripteurs de Fourier pour des formes convexes et concaves observées dans des images échographiques, ainsi que ceux obtenus avec la paramétrisation de la signature polaire. Nous pouvons évidemment remarquer que seule l'approche utilisant les descripteurs de Fourier que nous avons proposée permet de détecter et suivre les deux types de forme alors que le contour actif basé sur la signature polaire converge uniquement sur le contour de la section concave.

3.5.3 Initialisation du contour actif et changement de topologie

Un des problèmes majeurs des contours actifs paramétriques est leur sensibilité à leur initialisation. En effet, il est bien connu que le contour actif doit être initialisé proche du contour réel sous peine de le voir diverger, le terme de gradient introduit dans \mathbf{f} n'étant pertinent que dans un proche voisinage du contour réel. Différentes approches ont été proposées pour traiter ce problème. Signalons par exemple [Xu 1998] où une image de diffusion du gradient est utilisée. Il est également possible d'introduire des forces de pressions qui, de manière imagée, permettent de gonfler ou de dégonfler le contour actif pour l'amener vers le contour réel

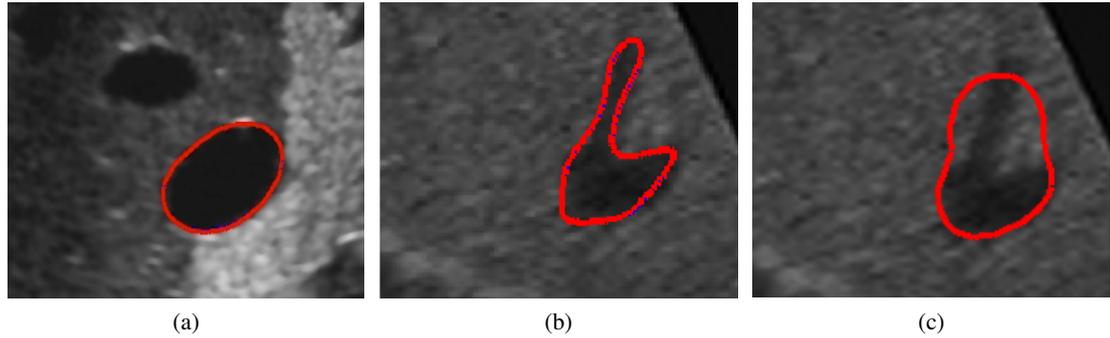


FIGURE 3.11 – Résultats du contour actif dans des images échographiques. (a) Contour obtenu pour une forme convexe à la convergence du contour actif basé soit sur les descripteurs de Fourier ou la signature polaire (nombre d'harmoniques $h = 6$). (b) Contour obtenu pour une forme concave à la convergence du contour actif basé sur les descripteurs de Fourier (nombre d'harmoniques $h = 6$). (c) Divergence du contour actif basé sur la signature polaire dans le cas d'une forme concave (nombre d'harmoniques $h = 6$).

[Cohen 1991]. Il est alors nécessaire de calculer la normale en tout point du contour actif. Néanmoins, ces deux approches restent très gourmandes en temps de calcul. Nous proposons ici une solution de faible complexité.

En effet, comme souligné dans [Cohen 1993], les forces de pression peuvent s'interpréter comme le gradient d'un terme d'énergie de surface E_{surf} et donc comme dérivant d'une énergie potentielle E_{surf} . Plus précisément, nous définissons E_{surf} tel que $E_{surf} = \alpha S_{\mathbf{q}}$ avec $S_{\mathbf{q}}$ la surface à l'intérieur du contour actif et α un scalaire. Selon le signe de α , la surface va croître ou décroître de façon à réaliser une initialisation à l'intérieur ou à l'extérieur du contour réel. Dans le cas général, le calcul de ce terme de surface est complexe et n'a donc pas d'intérêt. En revanche, dans le cas d'une description du contour par les descripteurs de Fourier, il est très simple à calculer par le théorème de Green-Riemann. Par conséquent,

$$S_{\mathbf{q}} = \pi \sum_{k=1}^h k(a_k d_k - b_k c_k), \quad (3.39)$$

et ainsi il devient très simple d'intégrer ce nouveau terme d'énergie potentielle E_{surf} pour obtenir la nouvelle équation d'évolution du contour actif en calculant le terme $\partial E_{surf} / \partial \mathbf{q}$ et en l'intégrant dans (3.24). Il vient en effet [Li 2011b]

$$\frac{\partial E_{surf}}{\partial \mathbf{q}} = \mathbf{S} \mathbf{p}, \quad (3.40)$$

avec $\mathbf{S} = \alpha \text{diag}(0, \pi \cdots h\pi, \pi \cdots h\pi, 0, \pi \cdots h\pi, \pi \cdots h\pi)$
 et $\mathbf{p} = (0, d_1 \cdots d_h, -c_1 \cdots -c_h, 0, -b_1 \cdots -b_h, a_1 \cdots a_h)$.
 Dans ces conditions, (3.29) devient

$$\mathbf{M} \ddot{\mathbf{q}} + \mathbf{C} \dot{\mathbf{q}} + \mathbf{K} \mathbf{q} + \mathbf{S} \mathbf{p} = - \int_{C_{\mathbf{q}}} \nabla \mathbf{E}_{im}^{\top}(u) \frac{\partial C(u)_{\mathbf{q}}}{\partial q_i} du, \quad (3.41)$$

et (3.31) s'écrit maintenant

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \Delta t \mathbf{C}^{-1}(\mathbf{Q}_{im}(\mathbf{q}_k) - \mathbf{K} \mathbf{q}_k - \mathbf{S} \mathbf{p}_k). \quad (3.42)$$

La figure 3.12 présente les résultats obtenus avec le terme d'énergie de surface qui permet d'effectuer ainsi une initialisation grossière à l'intérieur ou à l'extérieur du contour réel.

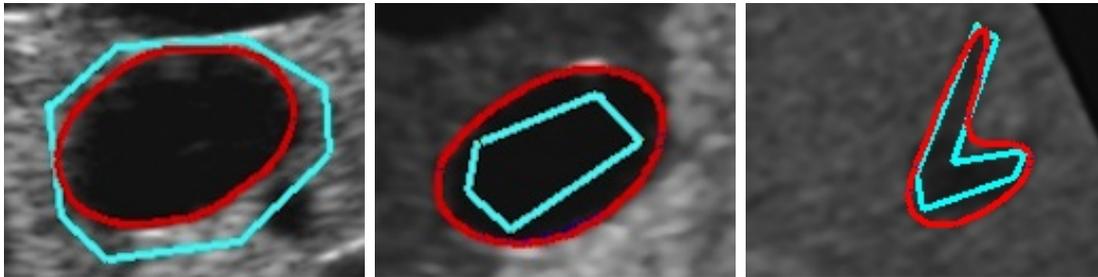


FIGURE 3.12 – Résultats du contour actif basé sur les descripteurs de Fourier dans des images échographiques avec le terme d'énergie de surface permettant d'augmenter la robustesse à une initialisation grossière (nombre d'harmoniques $h = 6$). Le contour initialisé avant convergence est affiché en bleu.

En pratique, la méthode du contour actif n'est cependant pas capable de suivre la section d'élément anatomique qui se modifie trop rapidement dans l'image lors du déplacement de la sonde. Par exemple, le suivi d'une artère est difficile à effectuer car la forme de sa section peut passer brutalement d'un cercle à un segment lorsqu'une rotation selon l'axe propre de la sonde est appliquée. Un autre problème qui peut être rencontré est un changement de topologie de la section observée et plus particulièrement une séparation de la section de l'élément anatomique en deux parties. Dans ce cas, un seul contour actif paramétrique n'est pas capable de suivre les deux sections. Il est donc nécessaire d'adapter l'algorithme de segmentation pour tenir compte de ce changement de topologie de l'organe observé.

Le principe général de l'algorithme que nous proposons dans [Li 2011b] est d'utiliser, en complément du contour actif, un algorithme d'extraction de régions afin de détecter les changements de topologie et générer automatiquement des points d'initialisation du contour actif. Cet algorithme consiste dans un premier temps à effectuer un seuillage adaptatif de l'image puis à détecter l'ensemble des pixels connexes appartenant à la région d'intérêt après initialisation d'un germe à l'intérieur de cette dernière par la méthode de codage de Freeman [Freeman 1977]. Etant donné que cette extraction de région est réalisée sur une image binaire, elle a l'avantage d'être très rapide mais l'inconvénient d'être peu précise. Cependant, cette méthode d'extraction n'est pas utilisée pour remplacer le contour actif, mais pour initialiser un ou plusieurs contours actifs en fonction du changement de topologie ou de variation brutale de la section. Le germe est tout d'abord indiqué dans l'image initiale par le médecin qui clique avec sa souris simplement sur un point appartenant à la section de l'élément anatomique qui l'intéresse. L'algorithme d'extraction de région fournit ensuite l'ensemble des points du contour initialisant ainsi le contour actif proche du contour réel de l'objet d'intérêt.

Afin de prendre en compte une modification éventuelle de la topologie de la section dans les images suivantes de la séquence, c'est à dire la séparation de la section en deux parties ou la fusion de deux sections, nous proposons d'utiliser deux contours actifs dont l'activation ou la désactivation est gérée automatiquement en fonction de l'évolution de la topologie observée. Le principe pour détecter le changement de topologie est de comparer le barycentre de 2 régions extraites à partir de 2 germes G1 et G2 (voir 3.13) qui ont été initialisés aux deux extrémités de l'axe principal de l'élément anatomique d'intérêt. Si les barycentres des deux

régions sont identiques, il n'y a pas de changement de topologie. Par contre, s'ils sont différents, il y a alors changement de topologie. Avec ce critère de détection, il est alors possible d'ajouter un contour actif supplémentaire dont l'initialisation est réalisée automatiquement à partir des points de contour fournis par l'algorithme d'extraction de région. Le ou les contours actifs sont également réinitialisés automatiquement si une variation brutale de la section est observée entre deux images consécutives. Les figures 3.13 et 3.14 présentent les résultats obtenus pour ces deux situations.

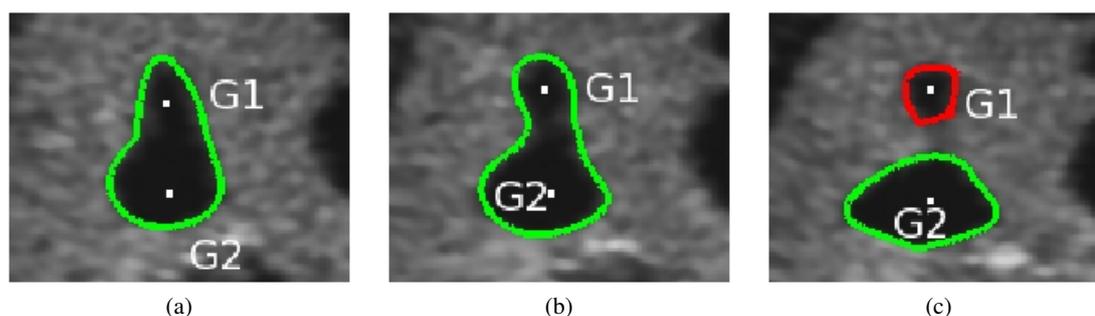


FIGURE 3.13 – Suivi de la veine porte par le contour actif basé sur les descripteurs de Fourier avec gestion du changement topologique (nombre d'harmoniques $h = 6$). (a) Contour détecté dans l'image initiale. (b) Contour détecté pour une image intermédiaire. (c) Contours détectés après séparation en deux sections.

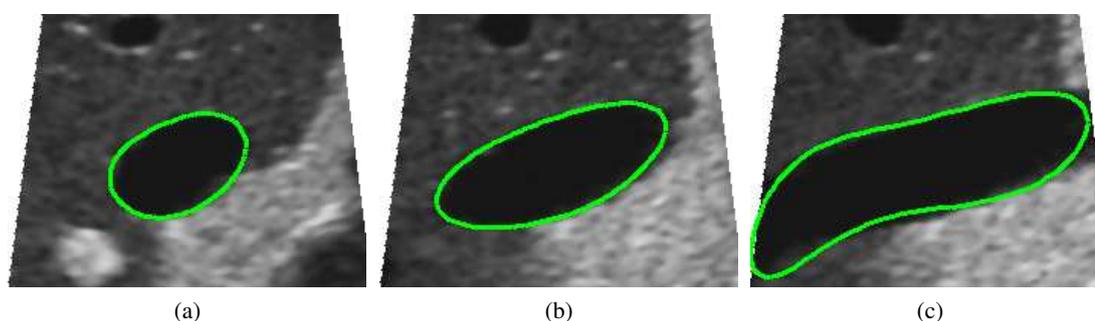


FIGURE 3.14 – Suivi de la vésicule biliaire par le contour actif basé sur les descripteurs de Fourier avec gestion de la variation brutale de section (nombre d'harmoniques $h = 6$). (a) Contour détecté dans l'image initiale. (b) Contour détecté pour une image intermédiaire. (c) Contour détecté après variation brutale.

3.6 Application à la télé-échographie robotisée

Apparue au milieu des années 90, la télé-échographie est une modalité d'imagerie qui permet à un médecin échographiste de réaliser l'examen d'un patient situé à distance dans un milieu isolé ou hostile. Le concept repose sur l'existence d'un site distant où se trouve le patient et d'un site expert où exerce le médecin. Ces deux sites sont connectés au moyen d'un réseau

de communication de type internet ou via une liaison satellite afin de retransmettre en direct le flux d'images échographiques au médecin. A ses débuts, la télé-échographie nécessitait l'emploi d'un technicien paramédical du côté du patient pour manipuler la sonde échographique sous les ordres oraux transmis par le médecin expert [Chimiak 1996]. Depuis une dizaine, d'années plusieurs équipes de recherche ont développé et proposé des systèmes robotiques de télé-opération permettant d'actionner la sonde à distance au moyen de robot légers placés sur le corps du patient [Masuda 2001, Vilchis 2003, Delgorgue 2005, Nouaille 2010]. Le principal avantage est de permettre au médecin de contrôler directement et efficacement la sonde par le biais d'une interface haptique matérialisée par une sonde fictive qu'il actionne manuellement et d'éviter ainsi les éventuels problèmes de désynchronisation entre ses ordres et les gestes du technicien. En pratique, les déplacements de la sonde fictive sont mesurés par cette interface haptique spécifique, également appelée robot maître, puis sont transmis en temps réel au robot esclave actionnant la sonde pour être dupliqués.

Le robot ESTELE commercialisé par la société française RoboSoft est un robot léger qui a été validé pour une utilisation en télé-échographie abdominale et fœtale suite à une campagne d'essais cliniques conduite dans 4 centres médicaux [Arbeille 2008]. Comme le montre la figure 3.15a, un assistant paramédical est présent pour porter le robot lors de l'examen et pour le pré-positionner initialement sur le patient de façon à fournir au médecin expert une première vue échographique de l'élément anatomique d'intérêt. Il existe cependant deux sources de perturbation qui peuvent entraîner la perte visuelle de l'élément anatomique d'intérêt. La première est le déplacement des organes dû au mouvement physiologique du patient et la seconde est le mouvement involontaire du patient ou du technicien portant le robot. Si cette perte de visibilité arrive, le médecin doit alors faire appel au technicien pour repositionner correctement le robot.

C'est dans ce contexte applicatif et dans le cadre du projet ANR PROSIT [Fonte 2010], dans lequel je suis responsable scientifique du partenaire INRIA, que nous avons développé de nouvelles fonctionnalités d'assistance à l'examen médical basées sur des asservissements visuels échographiques [Li 2012b]. Le système robotique utilisé dans ces travaux, dénommé robot PROSIT, est une version modifiée du robot ESTELE [Nouaille 2010] présenté à la figure 3.15b. Ce manipulateur esclave est un robot série à 4 ddl dont la cinématique à poignet sphérique permet d'orienter la sonde autour de son point de contact avec le patient selon les 3 angles de rotation. Une translation le long de l'axe principal de la sonde permet également de maintenir une force de contact entre la sonde et le patient.

Afin d'assister le médecin lors de l'examen, dans les travaux de thèse de Tao Li, nous avons proposé d'ajouter au système robotique PROSIT deux modes de pilotage automatique de la sonde qui sont activés sur demande du médecin. Le premier permet de positionner automatiquement la sonde afin de récupérer une section préalablement apprise et le second maintient automatiquement la visibilité d'un élément anatomique d'intérêt lors de la télé-opération. Ces deux fonctionnalités d'assistance sont décrites dans la suite de cette section.

3.6.1 Tâche de récupération de section

L'objectif de la tâche de récupération est d'orienter de manière autonome la sonde afin de retrouver la section d'un élément anatomique qui a été indiquée dans une image préalablement mémorisée durant la télé-opération. La méthode proposée repose sur un asservissement visuel

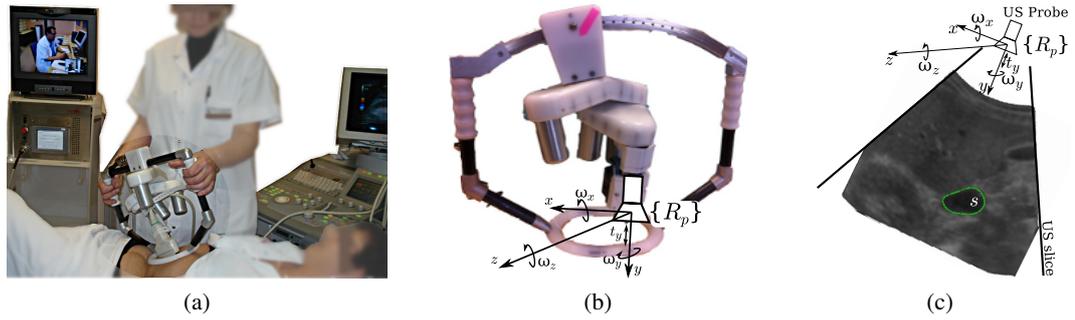


FIGURE 3.15 – (a) Le système robotique ESTELE de télé-échographie en action. (b) Le robot PROSIT à 4 ddl porteur de sonde (version modifiée du robot ESTELE). (c) La section de l'organe d'intérêt observé dans l'image échographique 2D.

qui utilise comme information d'entrée une combinaison des moments géométriques. Ces moments sont alors calculés à partir du contour d'un organe d'intérêt fourni par l'algorithme de segmentation par contour actif utilisant les descripteurs de Fourier (section 3.5).

Dans un premier temps, le médecin effectue son exploration par télé-opération pendant laquelle il peut enregistrer à sa guise une ou plusieurs coupes. Il peut ensuite à tout moment décider d'activer la tâche de positionnement automatique de la sonde pour retrouver une des coupes qu'il a préalablement enregistrées. La sonde est alors automatiquement orientée de manière à atteindre la coupe qui présente au mieux la section mémorisée de l'élément anatomique d'intérêt.

Comme nous l'avons indiqué, le robot PROSIT possède 4 ddl, mais la translation selon l'axe principal de la sonde est exclusivement contrôlée par la télé-opération ou par une commande en effort permettant de maintenir une force de contact désirée entre la sonde et la peau du patient. Le nombre de ddl restant à disposition étant de 3, par conséquent 3 informations visuelles ont été retenues pour caractériser la position de l'objet d'intérêt par rapport à la sonde. Nous avons plus particulièrement choisi le vecteur de primitives visuelles $\mathbf{s} = (x_g, \sqrt{a}, \alpha)$ représentant respectivement la coordonnée selon l'axe des abscisses dans l'image du barycentre de la section de l'objet, la racine carrée de l'aire de la section et l'orientation principale de cette dernière dans l'image afin de contrôler les 3 rotations du repère de la sonde. Ces informations visuelles sont calculées à partir des moments 2D de l'image. La matrice d'interaction \mathbf{L}_s de dimension 3×3 reliant la variation de \mathbf{s} aux vitesses angulaires de la sonde $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ est dérivée quant à elle à partir de celle des moments 2D que nous avons présentée dans la section 3.3. La tâche de positionnement est ensuite effectuée en appliquant la commande cinématique

$$\boldsymbol{\omega} = -\lambda \widehat{\mathbf{L}}_s^{-1} (\mathbf{s} - \mathbf{s}^*). \quad (3.43)$$

La figure 3.16 présente les résultats d'une tâche de récupération de section. Tout d'abord, le médecin télé-opère la sonde, enregistre une section d'intérêt puis poursuit la télé-opération (étape 1). Ensuite, à l'activation de la tâche de récupération, une commande en position utilisant uniquement la mesure d'orientation de la sonde est appliquée pour amener le plan de coupe à l'orientation enregistrée lors de la mémorisation de la section d'intérêt (étape 2). L'asservissement visuel est ensuite lancé pour affiner le positionnement et maintenir la section (étape

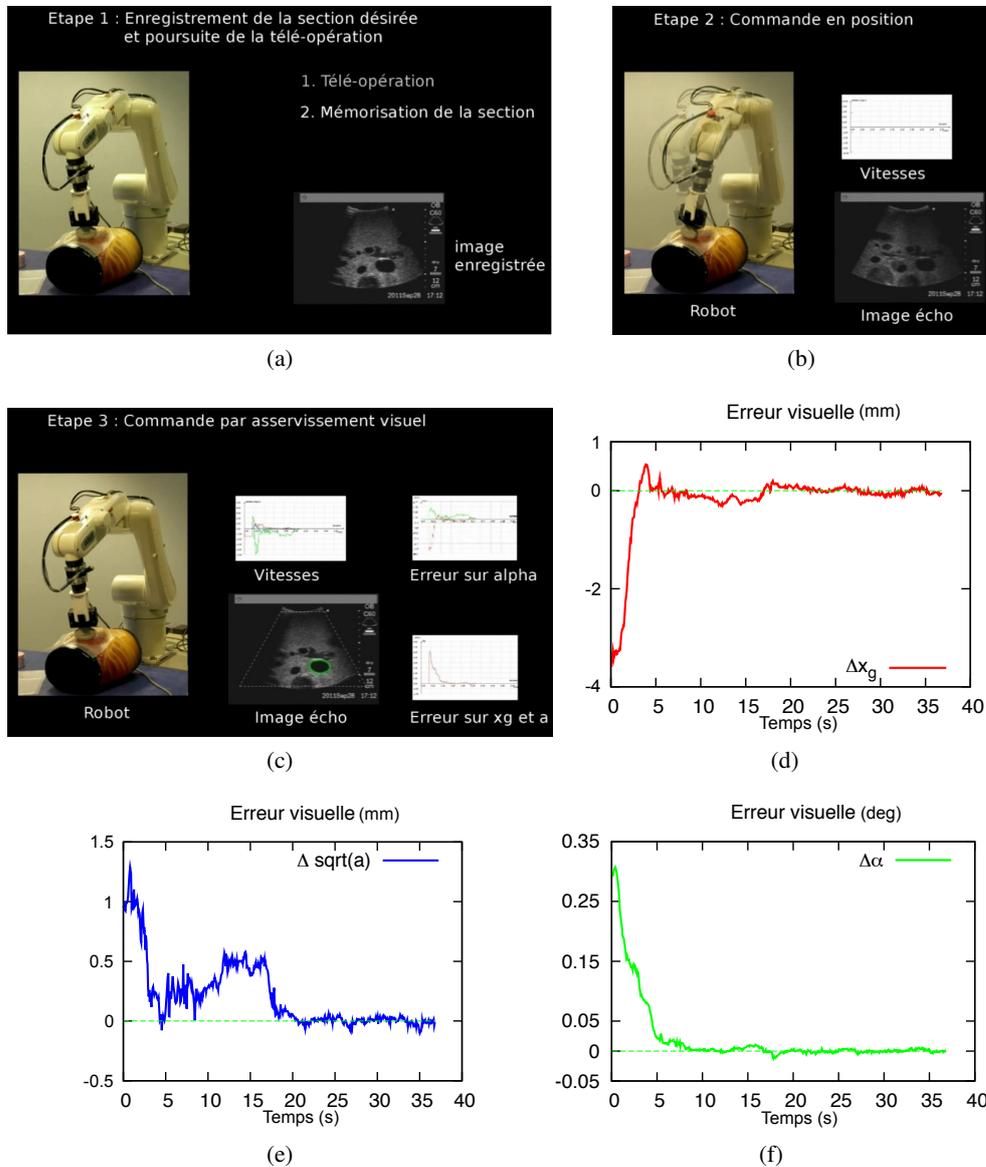


FIGURE 3.16 – Validation sur fantôme de la tâche de récupération automatique d'une section d'intérêt au moyen d'un robot manipulateur simulant la cinématique du robot PROSIT. (a)-(c) Les différentes étapes du positionnement. (d)-(f) Evolution des erreurs visuelles lors de l'étape d'asservissement visuel permettant d'affiner le positionnement.

3). Nous pouvons observer que le pré-positionnement effectué au moyen de la commande en position utilisant uniquement l'odométrie du robot (étape 2) ne permet pas d'obtenir une erreur visuelle nulle. En revanche, une fois l'asservissement visuel lancé (étape 3), l'erreur visuelle résiduelle converge vers zéro pour amener le système précisément à la section souhaitée.

3.6.2 Maintien automatique de la visibilité lors de la télé-opération

La tâche de visibilité permet d'assister le médecin lors de son exploration en maintenant automatiquement la visibilité de la section d'un élément anatomique d'intérêt dans le plan échographique durant la télé-opération. Pour lancer cette tâche, l'expert télé-opère dans un premier temps les 4 ddl du robot porteur de sonde pour atteindre la section d'un objet anatomique d'intérêt comme celle présentée à la figure 3.15c. Ensuite, il active la tâche d'assistance de visibilité en cliquant simplement avec sa souris à l'intérieur de la section de l'organe de son choix. La section est alors automatiquement segmentée par l'algorithme du contour actif basé sur les descripteurs de Fourier. Le médecin peut alors poursuivre l'exploration de l'organe en télé-opérant librement les 4 ddl de la sonde pendant que le système empêche la perte de visibilité de la section en agissant comme un guide mécanique virtuel. A tout moment, le médecin peut désactiver la tâche de maintien de visibilité pour aller explorer un autre organe.

La perte de la visibilité peut arriver lorsque la section d'intérêt sort des bords de l'image ou lorsque sa surface devient trop petite. Nous avons donc défini les contraintes de visibilité par 3 primitives visuelles $\mathbf{s} = (x_g, y_g, d)$ représentant respectivement les coordonnées dans l'image du barycentre de la section observée et la racine carrée de son aire $d = \sqrt{a}$. Ces primitives sont calculées à partir des moments 2D extraits du contour de la section. L'objectif de la tâche d'assistance est de maintenir automatiquement le barycentre de la section à l'intérieur d'une région de l'image tel que $x_g \in [x^-, x^+]$, $y_g \in [y^-, y^+]$, avec une aire de section minimale $d \geq d^-$ qui garantit une intersection entre l'organe et le plan de coupe échographique. La solution retenue pour assurer les contraintes de visibilité est d'activer un asservissement visuel lorsque au moins une primitive visuelle sort de son intervalle sûr défini par $x_g \in [x^{s-}, x^{s+}]$, $y_g \in [y^{s-}, y^{s+}]$ et $d \geq d^{s-}$ où la section reste visible, comme illustré à la figure 3.17a. L'idée est alors de ramener cette section à l'intérieur de cet intervalle si nécessaire. En revanche, si toutes les primitives sont dans leur intervalle sûr, l'asservissement visuel n'est pas activé afin de permettre au médecin de contrôler librement la sonde par télé-opération.

Ce concept d'activation et de désactivation automatique de l'asservissement visuel pour le maintien de contraintes a été introduit dans [Kermorgant 2011] pour commander les déplacements d'une caméra embarquée en combinant un asservissement visuel 2D avec un asservissement visuel 3D. En adaptant ce concept à notre application, nous avons proposé la loi de commande suivante pour garantir les contraintes de visibilité [Li 2012b] :

$$\boldsymbol{\omega} = -\lambda(\mathbf{HL}_s)^+ \mathbf{H} (\mathbf{s} - \mathbf{s}^*), \quad (3.44)$$

où $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ est la vitesse de rotation appliquée à la sonde, $\widehat{\mathbf{L}}_s$ est une estimation de la matrice 3×3 d'interaction construite à partir de la matrice d'interaction associée aux moments 2D qui relie la variation des primitives visuelles \mathbf{s} à la vitesse angulaire de la sonde, \mathbf{s}^* sont les primitives visuelles désirées que nous avons définies à l'intérieur de l'intervalle sûr et \mathbf{H} est une matrice diagonale de poids. Cette dernière est définie par $\mathbf{H} = \text{diag}(h_x, h_y, h_d)$ où $h_x, h_y,$

h_d sont des poids variables associés aux primitives visuelles. L'objectif n'est pas d'obtenir la convergence des primitives visuelles exactement vers leur valeur désirée mais de les faire revenir dans leur intervalle sûr. Pour cette raison, un poids variable est affecté à chaque primitive visuelle afin de l'asservir uniquement si elle quitte son intervalle fixé. La figure 3.17b présente la fonction de poids h considérée pour chaque primitive. A titre d'exemple, sa formulation h_x pour x_g , similaire pour les autres primitives, s'écrit

$$h_x(x_g) = \begin{cases} \frac{x_g - x^{s+}}{x^+ - x_g} & \text{si } x_g > x^{s+} \\ \frac{x_g - x^{s-}}{x^- - x_g} & \text{si } x_g < x^{s-} \\ 0 & \text{sinon.} \end{cases} \quad (3.45)$$

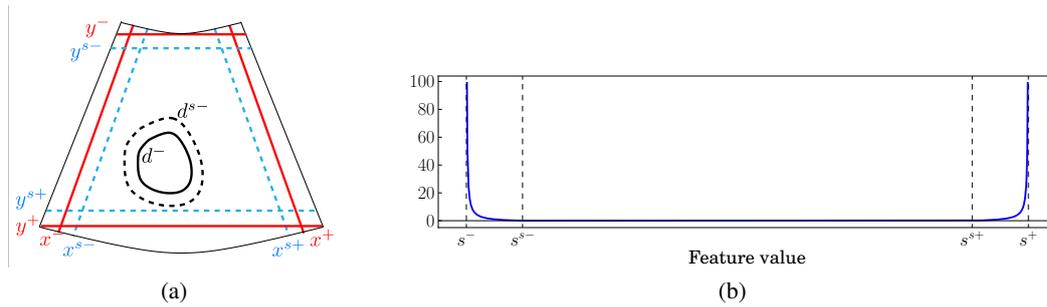


FIGURE 3.17 – (a) Contrainte de visibilité (ligne rouge) à ne pas dépasser et intervalles sûrs (ligne bleue) définis pour les coordonnées du barycentre x_g et y_g ; Valeur minimale à ne pas dépasser (contour noir) et sûre (contour en pointillé noir) définies pour l'aire de la section d . (b) Fonction de poids : h est nulle dans l'intervalle sûr et augmente aux bords.

Avec cette approche, la contrainte visuelle associée à une primitive est activée dès que son poids n'est pas nul, et désactivée lorsque la primitive considérée est revenue dans son intervalle sûr. De cette façon, uniquement le/les ddl du robot requis pour réguler la/les primitive(s) activée(s) est/sont contraint(s) par l'asservissement visuel. Il est par conséquent possible de contrôler les ddl restants par télé-opération grâce au formalisme de la redondance [Samson 1991] (section 2.2.5). La tâche de visibilité peut donc être fusionnée avec une tâche secondaire en vue de contrôler par télé-opération la sonde en appliquant la loi de commande

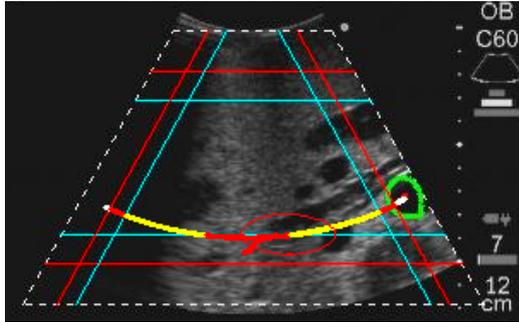
$$\boldsymbol{\omega} = -\lambda(\widehat{\mathbf{H}\mathbf{L}_s})^+ \mathbf{H}(\mathbf{s} - \mathbf{s}^*) + \mathbf{P} \boldsymbol{\omega}_m, \quad \text{avec } \mathbf{P} = \mathbf{I}_3 - (\widehat{\mathbf{H}\mathbf{L}_s})^+ \widehat{\mathbf{H}\mathbf{L}_s} \quad (3.46)$$

où $\boldsymbol{\omega}_m = (\omega_{xm}, \omega_{ym}, \omega_{zm})$ est la vitesse de rotation de la sonde fictive manipulée par le médecin durant la télé-opération. \mathbf{P} est l'opérateur qui projette les composantes de $\boldsymbol{\omega}_m$ sur le noyau de $(\widehat{\mathbf{H}\mathbf{L}_s})^+$ afin que la tâche de télé-opération ne compromette pas le succès de la tâche de maintien de visibilité si au moins une primitive visuelle est activée.

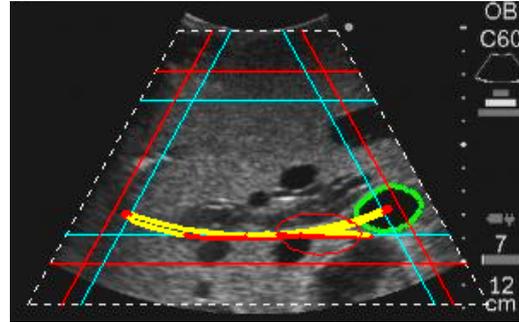
Nous pouvons noter que le rang de $\widehat{\mathbf{H}\mathbf{L}_s}$ est égal à 3 seulement si toutes les primitives visuelles sont activées dans la commande et chute progressivement à zéro lorsque l'ensemble des primitives revient dans la région sûre. Afin d'éviter l'obtention d'un comportement discontinu de la loi de commande induit par ce changement de rang, nous avons en pratique remplacé dans la loi de commande (3.46) l'opérateur de la pseudo inverse de Moore-Penrose par celui proposé dans le travail [Mansard 2009b] traitant la prise en compte d'un ensemble variable de primitives en asservissement visuel.

Comme proposé dans [Mansard 2009a], nous avons également utilisé l'opérateur général de projection qui permet de projeter partiellement les composantes de ω_m lorsque les poids de \mathbf{H} sont non nuls. Dans ce cas, la consigne de vitesse de télé-opération ω_m n'est pas ignorée mais appliquée dans les directions qui ne compromettent pas la tâche de visibilité. Une telle loi de commande a été uniquement utilisée pour des systèmes à redondance élevée [Mansard 2009a] qui considèrent un nombre limité de contraintes. Dans notre approche, la formulation retenue des poids assure les 3 contraintes tout en réalisant une tâche de télé-opération de dimension 3 avec un système possédant uniquement 3 ddl.

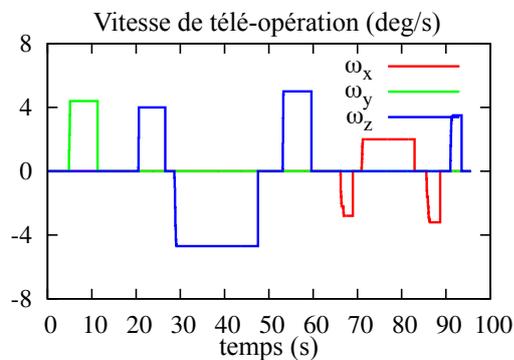
Les résultats de la télé-opération obtenus sur un fantôme abdominal sans et avec la tâche de maintien de visibilité active sont présentés à la figure 3.18. Comme nous l'avons signalé, cette tâche évite également la perte de la section d'un élément anatomique provoquée par le mouvement physiologique du patient ou un mouvement involontaire du technicien portant le robot. Cette assertion a été mise en évidence par les résultats présentés à la figure 3.19 obtenus sur un homme dans le cas où aucune consigne de télé-opération n'a été appliquée.



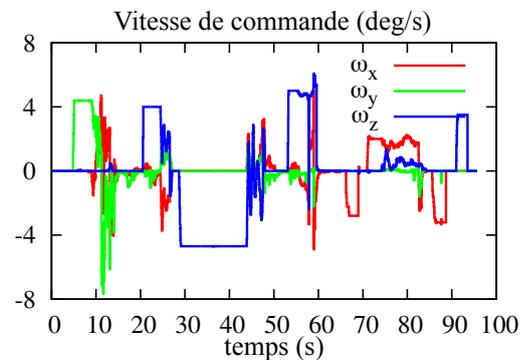
(a) Trajectoire obtenue dans l'image sans la tâche de visibilité. La portion blanche sort de l'intervalle de visibilité délimité par les droites rouges et l'aire de la section représentée à la position initiale par le contour rouge devient très faible (contour vert).



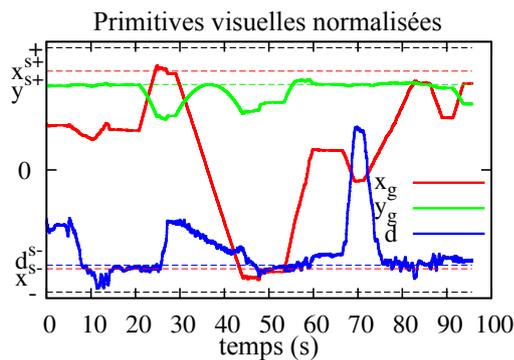
(b) Trajectoire obtenue dans l'image avec la tâche de visibilité active. La trajectoire reste dans l'intervalle de visibilité et l'aire de la section reste supérieure à la valeur minimale fixée.



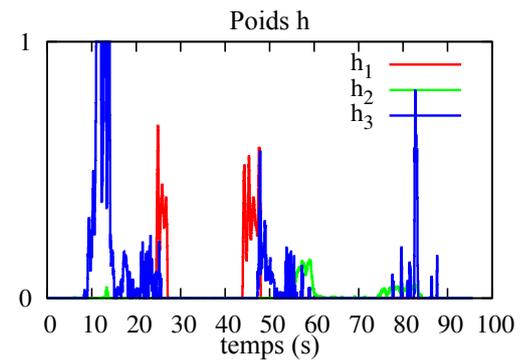
(c) Vitesse de télé-opération ω_m transmise en entrée de la commande ou appliquée directement à la sonde lorsque la tâche de visibilité n'est pas active.



(d) Vitesse appliquée à la sonde ω avec la tâche de visibilité active permettant de suivre au mieux la vitesse de consigne ω_m tout en respectant les contraintes de visibilité.



(e) Primitives visuelles normalisées respectant les contraintes durant la tâche de visibilité.



(f) Poids associés aux primitives visuelles s'activant lorsque ces dernières sortent de l'intervalle sûr.

FIGURE 3.18 – Résultats obtenus sur fantôme abdominal lors de l'exploration d'un organe d'intérêt par télé-opération sans (a),(c) et avec (b)-(f) la tâche de maintien de visibilité active.

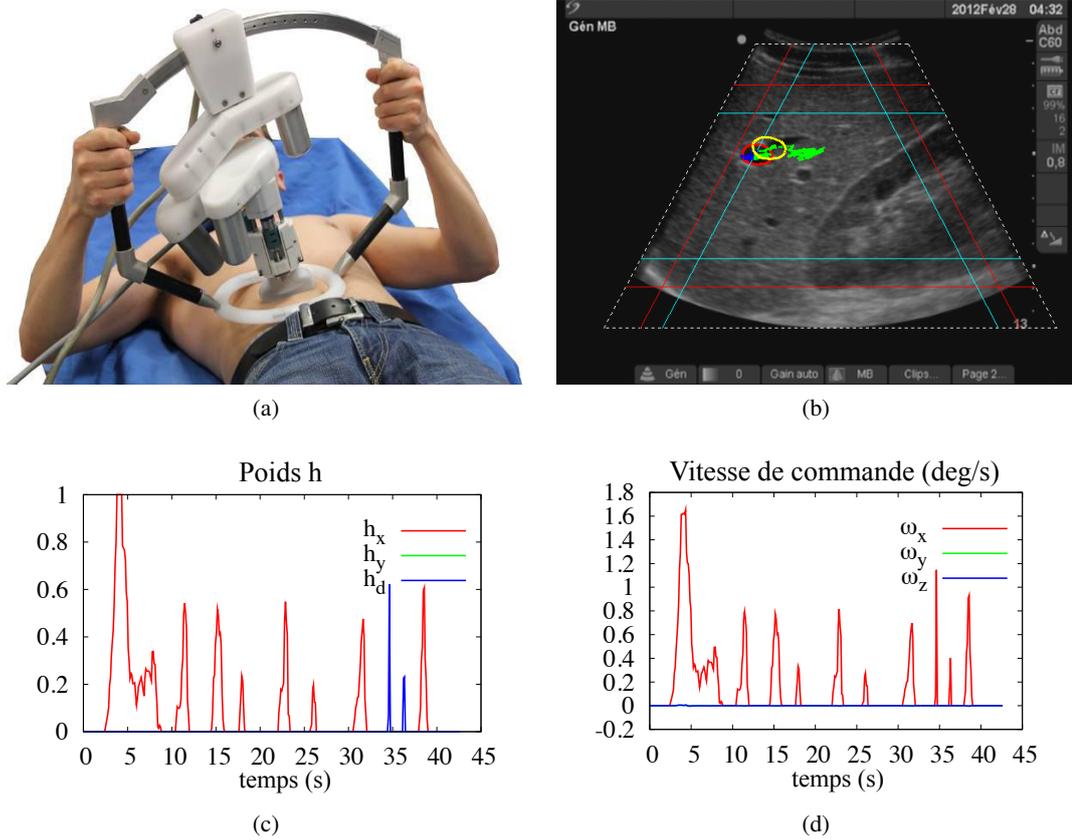


FIGURE 3.19 – Résultat obtenus sur volontaire humain. (a) Vue externe du robot au contact de l'abdomen. (b) Trajectoire de la section d'intérêt dont les positions initiales et finales sont représentées respectivement par le contour jaune et rouge. La portion bleue de la trajectoire correspond au déplacement du barycentre effectué automatiquement par la tâche de visibilité pour le ramener dans la région sûre de l'image. Dans cette région, la trajectoire verte représente alors uniquement la perturbation due au mouvement physiologique puisque aucune consigne de télé-opération n'a été appliquée durant l'expérimentation. (c) Poids associés aux primitives visuelles s'activant lorsque ces dernières sortent de l'intervalle sûr. Nous pouvons noter que le poids associé à l'aire de la section a également été activé à $t = 70s$ pour maintenir une valeur minimale de la surface, permettant ainsi de maintenir l'intersection entre le plan de coupe et l'objet d'intérêt. (d) Vitesse ω appliquée à la sonde.

3.7 Conclusion

Ce chapitre a porté sur l'étude de méthodes d'asservissement visuel 2D fondées sur l'exploitation d'informations géométriques extraites de l'image, et adaptées aux problématiques de l'imagerie échographique temps réel. Des solutions génériques ont été présentées en vue de positionner automatiquement, selon 6 degrés de liberté, une sonde ultrasonore par rapport à l'environnement observé. L'investigation a porté sur différents choix de primitives géométriques 2D afin d'obtenir un comportement efficace et robuste de la commande d'un robot porteur de sonde. Une modélisation complète de la matrice d'interaction, liant la variation de ces primitives visuelles dans l'image au mouvement complet de la sonde, a été définie pour des primitives de type point, contour et moments de la section. Ces travaux ont permis d'intégrer à un dispositif robotique de télé-échographie des fonctionnalités d'assistance au diagnostic, dont entre autres une tâche de récupération automatique de section et un mode de maintien de la visibilité d'un organe lors de la télé-opération de la sonde.

Des améliorations ont en outre été apportées par une phase de segmentation faiblement complexe, basée sur une approche paramétrique intégrant des descripteurs de Fourier. La chaîne de traitement ainsi conçue a prouvé son efficacité lors d'expérimentations. Toutefois, la succession d'étapes de calculs et tout particulièrement l'exécution de l'algorithme de segmentation de l'image complexifie le processus et atteint les limites d'un système temps réel. Afin de palier cet état de fait, des solutions alternatives présentées dans le chapitre suivant ont été élaborées.

Asservissement visuel échographique : approches utilisant l'information dense

Les approches décrites dans le chapitre précédent reposent toutes sur l'utilisation d'informations visuelles géométriques qui sont extraites de l'image par des techniques de segmentation. Ces informations visuelles sont données par des points, des contours ou les moments d'une section d'un organe d'intérêt. Cette phase de segmentation, indispensable au système, ne peut toutefois être utilisée pour des images très faiblement structurées dans la mesure où aucune information géométrique ne serait décelable dans un contexte temps réel.

Afin de s'affranchir de cette étape de segmentation et de considérer des images échographiques faiblement structurées, nous avons donc orienté nos travaux de recherche sur des approches totalement différentes qui utilisent directement l'information dense constituant l'image échographique de mode B.

Ce chapitre présente différentes solutions reposant sur l'utilisation de cette information dense. Une approche basée sur la corrélation spatiale de la texture de type «speckle» contenue dans l'image échographique est proposée à la section 4.1 en vue de réaliser une compensation automatique du mouvement d'une coupe de référence par un asservissement visuel hybride 2D-3D. La section 4.2 s'intéresse quant à elle à une approche utilisant directement l'intensité des pixels en entrée de la loi de commande. Les applications considérées concernent le positionnement du capteur par rapport à la scène et le suivi de coupes anatomiques en vue de stabiliser l'image observée. Cette méthode a en outre été étendue pour l'usage de différents types de sonde : 2D, 3D et bi-plans. La prise en compte des déformations des tissus mous est adressée à la section 4.3 où une approche, permettant d'extraire les déformations d'une région d'intérêt 3D dans une séquence de volumes échographiques, est proposée pour compenser automatiquement le mouvement rigide du volume observé par une sonde 3D.

4.1 Approche basée sur la corrélation spatiale du « speckle » échographique

Les collaborations que j'ai pu initier lors de ma mise à disposition à « The Johns Hopkins University » m'ont amené à proposer une méthode utilisant directement l'information du « speckle » présent dans l'image échographique pour réaliser la poursuite d'un organe en mouvement par une sonde robotisée. En effet, le speckle contenu dans les images échographiques ne peut être considéré comme un bruit car il est spatialement cohérent puisqu'il résulte de réflexions multiples de l'onde ultrasonore dans les tissus organiques. De plus, la coupe échographique n'est pas parfaite mais comporte une certaine largeur d'élévation (direction orthogonale au plan image) qui introduit une corrélation du speckle entre deux coupes relativement proches. La figure 4.1 présente l'évolution du coefficient de corrélation normalisée ρ mesuré en fonction de la distance d'élévation d du plan de coupe pour 25 motifs (patches) disposés dans l'image selon une grille de dimension 5×5 . Ces courbes ont été obtenues suite à l'acquisition de coupes parallèles sur un fantôme simulant des tissus mous et en calculant le coefficient de corrélation normalisée entre motifs (patches) par

$$\rho(d) = \frac{\sum_{i=1}^m \sum_{j=1}^n (I_0(i, j) - \bar{I}_0)(I_d(i, j) - \bar{I}_d)}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n (I_0(i, j) - \bar{I}_0)^2 \sum_{i=1}^m \sum_{j=1}^n (I_d(i, j) - \bar{I}_d)^2}}, \quad (4.1)$$

où \bar{I}_0 , \bar{I}_d représentent respectivement la valeur moyenne de l'intensité des patches \mathbf{I}_0 et \mathbf{I}_d de taille $m \times n$ pixels dans l'image.

En se référant à ces courbes il est ainsi possible d'estimer la distance d'élévation entre deux coupes à partir des mesures de corrélation entre motifs.

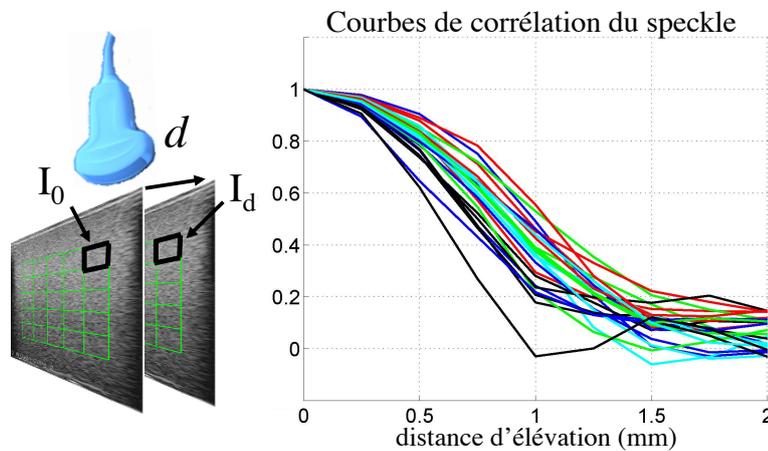


FIGURE 4.1 – Mesure de la corrélation entre motifs (patches I_0 et I_d) en fonction de la distance d'élévation d du plan de coupe pour 25 motifs disposés sur une grille de 5×5 .

Initialement, le speckle a été utilisé en imagerie ultrasonore pour estimer des écoulements multidimensionnels dans des images échographiques 2D [Bohs 2000]. Plus récemment, plusieurs auteurs [Chang 2003, Gee 2006] ont proposé des techniques de corrélation du speckle pour réaliser, avec une simple sonde 2D, une imagerie 3D de type main libre sans utiliser de capteur de position pour localiser la sonde. Un cadre probabiliste a également été défini dans

[Laporte 2007] pour estimer à partir de la corrélation du speckle la distance d'élévation entre des coupes éloignées. Toutes ces techniques sont basées sur l'utilisation de courbes expérimentales pré-enregistrées de la corrélation du speckle obtenues par des mesures effectuées sur des tissus mous réels ou des fantômes imitant le chatoiement du speckle. Dans [Boctor 2005], une méthode de suivi de speckle effectue l'étalonnage en ligne d'une sonde échographique 2D lors d'une application de chirurgie guidée par l'image. La corrélation du speckle a également été largement utilisée en élastographie ultrasonore pour estimer le déplacement de diffuseurs biologiques induit par une pression mécanique appliquée sur les tissus mous [Lubinski 1999, Revell 2005].

Dans nos travaux, l'idée est d'exploiter le speckle contenu dans l'image pour estimer la position relative entre la coupe courante fournie par la sonde échographique 2D et une coupe de référence apprise à un instant donné et évoluant avec le mouvement du patient. Le principe de la méthode d'estimation associée est présenté dans la section suivante. Nous avons ensuite développé un asservissement visuel mixte 2D-3D (section 4.1.2) pour commander les déplacements de la sonde de manière à minimiser la position relative entre la coupe courante et la coupe cible et réaliser ainsi une stabilisation automatique de l'image.

4.1.1 Estimation du mouvement

Pour localiser les coupes dans l'espace, nous définissons deux repères cartésiens \mathcal{R}_p et \mathcal{R}_t qui sont respectivement centrés sur la coupe courante notée $\{p\}$ et la coupe cible désignée par $\{t\}$. La position relative entre la coupe cible $\{t\}$ et la coupe courante $\{p\}$ est donnée par une combinaison successive de deux transformations homogènes telles que ${}^p\mathbf{H}_t = {}^p\mathbf{H}_c {}^c\mathbf{H}_t$, où ${}^p\mathbf{H}_c$ et ${}^c\mathbf{H}_t$ décrivent respectivement les déplacements dans le plan et en dehors du plan de la coupe cible par rapport à la coupe courante. \mathcal{R}_c désigne donc le repère d'une coupe virtuelle intermédiaire notée $\{c\}$. Le déplacement dans le plan est caractérisé par les translations t_x et t_y le long des axes X et Y du plan image observé et l'angle γ autour de l'axe Z orthogonal à l'image tel que

$${}^p\mathbf{H}_c = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 & t_x \\ \sin(\gamma) & \cos(\gamma) & 0 & t_y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.2)$$

Le déplacement relatif introduit par le mouvement en dehors du plan est quant à lui défini par la distance d'élévation t_z le long de l'axe Z et les rotations successives α et β autour des axes Y et X , de sorte que

$${}^c\mathbf{H}_t = \begin{pmatrix} \cos(\alpha) & \cos(\alpha)\sin(\beta) & \sin(\alpha)\cos(\beta) & 0 \\ 0 & \cos(\beta) & -\sin(\beta) & 0 \\ -\sin(\alpha) & \cos(\alpha)\sin(\beta) & \cos(\alpha)\cos(\beta) & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.3)$$

Mouvement dans le plan

Pour extraire le mouvement rigide observé dans le plan image, nous utilisons la technique de suivi différentiel proposée par [Hager 1998] dont nous rappelons brièvement le principe.

Cette technique consiste à estimer le vecteur de paramètres μ d'une fonction $f(\mathbf{x}; \mu)$ qui décrit la transformation géométrique des coordonnées des pixels $\mathbf{x} = (x, y)$ pour passer de l'image cible de référence acquise à l'instant $t_0 = 0$ à l'image courante observée après déplacement (figure 4.2). Pour un mouvement rigide, ce vecteur est défini par $\mu = (u_x, u_y, \gamma)$ où u_x, u_y

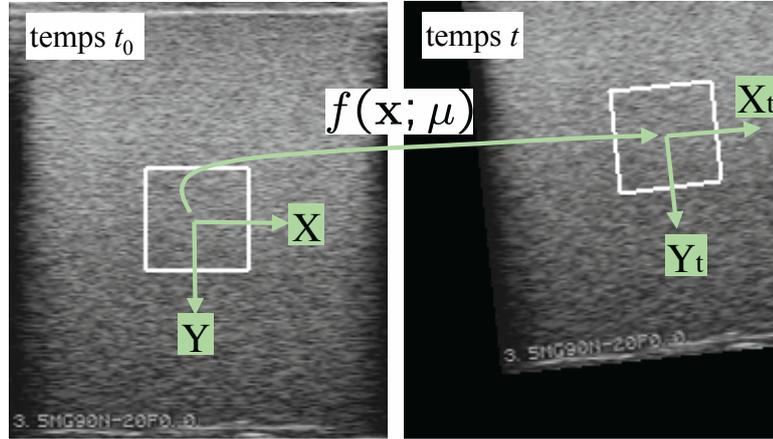


FIGURE 4.2 – (gauche) Image de référence acquise à l'instant $t_0 = 0$ avec la région d'intérêt à suivre. (droite) Image courante obtenue après application d'un mouvement dans le plan de coupe caractérisé par la transformation $f(\mathbf{x}; \mu)$.

sont les translations le long des axes X et Y de l'image et γ et l'angle autour de l'axe Z . La fonction de transformation est dans ce cas définie par

$$f(\mathbf{x}; u_x, u_y, \gamma) = \mathbf{R}(\gamma)\mathbf{x} + \mathbf{u}, \quad (4.4)$$

où $\mathbf{R}(\gamma)$ est une matrice de rotation 2×2 d'angle γ et $\mathbf{u} = (u_x, u_y)$ est le vecteur de translation. Le principe de la méthode consiste à estimer les paramètres de déplacement μ qui minimisent la fonction de coût

$$\mathcal{O}(\mu) = \|\mathbf{I}(\mu, t) - \mathbf{I}(0, t_0)\|^2, \quad (4.5)$$

où $\mathbf{I}(0, t_0)$ est un vecteur contenant la valeur des intensités de N pixels appartenant à une région d'intérêt cible de l'image à l'instant $t = 0$ et $\mathbf{I}(\mu, t)$ est un vecteur contenant les valeurs des intensités des N pixels mesurées à l'instant t après avoir appliqué la transformation (4.4) sur leurs coordonnées selon

$$\mathbf{I}(\mu, t) = \begin{bmatrix} I(f(\mathbf{x}_1, \mu), t) \\ \vdots \\ I(f(\mathbf{x}_N, \mu), t) \end{bmatrix}. \quad (4.6)$$

En récrivant (4.5) pour une image acquise à l'instant $t + \tau$ en fonction de la variation des paramètres $\delta\mu$ telle que $\mu(t + \tau) = \mu(t) + \delta\mu$, on a

$$\mathcal{O}(\delta\mu) = \|\mathbf{I}(\mu + \delta\mu, t + \tau) - \mathbf{I}(0, t_0)\|^2. \quad (4.7)$$

En approchant cette expression par une décomposition de Taylor au premier ordre, nous obtenons

$$\mathcal{O}(\delta\mu) \approx \|\mathbf{M}\delta\mu + \mathbf{I}(\mu, t + \tau) - \mathbf{I}(0, t_0)\|^2, \quad (4.8)$$

où \mathbf{M} est une matrice jacobienne reliant la variation de \mathbf{I} à celle du vecteur de paramètres $\boldsymbol{\mu}$. $\mathbf{M}(\boldsymbol{\mu})$ s'exprime alors par

$$\mathbf{M}(\boldsymbol{\mu}) = \begin{bmatrix} \nabla I(\mathbf{x}_1, t_0)^T f_{\mathbf{x}}(\mathbf{x}_1, \boldsymbol{\mu})^{-1} f_{\boldsymbol{\mu}}(\mathbf{x}_1, \boldsymbol{\mu}) \\ \vdots \\ \nabla I(\mathbf{x}_N, t_0)^T f_{\mathbf{x}}(\mathbf{x}_N, \boldsymbol{\mu})^{-1} f_{\boldsymbol{\mu}}(\mathbf{x}_N, \boldsymbol{\mu}) \end{bmatrix}, \quad (4.9)$$

où le vecteur $\nabla I(\mathbf{x}, t_0)^T$ correspond au gradient de l'intensité du pixel à la position $\mathbf{x} = (x, y)$ dans l'image de référence et $f_{\mathbf{x}}, f_{\boldsymbol{\mu}}$ sont respectivement les dérivées partielles de $f(\mathbf{x}; \boldsymbol{\mu})$ par rapport à \mathbf{x} et $\boldsymbol{\mu}$. Le produit $f_{\mathbf{x}}^{-1} f_{\boldsymbol{\mu}}$ de (4.9) est alors pour le cas du modèle de transformation retenu (4.4) donné par

$$f_{\mathbf{x}}^{-1} f_{\boldsymbol{\mu}} = \begin{bmatrix} 1 & 0 & -y \\ 0 & 1 & x \end{bmatrix} \begin{bmatrix} \mathbf{R}(-\gamma) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}. \quad (4.10)$$

La valeur de $\boldsymbol{\delta\mu}$ qui minimise la somme des différences au carré des intensités entre les deux images s'obtient en annulant le gradient de $\mathcal{O}(\boldsymbol{\delta\mu})$, ce qui donne

$$\boldsymbol{\delta\mu} = -\mathbf{M}^+(\mathbf{I}(\boldsymbol{\mu}, t + \tau) - \mathbf{I}(0, t_0)), \quad (4.11)$$

où \mathbf{M}^+ correspond à la pseudo-inverse de \mathbf{M} . Le vecteur des paramètres de déplacement est finalement donné par

$$\boldsymbol{\mu}(t + \tau) = \boldsymbol{\mu}(t) + \boldsymbol{\delta\mu}. \quad (4.12)$$

En pratique, les expressions (4.11) et (4.12) sont successivement calculées durant plusieurs itérations jusqu'à ce que $\|\boldsymbol{\delta\mu}\|^2$ devienne inférieur à un seuil ϵ . Nous obtenons ainsi les paramètres de la transformation homogène ${}^p\mathbf{H}_c$ avec $t_x = l_x u_x$ et $t_y = l_y u_y$ où l_x et l_y sont les facteurs d'échelle de l'image.

Pour obtenir plus de détails sur cette méthode, nous invitons le lecteur à se référer à l'article [Hager 1998]. D'autres méthodes basées sur le même principe ont également été proposées dans la littérature telle la technique de [Benhimane 2004] utilisant une minimisation du second ordre pour obtenir une convergence plus rapide des paramètres de déplacement. Un état de l'art approfondi comparant les différentes approches est notamment présenté dans [Baker 2004].

Mouvement hors du plan

Nous nous basons sur la propriété de corrélation du speckle pour estimer le déplacement hors plan de la coupe cible $\{t\}$ par rapport à la coupe intermédiaire $\{c\}$. Le principe consiste dans un premier temps à construire l'image en niveau de gris de la coupe intermédiaire en appliquant à l'image courante la transformation (4.4) du mouvement estimé dans le plan. Ensuite, la mesure de corrélation du speckle est utilisée pour estimer la distance d'élévation de chaque motif (patch) fixé sur la grille de l'image cible par rapport à leur correspondant positionné selon la même disposition dans l'image intermédiaire aux coordonnées fixées x, y . Cette estimation permet ainsi de fournir la coordonnée z manquante et d'obtenir avec les coordonnées 2D, la position 3D du centre de chaque patch de la coupe cible exprimée dans le repère \mathcal{R}_c de la coupe intermédiaire $\{c\}$. Les coordonnées 3D d'un minimum de 3 motifs non colinéaires permettent

ensuite de déterminer, à l'aide d'une résolution par moindres carrés, les paramètres (a, b, c, r) de l'équation du plan de coupe cible tels que

$$ax + by + cz + r = 0. \quad (4.13)$$

A partir de ces paramètres, il est alors possible de calculer les paramètres t_z , α et β définissant la transformation homogène ${}^c\mathbf{H}_t$ qui caractérise le déplacement hors du plan (voir [Krupa 2007a] pour le détail calculatoire).

Néanmoins, la difficulté porte sur l'affectation de la distance d'élévation de chaque patch en fonction de la mesure de corrélation du speckle. La plupart des travaux que nous avons mentionnés précédemment reposent sur l'utilisation de tables dans lesquelles sont mémorisées les courbes de corrélation du speckle pour toutes les positions 2D des patches dans l'image. Un mécanisme de recherche sur table (look up table) permet alors de fournir la valeur de la distance d'élévation du patch considéré en fonction de sa mesure de corrélation, de ses coordonnées 2D dans l'image et du type de tissu organique observé. Il est donc nécessaire à l'utilisateur d'effectuer un étalonnage hors ligne de ces courbes sur un échantillon du tissu ciblé avant chaque utilisation de la technique de corrélation du speckle. Afin de s'affranchir de cette étape laborieuse, la solution se base sur l'utilisation d'un modèle mathématique permettant de lier directement la distance d'élévation à la mesure de corrélation du speckle. En effet, en observant les courbes de corrélation de la figure 4.1, nous pouvons remarquer que leur évolution en fonction de la distance d peut être approchée par une fonction gaussienne

$$\rho(\mathbf{I}_0, \mathbf{I}_d) = \exp\left(\frac{-d^2}{2\sigma^2}\right), \quad (4.14)$$

où ρ est le coefficient de corrélation normalisée entre les patches \mathbf{I}_0 et \mathbf{I}_d , σ est une constante caractérisant la résolution spatiale de la granularité du speckle qui est dépendante de la profondeur du patch dans l'image échographique. A partir de ce modèle, il est ainsi possible d'obtenir une estimation $\hat{d}(\rho)$ de la distance d'élévation en fonction de la mesure de corrélation entre les deux patches et d'une estimation $\hat{\sigma}$ du paramètre σ , telle que

$$\hat{d}(\rho) = \sqrt{-2\hat{\sigma}^2 \ln(\rho)}, \quad (4.15)$$

Dans nos travaux, nous proposons de fixer le paramètre $\hat{\sigma}$ identique pour chaque courbe. Sa valeur est alors estimée à partir d'une courbe moyenne de corrélation mesurée sur un échantillon de tissu lors d'une seule étape d'étalonnage qui n'est pas réitérée par la suite. Cette approximation des courbes de corrélation n'est bien entendu pas appropriée pour être utilisée dans une application d'imagerie 3D comme cela fut l'objectif des travaux cités précédemment. En revanche, pour une application robotique de suivi de coupe, il n'est pas nécessaire d'obtenir une reconstruction précise des distances d'élévation entre patches car elles seront constamment proches de zéro lors de l'exécution de la tâche robotique de compensation du mouvement.

Cependant, cette méthode d'estimation du mouvement hors du plan ne peut pas être appliquée en l'état car la mesure du coefficient de corrélation ne permet pas d'associer un signe à la coordonnée en Z des patches. Nous avons par conséquent proposé différentes approches pour définir le signe de la distance d'élévation de chaque patch. La première méthode présentée dans [Krupa 2007a] consiste à affecter initialement un signe aléatoire à la coordonnée en Z de chaque patch avant estimation des paramètres de l'équation du plan de la coupe cible (4.13).

Un algorithme itératif permet ensuite de corriger les signes afin d'obtenir la plus faible erreur résiduelle sur l'estimation des paramètres de l'équation du plan. Deux solutions symétriques de la position du plan cible sont alors fournies à la convergence de l'algorithme (figure 4.3). La première correspond à une translation positive ($t_z > 0$) de la coupe cible selon l'axe Z du

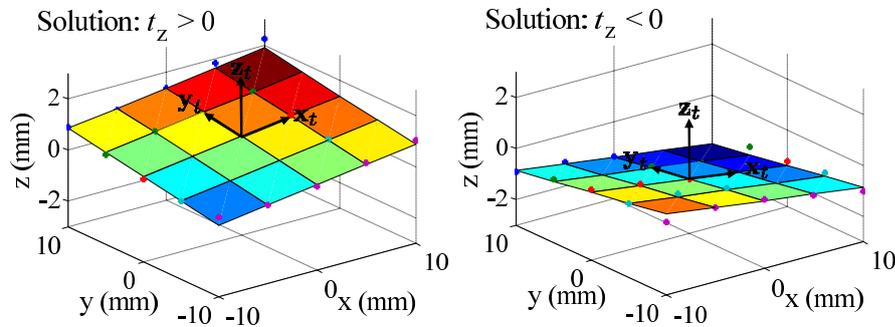


FIGURE 4.3 – Positions symétriques du plan cible fournies à la convergence de l'algorithme de correction du signe de la distance d'élévation des patches.

repère \mathcal{R}_c et la seconde à une translation négative ($t_z < 0$).

En pratique, afin de déterminer la solution adéquate, nous appliquons au démarrage de la tâche de suivi un faible déplacement en boucle ouverte de la sonde dans la direction d'élévation négative. Ce déplacement a pour effet d'imposer initialement la solution $t_z > 0$. Une machine à états (voir [Krupa 2007a] pour plus de détails) permet ensuite de suivre en permanence l'évolution du signe de t_z durant la tâche de suivi de coupe afin de déterminer la solution à retenir. Nous avons expérimentalement validé cette approche pour suivre à l'aide d'une sonde robotisée le déplacement d'un fantôme selon les 3 composantes de translation [Krupa 2007b]. Néanmoins, dans le cas de l'application de mouvements rapides et brutaux au fantôme, la faible bande passante du système robotique ne permet pas de compenser totalement l'erreur de suivi : il en résulte la possibilité d'une perte totale de corrélation du speckle entre la coupe courante et la coupe cible si l'erreur de suivi devient trop importante. Pour éviter cette situation qui met en défaut la méthode d'estimation de distances, nous avons proposé dans [Krupa 2009b] une seconde approche basée sur l'utilisation d'une série de coupes intermédiaires acquises à la volée de manière à garantir la contrainte de corrélation du speckle entre chaque coupe. La distance d'élévation d'un patch de la coupe cible est ensuite obtenue en sommant les distances relatives des patches appartenant aux coupes intermédiaires après avoir déterminé le signe de chaque distance inter-patch à partir de la résolution d'un système d'équations linéaires décrivant les relations d'égalité entre distances [Krupa 2009b]. Les paramètres de l'équation du plan de coupe cible sont ensuite estimés à partir des coordonnées 3D des patches et fournis en entrée de l'asservissement visuel présenté dans la section suivante.

4.1.2 Asservissement visuel 2D-3D pour une tâche de suivi

Nous présentons à présent la commande permettant d'actionner la sonde afin de suivre la coupe cible selon 6 ddl et de stabiliser ainsi l'image. La solution envisagée est de contrôler indépendamment les 3 mouvements dans le plan par un asservissement visuel 2D et les 3 mouvements hors du plan par un asservissement visuel 3D. Ces deux asservissements sont

alors appliqués simultanément.

Asservissement visuel 2D. Les informations visuelles retenues pour contrôler les mouvements dans le plan correspondent naturellement aux deux translations et à la rotation extraites dans l'image par la technique du suivi différentiel présentée précédemment. Elles sont représentées par le vecteur $\mathbf{s}_1 = (t_x, t_y, \gamma)$. Les valeurs désirées à atteindre pour compenser le mouvement dans le plan de la coupe cible correspondent dans ce cas au vecteur nul $\mathbf{s}_1^* = (0, 0, 0)$ et la matrice d'interaction \mathbf{L}_{s_1} qui relie la variation de \mathbf{s}_1 à la vitesse dans le plan de la sonde $\mathbf{v}_1 = (v_x, v_y, \omega_z)$ est simplement la matrice d'identité de taille 3×3 . Les composantes de la vitesse dans le plan à appliquer à la sonde sont alors fournies par la loi de commande

$$\mathbf{v}_1 = -\lambda_1 (\mathbf{s}_1 - \mathbf{s}_1^*), \text{ avec un gain } \lambda_1 > 0. \quad (4.16)$$

Asservissement visuel 3D. Pour stabiliser les mouvements hors du plan, le vecteur d'informations visuelles retenu est $\mathbf{s}_2 = (a, b, c, t_z)$. Les mesures associées sont directement les paramètres estimés de l'équation du plan de la coupe cible correspondant aux composantes de son vecteur normal ainsi que la distance d'élévation t_z de ce plan par rapport au plan courant observé par la sonde. Afin que le plan cible reste orthogonal et superposé au plan courant, les informations visuelles désirées à atteindre correspondent au vecteur $\mathbf{s}_2^* = (0, 0, 1, 0)$. La variation des informations visuelles en fonction de la vitesse de déplacement hors du plan de la sonde, à savoir $\mathbf{v}_2 = (v_z, \omega_x, \omega_y)$, est pour ce choix donnée par la relation

$$\dot{\mathbf{s}}_2 = \mathbf{L}_{s_2} \mathbf{v}_2 = \begin{bmatrix} 0 & 0 & -c \\ 0 & c & 0 \\ 0 & -b & a \\ -1 & 0 & 0 \end{bmatrix} \mathbf{v}_2, \quad (4.17)$$

où les coefficients de la matrice d'interaction \mathbf{L}_{s_2} ne dépendent que de la mesure des composantes du vecteur normal au plan cible. Afin de réguler de manière exponentielle l'erreur $\mathbf{e}_2 = \mathbf{s}_2 - \mathbf{s}_2^*$ vers zéro, il s'agit de déterminer les composantes hors du plan de la vitesse à appliquer à la sonde. La loi de commande appliquée s'exprime par

$$\mathbf{v}_2 = -\lambda_2 (\widehat{\mathbf{L}}_{s_2})^+ \mathbf{e}_2, \text{ avec un gain } \lambda_2 > 0, \quad (4.18)$$

et où $\widehat{\mathbf{L}}_{s_2}$ correspond à la valeur estimée courante de la matrice d'interaction.

Les résultats expérimentaux présentés dans [Krupa 2009b] ont démontré la faisabilité de l'approche utilisant le speckle pour réaliser la poursuite d'une coupe en mouvement selon les 6 degrés de liberté. La figure 4.4 illustre l'étude poursuivie, et démontre la capacité de la méthode à suivre la coupe désirée uniquement à partir de l'information du speckle.

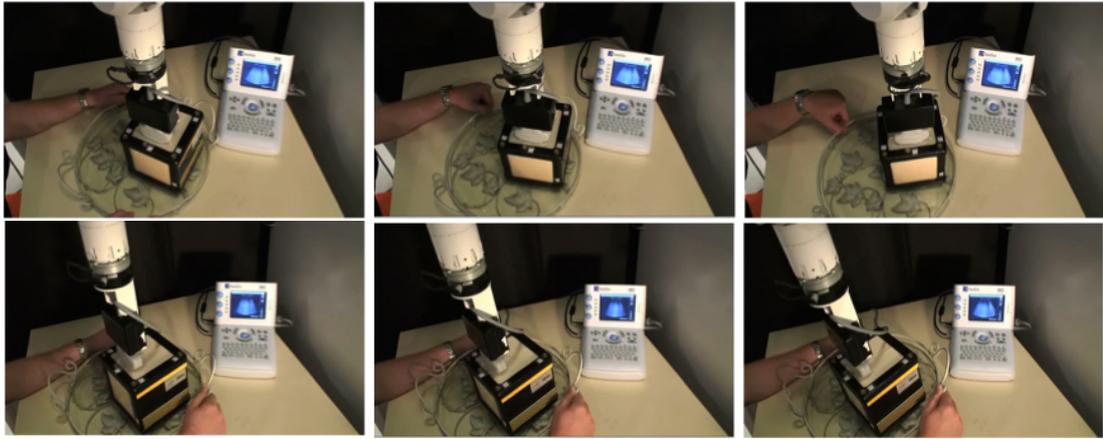


FIGURE 4.4 – Poursuite automatique d'une maquette anatomique, simulant le déplacement de tissus mous (selon 6 DDL), par une sonde échographique 2D embarquée sur un robot. L'asservissement visuel utilise uniquement l'information de speckle contenue dans l'image échographique pour réaliser la tâche.

4.2 Approche utilisant directement l'intensité des pixels

Une solution plus récente utilisant également l'information dense a été proposée dans [Collewet 2011] en vue d'effectuer un asservissement visuel à partir de l'image brute fournie par une caméra. Cette approche nommée « asservissement visuel photométrique » considère en entrée de la commande directement les niveaux de gris de l'image et permet ainsi de s'affranchir du traitement d'image nécessaire à l'extraction d'informations visuelles géométriques. Cependant, comme toutes les techniques d'asservissement visuel 2D fondées sur l'utilisation d'une caméra, elle n'est pas applicable directement au capteur échographique sans une adaptation en profondeur. Nous avons par conséquent dans le cadre des travaux de thèse de Caroline Nadeau [Nadeau 2011a] étendu cette stratégie basée sur l'intensité pour le cas de l'asservissement visuel échographique, en déterminant dans un premier temps la forme analytique de la matrice d'interaction associée à l'intensité de l'image ultrasonore [Nadeau 2011d, Nadeau 2011b].

4.2.1 Modélisation de l'information visuelle et de son interaction

Soit le vecteur d'informations visuelles défini par un ensemble de valeurs d'intensité de pixels tel que

$$\mathbf{s} = (I_{1,1}, \dots, I_{u,v}, \dots, I_{M,N}), \quad (4.19)$$

où $I_{u,v}$ représente l'intensité du pixel de coordonnées (u, v) et M et N correspondent à la largeur et à la hauteur d'une région d'intérêt de l'image échographique. Afin d'obtenir l'expression de la matrice d'interaction associée à \mathbf{s} , il est nécessaire de déterminer dans un premier temps la formulation analytique de la dérivée temporelle de l'intensité d'un pixel $I_{u,v}(t)$ en fonction de la vitesse de la sonde. Pour une image en mode B, l'intensité d'un pixel correspond à l'amplitude $I_{US}({}^o\mathbf{x}_O, t)$ de l'écho ultrasonore réfléchi par un point physique O de coordonnées 3D ${}^o\mathbf{x}_O$ exprimées dans le repère de la scène \mathcal{R}_o . Ainsi,

$$I_{u,v}(t) = I_{US}({}^o\mathbf{x}_O, t). \quad (4.20)$$

Comme l'illustre la figure 4.5, dans l'hypothèse où la sonde se déplace dans un environnement immobile, le point P observé aux mêmes coordonnées (u, v) de l'image à l'instant $t + dt$ représente alors un point physique O' différent de la scène. Les coordonnées 3D de ce nouveau point sont données par ${}^o\mathbf{x}_O + d\mathbf{x}$ où $d\mathbf{x}$ est le déplacement du point P rigidement attaché au repère de la sonde. Ce déplacement est induit par le mouvement de la sonde durant l'intervalle de temps dt . L'intensité du pixel $P(u, v)$ à l'instant $t + dt$ correspond par conséquent à l'amplitude de l'onde ultrasonore réfléchie par le point physique O' , soit

$$I_{u,v}(t + dt) = I_{US}({}^o\mathbf{x}_O + d\mathbf{x}, t + dt). \quad (4.21)$$

En considérant comme seconde hypothèse que la réflexion de l'onde ultrasonore par la scène

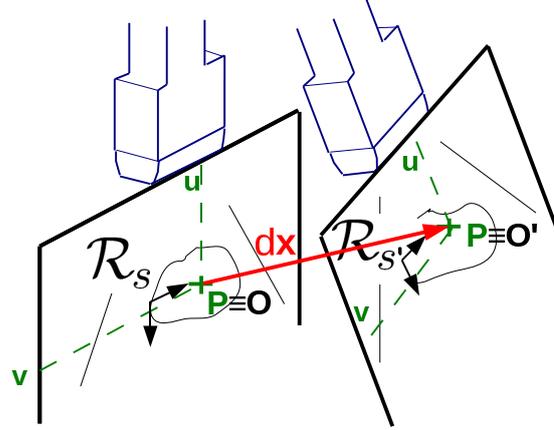


FIGURE 4.5 – Sous l'hypothèse d'un déplacement de la sonde dans un environnement immobile, le point P de coordonnées (u, v) rigidement attaché dans la coupe coïncide à un point 3D O à l'instant t et à un point 3D différent O' à l'instant $t + dt$.

comme constante, l'intensité de l'écho réfléchi par un point physique donné est alors invariante en fonction du temps, et donc

$$I_{US}({}^o\mathbf{x}_O, t) = I_{US}({}^o\mathbf{x}_O, t + dt) = I_{US}({}^o\mathbf{x}_O). \quad (4.22)$$

Cette seconde hypothèse est très forte car l'intensité de l'onde réfléchie est en réalité dépendante de l'incidence de l'onde par rapport à l'interface des structures rencontrées comme nous l'avons indiqué à la section 2.1.2. Néanmoins, cette équation (4.22) de conservation de l'intensité a été a posteriori validée à partir des résultats expérimentaux obtenus sur fantômes et tissus animaux. En soustrayant (4.20) à (4.21), on obtient

$$I_{u,v}(t + dt) - I_{u,v}(t) = I_{US}({}^o\mathbf{x}_O + d\mathbf{x}) - I_{US}({}^o\mathbf{x}_O). \quad (4.23)$$

En développant le premier terme de droite par une décomposition polynomiale de Taylor au premier ordre avec $d\mathbf{x} = (dx, dy, dz)$, telle que

$$I_{US}({}^o\mathbf{x}_O + d\mathbf{x}) \approx I_{US}({}^o\mathbf{x}_O) + \frac{\partial I_{US}({}^o\mathbf{x}_O)}{\partial x}(dx) + \frac{\partial I_{US}({}^o\mathbf{x}_O)}{\partial y}(dy) + \frac{\partial I_{US}({}^o\mathbf{x}_O)}{\partial z}(dz), \quad (4.24)$$

et en le substituant dans (4.23), nous obtenons la relation

$$I_{u,v}(t + dt) - I_{u,v}(t) \approx \frac{\partial I_{US}}{\partial x} dx + \frac{\partial I_{US}}{\partial y} dy + \frac{\partial I_{US}}{\partial z} dz. \quad (4.25)$$

Si maintenant tous les termes de l'équation sont divisés par dt , la variation temporelle de l'intensité d'un pixel $I_{u,v}$, en fonction de la vitesse de déplacement ${}^s\dot{\mathbf{x}}_P = (\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt})$ du point image P par rapport à la scène, s'exprime dans le repère de la sonde \mathcal{F}_s par

$$\dot{I}_{u,v} \approx \nabla I_{u,v} {}^s\dot{\mathbf{x}}_P, \quad (4.26)$$

où $\nabla I_{u,v} = [\nabla I_x \ \nabla I_y \ \nabla I_z]$ correspond au gradient 3D de l'image associé au pixel de coordonnées (u, v) . Il est constitué par les trois composantes $\nabla I_x = \frac{\partial I}{\partial x}$, $\nabla I_y = \frac{\partial I}{\partial y}$ et $\nabla I_z = \frac{\partial I}{\partial z}$ qui décrivent la variation de l'intensité du pixel (u, v) selon les trois axes du repère de l'image.

La vitesse par rapport à la scène du point P rigidement attaché à l'image est induite par la vitesse instantanée de la sonde \mathbf{v} selon la relation cinématique

$${}^s\dot{\mathbf{x}}_P = [\mathbf{I}_3 \quad -[{}^s\mathbf{x}_P]_{\times}] \mathbf{v}. \quad (4.27)$$

Rappelons que les coordonnées 3D du point P exprimées dans le repère de l'image sont données par $x = l_x(u - u_0)$, $y = l_y(v - v_0)$, $z = 0$ où (u_0, v_0) correspondent aux coordonnées en pixels du centre de l'image et l_x, l_y sont les deux facteurs d'échelle de l'image. A partir des équations (4.26) et (4.27), la matrice d'interaction $\mathbf{L}_{I_{u,v}}$ de taille 1×6 associée à l'intensité $I_{u,v}$ du pixel est donnée par

$$\mathbf{L}_{I_{u,v}} = [\nabla I_x \ \nabla I_y \ \nabla I_z \ y\nabla I_z \ -x\nabla I_z \ x\nabla I_y - y\nabla I_x]. \quad (4.28)$$

La matrice d'interaction complète \mathbf{L}_s associée au vecteur \mathbf{s} est finalement construite en empilant les $M \times N$ matrices $\mathbf{L}_{I_{u,v}}$ selon

$$\mathbf{L}_s = \begin{pmatrix} \mathbf{L}_{I_{1,1}} \\ \vdots \\ \mathbf{L}_{I_{M,N}} \end{pmatrix}. \quad (4.29)$$

4.2.2 Détermination du gradient 3D

L'expression (4.28) montre que la matrice d'interaction est dépendante du gradient 3D de l'image courante ainsi que des coordonnées 2D dans l'image des pixels constituant la région d'intérêt. Les positions de ces derniers sont évidemment connues car elles sont fixées à l'avance par la définition de la région d'intérêt dans l'image. Les composantes ∇I_x et ∇I_y du gradient peuvent donc être aisément calculées à partir de l'image courante à l'aide d'un filtrage de Sobel 2D. En revanche, la problématique de la détermination de la composante ∇I_z du gradient 3D selon l'axe orthogonal à l'image s'avère plus ardue. Une connaissance de ∇I_z est en effet indispensable pour contrôler les déplacements hors du plan de la sonde et permettre ainsi la mise en œuvre d'un asservissement visuel gérant les 6 ddl par l'application de la loi de commande (2.17). Nous avons par conséquent proposé différentes solutions présentées brièvement ici pour obtenir une mesure complète du gradient 3D.

Filtre image 3D. En traitement d'images, le gradient d'une image 2D est couramment calculé par l'application de deux filtres directionnels tels les filtres de Sobel qui permettent de calculer séparément les dérivées verticales et horizontales de l'image. Ces filtres sont construits à partir de deux opérateurs séparables qui permettent respectivement de lisser l'image à l'aide d'un opérateur gaussien perpendiculaire à la direction de dérivation et de calculer une différence centrale selon la direction de dérivation [Hadwiger 2006]. Si une image 3D construite à partir d'un ensemble de coupes parallèles est disponible, les composantes de son gradient 3D peuvent être directement obtenues en appliquant trois filtres 3D définis sur le même modèle que les filtres de Sobel.

Afin d'utiliser ces filtres 3D, la solution proposée dans [Nadeau 2011d] est d'appliquer, lors d'une phase d'initialisation, un faible mouvement de va-et-vient à une sonde 2D pour acquérir 4 coupes additionnelles parallèles à l'image courante et construire ainsi une image 3D permettant de calculer les trois composantes du gradient de la coupe centrale. Pour éviter de superposer continuellement ce mouvement de va-et-vient lors de l'asservissement visuel, nous avons considéré dans la loi de commande (2.17) une estimation de la matrice d'interaction à la position désirée. Cette stratégie de maintenir $\widehat{\mathbf{L}}_s = \widehat{\mathbf{L}}_{s|s=s^*}$ constante est cependant uniquement appropriée à des applications robotiques de suivi de coupe. Les résultats expérimentaux ont en effet montré qu'il est nécessaire d'effectuer une mise à jour en ligne de la matrice d'interaction pour des applications de positionnement où l'erreur de position initiale entre la coupe courante et la coupe cible est importante. Nous avons par conséquent proposé dans [Nadeau 2011a] d'utiliser une sonde 3D fournissant directement les coupes additionnelles nécessaires au calcul en ligne du gradient 3D.

Estimation en ligne du gradient. Une approche alternative a également été présentée dans [Nadeau 2011c] pour estimer le gradient 3D de l'image courante sans avoir recours à l'utilisation d'une sonde 3D ou l'acquisition d'images parallèles supplémentaires. Le principe de la méthode consiste à calculer à partir de l'image courante les 2 composantes du gradient ∇I_x et ∇I_y à l'aide des filtres dérivatifs de Sobel, puis d'estimer la composante ∇I_z à partir de l'observation de la variation d'intensité des pixels et de la mesure du déplacement de la sonde effectué lors de l'asservissement visuel. Cette estimation en ligne du gradient permet ainsi de mettre à jour la matrice d'interaction $\widehat{\mathbf{L}}_s$ considérée dans la loi de commande (2.17) en vue de réaliser une tâche de positionnement de la sonde. La section suivante s'intéresse particulièrement à cette application.

4.2.3 Application au positionnement de sondes 2D, 3D et bi-plans

L'approche utilisant l'intensité des pixels a été validée expérimentalement sur un fantôme abdominal pour réaliser des tâches de positionnement de la sonde. La figure 4.6 montre la plateforme expérimentale utilisée. Les valeurs des informations visuelles désirées à atteindre s^* sont fixées par apprentissage de l'image à la position désirée de la sonde. L'asservissement visuel est lancé à partir d'une position initiale différente en appliquant la loi de commande (2.17) où une estimation de la matrice d'interaction courante est considérée. La figure 4.7 présente les résultats d'un positionnement automatique obtenus avec une sonde 2D convexe et l'utilisation de l'algorithme d'estimation en ligne du gradient 3D proposé dans [Nadeau 2011c]

pour mettre à jour la matrice d'interaction. A la convergence de l'asservissement, l'erreur de pose nulle ainsi que la couleur grise uniforme de l'image de différence, qui correspond à la différence entre l'image courante et l'image désirée à atteindre, démontrent le succès de la tâche. Notons également qu'à la convergence, la norme euclidienne de l'erreur visuelle, définie par $\mathcal{C} = (\mathbf{s} - \mathbf{s}^*)^\top (\mathbf{s} - \mathbf{s}^*)$, converge vers une valeur faible qui représente uniquement le bruit du capteur.



FIGURE 4.6 – Plateforme expérimentale mise en œuvre pour les expérimentations de positionnement. Un fantôme abdominal (ABDFAN, Kyoto Kagaku) dédié à la formation des personnels médicaux est utilisé pour simuler de manière réaliste les organes digestifs et un robot 6 axes (Viper S850, Adept) est contrôlé pour actionner différentes sondes échographiques dont entre autres une sonde 2D convexe (C60, Sono-site), une sonde 3D motorisée (4DC7-3/40, Ultrasonix) et une sonde bi-plan (XYPA2.5, Ultrasonix).

Les résultats d'un positionnement d'une sonde 3D à balayage motorisée sont présentés à la figure 4.8. Dans ce cas, le gradient 3D est directement calculé à partir d'un volume constitué de 3 coupes et centré sur l'image 2D courante utilisée en entrée de l'asservissement visuel. Ici également, la tâche est réalisée avec succès et l'image de différence, obtenue à la convergence, contient uniquement le bruit du capteur 3D. Toutefois, la fréquence d'acquisition est limitée à 12 volumes/s pour un pas de moteur de 1,4 deg par coupe. Il en résulte une bande passante de l'asservissement visuel divisée par deux en comparaison à celle obtenue par l'emploi d'une sonde 2D dont la cadence d'acquisition est de 25 images/s. Bien entendu, l'utilisation d'une sonde 3D à technologie matricielle permettrait de supprimer cette limitation.

L'approche a également été étendue dans [Nadeau 2011c] pour considérer le cas d'une sonde bi-plans où l'information visuelle est extraite de deux plans orthogonaux. Le vecteur des informations visuelles est dans ce cas défini par

$$\mathbf{s} = (I_{1,1}, \dots, I_{u,v}, \dots, I_{M,N}, I'_{1,1}, \dots, I'_{u,v}, \dots, I'_{M,N}), \quad (4.30)$$

où $I'_{u,v}$ correspond à l'intensité du pixel de coordonnées (u, v) appartenant à la région d'intérêt

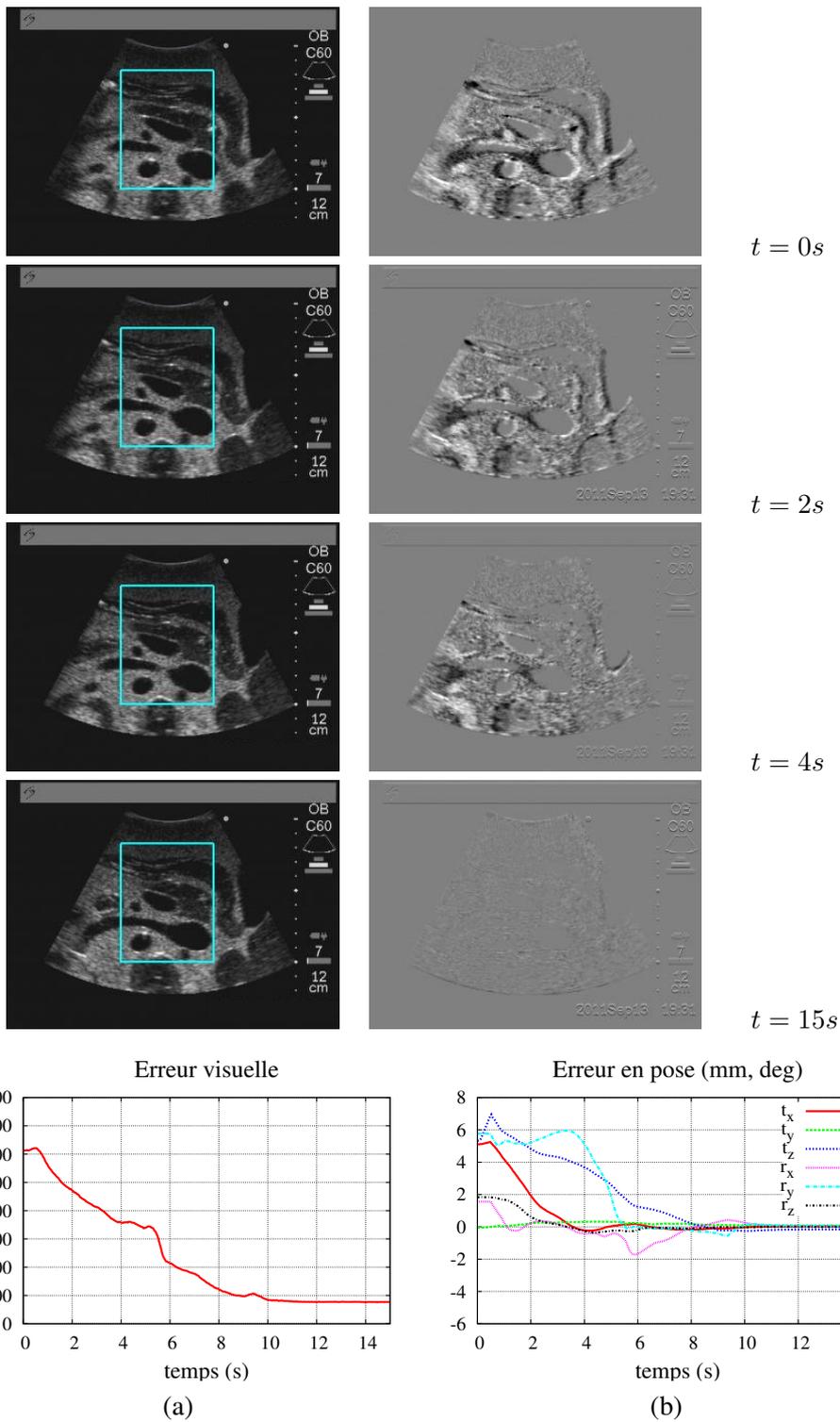


FIGURE 4.7 – Tâche de positionnement avec une sonde 2D. La colonne de gauche expose quelques images prises à des instants différents lors de l'asservissement visuel (avec la région d'intérêt délimitée en cyan). La colonne de droite présente la différence entre ces images et l'image désirée à atteindre. La couleur grise uniforme obtenue à la convergence démontre le succès de la tâche. (a) Evolution de la norme euclidienne de l'erreur visuelle $C = (s - s^*)^T (s - s^*)$. (b) Evolution de l'erreur en termes de pose.

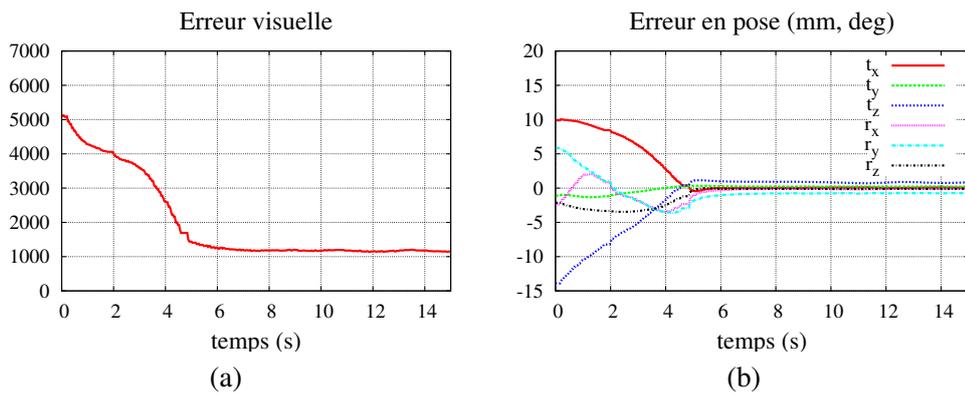
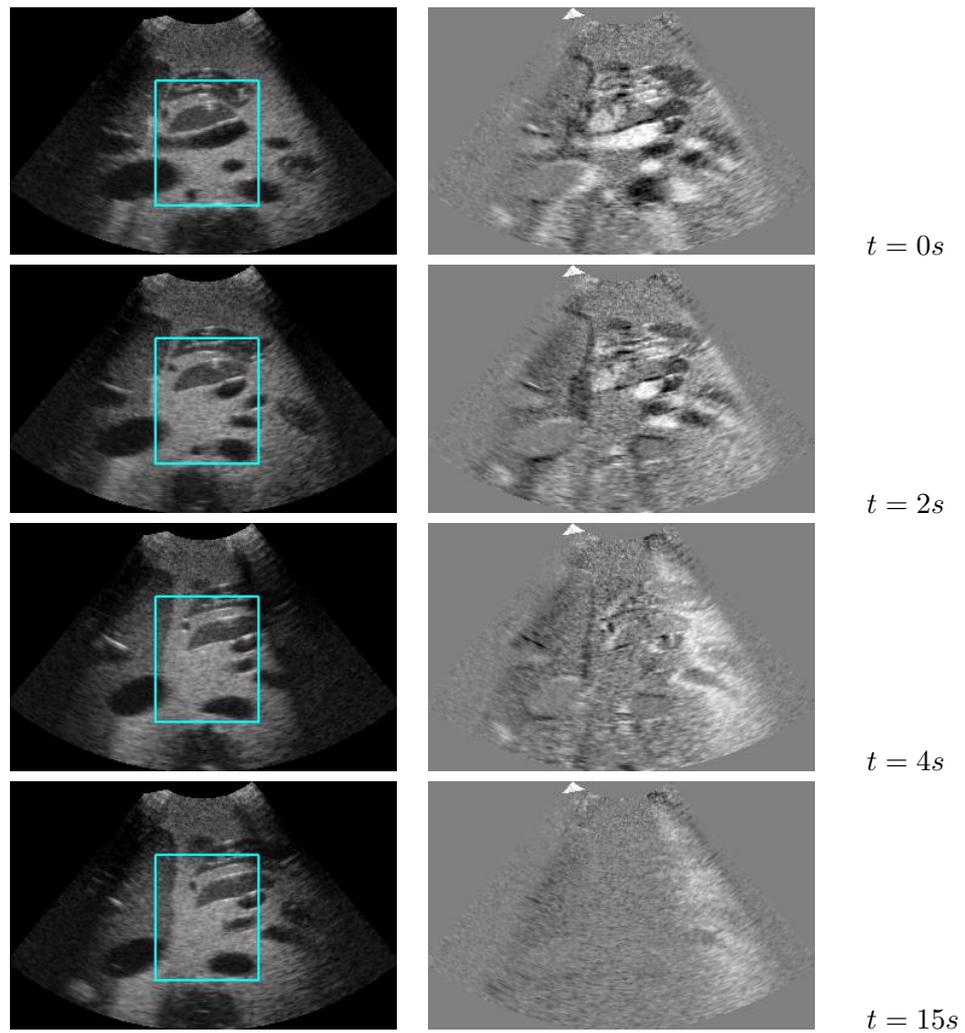


FIGURE 4.8 – Tâche de positionnement avec une sonde 3D. La colonne de gauche expose quelques images prises à des instants différents lors de l’asservissement visuel (avec la région d’intérêt délimitée en cyan). La colonne de droite présente la différence entre ces images et l’image désirée à atteindre. A la convergence l’image de différence contient uniquement le bruit du capteur 3D et la tâche est réalisée avec succès. (a) Evolution de la norme euclidienne de l’erreur visuelle $\mathcal{C} = (s - s^*)^T (s - s^*)$. (b) Evolution de l’erreur en termes de pose.

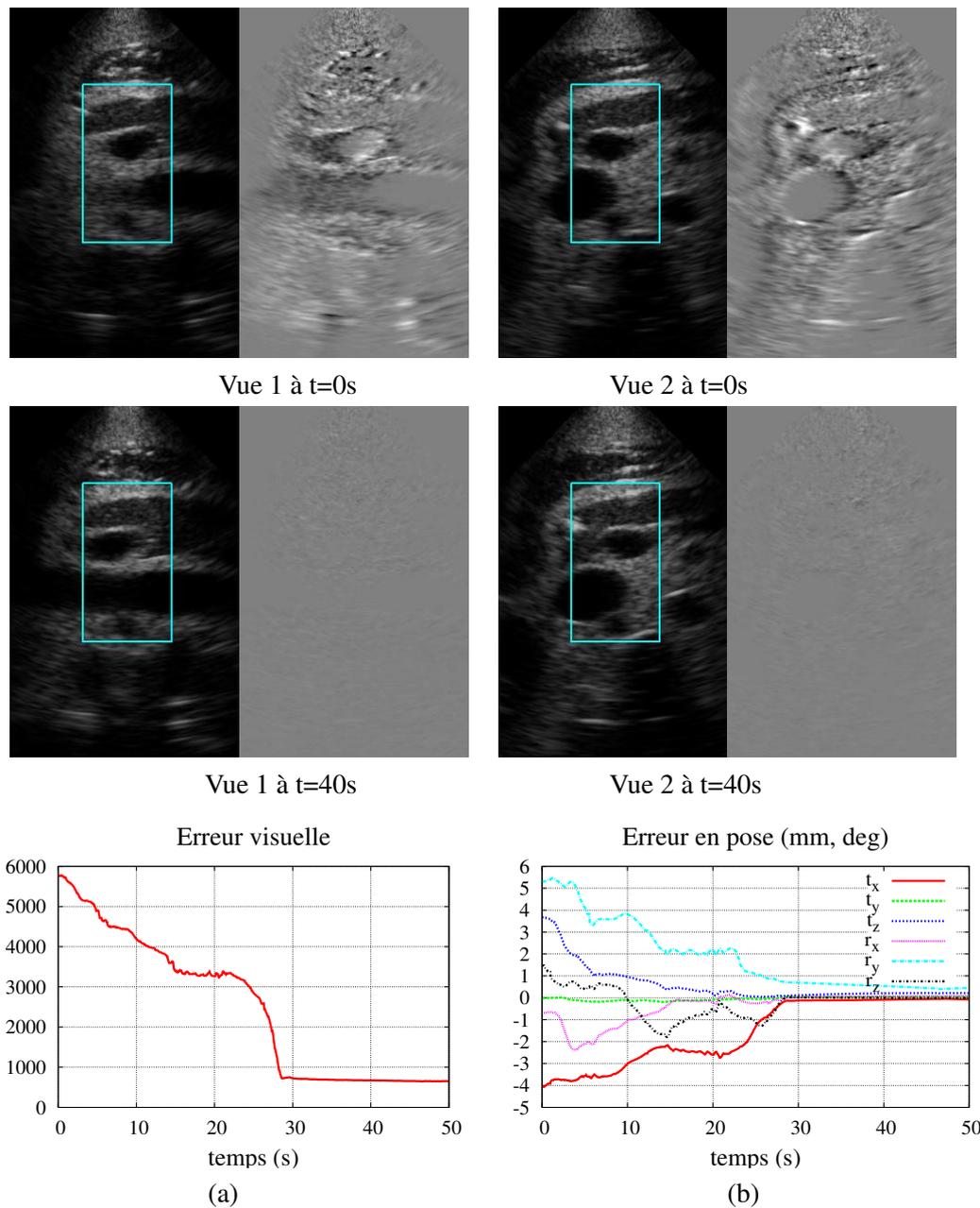


FIGURE 4.9 – Tâche de positionnement avec une sonde bi-plans. La ligne du haut présente les deux vues orthogonales (avec les régions d'intérêt délimitées en cyan) ainsi que les images de différence avec les vues désirées observées au démarrage de l'asservissement visuel. La couleur grise uniforme, des images de différence, obtenue à la convergence démontre le succès de la tâche. (a) Evolution de la norme euclidienne de l'erreur visuelle $C = (s - s^*)^T (s - s^*)$. (b) Evolution de l'erreur en termes de pose.

de la coupe orthogonale additionnelle. La matrice d'interaction complète correspond alors à

$$\mathbf{L}_s = \begin{bmatrix} \mathbf{L}_{I_{1,1}} \\ \vdots \\ \mathbf{L}_{I_{M,N}} \\ \mathbf{L}'_{I_{1,1}} \\ \vdots \\ \mathbf{L}'_{I_{M,N}} \end{bmatrix}, \quad (4.31)$$

où $\mathbf{L}_{I_{u,v}}$ est donnée par l'expression (4.28) et la matrice d'interaction $\mathbf{L}'_{I_{u,v}}$ associée à l'intensité des pixels supplémentaires de la coupe additionnelle est dérivée de $\mathbf{L}_{I_{u,v}}$ suite à une transformation de repère (voir [Nadeau 2011a] pour le détails des calculs). Ainsi,

$$\mathbf{L}'_{I_{u,v}} = [-\nabla I'_z \quad \nabla I'_y \quad \nabla I'_x \quad y' \nabla I'_x - x' \nabla I'_y \quad -x' \nabla I'_z \quad y' \nabla I'_z]. \quad (4.32)$$

Dans cette expression, $\nabla I'_{u,v} = [\nabla I'_x \quad \nabla I'_y \quad \nabla I'_z]$ correspond au gradient 3D de l'image additionnelle associé au pixel de coordonnées (u, v) , et (x', y') sont les coordonnées métriques de ce pixel dans le repère de l'image additionnelle. Les résultats d'une tâche de positionnement avec une sonde bi-plans sont présentés à la figure 4.9. Nous pouvons observer la convergence des deux images vers les images désirées et noter également le succès de l'asservissement en termes de pose. Le principal avantage d'utiliser une telle sonde est d'obtenir un domaine de convergence plus grand tout en garantissant une cadence de 25 Hz de la commande.

Néanmoins, quel que soit le type de sonde considéré, le domaine de convergence de l'approche utilisant l'intensité reste faible comparé à celui d'une approche géométrique telle celle utilisant les moments de la section d'un organe observé. Cette approche est donc à privilégier pour des applications de positionnement de faible amplitude ou pour affiner par la vision un positionnement peu précis préalablement réalisé à partir des mesures proprioceptives du robot. En contrepartie, cet asservissement visuel direct a l'avantage de s'appliquer à des images échographiques très faiblement structurées ne présentant pas de contour perceptible. Il s'avère également être très efficace pour des applications de suivi de coupe, comme nous le verrons dans la section suivante.

4.2.4 Application à la stabilisation de l'image échographique

Cette étude a été menée dans le cadre du projet ANR USComp que je dirige depuis fin 2008 et qui porte sur la compensation du mouvement physiologique du patient par une sonde échographique robotisée. L'objectif applicatif visé est de réaliser une compensation automatique du mouvement physiologique du patient en stabilisant l'image échographique par un système robotique manipulant la sonde. A plus long terme, la méthode proposée pourra être utilisée pour augmenter les performances des systèmes robotiques d'aide aux gestes chirurgicaux effectués sous imagerie ultrasonore interventionnelle. Nous pouvons par exemple signaler les applications robotiques d'assistance aux gestes de ponction (biopsie) ou d'ablation localisée de tumeur (cryo ou thermo-ablation) où les mouvements physiologiques induits par la respiration du patient sont fortement présents.

La loi de commande classique que nous avons appliquée jusqu'à présent pour effectuer des tâches de positionnement de la sonde n'intègre pas de connaissance a priori sur la nature du mouvement physiologique à compenser. En pratique, ce mouvement est fortement corrélé avec le cycle respiratoire du patient. Afin de tirer avantage du caractère répétitif du mouvement à compenser, dans [Nadeau 2011e], nous avons proposé d'intégrer un correcteur prédictif dans la boucle de commande exploitant l'intensité des pixels. Ce correcteur est basé sur le principe du GPC (Generalized Predictive Controller) décrit dans [Clarke 1987] et permet d'anticiper l'effet de la perturbation. Nous avons plus particulièrement mis en œuvre une variante de ce correcteur, proposée dans [Gangloff 2006], afin de considérer un modèle de bruit répétitif. Cette technique de contrôle consiste à décrire le système à commander par un modèle ARIMAX suivant

$$\mathbf{A}(z^{-1})\mathbf{y}(k) = \mathbf{B}(z^{-1})\mathbf{u}(k-1) + \frac{\mathbf{P}(z^{-1})}{\Delta(z^{-1})}\mathbf{b}(k) \quad (4.33)$$

où les vecteurs \mathbf{u} et \mathbf{y} correspondent respectivement aux entrées et sorties du système et les polynômes \mathbf{A} et \mathbf{B} représentent le modèle dynamique théorique du système robotique qui est dans notre cas approché par un intégrateur pur. Le vecteur d'entrée \mathbf{u} correspond donc au torseur cinématique de la sonde et la sortie \mathbf{y} représente les informations visuelles observées dans l'image courante. Afin d'éviter un nombre excessif de composantes visuelles et d'assurer ainsi la contrôlabilité du système, nous définissons un vecteur de sortie de dimension 6 par l'introduction d'une matrice de combinaison, soit $\mathbf{y} = \mathbf{C}\mathbf{s} = \widehat{\mathbf{L}}_s^+ \mathbf{s}$. Le bruit du système est modélisé par un terme $\mathbf{b}(k)$ caractérisant un bruit blanc qui est coloré par la matrice polynomiale \mathbf{P} et rendu non stationnaire par l'opérateur $\Delta(z^{-1})$. Pour le cas d'un GPC classique, cet opérateur est défini par un simple intégrateur : $\Delta(z^{-1}) = 1 - z^{-1}$. Par contre, si la perturbation est périodique avec une période constante et connue, la solution proposée dans [Gangloff 2006] est de redéfinir cet opérateur afin de considérer un modèle répétitif de la perturbation par $\Delta_R(z^{-1}) = 1 - \alpha z^{-T}$, avec $0 < \alpha \leq 1$. Dans cette nouvelle définition, α est un facteur d'oubli qui impose au modèle de bruit d'être répétitif avec une période T correspondant à la période du signal de perturbation mesurée en pas d'échantillonnage.

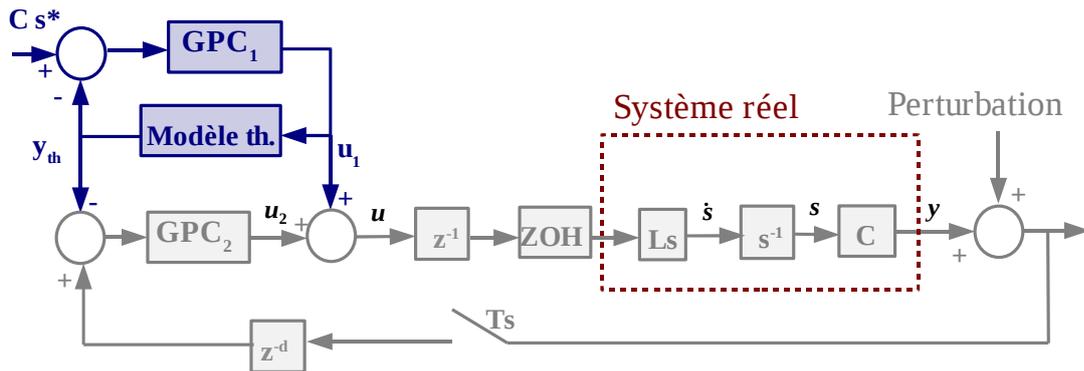


FIGURE 4.10 – Schéma bloc du contrôleur R-GPC qui présente un retard de commande d'une période d'échantillonnage et un retard de d périodes d'échantillonnage dû à la mesure.

Pour que ce modèle de bruit répétitif affecte uniquement le rejet de la perturbation et non la réponse à la consigne visuelle, ce nouveau correcteur appelé R-GPC (Repetitive Generalized

Predictive Controller) est construit sur une combinaison de deux GPC selon le schéma de commande illustré par la figure 4.10. Un premier correcteur GPC_1 est donc appliqué au modèle théorique du système, sans terme de bruit, pour anticiper les consignes futures $\mathbf{r} = \mathbf{C}\mathbf{s}^*$. Un second correcteur GPC_2 , introduisant le modèle de bruit répétitif, permet d'annuler l'effet de la perturbation sur le vecteur de sortie. La fonction de coût minimisée par l'approche R-GPC correspond ainsi à la combinaison des fonctions de coût des deux correcteurs GPC_1 et GPC_2 [Gangloff 2006] selon

$$\begin{aligned}
 J(\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2, k) = & \sum_{j=N_1}^{N_2} \|\hat{\mathbf{y}}_{th}(k+j) - \mathbf{r}(k+j)\|^2 + \sum_{j=N_1}^{N_2} \|\hat{\varepsilon}(k+j)\|^2 \\
 & + \mu_1 \sum_{j=1}^{N_u} \|\delta\mathbf{u}_1(k+j-1)\|^2 + \mu_2 \sum_{j=1}^{N_u} \|\delta\mathbf{u}_2(k+j-1)\|^2
 \end{aligned}
 \tag{4.34}$$

où N_1, N_2 sont les limites de l'horizon de prédiction, N_u est la taille de l'horizon de commande, $\hat{\mathbf{y}}_{th}(t)$ est la sortie prédite par le modèle sans bruit, $\mathbf{r}(t)$ est le signal de consigne à suivre et $\hat{\varepsilon}(t)$ est la prédiction de l'erreur due au bruit et à la perturbation. Les poids μ_1 et μ_2 permettent de pondérer respectivement les variations $\delta\mathbf{u}_1(t)$ et $\delta\mathbf{u}_2(t)$ des signaux de commande, représentant les efforts mis en jeu pour le suivi de la consigne $\mathbf{r}(t)$ et le rejet de l'erreur $\varepsilon(t)$. Le signal de commande $\mathbf{u}(t)$ appliqué au robot est ainsi composé par deux signaux $\mathbf{u}_1(t)$ et $\mathbf{u}_2(t)$ correspondant respectivement à l'entrée du modèle du système théorique sans perturbation et à la composante de la commande permettant de rejeter la perturbation et le bruit. Nous invitons le lecteur à se référer à l'article [Gangloff 2006] pour les détails concernant la minimisation de cette fonction de coût.

Le correcteur R-GPC a été utilisé afin de réaliser le suivi automatique d'une coupe anatomique en mouvement en fixant à l'aide de la matrice de combinaison la consigne par $\mathbf{r} = \mathbf{C}\mathbf{s}^* = \widehat{\mathbf{L}}_s^+ \mathbf{s}^*$, où \mathbf{s}^* correspond aux intensités des pixels de la coupe initiale observée par la sonde.



FIGURE 4.11 – Dispositif expérimental utilisé pour valider la tâche de compensation du mouvement sur un fantôme abdominal. Le mouvement de perturbation est réalisé en appliquant une rotation contrôlée à un plateau tournant sur lequel est posé le fantôme.

La figure 4.12 présente des résultats expérimentaux obtenus lors de l'application d'un mou-

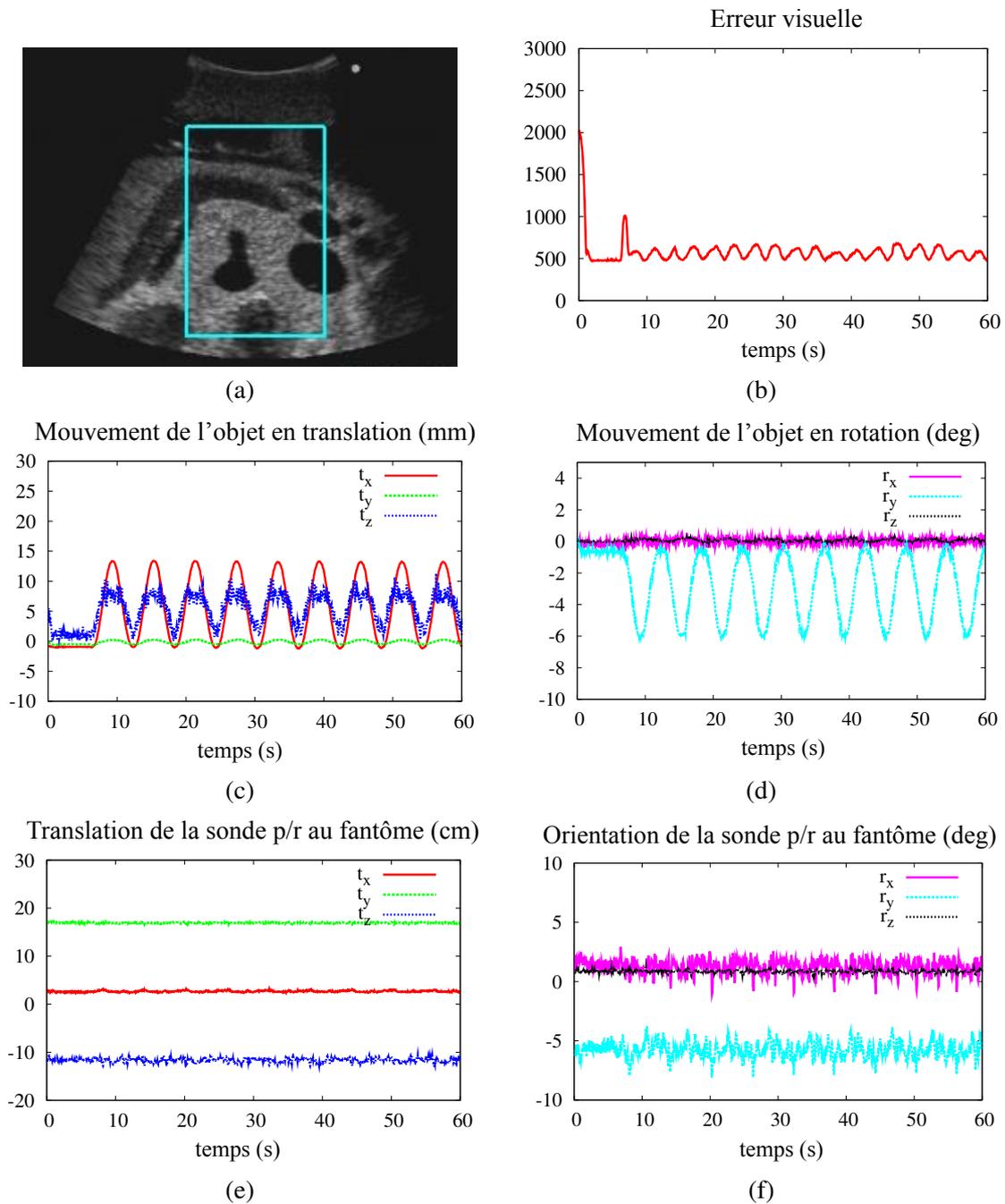


FIGURE 4.12 – (a) Section abdominale à suivre. (c)-(d) mouvement périodique appliqué au fantôme. (b) Evolution de l'erreur visuelle lors de la compensation du mouvement par asservissement visuel utilisant la commande prédictive (R-GPC). (e)-(f) Position et orientation de la sonde par rapport au fantôme lors du suivi.

vement sinusoïdal au fantôme abdominal avec une période de 6 secondes simulant le mouvement respiratoire d'un patient. Ce mouvement de perturbation a été généré à l'aide d'un dispositif (voir figure 4.11) utilisant un second manipulateur robotique pour déplacer le fantôme selon une rotation autour de l'axe Y de la sonde et deux translations le long des axes X et Z permettant ainsi d'appliquer des mouvements au fantôme dans le plan et hors du plan de coupe. La commande prédictive (R-GPC) assure une bonne précision du suivi qui se traduit par la valeur faible de l'erreur visuelle et la valeur globalement constante de la position et de l'orientation de la sonde par rapport au fantôme.

4.2.5 Application au suivi d'un outil

L'approche utilisant directement l'intensité ultrasonore a également été adaptée pour suivre le déplacement d'un outil chirurgical dans une séquence de volumes échographiques fournie par une sonde 3D. A cette fin, nous avons étendu la méthode en définissant le vecteur d'informations visuelles par les intensités des voxels constituant une région d'intérêt 3D. Ainsi, on a

$$\mathbf{s} = (I_{1,1,1}, \dots, I_{u,v,w}, \dots, I_{M,N,L}), \quad (4.35)$$

où M , N and L désignent respectivement la largeur, la hauteur et l'épaisseur de la région d'intérêt 3D et $I_{u,v,w}$ représente l'intensité du voxel de coordonnées (u, v, w) dans le volume. La forme analytique de matrice d'interaction $\mathbf{L}_{I_{u,v,w}}$ associée à l'intensité d'un voxel s'obtient simplement en considérant la coordonnée $z = l_z(w - w_0)$ du point P non nulle dans la modélisation présentée à la section 4.2.1 avec w_0 correspondant à la coordonné en Z (en voxel) du centre du volume et l_z étant le facteur d'échelle associée. La matrice d'interaction associée à $I_{u,v,w}$ s'écrit ainsi

$$\begin{aligned} \mathbf{L}_{I_{u,v,w}} &= \nabla I_{u,v,w} \begin{bmatrix} \mathbf{I}_3 & -[{}^s \mathbf{x}_P]_{\times} \end{bmatrix} \\ &= \begin{bmatrix} \nabla I_x & \nabla I_y & \nabla I_z & y \nabla I_z - z \nabla I_y & z \nabla I_x - x \nabla I_z & x \nabla I_y - y \nabla I_x \end{bmatrix}, \end{aligned} \quad (4.36)$$

où $\nabla I_{u,v,w}$ est le gradient 3D du voxel de coordonnées (u, v, w) obtenu en appliquant les filtres de Sobel 3D sur le volume. La matrice d'interaction \mathbf{L}_s est ensuite obtenue en empilant les $M \times N \times L$ matrices $\mathbf{L}_{I_{u,v,w}}$.

Basé sur ce choix d'informations visuelles, un asservissement visuel virtuel, dont le principe est similaire à celui présenté à la section 3.4.2 pour une application de recalage, a été mis en œuvre dans [Nadeau 2012] pour estimer dans une séquence de volumes échographiques les déplacements d'un outil chirurgical inséré dans le ventricule droit d'un cœur battant. Au démarrage, l'utilisateur définit dans le premier volume de la séquence une petite région d'intérêt 3D enfermant le bout de l'outil. Cette région d'intérêt constitue alors l'image 3D de référence à suivre dans la séquence de volumes par une sonde virtuelle 3D interagissant avec le volume complet réel et fournissant sa propre image 3D observée par technique d'interpolation. Le principe de la méthode consiste ensuite à appliquer, pour chaque volume de la séquence, un asservissement visuel à la sonde virtuelle afin que l'image 3D observée par celle-ci corresponde à l'image 3D de référence. Une estimation de la pose de l'outil est ainsi disponible en mesurant directement la pose de la sonde virtuelle à la convergence visuelle et l'asservissement est ensuite réitéré dans le volume suivant de la séquence pour suivre l'évolution de la cible.

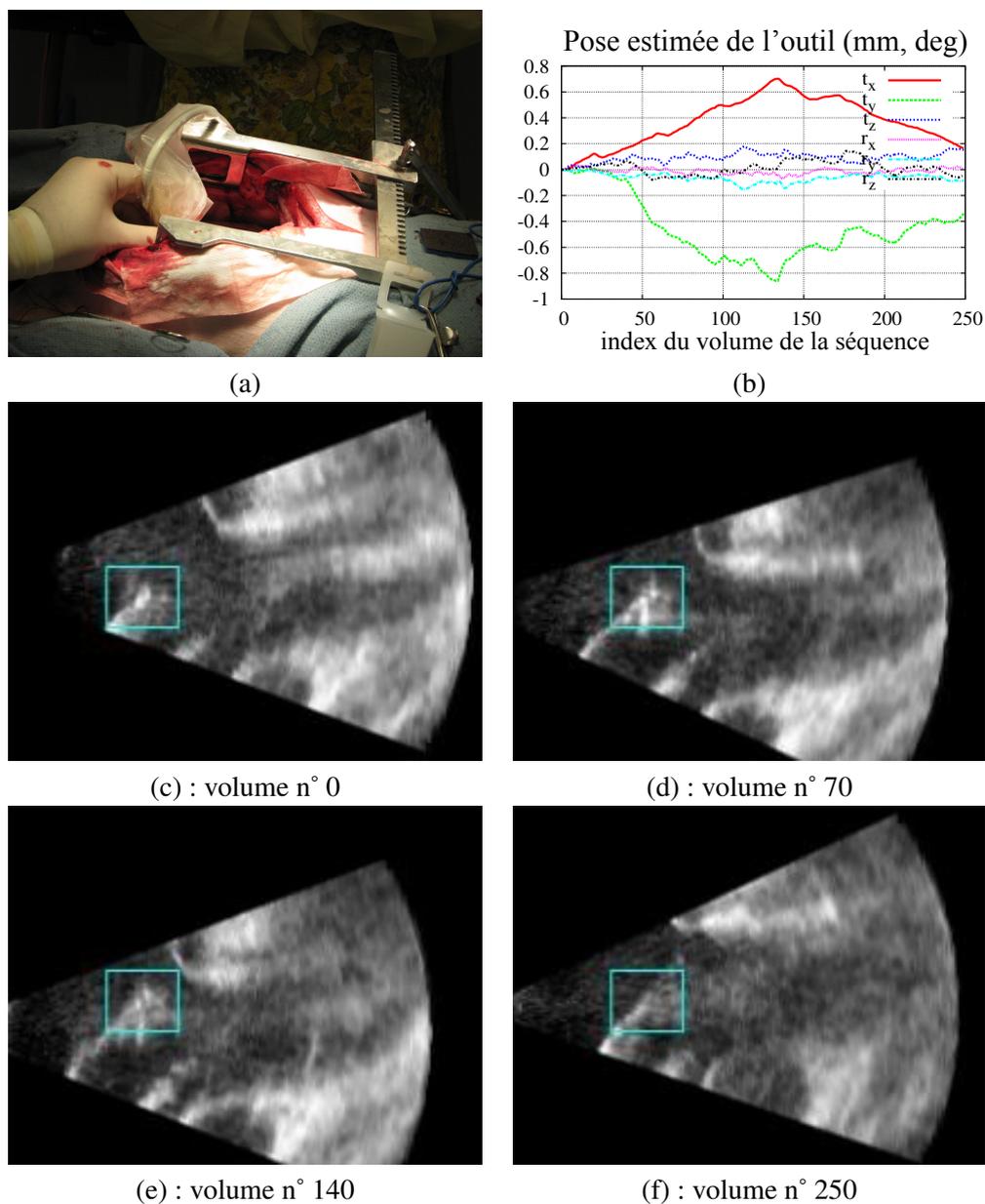


FIGURE 4.13 – Résultat du suivi d'un outil chirurgical par asservissement visuel virtuel. (a) Vue du champ opératoire lors de l'expérimentation réalisée sur animal porcin où une sonde 3D est maintenue sur la surface du cœur pour observer l'évolution de l'outil dans le ventricule droit lors de l'acquisition de 250 volumes pendant 10 secondes. (b) Evolution de la pose estimée de l'outil en fonction de l'index du volume de la séquence. (c)-(f) Coupes centrales extraites de l'image 3D observée par la sonde virtuelle à la convergence de l'asservissement visuel pour différents volumes de la séquence. Le centrage permanent de la région d'intérêt sur le bout de l'outil démontre le succès du suivi 3D.

La figure 4.13 présente les résultats de cette tâche de suivi d'outil obtenus en condition in vivo lors d'une expérimentation réalisée en collaboration avec l'équipe du Professeur Pierre Dupont du département de chirurgie cardiaque du Boston Children's Hospital. Un outil microscopique a été inséré initialement dans le ventricule droit du cœur battant d'un animal porcin sédaté à l'aide du robot à tubes concentriques présenté dans [P. Dupont 2010]. Ce robot permet d'accéder au cœur via la veine jugulaire à partir d'une petite incision effectuée au cou de l'animal. Lors de cette expérimentation, la séquence de volumes a été acquise à une cadence de 25 vol/s par une sonde 3D matricielle positionnée en contact avec la surface du cœur. Les résultats obtenus démontrent l'efficacité de la méthode pour réaliser un suivi 3D robuste de l'extrémité de l'outil. En effet, la région 3D cible contenant l'extrémité de l'outil est continuellement suivie lors de la séquence de volumes.

4.3 Approche 3D avec prise en compte des déformations des tissus

Les différentes approches présentées précédemment permettent de compenser le déplacement rigide d'une cible anatomique. Cependant, en conditions in vivo réelles les tissus mous sont également assujettis à des déformations internes induites en grande partie par le cycle respiratoire du patient. Bien entendu, ces déformations ne peuvent pas être compensées par le déplacement automatique de la sonde à moins de déformer la géométrie de l'empreinte de cette dernière au cours du suivi, ce qui est évidemment impossible. En revanche, cette compensation des déformations peut être effectuée numériquement sur l'image acquise en appliquant la transformation inverse des déformations internes si une mesure de ces dernières est disponible. Nous avons par conséquent proposé dans [Lee 2011] une méthode permettant d'extraire la déformation des tissus observés dans une séquence de volumes denses fournis par une sonde échographique 3D.

L'idée est de calculer, en temps réel, les mouvements rigides et non rigides d'une région d'intérêt qui a été définie par l'utilisateur dans un volume initial. Le principe de la méthode consiste à estimer, à partir des différences d'intensité des voxels mesurées entre volumes successifs, les déplacements d'un ensemble de points de contrôle 3D décrivant l'évolution de courbes 3D de type « Thin-Plate Splines » (TPS) [Bookstein 1989] qui modélisent les déformations de la région d'intérêt par une grille 3D déformable. Une approche basée sur les TPS avait déjà été proposée dans [Richa 2009] pour estimer les déformations de la surface d'un cœur battant à partir de l'information dense contenue dans deux images 2D fournies par un système de vision endoscopique stéréoscopique. Néanmoins, cette solution est dédiée à l'imagerie optique stéréoscopique.

Nos travaux ont donc porté sur le développement d'une méthode adaptée à l'imagerie échographique qui utilise, en entrée du système, une séquence de volumes fournie par une sonde échographique 3D. Cette méthode a en outre été parallélisée et implantée sur GPU (Graphics Processing Unit), afin de garantir une exécution temps réel. A partir de l'estimation du mouvement complet du volume contenu dans la grille, le mouvement rigide de la région d'intérêt est extrait pour ensuite être compensé automatiquement au moyen d'un asservissement visuel 3D. L'objectif de cet asservissement est de déplacer le capteur échographique de manière à minimiser le mouvement rigide estimé. La figure 4.14 présente quelques résultats de cette approche, qui ont été obtenus sur un fantôme mou auquel est appliqué une déformation périodique par

un dispositif mécanique. Nous pouvons observer que la grille 3D déformable, dont la position a été initialisée au milieu du premier volume échographique de la séquence, s'adapte aux déformations de la région d'intérêt lors des phases successives de pression et relaxation du fantôme. Tout au long de la séquence, la grille reste également centrée dans le volume courant sous l'effet de la compensation du mouvement rigide par asservissement visuel. Ces résultats démontrent ainsi la faisabilité d'une telle approche pour compenser le mouvement rigide de tissus mous présentant de fortes déformations internes.

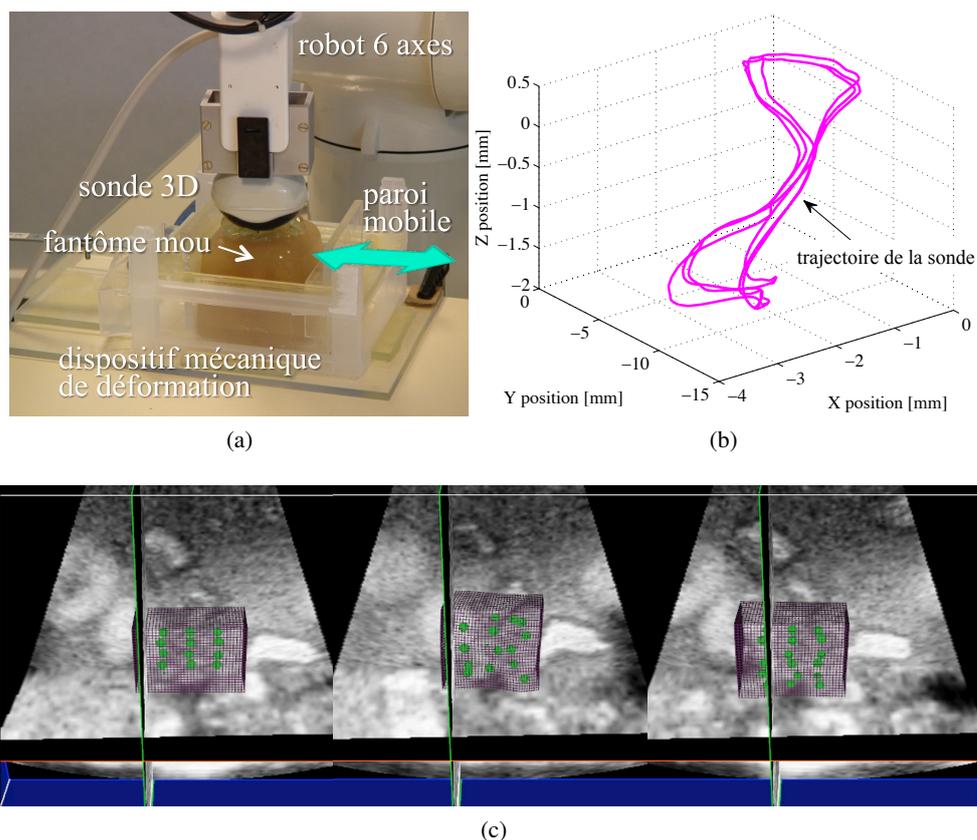


FIGURE 4.14 – Résultat de la méthode d'extraction des déformations dans une séquence de volumes fournie par une sonde à balayage motorisé. (a) Dispositif expérimental permettant d'appliquer des déformations périodiques à un fantôme mou par écrasement contre une paroi mobile. (c) Différents volumes observés à des phases successives de pression et relaxation du fantôme avec visualisation des déformations estimées de la grille contenant la région d'intérêt à suivre. (b) Trajectoire réalisée par la sonde durant l'asservissement visuel pour compenser automatiquement le mouvement rigide de la région d'intérêt.

Pour la mise en œuvre pratique, nous avons également développé un algorithme qui reconstruit, en temps réel, les volumes échographiques denses à partir d'un flux de données 2D qui est transmis à la volée par une sonde à balayage motorisé. Les données utilisées en entrée ne sont pas des images 2D formées (post-scan) mais directement les vecteurs de mesures brutes d'intensité (pre-scan) correspondant à chaque ligne de tir du capteur échographique. La reconstruction du volume échographique dense est alors réalisée en positionnant dans l'espace 3D les

données des vecteurs de mesures selon les différentes lignes de tir. La position et l'orientation de ces dernières sont définies par la géométrie de la sonde, l'élément piézo-électrique du capteur et la position angulaire du moteur pour le tir considéré. Les valeurs d'intensité des voxels du volume qui ne se trouvent pas sur une ligne de tir sont calculées par technique d'interpolation. Afin de permettre une reconstruction en temps réel du volume, c'est à dire à la cadence des mesures fournies par la sonde, l'algorithme a été implémenté sur GPU.

4.4 Conclusion

Ce chapitre a traité de solutions innovantes basées sur l'exploitation de l'information dense et adaptées aux spécificités des capteurs échographiques. D'une part, la corrélation spatiale inhérente au speckle propre à l'échographie apporte une information suffisamment riche pour être utilisée à des fins de contrôle de la sonde 2D pour le suivi d'une coupe d'intérêt. D'autre part, l'asservissement visuel peut directement s'appuyer sur l'intensité des pixels. Cette approche à l'avantage d'offrir un champ d'applications plus vaste que celle basée sur la corrélation du speckle. Ainsi, des tâches de suivi et de positionnement ont été envisagées. En outre, la problématique de la segmentation des images échographiques est ici éludée. Cet aspect garantit la généralité des méthodes par rapport à la nature des images. Par ailleurs, la prise en compte des déformations s'avère une problématique encore ouverte. Toutefois, les premiers résultats montrent la faisabilité d'une approche tenant compte de ces déformations.

Conclusion et perspectives

Mes principales contributions à l'asservissement visuel échographique ont été décrites dans ce mémoire. Conformément aux différents travaux rapportés dans ce document, mon objectif scientifique a été de fournir de nouvelles solutions méthodologiques pour la commande en boucle fermée de systèmes robotiques à partir d'informations issues des images échographiques. Jusqu'à présent, nous avons essentiellement considéré la configuration embarquée où la sonde est actionnée par un manipulateur robotique. Dans une telle configuration, la cible visuelle est généralement de nature anatomique et, contrairement aux applications de manipulation d'un instrument chirurgical, il n'est pas possible d'utiliser une connaissance a priori sur la géométrie de l'outil manipulé pour modéliser son interaction avec la sonde et pour faciliter sa détection dans l'image.

Les verrous scientifiques afférents que nous avons proposé de relever sont multiples. Tout d'abord, il s'agit de réaliser le contrôle de la totalité des degrés de liberté d'une sonde échographique 2D, c'est-à-dire également les déplacements effectués en dehors du plan de coupe. En effet, un capteur échographique 2D ne fournit aucune information en dehors du plan de coupe. Les premiers asservissements visuels échographiques proposés dans la littérature se sont donc limités à contrôler les trois degrés de liberté correspondant aux deux mouvements de translation et à celui de rotation dans le plan image.

Ensuite, à partir des images échographiques, l'extraction robuste et en temps réel des informations visuelles nécessaires à la commande du système robotique constitue une difficulté importante. Cette extraction peut être réalisée au moyen d'un algorithme de segmentation si les images échographiques contiennent des structures anatomiques fortement contrastées les unes par rapport aux autres, tels le système vasculaire du foie, les reins, la vésicule biliaire, des kystes, des tumeurs ou d'autres éléments discernables dans l'image. En revanche, lorsque les images échographiques sont très faiblement structurées, il est alors difficile d'extraire des informations visuelles géométriques de manière robuste.

Nous avons ainsi apporté des solutions génériques pour permettre le contrôle complet des déplacements de la sonde à partir d'images échographiques présentant ou non des structures géométriques décelables. Deux démarches principales ont été menées pour atteindre ces objectifs. La première a porté sur l'utilisation de primitives visuelles géométriques en entrée d'un asservissement visuel 2D, et la seconde, sur l'exploitation de l'information dense de l'image échographique.

5.1 Contributions : primitives visuelles géométriques

Nos travaux ont tout d'abord porté sur le choix et la modélisation des informations visuelles qu'il est possible d'extraire en temps réel à partir des images échographiques 2D et qui sont pertinentes pour la réalisation d'une tâche de positionnement d'une sonde 2D. Une tâche générique a été élaborée et consiste à déplacer automatiquement la sonde dans le but d'atteindre une section anatomique désirée préalablement mémorisée. Afin de contrôler également les déplacements de la sonde hors du plan de coupe, une solution consiste à considérer que la sonde interagit avec un élément anatomique de forme connue dont les paramètres 3D sont fournis par une imagerie préalable. Cette hypothèse a permis d'investiguer différentes primitives visuelles géométriques dont la matrice d'interaction associée a été déterminée analytiquement. Ces primitives ont été ensuite testées dans la commande. Ces études préliminaires nous ont permis d'aboutir au choix de primitives visuelles définies à partir des moments 2D de la section de l'objet considéré.

Par ailleurs, nous avons pu mettre analytiquement en évidence que la variation des moments 2D due au mouvement en dehors du plan de la sonde est directement dépendante des composantes du vecteur normal à la surface de l'objet pour chaque point du contour de la section observée dans l'image échographique. Afin de s'affranchir d'une connaissance de la forme de l'objet avec lequel la sonde interagit, nous avons proposé plusieurs méthodes permettant d'estimer en ligne le vecteur normal à la surface de l'objet à partir conjointement des mesures des points de contour observés dans les images successives et des mesures odométriques du robot. Un choix de 6 primitives visuelles indépendantes construites à partir des moments a été proposé pour obtenir une convergence de la commande en termes d'erreur visuelle et également en position.

Cependant dans le cas d'objets de forme très fortement symétrique, plusieurs minima globaux existent, et l'approche ainsi développée n'assure pas systématiquement la convergence de l'asservissement en position. En effet, bien que la convergence dans l'image soit atteinte, dans la mesure où une même valeur des moments choisis peut être observée pour des positions différentes de la sonde, la position désirée peut ne pas être obtenue. Pour caractériser au mieux la pose de la sonde avec des informations géométriques 2D, nous avons alors proposé un asservissement visuel tri-plans où le vecteur d'informations visuelles est composé de six primitives géométriques couplées principalement à un mouvement donné de la sonde. En raison de ce couplage, une forme simplifiée de la matrice d'interaction, entièrement calculée à partir des informations contenues dans trois plans image orthogonaux fournis par un capteur échographique 3D, a été définie.

Du point de vue applicatif, nous avons intégré cette tâche de positionnement automatique sur le dispositif de télé-échographie robotisée du projet ANR PROSIT afin de fournir au médecin des fonctionnalités d'assistances automatiques. Ainsi, une tâche de récupération automatique de section et une tâche de maintien de visibilité par asservissement visuel lors de la télé-opération ont été mises en pratique avec succès. A cette fin, nous avons également élaboré un algorithme de segmentation robuste capable de détecter et suivre en temps réel des éléments anatomiques d'intérêt dans l'image échographique. Cet algorithme est basé sur un contour actif dont une paramétrisation, utilisant les descripteurs de Fourier pour permettre la détection de section concave et convexe, a été réalisée.

5.2 Contributions : information dense

Toutefois, les méthodes basées sur des primitives visuelles géométriques ne sont pas adaptées pour tous les types d'images échographiques. En effet, elles nécessitent la présence de structures géométriques perceptibles et notamment des sections d'éléments anatomiques dont le contour est fermé dans l'image échographique observée. Dans le but de considérer des images échographiques les plus diverses possibles, nous avons dans une seconde démarche orienté nos travaux sur l'élaboration de méthodes d'asservissement visuel utilisant directement l'information dense contenue dans les images échographiques et permettant ainsi de s'affranchir de l'étape de segmentation.

L'information de texture de type « speckle » contenue dans l'image échographique peut être utilisée pour contrôler les déplacements d'une sonde 2D. Généralement perçu comme un bruit dans les images échographiques, le speckle est en fait une information qui est spatialement cohérente. En pratique, la coupe échographique n'est pas parfaite mais comporte une certaine largeur d'élévation (dans la direction orthogonale au plan image). Il en résulte une corrélation du speckle entre deux coupes relativement proches en fonction de la distance d'élévation. L'approche que nous avons proposée consiste à utiliser la mesure de corrélation du speckle pour reconstruire la position relative entre la coupe courante fournie par la sonde 2D et une coupe de référence apprise à un instant donné et évoluant avec les mouvements physiologiques du patient. Cette information de position est ensuite directement minimisée par un asservissement visuel de façon à réaliser une tâche de suivi d'une coupe de référence afin de stabiliser l'image échographique.

Plus récemment, nous avons proposé une nouvelle approche d'asservissement visuel échographique 2D dont l'originalité tient à l'utilisation directe de l'image échographique comme information visuelle. Le vecteur d'information visuelle est dans ce cas constitué des valeurs d'intensité de l'ensemble des pixels contenus dans une région d'intérêt. Par rapport à une approche géométrique, il s'agit néanmoins d'une méthode plus locale où la non linéarité du système et le nombre important d'informations visuelles considérées rendent l'asservissement visuel davantage sensible aux minima locaux. Néanmoins cette solution a pour avantage de convenir à un large éventail d'images anatomiques et se prête bien au suivi d'une section anatomique quelconque. Nous avons démontré que la forme analytique de la matrice d'interaction associée à ce vecteur d'informations visuelles dépend du gradient 3D de l'image. Ce dernier peut être maintenu constant dans le cas d'une tâche de poursuite et estimé lors d'une procédure

initiale qui consiste à acquérir des coupes parallèles voisines. Par contre, dans le cas d'une tâche de positionnement, il est souhaitable d'estimer le gradient en ligne par un algorithme adapté à l'utilisation d'une sonde 2D ou bi-plan, ou en calculant à la volée le gradient 3D à partir d'un volume fourni par une sonde motorisée.

5.3 Vers une prise en compte des déformations

Pour aller plus loin en termes d'applications, nous avons récemment orienté nos travaux pour prendre en compte la nature déformable des tissus organiques. Dans ce sens, une méthode a été développée de manière à estimer en temps réel les déformations des tissus mous observés dans une séquence de volumes échographiques fournie par une sonde échographique 3D. Ces déformations sont paramétrées par une grille 3D déformable dont l'évolution est estimée à partir de la différence d'intensité des voxels contenus dans les volumes successifs. A partir de la déformation estimée, le mouvement rigide de la grille est extrait afin de le compenser par un asservissement visuel 3D, et ceci en vue de réaliser une stabilisation active du volume observé. Les premiers résultats ont démontré la faisabilité d'une telle approche.

5.4 Perspectives

Les perspectives de ces travaux sur l'asservissement visuel échographique sont nombreuses dans la mesure où l'objectif final est de fournir des solutions robustes aux conditions expérimentales *in vivo*. En effet, jusqu'à présent, nous avons essentiellement validé ces approches sur des fantômes artificiels ou constitués de tissu animal auxquels nous avons appliqué des mouvements rigides ou des déformations externes. Il est donc souhaitable, lors de futurs travaux, d'augmenter le degré de réalisme afin de valider les approches en conditions *in vivo* avant d'envisager le développement d'applications médicales directement utilisable en routine clinique.

Par ailleurs, il persiste de nombreux verrous scientifiques auxquels des réponses devront être trouvées. Pour l'instant, nous n'avons pas pris en compte dans nos différentes approches les perturbations dues à la déformation interne des organes, à l'exception de nos travaux portant sur l'estimation des déformations des tissus mous à partir d'une séquence de volumes échographiques. Cette prise en compte des déformations est donc une perspective directe et nécessaire pour garantir la robustesse des approches. Bien entendu, la prise en compte des déformations n'est pas simple, que ce soit au niveau des approches de type géométrique ou dense.

Dans ce cadre, l'approche utilisant les moments 2D de la section observée nécessite une estimation en ligne du vecteur normal à la surface de l'objet. L'estimation du vecteur serait faussée dans le cas où l'objet se déforme. Il pourrait en résulter une erreur trop importante sur l'estimation de la matrice d'interaction qui conduirait à une instabilité du système. Néanmoins, ce problème ne concerne pas le cas de l'asservissement visuel tri-plans puisque la matrice d'interaction utilisée dans ce cas n'est pas construite à partir du vecteur normal à la surface. Une solution au problème rencontré dans le cas de l'utilisation d'une sonde 2D serait de considérer

un modèle dynamique de l'objet dont les paramètres seraient actualisés en fonction du temps à partir d'une phase d'apprentissage sur plusieurs cycles respiratoires du patient.

Quant à l'approche basée sur l'intensité, celle-ci montre déjà un bon comportement sur les expériences robotiques réalisées notamment dans le cadre d'applications de positionnement et de suivi. Néanmoins, à l'heure actuelle, seuls les mouvements rigides sont pris en compte dans la loi de commande et la modélisation des déformations semble être une perspective prometteuse pour améliorer la robustesse de la méthode dans le cas d'une application in vivo. La prise en compte d'un modèle de déformation pourrait être réalisée en intégrant au système des paramètres supplémentaires correspondant à ceux de la transformation non rigide. L'objectif de la loi de commande serait alors de minimiser l'erreur visuelle tout en optimisant ces paramètres additionnels.

Je souhaite également poursuivre les activités liées à la compensation du mouvement physiologique du patient par stabilisation active de l'image échographique. Jusqu'à présent, nous avons testé nos approches uniquement en condition ex vivo en appliquant au moyen d'un second système robotique des mouvements périodiques à des fantômes ou échantillons de tissus animaux. L'étape suivante consistera à expérimenter les méthodes en condition in vivo sur animal porcin sédaté. Cette validation expérimentale est une étape nécessaire pour la valorisation future des travaux en termes d'applications médicales.

Suite à des discussions avec des médecins radiologues des CHU de Rennes et de Nîmes à qui les résultats actuels ont été présentés, différentes cibles médicales qui pourraient bénéficier directement de ces résultats méthodologiques ont été définies. Les applications potentielles concernent en premier lieu les examens échographiques faisant intervenir des mesures de flux sanguin (Doppler). Pour réaliser de telles mesures dans l'image échographique, la section longitudinale de l'artère considérée doit être observée avec la sonde échographique pendant quelques minutes. Au cours de cette observation, une tâche de suivi de l'artère permettrait de garantir la validité de la mesure en compensant de manière active tout mouvement involontaire ou physiologique du patient durant la durée de l'examen. Par ailleurs, la compensation active peut également s'avérer très utile lors des procédures de mesure de la dilatation des vaisseaux induite par le flux sanguin. De telles mesures sont généralement réalisées au niveau de l'artère brachiale pour mesurer l'effet de différents facteurs sur les fonctions vasculaires du patient, tels la réaction à la prise de médicament, le changement de régime alimentaire ou les conséquences d'un effort physique. Le principe de cet examen est d'évaluer la réactivité de l'artère brachiale en mesurant en continu sur plusieurs dizaines de minutes l'évolution de son diamètre à un stress induit par un afflux sanguin. Cette étude est actuellement très délicate à réaliser, car il s'agit pour le médecin de synchroniser manuellement les déplacements de la sonde avec le mouvement de l'artère durant toute la durée de l'examen, afin qu'il puisse relever des mesures correctes du diamètre. Une stabilisation automatique de l'image par asservissement visuel pourrait sans aucun doute faciliter l'examen.

Des applications chirurgicales pourraient également bénéficier de nos contributions à plus long terme. En effet, la stabilisation automatique de l'image sur une tumeur mobile permettrait un traitement localisé plus efficace. Ce dernier pourrait ainsi être synchronisé avec le déplacement physiologique des organes. Typiquement, la radiothérapie s'avère être une application candidate : le point de focalisation du faisceau de traitement pourrait ainsi suivre automati-

quement le déplacement de la cible à l'aide d'un asservissement visuel échographique pour garantir un traitement optimal.

Par ailleurs, j'envisage dès à présent de considérer le cas où la sonde n'est pas embarquée en orientant plus particulièrement mes travaux de recherche vers l'étude de solutions pour le guidage d'aiguilles rigides ou déformables afin d'assister le geste de biopsie ou de thérapie par asservissement visuel échographique. L'aiguille est un outil utilisé dans de très nombreux gestes médicaux en radiologie interventionnelle comme la biopsie, la ponction et l'ablation localisée de tumeur (cryo- ou thermo-ablation). Afin d'assister le médecin lors de son geste, il serait intéressant de proposer une solution de guidage de l'aiguille vers la cible au moyen d'un système robotique utilisant directement l'information visuelle per-opératoire fournie par une sonde échographique. Un aspect de ces travaux concernera l'étude et la mise en œuvre de méthodes de traitement d'images permettant de détecter et suivre en temps réel, à partir d'une séquence de volumes échographiques, une aiguille déformable lors de son insertion. Le second aspect portera sur l'étude de méthodes d'asservissement visuel permettant de guider automatiquement un robot manipulant l'aiguille à partir des images fournies par la sonde et des mesures d'effort d'interaction entre l'aiguille et les tissus.

Pour l'instant, nous avons essentiellement considéré des images échographiques de mode B (Brillance). J'envisage d'étudier de nouvelles approches d'asservissement visuel échographique en utilisant des images en mode Doppler continu, et plus particulièrement la mesure du flux sanguin associée aux différents pixels de l'image. L'utilisation du mode Doppler couleur permettrait d'une part d'augmenter la robustesse des approches utilisant des primitives visuelles géométriques puisque la segmentation de l'image serait facilitée. D'autre part, l'information de vitesse permettrait d'introduire une information supplémentaire utile pour la prise en compte du caractère dynamique et déformable des structures d'intérêt. Cette modalité d'image serait également appropriée pour développer des méthodes d'asservissement visuel permettant de guider des robots navigant dans le réseau vasculaire, tels des micro-capsules ou des cathéters robotisés. Dans ce sens, j'ai déjà initié une collaboration scientifique avec le Children's Hospital à Boston dans le contexte applicatif de la chirurgie cardiaque robotisée.

A plus long terme, j'envisage également de développer des méthodes d'asservissement visuel combinant d'autres modalités d'imagerie, telle la résonance magnétique (IRM). En effet, depuis quelques années, des équipes de recherche ont développé des robots légers pouvant entrer dans les scanners. Ces systèmes, dont la plupart sont toujours en phase de prototype, sont destinés à la radiologie interventionnelle dans le but d'assister l'insertion d'aiguille dans une tumeur ou une zone d'intérêt. Actuellement, la procédure est réalisée au moyen de techniques de recalage qui utilisent généralement des marqueurs stéréotaxiques pour localiser la position du robot dans le repère de l'imageur. La cible à atteindre est définie manuellement par le médecin dans l'image à un instant donné. Dans ce contexte, il serait intéressant d'utiliser l'asservissement visuel pour commander les déplacements de l'aiguille directement à partir des images fournies par le scanner où la cible serait suivie en permanence, et de permettre ainsi un positionnement de l'aiguille avec une précision de l'ordre du pixel tout en compensant les mouvements physiologiques du patient.

Bien que ces perspectives de recherche se situent dans le contexte de la robotique médicale, je souhaite également orienter mes travaux vers d'autres champs applicatifs en considérant d'autres types de capteurs visuels « exotiques ».

L'imagerie sonar est une modalité très proche de l'échographie ; la formation de l'image résulte des mêmes principes physiques. De ce fait, il serait intéressant de mener des travaux d'investigation pour voir s'il est possible d'appliquer les méthodes proposées en échographie pour ce type de capteur en vue de réaliser, par exemple, le contrôle par asservissement visuel de véhicules autonomes sous-marins (AUV) ou marins.

La modélisation de l'interaction entre une sonde 2D et la scène peut également s'appliquer à d'autres capteurs fournissant une image correspondant à une coupe de l'environnement. En effet, les méthodes génériques proposées pourraient être utilisées, par exemple, dans des applications de métrologie industrielle « active » où l'asservissement visuel permettrait de positionner le capteur sur une coupe de l'objet à inspecter.

Nos approches pourraient également être étendues à l'usage de scanners laser de type lidar. A la différence de l'imagerie échographique, ces capteurs ne fournissent pas une section complète de l'environnement observé, mais une image représentant le profil extérieur visible de la scène selon le plan de balayage du capteur. Il serait donc intéressant de développer et mettre en œuvre des asservissements visuels utilisant ce type d'image dans le contexte de la robotique mobile.

Bibliographie

- [Abolmaesumi 2001] P. Abolmaesumi, S.E. Salcudean, W.H. Zhu, S.P. DiMaio et M.R. Siroospour. *A user interface for robot-assisted diagnostic ultrasound*. In IEEE Int. Conf. on Robotics and Automation, ICRA'01, volume 2, pages 1549–1554, 2001. [13](#), [38](#)
- [Abolmaesumi 2002] P. Abolmaesumi, S. Salcudean, W. Zhu, M. Siroospour et S. DiMaio. *Image-guided control of a robot for medical ultrasound*. IEEE Trans. on Robotics and Automation, vol. 18, no. 1, pages 11–23, feb 2002. [13](#), [38](#)
- [Allgower 1999] F. Allgower, T.A. Badgwell, S.J. Qin, J B. Rawlings et S.J. Wright. *Non-linear predictive control and moving horizon estimation - an introductory overview*. Advances in Control : Highlights of ECC'99, pages 391–449, 1999. [42](#)
- [Arbeille 2008] P. Arbeille, J. Ayoub, V. Kieffer, P. Ruiz, B. Combes, A. Coitrieux, P. Herve, S. Garnier, B. Lepartz, E. Lefbvre et F. Perrotin. *Realtime tele-operated abdominal and fetal echography in 4 medical centres, from one expert center, using a robotic arm & ISDN or satellite link*. In IEEE Int. Conf. on Automation, Quality and Testing, Robotics, AQTR'08, volume 1, pages 45–46, May 2008. [76](#)
- [Bachta 2006] W. Bachta et A. Krupa. *Towards ultrasound image-based visual servoing*. In IEEE Int. Conf. on Robotics and Automation, ICRA'2006, pages 4112–4117, Orlando, Florida, May 2006. [15](#), [52](#)
- [Baker 2004] S. Baker et I. Matthews. *Lucas-kanade 20 years on : A unifying framework*. Int. Journal of Computer Vision, vol. 56, no. 3, pages 221–255, 2004. [89](#)
- [Belaid 2009] A. Belaid, D. Boukerroui, Y. Maingourd et J.F. Lerallut. *Phase based level set segmentation of ultrasound images*. In 9th Int. Conf. on Information Technology and Applications in Biomedicine, pages 1–4, 2009. [68](#)
- [Benabid 1987] A. L. Benabid, P. Cinquin, S. Lavalley, J.F. Le Bas, J. Demongeot et J. de Rougemont. *Computer-driven robot for stereotactic surgery connected to CT scan and magnetic resonance imaging : Technological design and preliminary results*. Applied Neurophysiology, vol. 50, pages 153–154, 1987. [11](#)

- [Benhimane 2004] S. Benhimane et E. Malis. *Real-time image-based tracking of planes using efficient second-order minimization*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'04, pages 943–948, Sendai, Japan, September 2004. 89
- [Boctor 2005] E. Boctor, I. Iordachita, G. Fichtinger et G.D. Hager. *Real-time quality control of tracked ultrasound*. In Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, MICCAI'05, pages 621–630, Palm Springs, California, October 2005. 87
- [Bohs 2000] L.N. Bohs, B.J. Geiman, M.E. Anderson, S.C. Gebhart et G.E. Trahey. *Speckle tracking for multi-dimensional flow estimation*. Ultrasonics, vol. 28, no. 1, pages 369–375, 2000. 86
- [Bookstein 1989] F.L. Bookstein. *Principal Warps : Thin-Plate Splines and the decomposition of deformations*. Analysis, vol. 1, no. 6, 1989. 107
- [Bricault 2008] I. Bricault, N. Zemiti, E. Jouniaux, C. Fouard, E. Taillant, F. Dorandeu et P. Cinquin. *A light puncture robot for CT and MRI interventions*. IEEE Engineering in Medicine and Biology Magazine, pages 42–50, 2008. 12
- [Brigger 2000] P. Brigger, J. Hoeg et M. Unser. *B-spline snakes : a flexible tool for parametric contour detection*. IEEE Trans. on Image Processing, vol. 9, no. 9, pages 1484–1496, sep. 2000. 68, 69
- [Casal 1996] A. Casal, J. Amat et E. Laporte. *Automatic guidance of an assistant robot in laparoscopic surgery*. In IEEE Int. Conf. on Robotics and Automation, pages 895–900, 1996. 13
- [Chan 2001] T.F. Chan et L.A. Vese. *Active contours without edges*. IEEE Trans. on Image Processing, vol. 10, no. 2, pages 266–277, feb. 2001. 68
- [Chang 2003] R.F. Chang, W.J. Wu, D.R. Chen, W.M. Chen, W. Shu, J.H. Lee et L.B. Jeng. *3-D us frame positioning using speckle decorrelation and image registration*. Ultrasound in Medicine and Biology, vol. 29, no. 6, pages 801–812, 2003. 86
- [Chaumette 1990] F. Chaumette. *La relation vision-commande : théorie et application à des tâches robotiques*. PhD thesis, Université de Rennes 1, Mention informatique, July 1990. 36
- [Chaumette 2002] F. Chaumette. *Asservissement visuel*. In W. Khalil, editeur, La commande des robots manipulateurs, Traité IC2, chapitre 3, pages 105–150. Hermès, 2002. 35
- [Chaumette 2004] F. Chaumette. *Image moments : a general and useful set of features for visual servoing*. IEEE Trans. on Robotics, vol. 20, no. 4, pages 713–723, August 2004. 36, 55
- [Chaumette 2006] F. Chaumette et S. Hutchinson. *Visual servo control, Part I : Basic approaches*. IEEE Robotics and Automation Magazine, vol. 13, no. 4, pages 82–90, December 2006. 29, 33
- [Chaumette 2007] F. Chaumette et S. Hutchinson. *Visual servo control, Part II : Advanced approaches*. IEEE Robotics and Automation Magazine, vol. 14, no. 1, pages 109–118, March 2007. 29
- [Chimiak 1996] W.J. Chimiak, R.O. Rainer, N.T. Wolfman et W. Covitz. *Architecture for a high-performance tele-ultrasound system*. In Medical imaging : PACS design and evaluation : Engineering and clinical issues, volume 2711, pages 459–465, 1996. 76

- [Clarke 1987] D. W. Clarke, C. Mohtadi et P. S. Tuffs. *Generalized predictive control - Part I : The basic algorithm*. Automatica, vol. 23, pages 137–160, 1987. 102
- [Cohen 1991] L.D. Cohen. *On active contour models and balloons*. CVGIP : Image Understanding, vol. 53, no. 2, pages 211–218, 1991. 73
- [Cohen 1993] L.D. Cohen et I. Cohen. *Finite-element methods for active contour models and balloons for 2-D and 3-D images*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pages 1131–1147, November 1993. 73
- [Collewet 2000] C. Collewet et F. Chaumette. *A contour approach for image-based control of objects with complex shape*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'00, volume 1, pages 751–756, Takamatsu, Japon, November 2000. 36
- [Collewet 2009] C. Collewet. *Polar snakes : A fast and robust parametric active contour model*. In 16th IEEE Int. Conf. on Image Processing, pages 3013–3016, 2009. 68, 69, 71
- [Collewet 2011] C. Collewet et E. Marchand. *Photometric visual servoing*. IEEE Trans. on Robotics, vol. 27, no. 4, pages 828–834, August 2011. 93
- [Comport 2006] A.I. Comport, E. Marchand, M. Pressigout et F. Chaumette. *Real-time markerless tracking for augmented reality : the virtual visual servoing framework*. IEEE Trans. on Visualization and Computer Graphics, vol. 12, no. 4, pages 615–628, July 2006. 36
- [Crétual 2001] A. Crétual et F. Chaumette. *Visual servoing based on image motion*. Int. Journal of Robotics Research, vol. 20, no. 11, pages 857–877, November 2001. 37
- [Dagon 2004] N. Dagon, P. Frossard, M. Malik et N. Richli. *Techniques échographiques*, http://files.chuv.ch/internet-docs/rad/techniciens/rad_trm_us_technique_echo2.pdf, 2004. 25
- [Delgorge 2005] C. Delgorge, F. Courreges, L. A. Bassit, C. Novales, C. Rosenberger, N. Smith-Guerin, C. Bru, R. Gilabert, M. Vannoni, G. Poisson et P. Vieyres. *A tele-operated mobile ultrasound scanner using a light-weight robot*. IEEE Trans. on Information Technology in Biomedicine, vol. 9, no. 1, pages 50–58, Mars 2005. 11, 76
- [DeMenthon 1995] Daniel F. DeMenthon et Larry S. Davis. *Model-based object pose in 25 lines of code*. International Journal of Computer Vision, vol. 15, pages 123–141, 1995. 36
- [Detmer 1994] P.R. Detmer, G. Basheim, T. Hodges, K.W. Beach, E.P. Filer, D.H. Burns et D.E. Strandness. *3D ultrasonic image feature localization based on magnetic scanhead tracking : in vitro calibration and validation*. Ultrasound in Medicine and Biology, vol. 20, no. 9, pages 923–936, 1994. 50, 52
- [Drummond 2002] T. Drummond et R. Cipolla. *Real-time visual tracking of complex structures*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pages 932–946, July 2002. 36
- [Espiau 1992] B. Espiau, F. Chaumette et P. Rives. *A new approach to visual servoing in robotics*. IEEE Trans. on Robotics and Automation, vol. 8, no. 3, pages 313–326, 1992. 30, 31, 33, 34, 35

- [Felix 2005] N. Felix, D. Voisin, O. Clade et R. Dufait. *Biplane ultrasound arrays with integrated multiplexing solution for enhanced diagnostic accuracy in endorectal and transvaginal imaging*. In IEEE Int. Ultrasonics Symposium, Rotterdam, The Netherlands, 2005. 27
- [Fonte 2010] A. Fonte, T. Essomba, P. Vieyres, J. Canou, S. P. Fraisse, Zeghloul, A. Krupa et P. Arbeille. *Robotic Platform for an Interactive Tele-echographic System : The PROSIT ANR-2008 project*. In Hamlyn Symposium on Medical robotics, London, UK, May 2010. 76
- [Freeman 1977] H. Freeman et L.S. Davis. *A corner-finding algorithm for chain-coded curves*. IEEE Trans. on computers, vol. C-26, no. 3, pages 297–303, mar. 1977. 74
- [Gangloff 1999] J. Gangloff. *Asservissements visuels rapides d'un robot manipulateur à 6 degrés de liberté*. PhD thesis, Université Louis Pasteur - Strasbourg I, Strasbourg, January 1999. 30
- [Gangloff 2003] J. Gangloff et M. de Mathelin. *High speed visual servoing of a 6 DOF manipulator using multivariable predictive control*. Advanced Robotics, Special issue : advanced 3D vision and its application to robotics, December 2003. 30
- [Gangloff 2006] J. Gangloff, R. Ginhoux, M. De Mathelin, L. Soler et J. Marescaux. *Model predictive control for compensation of cyclic organ motions in teleoperated laparoscopic surgery*. IEEE Trans. on Control System Technology, vol. 14, pages 235–246, 2006. 102, 103
- [Gee 2006] A.H. Gee, R.J. Housden, P. Hassenpflug, G.M. Treece et R.W. Prager. *Sensorless freehand 3D ultrasound in real tissues : Speckle decorrelation without fully developed speckle*. Medical Image Analysis, vol. 10, no. 2, pages 137–149, 2006. 86
- [Ginhoux 2004] R. Ginhoux, J. Gangloff, M. de Mathelin, L. Soler, M. Sanchez et J. Marescaux. *Beating heart tracking in robotic surgery using 500 Hz visual servoing, model predictive control and an adaptive observer*. In IEEE Int. Conf. on Robotics and Automation, pages 274–279, 2004. 13
- [Glozman 2004] D. Glozman et M. Shoham. *Flexible needle steering and optimal trajectory planning for percutaneous therapies*. In Int. Conf. on Medical Image Computing and Computer Assisted Intervention, MICCAI'04, Saint-Malo, France, September 2004. 41
- [Guthart 2000] G.S. Guthart et J.K. Salisbury. *The Intuitive telesurgery system : Overview and application*. In IEEE Int. Conf. on Robotics and Automation, pages 618–621, San Francisco, USA, 2000. 11
- [Hadwiger 2006] M. Hadwiger, J.M. Kniss, C. Rezk-salama, D. Weiskopf et K. Engel. *Real-time volume graphics*. 2006. 96
- [Hager 1998] G.D. Hager et P.N. Belhumeur. *Efficient region tracking with parametric models of geometry and illumination*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 10, pages 1025–1039, 1998. 87, 89
- [Hong 2004] J. Hong, T. Dohi, M. Hashizume, K. Konishi et N. Hata. *An ultrasound-driven needle-insertion robot for percutaneous cholecystostomy*. Physics in Medicine and Biology, vol. 49, no. 3, pages 441–455, 2004. 13, 40

- [Hsu 2009] P.W. Hsu, R.W. Prager, A.H. Gee et G.M. Treece. *Freehand 3D ultrasound calibration : A review*. *Advanced Imaging in Biology and Medicine*, vol. 1, pages 47–84, 2009. [49](#)
- [Hu 1962] M.K. Hu. *Visual pattern recognition by moment invariants*. *IRE Trans. Information Theory*, vol. 8, pages 179–187, 1962. [55](#)
- [Hungr 2009] N. Hungr, J. Troccaz, N. Zemiti et N. Tripodi. *Design of an ultrasound-guided robotic brachytherapy needle-insertion system*. In *IEEE Int. Conf. on Engineering in Medicine and Biology Society*, pages 250–253, Minneapolis, USA, 2009. [12](#)
- [Impedovo 1978] S. Impedovo, B. Marangelli et A. M. Fanelli. *A fourier descriptor set for recognizing nonstylized numerals*. *IEEE Tran. on Systems, Man and Cybernetics*, vol. 8, no. 8, pages 640–645, August 1978. [69](#), [71](#)
- [Kass 1987] M. Kass, A. Witkin et D. Terzopoulos. *Snakes : Active contour models*. *Int. Journal Of Computer Vision*, vol. 1, no. 4, pages 321–331, 1987. [69](#)
- [Kermorgant 2011] O. Kermorgant et F. Chaumette. *Combining IBVS and PBVS to ensure the visibility constraint*. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'11*, pages 2849–2854, 2011. [79](#)
- [Krupa 2003a] A. Krupa. *Commande par vision d'un robot de chirurgie laparoscopique*. PhD thesis, Institut National Polytechnique de Lorraine, July 2003. [14](#)
- [Krupa 2003b] A. Krupa, J. Gangloff, C Doignon, M. de Mathelin, G. Morel, J. Leroy, L. Soler et J. Marescaux. *Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing*. *IEEE Trans. on Robotics and Automation*, vol. 19, no. 5, pages 842–853, October 2003. [13](#)
- [Krupa 2005] A. Krupa et F. Chaumette. *Control of an ultrasound probe by adaptive visual servoing*. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'05*, volume 2, pages 2007–2012, Edmonton, Canada, August 2005. [48](#)
- [Krupa 2006a] A. Krupa. *Automatic calibration of a robotized 3D ultrasound imaging system by visual servoing*. In *IEEE Int. Conf. on Robotics and Automation, ICRA'2006*, pages 4136–4141, Orlando, Florida, May 2006. [15](#), [50](#)
- [Krupa 2006b] A. Krupa et F. Chaumette. *Guidance of an ultrasound probe by visual servoing*. *Advanced Robotics*, vol. 20, no. 11, pages 1203–1218, November 2006. [15](#), [49](#)
- [Krupa 2007a] A. Krupa, G. Fichtinger et G. Hager. *Full motion tracking in ultrasound using image speckle information and visual servoing*. In *IEEE Int. Conf. on Robotics and Automation, ICRA'07*, pages 2458–2464, Roma, Italy, April 2007. [90](#), [91](#)
- [Krupa 2007b] A. Krupa, G. Fichtinger et G.D. Hager. *Real-time tissue tracking with B-mode ultrasound using speckle and visual servoing*. In *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, MICCAI'07*, volume 2, pages 1–8, Brisbane, Australia, October 2007. [91](#)
- [Krupa 2009a] A. Krupa. *Asservissement visuel par imagerie médicale*. In *Journées Nationales de la Recherche en Robotique, JNRR'09*, Neuvy-sur-Barangeon, France, November 2009. [13](#)
- [Krupa 2009b] A. Krupa, G. Fichtinger et G.D. Hager. *Real-time motion stabilization with B-mode ultrasound using image speckle information and visual servoing*. *The International Journal of Robotics Research, IJRR*, vol. 28, no. 10, pages 1334–1354, 2009. [16](#), [91](#), [92](#)

- [Laporte 2007] C. Laporte et T. Arbel. *Probabilistic speckle decorrelation for 3d ultrasound*. In Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, MIC-CAI'07, volume 1, pages 925–932, Brisbane, Australia, October 2007. [87](#)
- [Lawrence 1997] C. Lawrence, J.L. Zhou et A.L. Tits. *User's Guide for CFSQP Version 2.5 : A C code for solving (large Scale) constrained nonlinear (minimax) optimization problems, generating iterates satisfying all inequality constraints*. Rapport technique, Electrical Engineering Department and Institute for Systems Research, University of Maryland, 1997. [43](#)
- [Lee 2007] D. Lee, N. Koizumi, K. Ota, S. Yoshizawa, A. Ito, Y. Kaneko, Y. Matsumoto et M. Mitsuishi. *Ultrasound-based visual servoing system for lithotripsy*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'07, pages 877–882, San Diego, USA, November 2007. [39](#), [47](#)
- [Lee 2011] D. Lee et A. Krupa. *Intensity-based visual servoing for non-rigid motion compensation of soft tissue structures due to physiological motion using 4D ultrasound*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'11, pages 2831–2836, San Francisco, USA, September 2011. [16](#), [107](#)
- [Li 2011a] T. Li, A. Krupa et C. Collewet. *A robust parametric active contour based on Fourier descriptors*. In IEEE Int. Conf. on Image Processing, ICIP'11, Brussels, Belgium, September 2011. [15](#), [71](#)
- [Li 2011b] T. Li, A. Krupa et C. Collewet. *Un contour actif robuste basé sur les descripteurs de Fourier*. In ORASIS'11, journées francophones des jeunes chercheurs en vision par ordinateur, Praz-sur-Arly, France, June 2011. [73](#), [74](#)
- [Li 2012a] T. Li. *Commande d'un robot de télé-échographie par asservissement visuel*. PhD thesis, Université de Rennes 1, Mention Traitement du Signal, à paraître 2012. [15](#)
- [Li 2012b] T. Li, O. Kermorgant et A. Krupa. *Maintaining visibility constraints during tele-echography with ultrasound visual servoing*. In IEEE Int. Conf. on Robotics and Automation, ICRA'12, Saint Paul, USA, May 2012. [16](#), [76](#), [79](#)
- [Lie 1986] W. N. Lie et Y. C. Chen. *Shape representation and matching using polar signature*. In Proc. Int. Computer Symposium, pages 710–718, 1986. [69](#)
- [Liu 1990] Y. Liu, T.S. Huang et O.D. Faugeras. *Determination of camera location from 2D to 3D line and point correspondances*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pages 28–37, 1990. [36](#)
- [Lowe 1991] D.G. Lowe. *Fitting parameterized three-dimensional models to images*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 13, no. 5, pages 441–450, May 1991. [36](#)
- [Lubinski 1999] M.A. Lubinski, S.Y. Emelianov et M. O'Donnell. *Speckle tracking methods for ultrasonic elasticity imaging using short-time correlation*. IEEE Transactions on Ultrasonic, Ferroelectrics, and Frequency Control, vol. 46, no. 1, pages 82–96, January 1999. [87](#)
- [Maintz 1998] J.B. Maintz et M.A. Viergever. *A survey of medical image registration*. Medical Image Analysis, vol. 2, no. 1, pages 1–36, 1998. [12](#)
- [Malis 1999] E. Malis, F. Chaumette et S. Boudet. *2 1/2 D visual servoing*. IEEE Trans. on Robotics and Automation, vol. 15, no. 2, pages 238–250, April 1999. [37](#)

- [Malis 2002] E. Malis et F. Chaumette. *Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods*. IEEE Trans. on Robotics and Automation, vol. 18, no. 2, pages 176–186, April 2002. 37
- [Malladi 1995] R. Malladi, J.A. Sethian et B.C. Vemuri. *Shape modeling with front propagation : a level set approach*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 17, no. 2, pages 158–175, feb. 1995. 68
- [Mansard 2009a] N. Mansard, O. Khatib et A. Kheddar. *A unified approach to integrate unilateral constraints in the stack of tasks*. IEEE Trans. on Robotics and Automation, vol. 25, no. 3, pages 670–685, 2009. 81
- [Mansard 2009b] N. Mansard, A. Remazeilles et F. Chaumette. *Continuity of varying-feature-set control laws*. IEEE Trans. on Automatic Control, vol. 54, no. 11, pages 2493–2505, November 2009. 80
- [Marchand 2002] E. Marchand et F. Chaumette. *Virtual visual servoing : A framework for real-time augmented reality*. Computer Graphics Forum, vol. 21, no. 3, pages 289–298, September 2002. 65
- [Martinet 1996] P. Martinet, J. Gallice et D. Khadraoui. *Vision based control law using 3D visual features*. In World Automation Congress, WAC'96, Robotics and Manufacturing Systems, volume 3, pages 497–502, Montpellier, France, May 1996. 35
- [Masuda 2001] K. Masuda, E. Kimura, N. Tateishi et K. Ishihara. *Three dimensional motion mechanism of ultrasound probe and its application for tele-echography system*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'01, volume 2, pages 1112–1116, October 2001. 76
- [Maurin 2005] B. Maurin, B. Bayle, O. Piccin, J. Gangloff, M. de Mathelin, C. Doignon, P. Zanne et A. Gangi. *A patient-mounted robotic platform for CT-scan guided procedures*. IEEE Trans. on Biomedical Engineering, vol. 55, no. 10, pages 2417–2425, 2005. 12
- [Mebarki 2008a] R. Mebarki, A. Krupa et F. Chaumette. *Image moments-based ultrasound visual servoing*. In IEEE Int. Conf. on Robotics and Automation, ICRA'08, pages 113–119, Pasadena, California, May 2008. 55, 56
- [Mebarki 2008b] R. Mebarki, A. Krupa et C. Collewet. *Automatic guidance of an ultrasound probe by visual servoing based on B-mode image moments*. In Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, MICCAI'08, pages 339–346, New York, September 2008. 57, 71
- [Mebarki 2009] R. Mebarki, A. Krupa et F. Chaumette. *Modeling and 3D local estimation for in-plane and out-of-plane motion guidance by 2D ultrasound-based visual servoing*. In IEEE Int. Conf. on Robotics and Automation, ICRA'09, pages 319–325, Kobe, Japan, May 2009. 55, 59
- [Mebarki 2010a] R. Mebarki. *Automatic guidance of robotized 2D ultrasound probes with visual servoing based on image moments*. PhD thesis, Université de Rennes 1, Mention Traitement du Signal, March 2010. 15, 54, 60
- [Mebarki 2010b] R. Mebarki, A. Krupa et F. Chaumette. *2D ultrasound probe complete guidance by visual servoing using image moments*. IEEE Trans. on Robotics, vol. 26, no. 2, pages 296–306, April 2010. 15, 59, 60

- [Mitsubishi 2012] M. Mitsubishi, A. Morita, N. Sugita, S. Sora, R. Mochizuki, K. Tanimoto, Y. M. Baek, H. Takahashi et K. Harada. *Master-slave robotic platform and its feasibility study for micro-neurosurgery*. Int. Journal of Medical Robotics and Computer Aided Surgery, 2012. 11
- [Mukundan 1998] R. Mukundan et Ramakrishnan. K.R. *Moment functions in image analysis : Theory and application*. World Scientific, 1998. 55
- [Nadeau 2010] C. Nadeau et A. Krupa. *A multi-plane approach for ultrasound visual servoing : application to a registration task*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'10, pages 5706–5711, Taipei, Taiwan, October 2010. 62, 65
- [Nadeau 2011a] C. Nadeau. *Asservissement visuel échographique : application au positionnement et au suivi de coupes anatomiques*. PhD thesis, Université de Rennes 1, Mention Traitement du Signal, November 2011. 16, 62, 93, 96, 101
- [Nadeau 2011b] C. Nadeau et A. Krupa. *Asservissement visuel direct d'une sonde échographique*. In Recherche en Imagerie et Technologies pour la Santé, RITS'11, Rennes, France, April 2011. 93
- [Nadeau 2011c] C. Nadeau et A. Krupa. *Improving ultrasound intensity-based visual servoing : tracking and positioning tasks with 2D and bi-plane probes*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'11, pages 2837–2842, San Francisco, USA, September 2011. 96, 97
- [Nadeau 2011d] C. Nadeau et A. Krupa. *Intensity-based direct visual servoing of an ultrasound probe*. In IEEE Int. Conf. on Robotics and Automation, ICRA'11, pages 5677–5682, Shanghai, China, May 2011. 16, 93, 96
- [Nadeau 2011e] C. Nadeau, A. Krupa et J. Gangloff. *Automatic tracking of an organ section with an ultrasound probe : Compensation of respiratory motion*. In Int. Conf. on Medical Image Computing and Computer Assisted Intervention, MICCAI'11, pages 57–64, Toronto, Canada, September 2011. 102
- [Nadeau 2012] C. Nadeau, H. Ren, A. Krupa et P.E. Dupont. *Target tracking in 3D ultrasound volumes by direct visual servoing*. In Hamlyn Symposium on Medical robotics, London, UK, July 2012. 105
- [Nakadate 2011] R. Nakadate, J. Solis, A. Takanashi, E. Minagawa, M. Sugawara et K. Niki. *Out-of-plane visual servoing method for tracking the carotid artery with a robot-assisted ultrasound diagnostic system*. In IEEE Int. Conf. on Robotics and Automation, ICRA'11, Shanghai, China, may 2011. 39
- [Nakamura 2001] Y. Nakamura, K. Kishi et H. Kawakami. *Heartbeat synchronization for robotic cardiac surgery*. In IEEE Int. Conf. on Robotics and Automation, pages 2014–2019, 2001. 13
- [Nelson 1998] T. R. Nelson et D. H. Pretorius. *Three-dimensional ultrasound imaging*. Ultrasound in Medicine and Biology, vol. 24, no. 9, pages 1243–1270, 1998. 49
- [Neubach 2010] Z. Neubach et M. Shoham. *Ultrasound-guided robot for flexible needle steering*. IEEE Trans. on Biomedical Engineering, vol. 57, no. 4, pages 799–805, 2010. 13, 41
- [Nouaille 2010] L. Nouaille, N. Smith-Guerin, G. Poisson et P. Arbeille. *Optimization of a 4 dof tele-echography robot*. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'10, pages 3501–3506, October 2010. 76

- [Novotny 2007] P. Novotny, J. Stoll, P. Dupont et R.D. Howe. *Real-Time visual servoing of a robot using three-dimensional ultrasound*. In IEEE Int. Conf. on Robotics and Automation, ICRA'07, pages 2655–2660, April 2007. [43](#), [47](#)
- [Ortmaier 2005] T. Ortmaier, M.A. Vitrani et G. Morel. *Robust real-time instrument tracking in ultrasound images for visual servoing*. In IEEE Int. Conf. on Robotics and Automation, ICRA'05, pages 2179–2184, Barcelone, Espagne, 2005. [41](#)
- [P. Dupont 2010] B. Itkowitz E. Butler P. Dupont J. Lock. *Design and control of concentric-tube robots*. IEEE Trans. on Robotics, vol. 26, no. 2, pages 209–225, 2010. [107](#)
- [Pagoulatos 1999] N. Pagoulatos, W. Edwards, D.R. Haynor et Y. Kim. *Interactive 3D registration of ultrasound and magnetic resonance images based on a magnetic position sensor*. IEEE Trans. Information Technology in Biomedicine, vol. 3, no. 4, pages 278–288, 1999. [49](#)
- [Pagoulatos 2000] N. Pagoulatos, R.N. Rohling, W.S. Edwards et Y. Kim. *New spatial localizer based on fiber optics with applications in 3D ultrasound imaging*. Proceedings of SPIE - The International Society for Optical Engineering, vol. 3976, pages 595–602, 2000. [49](#)
- [Papanikolopoulos 1993] N.P. Papanikolopoulos, P.K. Khosla et Kanade T. *Visual tracking of a moving target by a camera mounted on a robot : combination of control and vision*. IEEE Trans. on Robotics and Automation, vol. 9, no. 1, pages 14–35, 1993. [37](#)
- [Pierrot 1999] F. Pierrot, E. Dombre, E. Degoulange, L. Urbain, P. Caron, S. Boudet, J. Garipey et J. Megnier. *Hippocrate : a safe robot arm for medical applications with force feedback*. Medical Image Analysis (MedIA), vol. 3, no. 3, pages 285–300, 1999. [49](#)
- [Pluim 2003] J.P. Pluim, J.B. Maintz et M.A. Viergever. *Mutual information based registration of medical images*. IEEE Trans. on Medical Imaging, vol. 22, no. 8, pages 986–1004, 2003. [12](#)
- [Prager 1999] R.W. Prager, A.H. Gee et L. Berman. *Stradx : real-time acquisition and visualisation of freehand three-dimensional ultrasound*. Medical Image Analysis, vol. 3, no. 2, pages 129–140, 1999. [52](#)
- [Precioso 2005] F. Precioso, M. Barlaud, T. Blu et M. Unser. *Robust real-time segmentation of images and videos using a smooth-spline snake-based algorithm*. IEEE Trans. on Image Processing, vol. 14, no. 7, pages 910–924, july 2005. [68](#)
- [Prokop 1992] J. Prokop et A.P. Reeves. *A survey of moment-based techniques for unoccluded object representation and recognition*. In CVGIP : Graphical Models and Image Processing, pages 438–460, 1992. [55](#)
- [Revell 2005] J. Revell, M. Mirmehdi et D. McNally. *Computer vision elastography : speckle adaptive motion estimation for elastography using ultrasound sequences*. IEEE Transactions on Medical Imaging, vol. 24, no. 6, pages 755–766, June 2005. [87](#)
- [Richa 2009] R. Richa, P. Poignet et C. Liu. *Three-dimensional motion tracking for beating heart surgery using a Thin-plate Spline deformable model*. Int. Journal of Robotics Research, vol. 29, pages 218–230, 2009. [107](#)
- [Rohling 1998] R. Rohling. *3D freehand ultrasound : reconstruction and spatial compounding*. PhD thesis, University of Cambridge, Department of Engineering, 1998. [28](#)

- [Rousseau 2005] F. Rousseau, P. Hellier et C. Barillot. *Confhustus : a robust and fully automatic calibration method for 3D freehand ultrasound*. Medical Image Analysis, vol. 9, no. 1, pages 25–38, 2005. [50](#)
- [Salcudean 1999] S.E. Salcudean, W. H. Zhu, P. Abolmaesumi, S. Bachmann et P.D. Lawrence. *A robot system for medical ultrasound*. In 9th International Symposium of Robotics Research (ISRR'99), Snowbird, pages 152–159, 1999. [38](#)
- [Samson 1991] C. Samson, B. Espiau et M. Le Borgne. *Robot control : The task function approach*. Oxford University Press, 1991. [30](#), [34](#), [35](#), [80](#)
- [Sauvée 2008] M. Sauvée, P. Poignet et E. Dombre. *Ultrasound image-based visual servoing of a surgical instrument through nonlinear model predictive control*. Int. Journal of Robotic Research, vol. 27, pages 25–40, January 2008. [42](#)
- [Slabaugh 2006] G. Slabaugh, G. Unal, Tong Fang et M. Wels. *Ultrasound-specific segmentation via decorrelation and statistical region-based active contours*. In IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, volume 1, pages 45–53, 17-22 2006. [68](#)
- [Stewart 1991] J. Stewart. *Calculus, 2nd ed.* Pacific Grove, CA : Brooks/Cole, 1991. [55](#)
- [Stoll 2006] J.A. Stoll, P.M. Novotny, R.D. Howe et P.E. Dupont. *Real-time 3D ultrasound-based servoing of a surgical instrument*. In IEEE Int. Conf. on Robotics and Automation, ICRA'06, 2006. [43](#)
- [Subasic 2003] M. Subasic, S. Loncaric et E. Sorantin. *Region-based deformable model for aortic wall segmentation*. In Proceedings of the 3rd Int. Symposium on Image and Signal Processing and Analysis, volume 2, pages 731–735 Vol.2, 2003. [68](#)
- [Tahri 2005] O. Tahri et F. Chaumette. *Point-based and region-based image moments for visual servoing of planar objects*. IEEE Trans. on Robotics, vol. 21, no. 6, pages 1116–1127, December 2005. [36](#), [55](#), [60](#)
- [Taylor 1994] R.H. Taylor, B.D. Mittelstadt, H.A. Paul, W. Hanson, P. Kazanzides, J.F. Zuhars, B. Williamson, B.L. Musits, E. Glassman et W.L. Bargar. *An image-directed robotic system for precise orthopaedic surgery*. IEEE Trans. on Robotics and Automation, vol. 10, no. 3, pages 261–275, jun 1994. [11](#)
- [Uecker 1995] D.R. Uecker, C. Lee, Y.F. Wang et Y. Wang. *Automated instrument tracking in robotically assisted laparoscopic surgery*. Journal of Image Guided Surgery, vol. 1, no. 6, pages 308–325, 1995. [13](#)
- [Vilchis 2003] A. Vilchis, J. Troccaz, P. Cinquin, K. Masuda et F. Pellissier. *A new robot architecture for tele-echography*. IEEE Trans. on Robotics and Automation, vol. 19, no. 5, pages 922–926, october 2003. [11](#), [76](#)
- [Vitrani 2005] M.A. Vitrani, G. Morel et T. Ortmaier. *Automatic guidance of a surgical instrument with ultrasound based visual servoing*. In IEEE Int. Conf. on Robotics and Automation, ICRA'05, pages 508–513, 2005. [41](#)
- [Vitrani 2007] M.A. Vitrani, H. Mitterhofer, G. Morel et N. Bonnet. *Robust ultrasound-based visual servoing for beating heart intracardiac surgery*. In IEEE Int. Conf. on Robotics and Automation, ICRA'07, pages 3021–3027, Rome, Italie, 2007. [42](#)

- [Vitrani 2008] M.A. Vitrani et G. Morel. *Hand-eye self-calibration of an ultrasound image-based robotic system*. In IEEE/RSJ IROS'08 International Conference on Intelligent Robots and Systems, pages 1179–1185, Nice, France, September 2008. [42](#)
- [Wei 1997] G.K. Wei, K. Arbter et G. Hirzinger. *Real-time visual servoing for laparoscopic surgery*. IEEE Trans. in Engineering in Medecine and Biology, vol. 16, no. 1, pages 40–45, 1997. [13](#)
- [Wilson 1996] W. Wilson, C. Hulls et G. Bell. *Relative end-effector control using cartesian position-based visual servoing*. IEEE Trans. on Robotics and Automation, vol. 12, no. 5, pages 684–696, October 1996. [35](#)
- [Wyawahare 2009] M.V. Wyawahare, P.M. Patil et H.K. Abhyankar. *Image registration techniques : An overview*. Int. Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 2, no. 3, pages 11–28, 2009. [12](#)
- [Xu 1998] C. Xu et L. Prince. *Generalized gradient vector flow external forces for active contours*. Signal Processing, vol. 71, pages 131–139, 1998. [72](#)

Résumé : Ces travaux portent sur l'étude et l'élaboration de méthodes d'asservissement visuel utilisant les images échographiques. Dans ce contexte, très peu de travaux avaient été menés sur l'utilisation de l'image fournie par un capteur échographique et les méthodes existantes permettaient uniquement de contrôler les mouvements dans le plan de coupe du capteur. En effet, il est important de noter qu'une sonde échographique 2D a la particularité de fournir une information complète dans son plan d'observation mais de ne fournir aucune information en dehors de ce plan. A la différence, une caméra fournit une projection de la scène 3D vers une image 2D. De ce fait, les méthodes d'asservissement visuel basées sur la modélisation de l'interaction entre une caméra et son environnement ne peuvent être appliquées directement à la modalité échographique. Il faut également noter qu'une problématique importante est l'extraction en temps réel des informations visuelles, nécessaires à la commande d'un système robotique, à partir des images échographiques qui sont par nature très fortement bruitées. Nous avons par conséquent apporté des solutions génériques pour permettre le contrôle complet des déplacements d'une sonde embarquée sur un robot à 6 degrés de liberté en utilisant directement les images échographiques. Deux démarches principales ont été menées pour atteindre cet objectif. La première a porté sur le choix et la modélisation exacte des informations visuelles géométriques qu'il est possible d'extraire en temps réel à partir des images échographiques 2D et qui sont pertinentes pour la réalisation d'une tâche de positionnement d'une sonde robotisée. Nous avons plus particulièrement considéré des informations géométriques de type points, contours et moments de la section d'un organe d'intérêt. Des modèles simplifiés ou estimés en ligne de la forme des objets observés ont été considérés pour déterminer la variation des informations géométriques introduite par les mouvements effectués en dehors du plan de coupe de la sonde. Cette modélisation a permis de déterminer la loi de commande à appliquer au robot porteur de sonde pour atteindre automatiquement la coupe échographique présentant l'information visuelle désirée. La seconde démarche a porté sur l'exploitation de l'information dense de l'image échographique en vue de s'affranchir de l'étape de segmentation. Dans une première approche, l'information de corrélation de la texture de type « speckle » présente dans l'image échographique a été considérée pour réaliser la poursuite de tissus en mouvement par la sonde échographique. Une approche alternative a également été proposée où les informations visuelles choisies en entrée de la commande du système correspondent directement à la valeur d'intensité d'un ensemble de pixels de l'image. La variation de ces informations en fonction du mouvement de la sonde a été modélisée afin de mettre en œuvre des asservissements visuels capables de réaliser des tâches de positionnement ou de suivi de coupes anatomiques. Cette méthode a en outre été étendue pour différents types de capteurs ultrasonores (2D, 3D, bi-plans). Les applications qui découlent de ces travaux se situent principalement dans le domaine de l'assistance à l'examen échographique. Elles concernent d'une part le positionnement automatique de la sonde sur une section anatomique désirée. D'autre part, les applications traitées portent également sur la stabilisation active de l'image échographique. A cette fin, différentes approches ont été mises en œuvre pour compenser le mouvement de tissus mous en synchronisant les déplacements de la sonde par asservissement visuel échographique.

Mots clés : Asservissement visuel, imagerie échographique, robotique médicale, vision par ordinateur