



TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Informatik
Lehrstuhl für Informatikanwendung in der Medizin & Augmented Reality

Doppelpromotion / Joint doctorate – Université de Rennes 1 / TUM

Quality-Driven Control of a Robotized Ultrasound Probe

Pierre Chatelain

Vollständiger Abdruck der von der promotionsführenden Einrichtung Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende/-r: Prof. Dr. Philippe Poignet

Prüfende/-r der Dissertation:

1. Dr. Alexandre Krupa
2. Prof. Dr. Nassir Navab
3. Prof. Dr. Purang Abolmaesumi
4. Dr. Céline Fouard

Die Dissertation wurde am 10.10.2016 bei der Technischen Universität München eingereicht und durch die promotionsführende Einrichtung Fakultät für Informatik angenommen.

À mes parents

Acknowledgements

I am deeply thankful to my supervisors Alexandre Krupa and Nassir Navab for the invaluable support they have provided me throughout the preparation of my thesis, and for the incredible autonomy they have given me in the conduct of my research. Their friendly and encouraging guidance has been crucial at all stages of my doctoral studies.

My most sincere gratitude goes to Philippe Poinet and Purang Abolmaesumi for reviewing my manuscript, and also to Céline Fouard for taking part to my thesis committee. I have truly appreciated the insightful feedback the committee has provided on my work.

During the preparation of this thesis, I split time between IRISA at the University of Rennes 1 and the Chair for Computer-Aided Medical Procedures at the Technical University of Munich, and I have been lucky to meet exceptional people in both places. I would first like to warmly thank all members of the Lagadic team at IRISA. In particular François Chaumette, who first welcomed me in the team in 2010, Fabien for his regular support with the robotics platform and software, Céline and Hélène for their constant help with administrative matters. And of course a big thank you to Aurélien, Mani, Bertrand, François Pasteau, Le, Riccardo, Lucas, Aly, Vishnu, Suman, Nicolas, Fabrizio, Bryan, Pedro, Noël, Lesley-Ann, Quentin, Jason, Firas, Souriya, Thomas, Marc, Giovanni, Don Joven, for so many good memories.

I am grateful to my friends and colleagues at CAMP as well. More especially to the IFL group, and in particular to Oliver, Jack, Marco, and Rüdiger for their precious help with some experiments. To Loïc, for countless constructive (and also less constructive) discussions, and Ben, for exciting competitions at various precision sports. And also to Olivier, Ahmad, Diana, Vasilis, Sebastian, Ralf, Pascal, Christian, Séverine, Christoph, Nicola, Philipp, for some very nice moments.

Je tiens également à remercier certaines personnes qui ont contribué de manière significative — volontairement ou non — à mon parcours jusqu'à cette thèse. M^{me} Martin, pour une introduction très précoce à la robotique. M^{me} Osouf, pour avoir stimulé ma passion pour les sciences.

ACKNOWLEDGEMENTS

Et plus récemment Stéphane Legros, Luc Bougé, François Chaumette, et Nicolas Padoy.

Pour leur présence et leur soutien, je remercie très chaleureusement Guillaume, Lucile, Paulin, Anaïs, Marguerite, Paul, Antoine, Boubou, Maxence, les Poukis, M. Rogler, Guigui, et Nico.

Enfin, je suis infiniment reconnaissant à mes parents, à Clément, Laïs, et tout particulièrement à Anne pour m'avoir soutenu (et supporté. . .) pendant cette période.

Résumé en Français

L'imagerie médicale a été l'une des plus importantes révolutions dans le domaine de la santé. En effet, la possibilité de visualiser l'intérieur du corps humain sans recourir à une opération chirurgicale a permis un changement de paradigme pour le diagnostic médical. D'une part, parce que la chirurgie invasive, nécessaire jusqu'à la fin du dix-neuvième siècle pour voir et analyser le fonctionnement interne du corps humain, produit de larges plaies qui peuvent être douloureuses et se résorbent lentement. A contrario, l'imagerie médicale permet d'effectuer un diagnostic non invasif, ainsi que des procédures chirurgicales minimalement invasives. D'autre part, l'imagerie médicale permet de visualiser les organes internes dans leur état normal de fonctionnement. L'avènement de l'imagerie médicale remonte à la découverte des rayons X par Wilhelm Röntgen [Röntgen, 1898]. Depuis, l'imagerie médicale a connu un développement croissant, avec notamment l'invention de la tomодensitométrie (TDM) par Allan McLeod Cormack et Godfrey Hounsfield [Hounsfield, 1973], et de l'imagerie par résonance magnétique par Paul Lauterbur et Peter Mansfield [Lauterbur, 1973; Mansfield, 1977].

L'échographie, dont le principe repose sur la propagation d'ondes ultrasonores dans le corps, a été développée à la même période. Contrairement aux autres modalités d'imagerie, l'échographie a la capacité d'imager en temps réel. Pour cette raison, elle est la modalité idéale pour observer des organes en mouvement. Elle est en particulier utilisée pour l'imagerie du coeur, et pour l'imagerie per-opératoire. L'échographie est également inoffensive, à la différence de l'imagerie par rayons X qui produit des radiations ionisantes [Frush, 2004]. De plus, les équipements d'imagerie échographique sont compacts et peu coûteux, en comparaison à un scanner TDM ou IRM. Cependant, l'utilisation clinique de l'échographie s'est développée relativement lentement par rapport à la radiologie. Ceci s'explique en partie par la qualité moyenne des images échographiques, ainsi que par le champ visuel limité des sondes échographiques. Récemment, des améliorations techniques dans la conception des transducteurs piézoélectriques ont permis d'améliorer significative-

ment la qualité des images échographiques. D'autre part, de plus en plus d'importance est attachée à la rentabilité des procédures médicales. Il en résulte que l'échographie devient progressivement recommandée comme modalité d'imagerie de premier recours pour de nombreuses applications. Toutefois, les évaluations cliniques fondées sur l'imagerie échographique manuelle sont sujettes à une forte variabilité inter- et intra-expert. De plus, la manipulation fréquente de la sonde est une source de trouble musculo-squelettiques pour les échographistes. L'échographie robotisée est une technologie émergente qui permettrait de palier à ces problèmes, en assistant le geste de l'échographiste, tout en améliorant la fiabilité et la répétabilité des examens et des procédures chirurgicales sous imagerie échographique.

En dépit des bénéfices potentiels que la robotique médicale pourrait apporter pour de nombreuses procédures médicales, l'introduction de systèmes robotiques dans le domaine clinique a été jusqu'à présent très limitée. En effet, les procédures médicales, et en particulier les interventions chirurgicales, sont des environnements infiniment complexes, qui requièrent une interaction sécurisée et transparente entre le robot, le personnel médical, et le patient. Navab et al. [2016] retiennent, pour décrire les critères nécessaires au développement de systèmes d'imagerie per-opératoire, la pertinence, la vitesse, la flexibilité, la facilité d'utilisation, la fiabilité, la reproductibilité, et la sécurité. Tous ces critères doivent être réunis afin de pouvoir intégrer un système robotique dans la salle d'opération. En particulier, le but d'un robot médical ne doit pas être de remplacer le chirurgien, mais plutôt de faire office d'outil pour l'assister durant la procédure. Il est donc important de considérer un robot médical comme un composant d'un système plus large de chirurgie assistée par ordinateur [Taylor and Stoianovici, 2003].

En raison de sa capacité d'imagerie en temps réel, l'échographie est la modalité d'imagerie la plus indiquée pour les procédures médicales assistées par un robot. En effet, l'information fournie en temps réel permet d'atteindre la flexibilité et la réactivité nécessaires à la réalisation de procédures personnalisées. La flexibilité est cruciale pour l'imagerie per-opératoire, afin d'ajuster la procédure aux spécificités de chaque patient. La réactivité doit permettre au système de s'adapter de manière dynamique aux commandes du chirurgien, ainsi qu'à la situation courante de la procédure. À cette fin, l'échographie per-opératoire est capable de fournir sur le patient des informations précieuses, qui permettent d'adapter le comportement du système en prenant en compte, par exemple, les mouvements du patient ou la déformation des organes. Ceci est impossible pour les procédures utilisant uniquement une image pré-opératoire, puisque le chirurgien ne peut pas observer l'état courant des tissus in-

ternes. Les systèmes robotisés guidés par échographie sont donc une technologie prometteuse pour un large champ d'applications cliniques.

Motivations cliniques

Diagnostic échographique télé-opéré

L'idée de la télé-échographie a émergé à la fin du vingtième siècle comme une solution permettant à un expert d'effectuer un diagnostic clinique à distance [Sublett et al., 1995; Coleman et al., 1996]. Le développement de la télé-médecine a deux motivations principales. Premièrement, l'accessibilité aux soins médicaux est très inégalement répartie géographiquement. Les spécialistes sont surtout présents dans les régions fortement peuplées, quand les zones rurales et les régions isolées ont seulement accès à un service médical minimal, voire inexistant. La télé-médecine peut donc être considérée comme une solution à bas coût pour permettre l'accès à des soins médicaux de haute qualité. En second lieu, les visites médicales à effectuer dans un hôpital éloigné peuvent s'avérer épuisantes pour les patients. Par exemple, les patients sous hémodialyse doivent être régulièrement contrôlés afin de détecter une éventuelle arthropathie amyloïde. Dans ce contexte, la télé-échographie permettrait d'obtenir une expertise médicale dans un centre médical local avec des ressources limitées, voire à domicile. La télé-échographie consistait initialement en un système de visio-conférence, où une communication vidéo permettait à un expert d'effectuer un diagnostic à distance, avec l'aide d'un technicien sur site [Kontaxakis et al., 2000]. À l'aide d'un système robotique télé-opéré, l'expert peut également manipuler la sonde échographique lui-même [Vilchis et al., 2003; Arbeille et al., 2003; Courreges et al., 2004]. La télé-échographie robotisée a notamment été proposée pour l'examen des artères [Pierrot et al., 1999; Abolmaesumi et al., 2002], en obstétrique [Arbeille et al., 2005], et en échocardiographie [Boman et al., 2009].

Imagerie interventionnelle

Biopsie de la prostate

Le cancer de la prostate est le second type de cancer le plus répandu chez l'homme dans le monde, et le plus fréquent en occident [Brody, 2015]. Lorsqu'il y a suspicion, le diagnostic du cancer de la prostate nécessite le plus souvent d'effectuer une biopsie. La biopsie de la prostate est généralement guidée par échographie transrectale (ETR) afin de permettre un positionnement précis de l'aiguille de biopsie. Cependant, la qualité des

biopsies réalisées sous ETR manuelle est dépendante de l'expertise du médecin. L'utilisation d'un système robotisé présente plusieurs avantages [Kaye et al., 2014]. Par exemple, l'aiguille peut être alignée automatiquement avec la cible afin d'obtenir une meilleure précision du placement. Il a été démontré que la biopsie sous ETR avec assistance robotique permet d'améliorer le taux de détection du cancer, par rapport à la biopsie manuelle sous ETR [Han et al., 2012; Vitrani et al., 2016].

Curiethérapie

La curiethérapie (ou brachythérapie) consiste à implanter des sources radioactives à des emplacements spécifiques pour le traitement de certains cancers. Tout comme la biopsie, la curiethérapie est le plus souvent réalisée sous ETR. Dans ce contexte, un système robotisé permet de faciliter le placement de la sonde échographique, de suivre la cible et l'aiguille, et d'assister le médecin dans le positionnement de l'aiguille [Wei et al., 2005; Fichtinger et al., 2008].

Prostatectomie

Des systèmes robotiques sont également utilisés pour la prostatectomie, qui est une alternative à la curiethérapie consistant en une ablation chirurgicale de la prostate. L'imagerie ETR permet de visualiser les faisceaux nerveux qui doivent être préservés [Long et al., 2012; Hung et al., 2012].

Biopsie du cancer du sein

Le cancer du sein, provoquant plus de 500 000 décès chaque année dans le monde [Stewart and Wild, 2014], est une des premières causes de mortalité liée à un cancer chez les femmes. La procédure standard pour le diagnostic du cancer du sein est une biopsie, qui est généralement guidée par échographie afin de permettre une visualisation de la cible en temps réel. Ceci permet de détecter un éventuel déplacement de la cible dû à l'insertion de l'aiguille. Un système robotique peut permettre de faciliter le positionnement de l'aiguille [Kobayashi et al., 2012] ou la stabilisation des tissus [Mallapragada et al., 2011; Wojcinski et al., 2011].

Ultrasons focalisés de haute intensité

Les ultrasons focalisés de haute intensité (HIFU, de l'anglais *high intensity focused ultrasound*) sont une méthode d'ablation tumorale reposant

sur la génération d'ondes ultrasonores focalisées, qui provoquent une lésion thermique localisée. La thérapie HIFU est une technologie prometteuse, qui a l'avantage d'être minimalement invasive [Tempany et al., 2011]. L'assistance robotique à la thérapie HIFU a été proposée pour la première fois afin de permettre un placement précis du transducteur HIFU pour la neurochirurgie [Davies et al., 1998]. Cependant, un traitement HIFU planifié uniquement à partir d'une image pré-opératoire a une précision limitée, en raison des potentielles erreurs de recalage, et des déplacements de tissus. Plus récemment a été proposé un traitement HIFU avec assistance robotique en boucle fermée, permettant d'effectuer une compensation des mouvements par asservissement visuel échographique [Seo et al., 2011; Chanel et al., 2015].

Contributions

La qualité des images échographiques dépend de plusieurs facteurs, qui ne sont pas toujours contrôlés. En particulier, la qualité des images dépend du couplage acoustique entre la sonde et la peau du patient. C'est pourquoi l'utilisation de gel acoustique et la force de contact avec le patient sont des facteurs cruciaux pour obtenir une qualité d'image satisfaisante. De plus, la position et l'orientation de la sonde par rapport aux tissus influencent la qualité du signal acoustique, car l'intensité de l'écho ultrasonore dépend du chemin parcouru par l'onde pour atteindre un certain point et être réfléchi jusqu'au transducteur. L'amplitude du signal acoustique peut décroître fortement en présence d'importantes différences d'impédance acoustique, par exemple aux interfaces tissu/os ou tissu/gaz. Il en résulte que la qualité du signal ultrasonore peut être très hétérogène au sein d'une même image. Il est donc important de trouver une bonne fenêtre acoustique pour avoir une image nette de l'anatomie ciblée. Pour l'asservissement visuel échographique, l'influence du positionnement de la sonde sur la qualité de l'image est habituellement ignorée, et la visibilité de la cible n'est pas garantie.

Les travaux présentés dans ce manuscrit traitent spécifiquement du contrôle de la qualité de l'image échographique. La qualité du signal ultrasonore, représentée par une carte de confiance, est utilisée comme signal sensoriel pour asservir une sonde échographique robotisée, en vue d'optimiser son positionnement. Les principales contributions de cette thèse sont :

- Une méthode de suivi d'aiguille de biopsie flexible dans les images échographiques 3D.

- Une nouvelle méthode d'estimation de la qualité du signal acoustique pour les images échographiques. Cette méthode, qui repose sur une intégration du signal ultrasonore le long des lignes de tir, fournit une estimation en temps réel et pixel par pixel de la qualité d'image.
- Une comparaison du temps de calcul et de la régularité entre cette nouvelle méthode d'estimation de la qualité et une approche existante.
- Une étude de la relation entre la position de la sonde et la distribution de la qualité du signal acoustique au sein de l'image.
- L'utilisation de la carte de confiance échographique comme modalité visuelle pour réaliser une commande en boucle fermée d'une sonde robotisée. Nous proposons des lois de commandes permettant d'optimiser la qualité d'image, soit globalement, soit par rapport à une cible anatomique spécifique.
- Une commande hybride, combinant la commande guidée par qualité avec d'autres tâches, telles que la commande en effort et le centrage d'une cible dans l'image.
- Une validation des méthodes proposées par des expériences illustrant différents scénarios. En particulier, nous présentons des résultats expérimentaux obtenus sur un volontaire humain.

Nous détaillons à présent le contenu de ce manuscrit, chapitre par chapitre.

Chapitre 1

Ce premier chapitre est consacré au traitement et à l'analyse des images échographiques. Nous proposons tout d'abord une introduction au principe de l'imagerie échographique. Puis, nous présentons un état de l'art des méthodes d'analyse des images échographiques. Une attention particulière est apportée aux méthodes de suivi en temps réel de tissus mous et d'instruments chirurgicaux. Nous proposons également une première contribution portant sur le suivi d'aiguille flexible dans les images échographiques 3D. Enfin, nous introduisons différentes méthodes d'estimation de la qualité des images échographiques.

Chapitre 2

Dans ce chapitre, nous abordons le sujet de la commande robotique guidée par échographie, et plus particulièrement de l’asservissement visuel échographique. Après une introduction générale à la commande par asservissement visuel, nous proposons un état de l’art des méthodes de commande d’une sonde échographique, puis des méthodes d’asservissement d’un instrument chirurgical sous imagerie échographique. Nous présentons également une première contribution personnelle sur ce sujet, qui consiste en un asservissement visuel par échographie tridimensionnelle pour le guidage d’une aiguille de biopsie flexible.

Chapitre 3

Dans ce chapitre, nous présentons la contribution principale de cette thèse. En partant des notions introduites dans les deux précédents chapitres, nous proposons un modèle de l’interaction entre le positionnement d’une sonde échographique et la qualité des images. Nous utilisons ensuite ce modèle pour définir des primitives visuelles adaptées à la conception de lois de commandes, et nous proposons deux stratégies de commande d’une sonde portée par un bras robotique. Une première méthode permet d’asservir la sonde de manière à optimiser globalement la qualité d’image. Des degrés de liberté supplémentaires sont alors disponibles pour téléopérer la sonde. Dans une seconde méthode, nous considérons une cible anatomique spécifique. Nous proposons alors une loi de commande permettant de réaliser simultanément le centrage de la cible dans l’image (par un déplacement de la sonde), et l’optimisation de la qualité d’image pour cette cible.

Chapitre 4

Le dernier chapitre est consacré à la validation expérimentale des méthodes proposées dans le chapitre 3. Nous reportons les résultats d’une analyse de la convergence et de la réaction aux perturbations de notre système. Nous proposons ensuite une illustration expérimentale de l’utilisation de notre méthode pour la télé-échographie, avec une commande partagée entre l’ordinateur et un utilisateur. En particulier, nous présentons une validation expérimentale de notre système sur un sujet humain volontaire.

Contents

Acknowledgements	iii
Résumé en Français	v
Introduction	1
1 Ultrasound Image Analysis	9
1.1 Ultrasound imaging	10
1.1.1 Piezoelectricity	10
1.1.2 Ultrasound image formation	10
1.1.2.1 Ultrasound propagation	10
1.1.2.2 Radio frequency ultrasound	12
1.1.2.3 B-mode ultrasound	15
1.1.2.4 3D ultrasound	19
1.2 Soft tissue tracking	21
1.2.1 Motion tracking	21
1.2.1.1 Block matching	21
1.2.1.2 Deformable block matching	25
1.2.1.3 Elastic registration	26
1.2.2 Deformable shape models	26
1.2.2.1 Active contour	26
1.2.2.2 Active shape models	28
1.3 Instrument tracking	28
1.3.1 Hardware approaches	29
1.3.2 Ultrasound-based tracking	31
1.3.2.1 Curve fitting	33
1.3.2.2 Random sample consensus	35
1.4 Needle tracking via particle filtering	39
1.4.1 Bayesian tracking	40
1.4.2 Needle dynamics	40
1.4.3 Particle filtering	41
1.4.3.1 Appearance model	42

1.4.3.2	Sequential importance re-sampling . . .	42
1.5	Quality estimation	44
1.5.1	Random walks confidence map	48
1.5.1.1	Graphical representation	48
1.5.1.2	Random walks	49
1.5.2	Scan line integration	50
1.5.3	Comparison	53
1.5.3.1	Computation time	53
1.5.3.2	Temporal regularity	55
1.6	Conclusion	57
2	Ultrasound-Based Visual Servoing	59
2.1	Visual servoing	60
2.1.1	Introduction	60
2.1.2	Interaction matrix	61
2.1.3	Visual servo control	62
2.1.4	Hybrid tasks	64
2.2	Ultrasound probe control	65
2.2.1	Notations	68
2.2.2	Geometric visual servoing	68
2.2.2.1	Point-based in-plane control	68
2.2.2.2	3D point-based control	69
2.2.2.3	Moments-based control	70
2.2.3	Intensity-based visual servoing	72
2.2.3.1	Case of a 2D probe	74
2.2.3.2	Case of a 3D probe	75
2.3	Ultrasound-guided needle control	75
2.3.1	Needle steering	76
2.3.2	Path planning	79
2.3.3	Ultrasound-guided needle steering	80
2.3.3.1	2D probe, in-plane needle	80
2.3.3.2	2D probe, out-of-plane needle	81
2.3.3.3	3D probe	81
2.3.3.4	3D US-guided needle steering via visual servoing	81
2.4	Conclusion	86
3	Quality-Driven Visual Servoing	87
3.1	Dynamics of the confidence map	89
3.1.1	General considerations	89
3.1.1.1	Effect of the contact force	89
3.1.1.2	Effect of orientation	91

3.1.1.3	Approximated interaction matrix	91
3.1.2	Analytic solution for scan line integration	92
3.2	Geometric confidence features	95
3.2.1	2D Case	95
3.2.1.1	Definition	95
3.2.1.2	Experimental comparison	96
3.2.1.3	Feature dynamics	97
3.2.2	3D Case	100
3.3	Global confidence-driven control	101
3.3.1	Force control	102
3.3.2	Confidence control	106
3.3.3	Tertiary task	107
3.4	Target-specific confidence-driven control	108
3.4.1	Target centering	109
3.4.2	Confidence control	111
3.5	Conclusion	113
4	Experimental Results	115
4.1	Experimental setup	115
4.1.1	Equipment	116
4.1.1.1	Ultrasound systems	117
4.1.1.2	Robots	118
4.1.1.3	Phantoms	118
4.1.2	Implementation details	119
4.2	Convergence analysis	119
4.2.1	Global confidence control	120
4.2.2	Target-specific confidence control	122
4.3	Reaction to disturbances	126
4.3.1	Global confidence-driven control	126
4.3.2	Target-specific confidence-driven control	128
4.3.2.1	Evaluation of target tracking quality	131
4.4	Confidence-optimized tele-echography	133
4.4.1	Phantom experiment	133
4.4.2	In vivo experiment	136
4.5	Conclusion	139
	Conclusion	143
	List of Publications	151
	Abstracts of Additional Contributions	153

CONTENTS

Bibliography	157
List of Figures	179
Acronyms	183

Introduction

Medical imaging has been one of the most important revolutions in health care. Indeed, visualizing the interior of the body without the need to open it has enabled a change of paradigm for medical diagnosis. First, because until the end of the nineteenth century, seeing and analyzing the inside of the human body required to perform an open surgery. Open surgery leaves large wounds on the body, which can be painful, and heal slowly with a risk of infection. In this regard, medical imaging enables noninvasive diagnosis as well as minimally invasive procedures, where only a small incision is made in order to insert a surgical tool. Second, because medical imaging provides a way to visualize any part of the human body in its functional state. The advent of medical imaging dates back to the discovery of X-rays by Wilhelm Röntgen in 1895. Noticing the surprising properties of these rays capable of propagating through opaque media, Röntgen had the idea to experiment it on the human body [Röntgen, 1898]. One of the first pictures, now famous, produced with his system displays the left hand of his wife. Since then, medical imaging has been continuously improved, with the development of Computerized Tomography (CT) by Allan McLeod Cormack and Godfrey Hounsfield [Hounsfield, 1973], and Magnetic Resonance imaging (MRI) by Paul Lauterbur and Peter Mansfield [Lauterbur, 1973; Mansfield, 1977].

Medical ultrasound (US) imaging, based on the propagation of ultrasonic waves in the body, was also developed in the same period. In contrast with other imaging modalities, ultrasonography provides real-time imaging capabilities. Consequently, ultrasound is the modality of choice for imaging moving organs, such as in echocardiography, and for intraoperative imaging. Ultrasound is also harmless, unlike X-ray imaging, which emits ionizing radiations [Frush, 2004]. Moreover, ultrasound imaging devices are lightweight and inexpensive compared to a CT or MRI scanner. This is another advantage for its use in an operating room. However, the clinical use of ultrasound has developed quite slowly compared to radiology, and was, until recently, quite marginal. An explanation for this relatively poor use of ultrasound was the low quality of

the images it produces, together with the limited field of view it provides. With the recent technological improvements on the design of ultrasound transducers, the quality of ultrasound imaging has increased steadily. In addition, an ever greater importance is given to the cost effectiveness of medical procedures. For these reasons, ultrasound is now recommended as a frontline imaging modality for more and more applications. However, clinical assessment based on free-hand ultrasound imaging is subject to a large inter- and intraoperator variability. The frequent manual manipulation of an ultrasound probe is also a source of musculoskeletal disorders for sonographers. Consequently, robotized ultrasonography is emerging as a solution to assist the sonographer and to improve the reliability and repeatability of examinations and surgical procedures.

Despite the potential benefits that medical robotics can bring to a wide range of medical procedures, the introduction of robotic systems into the clinic has been quite limited, compared to the rapid development of robotics in the industrial sector. A reason for the relatively slow development of medical robots is the complexity of operative procedures, where the robot has to interact safely and transparently with the medical staff and the patient. [Navab et al., 2016] discuss the requirements for intraoperative imaging systems, and retain the criteria of relevance, speed, flexibility, usability, reliability, reproducibility, and safety. These are all critical requirements that have to be met for a robotic system to be accepted in the operating room. In particular, the purpose of a medical robot should not be to replace the surgeon. It should rather serve as a tool, among others, to assist him during the medical procedure. In this regard, it is important to integrate a medical robotic system as a component of a global Computer-Integrated Surgery (CIS) workflow [Taylor and Stoianovici, 2003].

Due to its real-time imaging capability, ultrasound is a modality of choice for robot-assisted medical procedures. Indeed, real-time ultrasound feedback can provide the system with the flexibility and reactivity required to perform personalized and adaptive procedures. Flexibility is crucial in intraoperative imaging in order to adjust the procedure to the characteristics of each patient. Reactivity should allow the system to adapt dynamically to the surgeon's commands, but also to the current state of the process. In this regard, intraoperative ultrasound can provide precious information on the patient. It can be used to adapt the process by taking into account, for instance, patient motion or organ deformation. This contrasts with procedures based on preoperative imaging alone, where the surgeon cannot observe the current state of the body. Thus, robot-assisted systems with ultrasound guidance are promising for a wide range of clinical applications.

Clinical motivations

Teleoperated diagnostic ultrasound

The idea of tele-echography (or teleultrasound) has emerged at the end of the twentieth century as a solution to enable remote expert clinical diagnosis [Sublett et al., 1995; Coleman et al., 1996]. The motivation behind the development of telemedicine is twofold. First, the accessibility to quality health care is unequally distributed. Specialists tend to be located in highly populated areas, whereas rural and remote regions have only access to minimal, if any, medical care. Thus, telemedicine is a low cost solution to increase the accessibility to high-standard medical care. Second, regular visits to remote hospitals can be exhausting for patients. For instance, patients under long-term hemodialysis need to be checked regularly for amyloid arthropathy. In this context, tele-echography has the potential to provide expert diagnosis in a local hospital with limited resources, or even at home. Tele-echography initially relied on a video conferencing system. Video communication allowed an expert to perform a diagnosis remotely, with the help of a technician on site [Kontaxakis et al., 2000]. With a teleoperated robotic ultrasound system, the expert clinician can manipulate the ultrasound probe himself [Vilchis et al., 2003; Arbeille et al., 2003; Courreges et al., 2004]. Teleoperated ultrasound has been considered for the examination of arteries [Pierrot et al., 1999; Abolmaesumi et al., 2002], in obstetrics [Arbeille et al., 2005], and in echocardiography [Boman et al., 2009].

Interventional imaging

Prostate cancer biopsy

Prostate cancer is the second most frequent cancer in men in the world, and the most frequent in occidental countries [Brody, 2015]. When suspected, prostate cancer is typically diagnosed by performing a biopsy, which consists in the insertion of a needle to extract tissue samples to be analyzed. Prostate cancer biopsy is most commonly conducted under Transrectal Ultrasound (TRUS) guidance to allow for a precise positioning of the needle. However, the quality of the biopsy under freehand TRUS is dependent on the clinician's expertise. The use of a robotic system in this context has several advantages, which are discussed, e.g., in [Kaye et al., 2014]. For instance, the needle can be automatically aligned with the target to allow for a greater precision, by accounting for tissue displacement. It has been demonstrated that robot-assisted

TRUS biopsy can improve the cancer detection rate, compared to free-hand TRUS biopsies [Han et al., 2012].

Brachytherapy

Brachytherapy consists in the implantation of radioactive seeds at a precise location for the treatment of cancer. As for biopsies, brachytherapy is most frequently performed under TRUS guidance. Therefore, robotic systems can assist in the placement of the probe, tracking of the target and needle, as well as needle placement [Wei et al., 2005; Fichtinger et al., 2008].

Laparoscopic prostatectomy

Robotic systems are also used for laparoscopic prostatectomy (prostate removal). In this context, intraoperative TRUS imaging is required in order to visualize neurovascular bundles that have to be preserved [Long et al., 2012; Hung et al., 2012].

Breast cancer biopsy

Breast cancer is one of the first causes of cancer-related death in women, with about 500 000 deaths each year worldwide [Stewart and Wild, 2014]. Breast biopsy is the standard protocol for cancer diagnosis. Ultrasound guidance is widely used for breast biopsy, because it provides a real-time visualization of the target, which can be displaced due to the insertion of the needle. However, freehand ultrasound-guided biopsy is strongly dependent on the expertise of the clinician. Therefore, a robotic system can assist in breast biopsy for needle placement [Kobayashi et al., 2012] or tissue stabilization [Mallapragada et al., 2011; Wojcinski et al., 2011].

High intensity focused ultrasound

High Intensity Focused Ultrasound (HIFU) is an emerging method for tumor ablation. Based on the generation of focalized ultrasound to generate a local thermal lesion, HIFU therapy is a promising technology, because it is minimally invasive [Tempany et al., 2011]. Robotic HIFU was proposed early for the placement of the robotic transducer in neurosurgery [Davies et al., 1998]. However, HIFU therapy based on pre-operative planning has a limited precision, due to registration errors and tissue displacements. Recently, closed-loop robot-assisted HIFU was considered to allow for motion compensation using ultrasound-based visual servoing [Seo et al., 2011; Chanel et al., 2015].

Challenges

Several challenges are associated to the design of robotic ultrasound systems. We detail thereafter the main challenges in terms of image analysis, real-time requirements, modeling, and image quality.

Image analysis First, a reactive robotic ultrasound system should be able to exploit the information available in the ultrasound images. The noisy nature of ultrasound images can make this task extremely difficult, compared to other imaging modalities. As a result, the analysis of ultrasound images, which comprises segmentation, tracking, registration, and feature extraction, is a whole field of research by itself.

Speed Robotized ultrasound brings an additional constraint to the design of image processing algorithms, in terms of real-time requirements. Indeed, the processing of the images should be fast enough to enable a close-loop control of the robot. In the image processing literature, the term *real-time* is often used loosely, to indicate that a process is fast enough to allow interaction with a human. It is important, however, to always look at the targeted application before to consider a process as real-time. An acceptable definition would be that a real-time system is able to process information at the same rate as the said information is produced. In ultrasound-guided robotics, a real-time system should also allow a reactive control of the robot. As a result, while many elaborated ultrasound image analysis methods are available, the priority in ultrasound-guided robotics is the speed and the robustness of the algorithms.

Interaction modeling Controlling a robot under ultrasound guidance requires a model of the interaction between the motion of the robot and the image contents. While efficient solutions exist for image-guided control, and, in particular, visual servoing, the characteristics of ultrasound imaging makes the modeling of interactions more difficult. Specific challenges are the noise of the images, tissue deformations, but also the fact that 2D ultrasound probes only provide information in their observation plane.

Image quality Finally, the quality of ultrasound images depends on several factors, which are not always controlled. In particular, ultrasound image quality depends on the acoustic coupling between the probe and the patient's skin. For this reason, the use of acoustic gel and the contact

force with the body are crucial factors of image quality. Moreover, the position and orientation of the probe with respect to the tissues influences the quality of the ultrasound signal. Indeed, the intensity of the ultrasound echo depends on the path traveled by the wave to reach a certain point and to propagate back to the transducer. The amplitude of the sound signal can decrease drastically when there are strong changes in acoustic impedance, such as at tissue/bone or tissue/gas interfaces. Consequently, the ultrasound signal quality can be highly heterogeneous within the image, and it is important to position the probe on a good acoustic window in order to have a clear image of the target anatomy. In ultrasound-based visual servoing, the impact of probe positioning on the image quality is usually ignored. As a result, the visibility of the target is not guaranteed.

Context and objectives

This thesis was conducted in a co-supervision scheme between the Lagadic team at IRISA, Université de Rennes 1 and the CAMP group at the Technische Universität München. Starting from the observation that ultrasound image quality has received little attention, the purpose of this thesis was to investigate the possibility to optimize the quality of ultrasound acquisitions via a dedicated control strategy. Thus, the aim of this thesis was to answer the questions:

- *How can one represent the quality of the ultrasound signal?*
- *How can one control the position of an ultrasound probe so as to maximize the quality of acquired images?*

Contributions

The main contributions of this thesis are:

- A method for tracking a flexible biopsy needle in 3D ultrasound images.
- A new method for estimating the quality of the acoustic signal in ultrasound images. This method, based on an integration of the ultrasound signal along the scan lines, provides a real-time pixel-wise estimation of image quality.

- A comparison between this new quality estimation method and an existing method based on the random walks algorithm, in terms of computational cost and regularity.
- A study of the relation between the position of an ultrasound probe and the distribution of the ultrasound signal quality within the image.
- The use of the ultrasound confidence map as a visual modality to perform a closed-loop control of a robot-held probe. Control strategies are proposed for optimizing the image quality either globally, or with respect to a specific anatomic target. This contribution addresses the different challenges described above, in particular in terms of modeling and real-time processing requirements.
- A control fusion approach, to combine the proposed quality-driven control with other tasks, such as force control and target tracking.
- A validation of the proposed methods via experiments which illustrate different use cases.

The contributions on the topic of quality-driven control were partly published in two articles in the proceedings of the International conference on Robotics and Automation (ICRA) [Chatelain et al., 2015b, 2016]. The contribution on the tracking of a flexible biopsy needle under 3D ultrasound guidance was also published in an ICRA paper [Chatelain et al., 2015a].

During the preparation of this thesis, other contributions were made that are not discussed herein, because they are out of the topic of this dissertation. We provide the abstracts of the corresponding publications in the appendix.

Thesis outline

This manuscript is organized as follows.

Chapter 1 We present an overview of ultrasound image analysis techniques. The chapter starts with an introduction to ultrasound imaging, which provides a basic background on this imaging modality. Then, we present a review of ultrasound image analysis algorithms, focusing on soft tissue and instrument tracking methods with real-time capability. In this context, we provide a first personal contribution on the tracking

of a flexible needle in 3D ultrasound images. Finally, we introduce different methods for the estimation of ultrasound signal quality, and we propose a new algorithm that presents some advantages compared to the state of the art.

Chapter 2 We address the topic of ultrasound-guided robot control, with a particular focus on ultrasound-based visual servoing methods. After a general introduction to visual servoing, we provide a review of the state-of-the-art on (i) methods for controlling an ultrasound probe, and (ii) methods for controlling a surgical instrument, under ultrasound guidance. We also propose a personal contribution on this second topic, which consists in a 3D ultrasound-based visual servoing method for steering a flexible needle.

Chapter 3 We present the main contribution of this thesis. Based on the notions introduced in the two previous chapters, we propose a model of the interaction between the probe positioning and the quality of ultrasound images. Then, we use this model to design different control approaches aimed at optimizing the quality of ultrasound imaging.

Chapter 4 We provide experimental results which validate the framework introduced in chapter 3. We illustrate the application of our method to a teleoperation scenario, with a shared control between the automatic controller and a human operator. In particular, we provide the results of experiments performed with our framework on a human subject.

Conclusion Finally, we draw the conclusions of this dissertation, and we propose perspectives for further developments.

Chapter 1

Ultrasound Image Analysis

This chapter provides an overview of ultrasound image analysis methods. The automatic processing of ultrasound images, ranging from low-level signal processing to computer-aided diagnosis, provides ways to improve the image quality, to assist the physician in interpreting the images, and to extract quantitative clinical features that are valuable for diagnosis or treatment planning. While physicians are well trained in the interpretation of ultrasound images, computer-aided analysis is desirable to perform tedious and time-consuming tasks, such as segmentation or registration. In addition, modern ultrasound technologies such as three-dimensional ultrasound and elastography generate data that are more challenging to analyze than conventional B-mode ultrasound. The analysis of ultrasound images is also subject to inter- and intra-observer variability [Tong et al., 1998]. In this regard, automatic or semi-automatic image analysis can be a tool to help standardizing and improving the reliability of diagnosis and treatment planning. Moreover, automatically interpreting the contents of ultrasound images is necessary to the development of robot-assisted imaging and ultrasound-guided robotic surgery.

In the following, we provide an introduction to ultrasound imaging (section 1.1). Then, we present a review of real-time algorithms for soft tissue tracking (section 1.2) and needle tracking (section 1.3). On this topic, we propose a new solution for tracking flexible needles in 3D ultrasound via particle filtering (section 1.4). Finally, we address the issue of estimating the quality of ultrasound images (section 1.5).

1.1 Ultrasound imaging

1.1.1 Piezoelectricity

The development of ultrasound imaging was enabled by the discovery of the piezoelectric effect at the end of the 19th century. Piezoelectricity was first evidenced by the brothers Jacques and Pierre Curie, who observed a production of electricity in hemihedral crystals, induced by mechanical compression and decompression [Curie and Curie, 1880]. This phenomenon is analogous to another property of these crystals which was already well known at this time: the production of electricity by a change of temperature, or pyroelectricity. In 1881, Lippmann predicted, through his analytic formulation of the electricity conservation law, that the reverse effect should also occur. That is, a hemihedral crystal would deform under the action of electricity [Lippmann, 1881]. The reverse piezoelectric effect was confirmed experimentally in the same year by the Curie brothers [Curie and Curie, 1881].

One of the first notable applications of piezoelectricity was an ultrasonic underwater detector for submarines [Chilowsky and Langevin, 1916]. Therapeutic applications of ultrasound were investigated in the middle of the 20th century, with, for example, the work of [Bierman, 1954] on the treatment of scars. A first attempt of application to diagnostic ultrasound imaging was made by [Dussik, 1942]. Dussik proposed an imaging system based on the transmission of ultrasound between an emitter and a receiver, and presented an image resulting from the scan of a brain. However, the structures appearing in the image were, in fact, imaging artifacts due to reflections of the skull [Güttner et al., 1952]. The first successful ultrasound diagnosis experiments are due to Ludwig and Struthers, who reported the detection of gallstones in biological tissues [Ludwig and Struthers, 1949].

[Wild and Reid, 1952] combined a piezoelectric transducer with a mechanical scanning system in order to create a two-dimensional ultrasound image. This invention has led to what is now known as B-mode ultrasound imaging. B-mode imaging has been later improved by [Howry et al., 1955], who evidenced the presence of ultrasonic interfaces between different tissues.

1.1.2 Ultrasound image formation

1.1.2.1 Ultrasound propagation

Sound is a mechanical wave of pressure and displacement that propagates through a medium. The human ear is sensible to sound with frequencies

between 20 Hz and 20 kHz [Davis, 2007]. Ultrasound is a sound wave with frequency higher than 20 kHz, i.e., beyond the audible limit of humans. Conventional medical ultrasound systems operate in the range between 1 MHz and 20 MHz [Chan and Perlas, 2011].

The speed of sound depends on the properties of the medium it propagates through. It can be described by the Newton-Laplace equation [Biot, 1802]

$$c = \sqrt{\frac{K}{\rho}}, \quad (1.1)$$

where c is the speed of sound, K is the bulk modulus (a coefficient of stiffness) of the medium, and ρ is the density of the medium. Thus, the speed of sound increases with the stiffness of the tissues, and decreases with the density. The speed of sound is also related to the frequency f and the wavelength λ by

$$c = f\lambda, \quad (1.2)$$

so that the wavelength is inversely proportional to the frequency. Since the wavelength indicates the resolution at which adjacent objects can be distinguished, the axial resolution of ultrasound images is proportional to the frequency. For this reason, superficial structures are imaged at high frequencies (from 7 MHz to 20 MHz), while deeper structures are imaged at lower frequencies (from 1 MHz to 6 MHz) to allow for a greater penetration, at the cost of a lower resolution. The speed of sound in soft tissues is generally assumed to be constant at 1540 m s^{-1} . Thus, the wavelength for medical ultrasound lies between $77 \mu\text{m}$ (at 20 MHz) and 1.54 mm (at 1 MHz). Therefore, ultrasound imaging has the capability to distinguish structures at a submillimeter accuracy. This high resolution is, together with the real-time imaging capability, one of the main advantages of ultrasound, compared to other medical imaging modalities.

When traveling through the tissues, ultrasound is subject to reflections at tissue interfaces. Assuming that the ultrasound wavelength is small with respect to the structure, the interaction at a boundary between two media with different acoustic properties can be described with Snell's law, which links the incidence, reflection and transmission directions to the acoustic velocity in the two media. Let us consider a medium 1 with sound velocity c_1 , a medium 2 with sound velocity c_2 , and a sound wave hitting the interface with an incidence angle θ_i with respect to the normal to the interface (see Figure 1.1). When $\theta_i \neq 0$, the direction of incidence and the normal to the interface define a plane, which is referred to as the plane of incidence. Snell's law states that the sound wave is reflected (resp. transmitted) in the plane of incidence with an angle θ_r .

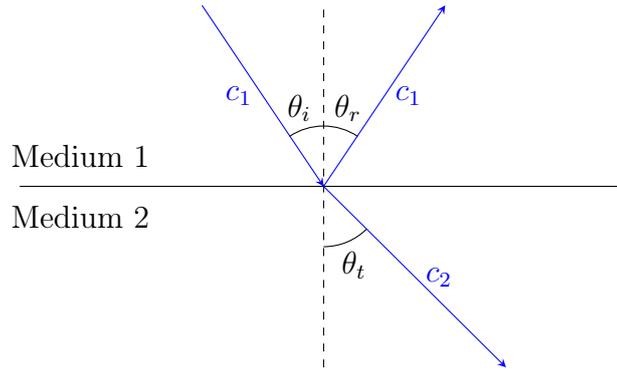


Figure 1.1 – Specular reflection and transmission of a sound wave at an interface between two media.

(resp. θ_t), such that

$$\frac{\sin \theta_i}{c_1} = \frac{\sin \theta_r}{c_1} = \frac{\sin \theta_t}{c_2}. \quad (1.3)$$

In particular, when an ultrasonic wave intersects a tissue boundary at normal incidence, part of the wave is reflected directly in the opposite direction. This is one of the most important phenomena underlying the formation of ultrasound images, as the reflected echo indicates the presence of a tissue boundary. At non-normal incidence angles, however, the ultrasound is reflected away from the source.

When the ultrasound wave encounters objects with a scale that is comparable to or smaller than its wavelength, the wave is reflected in all directions. This phenomenon is referred to as *scattering*, or diffuse reflection, as opposed to specular reflection. Scattering is responsible for the texture patterns characterizing ultrasound images. It is also one of the sources of energy loss for the sound wave.

Soft tissues also have a certain viscosity, which causes the conversion of acoustic energy to thermal or chemical energy. Therefore, part of the energy of the ultrasound wave is absorbed by soft tissues and converted to heat. Acoustic attenuation in soft tissues can be expressed as a frequency-dependent power law [Wells, 1975]

$$p(x + \delta x) = p(x)e^{-\alpha \delta x}, \quad (1.4)$$

where p is the sound pressure amplitude, x is the position, δx is the displacement, and α is a frequency-dependent attenuation coefficient.

1.1.2.2 Radio frequency ultrasound

Ultrasound transducers consist in an array of piezoelectric crystals, with typically 128 elements. The piezoelectric elements can be triggered by an

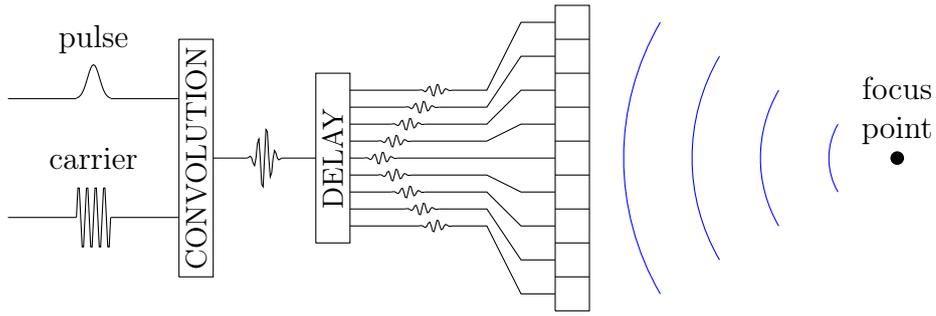


Figure 1.2 – Generation of a focused ultrasound beam. A Gaussian pulse is modulated by a sinusoidal carrier to produce a pulse wave. The transmission pulse is delayed and sent to trigger an array of piezoelectric crystals.

electric signal to generate an ultrasound wave by vibration of the crystal. In practice, a small group of adjacent elements is triggered synchronously, each with a specific time delay, in order to generate a focused ultrasound beam. The electric signal applied to the transducer elements consists in a sinusoidal wave convolved with a Gaussian modulator, in order to generate a pulse. This ultrasound beam generation process is illustrated in Figure 1.2.

Conversely, the transducer elements allow the reception of reflected ultrasound echoes. The returning ultrasound wave triggers the crystals, which, in turn, generate an electric signal. Thus, the acquisition of a scan line consists in a transmission phase, where an ultrasound pulse is generated, and a listening phase, where the returning signal is processed. The listening phase lasts for a predefined period of time, which depends on the desired imaging depth. Indeed, imaging at a depth d requires to listen for a time

$$T = \frac{2d}{c}, \quad (1.5)$$

where c is the speed of sound. For instance, the acquisition of a scan line at a depth $d = 10$ cm, assuming a speed of sound $c = 1540$ m.s⁻¹, requires to listen to the echo during $T \approx 0.13$ ms.

The signal received during the listening phase is then amplified, and compensated for attenuation by a *time gain compensation* function, i.e., a depth-dependent gain adjustment. The resulting signal is referred to as the *radio frequency* (RF) signal. An example of RF line measure is shown in Figure 1.3.

The acquired RF signal takes the form

$$u(t) = A(t) \cos(2\pi f_c t + \phi(t)), \quad (1.6)$$

where A is the amplitude, f_c is the carrier frequency, and ϕ is the phase

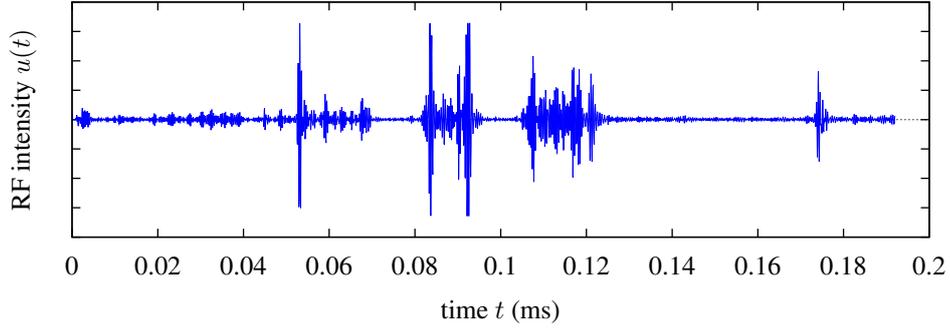


Figure 1.3 – Radio frequency ultrasound signal.

function. Alternatively, to simplify subsequent mathematical manipulations, the RF signal can be expressed in its analytic form

$$z(t) = u(t) + jH(u)(t), \quad (1.7)$$

where $H(\cdot)(t)$ is the Hilbert transform, defined as

$$H(u)(t) = -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^{\infty} \frac{u(t+\tau) - u(t-\tau)}{\tau} d\tau. \quad (1.8)$$

In practice, the Hilbert transform of the signal is computed in the Fourier domain.

RF demodulation consists in extracting the signal envelope, by removing the sinusoidal carrier component. This processing step can be performed by computing the module of the analytic signal, defined as

$$U_{\text{env}} = \sqrt{u(t)^2 + H(u)(t)^2}, \quad (1.9)$$

which is referred to as the *envelope-detected RF signal*. Figure 1.4 shows the envelope-detected signal computed from the RF line of Figure 1.3.

The RF signal is sampled at the *sampling frequency* f_s . Following the Nyquist criterion, the sampling frequency guarantees a perfect reconstruction of the signal if it satisfies the inequality

$$f_s > 2f_{\text{max}}, \quad (1.10)$$

where f_{max} is the highest frequency of the signal [Shannon, 1949]. The sampling frequency is typically set to 20 MHz for convex probes, which work at frequencies below 10 MHz, and to 40 MHz for linear probes, which use higher frequencies up to 20 MHz.

Given a sampling frequency f_s , an acquisition up to depth d provides

$$n_s = \frac{f_s d}{c} \quad (1.11)$$

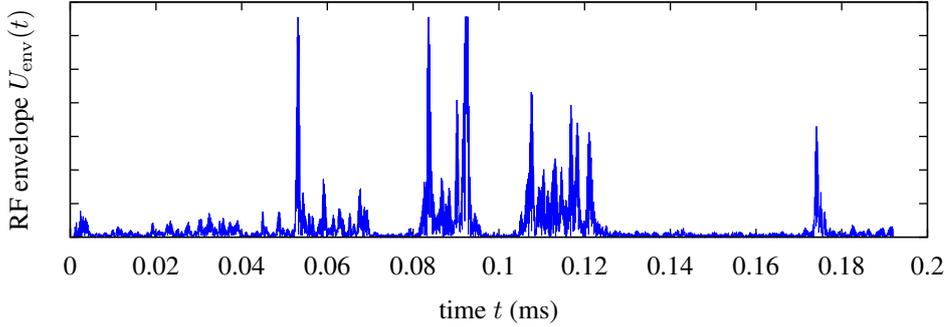


Figure 1.4 – Envelope-detected ultrasound signal.

samples. For instance, sampling at depth $d = 10$ cm with frequency $f_s = 40$ MHz provides $n_s = 2597$ samples. This number is larger than necessary for display. For this reason, the envelope detected RF signal is usually decimated, that is, subsampled by a factor 10. We define

$$f_A = \frac{f_s}{10} \quad (1.12)$$

the A-sample frequency. This corresponds to the final sampling frequency. Note that decimation is done after envelope-detection, so that the Nyquist criterion is respected for the envelope-detection. Then, the number of A-samples in a scan line is

$$AN = \frac{n_s}{10} = \frac{f_A d}{c}. \quad (1.13)$$

1.1.2.3 B-mode ultrasound

Envelope detected RF data is conventionally represented numerically by 16-bits integers. Therefore, the dynamic range has to be reduced to 8-bits for display. In order to preserve low-intensity values, dynamic range compression is typically done by a non-linear mapping, such as the logarithmic compression

$$f_{\log}(x) = A \log(x) + B, \quad (1.14)$$

or the square-root operator

$$f_{\text{sqrt}}(x) = Ax^{0.5} + B, \quad (1.15)$$

where A is the amplification parameter, and B is a linear gain parameter.

For the acquisition of a 2D section, the same procedure is repeated for each transducer element, which results in a set of LN (line number)

scan lines. A typical line number, found in most of the commercial transducers, is $LN = 128$. Note that, following the previous example of a scan line acquisition lasting for $130 \mu\text{s}$, the acquisition of a 2D section will take about 17 ms , so that the frame rate of B-mode ultrasound, with these settings, is $f_{2D} \approx 60 \text{ Hz}$. This frame rate is higher than necessary to provide a fluid display to the user. As a result, B-mode ultrasound is considered as a real-time imaging modality.

Ultrasound probes come in different shapes to fit various medical purposes. The transducer geometries can be roughly categorized in two sets. In *linear* probes, the transducer elements are aligned, and the scan lines are parallel. In *convex* (or curved) probes, the transducer elements are placed along a circular arc, and the scan lines are fan-shaped. In both cases, the scan lines are coplanar. Objects and directions contained in the plane defined by the scan lines is referred to as *in-plane*, while objects and directions not contained in that plane are referred to as *out-of-plane*. In the following, we detail the imaging geometry of linear and convex probes.

Linear probes For linear probes, the scan lines are parallel, and the imaging region is rectangular, as represented in Figure 1.5. Therefore, the processed scan lines can be directly aggregated to form the B-mode image. The relation between a physical point located at (x, y) in the field of view and a pixel (i, j) in the image is given by

$$x = L_{\text{pitch}} \times (j - j_0), \quad (1.16)$$

$$y = A_{\text{pitch}} \times i, \quad (1.17)$$

where A_{pitch} is the A-pitch, or axial resolution, defined as the distance between two consecutive samples along a scan line. Given the A-sample frequency f_A , the A-pitch can be computed as

$$A_{\text{pitch}} = \frac{c}{f_A}. \quad (1.18)$$

Similarly, the L-pitch L_{pitch} is defined as the distance between two consecutive transducer elements.

Convex probes For convex probes, however, the imaging region is a fan-shaped sector, and the conversion to B-mode requires a geometrical conversion from polar to Cartesian coordinates, which is referred to as *scan conversion*. Following the notations of Figure 1.6, let \mathcal{F}_p be the frame attached to the imaging center O of the probe, such that the x - y plane coincides with the imaging plane, and the y direction passes

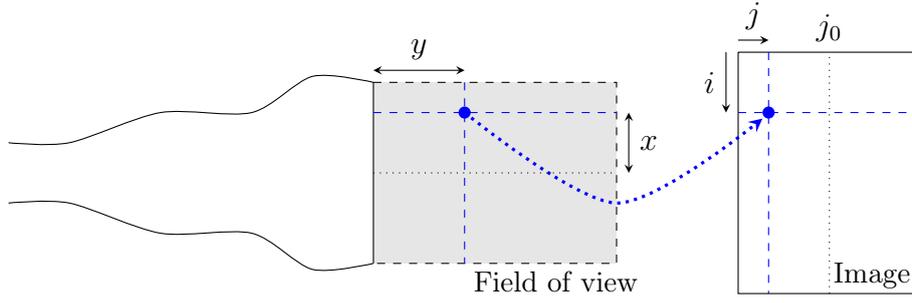


Figure 1.5 – Geometry of a 2D linear ultrasound probe.

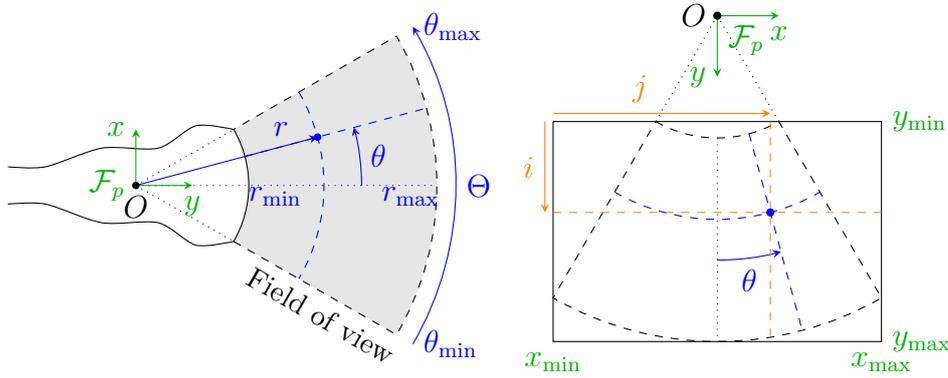


Figure 1.6 – Geometry of a 2D convex ultrasound probe.

through the center of the imaging sector. In polar coordinates, we note r the distance to O , and θ the angle with respect to the y -axis. The polar and Cartesian coordinates are related by the equations

$$x = r \sin \theta, \quad (1.19)$$

$$y = r \cos \theta. \quad (1.20)$$

Let l be the scan line index, and a the sample index along each scan line, in the order of acquisition, starting from 0. We write r_{\min} the *probe radius*, i.e., the distance from the imaging center to the first sample. Similarly, we define r_{\max} as the distance from the imaging center to the farthest sample. Under the assumption of a constant velocity of sound c , the a -th sample in each scan line corresponds to the echo produced by the tissue element located at a distance

$$r = r_{\min} + A_{\text{pitch}} a, \quad (1.21)$$

where the A-pitch A_{pitch} is defined as in (1.18).

Knowing the L-pitch (distance between two adjacent transducer elements) L_{pitch} and the probe radius r_{\min} , one can compute the angular

resolution

$$\delta\theta = \frac{L_{\text{pitch}}}{r_{\text{min}}}, \quad (1.22)$$

and the angular field of view

$$\Theta = LN\delta\theta. \quad (1.23)$$

Then, noting $\theta_{\text{min}} = -\frac{\Theta}{2}$ and $\theta_{\text{max}} = \frac{\Theta}{2}$ the limits of the angular field of view, the angular position of the l -th scan line can be computed as

$$\theta = \theta_{\text{min}} + l\delta\theta. \quad (1.24)$$

The B-mode image, with pixels indexed by (i, j) , can be reconstructed at an arbitrary resolution s with

$$i = \frac{y - y_{\text{min}}}{s}, \quad (1.25)$$

$$j = \frac{x - x_{\text{min}}}{s}, \quad (1.26)$$

where x_{min} and y_{min} are the smallest x - and y - coordinates of the field of view, respectively. These values can be easily computed as

$$x_{\text{min}} = r_{\text{max}} \sin \theta_{\text{min}}, \quad (1.27)$$

and

$$y_{\text{min}} = r_{\text{min}} \cos \theta_{\text{min}}. \quad (1.28)$$

Finally, the B-mode image is constructed by interpolating for each (i, j) the prescan image at

$$a = \frac{\sqrt{(r_{\text{max}} \sin \theta_{\text{min}} + s.j)^2 + (r_{\text{min}} \cos \theta_{\text{min}} + s.i)^2} - r_{\text{min}}}{A_{\text{pitch}}} \quad (1.29)$$

$$l = \frac{\arctan \frac{r_{\text{max}} \sin \theta_{\text{min}} + s.j}{r_{\text{min}} \cos \theta_{\text{min}} + s.i} - \theta_{\text{min}}}{\delta\theta} \quad (1.30)$$

Different techniques can be used for interpolation:

- Nearest neighbor: the value of the closest prescan data point to (a, l) is used:

$$U_{\text{postscan}}(i, j) = U_{\text{prescan}}([\![a]\!], [\![l]\!]), \quad (1.31)$$

where $[\cdot]$ is the round to nearest function. This is the simplest and fastest method. However, the resulting image presents undesirable interpolation artifacts.

- Bilinear interpolation: the value of the current pixel is interpolated linearly between the four prescan data points adjacent to (a, l) :

$$U_{\text{postscan}}(i, j) = (l - \lfloor l \rfloor)U_1 + (\lceil l \rceil - l)U_2, \quad (1.32)$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the round down and round up function, respectively,

$$U_1 = (a - \lfloor a \rfloor)U_{\text{prescan}}(\lfloor a \rfloor, \lfloor l \rfloor) + (a - \lceil a \rceil)U_{\text{prescan}}(\lceil a \rceil, \lfloor l \rfloor), \quad (1.33)$$

and

$$U_2 = (a - \lfloor a \rfloor)U_{\text{prescan}}(\lfloor a \rfloor, \lceil l \rceil) + (a - \lceil a \rceil)U_{\text{prescan}}(\lceil a \rceil, \lceil l \rceil). \quad (1.34)$$

- Bicubic interpolation: the value of the current pixel is interpolated from 16 neighboring pixels, using Lagrange polynomials or cubic splines. This technique leads to a smoother result, at the cost of a longer processing time.

1.1.2.4 3D ultrasound

Two-dimensional ultrasound probes are limited to the visualization of a planar section of the body, which complicates real-time ultrasound-based diagnosis. First, 2D ultrasound requires expertise to understand and mentally visualize three-dimensional structures of the body. Then, 2D ultrasound does not allow any direct quantitative measurement of anatomical objects or diseases. Finally, for robotic guidance, which is the main topic of this thesis, the lack of information in the out-of-plane direction makes the 3D control of an ultrasound probe challenging.

Three-dimensional ultrasound imaging has been a subject of research since the late 1990s, mainly in the context of obstetrics [Pretorius and Nelson, 1995; Nelson et al., 1996]. Two main 3D ultrasound techniques have been developed and commercialized: motorized transducers and matrix array transducers. We briefly describe these thereafter.

Remark: *Volumetric ultrasound imaging is sometimes referred to as 4D ultrasound in the literature. Then, the forth dimension represents the temporal dimension, which is a way to insist on the real-time aspect of ultrasound imaging. In order to avoid any confusion, we reserve in this dissertation the term 2D ultrasound to 2D ultrasound probes, and 3D ultrasound to 3D ultrasound probes.*

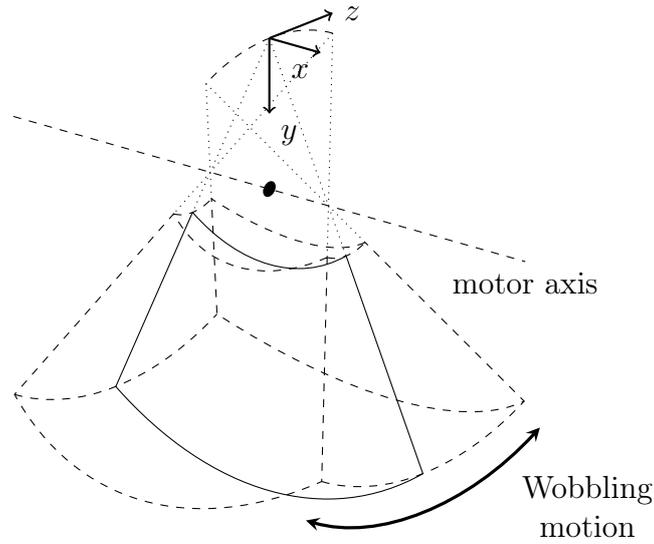


Figure 1.7 – Geometry of a motorized convex 3D ultrasound probe.

Motorized transducer A motorized 2D transducer, sometimes called *wobbler probe*, is a 2D transducer that mechanically scans a volume of interest with a back-and-forth motion. The geometry of such a probe is illustrated in Figure 1.7. This system acquires a series of B-mode frames, which can be reconstructed into a volume, based on the known scanning pattern.

Since motorized 3D ultrasound imaging relies on the successive acquisition of 2D frames, the 3D acquisition rate f_{3D} is at most equal to the frame rate f_{2D} divided by the number of frames FN :

$$f_{3D} < \frac{f_{2D}}{FN}. \quad (1.35)$$

For instance, for 33 frames per volume, a SonixTOUCH scanner (BK Ultrasound, MA) with the 4DC7-3/40 Convex 4D transducer acquires 1.1 volumes per second. To limit the artifacts due to motion, the sweeping motion has to be slow enough, so as to consider the transducer static during the acquisition of a frame. As a result, the volume rate of such a probe is limited.

Matrix array transducer A more recent design for 3D ultrasound imaging consists in a 2D matrix array of transducer elements, which allows direct volume acquisition [Woo and Roh, 2012]. In this case, the transducer elements are arranged in a 2D regular grid, that can be either planar or bi-convex.

This design alleviates the limitations of motorized scanning, by increasing the acquisition speed, while eliminating the artifacts due to the sweeping motion. However, the number of piezoelectric elements embedded in matrix array transducers is usually limited, which leads to a relatively low resolution, compared to motorized transducers.

1.2 Soft tissue tracking

The localization of anatomical landmarks is an important step in the analysis of medical images. Automatic segmentation, which consists in delineating the contours of a structure of interest, offers a mean to ease the interpretation of medical images, and to obtain a reliable diagnosis by the direct extraction of quantitative features. In addition, detecting anatomical features in real-time is necessary for the development of intelligent robot-assisted procedures. Ultrasound image segmentation has been the subject of intensive research for various medical applications [Noble and Boukerroui, 2006]. Having in mind the context of robotized ultrasound examination, we focus, in this section, on methods capable of providing real-time tracking, rather than image-by-image segmentation. On the other hand, we do not limit the scope of tracking to the precise delineation of an object's contour, but we also consider the tracking of a region of interest. Therefore, instead of an exhaustive review of ultrasound image segmentation techniques, the purpose of this section is to provide an overview of real-time ultrasound tracking algorithms, that can be useful to the design of ultrasound-based robot control strategies.

1.2.1 Motion tracking

1.2.1.1 Block matching

Block matching is a motion estimation technique, that consists in finding corresponding blocks (small regions of interest) between two consecutive images of a sequence. The underlying assumption is that the motion inside each block is rigid, so that pixels contained in a pair of matching blocks have a one-to-one correspondence.

Sliding window Given a block defined in one frame, the localization of the corresponding block in the next frame can be performed by the optimization of a similarity measure with a sliding window approach, as illustrated in Figure 1.8.

For instance, [Trahey et al., 1987] use the B-mode ultrasound intensity correlation as similarity measure to track the motion of speckle pattern

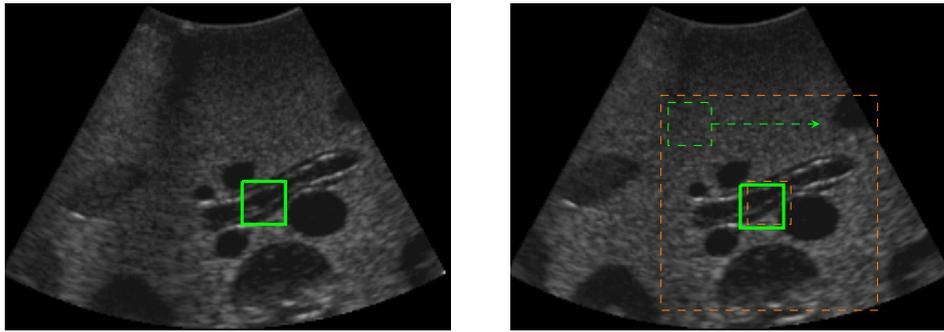


Figure 1.8 – Sliding window approach for block matching.

for blood flow estimation, and [Golemati et al., 2003] use the normalized correlation to track the carotid artery wall. Given two blocks containing N pixels with intensity $(I_1(i))_{i=1}^N$ and $(I_2(i))_{i=1}^N$, the normalized correlation is defined as

$$\rho(I_1, I_2) = \sum_{i=1}^N \frac{(I_1(i) - \bar{I}_1)(I_2(i) - \bar{I}_2)}{\sigma(I_1)\sigma(I_2)}, \quad (1.36)$$

where $\bar{\cdot}$ is the mean operator, and $\sigma(\cdot)$ is the standard deviation. The normalized correlation ρ takes values in the interval $[-1, 1]$, where $\rho = 1$ corresponds to a perfect correlation ($I_1 = kI_2$, where k is a positive scalar), $\rho = 0$ denotes a total decorrelation, and $\rho = -1$ denotes a perfect anti-correlation ($I_1 = kI_2$, where k is a negative scalar). Due to the normalization by the standard deviation of the signals, this similarity measure is independent of global intensity variations.

Other similarity measures used for block matching in ultrasound images, and inspired from the computer vision community, include the sum of absolute differences [Bohs and Trahey, 1991], and the sum of squared differences (SSD) [Yeung et al., 1998; Ortmaier et al., 2005; Zahnd et al., 2011]. A review and comparison of block matching algorithms for ultrasound images can be found in [Golemati et al., 2012].

In addition to generic similarity measures, some authors have also designed ultrasound-specific similarity measures, in order to better model the specificities of the ultrasound signal. [Cohen and Dinstein, 2002] and [Boukerroui et al., 2003] use the Rayleigh distribution to model the speckle noise statistics. This model is used for speckle tracking in 3D echocardiography in [Song et al., 2007] and [Linguraru et al., 2008]. [Myronenko et al., 2009] introduce similarity measures based on the bivariate Rayleigh and the Nakagami distributions. A Nakagami-based similarity measure is used in [Wachinger et al., 2012].

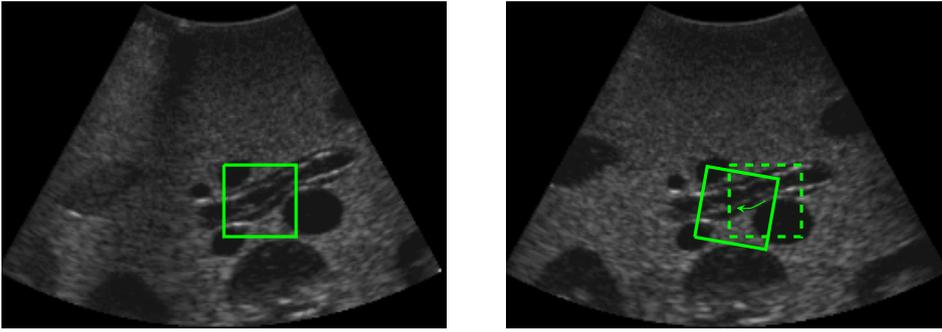


Figure 1.9 – Gauss-Newton optimization approach for block matching.

Instead of an exhaustive sliding window search, the optimization can be performed in a neighborhood around the previous block location, when a bound on the displacement is known. Reducing the search space has the double advantage to reduce the computational cost, and to limit the risk of mismatch with similar patterns at other locations.

However, the sliding window approach is limited to the tracking of pure translation motion. To account for a rotation of the target, it is necessary to optimize the orientation of the Region of Interest (ROI) as well. An exhaustive search for the orientation can become computationally costly, especially in 3D, so that smarter optimization approaches should be considered.

Gauss-Newton optimization [Hager and Belhumeur, 1998] describe a linear optimization approach based on the Gauss-Newton algorithm. This method is used in [Krupa et al., 2009] and [Nadeau et al., 2015] to track a region of interest in 2D and 3D ultrasound, respectively. It consists in iteratively updating the position of the ROI to converge towards a position that minimizes the difference with the initial template, as illustrated in Figure 1.9.

Let $I : \Omega \rightarrow \mathbb{R}$ represent the ultrasound image where the template is searched for, and \mathbf{P} a point with coordinates (x, y, z) in a frame attached to the ROI \mathcal{R} . The pose of the ROI is represented by a vector of parameters $\mathbf{r} = (t_x, t_y, t_z, \theta u_x, \theta u_y, \theta u_z)$, encoding its position and orientation with respect to the image frame. Here, the terms (t_x, t_y, t_z) correspond the translation between the image frame and the ROI frame, and $(\theta u_x, \theta u_y, \theta u_z)$ correspond the axis-angle representation of the rotation between these frames. The variation of the intensity $I_{\mathbf{P}}$ at \mathbf{P} can be related to the displacement of the ROI by an interaction matrix

$$\mathbf{L}_{I_{\mathbf{P}}} = \frac{\partial I_{\mathbf{P}}}{\partial \mathbf{r}}, \quad (1.37)$$

which can be written

$$\mathbf{L}_{I_{\mathbf{P}}} = \left(\frac{\partial I_{\mathbf{P}}}{\partial \mathbf{P}} \right)^{\top} \frac{\partial \mathbf{P}}{\partial \mathbf{r}}, \quad (1.38)$$

where the first term corresponds to the 3D image gradient

$$\frac{\partial I_{\mathbf{P}}}{\partial \mathbf{P}} = (\nabla I_x \quad \nabla I_y \quad \nabla I_z). \quad (1.39)$$

The second term is given by Varignon's formula for velocity composition:

$$\frac{\partial \mathbf{P}}{\partial \mathbf{r}} = \begin{bmatrix} 1 & 0 & 0 & 0 & z & -y \\ 0 & 1 & 0 & -z & 0 & x \\ 0 & 0 & 1 & y & -x & 0 \end{bmatrix}, \quad (1.40)$$

so that the interaction matrix $\mathbf{L}_{I_{\mathbf{P}}}$ takes the form

$$\mathbf{L}_{I_{\mathbf{P}}} = \begin{bmatrix} \nabla I_x & \nabla I_y & \nabla I_z \\ y \nabla I_z - z \nabla I_y & -x \nabla I_z + z \nabla I_x & x \nabla I_y - y \nabla I_x \end{bmatrix}. \quad (1.41)$$

The intensity pattern inside the region \mathcal{R} can be represented by a vector $\mathbf{s}_{\mathcal{R}}$ of pixel intensities interpolated on a regular grid $G(\mathcal{R})$:

$$\mathbf{s}_{\mathcal{R}} = (I_{\mathbf{P}})_{\mathbf{P} \in G(\mathcal{R})}, \quad (1.42)$$

and the interaction matrix for this vector is obtained by stacking the interaction matrices associated to all points in $G(\mathcal{R})$:

$$\mathbf{L}_{\mathbf{s}_{\mathcal{R}}} = \begin{bmatrix} \mathbf{L}_{I_{\mathbf{P}_1}} \\ \vdots \\ \mathbf{L}_{I_{\mathbf{P}_N}} \end{bmatrix}, \quad (1.43)$$

where N is the size of $\mathbf{s}_{\mathcal{R}}$.

Noting $\mathbf{s}_{\mathcal{R}}^*$ the reference template, the optimization process aims at minimizing the sum of squared errors

$$SSE(\mathbf{s}_{\mathcal{R}}, \mathbf{s}_{\mathcal{R}}^*) = \sum_{i=1}^N (\mathbf{s}_{\mathcal{R}} - \mathbf{s}_{\mathcal{R}}^*)^2. \quad (1.44)$$

The Gauss-Newton algorithm performs the least squares optimization iteratively, with updates

$$\delta \mathbf{r} = -\mathbf{L}_{\mathbf{s}_{\mathcal{R}}}^+ (\mathbf{s}_{\mathcal{R}} - \mathbf{s}_{\mathcal{R}}^*), \quad (1.45)$$

where $\mathbf{L}_{\mathbf{s}_{\mathcal{R}}}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{L}_{\mathbf{s}_{\mathcal{R}}}$, defined as

$$\mathbf{L}_{\mathbf{s}_{\mathcal{R}}}^+ = (\mathbf{L}_{\mathbf{s}_{\mathcal{R}}}^{\top} \mathbf{L}_{\mathbf{s}_{\mathcal{R}}}) \mathbf{L}_{\mathbf{s}_{\mathcal{R}}}^{\top}. \quad (1.46)$$

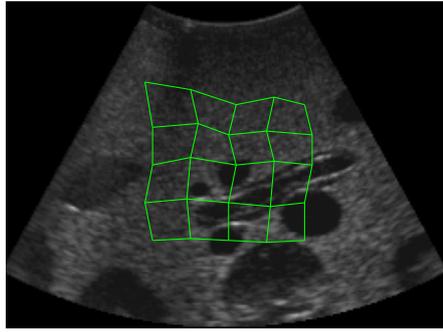


Figure 1.10 – Deformable tracking with an irregular mesh.

To obtain a more stable convergence, it is common to introduce a convergence rate parameter $\lambda \in [0, 1]$, and to perform updates

$$\delta \mathbf{r} = -\lambda \mathbf{L}_{\mathbf{s}_{\mathcal{R}}}^+ (\mathbf{s}_{\mathcal{R}} - \mathbf{s}_{\mathcal{R}}^*). \quad (1.47)$$

Note that this optimization approach requires the computation of the image gradients and the inversion of the interaction matrix $\mathbf{L}_{\mathbf{s}_{\mathcal{R}}}$ for each image. Instead, one can use the interaction matrix corresponding to the reference template in the optimization [Nadeau et al., 2015]. The update equation is in this case

$$\delta \mathbf{r} = -\lambda \mathbf{L}_{\mathbf{s}_{\mathcal{R}}^*}^+ (\mathbf{s}_{\mathcal{R}} - \mathbf{s}_{\mathcal{R}}^*), \quad (1.48)$$

which only requires the computation of the vector error $\mathbf{s}_{\mathcal{R}} - \mathbf{s}_{\mathcal{R}}^*$ and the multiplication by the constant interaction matrix $\mathbf{L}_{\mathbf{s}_{\mathcal{R}}^*}^+$.

1.2.1.2 Deformable block matching

The block matching techniques described above can only model rigid motion. However, soft tissues are subject to deformations due to physiological motion and compression originating from the application of the ultrasound probe. Deformable block matching, initially proposed in [Seferidis and Ghanbari, 1994] under the term generalized block matching, consists in modeling the region of interest with an irregular mesh (Figure 1.10).

This technique has been applied to ultrasound images for thyroid nodule imaging [Basarab et al., 2008] and echocardiography [Touil et al., 2010]. In these methods, a bilinear model is used to describe the displacement, and rigid block matching is used to estimate the position of the mesh nodes. A high motion estimation accuracy can be reached using a multiscale approach, where the mesh resolution is refined iteratively [Touil et al., 2010]. [Richa et al., 2010] use a thin-plate spline model

to cope with large deformations of cardiac tissues in stereoscopic images for beating heart surgery. A similar method is used in [Lee and Krupa, 2011] for non-rigid motion compensation of soft tissues with 3D ultrasound images. [Royer et al., 2017] propose a 3D mesh-based deformable target tracking combining image information with physical constraints, in order to reduce the sensitivity to image noise. The method has real-time capability and was shown to be efficient in the context of liver tracking [Royer et al., 2015].

1.2.1.3 Elastic registration

More generally, finding correspondences in successive images can be seen as a registration problem. For instance, [Foroughi et al., 2006] propose an elastic registration method to perform tracking in sequences of 3D ultrasound images of the liver. [Leung et al., 2009] design a dynamic feature-based elastic registration algorithm for 2D ultrasound images. [Khallaghi et al., 2012] generalize this method to 3D ultrasound images, and present an experimental evaluation on carotid artery, liver, and kidney images.

For a more detailed introduction to motion estimation in ultrasound images, the reader is referred to the recent survey [De Luca et al., 2015].

1.2.2 Deformable shape models

Deformable shape methods consist in fitting a parameterized shape to the image data. These methods typically rely on the presence of salient features, such as edges, to optimize a predefined shape template so that its boundaries match tissue interfaces. This can be seen as a top-down approach, where the shape template is defined *a priori*, and optimized globally based on image data. On the other hand, aforementioned motion tracking methods work bottom-up, by estimating first local displacements before to infer the global deformation.

1.2.2.1 Active contour

Perhaps one of the most popular classes of deformable shape models is the active contour model [Kass et al., 1988], also known as *snake*. The active contour model consists in the representation of the contour by a parameterized curve, such as a spline, represented by a set of control points. The contour is assigned an energy, defined as the combination of an internal energy and an external energy. The internal energy controls, for instance, the length and smoothness of the curve, while the external

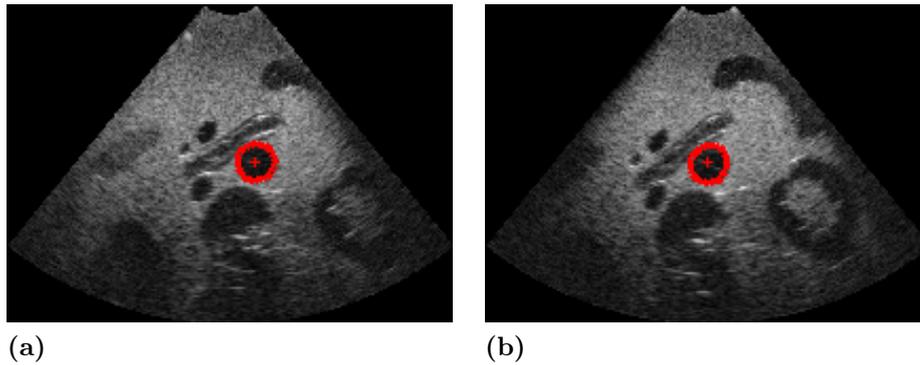


Figure 1.11 – Vessel tracking in ultrasound images using an active contour algorithm. These results were obtained with the algorithm of [Li et al., 2011].

energy depends on image features, and is usually designed to be minimal when the curve follows image edges. Active contour segmentation consists in minimizing the energy of the contour in an iterative manner. This iterative optimization is well suited for real-time tracking, because it can be performed continuously in successive frames. In this case, the detected contour of each frame is used as initialization in the next frame. Figure 1.11 provides an illustration of active contour for tracking the section of a vessel in ultrasound images, using the method proposed in [Li et al., 2011].

Active contour models have been used early in the development of ultrasound image segmentation, in particular in the context of echocardiography. Indeed, the segmentation of the borders of the left ventricle for the estimation of the left ventricular volume is of interest for the quantification of cardiac function and the diagnosis of cardiovascular diseases [Frangi et al., 2001]. In one of the earliest papers on the subject of echocardiographic image segmentation, [Herlin and Ayache, 1992] propose an active contour detection, that uses edge detection and morphological operations as preprocessing steps, to track the left auricle and the mitral valve in a 2D echocardiographic sequence. [Chalana et al., 1996] use a multiple active contours model to detect the epicardial and endocardial borders in 2D echocardiography. [Mikic et al., 1998] incorporate optical flow estimation in the contour model to cope with fast-moving boundaries, and apply their method to the tracking of the mitral valve leaflets, aortic root, and left ventricle endocardial borders. [Jacob et al., 1999] constrain the contour evolution based on learned motion priors.

Active contour models have also been used for prostate segmentation. [Knoll et al., 1999] propose a contour parameterization based on the

dyadic wavelet transform to allow a local analysis of the contour shape. The contour parameters are constrained by a shape prior.

1.2.2.2 Active shape models

Active shape models [Cootes and Taylor, 1992] are statistical models (or templates) of objects, pretrained from a set of example data, that are deformed to fit the image at hand. Although these methods are closely related to active contour models, an important difference is that active shape models incorporate a strong shape prior which eases the initialization of the shape fitting procedure, and constrains the deformations. Active shape models can also refer to a wide range of 2D and 3D shape models, while active contours are most commonly restricted to 2D curves. Thereafter, we provide some examples of the use of active shape models for contour tracking in echocardiography. For a more comprehensive review of active shape models for medical image segmentation in general, the interested reader is referred to [Carneiro and Nascimento, 2013].

[Mignotte et al., 2001] propose a deformable template model coupled with a Markov Random Fields model to perform a Bayesian segmentation of the endocardial boundary. [Jacob et al., 2002] use principal component analysis in the shape optimization for tracking the myocardial borders in echocardiography. [Comaniciu et al., 2004] incorporate a measurement uncertainty estimation to improve the robustness of shape tracking for the myocardial borders. Measurement uncertainties are also considered in [Nascimento and Marques, 2008], where a combination of multiple dynamic models is used.

1.3 Instrument tracking

Minimally invasive surgical procedures such as biopsy or localized tumor ablation require the insertion of a thin needle towards a precise anatomical target. The needle positioning accuracy has a critical impact on the outcome of the intervention. Studies have shown the potential of ultrasound imaging for guiding the needle insertion when direct vision is not possible [Chapman et al., 2006]. Thus, ultrasound is being increasingly used clinically for needle guidance. However, the visualization of the needle in ultrasound images is challenging, and manual needle guidance requires a perfect coordination of the clinician between the positioning of the ultrasound probe and the insertion of the needle.

Robust image analysis techniques can greatly help the clinician during the insertion procedure, by enhancing the visibility of the needle, or by

providing additional cues on the state of the intervention. In addition, an automatic detection of the position of the needle enables the use of a robot to assist the insertion. It was recalled in a recent study that robotic needle guidance can improve diagnostic and therapeutic accuracy [Kaye et al., 2014]. For this reason, the detection of biopsy needles in ultrasound images has become an intensive research subject, both in the medical imaging and medical robotics communities.

This section offers a review of ultrasound-based needle tracking methods, and details some contributions made in the context of this thesis. As previously done for soft tissue tracking methods, a focus is put on methods with real-time tracking capability. Two general trends can be distinguished among the wide range of methods that have been developed for needle tracking. First, hardware approaches, where additional equipment, or a hardware modification of the needle, is used in order to ease the detection of the needle. These methods have the advantage to bypass, or at least to ease, the challenging task of detecting the needle in ultrasound images. Section 1.3.1 gives a short overview of these methods. Other needle tracking methods are purely based on the analysis of ultrasound images. In spite of the difficulty of this task, recent approaches have demonstrated the feasibility of image-based needle tracking, therefore avoiding the need for complex hardware additions to the intervention procedure. These methods are detailed in section 1.3.2.

1.3.1 Hardware approaches

Rigid needle guide A simple approach to control the insertion path of the needle consists in having a mechanical guide rigidly attached to the ultrasound probe. This way, the needle is inserted through the guide, and the insertion axis is perfectly known in a frame attached to the ultrasound image. However, such an approach drastically limits the possible insertion trajectories.

Optical tracking Optical tracking relies on one or several cameras observing the instruments. Using a calibration between the camera frame and the ultrasound probe frame, and the detected position of the instrument, one can display the needle trajectory on the ultrasound image (Figure 1.12). Markers can be attached to the instrument to facilitate the localization. These can be either passive or active (emission of infrared light). [Chan et al., 2005] propose a markerless optical tracking system, where two cameras are used to detect the needle shaft. The use of external cameras to track the needle his limited by the line-of-sight requirement: the line between the camera and the instrument has

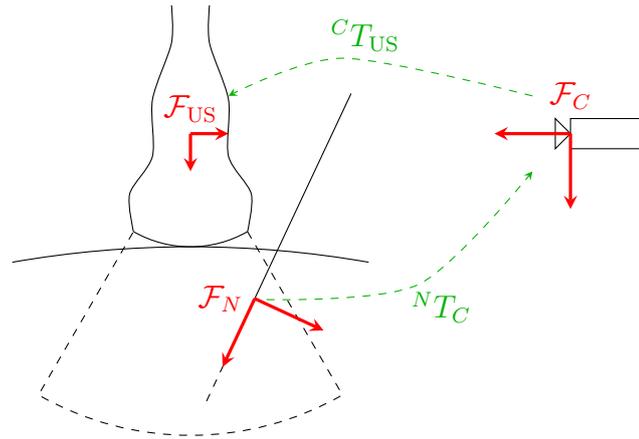


Figure 1.12 – Optical needle tracking system. Via a calibration between the camera frame \mathcal{F}_C and the ultrasound probe frame \mathcal{F}_{US} , and an estimation of the pose of the needle frame \mathcal{F}_N in \mathcal{F}_C , one can infer the pose of the needle in the probe frame.

to be free from occlusions, which is not realistic in the cluttered environment of an operating room. [Najafi et al., 2015] use a single camera mounted on the ultrasound probe to detect the needle insertion trajectory. The advantage of placing the camera directly on the ultrasound probe is twofold. First, the line-of-sight issue is reduced. Second, the calibration between the camera and the ultrasound image is simplified. [Stolka et al., 2014] propose the stereo imaging system Clear Guide ONE (Clear Guide Medical, MD). This accessory is mounted on the ultrasound probe, and it allows the detection of the needle and its registration with the ultrasound image.

Electromagnetic tracking Electromagnetic tracking can be used for tracking a needle during insertion. An electromagnetic field is generated by a device placed next to the patient, and a magnetic sensor embedded into the needle allows its localization within the field. [Zaaroor et al., 2001] use the MAGELLAN electromagnetic tracking system (Biosense Webster, Inc.) for neurosurgical procedures. [Howard et al., 2001] and [Krombach et al., 2001] use the UltraGuide 1000 system, which provides electromagnetic needle tracking for ultrasound-guided interventions. [Banovac et al., 2002] demonstrate the use of the AURORA tracking device (Northern Digital Inc., Ontario, Canada) to perform liver tumor biopsy. [Boctor et al., 2008] use the Flock of Birds tracking system (Ascension Technology Corp, VT) to track a hand-held 2D ultrasound probe. The tracking information is used to reconstruct a 3D ultrasound

image in which the needle can be visualized. However, electromagnetic tracking only offers a limited accuracy, and requires specialized needles. In addition, the incorporation of the electromagnetic device into the operating room necessitates special attention to ensure that no ferro-magnetic objects can disturb the tracking system.

Both optical and electromagnetic tracking require a calibration between the tracking device and the ultrasound image, so that they are subject to registration errors.

Needle motion detection Color Doppler imaging can facilitate the detection of the needle when it is in motion. The use of color Doppler imaging to visualize a needle was first proposed in [Kurohiji et al., 1990]. The detection can be improved by applying an additional motion to the needle or its inner stylet [Hamper et al., 1991]. [Fronheiser et al., 2008] propose a design where a piezoelectric buzzer vibrates the needle, and detection is done with 3D Doppler. [Adebar et al., 2014] use 3D Doppler to detect a vibrating flexible needle, and perform needle steering in biological tissues. However, the authors also report that the vibrations, originating from the needle base, are damped by the tissues when the needle is deeply inserted, which results in a lower visibility of the needle. [Mignon et al., 2015] use the rotation of the needle shaft to generate vibrations that can be detected with 3D Doppler imaging. This approach alleviates some of the limitations of previous methods, since it does not require a vibrating device to be mounted on the needle base, and vibrations are naturally generated along the needle shaft.

A limitation of Doppler imaging is that the frame rate is slower than with B-mode imaging. This impacts the real-time tracking and control capability of Doppler-based needle guidance. In a recent article, [Beigi et al., 2016] propose a method to detect a needle in B-mode images, using an analysis of tremor motion during manual insertion. This is a promising direction, since the results show that Doppler imaging is not necessary to detect needle motion. However, the tremor motion analysis relies on a spatiotemporal sampling, which is performed on a sequence of frame, and not in real-time.

1.3.2 Ultrasound-based tracking

The methods presented above require either a specific needle design, or the use of additional equipment (tracking system, vibrating device). Optical and electromagnetic tracking also necessitate a calibration with the ultrasound probe, in order to register the detected needle position to the ultrasound image. Alternatively, the needle can be directly detected in

the ultrasound image. Indeed, the needle, being metallic, generates a discernible strong echo in the image. Ultrasound-based tracking has the advantage of being free from additional equipment, and it can be used with conventional needles. In addition, the needle is directly localized in the image frame. The curvature of flexible needles can also be detected in ultrasound, which is not possible via optical tracking.

Early work on needle detection is based on parallel projection methods, which are designed to find imperfect instances of parameterized shapes by optimizing the integral of the image along parallel curves. [Hong et al., 2002] use the Hough transform [Hough, 1962] to detect straight needles in 2D ultrasound images. [Aboofazeli et al., 2009] adapt this method to the detection of mildly curved needles in a 3D volume, by first projecting the volume onto 2D planes by a ray casting process. The Parallel Integral Projection (PIP), an alternative algorithm to the Hough transform, is used in [Cachard et al., 2006] to localize straight needles in 3D volumes. [Uherčík et al., 2008] propose a mutli-resolution scheme to speed-up the computation. [Novotny et al., 2007b] optimize the Radon transform [Radon, 1917] for an implementation on a Graphics Processing Unit (GPU) to achieve real-time detection. The method is applied to the detection of endoscopic instruments with passive reflective markers. The generalized Hough transform allows the detection of more complex instrument shapes, such as flexible needles. The shape can be modeled, for instance, by a Bézier polynomial [Neshat and Patel, 2008]. However, the estimation of the generalized Hough transform from large 3D volumes is computationally expensive, even with an implementation on a GPU. Several variants of the Hough transform can reduce the computation time, using coarse-to-fine strategies [Zhou et al., 2008] or randomization [Qiu et al., 2008].

Projection-based methods have a relatively high computational complexity, and an accurate needle localization in large volumes is time-consuming, even with a parallel implementation. These methods also lack robustness with cluttered background. Alternatively, the Random Sample Consensus (RANSAC) algorithm can be used to detect polynomial curves in 2D or 3D ultrasound volumes [Uherčík et al., 2010]. Instead of an exhaustive search for the needle shape configuration, the RANSAC algorithm is driven by the image information to sample candidate shapes that are likely to correspond to the needle position. Thereafter, we detail this approach following the description in [Chatelain et al., 2013].

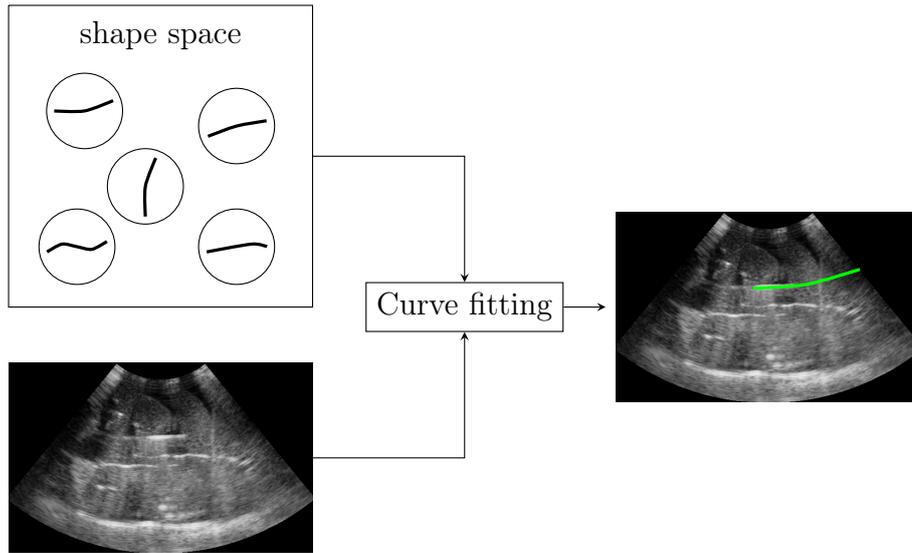


Figure 1.13 – Curve fitting for needle detection in ultrasound images. The curve fitting algorithm estimates the shape that best fits the current image, from a candidate shape space.

1.3.2.1 Curve fitting

Flexible needles can bend during the insertion due to contact forces with surrounding tissues. Even needles considered as rigid can be subject to a slight bending, so that a straight trajectory assumption is likely to be violated. For this reason, it is of interest to model the needle with a deformable shape. Curve fitting is a promising approach for needle detection in ultrasound images, where the shape of the needle is modeled by a parametric curve, such as a polynomial curve or a spline. Curve fitting consists in estimating the parameters of the curve that best fit to the image at hand. Consequently, this approach requires two main components. First, a *shape space* that describes the possible configurations of the curves. This consists in a type of curve, and its associated parameters. Second, a cost function, which evaluates how well a given curve fits to the ultrasound data. This framework is illustrated in Figure 1.13.

In the following, we start by describing a polynomial curve model, which is widely used in the literature. Then, we propose two different approaches to curve fitting.

Polynomial curve model Polynomial curve models allow a fairly adaptive description of the needle shape at a low computational expense.

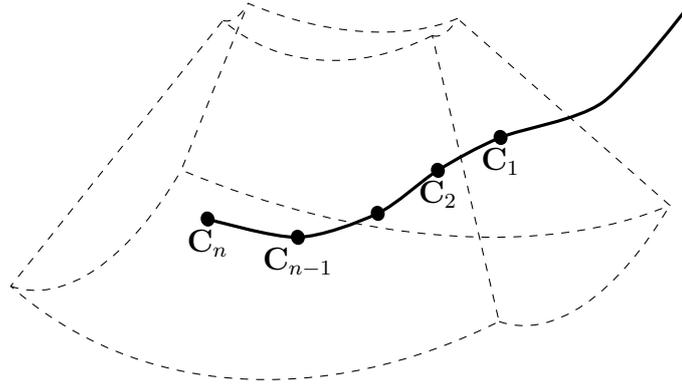


Figure 1.14 – Polynomial curve model of a flexible needle.

A 3D polynomial curve of order $n - 1$ can be defined as

$$\mathcal{C}(a, \mathbf{H}) = \mathbf{H} [1 \quad a \quad \cdots \quad a^{n-1}]^\top, \quad (1.49)$$

where $\mathbf{H} \in \mathbb{R}^{3 \times n}$ is the matrix of polynomial coefficients, and a is the curve parameter. The start and end points of the curve are obtained by constraining the parameter a to an interval $[0, 1]$ such that, by convention, $a = 0$ corresponds to the starting point, and $a = 1$ to the end point (the tip of the needle). The order $n - 1$ of the polynomial determines the degree of flexibility of the shape. Possible values are, for example, $n = 2$ for a straight line, $n = 3$ for a parabola, and $n = 4$ for a cubic curve. Note that, to enable a modeling of the curvature in 3D, n should be at least equal to 4 (with $n = 3$, the curve is always planar).

The polynomial curve can be equivalently defined by a set of n control points $\mathbf{C}_i \in \mathbb{R}^3$ along the curve (Figure 1.14). Under the constraint of equally spaced control points, the control points are defined as

$$\mathbf{C}_i = \mathcal{C} \left(\frac{i-1}{n-1}, \mathbf{H} \right). \quad (1.50)$$

The vector of control points $\mathbf{C} \in \mathbb{R}^{3 \times n}$ can be computed via the matrix multiplication

$$\mathbf{C} = \mathbf{H}\mathbf{V}^\top, \quad (1.51)$$

where \mathbf{V} is the Vandermonde matrix [Pascal, 1900] associated to the

vector of parameters $\mathbf{a} = (0, \frac{1}{n-1}, \frac{2}{n-1}, \dots, 1)$:

$$\mathbf{V} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & \frac{1}{n-1} & \cdots & (\frac{1}{n-1})^{n-1} \\ 1 & \frac{2}{n-1} & \cdots & (\frac{2}{n-1})^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}. \quad (1.52)$$

Since the elements of \mathbf{a} are distinct, \mathbf{V} is invertible, and the matrix \mathbf{H} can be computed from the control points by

$$\mathbf{H} = \mathbf{C} (\mathbf{V}^\top)^{-1}. \quad (1.53)$$

1.3.2.2 Random sample consensus

We have presented a polynomial curve model that can be used to represent the shape of a needle. We now introduce a first method to optimize the shape model for a given ultrasound image. This method is based on the RANSAC algorithm.

Candidate set The RANSAC algorithm is a data-driven iterative parameter estimation method designed to be robust to outliers. In the context of curve fitting in images, this method relies on the definition of a candidate set of pixels (or voxels) \mathcal{X} , as represented in Figure 1.15. Let $I : \Omega \rightarrow \mathbb{R}$ be an ultrasound image. The candidate set \mathcal{X} can be defined as the output of a binary filter $f(I) : \Omega \rightarrow \{0, 1\}$ which is tailored to distinguish pixels belonging to the needle:

$$\mathcal{X} = \{\mathbf{p} \in \Omega : f(I)(\mathbf{p}) = 1\}. \quad (1.54)$$

A simple selection criterion, based on the assumption that the needle shaft appears as high-intensity pixels in the image, consists in defining a threshold τ on the intensity:

$$f(I)(\mathbf{p}) = \begin{cases} 1 & \text{if } I(\mathbf{p}) \geq \tau, \\ 0 & \text{if } I(\mathbf{p}) < \tau. \end{cases} \quad (1.55)$$

Inlier set Given a curve $\mathcal{C} : [0, 1] \rightarrow \mathbb{R}^3$, one can define a set of inliers from \mathcal{X} , i.e., pixels that belong to the curve, with a certain deviation tolerance. To this end, we define a cost function $h_{\mathcal{C}} : \Omega \rightarrow \mathbb{R}$ associated to the curve, and we define the set of inliers as those pixels in \mathcal{X} for which the cost function is lower than a threshold h_0 . Noting \mathcal{X}_{in} the set of inliers, we have

$$\mathcal{X}_{\text{in}} = \{\mathbf{p} \in \mathcal{X} : h_{\mathcal{C}}(\mathbf{p}) < h_0\}. \quad (1.56)$$

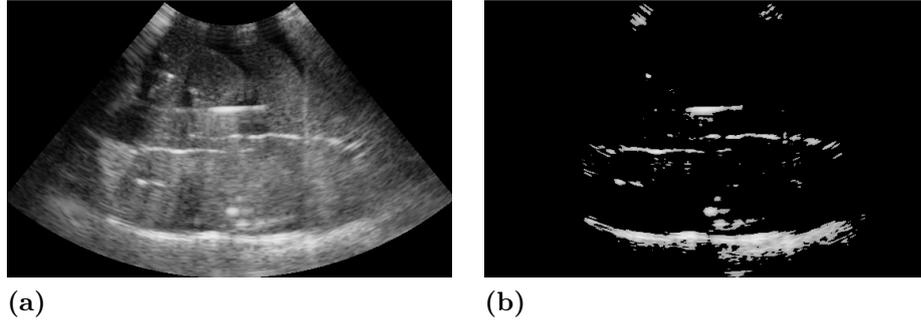


Figure 1.15 – Needle detection using the RANSAC algorithm. (a) Ultrasound image. (b) Candidate set obtained by thresholding.

Iterative sampling The RANSAC algorithm consists in iteratively sampling curves from the candidate set \mathcal{X} in order to optimize the number of inliers. A polynomial curve of order $n - 1$ can be minimally defined by a set of n distinct points. Therefore, any set of n distinct points in \mathcal{X} is referred to as a Minimal Sampling Set (MSS). At each iteration of the RANSAC algorithm, an MSS \mathcal{S} is uniformly sampled from \mathcal{X} . It defines a curve $\mathcal{C}_{\mathcal{S}}$ with parameters $\mathbf{H}_{\mathcal{S}}$. Then, the corresponding set of inliers $\mathcal{X}_{\text{in}}(\mathbf{H}_{\mathcal{S}})$ is computed, and it is retained if it is the largest set seen so far. This procedure is described in Algorithm 1.

Stopping criterion The stopping criterion is defined so has to guarantee a probability of failure of at most ϵ . Given the probability p_{in} to select an inlier when sampling uniformly \mathcal{X} , the required number of iterations to reach a failure probability below ϵ is

$$T = \frac{\log \epsilon}{\log(1 - p_{\text{in}}^n)}. \quad (1.57)$$

On the other hand, p_{in} can be bounded by the largest proportion of inliers found so far, that is,

$$p_{\text{in}} \leq \frac{|\mathcal{X}_{\text{in}}|}{|\mathcal{X}|}. \quad (1.58)$$

Thus, the stopping criterion can be defined as performing at least

$$\hat{T} = \frac{\log \epsilon}{\log \left(1 - \left(\frac{|\mathcal{X}_{\text{in}}|}{|\mathcal{X}|} \right)^n \right)} \quad (1.59)$$

iterations.

Algorithm 1 The RANSAC algorithm

```

1: procedure RANSAC( $\mathcal{X}, n, \epsilon$ )
2:    $\mathcal{X}_{\text{in}} \leftarrow \emptyset$ 
3:    $T \leftarrow \infty$ 
4:    $t \leftarrow 0$ 
5:   while  $t < T$  do
6:      $\mathcal{S} \leftarrow \emptyset$ 
7:     while  $|\mathcal{S}| < n$  do
8:        $i \sim \mathcal{U}([1, n])$ 
9:       if  $\mathcal{X}_i \notin \mathcal{S}$  then
10:         $\mathcal{S} \leftarrow \mathcal{S} \cup \{\mathcal{X}_i\}$ 
11:      end if
12:    end while
13:     $\mathbf{H} \leftarrow \mathbf{S} (\mathbf{V}^\top)^{-1}$ 
14:     $\mathcal{X}_{\text{in}} \leftarrow \{\mathbf{p} \in \mathcal{X} : h_{\mathcal{C}}(\mathbf{p}) < h_0\}$ 
15:    if  $|\mathcal{X}_{\text{in}}| > |\mathcal{X}_{\text{in}}^*|$  then
16:       $\mathcal{X}_{\text{in}}^* \leftarrow \mathcal{X}_{\text{in}}$ 
17:       $\mathbf{H}^* \leftarrow \mathbf{H}$ 
18:       $T \leftarrow \frac{\log \epsilon}{\log(1 - (|\mathcal{X}_{\text{in}}|/|\mathcal{X}|)^n)}$ 
19:    end if
20:     $t \leftarrow t + 1$ 
21:  end while
22:  return  $(\mathbf{H}^*, \mathcal{X}_{\text{in}}^*)$ 
23: end procedure

```

Local optimization The RANSAC algorithm provides a robust estimate of the shape model, but it has a relatively poor accuracy, since its parameters are estimated from a minimal sample set of n points only. In order to reach a higher accuracy, [Uherčík et al., 2010] perform a local optimization of the parameters from the detected set of inliers \mathcal{X}_{in} , using a derivative-free Nelder-Mead downhill simplex method [Nelder and Mead, 1965]. [Zhao et al., 2012] use a Kalman filter on the resulting detection in order to improve the stability of the tracking in a sequence of volumes.

Early pruning The RANSAC algorithm, as presented in Algorithm 1, involves the computation of the inlier set \mathcal{X}_{in} for a large number of curve models, most of which are physically incoherent. In order to avoid the computation of the cost function for these unlikely configurations, [Uherčík et al., 2010] use a constrained version of the RANSAC algorithm, where the inlier set computation step is not performed for models with excessively high curvature. In [Chatelain et al., 2013], we propose, in addition, to prune unlikely configurations by considering the angle between the candidate curve and the prediction of a Kalman filter. We define this angle as

$$\alpha = \arccos \left(\frac{\mathbf{k}_0 \cdot \tilde{\mathbf{k}}_0}{\|\mathbf{k}_0\| \times \|\tilde{\mathbf{k}}_0\|} \right), \quad (1.60)$$

where \mathbf{k}_0 is the principal direction of the candidate model, and $\tilde{\mathbf{k}}_0$ is the principal direction of the Kalman-predicted model. Such a pruning has the double advantage to speed-up the convergence, and to increase the robustness to outliers. More generally, the pruning step can be represented by a binary shape acceptance function $g(\mathcal{C}_t, \{\mathcal{C}_0, \dots, \mathcal{C}_{t-1}\})$, which decides whether the current candidate shape \mathcal{C}_t is acceptable, based on the past detections $\{\mathcal{C}_0, \dots, \mathcal{C}_{t-1}\}$.

Tip localization The result of the line fitting procedure is a curve in space. In order to complete the needle detection, one still needs to estimate the start and end points of the shape along the curve. The start point can be simply defined as the intersection between the curve and the image border. For the tip localization, [Barva et al., 2008] propose to sample the voxel intensities along the curve, and to identify an important intensity drop (see Figure 1.16). This is motivated by the fact that the needle tip appears as a bright spot in the ultrasound image.

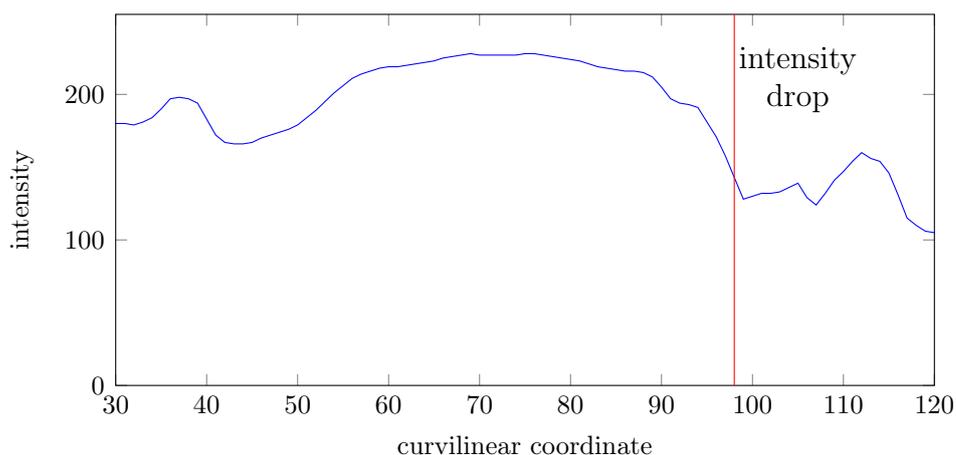


Figure 1.16 – Needle tip localization by intensity drop detection. The tip of the needle can be identified by detecting an important drop in intensity along the curve.

Limitations The RANSAC-based needle detection method relies on the definition of a filter f that provides a set of candidate voxels, a cost function h that determines the set of inliers with respect to a curve, and an acceptance function g that filters the candidate shapes. As a result, the success of the RANSAC algorithm is highly dependent on the definition of these three functions. In particular, the definition of the filter f has a critical impact on the algorithm, since it defines the initial candidate set used for RANSAC.

1.4 Needle tracking via particle filtering

As discussed in the previous section, the RANSAC-based needle detection method is sensitive to the parameters used in the appearance and shape models. In order to overcome this limitation, we propose in this section a new algorithm for flexible needle tracking, based on particle filtering [Gordon et al., 1993]. Instead of sampling candidate shapes from a set of filtered voxels (image-driven) we sample the curves in the space of possible shapes, based on a motion model (dynamics-driven). The image data is then used *a posteriori* to compute a likelihood for each candidate shape. As a result, the shape constraints are directly driving the sampling process.

1.4.1 Bayesian tracking

In the Bayesian tracking framework, the target is defined as a state vector \mathbf{x}_t of size N following a dynamics equation of the form

$$\mathbf{x}_t = f_t(\mathbf{x}_{t-1}, \nu_{t-1}), \quad (1.61)$$

where $f_t : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ is the dynamics function, modeling the evolution of the target, and ν_{t-1} is an independently and identically distributed (i.i.d.) state noise. The purpose of tracking is to estimate the state vector \mathbf{x}_t based on some measurement $\mathbf{z}_t = h_t(\mathbf{x}_t, \mu_t)$, where μ_{t-1} is an i.i.d. measurement noise.

In our case, the state vector corresponds to the needle's position, which we model as the concatenation of all control points defining the needle:

$$\mathbf{x}_t = \left[\mathbf{C}_t^{(1)\top} \quad \dots \quad \mathbf{C}_t^{(n)\top} \right]^\top \quad (1.62)$$

1.4.2 Needle dynamics

The dynamics of the system is supposed to be known through an estimation of the external needle insertion velocity. For instance, such an estimation can be made by the odometry of a robot inserting the needle, or via an external tracking device. We assume that the needle does not buckle during insertion, so that the scalar insertion velocity at the needle tip corresponds to the insertion velocity component of the insertion device. Given an insertion velocity $v_t \in \mathbb{R}$ over a period δt , the state update equation for the needle tip can be written as:

$$\mathbf{C}_t^{(n)} = \mathbf{C}_{t-1}^{(n)} + v_t \delta t \mathbf{u}_{t-1}(1) + \nu_{t-1}^{(n)} \quad (1.63)$$

where $\nu_{t-1}^{(n)} \in \mathbb{R}^3$ is the tip's state noise and $\mathbf{u}_{t-1}(1)$ is the unitary direction vector of the curve at the needle tip, as represented in Figure 1.17. The unitary direction vector at a position a along the curve can be computed analytically as

$$\mathbf{u}_{t-1}(a) = \frac{\mathbf{H}_{t-1} [0 \ 1 \ 2a \ \dots \ (n-1)a^{(n-2)}]^\top}{\| \mathbf{H}_{t-1} [0 \ 1 \ 2a \ \dots \ (n-1)a^{(n-2)}]^\top \|_2}. \quad (1.64)$$

Considering that the ultrasound probe does not move with respect to the scene, the entry point is defined as a fixed point in the volume frame. Therefore, its update equation is simply $\mathbf{C}_t^{(1)} = \mathbf{C}_{t-1}^{(1)} + \nu_{t-1}^{(1)}$, where $\nu_{t-1}^{(1)} \in \mathbb{R}^3$ is the entry point's state noise. The noise model for the entry point allows the system to correct its position if the needle drifts due to tissue deformation, or if the ultrasound probe moves.

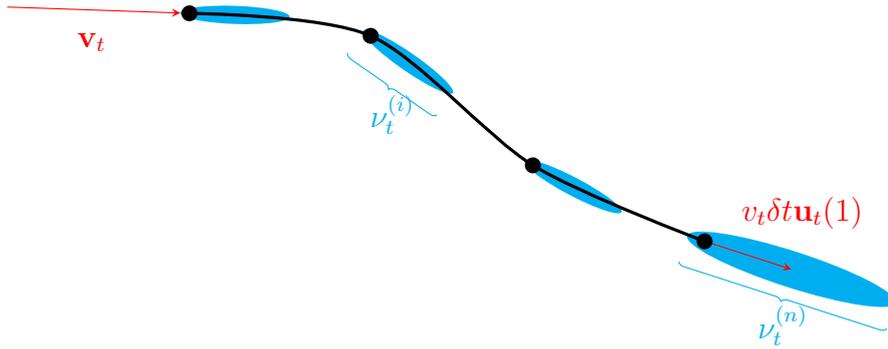


Figure 1.17 – Dynamic model of a flexible needle. The insertion velocity \mathbf{v}_t during a time δt induces a displacement $v_t \delta t \mathbf{u}_t(1)$ of the needle tip. The blue ellipses represent the uncertainty on the estimation of the next state.

Since all other control points are defined to be equally spaced between the entry point and the needle tip, their state update equation can be approximated with:

$$\mathbf{C}_t^{(i)} = \mathbf{C}_{t-1}^{(i)} + \frac{i-1}{n-1} v_t \delta t \mathbf{u}_{t-1} \left(\frac{i-1}{n-1} \right) + \nu_{t-1}^{(i)} \quad (1.65)$$

where $\nu_{t-1}^{(i)}$ is the state noise for the control point $\mathbf{C}^{(i)}$, which allows the detection of needle shaft drifting. The motion of the ultrasound probe could also easily be incorporated in the system model by modifying (1.63) and (1.65).

1.4.3 Particle filtering

The key idea behind particle filtering is to model the posterior distribution of the tracked object with a large number M of random samples \mathbf{x}_t^m , $m = 1, \dots, M$ (the particles). To each particle is associated a weight w_t^m that corresponds to its likelihood. In our case, each particle represents the shape of a needle modeled as a polynomial curve, and the weights are updated according to the likelihood of the particle given the current ultrasound image. With the notations introduced above, this can be expressed by the relation

$$w_t^m \propto p(\mathbf{z}_t | \mathbf{x}_t^m), \quad (1.66)$$

with the constraint $\sum_{m=1}^M w_t^m = 1$. The term \mathbf{z}_t corresponds to the ultrasound image observed at time t .

The current state is approximated from this discrete posterior distribution as:

$$\hat{\mathbf{x}}_t = \sum_{m=1}^M w_t^m \mathbf{x}_t^m. \quad (1.67)$$

1.4.3.1 Appearance model

The computation of the weights requires some model of the needle's appearance in the image in order to approximate the likelihood $p(\mathbf{z}_t | \mathbf{x}_t^m)$. For example, one could apply statistical learning techniques to estimate this likelihood. In this paper we consider a simple model which presents the advantage to be very cost-effective, and that does not require any prior learning. We propose to use directly the intensity of the voxels along the needle's shaft. More specifically, we consider the mean intensity \bar{I}_t^m along the curve defined by the particle \mathbf{x}_t^m :

$$\bar{I}_t^m = \frac{1}{L} \int_{a=0}^1 I(\mathcal{C}(a, \mathbf{H}_t^m)) da, \quad (1.68)$$

where L is the length of the curve and $I(\mathbf{P})$ is the intensity at the position \mathbf{P} . Indeed, the intensity of the echos reflected by the needle is in general higher than that of soft tissues. Even if the needle's appearance can be discontinuous (the shaft does not always reflect very well, and the corresponding intensities can be in the same range as surrounding speckle), integrating the intensities along the curve allows discrimination between the different particles. In addition, the needle tip typically appears as a very bright spot due to reflections on the bevel. We take this observation into account in our model, and define the weight w_t^m as a trade-off between the mean shaft intensity and the tip intensity:

$$w_t^m \propto \bar{I}_t^m + \alpha I(\mathcal{C}(1, \mathbf{H}_t^m)), \quad (1.69)$$

where α is a positive constant. Giving more weight (higher α) to the tip intensity ensures a precise detection of the needle tip. An advantage of this measure in terms of computation time is that we only need to access the intensities of the volume along the curves defined by the particles, and not in the whole volume.

1.4.3.2 Sequential importance re-sampling

Then, we use a Sequential Importance Re-Sampling (SIR) filter to update the particles during tracking. The initial state is determined manually via an interface that allows the user to browse the ultrasound volume and

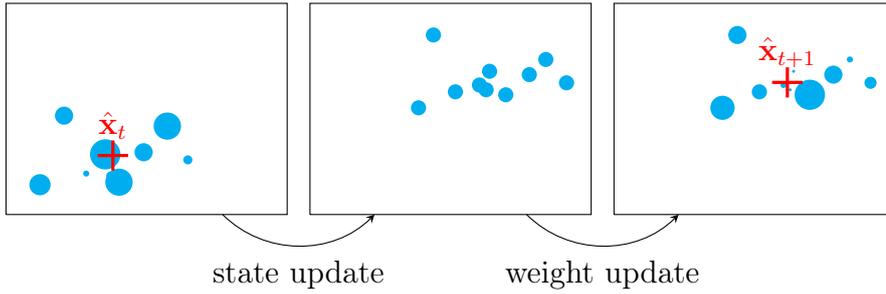


Figure 1.18 – Particle filtering on a simple 2D example with 10 particles. The particles are represented by the blue points, with a size proportional to their weight. Each particle is updated according to the state update model. Then, the weight of each particle is updated according to the likelihood of the observation. The state estimate $\hat{\mathbf{x}}$ corresponds to the weighted barycenter of the particles.

to select the location of the needle. The weights are initialized equally as $w_0^m = \frac{1}{M}$. At each iteration, the SIR filter performs the following operations, which are illustrated in Figure 1.18 on a simple 2D example:

1. Update the particles according to the model’s dynamics.
2. Re-compute the weights w_t^m according to (1.69).
3. Estimate the current state according to (1.67).
4. Re-sample.

The aim of the re-sampling step is to avoid the *degeneracy phenomenon*, where the weight of one particle tends to 1 while the others are negligible. The degeneracy of the system can be detected by computing the *effective* number of particles

$$N_{\text{eff}} = \frac{1}{\sum_{m=1}^M (w_t^m)^2}. \quad (1.70)$$

In order to avoid degeneracy, the particles are re-sampled when the effective number of particles falls below a threshold N_0 . In this case, the particles are re-sampled M times with replacement with probability proportional to their weight. The weights are then reset to $\frac{1}{M}$. The objective of re-sampling is to eliminate particles with small weight while generating more particles from those with higher weights.

The proposed SIR-based needle tracking algorithm is designed to track a curved needle in ultrasound images using an estimation of the insertion velocity v_t applied at the needle base, provided by the robot

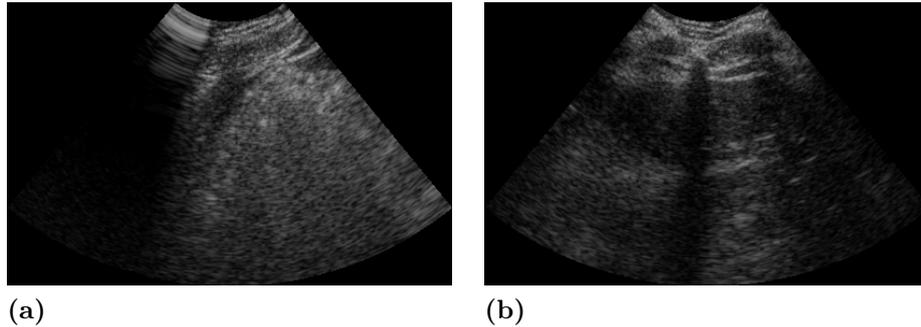


Figure 1.19 – Ultrasound images of a human abdomen showing (a) a drop-out artifact on the left, and (b) a local acoustic shadow in the center.

odometry or an external tracking device. The particle filter scheme allows the system to correct for imprecision in the velocity estimation and to detect the bending of the needle, based on ultrasound measurements.

1.5 Quality estimation

In sections 1.2 and 1.3, we have described different methods for tracking soft tissues or surgical instruments in ultrasound images. These methods generally assume that the ultrasound signal is equally reliable within the image. However, ultrasound images are heterogeneous in quality, so that parts of the image may not provide reliable information. This can lead to tracking or interpretation errors, and might cause the previously described algorithms to fail. In this section, we discuss the issue of ultrasound image quality estimation.

We can define two main types of ultrasound-specific artifacts:

- the *drop-out artifact*, when the coupling is not sufficient for the sound emitted by some transducer elements to enter the body. In this case, most of the amplitude of the sound wave is reflected at the patient's surface, which results in black scan lines in the ultrasound image. This artifact is illustrated in Figure 1.19(a).
- *shadowing*, when the ultrasound wave travels through the tissues for a certain distance, and is then stopped at an important interface between different tissue types. This is illustrated in Figure 1.19(b).

Different approaches have been proposed to estimate the quality of ultrasound images, with different motivations. We can broadly distinguish four categories of ultrasound quality estimation methods. Global

quality measures are used to evaluate the overall quality of the image. Acoustic attenuation estimation (or signal loss estimation) methods aim at estimating the physical attenuation properties of the tissues from the RF signal. Binary shadow detection simplifies this problem to the classification of pixels as shadowed or not. Finally, confidence maps provide a smooth pixelwise quality measure, based on a simplified model of sound propagation. These methods offer a good trade-off between the complex acoustic attenuation estimation, and binary shadow detection. We briefly review each of these categories.

Global quality measures Global quality measures have been investigated for the purpose of providing a quality index to ultrasound image. This global quality index can be used, for instance, to evaluate the performance of image enhancement techniques (such as speckle reduction), or simply to estimate the reliability of the image for subsequent diagnosis. [Loizou et al., 2006] study various quality measures based on statistical and texture analysis, mostly derived from general image quality measures. Similarly, [Kuhlemann, 2013] uses the entropy of the ultrasound intensity distribution, and the sum of gradients, as global image measures. Such quality measures are quite general, and do not account for the characteristics of ultrasound imaging. In an interesting alternative approach, a measure of ultrasound image quality is inferred from expert ratings, using a machine learning algorithm [El-Zehiry et al., 2013]. The authors use the learned quality metric to optimize the imaging parameters of an ultrasound scanner, in an autofocus fashion. However, the quality measure remains global, i.e., at the image level.

Acoustic attenuation estimation Because of the physics underlying their acquisition, ultrasound images are heterogeneous in quality. Specifically, the amplitude of the ultrasound wave is attenuated at tissue interfaces, which can result in local *shadows* in the image below strong attenuators, such as bone or gas. Ultrasound image quality also depends on the contact properties between the ultrasound probe and the patient, such as the contact force, and the presence of acoustic coupling gel.

The automatic detection of such artifacts is a crucial prerequisite for subsequent processing and analysis steps. In addition, acoustic shadowing may be of diagnostic value, since it can help characterizing tissue types. Shadow detection was even originally considered mainly as a tissue characterization technique for diagnosis purposes. For instance, [Kuc and Schwartz, 1979] use an estimation of the acoustic attenuation coefficient for the diagnosis of liver condition. The acoustic attenua-

tion coefficient is estimated via a maximum likelihood estimator based on the log spectral difference of the ultrasound signal. [Drukker et al., 2003] detect posterior acoustic shadowing in breast ultrasound to help distinguishing between benign and malignant breast lesions, which have a different acoustic shadow signature. The authors use the skewness of the ultrasound intensity distribution within small regions of interest to detect shadowing. [Madabhushi et al., 2006] use the Adaboost algorithm to discriminate between lesions and shadowing, based on the local statistical distribution of the ultrasound signal.

Shadow detection Aside from tissue characterization, acoustic shadowing can be seen as undesirable, because it affects subsequent image processing tasks. This is the case in ultrasound image segmentation, tracking, registration, or 3D reconstruction, where it is important to localize shadowing artifacts that could impair the accuracy of the algorithm. Penney et al. use shadow detection to discard shadowed areas during US/MRI [Penney et al., 2004] and US/CT [Penney et al., 2006] registration. The authors define the shadow along each scan line as the area below which the ultrasound intensity is smaller than a fixed threshold T_{art} . This definition is motivated by the fact that ultrasound propagates mainly in the direction of the scan line. Figures 1.20(b), (e), and (h) show some results obtained with a reimplementation of this method, using the threshold $T_{\text{art}} = 40$ as recommended by the authors. [Leroy et al., 2004] fit a decreasing exponential function to each scan line, and define a shadow as an area in which the correlation with the exponential function is higher than a threshold. A similar method is also used in [Wein et al., 2007], where the authors define a threshold σ_y on the variance of the ultrasound intensity to detect the shadow region in each scan line. The starting indices of shadowing along the scan lines are regularized laterally using a median filter. Figures 1.20(c), (f), and (i) show some results obtained with a reimplementation of this method. The authors do not specify the value of σ_y used in their work. We have used $\sigma_y = 200$. [Hellier et al., 2010] use a statistical noise model to detect shadowing along each scan line. Again, a median filter is used to regularize the boundary of the shadow region. This method is applied in [Berton et al., 2016] to the detection of vertebrae.

Confidence maps The methods mentioned above provide a binary identification of shadow areas. While this is useful for some applications, it does not provide a precise estimate of ultrasound signal quality. In this regard, it is interesting to define the quality of the signal as the relative

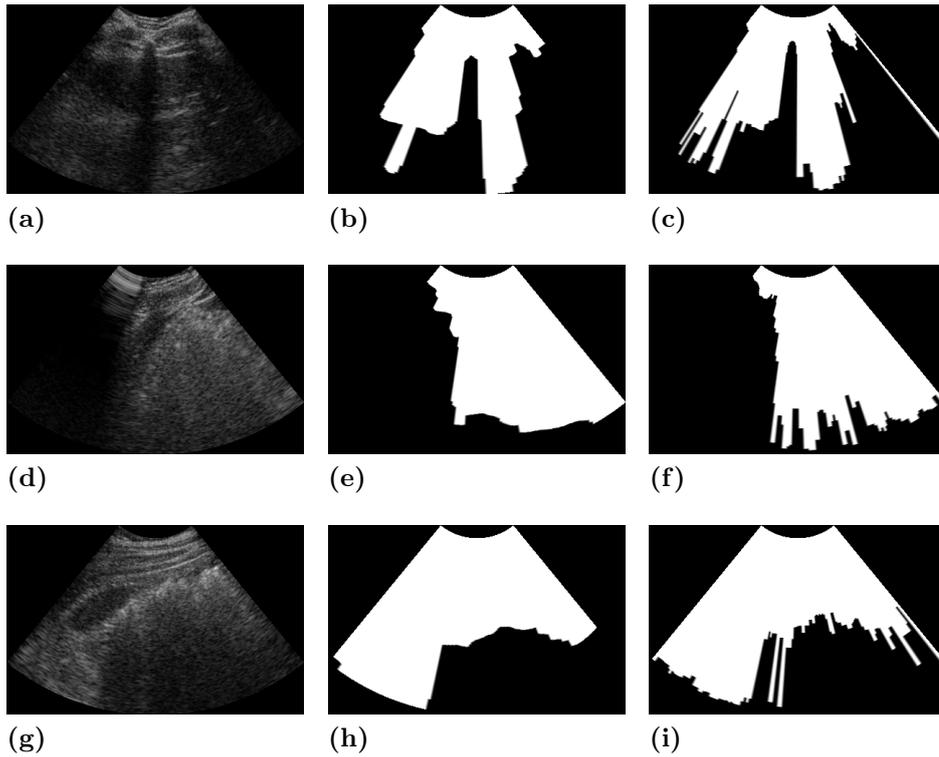


Figure 1.20 – Acoustic shadow detection. First column: ultrasound images of a human abdomen. Second column: shadow segmentation using the method of [Penney et al., 2004]. Third column: shadow segmentation using the method of [Wein et al., 2007]. First row: local shadow. Second row: drop-out artifact. Third row: diffuse shadow.

amplitude of the ultrasound wave that is transmitted up to a certain depth, with respect to the amplitude of the emitted signal. [Wein et al., 2008] simulate ultrasound images from CT, based on the assumption that X-ray attenuation is proportional to the acoustic impedance. An interesting by-product of this framework is an ultrasound transmission map, which can be seen as a pixelwise measure of quality. [Gill et al., 2012] use a similar method for US/CT registration of the spine. Pixels detected as shadowed are discarded from the registration procedure.

Recently, [Karamalis et al., 2012] introduced the concept of *confidence map*, a pixel-wise representation of ultrasound quality, where the measure of quality takes values in the interval $[0, 1]$. The quality estimation is performed globally, using a Random Walks (RW) algorithm [Grady, 2006] based on a set of image-defined propagation constraints. The method is applied to improved ultrasound reconstruction and mosaicing. [Klein and Wells III, 2015] propose an extension of the method to RF ultrasound data, where the propagation constraints are based on a similarity measure specific to the RF intensity distribution. The use of the confidence map was shown useful to improve atherosclerotic tissue labeling [Katouzian et al., 2012; Sheet et al., 2014], bone detection [Wein et al., 2015], and registration [Kutarnia and Pedersen, 2015]. It has also been considered as a tool for enhanced ultrasound visualization [Schulte zu Berge et al., 2014, 2015].

In the following, we describe two different approaches to estimate an ultrasound confidence map. First, the random walks method initially introduced in [Karamalis et al., 2012]. Second, a simple but more computationally efficient method based on Scan Line Integration (SLI). Then, we propose a comparison between these two methods in terms of computational cost and temporal regularity.

1.5.1 Random walks confidence map

1.5.1.1 Graphical representation

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected weighted graph, with vertices \mathcal{V} the set of image pixels, and edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ representing the pixel neighborhood relation. For instance, we assume that \mathcal{E} represents pixel adjacency in terms of 4-connectivity for 2D images (see Figure 1.21). Each edge $(\mathbf{p}, \mathbf{q}) \in \mathcal{E}$ is assigned a weight $w(\mathbf{p}, \mathbf{q}) > 0$ that represents the probability of propagation through that edge. To model the ultrasound beam propagation direction, \mathcal{E} is separated in two subsets \mathcal{E}_r and \mathcal{E}_l , where \mathcal{E}_r contains pixels that are radially adjacent (they belong to the same scan line), and \mathcal{E}_l contains pixels that are laterally adjacent (they belong to

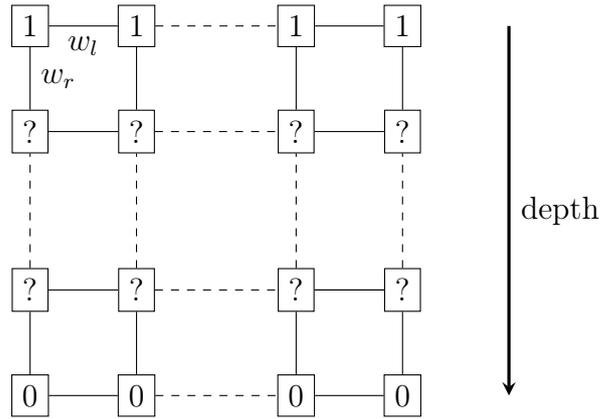


Figure 1.21 – Graph layout for the computation of the random walks confidence map.

adjacent scan lines, and have the same depth). The weights for edges in \mathcal{E}_r , noted w_r , model the ultrasound attenuation due to tissue interfaces. For an edge $(\mathbf{p}, \mathbf{q}) \in \mathcal{E}_r$, we define

$$w_r(\mathbf{p}, \mathbf{q}) = \exp(-\beta|\check{I}(\mathbf{p}) - \check{I}(\mathbf{q})|), \quad (1.71)$$

where \check{I} is the attenuated intensity, defined as

$$\check{I}(\mathbf{p}) = I(\mathbf{p}) \exp(-\alpha d(\mathbf{p})), \quad (1.72)$$

$d(\mathbf{p})$ being the depth of \mathbf{p} . The parameters α and β control the importance of acoustic attenuation and absorption, respectively.

The ultrasound beam width is modeled by allowing lateral propagation as well. For an edge $(\mathbf{p}, \mathbf{q}) \in \mathcal{E}_l$, the propagation weight between \mathbf{p} and \mathbf{q} is defined as

$$w_l(\mathbf{p}, \mathbf{q}) = \exp(-\beta(|\check{I}(\mathbf{p}) - \check{I}(\mathbf{q})| + \gamma\|\mathbf{p} - \mathbf{q}\|)), \quad (1.73)$$

where γ is a parameter related to the beam width.

1.5.1.2 Random walks

Once the propagation graph is constructed, the confidence estimation problem is solved using the random walks algorithm [Grady, 2006]. This algorithm requires boundary conditions, which are defined as follows (see Figure 1.21):

- The confidence is 1 at the minimal depth.
- The confidence is 0 at the maximal depth.

The weights of (1.71) and (1.73) are used to define the graph Laplacian matrix

$$\Lambda(\mathbf{p}, \mathbf{q}) = \begin{cases} \sum_{\mathbf{r} \in \mathcal{N}(\mathbf{p})} w(\mathbf{p}, \mathbf{r}) & \text{if } \mathbf{p} = \mathbf{q} \\ -w(\mathbf{p}, \mathbf{q}) & \text{if } \mathbf{q} \in \mathcal{N}(\mathbf{p}) \\ 0 & \text{otherwise} \end{cases}, \quad (1.74)$$

where $w(\mathbf{p}, \mathbf{q})$ refers to w_r or w_l according to the configuration of \mathbf{p} and \mathbf{q} , and $\mathcal{N}(\mathbf{p})$ is the set of neighbors of \mathbf{p} . Let \mathbf{c} be a vector representation of the confidence map. We assume that \mathbf{c} is ordered in such a way that known (K) values (the boundary conditions) appear before unknown (U) values: $\mathbf{c} = (\mathbf{c}_K^\top \ \mathbf{c}_U^\top)^\top$. Thus, the Laplacian matrix can be written

$$\Lambda = \begin{bmatrix} \Lambda_K & \mathbf{B} \\ \mathbf{B}^\top & \Lambda_U \end{bmatrix}, \quad (1.75)$$

and the confidence map is obtained by solving for \mathbf{c}_U the sparse linear system

$$\Lambda_U \mathbf{c}_U = -\mathbf{B}^\top \mathbf{c}_K. \quad (1.76)$$

Figure 1.22 shows some examples of confidence maps estimated with the random walks method, from ultrasound images of a human abdomen. The ultrasound images exhibit a local shadow, a drop-out artifact, and a diffuse shadow, which are clearly visible in the confidence map (dark areas). Since the confidence map is estimated via a global optimization with a two-dimensional lattice, sound propagation in lateral directions is considered, which results in spatially smooth confidence maps.

1.5.2 Scan line integration

The random walks method provides a good representation of ultrasound signal quality. However, this algorithm has two important limitations. First of all, it is quite computationally expensive, as it involves the resolution of large sparse linear system. In addition, while the confidence maps estimated via this method are spatially smooth, they lack a temporal regularity. Indeed, the confidence values at a given location can be subject to variations in successive frames, even if the content of the ultrasound image only varies slowly.

In order to overcome these limitations, we propose thereafter an alternative method for the estimation of the confidence map. This new method is based on an estimation of the acoustic loss along each scan line individually. Therefore, we do not perform any lateral regularization of the confidence map. We will then show in section 1.5.3 that the scan line integration method is more computationally efficient and more regular in time than the random walks method.

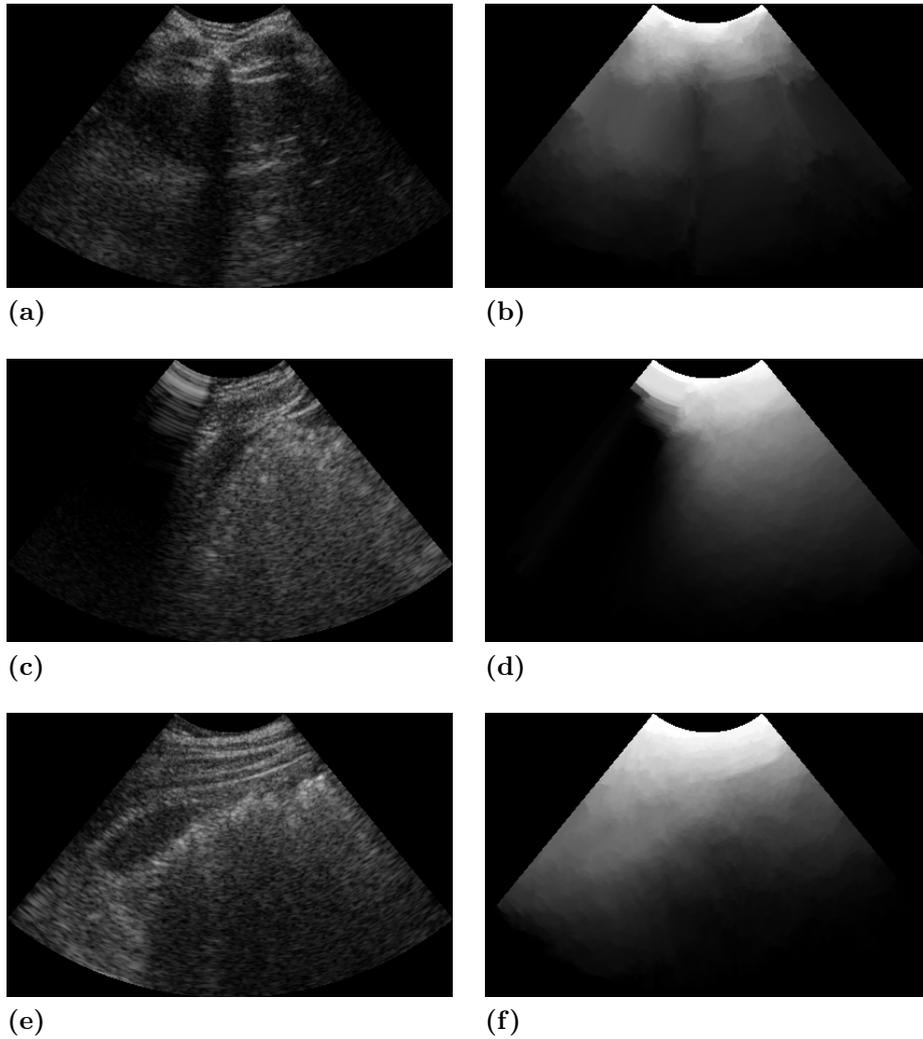


Figure 1.22 – Confidence maps estimated via the random walks method. Left: ultrasound image of a human abdomen. Right: the corresponding confidence maps. From top to bottom: local shadow, drop-out artifact, diffuse shadow.

Let $U(r, \theta)$ be the intensity of the envelope-detected RF signal corresponding to the position (r, θ) in the transducer's field of view. A simple model for ultrasound signal attenuation consists in assuming that the amplitude of the signal decreases proportionally to the difference in acoustic impedance of the tissues around the interface it propagates through. Since acoustic impedance difference is also related to the RF image intensity, we can estimate the ultrasound signal confidence via an integration along each scan line as follows:

$$C(R, \theta) = 1 - \left(\int_{R_{min}}^{R_{max}} U(r, \theta) dr \right)^{-1} \int_{R_{min}}^R U(r, \theta) dr, \quad (1.77)$$

where $\int_{R_{min}}^{R_{max}} U(r, \theta) dr$ is a normalization factor designed to scale the confidence values between 0 and 1. This model is similar to the one introduced in [Karamalis et al., 2012] as a baseline for the evaluation of their confidence map.

Most ultrasound scanners do not provide access to the RF signal, and only provide the B-mode image. As mentioned in section 1.1.2.3, the conversion from RF to B-mode consists in two main steps. First, the intensities of the envelope-detected RF signal are compressed via a non-linear mapping $I_{pre} = f(U)$ in order to adjust the dynamic range of the image for display. For instance, commonly used mappings are the logarithmic compression

$$f(x) = A \log x + B, \quad (1.78)$$

or the square-root operator

$$f(x) = Ax^{-0.5} + B, \quad (1.79)$$

where A is the amplification parameter, and B a linear gain parameter. Then, the prescan image I_{pre} is converted to Cartesian coordinates according to the geometry of the transducer. In order to estimate the confidence map from a B-mode image, one can simply revert these steps to obtain an approximation of the envelope-detected RF signal. For the square-root operator, this amounts to computing the confidence map from the estimated signal

$$\hat{U} = I_{pre}^2. \quad (1.80)$$

Figure 1.23 shows some examples of confidence maps obtained with the scan line integration method, from the same ultrasound images as in section 1.5.1. As can be observed by comparing these images to Figure 1.22, the confidence maps obtained by the scan line integration

method are less smooth spatially than with the random walks method. However, as we show in section 1.5.3, we obtain a better temporal regularity. For some applications, in particular for robot control, temporal regularity can be more important than spatial regularity.

1.5.3 Comparison

1.5.3.1 Computation time

The computation of the scan line integrated confidence map can be implemented as in Algorithm 2. The complexity of this algorithm is $\mathcal{O}(AN.LN)$, i.e., linear in the size of the image. This makes the scan line integration method computationally more efficient than the random walks method.

Algorithm 2 Scan line integration

```
1: procedure SCANLINEINTEGRATION( $I, AN, LN$ )           ▷  $I$  of size  
    $AN \times LN$   
2:   for  $j : 0 \rightarrow LN - 1$  do  
3:      $s \leftarrow 0$                                      ▷ Cumulative intensity  
4:     for  $i : 0 \rightarrow AN - 1$  do  
5:        $C(i, j) \leftarrow s$   
6:        $s \leftarrow s - I(i, j)$   
7:     end for  
8:     for  $i : 0 \rightarrow AN - 1$  do           ▷ Normalization between 0 and 1  
9:        $C(i, j) \leftarrow (C(i, j) + s)/s$   
10:    end for  
11:  end for  
12:  return  $C$                                          ▷  $C$  confidence map of size  $AN \times LN$   
13: end procedure
```

To illustrate the difference in processing time between the two methods, we ran a benchmark test with images of size ranging from 30×8 to 960×256 . The results of this experiment are presented in Figure 1.24. For reference, we show, on the same plot, the setting corresponding to an acquisition frame rate of 25 frames per seconds, for an image size of 480×128 (intersection between the two dotted lines). As expected, the plot for the scan line method is straight with a slope of 1, which corresponds to a linear complexity. On the other hand, the algorithm based on random walks is supralinear, and is approximately 200 times slower at resolution 480×128 . The random walks method is also too slow to process the images in real-time at full resolution. With the tested imple-

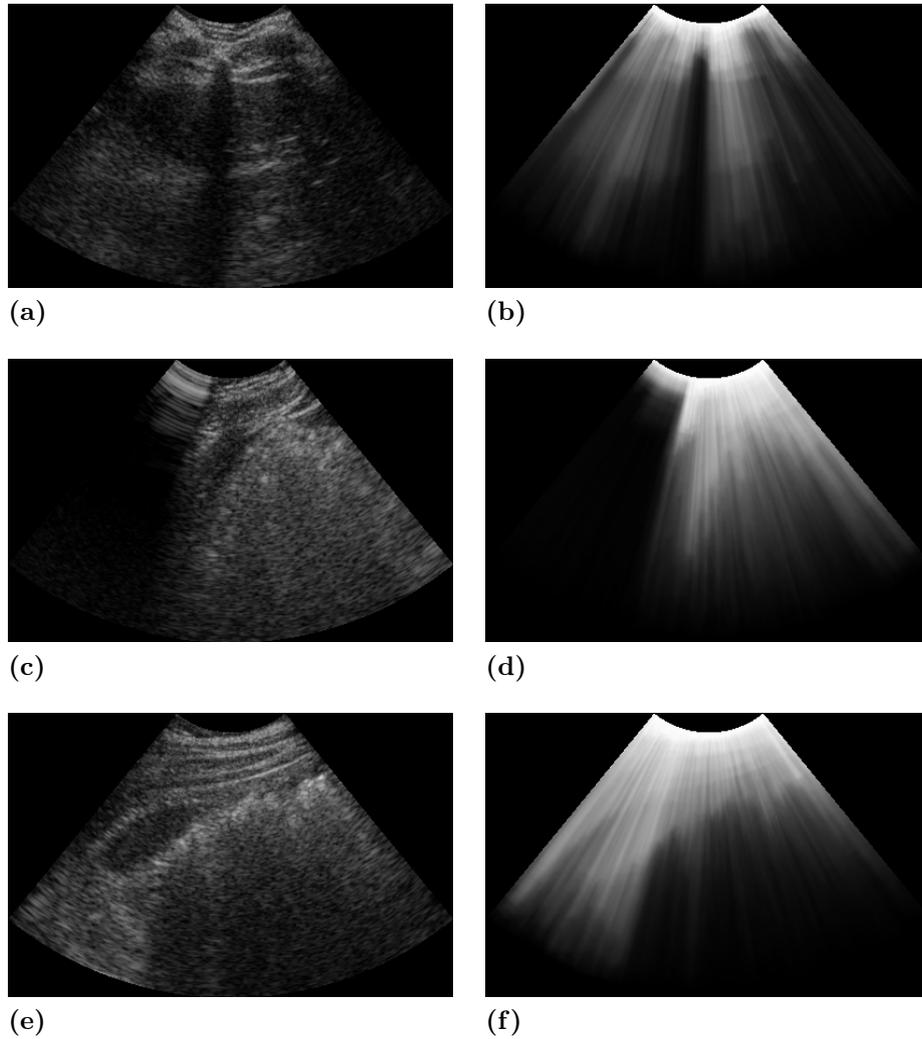


Figure 1.23 – Confidence maps estimated via the scan line integration method. Left: ultrasound image of a human body. Right: the corresponding confidence maps. From top to bottom: local shadow, drop-out artifact, diffuse shadow.

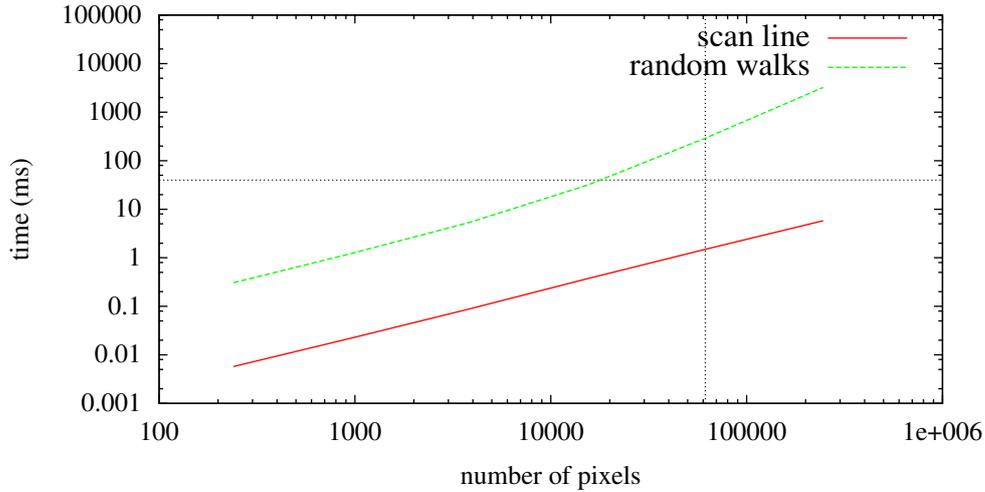


Figure 1.24 – Computation time for the random walks confidence map and the scan line integration methods. For reference, the dotted lines correspond to a standard acquisition rate of 25 frames per second and a prescan image size of 480×128 . Both axes are in logarithmic scale.

mentation, downsampling by a factor of 4 is required to allow real-time processing.

1.5.3.2 Temporal regularity

We have compared the random walks confidence map and the scan line integration methods in terms of computational complexity. Another important factor for the usability of the confidence map in a control framework is the smoothness of the estimated confidence values. The random walks confidence map has been designed so as to be spatially smooth. This is clearly visible when comparing Figures 1.23 and 1.22. However, the temporal regularity of the confidence map has not been studied up to now.

In order to measure the temporal regularity of the estimated confidence values, we propose two different measures. The first one corresponds to the temporal gradient of the confidence. The second one is based on the local standard deviation of the confidence values.

Temporal gradient The temporal gradient is measured by computing the absolute difference between confidence values in successive frames. Let $C(\mathbf{p}, t)$ be the confidence at a location \mathbf{p} and time t . The local

temporal gradient in \mathbf{p} at time t is defined as

$$G(\mathbf{p}, t) = |C(\mathbf{p}, t) - C(\mathbf{p}, t - 1)|, \quad (1.81)$$

which is then averaged over space and time, providing the mean temporal gradient

$$G = \frac{1}{|\Omega| \times (t_f - t_0)} \sum_{t=t_0+1}^{t_f} \iint_{\mathbf{p} \in \Omega} G(\mathbf{p}, t), \quad (1.82)$$

where $|\Omega|$ is the image size, t_0 is the starting time index, and t_f is the final time index.

Irregularity We define a second measure of irregularity, based on the local standard deviation of the confidence in a sliding time window. Let $C(\mathbf{p}, t)$ be the confidence at a location \mathbf{p} and time t . For a time window of size T , we define the local instantaneous standard deviation in \mathbf{p} at time t as

$$\sigma(\mathbf{p}, t) = \sqrt{\frac{1}{T} \sum_{i=0}^{T-1} (C(\mathbf{p}, t - i) - \mu(\mathbf{p}, t))^2}, \quad (1.83)$$

where

$$\mu(\mathbf{p}, t) = \frac{1}{T} \sum_{i=0}^{T-1} C(\mathbf{p}, t - i) \quad (1.84)$$

is the instantaneous mean. Then, we define the mean instantaneous standard deviation as

$$\sigma(t) = \frac{1}{|\Omega|} \iint_{\mathbf{p} \in \Omega} \sigma(\mathbf{p}, t), \quad (1.85)$$

where $|\Omega|$ is the image size. Finally, the mean standard deviation of an image sequence ranging from t_0 to t_f is defined as

$$\sigma = \frac{1}{t_f - t_0 - T + 2} \sum_{t=t_0+T-1}^{t_f} \sigma(t), \quad (1.86)$$

which we use as a measure of temporal irregularity.

The two measures defined above are used to compare the random walks and the scan line integration confidence estimation methods. We have estimated the confidence regularity on a sequence of 2159 ultrasound images acquired in vivo on a human volunteer. The time window for the σ irregularity measure was set to $T = 10$ images. The results are shown in Figure 1.25. The two measures consistently indicate that the scan

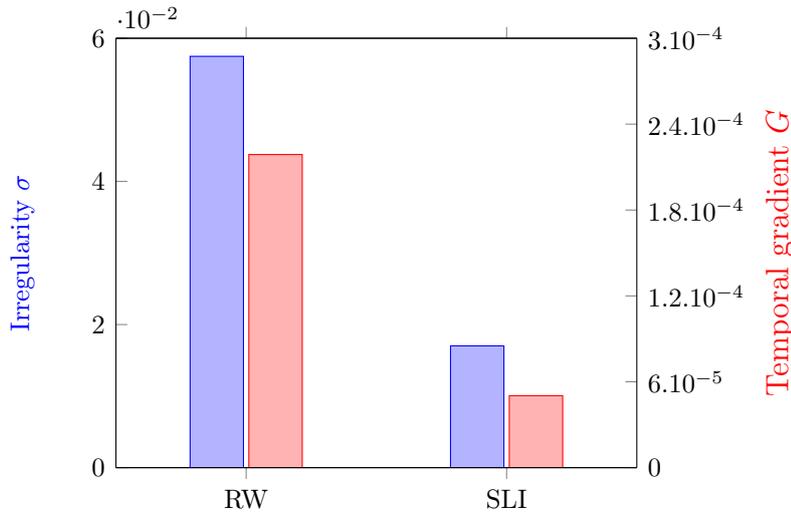


Figure 1.25 – Comparison of the temporal regularity between the random walks (RW) and scan line integration (SLI) confidence estimation methods.

line integration method results in temporally more regular confidence estimates than the random walks method. As a conclusion, while the random walks method may be more suited to compare confidence values within a single image, the scan line integration method provides more reliable confidence estimates over time.

1.6 Conclusion

We presented a review of ultrasound image analysis methods. This review is necessarily non-exhaustive, owing to the wide range of approaches that have been proposed in this field. We chose to focus on soft tissue and needle tracking algorithms with real-time capability. Indeed, this thesis relates to robotic ultrasound imaging, and real-time processing is a key requirement in order to be able to implement ultrasound-driven control solutions. Thus, the methods we have presented in this chapter constitute a basis for the design of ultrasound-based visual servoing approaches, which we describe in chapter 2.

In section 1.4, we have proposed a new method for tracking a flexible needle in 3D ultrasound images, using particle filtering. This approach provides a good control on the definition of the appearance and kinematic models of the needle.

In section 1.5, we have introduced the notion of ultrasound image quality, and we have presented different methods for estimating the quality of the ultrasound signal. We have introduced a simple but computa-

tionally inexpensive confidence map model, which we have compared with the random walks confidence map in terms of processing time and temporal regularity. As a result, it appears that the proposed scan line integration method provides temporally smoother confidence estimates than the random walks confidence map. We will use the concept of confidence maps in chapter 3 to design confidence-driven control approaches.

Chapter 2

Ultrasound-Based Visual Servoing

In chapter 1, we presented a review of ultrasound image analysis methods. These methods provide means to extract quantitative features from the ultrasound images. In this chapter, we address the topics of robotic ultrasound imaging and ultrasound-guided robot-assisted interventions. The overall goal of robot-assisted ultrasound is to help sonographers in performing diagnosis or surgery. More precisely, we focus in this chapter on visual servoing methods, where the content of the ultrasound image is used in order to control the robotic system.

Visual servoing refers to the closed-loop control of a dynamical system using feedback information provided by a visual sensor. Although ultrasound relies on acoustics, instead of vision, ultrasound-based control is generally assimilated to visual servoing. Indeed, the numerical representation of B-mode ultrasound data by an image creates an analogy with a camera image. An important difference, however, lies in the geometrical aspect of the two modalities. While a standard camera image corresponds to a perspective projection of the scene on a image plane, a B-mode image corresponds to a section of the body. Thus, although a wide range of methods designed for visual servoing can be transferred to ultrasound-based control, ultrasound imaging also has specificities that have to be taken into account.

We start with an introduction to the concept of visual servoing (section 2.1). Then, we present a review of ultrasound-based visual servoing methods for controlling an ultrasound probe (section 2.2) or a surgical needle (section 2.3). The notions introduced in this chapter will serve as a basis for the further developments proposed in chapter 3.

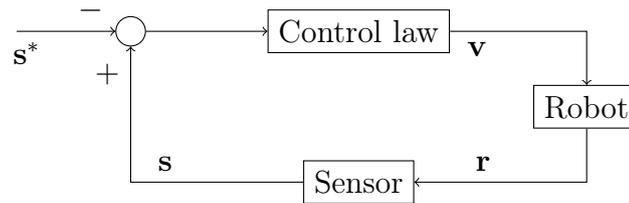


Figure 2.1 – Visual servoing control loop. The difference between the current visual feature \mathbf{s} and the desired visual feature \mathbf{s}^* is used to compute the control velocity \mathbf{v} . The visual feature \mathbf{s} depends on the configuration \mathbf{r} of the robot.

2.1 Visual servoing

This section provides an overview of the basic concepts of visual servoing. For a more detailed introduction to the topic, the interested reader is invited to refer to the two-parts tutorial from François Chaumette and Seth Hutchinson. The first part [Chaumette and Hutchinson, 2006] describes the basic approaches of image-based and position-based visual servoing, and discusses performance and stability issues. The second part [Chaumette and Hutchinson, 2007] addresses more advanced concepts related to the numerical estimation of the interaction matrix, and to the tracking of a moving target.

2.1.1 Introduction

Visual servoing consists in a closed control loop that involves a robot and a visual sensor (Figure 2.1). The sensor provides some visual information, which can be represented by a feature vector \mathbf{s} . This vector depends on the configuration of the robot, and visual servoing can be defined as providing the adequate input velocities to the robot, so as to obtain a desired visual feature \mathbf{s}^* .

We can distinguish two types of configurations, depending on the relation between the sensor and the robot. In the *eye-in-hand* configuration, the sensor is attached to the robot, so that its motion is directly guided by that of the robot. Thus, visual servoing typically consists in controlling the robot so that the visual sensor provides a desired view of an object of interest. In the *eye-to-hand* configuration, the sensor is fixed, and it observes externally the position of the end-effector of the robot as well as the scene. In this case, visual servoing consists in controlling the robot based on the external view of its configuration and its environment.

From the controller point of view, visual servoing can also be classified in two broad classes of approaches, depending on the nature of the visual

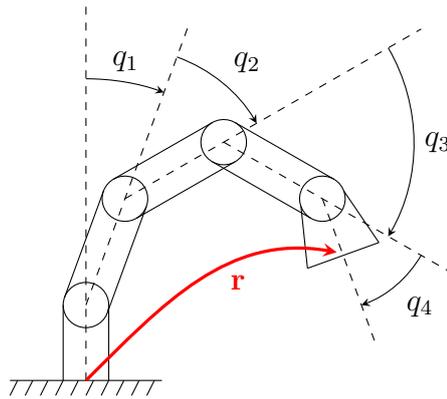


Figure 2.2 – Configuration of a robotic manipulator with 4 degrees of freedom. The pose \mathbf{r} of the end-effector depends on the configuration of all joints q_1, \dots, q_4 .

feature \mathbf{s} . In *position-based* visual servoing, the visual feature corresponds to the pose of the sensor with respect to some reference frame. In *image-based* visual servoing, the visual feature is defined directly in the image space. This second approach has the advantage of avoiding the need for a reconstruction of the robot’s pose from the visual data. On the other hand, the design of an appropriate control law can be more challenging.

2.1.2 Interaction matrix

Let us note $\mathbf{q}(t) \in \mathbb{R}^n$ the n -dimensional configuration of the robot’s actuators at time t . This vector can represent, for example, the angular position of the robot’s joints, for a manipulator with revolute joints (see Figure 2.2). The number n of actuators is referred to as the number of Degrees of Freedom (DOF) of the robot. For a given configuration $\mathbf{q}(t)$, the end-effector of the robot is in a pose $\mathbf{r}(\mathbf{q}, t) \in \text{SE}_3$, where SE_3 is the special Euclidean group, representing the configuration space of a rigid body. The pose \mathbf{r} can be defined, for instance, in a frame attached to the environment. The motion of the robot’s end-effector is represented by a velocity screw $\mathbf{v} \in \mathfrak{se}_3$, i.e., an element of the Lie algebra of equiprojective fields in \mathbb{R}^3 .

The basic property behind sensor-based control is that a sensor is defined by a differential mapping from SE_3 to \mathbb{R}^p , where p is the dimension of the sensor signal. Noting $\mathbf{s}(\mathbf{r}, t) \in \mathbb{R}^p$ the output of the sensor at time t , its variation with respect to time can be expressed as

$$\dot{\mathbf{s}} = \frac{\partial \mathbf{s}}{\partial \mathbf{r}} \mathbf{v} + \frac{\partial \mathbf{s}}{\partial t}. \quad (2.1)$$

Note that this equation relies on a simplification in notations, introduced in [Espiau et al., 1992], that consists in considering the velocity screw $\mathbf{v} \in \mathfrak{se}_3$ as a vector in \mathbb{R}^6 . This simplification amounts to fixing the frame in which the screw is expressed.

The term

$$\mathbf{L}_s = \frac{\partial \mathbf{s}}{\partial \mathbf{r}} \quad (2.2)$$

is the *interaction matrix*, of size $p \times 6$, that relates the variations of the sensor output \mathbf{s} to the velocity screw \mathbf{v} . The study of the interaction matrix is a key issue in the design of visual servoing approaches. The term $\frac{\partial \mathbf{s}}{\partial t}$ represents the contribution of scene modifications to the variation of \mathbf{s} . For a static environment, we have $\frac{\partial \mathbf{s}}{\partial t} = 0$, so that the variation of \mathbf{s} is completely defined by the interaction matrix via

$$\dot{\mathbf{s}} = \mathbf{L}_s \mathbf{v}. \quad (2.3)$$

Depending on the application, it is not unusual to assume that this relation holds even when the environment is not perfectly immobile, due to the difficulty of predicting the term $\frac{\partial \mathbf{s}}{\partial t}$. If predicting the evolution of the environment is important, e.g., for tracking a moving object, then it is necessary to estimate $\frac{\partial \mathbf{s}}{\partial t}$.

2.1.3 Visual servo control

Visual servo control consists in minimizing a p -dimensional visual error $\mathbf{e}(\mathbf{q}, t)$. This visual error depends on the configuration $\mathbf{q} \in \mathbb{R}^n$ of the robot, and the time t . It can be defined as the difference between the current visual feature vector $\mathbf{s}(\mathbf{r}(\mathbf{q}, t), t)$ and the desired feature vector $\mathbf{s}^*(t)$, i.e.,

$$\mathbf{e}(\mathbf{q}, t) = \mathbf{s}(\mathbf{r}(\mathbf{q}, t), t) - \mathbf{s}^*(t), \quad (2.4)$$

where $\mathbf{r}(\mathbf{q}, t) \in \text{SE}_3$ is the pose of the robot's end-effector.

From (2.1) and (2.4), we can express the variation of the visual error as

$$\dot{\mathbf{e}} = \mathbf{L}_s \mathbf{v} + \frac{\partial \mathbf{e}}{\partial t}, \quad (2.5)$$

where

$$\frac{\partial \mathbf{e}}{\partial t} = \frac{\partial \mathbf{s}}{\partial t} - \dot{\mathbf{s}}^* \quad (2.6)$$

represents the variation of \mathbf{e} due to modifications of the environment and to a change in the desired feature \mathbf{s}^* .

A typical approach to the regulation of the error is to enforce an exponential decay of \mathbf{e} . This constraint can be modeled by the first-order differential equation

$$\dot{\mathbf{e}} = -\lambda\mathbf{e}, \quad (2.7)$$

where $\lambda > 0$ is a control gain, that determines the convergence speed of the system. Incorporating (2.7) in (2.5) gives

$$-\lambda\mathbf{e} = \mathbf{L}_s\mathbf{v} + \frac{\partial\mathbf{e}}{\partial t}. \quad (2.8)$$

Let us first assume that the interaction matrix \mathbf{L}_s is squared ($p = 6$) and invertible. Then, the theoretic solution to (2.8) is

$$\mathbf{v} = -\mathbf{L}_s^{-1} \left(\lambda\mathbf{e} + \frac{\partial\mathbf{e}}{\partial t} \right). \quad (2.9)$$

In practice, however, it is impossible to know perfectly the interaction matrix \mathbf{L}_s or the task variation $\frac{\partial\mathbf{e}}{\partial t}$. Instead, the applied control law will be

$$\mathbf{v} = -\widehat{\mathbf{L}}_s^{-1} \left(\lambda\mathbf{e} + \frac{\widehat{\partial\mathbf{e}}}{\partial t} \right), \quad (2.10)$$

where the notation $\widehat{}$ denotes an approximation, or numerical estimation. An interesting property is that the local asymptotic stability of the system is ensured as long as the condition

$$\mathbf{L}_s\widehat{\mathbf{L}}_s^{-1} > 0 \quad (2.11)$$

is satisfied [Chaumette and Hutchinson, 2006]. As a result, local asymptotic stability can be guaranteed even when strong hypotheses are made in the approximation of \mathbf{L}_s .

If $p < 6$, the system has redundant degrees of freedom, so that it is possible to define a secondary task in addition to \mathbf{e} . This is further detailed in section 2.1.4. If $p > 6$, however, the system is underactuated, and it is in general not possible to fulfill the task, because there are too many visual constraints. In this case, a solution consists in using the Moore-Penrose pseudoinverse of the interaction matrix to solve (2.8) [Chaumette and Hutchinson, 2006]. The control law is then defined as

$$\mathbf{v} = -\widehat{\mathbf{L}}_s^+ \left(\lambda\mathbf{e} + \frac{\widehat{\partial\mathbf{e}}}{\partial t} \right), \quad (2.12)$$

where the Moore-Penrose pseudoinverse of a matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ of full rank m is defined as [Penrose, 1955]

$$\mathbf{A}^+ = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top. \quad (2.13)$$

2.1.4 Hybrid tasks

When a task constrains less degrees of freedom than available in the system, the system is said redundant, and it is possible to define a secondary task to be realized jointly with the first one. This is interesting, for instance, to mix different types of sensor data. Such combination of different tasks is addressed in the *redundancy formalism* [Samson et al., 1991], where the secondary task is defined in the null-space of the first task.

Let us consider a main task $\mathbf{e}_1 = \mathbf{s}_1 - \mathbf{s}_1^*$ defined from a feature vector $\mathbf{s}_1 \in \mathbb{R}^{p_1}$ with desired value \mathbf{s}_1^* . The variation of \mathbf{e}_1 is defined by the interaction matrix \mathbf{L}_1 , such that

$$\dot{\mathbf{e}}_1 = \mathbf{L}_1 \mathbf{v}, \quad (2.14)$$

where, for the sake of simplicity, we have considered $\frac{\partial \mathbf{e}_1}{\partial t} = 0$. Assuming that $p_1 < 6$, the general form of a control law realizing this task is

$$\mathbf{v} = -\lambda_1 \mathbf{L}_1^+ \mathbf{e}_1 + \mathbf{v}', \quad (2.15)$$

where \mathbf{L}_1^+ is the pseudoinverse of \mathbf{L}_1 , \mathbf{v}' is an element of $\ker(\mathbf{L}_1)$, the null-space of \mathbf{L}_1 , and λ_1 is a control gain.

\mathbf{v}' can be chosen so as to realize a secondary task \mathbf{e}_2 . As long as \mathbf{v}' lies in $\ker(\mathbf{L}_1)$, this secondary task will not disturb the main task. This is illustrated in a simple two-dimensional example in Figure 2.3. Let us note \mathbf{L}_2 the interaction matrix associated to the secondary task, such that $\dot{\mathbf{e}}_2 = \mathbf{L}_2 \mathbf{v}$. From (2.15), we have

$$\dot{\mathbf{e}}_2 = -\lambda_1 \mathbf{L}_2 \mathbf{L}_1^+ \mathbf{e}_1 + \mathbf{L}_2 \mathbf{v}', \quad (2.16)$$

where $-\lambda_1 \mathbf{L}_2 \mathbf{L}_1^+ \mathbf{e}_1$ is the contribution of the main task to the variation of \mathbf{e}_2 . \mathbf{v}' can be defined as

$$\mathbf{v}' = \mathbf{P}(\mathbf{L}_2 \mathbf{P})^+ (-\lambda_2 \mathbf{e}_2 - \lambda_1 \mathbf{L}_2 \mathbf{L}_1^+ \mathbf{e}_1), \quad (2.17)$$

where $\lambda_2 > 0$ is the control gain for the secondary task \mathbf{e}_2 , and

$$\mathbf{P} = \mathbf{I}_6 - \mathbf{L}_1^+ \mathbf{L}_1 \quad (2.18)$$

is an orthogonal projection operator on the null-space of \mathbf{L}_1 , so that $\mathbf{v}' \in \ker(\mathbf{L}_1)$. One can easily check that we have $\mathbf{L}_1 \mathbf{v}' = 0$ and $\mathbf{L}_2 \mathbf{v}' = -\lambda_2 \mathbf{e}_2$.

This framework can be extended to a hierarchical stack of tasks, where each task is constructed on top of the previous ones in an iterative manner [Siciliano and Slotine, 1991]. Let $(\dot{\mathbf{e}}_i^*, \mathbf{L}_i)_{i \in \llbracket 1, N \rrbracket}$ be a stack of N tasks, ordered according to priority, where $\dot{\mathbf{e}}_i^*$ is the desired variation of the i -th

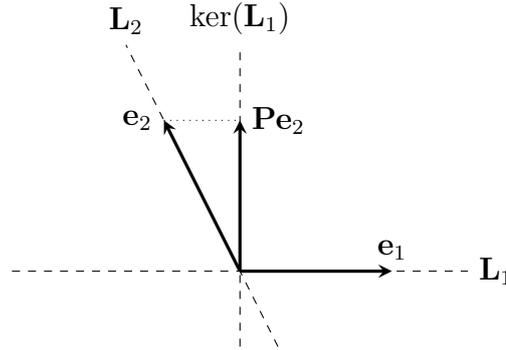


Figure 2.3 – Null-space projection of a secondary task, illustrated in a simple 2D case. By projecting the secondary task \mathbf{e}_2 onto the null-space $\ker(\mathbf{L}_1)$ of the first task \mathbf{e}_1 , the secondary task does not disturb the first task.

error, and \mathbf{L}_i its interaction matrix. A control law for this stack of tasks can be constructed recursively as

$$\mathbf{v}_0 = \mathbf{0}, \quad (2.19)$$

$$\mathbf{v}_i = \mathbf{v}_{i-1} + \mathbf{P}_{i-1}(\mathbf{L}_i \mathbf{P}_{i-1})^+ (\dot{\mathbf{e}}_i^* - \mathbf{L}_i \mathbf{v}_{i-1}), \quad (2.20)$$

with

$$\mathbf{P}_0 = \mathbf{I}_6, \quad (2.21)$$

$$\mathbf{P}_i = \mathbf{P}_{i-1} - \mathbf{P}_{i-1}(\mathbf{L}_i \mathbf{P}_{i-1})^+ \mathbf{L}_i \mathbf{P}_{i-1}. \quad (2.22)$$

A limitation of the redundancy formalism is that the secondary tasks are strongly constrained by the main task. [Mansard and Chaumette, 2009] propose another solution that releases some constraints on the secondary task, by enabling motions that could contribute to first task as well.

2.2 Ultrasound probe control

The development of robotized ultrasonography is motivated by its potential to improve the reliability and repeatability of examinations, but also by the fact that manual manipulation of an ultrasound probe is cumbersome. Indeed, several studies have reported a prevalence of musculoskeletal disorders among sonographers [Vanderpool et al., 1993; Magnavita et al., 1999]. The first robotic ultrasound systems were developed for the purpose of artery screening, in order to detect atheromatous plaques [Boudet et al., 1997; Pierrot et al., 1999]. In parallel, several solutions for robotized tele-echography have been proposed to allow specialists to

remotely examine patients located in medically isolated sites [De Cunha et al., 1998; Masuda et al., 2001; Vilchis et al., 2003; Delgorge et al., 2005]. Initially, these robotic systems were designed to follow the commands of a teleoperator, with only force control to maintain a desired contact force with the patient. With the advent of sophisticated real-time ultrasound image analysis techniques, it has then become possible to further assist the sonographer using image-guided control techniques.

The first application of the visual servoing framework to the control of an ultrasound probe was proposed by a team from the University of British Columbia in Vancouver [Salcudean et al., 1999]. The authors describe a system for teleoperated ultrasound examination of the carotid artery, where the control is shared between a human operator and the robot controller. First, teleoperation experiments with force and velocity control are reported. Then, the authors demonstrate the feasibility of tracking the carotid artery using a block matching technique based on normalized cross-correlation, and derive the interaction matrix for the coordinates of a point in the ultrasound image. An experimental validation of ultrasound-based visual servoing is proposed in [Abolmaesumi et al., 2002], using a phantom containing three plastic tubes. Various methods for tracking the tubes are compared. Using the tracked center of the tube sections, the authors servo the in-plane motion (3 DOF) of an ultrasound probe in order to maintain a selected pipe in the center of the image, while an operator moves the probe in the out-of-plane direction. We recall that *in-plane* and *out-of-plane* directions are defined with respect to the plane of imaging of a 2D probe, as illustrated in Figure 2.4.

[Krupa and Chaumette, 2006] propose a 6-DOF ultrasound-based visual servoing method to calibrate an ultrasound imaging system, using a cross-wire phantom. Geometric visual servoing is performed based on the detection of two points, corresponding to the intersection between the imaging plane and the two wires. [Krupa et al., 2007] use ultrasound speckle correlation to servo the out-of-plane motion of a 2D ultrasound probe for motion compensation. The in-plane motion tracking is based on the estimation of the rigid displacement of a region of interest. This method is interesting, because it allows motion compensation out of the imaging plane. However, the proposed speckle correlation technique is not robust to tissue deformations, and has a limited convergence domain. [Mebarki et al., 2008] propose a visual servoing approach based on 2D ultrasound image moments to control both the in-plane and out-of-plane motions (6 DOF) of a 2D ultrasound probe. Image moments up to order 3 are extracted from a tracked object contour and directly used as geometric features for visual servoing. [Nakadate et al., 2011] use a 1-DOF out-of-plane visual servoing approach to track a section of the carotid

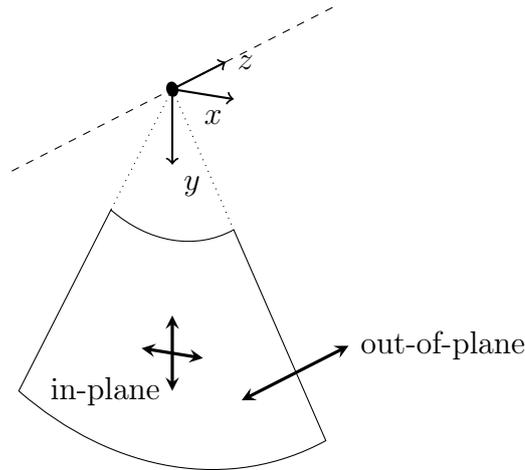


Figure 2.4 – In-plane and out-of-plane directions for a 2D ultrasound probe. In-plane corresponds to the x, y plane, i.e., the imaging plane of the probe. Out-of-plane corresponds to the z direction.

artery. The out-of-plane translation motion is detected using a block matching technique based on the SSD similarity measure. The method is validated on a human volunteer. In addition to motion compensation, it has been proposed to use visual servoing for tele-echography with shared control, to assist the teleoperator. [Krupa et al., 2016] use visual servoing to constrain the robot’s motion so as to maintain an object of interest within the field of view of the probe. In this setup, the robot is teleoperated by a clinician, an visual servoing is activated when the object of interest gets close to the image borders, in order to keep it from leaving the field of view. [Zettinig et al., 2016] propose a 3D registration-based visual servoing approach. 3D to 3D registration is used to find a transformation between the current ultrasound volume and a volume of interest defined pre-operatively. The rigid transformation found by the algorithm is directly used for servo control. The authors show a submillimeter positioning accuracy on a gelatin phantom. However, this method is limited by the low volume rate provided by the ultrasound probe.

The methods described above belong to the class of geometric visual servoing. These methods rely on the extraction of geometric features from the ultrasound image, using tracking or segmentation algorithms, or on the registration with a desired view (position-based visual servoing). The robust detection and tracking of anatomical landmarks can be challenging, in particular in the case of weakly structured ultrasound images. In order to avoid the need for image segmentation, different

intensity-based visual servoing approaches have been designed. [Nadeau and Krupa, 2013] propose an intensity-based ultrasound visual servoing method for 2D and 3D probes. The authors define the visual features directly as the vector of the pixel intensities, and derive the interaction matrix associated to the ultrasound image intensity. [Duflot et al., 2016] use a shearlet decomposition [Easley et al., 2008] of the ultrasound image instead of image intensities to design an ultrasound-based visual servoing method. The use of shearlet coefficients provides a sparse representation of the image contents, which increases the robustness to noise.

Thereafter, we introduce some general notations that will be used in the description of the control laws. Then, we describe the basic frameworks of geometric visual servoing (section 2.2.2) and intensity-based visual servoing (section 2.2.3), for the control of an ultrasound probe.

2.2.1 Notations

Let us first introduce some general conventions for the control of an ultrasound probe. Unless otherwise stated, we define the control velocity screw $\mathbf{v} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)$ in the *probe frame* \mathcal{F}_p attached to the imaging center of the probe, as illustrated in Figure 2.5. The terms v_x and v_y correspond to in-plane translations, and ω_z corresponds to the in-plane rotation (around the z -axis). The terms v_z , ω_x and ω_y correspond to out-of-plane motions. The same conventions are used for the case of a 3D probe, where \mathcal{F}_p is defined as the imaging center of the central frame.

2.2.2 Geometric visual servoing

Geometric visual servoing refers to the methods which rely on the extraction of geometric features from the ultrasound image, such as points, lines, or contours resulting from a segmentation of the image. The extraction of such features can be performed by the algorithms described in section 1.2. We start this section with the simple problem of controlling the in-plane motion of an ultrasound probe based on the position of a landmark.

2.2.2.1 Point-based in-plane control

The simplest approach to in-plane ultrasound-based visual servoing, initially proposed in [Salcudean et al., 1999], consists in using a point in the ultrasound image as feature for image-based visual servoing. Let \mathbf{p} be a point in the image with coordinates (u, v) . This image point corresponds to a physical point \mathbf{P} with 3D coordinates $(x_{\mathbf{p}}, y_{\mathbf{p}}, z_{\mathbf{p}})$ in the probe frame

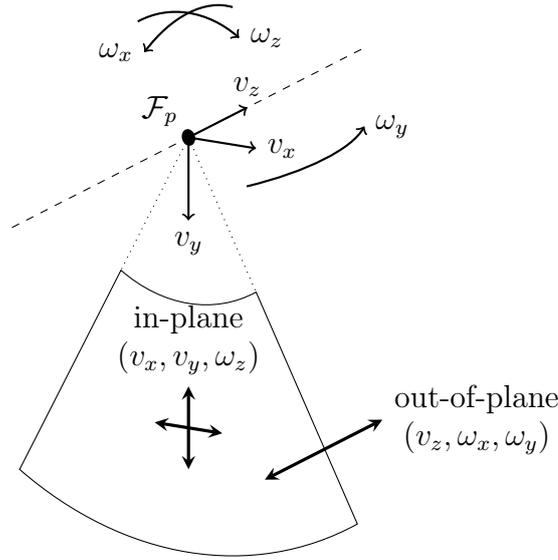


Figure 2.5 – Control frame attached to an ultrasound probe. The control frame \mathcal{F}_p is attached to the imaging center of the probe. The control screw components v_x, v_y, ω_z correspond to in-plan motions, and v_z, ω_x, ω_y correspond to out-of-plane motions.

(Figure 2.6). With the convention that the z -axis of the probe frame is orthogonal to the imaging plane, we have $z_{\mathbf{p}} = 0$ and $(x_{\mathbf{p}}, y_{\mathbf{p}}) = a(u, v)$, where a is the image scaling factor. Therefore, the variation of the point coordinates is linked to the velocity screw of the probe by the interaction matrix

$$\mathbf{L}_{\mathbf{p}} = \begin{bmatrix} -a & 0 & 0 & 0 & 0 & v \\ 0 & -a & 0 & 0 & 0 & -u \end{bmatrix}. \quad (2.23)$$

Following the classical visual servoing framework, the point \mathbf{p} can be brought to a desired position (u^*, v^*) in the image with the control law

$$\mathbf{v} = -\lambda \mathbf{L}_{\mathbf{p}}^+ \begin{pmatrix} u - u^* \\ v - v^* \end{pmatrix}, \quad (2.24)$$

where $\lambda > 0$ is a control gain.

2.2.2.2 3D point-based control

The previous method can be easily extended in the case where 3D ultrasound imaging is available. Let $\mathbf{p} = (u, v, w)$ a point in the ultrasound volume, expressed in voxel coordinates. This point corresponds to a physical point $\mathbf{P} = (x_{\mathbf{p}}, y_{\mathbf{p}}, z_{\mathbf{p}})$, whose coordinates are expressed in the probe frame. Assuming that the probe frame is defined in such a way that

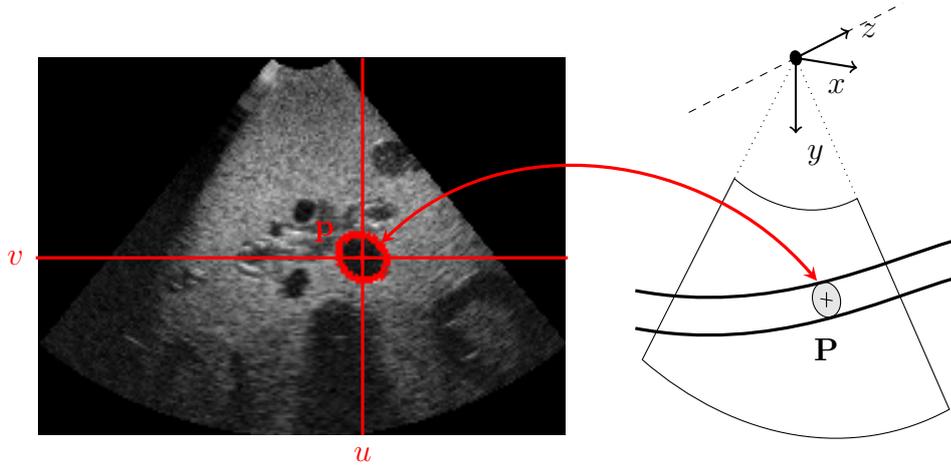


Figure 2.6 – Relation between a target detected in a 2D ultrasound image and a physical point.

$\mathbf{P} = a\mathbf{p}$, where a is a scaling factor, the interaction matrix associated to \mathbf{p} is

$$\mathbf{L}_{\mathbf{p}} = \begin{bmatrix} -a & 0 & 0 & 0 & -w & v \\ 0 & -a & 0 & w & 0 & -u \\ 0 & 0 & -a & -v & u & 0 \end{bmatrix}. \quad (2.25)$$

Noting (u^*, v^*, w^*) the desired position of the point in the volume, the visual servoing control law can be defined as

$$\mathbf{v} = -\lambda \mathbf{L}_{\mathbf{p}}^+ \begin{pmatrix} u - u^* \\ v - v^* \\ w - w^* \end{pmatrix}, \quad (2.26)$$

where $\lambda > 0$ is a control gain.

However, 3D ultrasound imaging may not always be available, or its imaging quality might be too low for the application at hand. Therefore, the out-of-plane control of a 2D ultrasound probe is also of interest. The main limitation of 2D ultrasound imaging is that it does not provide any direct information on the probe positioning out of the imaging plane. For this reason, alternative methods have been developed to control the out-of-plane motion of the probe. In the following, we detail a method based on 2D image moments [Mebarki et al., 2010].

2.2.2.3 Moments-based control

Let us assume that some algorithm provides a segmentation of an object of interest in the ultrasound image. We note \mathcal{S} the surface covered by

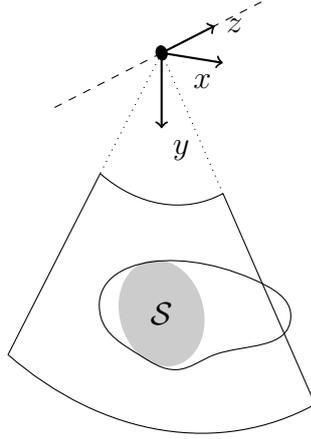


Figure 2.7 – Intersection between an object and the image plane. Geometric moments are extracted from the surface of intersection \mathcal{S} .

the object in the image (Figure 2.7). The 2D moments of \mathcal{S} of order $i + j$, with $i, j \in \mathbb{N}$, are defined as [Mukundan and Ramakrishnan, 1998]

$$m_{i,j} = \iint_{\mathcal{S}} x^i y^j dx dy, \quad (2.27)$$

where (x, y) are the image coordinates of points in \mathcal{S} . For instance, the moment of order 0 of \mathcal{S} corresponds to the area

$$a = m_{0,0} = \iint_{\mathcal{S}} 1 dx dy \quad (2.28)$$

of \mathcal{S} , and the moments of order 1 are related to the coordinates (x_g, y_g) of its barycenter via

$$x_g = \frac{m_{1,0}}{m_{0,0}} \quad (2.29)$$

and

$$y_g = \frac{m_{0,1}}{m_{0,0}}. \quad (2.30)$$

To control the 6 DOF of the robot-held probe, [Mebarki et al., 2010] propose to use the visual features

$$\mathbf{s} = (x_g, y_g, \alpha, \sqrt{a}, \phi_1, \phi_2), \quad (2.31)$$

where x_g, y_g and a have already been defined in equations (2.28) to (2.30). The feature α corresponds to the main orientation of the object in the image, and is defined as

$$\alpha = \frac{1}{2} \arctan \left(\frac{2\mu_{1,1}}{\mu_{2,0} + \mu_{0,2}} \right), \quad (2.32)$$

where

$$\mu_{i,j} = \iint_S (x - x_g)^i (y - y_g)^j dx dy \quad (2.33)$$

is a centered image moment of order $i + j$. The features ϕ_1 and ϕ_2 are defined from the centered moments of order 2 and 3 and are invariant to image translation and rotation.

The interaction matrix associated to the visual feature \mathbf{s} is of the form

$$\mathbf{L}_s = \begin{bmatrix} -1 & 0 & x_{gvz} & x_{g\omega_x} & x_{g\omega_y} & y_g \\ 0 & -1 & y_{gvz} & y_{g\omega_x} & y_{g\omega_y} & -x_g \\ 0 & 0 & \alpha_{v_z} & \alpha_{\omega_x} & \alpha_{\omega_y} & -1 \\ 0 & 0 & \frac{a_{v_z}}{2\sqrt{a}} & \frac{a_{\omega_x}}{2\sqrt{a}} & \frac{a_{\omega_y}}{2\sqrt{a}} & 0 \\ 0 & 0 & \phi_{1v_z} & \phi_{1\omega_x} & \phi_{1\omega_y} & 0 \\ 0 & 0 & \phi_{2v_z} & \phi_{2\omega_x} & \phi_{2\omega_y} & 0 \end{bmatrix}. \quad (2.34)$$

Without detailing the expression of all the elements in \mathbf{L}_s , we can see that \sqrt{a} , ϕ_1 and ϕ_2 are invariant to the in-plane motions (v_x, v_y, ω_z) . The terms in the third, fourth and fifth columns of \mathbf{L}_s depend on 3D information about the observed object. More precisely, these terms require an estimation of the normal vector to the object's surface along the contour resulting from the intersection between the image plane and the object. In practice, the normal vectors are estimated online from successive frames, using the robot's odometry. This method results in a local reconstruction of the object's surface.

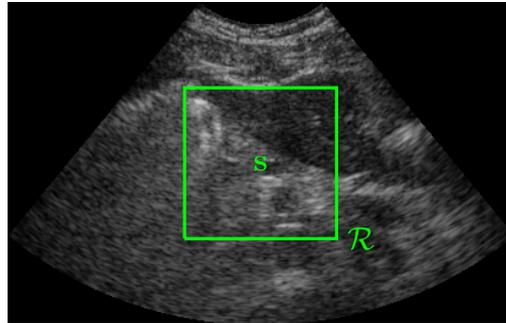
2.2.3 Intensity-based visual servoing

We have seen that geometric visual servoing requires a segmentation or tracking step to extract appropriate geometric features to be used in the control law. An interesting alternative is to define the visual feature \mathbf{s} directly as a vector of ultrasound image intensities [Nadeau and Krupa, 2013]. In this case, the visual information is *dense*. We start by describing the method for the in-plane control of a 2D ultrasound probe. Then, we show how this method extends to the full control of a 3D ultrasound probe. Finally, we present an approach to control the out-of-plane motion of a 2D ultrasound probe.

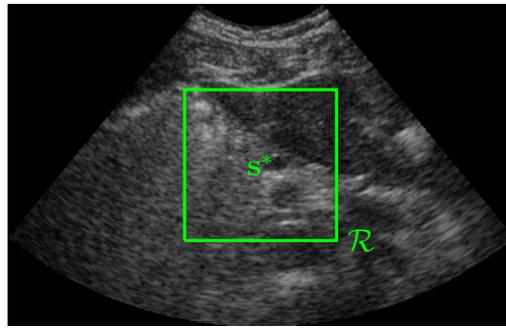
Let $I : \Omega \rightarrow \mathbb{R}$ represent the current ultrasound image. We consider a fixed region of interest $\mathcal{R} \subset \Omega$ containing n pixels, and we define the visual feature

$$\mathbf{s} = (I(\mathbf{p}))_{\mathbf{p} \in \mathcal{R}} \quad (2.35)$$

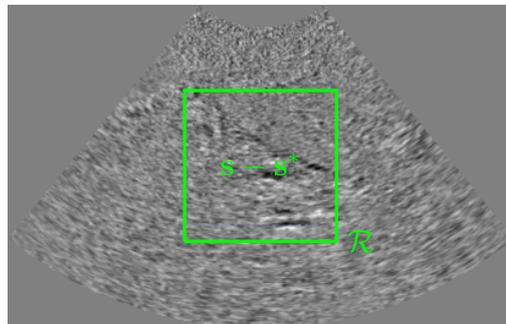
as the vector of ultrasound intensities $I(\mathbf{p})$ for all pixels \mathbf{p} in \mathcal{R} , as illustrated in Figure 2.8.



(a)



(b)



(c)

Figure 2.8 – Dense visual feature for intensity-based visual servoing. The visual feature \mathbf{s} is defined as the vector of intensities in a region \mathcal{R} . (a) Current image. (b) Desired image, defining the desired feature \mathbf{s}^* . (c) Image difference, defining the feature error $\mathbf{s} - \mathbf{s}^*$.

The derivation of the interaction matrix associated to \mathbf{s} requires the expression of the time derivative of $I(\mathbf{p})$. The ultrasound intensity at \mathbf{p} corresponds to the intensity $I_{\text{US}}(\mathbf{x}, t)$ originating from a physical point \mathbf{P} with coordinates $\mathbf{x} = (x, y, z)$ in the probe frame. The total derivative of I_{US} is

$$\dot{I}_{\text{US}}(\mathbf{x}, t) = \left(\frac{\partial I_{\text{US}}(\mathbf{x}, t)}{\partial \mathbf{x}} \right)^\top \dot{\mathbf{x}} + \frac{\partial I_{\text{US}}(\mathbf{x}, t)}{\partial t} \quad (2.36)$$

where

$$\frac{\partial I_{\text{US}}(\mathbf{x}, t)}{\partial \mathbf{x}} = \nabla I(\mathbf{p})^\top = (\nabla I_x, \nabla I_y, \nabla I_z)^\top \quad (2.37)$$

is the image gradient, $\dot{\mathbf{x}}$ is the variation of the point's coordinates, and $\frac{\partial I_{\text{US}}(\mathbf{x}, t)}{\partial t}$ is the variation of the intensity originating from \mathbf{P} . This intensity variation is hardly predictable, since it depends on the physical properties of the tissues lying on the path between the point \mathbf{P} and the ultrasound transducer. A standard hypothesis consists in assuming that the ultrasound intensity reflected from a physical point is constant, i.e.,

$$\frac{\partial I_{\text{US}}(\mathbf{x}, t)}{\partial t} = 0. \quad (2.38)$$

This hypothesis, which can seem rather strong *a priori*, has been validated empirically in [Nadeau and Krupa, 2013]. Using (2.38) in (2.36), we obtain

$$\dot{I}_{\text{US}}(\mathbf{x}, t) = \nabla I(\mathbf{p}) \dot{\mathbf{x}}. \quad (2.39)$$

The variation of \mathbf{x} can be obtained from Varignon's formula for the composition of velocities in a solid. Noting \mathbf{v} the velocity screw of the end-effector in the probe frame, we have

$$\dot{\mathbf{x}} = \begin{bmatrix} 1 & 0 & 0 & 0 & z & -y \\ 0 & 1 & 0 & -z & 0 & x \\ 0 & 0 & 1 & y & -x & 0 \end{bmatrix} \mathbf{v}. \quad (2.40)$$

2.2.3.1 Case of a 2D probe

For a 2D ultrasound probe, the observed section lies in a plane with $z = 0$. As a result, the interaction matrix associated to $I(\mathbf{p})$ is obtained from (2.39) and (2.40) as

$$\mathbf{L}_{I(\mathbf{p})} = \left[\nabla I_x \quad \nabla I_y \quad \nabla I_z \quad y \nabla I_z \quad -x \nabla I_z \quad x \nabla I_y - y \nabla I_x \right], \quad (2.41)$$

where ∇I_x and ∇I_y can be computed directly from the image. The term ∇I_z , corresponding to the intensity gradient in the direction orthogonal to the imaging plane, is not directly available. [Nadeau and Krupa, 2013] propose two different approaches to estimate ∇I_z :

- By applying a small back-and-forth translation motion in the out-of-plane direction, one can compute ∇I_z from a set of parallel frames. This approach can be useful for a motion compensation task, when the initial image is the desired one.
- Alternatively, the gradient can also be estimated online using the difference in intensities between successive frames that are not coplanar. From these intensity differences and the robot's odometry, one can estimate the gradient via a least squares methods.

The full interaction matrix associated to \mathbf{s} is obtained by stacking the 1×6 interaction matrices for all pixels $\mathbf{p}_1, \dots, \mathbf{p}_n$ in \mathcal{R} :

$$\mathbf{L}_s = \begin{bmatrix} L_{I(\mathbf{p}_1)} \\ \vdots \\ L_{I(\mathbf{p}_n)} \end{bmatrix}. \quad (2.42)$$

2.2.3.2 Case of a 3D probe

With a 3D probe, the region of interest \mathcal{R} is a volume, and the three components of the gradient are directly available from the ultrasound volume. In this case, the interaction matrix for a point \mathbf{p} with coordinates (x, y, z) is

$$\mathbf{L}_{I(\mathbf{p})} = \begin{bmatrix} \nabla I_x & \nabla I_y & \nabla I_z \\ y\nabla I_z - z\nabla I_y & z\nabla I_x - x\nabla I_z & x\nabla I_y - y\nabla I_z \end{bmatrix}. \quad (2.43)$$

2.3 Ultrasound-guided needle control

In section 2.2, we reviewed methods for controlling an ultrasound probe held by a robotic system. The main applications of these methods are tele-echography, motion compensation, and (semi-)autonomous ultrasound acquisition. Another useful configuration using ultrasound-based visual servoing is to have a robotized surgical instrument guided by ultrasound imaging. This corresponds to an eye-to-hand visual servoing framework. In particular, as already mentioned in chapter 1, ultrasound-guided robotic needle guidance is of interest to improve the precision of minimally invasive surgical procedures such as biopsy or localized tumor ablation.

Robot-assisted needle insertion under ultrasound guidance has gained interest from the mid 2000s, and has now become an important subject of research in the medical robotics community. The first ultrasound-guided robotic needle insertion method was presented in [Hong et al., 2004].

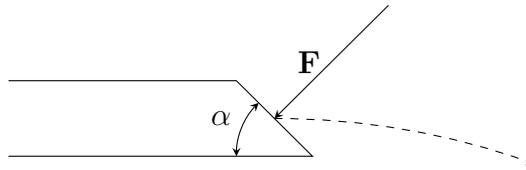


Figure 2.9 – Schematic of a bevel-tip needle inserted in soft tissues. The asymmetry of the contact forces at the tip results in a bending of the needle. The radius of curvature depends on the bevel angle α , and on the stiffness of surrounding tissues.

The authors propose a robotic system to drive a rigid needle towards the gallbladder under 2D ultrasound guidance. The considered clinical application is percutaneous cholecystostomy, where a precise alignment with respect to the preoperative plan is of critical importance. This method requires the rigid needle to be aligned with the ultrasound imaging plane. [Wei et al., 2005] design a robot-assisted insertion system under 3D transrectal ultrasound guidance for prostate brachytherapy. The use of 3D ultrasound allows oblique needle insertion trajectories, which provides more control on the positioning of the needle. [Novotny et al., 2007a] use 3D ultrasound-based visual servoing to guide a surgical instrument with passive markers to ease its detection in the ultrasound images. Experiments are performed in a water bath. The targeted application is cardiac surgery.

While the previous approaches consider a straight insertion path, needles can be subject to deformation due to contact forces with the surrounding tissues. This effect is particularly strong for flexible needles, which can undergo relatively large deformations during the insertion. An interesting case is that of bevel-tip needles (Figure 2.9), where the contact forces at the tip are asymmetrical, leading to a natural bending [Misra et al., 2010].

2.3.1 Needle steering

Needle steering, such as proposed in [DiMaio and Salcudean, 2005], consists in taking advantage of the needle's flexibility to follow curvilinear insertion trajectories. The main advantage of this approach is that the needle can be guided precisely towards its target without having to perform multiple retractions and re-insertions to correct its position. As a result, the duration of the intervention can be reduced. In addition, flexible needle steering allows the system to reach otherwise unattainable targets by avoiding obstacles such as vessels or bones [Alterovitz et al., 2005; Bernardes et al., 2014], as illustrated in Figure 2.10.

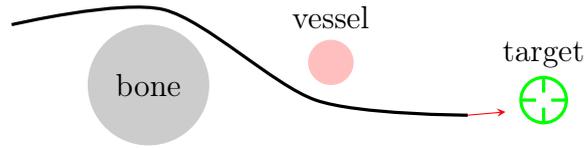


Figure 2.10 – Needle steering for obstacle avoidance. By steering a flexible needle, one can reach a target that would be otherwise unattainable.

Base control The approach initially proposed to steer a flexible needle was to manipulate it from its base, in order to induce a deformation of the tissues and a bending of the needle [DiMaio and Salcudean, 2005]. This method requires a model of the interaction forces between the needle and the tissues. Such a model should also be simple enough to be updated in real-time. [Glozman and Shoham, 2007] propose a virtual springs model to estimate the needle-tissues interactions locally.

Duty cycling Controlling the needle from its base offers a good steerability near the body surface. However, because of the deformations applied on the tissues, this technique can potentially damage the tissues when the needle is inserted deeper. An alternative approach consists in exploiting the natural bending property of bevel-tip needles. The kinematic model generally adopted for bevel-tip needle steering is that of a nonholonomic bicycle model [Webster III et al., 2006]. A desired needle trajectory can be achieved thanks to a duty-cycling control strategy [Minhas et al., 2007], which enables a precise control of the needle curvature. Duty-cycling consists in applying alternately two types of motions to the needle:

- Pure insertion, where the needle follows an arc of natural curvature K [Figure 2.11(a)].
- Simultaneous insertion and rotation around the needle axis, which results in a helical trajectory. If the rotation velocity is fast enough compared to the insertion velocity, then the trajectory can be approximated to a straight line [Figure 2.11(b)].

As a result, one can obtain a desired effective curvature K_{eff} of the needle by alternating the two phases with a specific ratio between the duration of each phase. Noting T_{trans} the duration of pure insertion and T_{rot} the duration of the combined insertion/rotation, the duty-cycle ratio is defined as

$$DC = \frac{T_{\text{rot}}}{T_{\text{rot}} + T_{\text{trans}}}, \quad (2.44)$$

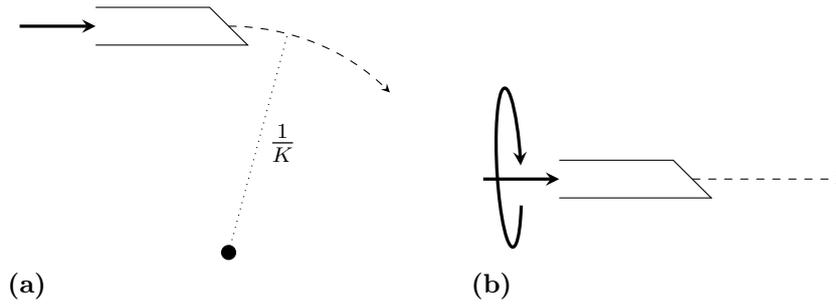


Figure 2.11 – Insertion types for bevel-tip needles. (a) Under pure insertion, the needle tip follows a circular trajectory with curvature K . (b) Under simultaneous insertion and rotation, the needle follows an helical trajectory, that can be approximated to a straight line if the rotation velocity is high enough.

and the resulting motion can be approximated to a circular trajectory with effective curvature

$$K_{\text{eff}} = K(1 - DC), \quad (2.45)$$

where K is the natural curvature of the needle. Note that we always have $K_{\text{eff}} < K$, so that the duty-cycling technique cannot increase the curvature of the insertion path. During the pure insertion phase, the bending direction can be adjusted by controlling the orientation of the tip bevel.

Recently, [Chevrie et al., 2016a] proposed to fuse duty-cycling with base control to exploit the advantages of both methods. The authors use base control at the beginning of the insertion, in order to correct the orientation of the needle tip. When the needle reaches a certain depth, the system switches to duty-cycling for the final approach, in order to reduce the risk of damaging tissues. In [Chevrie et al., 2016b], the authors extend this method with visual feedback to estimate the displacement of the tissues. The estimated displacement is used to update the needle/tissue interaction model, which leads to a higher robustness to tissue motion.

Actuated needles The steering methods we have presented above rely on the properties of conventional needles. Alternatively, some actuated needle systems have been recently proposed. These provide an enhanced steerability of the needle, while reducing the risk of tissue damage. [Okazawa et al., 2005] propose a steerable needle system based on a curved-tip stylet embedded in a cannula. The geometry of the stylet provides a way to steer the needle by controlling the portion of the stylet that extends out of the cannula. [Sears and Dupont, 2006] introduce a

concentric tubes design that allows more control on the needle shape. Using a series of curved concentric tubes, the needle can be steered by controlling the orientation and extension of each tube. The authors also propose a mechanics model for estimating the needle shape and computing forward and inverse dynamics. [Ko and Rodriguez y Baena, 2013] propose a biologically-inspired steerable needle composed of adjacent segments that can slide along one another. By controlling the respective motion of each segment, the needle can be steered along arbitrary trajectories. [Roesthuis et al., 2015] present a tip actuation system that provides control on the orientation of the needle tip. This system consists in a tendon-actuated conical tip mounted on a ball joint. [Shahriari et al., 2016] use Fiber Bragg Grating (FBG) sensor data and ultrasound images to track the position of an actuated-tip flexible needle.

2.3.2 Path planning

Various methods have been developed using needle steering models to plan the needle insertion trajectory. Most approaches use inverse kinematics for preoperative path planning [Duindam et al., 2010]. However, preoperative path planning is subject to errors due to tissue deformations. Some methods have been proposed to adjust the needle path intraoperatively. [Wood et al., 2010] use a succession of planar arc-trajectories to steer the needle in a 2D plane. The current position of the needle tip is estimated from camera images thanks to an extended Kalman filter. Obstacle avoidance can be achieved using probabilistic methods, such as the Rapidly-Exploring Random Tree (RRT) algorithm [LaValle and Kuffner, 1999]. [Xu et al., 2008] use the RRT algorithm and back-chaining to steer the needle in a 3D environment with obstacles. [Bernardes et al., 2013] combine RRT-based path planning with a duty-cycling control strategy to steer a needle in a 2D plane with closed-loop visual feedback from a camera. The authors have later adapted this method to steer a needle along a 3D trajectory composed of a succession of 2D planar arcs [Bernardes et al., 2014]. This method was tested in a simulation environment. [Dorileo et al., 2015] use a model of needle and tissue stiffness to predict the deflection of the needle tip, and perform adaptive path planning under CT/MRI guidance. [Krupa, 2014] proposes a new duty-cycling approach which permits non-planar 3D trajectories. Based on the visual servoing framework, this approach does not require any path planning. Visual feedback was provided by an optical camera observing the needle in a translucent phantom.

2.3.3 Ultrasound-guided needle steering

The control strategies mentioned above have been tested in simulation, or with visual feedback provided by a camera observing the needle in a translucent phantom. [Dorileo et al., 2015] use CT/MRI guidance. Recent studies have tackled the problem of using ultrasound imaging to track the needle during robotic insertion. We can distinguish three different approaches to ultrasound-guided needle insertion, depending on the probe type and its orientation with respect to the needle:

- A 2D ultrasound probe oriented so as to observe the needle trajectory. This configuration provides a mean to estimate the position of the needle shaft and the target, provided that the needle has a straight (or, at least, planar) trajectory, and that the target is coplanar with the needle trajectory.
- A 2D ultrasound probe placed perpendicular to the needle insertion direction, so as to visualize the needle tip. This approach has the advantage that the detection of the needle tip is easier. However, it requires to control the position of the probe to ensure that the imaging plane always intersects the needle at its tip. In addition, this technique requires a mean to estimate the shaft orientation, and does not allow to visualize the target.
- A 3D ultrasound probe. With this type of probe, the shape of the needle can be directly extracted from the ultrasound volume, even for deformable needles.

Thereafter, we provide a short review of the methods proposed for each of these configurations.

2.3.3.1 2D probe, in-plane needle

This configuration was the one initially considered in [Hong et al., 2004] for rigid needles. [Neubach and Shoham, 2010] perform a closed-loop insertion of a flexible needle. The authors estimate the stiffness of tissues and model the needle with virtual springs. This physical model is used to compute the manipulations to apply to the needle base in order to steer it in a desired direction, based on an inverse kinematics algorithm. [Kojcev et al., 2016] use two independent robots to control the position of the ultrasound probe and the needle insertion.

2.3.3.2 2D probe, out-of-plane needle

[Vrooijink et al., 2014] place a 2D ultrasound transducer perpendicular to the direction of insertion, and automatically re-position the transducer during insertion so that the needle tip stays in the imaging plane. In this work, the needle is steered in 3D with moving obstacle avoidance using the RRT algorithm and a duty-cycle controller. In this configuration, the 2D ultrasound image does not display the needle shaft, so that the orientation of the needle tip has to be estimated by integrating the displacements of the detected tip in the ultrasound image plane and the out-of-plane motion of the 2D ultrasound probe. [Abayazid et al., 2014] propose an experimental evaluation of this method in a gelatin phantom with moving obstacles, and in chicken breast. [Waine et al., 2016] use a model of needle-tissues interactions to predict the needle deflection during insertion. [Abayazid et al., 2016] perform three-dimensional needle steering using the Automated Breast Volume Scanner (ABVS) [Wojcinski et al., 2011], with obstacle avoidance. In this case, the scanning velocity is constant, and the needle tip velocity is adapted in order to keep the needle tip in the imaging plane. The needle trajectory is replanned every second using the RRT algorithm.

2.3.3.3 3D probe

[Hungu et al., 2012] use 3D ultrasound imaging to track the motion of the prostate, and propose a robotic needle manipulator to align insertion path of the needle with a target defined in the ultrasound volume. Needle insertion is done by iterative re-planning until the target is reached. [Adebar et al., 2014] use 3D Doppler ultrasound to detect a bent-tip needle which is vibrated, thus highlighting the needle shape in the Doppler image. The needle is steered into *ex vivo* bovine tissues towards a fixed target using a duty-cycling controller with intraoperative replanning based on the detected needle pose.

2.3.3.4 3D US-guided needle steering via visual servoing

Most of the methods presented above rely on an iterative replanning strategy to update the desired needle trajectory. We now propose a visual servoing approach, where needle steering is performed in closed loop under 3D ultrasound guidance. We have already presented a method for tracking a flexible needle in ultrasound images via particle filtering (section 1.4). In this section, we detail our needle steering approach, which relies on the detection of the needle shaft.

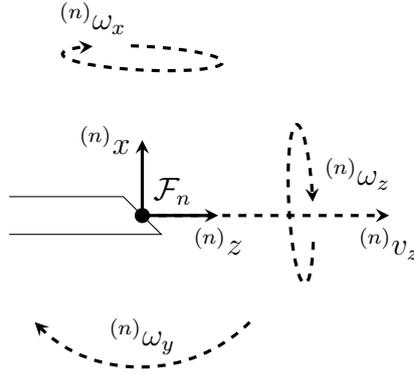


Figure 2.12 – Control velocities in the needle tip frame \mathcal{F}_n .

Duty-cycling We control continuously (at the rate of the duty-cycle period) the two lateral angular velocities ${}^{(n)}\omega_x$, ${}^{(n)}\omega_y$ and the insertion velocity ${}^{(n)}v_z$ of the needle in the needle tip frame \mathcal{F}_n , represented in Figure 2.12. These velocities are linked by the equations

$${}^{(n)}\omega_x = \cos(\theta)K_{\text{eff}}{}^{(n)}v_z, \quad (2.46)$$

$${}^{(n)}\omega_y = \sin(\theta)K_{\text{eff}}{}^{(n)}v_z, \quad (2.47)$$

where θ is an accumulation angle applied to the needle around its axis during the current duty-cycle period, such that the absolute angle θ_t obtained at the end of the cycle becomes $\theta_t = \theta_{t-T} + 2\pi + \theta$. Therefore, controlling the effective curvature K_{eff} acts on the amplitude of ${}^{(n)}\omega_x$ and ${}^{(n)}\omega_y$, while controlling θ acts on the ratio between ${}^{(n)}\omega_x$ and ${}^{(n)}\omega_y$.

Visual features The control task consists in guiding the needle towards a user-defined target $\mathbf{O} \in \mathbb{R}^3$ whose coordinates are expressed in the ultrasound volume frame \mathcal{F}_{us} . We use the coordinates of the needle tip $\mathbf{N} = \mathbf{C}_t^{(n)}$ detected by the SIR-based needle tracker described in section 1.4, as well as the unitary direction vector $\mathbf{u}_t(1)$ defined in (1.64), to compute the visual features $\mathbf{s} = [l_{xz} \ l_{yz}]^\top$. These features correspond to the distance l_{xz} (resp. l_{yz}) between the target point \mathbf{O} and its projection onto the needle direction line projected on the x - z (resp. y - z) plane of \mathcal{F}_{us} [Figure 2.13(a)]. They can be computed from measurements made in the ultrasound volume, via the relations

$$l_{xz} = d_{xz}\sin(\theta_y) \quad (2.48)$$

and

$$l_{yz} = -d_{yz}\sin(\theta_x), \quad (2.49)$$

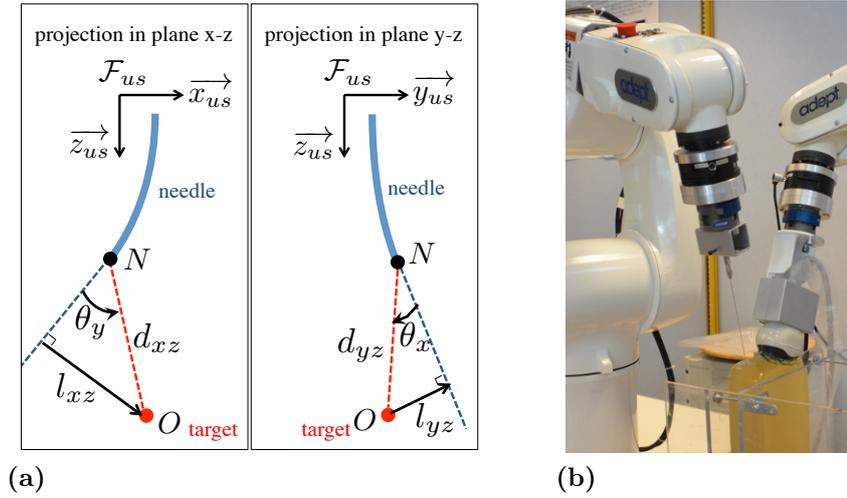


Figure 2.13 – Visual features for US-based needle steering. (a) Projections of the needle and target in the 3D image frame. (b) Experimental setup.

where d_{xz} and d_{yz} denote the length of the projection of the vector \overrightarrow{NO} on the x - z and y - z plane, respectively, and θ_y and θ_x are the projected angles between the direction vector of the needle at its tip and \overrightarrow{NO} . Note that, since these quantities can be measured directly in the ultrasound volume, no calibration is required between the robot frame and the volume frame.

Control law Finally, we use the control law

$$\begin{bmatrix} {}^{(us)}\omega_x \\ {}^{(us)}\omega_y \end{bmatrix} = -\lambda \mathbf{J}_{\omega_{xy}}^{-1} (\mathbf{s} - \mathbf{s}^*) \quad (2.50)$$

that was proposed in [Krupa, 2014], where $\mathbf{s}^* = [0 \ 0]^\top$ is the desired value of the visual feature vector, $\lambda > 0$ is the control gain, and $\mathbf{J}_{\omega_{xy}}$ is the Jacobian that relates the variations of the visual features to the needle's lateral velocities ${}^{us}\omega_x$ and ${}^{us}\omega_y$ expressed in the ultrasound volume frame \mathcal{F}_{us} . This Jacobian only depends on measures extracted from the ultrasound volume by the needle detector:

$$\mathbf{J}_{\omega_{xy}} = \begin{bmatrix} 0 & d_{xz} \cos(\theta_y) \\ -d_{yz} \cos(\theta_x) & 0 \end{bmatrix}. \quad (2.51)$$

In practice, we then express the lateral control velocities ${}^{(n)}\omega_x$ and ${}^{(n)}\omega_y$ of the needle's tip in the frame \mathcal{F}_n , thanks to a velocity transformation that depends on the direction vector $\mathbf{u}_t(1)$ defined in (1.64) and the angle θ_t , provided respectively by the tracking algorithm and the

odometry of the robot. This provides the effective curvature K_{eff} and the accumulation angle θ to use in the duty-cycling control:

$$K_{\text{eff}} = \frac{1}{({}^n)v_z} \sqrt{{}^n\omega_x^2 + {}^n\omega_y^2}, \quad (2.52)$$

$$\theta = \text{atan} \left(\frac{{}^n\omega_y}{{}^n\omega_x} \right). \quad (2.53)$$

Since the proposed features are invariant to the insertion velocity ${}^n v_z$, we can fix this velocity to a constant, or leave it to the appreciation of the clinician. The insertion velocity $v_t = {}^n v_z$ is used to update the tracker state model in (1.65).

Experimental results To illustrate the proposed 3D ultrasound-based needle steering method, we conducted experiments with a home-made gelatin phantom. These experiments were performed with a 6-DOF Viper s650 robot (Omron Adept Technologies, Inc., CA) holding a 22 gauge bevel-tip flexible needle (Angiotech Chiba MCN2208). Ultrasound imaging was provided by a 4DC7-3/40 motorized ultrasound transducer (BK Ultrasound, MA) maintained still [see Figure 2.13(b)].

The needle was initially inserted into the phantom until the tip was visible in the 3D ultrasound image. The initial position of the needle was detected manually by clicking in the x - z and y - z planes. Given this initial detection, the target was defined 8 cm away from the needle tip and with an offset of 7 mm with respect to the direction vector of the needle. Needle tracking was performed with the particle filtering approach proposed in section 1.4. As parameters, we used $n = 4$ control points (cubic curve), $M = 1000$ particles, and an isotropic Gaussian noise model with a standard deviation $\sigma^{(n)} = 1.0$ for the tip, $\sigma^{(2)} = \sigma^{(3)} = 0.25$ for the inner control points, and $\sigma^{(1)} = 0.0$ for the entry point.

Then, the duty-cycling control law presented in (2.52) and (2.53) was activated. The needle was inserted with a constant insertion velocity $v_z^* = 0.5 \text{ mm s}^{-1}$, and the control gain was set to $\lambda = 0.01 \text{ s}^{-1}$. The natural curvature of the needle was estimated experimentally to $K = 2 \text{ m}^{-1}$. The position of the needle in the ultrasound volume at different instants is presented in Figure 2.14. The target was successfully reached by the needle, with a final positioning error of 1.08 mm between the needle tip and the target. The final feature errors were 0.453 mm for l_{xz} and 1.07 mm for l_{yz} . These errors are close to the resolution of the reconstructed volume (0.65 mm). The difference of precision between the two features can be explained by the lower resolution of the prescan ultrasound volume in the direction corresponding to the probe sweep, which is less than half the angular resolution of the frames.

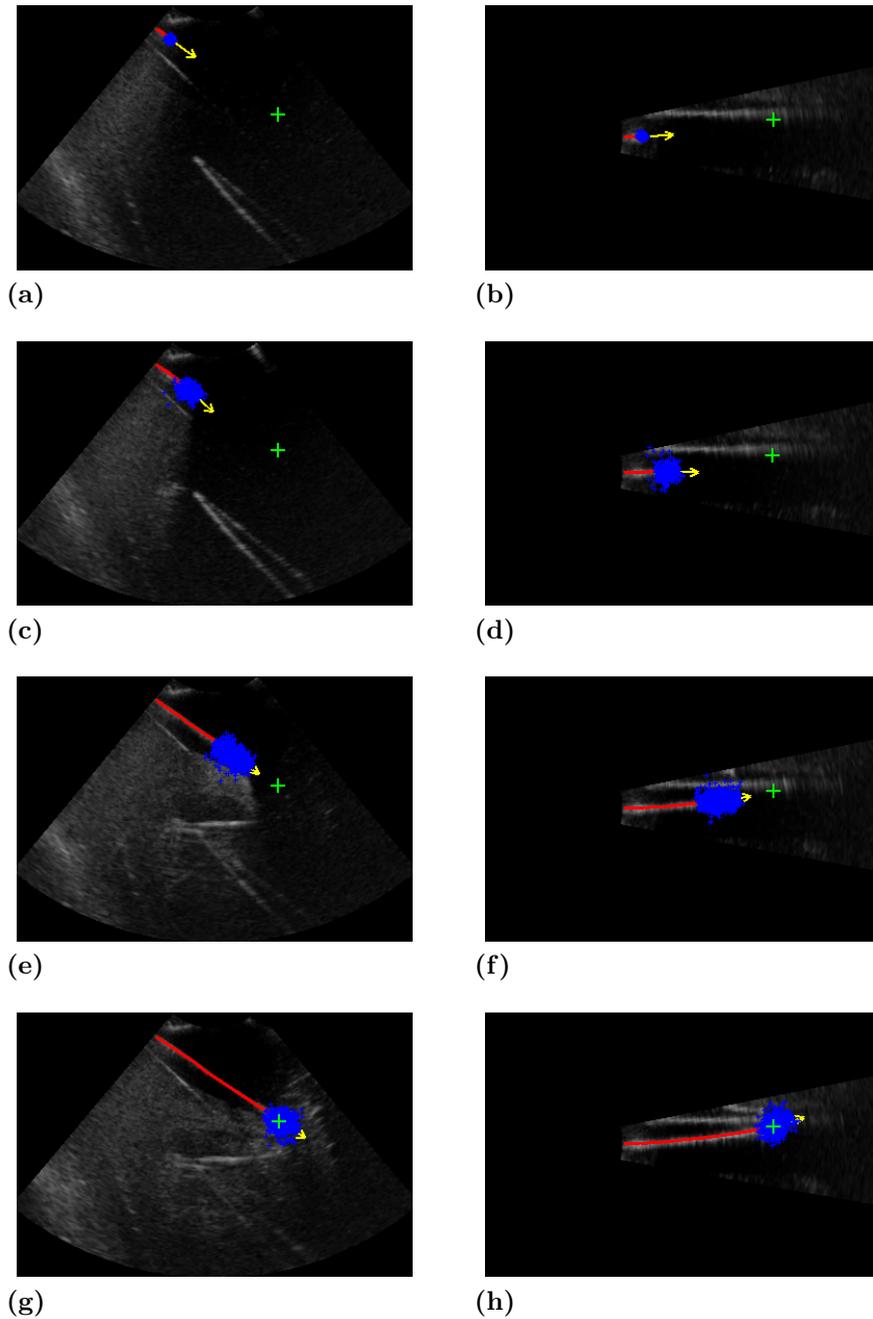


Figure 2.14 – Needle steering under 3D ultrasound guidance. The needle is represented projected onto the x - z (left column) and y - z (right column) planes. The needle detected by the tracker is represented as a red line, and the target as a green cross. The yellow arrow corresponds to the direction vector of the needle at the tip, and the blue points represent the candidate position of the needle tip for all particles.

2.4 Conclusion

We presented a review of ultrasound-based visual servoing methods. In particular, we described different approaches for servoing an ultrasound probe to the image content. These approaches allow automatic motion compensation or target tracking schemes that can assist the sonographer in performing an exam. Then, we provided an overview of the state-of-the-art on ultrasound-guided robot-assisted needle insertion. Robot-assisted needle insertion has the potential to improve the flexibility and accuracy of biopsy and drug delivery procedures. Ultrasound guidance enables the design of adaptive insertion strategies with a real-time visual feedback. We proposed a closed-loop 3D ultrasound-guided needle steering method, using visual servoing to perform continuous trajectory adaptation.

In chapter 3, we will introduce a new control framework, aimed at optimizing the quality of ultrasound images. This new framework will build on the notions introduced in chapter 1 on ultrasound image quality, and in chapter 2 on visual servoing.

Chapter 3

Quality-Driven Visual Servoing

In chapter 2, we reviewed several visual servoing strategies for the control of an ultrasound probe. Various methods have been proposed to control the position of an ultrasound probe based on geometric or photometric features extracted from the ultrasound images. These methods typically aim at controlling the position of an anatomical target within the image, or at compensating for soft tissue/respiratory motion. However, such methods usually assume that image quality is uniform, i.e., that the information contained in all pixels is equally reliable. More importantly, existing ultrasound-based visual servoing methods do not take into account the relation between the positioning of the probe and the image quality. It is generally assumed that a good acoustic coupling is maintained throughout the acquisition.

As we have discussed in section 1.5, the content of an ultrasound image is not uniform in quality. Several factors, such as the contact force, the probe orientation, and the presence of acoustic coupling gel, can influence the image quality. In addition, abrupt changes in the acoustic impedance of tissues, such as between soft tissues and bones, can result in an acoustic shadow masking a part of the image. An important concept in ultrasound imaging is that of *acoustic window*. The acoustic window for a target can be defined by the area of the body surface covered by the ultrasound transducer and the direction through which the ultrasound waves are transmitted towards the target. This is illustrated in Figure 3.1. Due to the presence of tissue interfaces, the choice of the acoustic window affects the quality of the image. Finding a good acoustic window is not only critical for subsequent diagnosis, but it is also necessary for a robust ultrasound-based control, in order to ensure that the image content remains reliable.

In this chapter, which constitutes the main contribution of this thesis, we propose a control framework for optimizing the quality of ultrasound

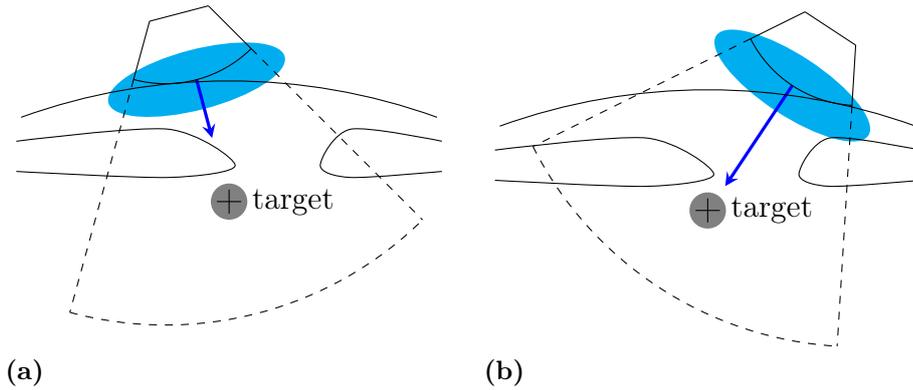


Figure 3.1 – Acoustic window in ultrasound imaging. For a given target, the acoustic window (in blue) is defined as the surface position and orientation from which the ultrasound waves are emitted to reach the target. (a) Sub-optimal acoustic window: the presence of an obstacle masks the target. (b) Good acoustic window: the path between the probe and the target is free from strong attenuators.

images. To this end, we use the concept of confidence map, which was introduced in section 1.5. We consider the confidence map as a new sensory modality to design a control law aimed at improving the quality of acquired images. More specifically, we derive from the confidence map an angular confidence feature that indicates the optimal acoustic window. We use this angular confidence feature as input to a control law, following the visual servoing framework, to control the orientation of an ultrasound probe. We first propose a general confidence-driven control strategy, where the task is to maintain an optimal acoustic window, while the probe is being teleoperated by a user. Second, we propose a target-specific control strategy, where the acoustic window is optimized for a specific anatomical target tracked in the ultrasound stream.

This chapter is organized as follows. In section 3.1, we study the dynamics of the confidence map. That is, the relation between the position of the probe and the quality of the ultrasound images. We first provide some general observations on the impact of contact force and probe orientation on the image quality. Then, we propose a mathematical analysis of the interaction for the specific case of the scan line integrated confidence map, which we defined in section 1.5.2.

In section 3.2, we propose a definition of geometric confidence features which are adapted to the design of a confidence-driven control law, and we derive the relation between these features and the rotation velocity of the ultrasound probe.

In section 3.3, we describe a global confidence-driven control strategy, by which the orientation of the ultrasound probe is controlled so as to globally optimize the image quality. We combine the confidence-control with force control, to ensure a constant contact force with the body, and teleoperation.

In section 3.4, we propose a target-specific confidence-driven control strategy, where the acoustic window is optimized with respect to a tracked anatomical target. We use a hybrid tasks approach to control simultaneously the contact force, the position of the target, and the confidence.

3.1 Dynamics of the confidence map

In this section, we study the variations of the confidence map with respect to the velocity of the ultrasound probe. We first provide some general observations on the impact of contact force and probe orientation, which are independent of the definition of the confidence map. Then, we propose a more detailed analysis of the dynamics of the confidence map for the scan line integration method we introduced in section 1.5.2.

3.1.1 General considerations

Let us start with some general observations that can be made without considering any specific definition for the confidence map. Due to the formation process of ultrasound images, the confidence $C(\mathbf{p})$ at a given location \mathbf{p} in the image depends on the path traveled by the ultrasound wave to reach \mathbf{p} , in addition to the local properties of the tissues at \mathbf{p} . In particular, the quality of the ultrasound image is strongly dependent on the acoustic coupling between the probe and the patient's skin. In order to ensure an efficient transmission into the body, an acoustic coupling medium (water-based gel) with controlled acoustic properties is placed on the skin. In addition, it is necessary to apply a sufficient contact force between the probe and the patient's skin to minimize the attenuation of sound amplitude at the patient's surface.

3.1.1.1 Effect of the contact force

While force control is mainly presented as a way to ensure the safety of the patient during robotized ultrasonography, the contact force is also a crucial factor for ultrasound image quality. Therefore, we propose to study the relation between contact force and image quality. This relation

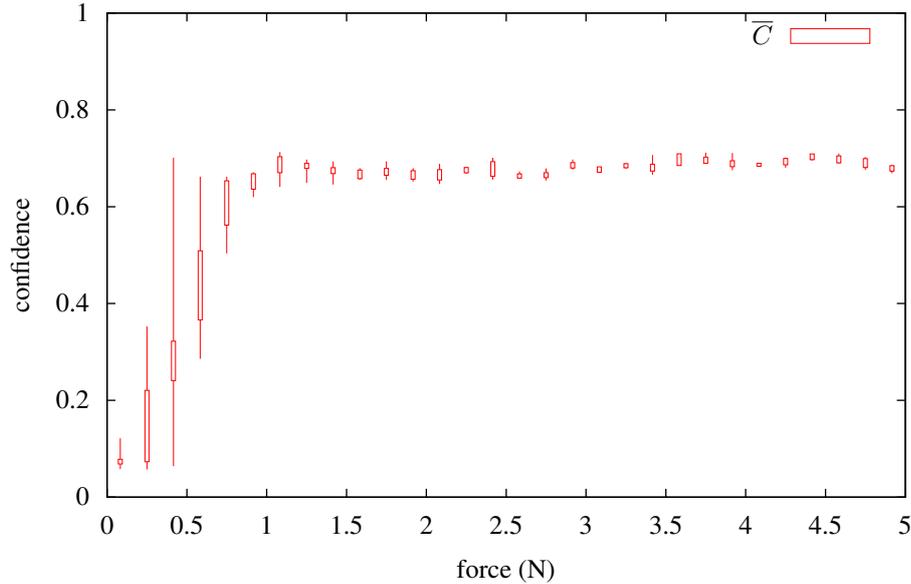


Figure 3.2 – Relation between the mean image confidence and the contact force.

has been briefly studied in [Kuhlemann, 2013], where the sum of image gradients is used as a measure of image quality. The author observes that this quality measure increases with the contact force up to 7.5 N, and is stable for higher contact forces.

We propose an experiment to evaluate the relation between contact force and image quality, using the random walks confidence map and the scan line integration methods described in section 1.5.2. Let $C(\mathbf{p})$ represent the signal confidence at a pixel $\mathbf{p} \in \Omega$, estimated with either of the two methods. We measure the mean image confidence

$$\bar{C} = \frac{1}{|\Omega|} \iint_{\mathbf{p} \in \Omega} C(\mathbf{p}) d\mathbf{p}, \quad (3.1)$$

where $|\Omega|$ is the size of the field of view Ω . We estimate \bar{C} for various contact forces ranging from 0 N to 5 N. The results obtained on phantom are presented in Figure 3.2. We observe that the mean image confidence increases rapidly while contact is made, from 0 N to 1 N, and varies only slightly when the contact force continues to increase. This behavior is consistent with what was observed in [Kuhlemann, 2013]. Note that the force threshold beyond which the image quality is constant depends on the characteristics of the patient’s body.

3.1.1.2 Effect of orientation

Let us consider a 2D convex ultrasound probe. Following the notations introduced in chapter 1, we represent a point in the image with its polar coordinates (r, θ) , where r is the distance to the imaging center, and θ is the angle with respect to the central scan line. A rotation around the imaging center of the probe is equivalent to a shift in the scan lines. Indeed, if the probe is rotated by angle $d\theta$ between the times t and $t + dt$, then a point with polar coordinates (r, θ) in the image at time t corresponds to the point with coordinates $(r, \theta + d\theta)$ at $t + dt$ (see Figure 3.3). As a result, we have

$$C(r, \theta, t) = C(r, \theta + d\theta, t + dt), \quad (3.2)$$

from which we can deduce the variation of confidence for a given image pixel at (r, θ) .

Following the conventions introduced in Figure 2.5, we note ω_z the instantaneous in-plane angular velocity, such that $d\theta = \omega_z dt$. Then, we have

$$\frac{C(r, \theta, t + dt) - C(r, \theta, t)}{dt} = \frac{C(r, \theta - \omega_z dt, t) - C(r, \theta, t)}{dt}, \quad (3.3)$$

which gives, when $dt \rightarrow 0$,

$$\frac{dC}{dt}(r, \theta, t) = -\omega_z \frac{\partial C}{\partial \theta}(r, \theta, t). \quad (3.4)$$

Therefore, the instantaneous variation of the confidence at a pixel under a pure rotation motion around the imaging center is proportional to the angular confidence gradient.

3.1.1.3 Approximated interaction matrix

We will now propose a general approximation of the interaction matrix associated to the confidence values. To this end, let us consider the impact of translations on the confidence.

A translation in the direction normal to the surface influences the global quality of the image, because it changes the contact force, and so the acoustic coupling between the probe and the skin. However, as we have seen in section 3.1.1.1, the influence of contact force on image quality is low above a certain force threshold.

The effect of a translation at the patient's surface is more difficult to predict without any further assumption on the confidence map. In the case of a 2D probe, out-of-plane motions also have an impact that

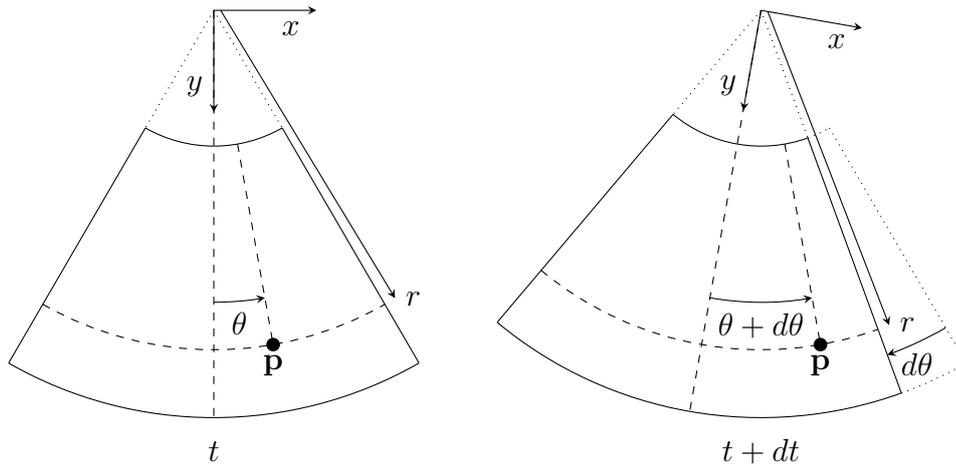


Figure 3.3 – Relation between the in-plane orientation of a probe and the polar coordinates of a point. The point \mathbf{p} is located at (r, θ) in polar coordinates at time t . At time $t + dt$, after a rotation $d\theta$ around the z -axis, \mathbf{p} is located at $(r, \theta + d\theta)$.

cannot be directly predicted. When scanning a relatively uniform region, translations at the surface will not have a large impact on the confidence map. However, in the presence of strong acoustic attenuators, the impact of such translations could be larger. In practice, the influence of the surface contact and near-surface interfaces is more important than that of deeper interfaces. Indeed, a loss of signal occurring near the surface impacts a large part of the image, while a loss of signal occurring deeper only impacts the area located below the attenuating interface. We propose, in first approximation, to neglect the effect of surface translations in the design of the control law.

Consequently, we use the following approximation of the interaction matrix, where only the effect of rotation around the imaging center is considered:

$$\widehat{\mathbf{L}}_{C(r,\theta)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -\frac{\partial C}{\partial \theta}(r, \theta) \end{bmatrix}. \quad (3.5)$$

3.1.2 Analytic solution for scan line integration

We now propose an analysis of the confidence dynamics in the specific case of the scan line integrated confidence map that we proposed in section 1.5.2. Let us first recall the definition of this confidence map. We note $U : \Omega \rightarrow \mathbb{R}$ an ultrasound image, where the field of view $\Omega = [r_{\min}, r_{\max}] \times [\theta_{\min}, \theta_{\max}]$ is defined in polar coordinates following the notations of Figure 1.6. Then, the scan line integrated confidence of the

ultrasound signal at (r, θ) is defined as

$$C(r, \theta) = 1 - \left(\int_{r_{\min}}^{r_{\max}} U(u, \theta) du \right)^{-1} \int_{r_{\min}}^r U(u, \theta) du. \quad (3.6)$$

Interaction matrix We have already seen in section 2.2.3 that the variation of ultrasound intensity at a position $\mathbf{x} = (x, y)$ in Cartesian coordinates can be expressed as

$$\dot{U} = \nabla U \dot{\mathbf{x}}, \quad (3.7)$$

where

$$\nabla U = \left(\frac{\partial U}{\partial x} \quad \frac{\partial U}{\partial y} \right) \quad (3.8)$$

is the ultrasound image gradient, and

$$\dot{\mathbf{x}} = \mathbf{L}_{\mathbf{x}} \mathbf{v} = \begin{bmatrix} 1 & 0 & -y \\ 0 & 1 & x \end{bmatrix} \begin{pmatrix} v_x \\ v_y \\ \omega_z \end{pmatrix} \quad (3.9)$$

is the variation of the coordinates of the physical point located at \mathbf{x} , expressed in the probe frame.

Noting $A(\theta) = \int_{r_{\min}}^{r_{\max}} U(u, \theta) du$ the sum of intensities along the scan line with angle θ and $a(r, \theta) = \int_{r_{\min}}^r U(u, \theta) du$ the partial sum up to r , the derivation of (3.6) gives

$$\dot{C}(r, \theta) = \dot{A}(\theta) A(\theta)^{-2} a(r, \theta) - A(\theta)^{-1} \dot{a}(r, \theta), \quad (3.10)$$

which we can rewrite, noticing that $A(\theta)^{-1} a(r, \theta) = 1 - C(r, \theta)$,

$$\dot{C}(r, \theta) = A(\theta)^{-1} \left((1 - C(r, \theta)) \dot{A}(\theta) - \dot{a}(r, \theta) \right). \quad (3.11)$$

Using (3.7), the derivatives of A and a can be expressed as

$$\dot{A}(r, \theta) = \int_{r_{\min}}^{r_{\max}} \nabla U(u, \theta) \dot{\mathbf{x}}(u, \theta) du \quad (3.12)$$

and

$$\dot{a}(r, \theta) = \int_{r_{\min}}^r \nabla U(u, \theta) \dot{\mathbf{x}}(u, \theta) du, \quad (3.13)$$

so that we can rewrite (3.11) as

$$\begin{aligned} \dot{C}(r, \theta) = A(\theta)^{-1} & \left((1 - C(r, \theta)) \int_{r_{\min}}^{r_{\max}} \nabla U(u, \theta) \dot{\mathbf{x}}(u, \theta) du \right. \\ & \left. - \int_{r_{\min}}^r \nabla U(u, \theta) \dot{\mathbf{x}}(u, \theta) du \right). \end{aligned} \quad (3.14)$$

Finally, we can rearrange slightly this equation using Chasles' theorem, to obtain

$$\begin{aligned} \dot{C}(r, \theta) = A(\theta)^{-1} & \left((1 - C(r, \theta)) \int_r^{r_{\max}} \nabla U(u, \theta) \dot{\mathbf{x}}(u, \theta) du \right. \\ & \left. - C(r, \theta) \int_{r_{\min}}^r \nabla U(u, \theta) \dot{\mathbf{x}}(u, \theta) du \right). \end{aligned} \quad (3.15)$$

As a result, the interaction matrix for the scan line integrated confidence at a location (r, θ) is

$$\mathbf{L}_C = A^{-1} \left((1 - C) \int_r^{r_{\max}} \nabla U \mathbf{L}_x - C \int_{r_{\min}}^r \nabla U \mathbf{L}_x \right), \quad (3.16)$$

where we have omitted the function arguments r and θ and the variable of integration for the sake of readability.

In-plane rotation component In section 3.1.1.2, we demonstrated geometrically that the ω_z (sixth) component of \mathbf{L}_C is equal to $-\frac{\partial C}{\partial \theta}(r, \theta)$. This term links the variation of confidence to the angular velocity of the probe around the imaging center. We now propose an alternative analytic demonstration using the expressions derived above. Let us note $\mathbf{L}_{U\omega_z}$ the sixth term of the interaction matrix \mathbf{L}_U that links the variation of a pixel intensity to the angular velocity ω_z around the imaging center. From (3.7), we have

$$\mathbf{L}_{U\omega_z} = -y \frac{\partial U}{\partial x} + x \frac{\partial U}{\partial y}, \quad (3.17)$$

which we can also express from the gradient in polar coordinates as

$$\mathbf{L}_{U\omega_z} = -\frac{\partial U}{\partial \theta}. \quad (3.18)$$

On the other hand, the partial derivative of (3.6) with respect to θ gives

$$\frac{\partial C}{\partial \theta}(r, \theta) = A(\theta)^{-1} \left((1 - C(r, \theta)) \frac{\partial A}{\partial \theta}(\theta) - \frac{\partial a}{\partial \theta}(r, \theta) \right), \quad (3.19)$$

where

$$\frac{\partial A}{\partial \theta}(\theta) = \frac{\partial}{\partial \theta} \int_{r_{\min}}^{r_{\max}} U(u, \theta) du = \int_{r_{\min}}^{r_{\max}} \frac{\partial U}{\partial \theta}(u, \theta) du \quad (3.20)$$

and

$$\frac{\partial a}{\partial \theta}(r, \theta) = \frac{\partial}{\partial \theta} \int_{r_{\min}}^r U(u, \theta) du = \int_{r_{\min}}^r \frac{\partial U}{\partial \theta}(u, \theta) du. \quad (3.21)$$

From (3.16) and (3.18)–(3.21), we finally obtain

$$\mathbf{L}_{C(r,\theta)\omega_z} = -\frac{\partial C}{\partial \theta}(r, \theta), \quad (3.22)$$

which corresponds to the result (3.5) demonstrated in the general case.

3.2 Geometric confidence features

In this section, we propose the definition of visual features from the confidence map. We consider a generic confidence map, which can be implemented by any of the methods presented in section 1.5. These features are inspired from the observations made, in the previous section, on the relation between the variation of the confidence and the rotation of the probe.

3.2.1 2D Case

Let us first consider the acquisition of a single ultrasound frame by a convex probe. We have seen that, independently of the definition of the confidence map, the relation between the variation of confidence at a location (r, θ) and a rotation motion around the z -axis is given by

$$\mathbf{L}_{C(r,\theta)\omega_z} = -\frac{\partial C}{\partial \theta}. \quad (3.23)$$

3.2.1.1 Definition

We propose to define a geometric feature that represents the angular distribution of the confidence within the field of view. A natural choice would be to use the direction in which the confidence is maximum. However, this feature would be unstable, due to possible discontinuities of the argmax function. Instead, we consider the angular coordinate of the confidence-weighted barycenter. Inspired from the definition of photometric moments [Bakthavatchalam et al., 2013], this feature consists in a weighted average of all values in the confidence map. As a result, it is more robust to noise. We define the angular confidence feature as

$$\theta_c = \frac{1}{C_\Omega} \iint_{(r,\theta) \in \Omega} \theta C(r, \theta, t) r dr d\theta, \quad (3.24)$$

where

$$C_\Omega = \iint_{(r,\theta) \in \Omega} C(r, \theta) r dr d\theta \quad (3.25)$$

is the sum of the confidence values. The angular confidence feature is represented in Figure 3.4 for three different ultrasound images.

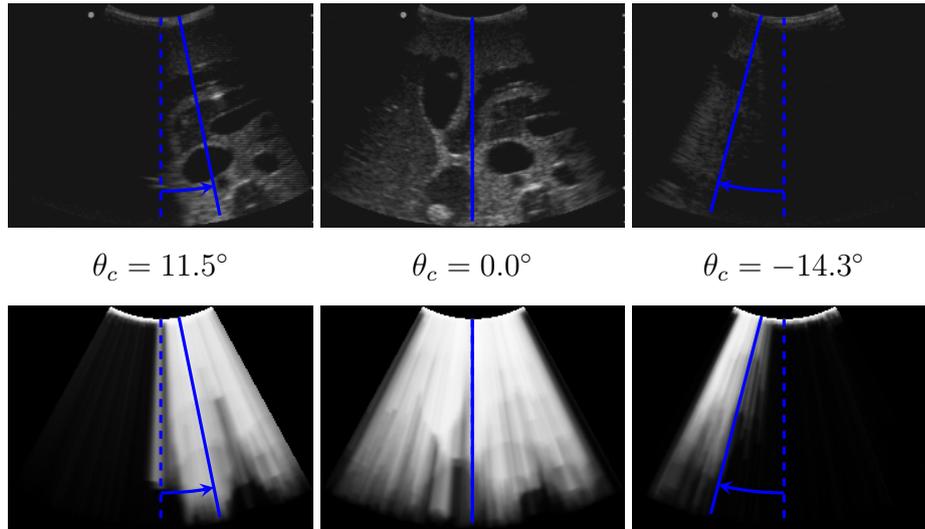


Figure 3.4 – Angular confidence feature for three different positions of an ultrasound probe. On the top row, the B-mode ultrasound images. On the bottom row, the corresponding scan line integrated confidence maps. The value of the angular confidence feature θ_c is indicated for each image.

3.2.1.2 Experimental comparison

We have claimed above that the angular confidence feature defined in (3.24) is more stable than the direction of maximum confidence. We propose a comparison between the two approaches to support this assertion. In addition, we also compare the angular confidence features obtained from the Random Walks (RW) confidence map and the Scan Line Integration (SLI) confidence map.

Direction of maximum confidence A natural approach to estimate the best acoustic direction would be to find the direction in which the confidence is maximized. This can be done by finding the angular coordinate θ_c^{\max} of the scan line that has the highest mean confidence:

$$\theta_c^{\max} = \arg \max_{\theta} \int_r C(r, \theta, t) dr. \quad (3.26)$$

We estimated the value of θ_c^{\max} for each frame of a sequence of ultrasound images acquired on a human subject. The resulting curve is presented in Figure 3.5. As expected, the curve has important discontinuities.

Confidence-weighted angular feature with SLI We provide in Figure 3.6 the results obtained on the same ultrasound sequence, for the pro-

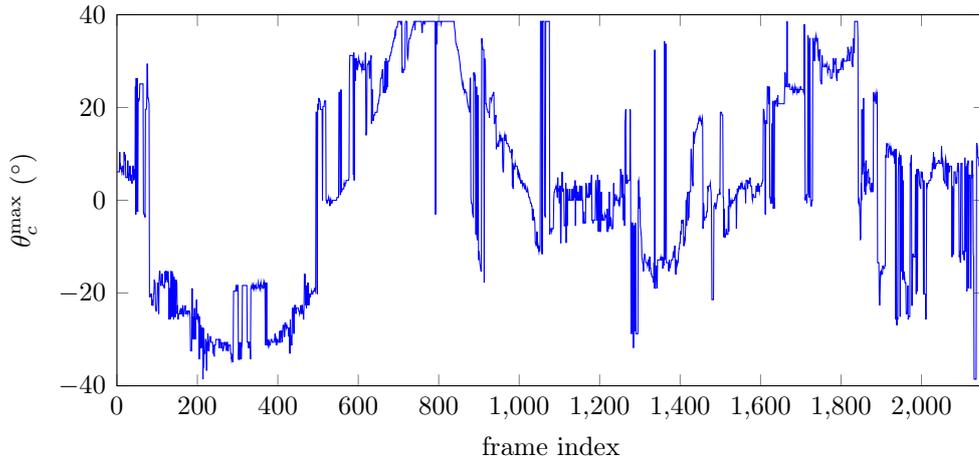


Figure 3.5 – Estimation of the direction of maximum confidence θ_c^{\max} on a sequence of ultrasound images. The confidence map was computed via the scan line integration method.

posed confidence-weighted angular feature θ_c , estimated from the scan line integrated confidence map. The obtained curve is much smoother than with the direction of maximum confidence. This suggests that the confidence-weighted angular feature is more robust. Therefore, it is more reliable for being used as input to a control law.

Confidence-weighted angular feature with RW Finally, we report in Figure 3.7 the results obtained with the confidence-weighted angular feature, when estimating the confidence map via the random walks method. The shape of the curve is similar to the one obtained with the scan line integration method, but it is noisier. We had already noted in section 1.5.3.2 that the scan line integration method provides locally more regular confidence estimates than the random walks method. We now see that this property propagates to the estimation of the confidence-weighted angular feature.

3.2.1.3 Feature dynamics

Let us now study the dynamics of the angular confidence feature θ_c . It is reasonable to assume that the mean image confidence is constant, i.e., $\dot{C}_\Omega = 0$. Under this assumption, derivating (3.24) yields

$$\dot{\theta}_c = \frac{1}{C_\Omega} \iint_{(r,\theta) \in \Omega} \theta \dot{C}(r, \theta) r dr d\theta. \quad (3.27)$$

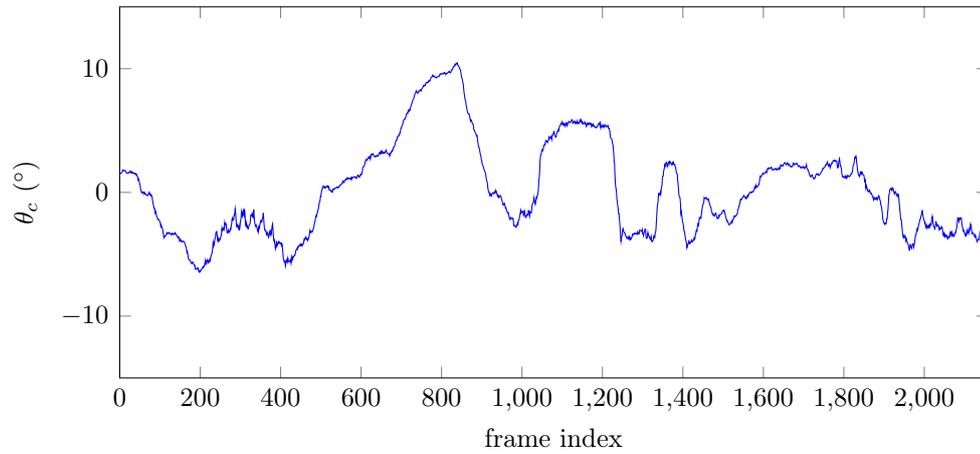


Figure 3.6 – Estimation of the confidence-weighted angular feature θ_c on a sequence of ultrasound images. The confidence map was computed via the scan line integration method.

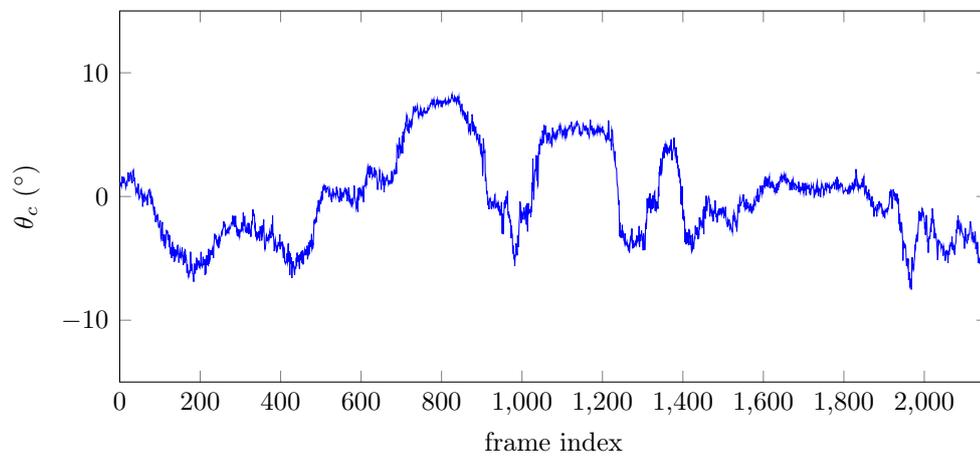


Figure 3.7 – Estimation of the confidence-weighted angular feature θ_c on a sequence of ultrasound images. The confidence map was computed via the random walks method.

Using the expression of \dot{C} given in (3.4), we can write

$$\dot{\theta}_c = -\frac{1}{C_\Omega} \iint_{(r,\theta) \in \Omega} \theta \omega_z \frac{\partial C}{\partial \theta}(r, \theta) r dr d\theta, \quad (3.28)$$

where ω_z is the angular velocity around the z -axis. We can now rearrange the double integral by factoring the constant terms, such as

$$\dot{\theta}_c = -\frac{\omega_z}{C_\Omega} \int_{r_{\min}}^{r_{\max}} r \left(\int_{\theta_{\min}}^{\theta_{\max}} \theta \frac{\partial C}{\partial \theta}(r, \theta) d\theta \right) dr. \quad (3.29)$$

Then, an integration by parts on the inner integral gives

$$\int_{\theta_{\min}}^{\theta_{\max}} \theta \frac{\partial C}{\partial \theta}(r, \theta) d\theta = [\theta C(r, \theta)]_{\theta_{\min}}^{\theta_{\max}} - \int_{\theta_{\min}}^{\theta_{\max}} C(r, \theta) d\theta, \quad (3.30)$$

where the term

$$[\theta C(r, \theta)]_{\theta_{\min}}^{\theta_{\max}} = \frac{\Theta}{2} (C(r, \theta_{\min}) + C(r, \theta_{\max})) \quad (3.31)$$

corresponds to the right and left image borders. $\Theta = \theta_{\max} - \theta_{\min}$ is the angular extent of the field of view. We note that the confidence at the limits of the field of view is typically lower than inside the field of view, and $\Theta < \pi$, so that we can safely assume

$$[\theta C(r, \theta)]_{\theta_{\min}}^{\theta_{\max}} \ll \int_{\theta_{\min}}^{\theta_{\max}} C(r, \theta) d\theta. \quad (3.32)$$

Consequently, by neglecting $[\theta C(r, \theta)]_{\theta_{\min}}^{\theta_{\max}}$ in (3.30), we can rewrite (3.29) as

$$\dot{\theta}_c = \frac{\omega_z}{C_\Omega} \int_{r_{\min}}^{r_{\max}} r \left(\int_{\theta_{\min}}^{\theta_{\max}} C(r, \theta) d\theta \right) dr, \quad (3.33)$$

where we recognize

$$\int_{r_{\min}}^{r_{\max}} r \left(\int_{\theta_{\min}}^{\theta_{\max}} C(r, \theta) d\theta \right) dr = C_\Omega, \quad (3.34)$$

so that we finally obtain

$$\dot{\theta}_c = \omega_z. \quad (3.35)$$

Therefore, the time variation of the angular confidence feature is equal to the angular velocity around the imaging center.

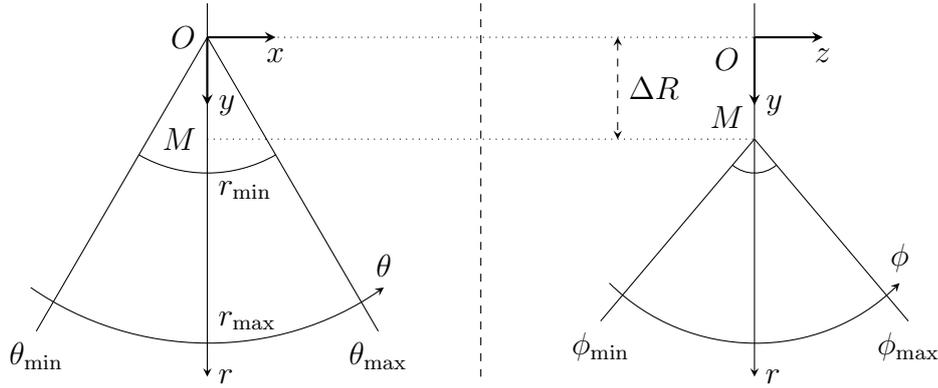


Figure 3.8 – Imaging geometry of a motorized 3D convex ultrasound probe. Left: front view, also corresponding to the imaging geometry of a 2D probe. Right: side view, showing the range of the motor.

3.2.2 3D Case

We will now provide a generalization to 3D ultrasound imaging. Let us consider, for instance, a wobbler probe, which sweeps the transducer back and forth around the motor's axis. Given the geometry of the probe, the Cartesian coordinates (x, y, z) of a point in the field of view can be expressed in a pseudo-spherical coordinate system (r, θ, ϕ) as [Lee and Krupa, 2011]

$$x = r \sin \theta, \quad (3.36)$$

$$y = (r \cos \theta - \Delta R) \cos \phi + \Delta R, \quad (3.37)$$

$$z = (r \cos \theta - \Delta R) \sin \phi, \quad (3.38)$$

where ΔR is the offset between the imaging center O and the axis of the probe's motor (see Figure 3.8).

We have seen before that a rotation around the axis Oz is equivalent to a shift in the angular coordinate θ for the confidence map. Similarly, for a 3D probe, a rotation around the motor axis Mx is equivalent to a shift in the frames. Following this observation, we define two angular features in the same way as for the 2D case, using the pseudo-spherical coordinates of the confidence-weighted barycenter in 3D:

$$\theta_c = \frac{1}{C_\Omega} \iiint_{\Omega} \theta C(r, \theta, \phi, t) |\det(J)| dr d\theta d\phi, \quad (3.39)$$

$$\phi_c = \frac{1}{C_\Omega} \iiint_{\Omega} \phi C(r, \theta, \phi, t) |\det(J)| dr d\theta d\phi, \quad (3.40)$$

where

$$|\det(J)| = r \cos^2 \theta \cos^2 \phi (r \cos \theta - \Delta R) \quad (3.41)$$

is the determinant of the Jacobian of the transformation from Cartesian to pseudo-spherical coordinates, and

$$C_\Omega = \iiint_{\Omega} C(r, \theta, \phi) |\det(J)| dr d\theta d\phi \quad (3.42)$$

is the sum of the confidence values in the 3D field of view. Under the same assumptions as in the 2D case, the time variation of the features θ_c and ϕ_c can be written:

$$\dot{\theta}_c = \omega_z, \quad (3.43)$$

$$\dot{\phi}_c = -\omega_x, \quad (3.44)$$

where ω_x and ω_z are the angular velocities around the x - and y -axes, respectively.

3.3 Global confidence-driven control

In the previous sections, we have studied the variations of the ultrasound confidence map with respect to the velocity of the probe. We have also defined angular confidence features, whose variations are linked to the rotation velocity components ω_x and ω_z of the ultrasound probe. In this section, we propose the design of a first control law based on the confidence map. The aim of this control law is to optimize the orientation of the ultrasound probe, so as to obtain the best possible image quality. A typical application case will be tele-echography, where the physician controls the translations of the probe along the patient's surface, as well as the rotation around the y -axis. In this scenario, the confidence-driven controller aims at maintaining a high image quality globally. We present our control framework in the generic case of a 3D ultrasound probe. It can be easily adapted to the in-plane control of a 2D probe, by simply omitting the components v_z , ω_x and ω_y in the control law.

The proposed framework, illustrated in Figure 3.9, can be modeled as a hierarchical series of three tasks. These tasks are, in decreasing order of priority, the force control (\mathbf{e}_f^* , \mathbf{L}_f), the confidence-driven control (\mathbf{e}_c^* , \mathbf{L}_c), and the operator's commands (\mathbf{e}_{op}^* , \mathbf{L}_{op}). We have noted, following the conventions for hybrid tasks defined in section 2.1.4, \mathbf{e}_i^* the desired evolution of the feature error i , and \mathbf{L}_i the interaction matrix associated to this error. The full control screw \mathbf{v} can be decomposed as

$$\mathbf{v} = \mathbf{v}_f + \mathbf{v}_c + \mathbf{v}_{\text{op}}, \quad (3.45)$$

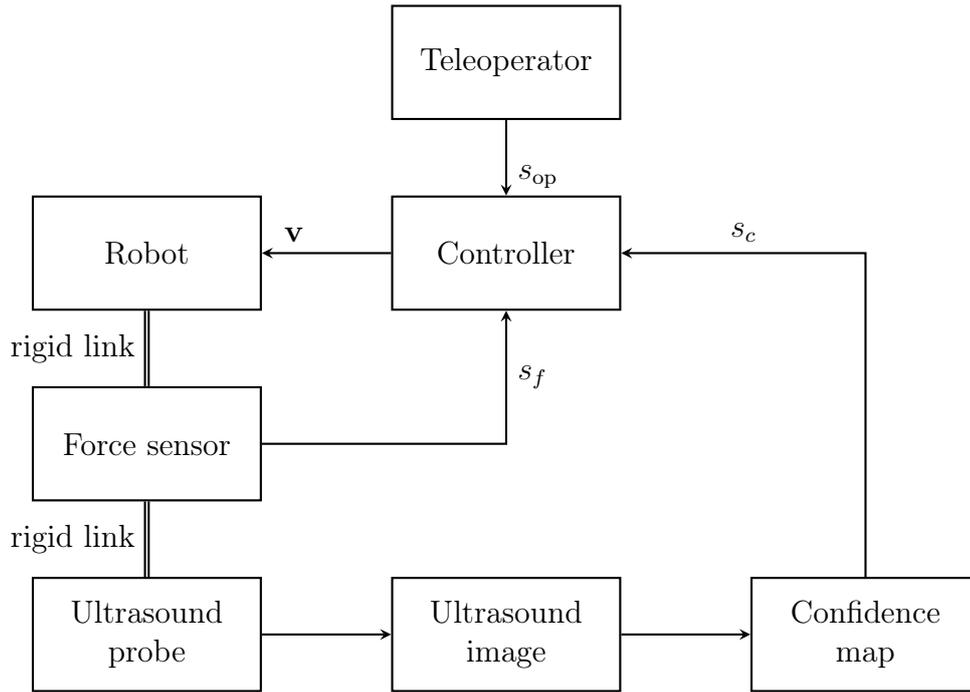


Figure 3.9 – Overview of the global confidence-driven control framework.

where \mathbf{v}_f , \mathbf{v}_c , \mathbf{v}_{op} are the contributions of the control screw dedicated to the force control task, confidence control task, and operator task, respectively.

Following the redundancy formalism, we will build the control screw incrementally, starting with force control.

3.3.1 Force control

Controlling the contact force applied to the body is crucial for the patient’s safety. In addition, the contact force influences the acoustic coupling between the transducer and the body, which is an important factor for the image quality. Therefore, force control is defined as the task of highest priority. The aim is to maintain a desired contact force with the patient. For the purpose of presenting a complete solution for our hybrid control design, we recall in this section the force control approach proposed in [Nadeau and Krupa, 2013]. This approach has the merit of integrating well with the visual servoing formulation. However, other force control methods could be considered as well. See for instance [Gilbertson and Anthony, 2015] or [Pappalardo et al., 2016].

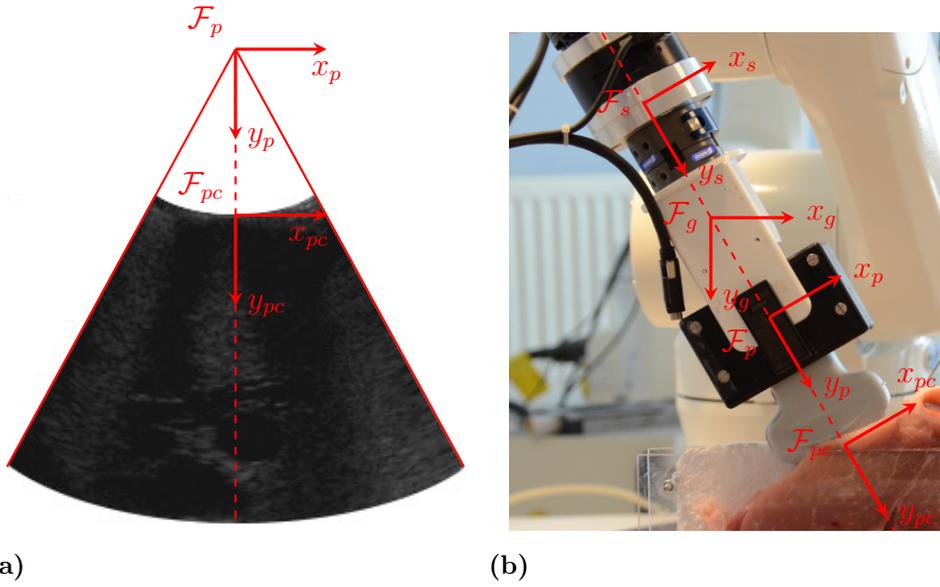


Figure 3.10 – Configuration of the probe frame, contact frame, sensor frame, and gravity frame. (a) Ultrasound image. (b) End-effector of the robot with the force sensor and ultrasound probe, positioned in contact with a piece of meat.

Force sensor Let us assume that the robot is equipped with a force sensing system, that provides a measure of the external forces \mathbf{f} and torques $\boldsymbol{\tau}$ expressed in a sensor frame \mathcal{F}_s . Figure 3.10 presents a possible configuration of the force sensor and ultrasound probe attached to the end-effector of a robot, as well as the different frames which will be used in the design of the control law:

- The *probe frame* \mathcal{F}_p is attached to the imaging center O of the convex probe, such that its y -axis coincides with the central scan line, oriented towards the body, and the plane (O, x_p, y_p) corresponds to the central frame.
- The *contact frame* \mathcal{F}_{pc} is attached to contact point between the probe and the patient, and oriented like \mathcal{F}_p .
- The *sensor frame* \mathcal{F}_s is the one in which the measured force/torque tensor is expressed, and is oriented as well like \mathcal{F}_p .
- The *gravity frame* \mathcal{F}_g , is attached to the probe's center of mass, with its y -axis in the direction of the gravity force.

We represent the wrench $\{\mathbf{f}, \boldsymbol{\tau}\}$ with a 6-dimensional force/torque

vector

$${}^s\mathbf{H}_s = \begin{pmatrix} \mathbf{f} \\ \boldsymbol{\tau} \end{pmatrix}. \quad (3.46)$$

Contact force The force sensor provides a ${}^s\mathbf{H}_s$ measure of external forces and torques expressed in the sensor frame \mathcal{F}_s . In order to obtain the actual contact force between the probe and the body, we need to convert this measure to the contact frame \mathcal{F}_{pc} , and to subtract the contribution of the probe's weight to the external forces. Given the twist transformation matrices ${}^{pc}\mathbf{F}_s$ and ${}^s\mathbf{F}_g$ allowing the transformation of a force/torque vector respectively from the sensor frame \mathcal{F}_s to the probe contact frame \mathcal{F}_{pc} , and from the probe's inertial frame \mathcal{F}_g to \mathcal{F}_s , the force/torque vector in the contact frame can be written

$${}^{pc}\mathbf{H}_{pc} = {}^{pc}\mathbf{F}_s({}^s\mathbf{H}_s - {}^s\mathbf{F}_g{}^g\mathbf{H}_g), \quad (3.47)$$

where

$${}^g\mathbf{H}_g = (0 \ 0 \ m_p g \ 0 \ 0 \ 0)^\top, \quad (3.48)$$

is the gravity force expressed in \mathcal{F}_g , with m_p the mass of the probe, and $g \approx 9.81 \text{ m/s}^2$ the standard acceleration due to gravity.

Since we are interested in controlling the force only along the y -axis of the probe, we define the force feature as the y -component of the force tensor in the probe contact frame:

$$s_f = \mathbf{S}_y{}^{pc}\mathbf{H}_{pc}, \quad (3.49)$$

where $\mathbf{S}_y = (0 \ 1 \ 0 \ 0 \ 0 \ 0)$.

Elastic contact model Considering soft tissues as an elastic body, we can approximate the force interaction between the probe and the body using Hooke's law, which states that the force needed to compress an elastic body is proportional to the distance of compression. This law can be written

$$F_y = k\delta y, \quad (3.50)$$

where k is the stiffness of the body, and δy is the displacement of the body's surface with respect to its relaxed position (Figure 3.11). Consequently, the interaction matrix for the force feature s_f is

$$\mathbf{L}_f = [0 \ k \ 0 \ 0 \ 0 \ 0]. \quad (3.51)$$

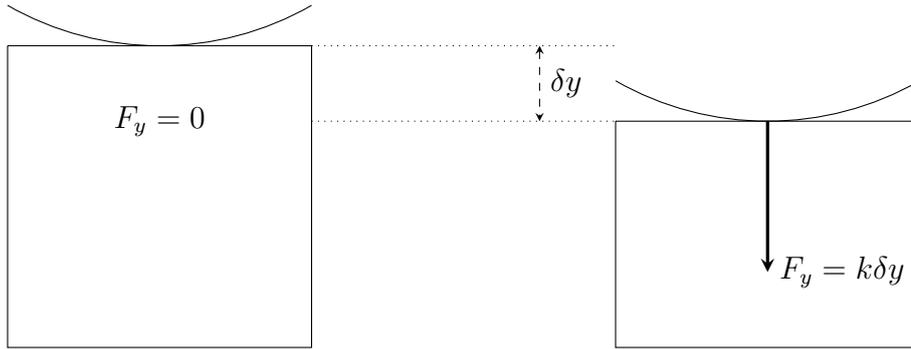


Figure 3.11 – Elastic compression model following Hooke’s law. Left: relaxed position. Right: compressed position.

Control law Let us define a constant desired contact force s_f^* , and the force feature error $e_f = s_f - s_f^*$. The visual servoing formulation leads to the following control law for the force control task:

$$\mathbf{v}_f = -\lambda_f \mathbf{L}_f^+ e_f, \quad (3.52)$$

where λ_f is the force control gain, and

$$\mathbf{L}_f^+ = \begin{bmatrix} 0 & \frac{1}{k} & 0 & 0 & 0 & 0 \end{bmatrix}^\top. \quad (3.53)$$

This control law, which can also be written as

$$\mathbf{v}_f = -\frac{\lambda_f}{k} (\mathbf{S}_y^{pc} \mathbf{F}_s ({}^s \mathbf{H}_s - {}^s \mathbf{F}_g {}^g \mathbf{H}_g) - s_f^*) \mathbf{S}_y^\top, \quad (3.54)$$

requires an estimation of the probe’s mass m_p , the transformations ${}^{pc} \mathbf{F}_s$ and ${}^s \mathbf{F}_g$, and the contact stiffness k .

According to the redundancy formalism (section 2.1.4), the force control task can be augmented with a secondary task by adding any velocity screw projected on the null-space of \mathbf{L}_f , via the projection operator

$$\mathbf{P}_f = \mathbf{I}_6 - \mathbf{L}_f^+ \mathbf{L}_f = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.55)$$

In other words, the force control task defines a constraint that allows any additional motion, except those which translate the probe along the y -axis.

3.3.2 Confidence control

Let us now design the secondary task, which consists in the confidence-driven control. The aim of this task is to maintain an optimal image quality. To this end, we define the confidence feature vector as

$$\mathbf{s}_c = \begin{pmatrix} \theta_c \\ \phi_c \end{pmatrix}, \quad (3.56)$$

where θ_c and ϕ_c are the angular confidence features defined in (3.39) and (3.40). The desired value of this feature vector is

$$\mathbf{s}_c^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (3.57)$$

The purpose of this choice is to balance the distribution of the confidence within the ultrasound volume, bringing high-confidence areas towards the center (which corresponds to $\theta_c = \phi_c = 0$).

According to (3.43) and (3.44), the interaction matrix associated to the confidence feature vector \mathbf{s}_c is

$$\mathbf{L}_c = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}. \quad (3.58)$$

From (3.52), (3.55), and (3.58), we notice that

$$\mathbf{L}_c \mathbf{P}_f = \mathbf{L}_c \quad (3.59)$$

and

$$\mathbf{L}_c \mathbf{v}_f = 0, \quad (3.60)$$

which means that the two tasks are decoupled. Consequently, the contribution of the confidence control task to the velocity screw can, in this case, be defined directly as

$$\mathbf{v}_c = -\lambda_c \mathbf{L}_c^+ (\mathbf{s}_c - \mathbf{s}_c^*), \quad (3.61)$$

where λ_c is the confidence control gain, and

$$\mathbf{L}_c^+ = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}^\top \quad (3.62)$$

is the pseudoinverse of \mathbf{L}_c .

The behavior induced by this control law is twofold:

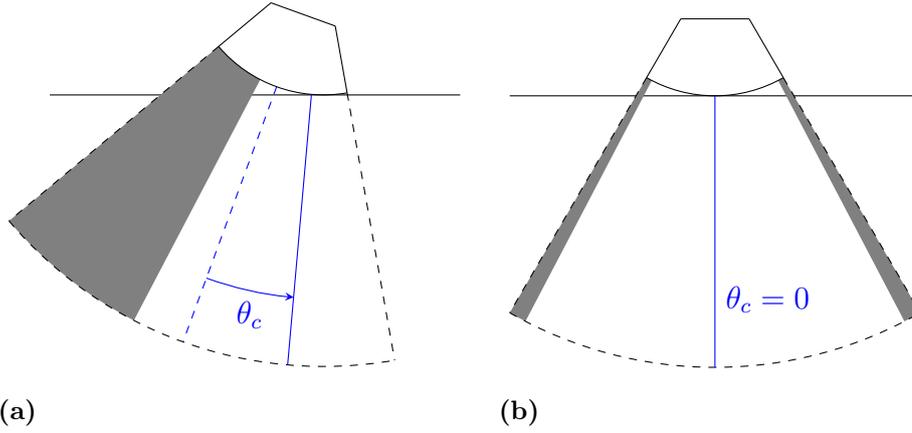


Figure 3.12 – 2D illustration of the system’s behavior with an homogeneous medium. (a) Initial state, the orientation of the probe induces a low confidence (gray area) in the left part of the image, due to a weak contact between the probe and the skin. The confidence is high in the right part of the image. (b) At convergence, the probe is orthogonal to the surface, so that the confidence is balanced in the image.

- While scanning an homogeneous region, the probe will be oriented in such a way that the contact between the transducer and the body surface is symmetrical with respect to the depth axis (see Figure 3.12). As a result, the probe will remain orthogonal to the object’s surface during the navigation.
- While scanning a region which contains a strongly shadowing object, such as a bone, the probe will be oriented in such a way that the ultrasound beam does not go through this object (see Figure 3.13). This provides a kind of obstacle-avoidance behavior during the navigation.

3.3.3 Tertiary task

The projection operator on the null-space of $(\mathbf{L}_f, \mathbf{L}_c)$ is

$$\mathbf{P}_{f,c} = \mathbf{P}_f - \mathbf{L}_c^+ \mathbf{L}_c = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3.63)$$

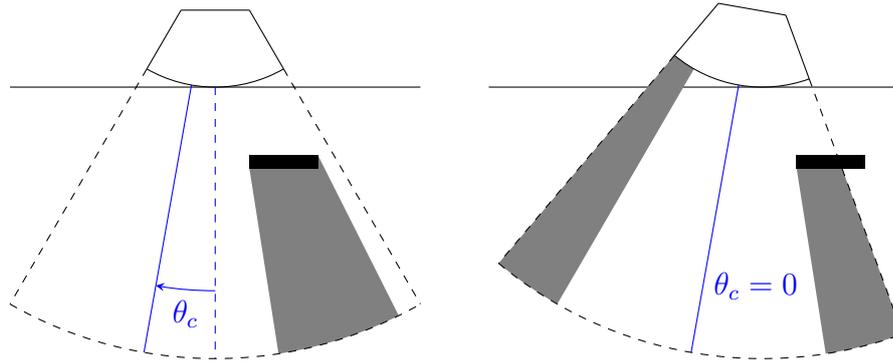


Figure 3.13 – 2D illustration of the system’s behavior with the presence of a strong acoustic attenuator. Left: initial state, the presence of an attenuator induces a low confidence (gray area) in the right part of the image, while the left part has a high confidence. Right: at convergence, the probe is oriented so as to avoid the obstacle.

so that three degrees of freedom remain available for the definition of a tertiary task. These degrees of freedom correspond to the translations along the patient’s surface, and the rotation around the y -axis. For instance, in the case of tele-echography, these motions can be controlled by the sonographer.

3.4 Target-specific confidence-driven control

In the previous section, we have proposed a framework for controlling the orientation of an ultrasound probe using the confidence map, in order to optimize the image quality. This quality optimization is global, since it does not consider any specific anatomical target. It is up to the sonographer to guide the probe towards a structure of interest, while the orientation of the probe around the x - and z -axes is automatic. However, it might also be interesting to optimize the image quality with respect to a specific target. In this section, we consider a target-specific confidence-driven control scenario, which involves three control tasks summarized as follows:

1. Maintaining a constant contact force between the probe and the patient.
2. Maintaining the target centered in the image.
3. Optimizing the image quality at the target location.

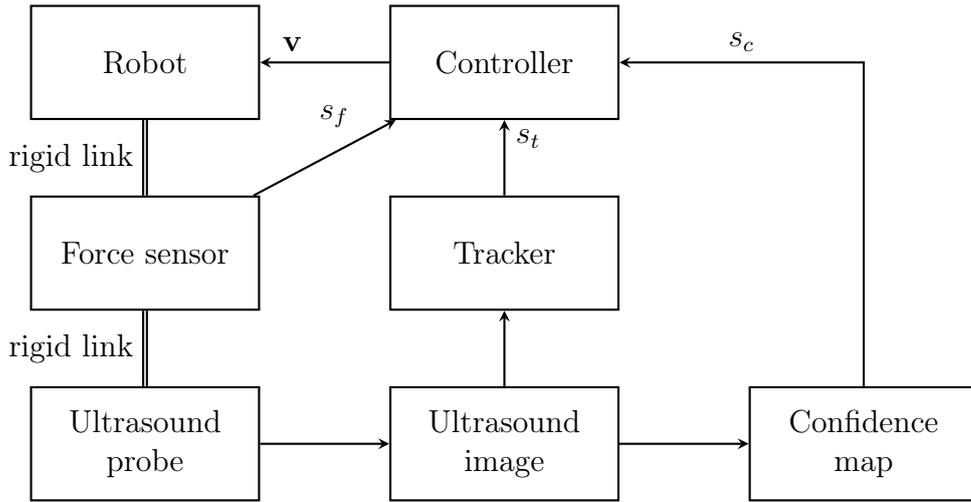


Figure 3.14 – Overview of the target-specific confidence-driven control framework.

This framework is illustrated in Figure 3.14

Again, these tasks are modeled by a hierarchical stack of tasks, with, in decreasing order of priority, the force control task $(\mathbf{e}_f^*, \mathbf{L}_f)$, the target centering task $(\mathbf{e}_t^*, \mathbf{L}_t)$, and the confidence control task $(\mathbf{e}_c^*, \mathbf{L}_c)$. The full control screw \mathbf{v} is decomposed as

$$\mathbf{v} = \mathbf{v}_f + \mathbf{v}_t + \mathbf{v}_c, \quad (3.64)$$

where we will build successively the contributions \mathbf{v}_f for force control, \mathbf{v}_t for target centering, and \mathbf{v}_c for confidence control.

We have already described the force control task in section 3.3.1. This task remains unchanged, so that we will start with the design of the target centering task.

3.4.1 Target centering

Let us consider a physical target tracked in the ultrasound image. We have reviewed in chapter 1 different algorithms for target tracking. Any of those can be used indifferently, as long as it provides the position $\mathbf{p} = (x_t, y_t, z_t)$ of the target, expressed in the probe frame, in real-time. As we have seen in section 2.2.2, the interaction matrix associated to a 3D point \mathbf{p} is

$$\mathbf{L}_p = \begin{bmatrix} -1 & 0 & 0 & 0 & -z_t & y_t \\ 0 & -1 & 0 & z_t & 0 & -x_t \\ 0 & 0 & -1 & -y_t & x_t & 0 \end{bmatrix}. \quad (3.65)$$

The target centering task can be formalized as the regulation of the visual feature vector

$$\mathbf{s}_t = \begin{pmatrix} x \\ z \end{pmatrix} \quad (3.66)$$

to $\mathbf{s}_t^* = \mathbf{0}$, corresponding to the central scan line of the volume. Note that the y -axis corresponds to the depth direction, so that it is not desirable to constrain the y -coordinate of the target.

The interaction matrix associated to \mathbf{s}_t is simply deduced from (3.65) as

$$\mathbf{L}_t = \begin{bmatrix} -1 & 0 & 0 & 0 & -z_t & y_t \\ 0 & 0 & -1 & -y_t & x_t & 0 \end{bmatrix}. \quad (3.67)$$

Projecting the interaction matrix \mathbf{L}_t onto $\ker \mathbf{L}_f$, using the projection operator \mathbf{P}_f defined in (3.55), gives

$$\mathbf{L}_t \mathbf{P}_f = \mathbf{L}_t, \quad (3.68)$$

so that the two tasks are decoupled (we also have $\mathbf{L}_t \mathbf{v}_f = \mathbf{0}$). Instead of the Moore-Penrose pseudoinverse, let us choose, among the possible generalized inverses of \mathbf{L}_t ,

$$\mathbf{L}_t^g = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (3.69)$$

which corresponds to a pure translation motion. This choice is arbitrary, and does not influence the final result. However, we will see that it allows an intuitive illustration of the contribution of the different tasks.

Noting $\mathbf{e}_t = \mathbf{s}_t - \mathbf{s}_t^*$ the target centering feature error, the contribution of the target centering task to the control screw can be defined as

$$\mathbf{v}_t = -\lambda_t \mathbf{L}_t^g \mathbf{e}_t, \quad (3.70)$$

where λ_t is the control gain of the target centering task.

Using (3.55), (3.65), and (3.69), the projection operator on $\ker(\mathbf{L}_f, \mathbf{L}_t)$ is

$$\mathbf{P}_{f,t} = \mathbf{P}_f - \mathbf{L}_t^g \mathbf{L}_t = \begin{bmatrix} 0 & 0 & 0 & 0 & -z_t & y_t \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -y_t & x_t & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3.71)$$

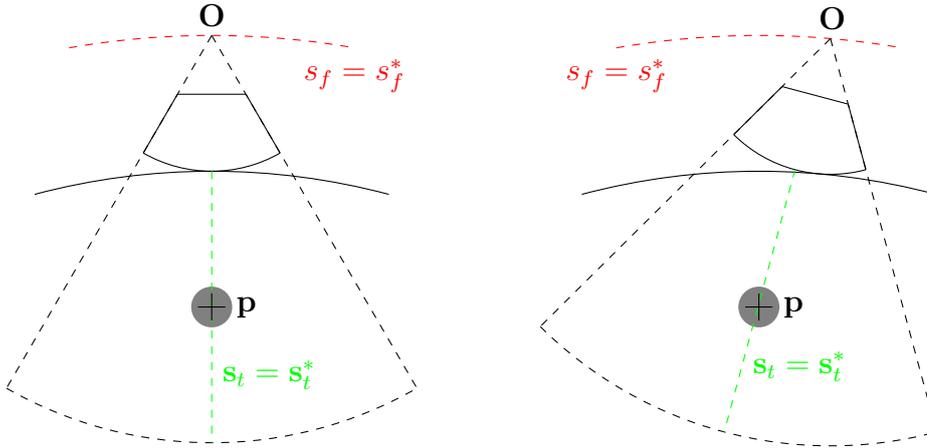


Figure 3.15 – 2D illustration of the constraints induced by the force control and target centering tasks.

which is of rank 3. The constraints induced by the force control and the target centering tasks are illustrated in Figure 3.15. The remaining degrees of freedom correspond to all the motions that leave the force contact invariant, and that keep the target centered. In other words, the allowed motions are sliding motions along the surface that keep \mathbf{p} in the central scan line.

3.4.2 Confidence control

We now address the problem of optimizing the image quality for the target \mathbf{p} . Let us note (r_t, θ_t, ϕ_t) the pseudo-spherical coordinates of \mathbf{p} . The task consists in maintaining a good image quality at the location of the target. Therefore, we define the feature vector for target-specific confidence control as

$$\mathbf{s}_c = \begin{pmatrix} \theta_c - \theta_t \\ \phi_c - \phi_t \end{pmatrix}, \quad (3.72)$$

which corresponds to the two angles between the target and the confidence barycenter. The desired value of this feature vector is $\mathbf{s}_c^* = \mathbf{0}$. With this new choice of visual features, the interaction matrix \mathbf{L}_c has to

be modified to account for the variations of θ_t and ϕ_t , so that

$$\mathbf{L}_c = \begin{bmatrix} \frac{\cos \theta_t}{r_t} & 0 \\ -\frac{\sin \theta_t \cos \phi_t}{r_t} & -\frac{\sin \phi_t}{r_t \cos \theta_t - \Delta R} \\ -\frac{\sin \theta_t \sin \phi_t}{r_t} & \frac{r_t \cos \theta_t - \Delta R}{\cos \phi_t} \\ \frac{\Delta R \sin \theta_t \sin \phi_t}{r_t} & \frac{\Delta R \cos \theta_t}{r_t \cos \theta_t - \Delta R} \\ r_t - \Delta R \cos \theta_t & \frac{r_t \cos \theta_t - \Delta R}{r_t \sin \theta_t \cos \phi_t} \\ \frac{\Delta R \cos \theta_t - r_t}{r_t} (\cos \phi_t - 1) & -\frac{r_t \sin \theta_t \cos \phi_t}{r_t \cos \theta_t - \Delta R} \\ & -\frac{r_t \sin \theta_t \sin \phi_t}{r_t \cos \theta_t - \Delta R} \end{bmatrix}^\top. \quad (3.73)$$

Then, the component of the control screw corresponding to the quality optimization task can be written:

$$\mathbf{v}_c = \mathbf{P}_{f,t}(\mathbf{L}_c \mathbf{P}_{f,t})^g (\dot{e}_c^* - \mathbf{L}_c(\mathbf{v}_f + \mathbf{v}_t)). \quad (3.74)$$

In order to simplify the resulting expression, let us consider the value of the interaction matrix at the desired pose $\mathbf{s}_t = \mathbf{s}_t^*$, as it is usually done in the literature [Chaumette and Hutchinson, 2006]. We define

$$\mathbf{L}_c^* = \mathbf{L}_c \Big|_{\substack{\theta_t = 0 \\ \phi_t = 0}} = \begin{bmatrix} \frac{-1}{r_t} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta R - r_t} & \frac{\Delta R}{\Delta R - r_t} & 0 & 0 \end{bmatrix} \quad (3.75)$$

to replace \mathbf{L}_c in (3.74), and we obtain the final velocity screw for the stack of tasks:

$$\mathbf{v} = \begin{pmatrix} \lambda_t x_t + \lambda_c (\theta_c - \theta_t) y_t \\ -\frac{\lambda_f e_f}{k} \\ \lambda_t z_t + \lambda_c (\phi_c - \phi_t) y_t \\ \lambda_c (\phi_c - \phi_t) \\ 0 \\ \lambda_c (\theta_t - \theta_c) \end{pmatrix} \quad (3.76)$$

Note that this is only a particular solution to the problem, to which one can add any element of the null space $\ker(\mathbf{L}_f, \mathbf{L}_t, \mathbf{L}_c^*)$. That is to say, any scalar multiple of $(-z_t \ 0 \ x_t \ 0 \ 1 \ 0)^\top$ can be added to \mathbf{v} without disturbing the three tasks. This remaining degree of freedom corresponds to a rotation around the axis passing through the imaging center and the target. It can be used to build a fourth control task. For instance, in the context of ultrasound-guided needle insertion, the rotation around the center-target axis could be exploited to align the lateral axis of the probe with the needle shaft, in order to improve its visibility.

3.5 Conclusion

We proposed a general framework for controlling an ultrasound probe based on ultrasound confidence maps. By studying the dynamics of the confidence map, we showed that the distribution of the confidence within the image was strongly linked to the orientation of the probe. Based on this observation, we defined adequate visual features computed from the confidence map, and proposed a first control law for the orientation of the probe, aiming at optimizing the image quality. Then, we proposed a second control law which takes into account the position of an anatomical target of interest. We showed that, using the redundancy formalism, it is possible to obtain a control law that simultaneously centers the target in the image and optimizes the acoustic window for this target. The next chapter will be dedicated to the experimental validation and analysis of the proposed control frameworks.

Chapter 4

Experimental Results

In chapter 3, we have proposed a control framework for optimizing the quality of ultrasound imaging, using a visual servoing approach based on ultrasound confidence maps. In particular, we have introduced a global confidence-driven visual servoing method, aimed at optimizing the image quality globally. Then, we have proposed a target-specific confidence-driven visual servoing method, aimed at optimizing the image quality relative to a specific target which is tracked in the ultrasound images.

In this chapter, we report some experimental results that validate the proposed control methods, through (i) an experimental study of the convergence of the system, and (ii) tests of the system's reaction to disturbances. We also provide experimental demonstrations of the use of our method in tele-echography. Experiments were conducted on two different robotic platforms, and include both experiments on phantoms and a validation on a human volunteer.

This chapter is organized as follows. In section 4.1, we describe the experimental setup used in our study, and we provide some details on the implementation of the methods introduced in chapter 3. In section 4.2, we report the results of a series of experiments dedicated to the analysis of the system's convergence. In section 4.3, we present another series of experiments, dedicated to the analysis of the system's reaction to disturbances such as motion and shadowing. In section 4.4, we provide an illustration for the application of our framework to tele-echography, and we report experimental results obtained on a human volunteer.

4.1 Experimental setup

We summarize in this section the different components used in the experiments. The general setup is illustrated in Figure 4.1. It consists in

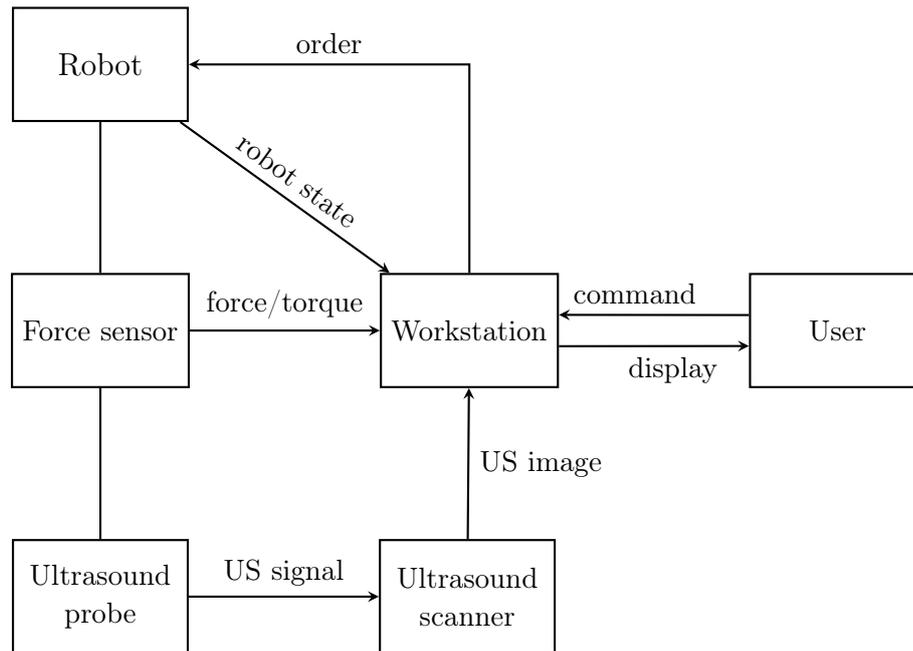


Figure 4.1 – Experimental setup.

an ultrasound probe rigidly attached to the end-effector of a robotic manipulator. The robot is also equipped with a force sensor. A workstation receives and processes the ultrasound images transmitted by the ultrasound scanner, the state of the robot, and the force measurements. The workstation is responsible for computing the control law and sending it to the robot. It also allows interaction with the user. We describe in section 4.1.1 the equipment used in the experiments, and we provide in section 4.1.2 some details on the software implementation.

4.1.1 Equipment

This thesis was conducted in the context of a collaboration between the Université de Rennes 1 and the Technische Universität of München. As a result, experiments could be performed in laboratories attached to both institutions, each with a different experimental setup. This was an opportunity to test our framework on different platforms. To this end, the implementation of the method was made generic enough, so that the same applications can be run independently of the physical equipment. Thereafter, we provide a short description of hardware components that were used in the experiments. Then, for each experiment in the following sections, we will recall in a table which specific equipment (robot, scanner, probe, phantom) was used in that experiment.

4.1.1.1 Ultrasound systems

Three different ultrasound scanners were used in the experiments:

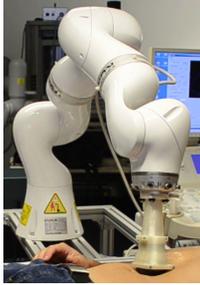
<p>The Sonix RP from BK Ultrasound, MA (formerly Ultrasonix Medical Corporation, Canada). This ultrasound system provides an Ultrasound Research Interface, which allows the development of custom acquisition protocols directly on the scanner's computer. In particular, it provides access to raw data, such as radio frequency and pre-scan converted B-mode ultrasound. Data can be transmitted on a local network through Ethernet.</p>	
<p>The SonixTOUCH, another scanner from BK Ultrasound. This model has a touch screen interface.</p>	
<p>The SonoSite 180PLUS (FUJIFILM SonoSite Inc., USA), a portable ultrasound system. The scanner's screen is captured on the workstation using a video grabber.</p>	

These systems were used with the following probes:

<p>The 4DC7-3/40 Convex 4D ultrasound probe from BK Ultrasound. This is a motorized probe, such as described in section 1.1.2.4. It can be used either in 2D or 3D mode. The curved array transducer has a radius of 39.8 mm and a frequency range from 3 MHz to 7 MHz. The motor radius is 27.25 mm.</p>	
<p>The SonoSite C60 convex ultrasound probe from SonoSite. This is a curved array ultrasound transducer with a radius of 60 mm and a frequency range from 2 MHz to 5 MHz.</p>	

4.1.1.2 Robots

Two different robotic manipulators were used in the experiments:

<p>The Viper s850 (Omron Adept Technologies, Inc., CA), a 6-axis industrial manipulator. This robot does not have any integrated force sensor. Therefore, we used a 6-axis force/torque sensor rigidly attached to the robot's end-effector. The force/torque sensor was a Gamma IP65 SI-65-5 (ATI Industrial Automation, NC), with a force sensing resolution of 25 mN and a torque sensing resolution of 75 μN m.</p>	
<p>The KUKA LWR iiwa R800 (KUKA Roboter GmbH, Germany), a 7-axis lightweight manipulator. This robot has internal torque sensors in each joint, so that it was not necessary to use an additional force sensor.</p>	

The ultrasound probes were rigidly attached to the end-effector of the robot. For the Viper s850 robot, we used a metallic mount. For the KUKA robot, we used a 3D printed plastic mount.

4.1.1.3 Phantoms

For phantom experiments, we used the following phantoms:

<p>A CIRS Multi-Purpose Multi-Tissue Ultrasound Phantom (model 040GSE, Computerized Imaging Reference Systems, Inc. (CIRS), USA), a calibration phantom.</p>	
<p>An ABDFAN ultrasound examination training model (Kyoto Kagagu Co., Japan), simulating a patient's abdomen with internal organs.</p>	

4.1.2 Implementation details

Let us now provide some details on the implementation of the methods proposed in chapter 3.

The image processing and control algorithms were implemented in C++, using the Visual Servoing Platform (ViSP) library [Marchand et al., 2005]. The application, with user interface and thread management, was implemented using the Qt library. For the KUKA robot, communication with the robot's controller was achieved through the Robot Operating System (ROS) software platform [Quigley et al., 2009].

In order to achieve a real-time behavior, we use a multi-threaded architecture. The following tasks take place in parallel threads, as illustrated in Figure 4.2:

- **Grabber:** this thread is responsible for grabbing the prescan ultrasound images. For the Sonix scanners, images are acquired as soon as they are received through the network. For the SonoSite scanner, images are grabbed at the rate of 25 frames per second. When a new prescan image is available, the grabber sends it to the main thread. Thus, the visual servoing control law is updated at the same rate as the data acquisition rate.
- **Main:** the main thread is in charge of user interaction, display, computation of the visual features and control law. It also takes care of the management of the other threads.
- **Converter:** this thread takes care of the conversion of ultrasound images and confidence map from prescan to postscan format.
- **Tracker:** this thread is responsible for target tracking.
- **Confidence:** this thread performs the estimation of the confidence map.
- **Control:** this thread implements the control of the robot. It translates the desired Cartesian control velocity into joint velocities, and it is in charge of force control. In order to have a responsive force control, the joint velocities are updated every 5 ms.

4.2 Convergence analysis

We start by presenting an experimental evaluation of the convergence of the proposed control laws in a static environment.

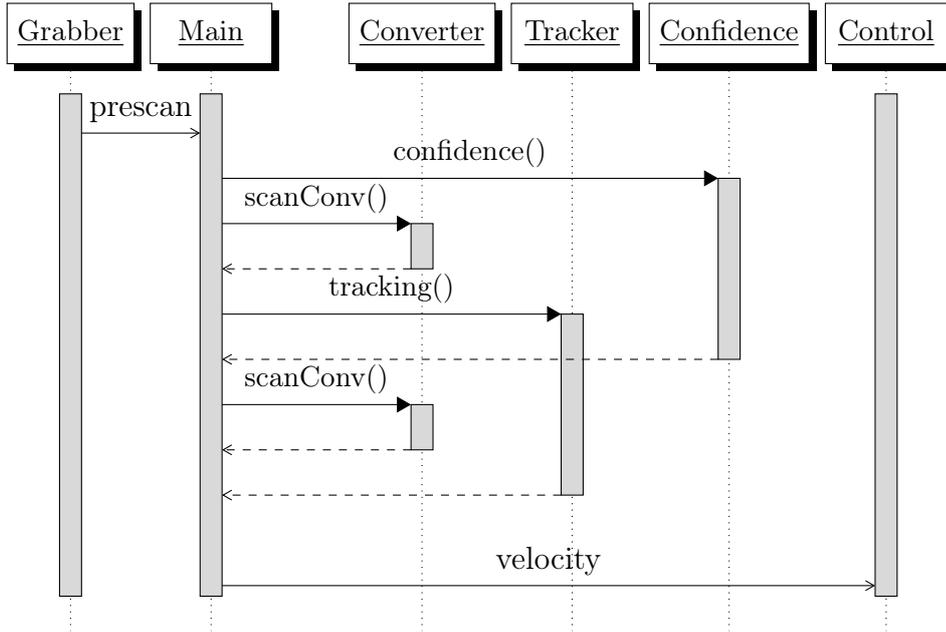


Figure 4.2 – Multi-threaded processing of the control application. Each column corresponds to a different thread.

4.2.1 Global confidence control

Experimental setup			
robot	scanner	probe	phantom
KUKA iiwa	SonixRP	4DC7 (3D mode)	CIRS

First, let us consider the global confidence-driven control framework presented in section 3.3. We recall that the aim of this control framework is to globally optimize the image quality, by adjusting the orientation of the probe. No target tracking was performed in this experiment.

To assess the convergence of the proposed control law, the probe was initially placed at an angle from the normal to the phantom surface, barely in contact with it, as in Figure 4.3(a). Then, the control law was activated, and we let the system converge in a static environment. The control gains were set to $\frac{\lambda_f}{k} = 5.0 \times 10^{-4} \text{ m N}^{-1} \text{ s}^{-1}$ for the force control task and $\lambda_c = 0.8 \text{ s}^{-1}$ for the confidence control task. The system systematically converged to such a position that the probe was roughly orthogonal to the surface of the phantom, as shown in Figure 4.3(b).

The evolution of the contact force and confidence feature errors during a convergence experiment is presented in Figure 4.4. The mean absolute final confidence feature errors were 0.12° for θ_c and 0.23° for ϕ_c .

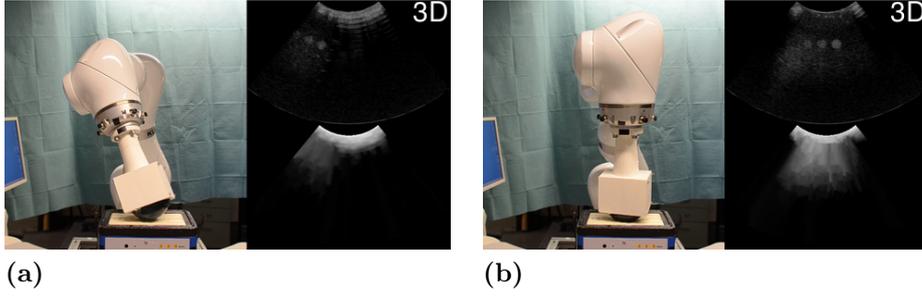


Figure 4.3 – 3D global confidence-driven control. (a) Initial configuration – the probe is not fully in contact with the phantom. (b) The system has converged, and the image quality is higher.

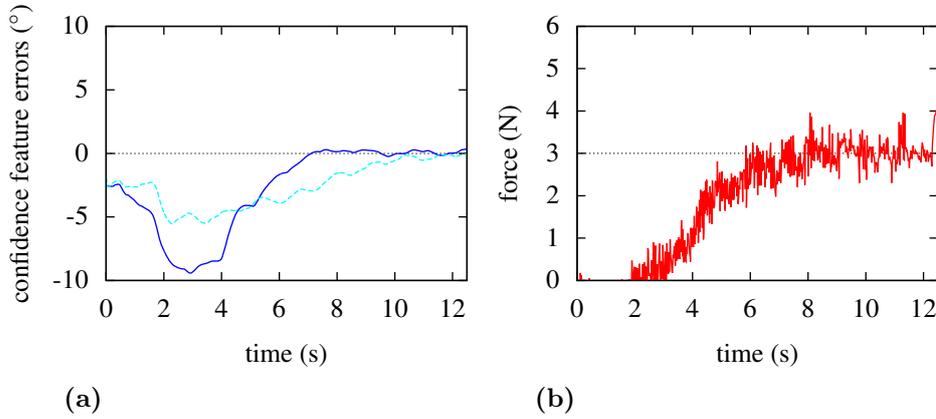


Figure 4.4 – Convergence of the global confidence-driven control in a static environment. (a) Confidence feature errors e_θ (—) and e_ϕ (- - -) in degrees. (b) Contact force s_f in Newtons.

Therefore, our system successfully regulates the angular features to their desired value with an excellent precision. From Figure 4.4(a), one can note that the confidence feature errors are initially underestimated. At $t = 0$ s, we have $\theta_c \approx \phi_c \approx 2.5^\circ$, while the probe is oriented at an angle of approximately 10° from the normal to the surface. However, the estimation gets more precise between $t = 0$ s and $t = 3$ s, which explains the apparent increase of the confidence feature errors during that time interval. After $t = 3$ s, both feature errors decrease towards 0° . As a result, the proposed control law leads to the expected behavior. That is to say, the ultrasound probe is automatically re-oriented perpendicular to the surface of the phantom, which maximizes the image quality.

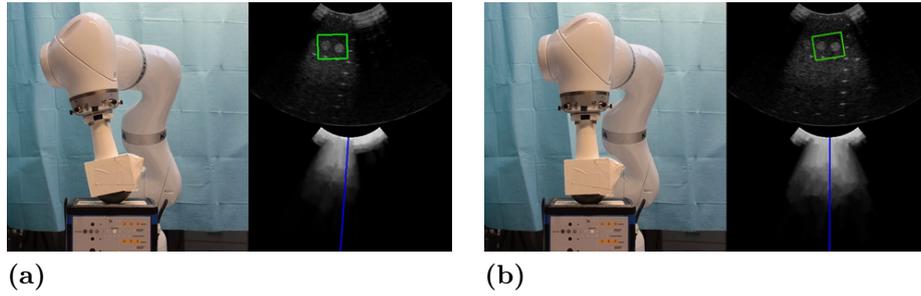


Figure 4.5 – Target-specific confidence-driven control. The green rectangle displayed on the ultrasound image represents the target. The blue line displayed on the confidence map represents the angular confidence feature. (a) Initial configuration. (b) The system has converged. The target is centered, and the image quality is higher.

4.2.2 Target-specific confidence control

Experimental setup			
robot	scanner	probe	phantom
Viper s850	SonixTOUCH	4DC7 (2D mode)	ABDFAN

Let us now study the convergence of the target-specific confidence-driven control law, which was presented in section 3.4. For this experiment, the probe was placed at an angle from the normal to the surface, such that the phantom contents was partly visible in the image, as in Figure 4.5(a). Then, a rectangular region of interest was selected manually by clicking in the live ultrasound view, and the control law was activated. The region of interest was tracked using the intensity-based approach described in section 1.2.1.1. The system converged, as expected, to such a position that the region of interest was centered in the field of view, and the probe was roughly orthogonal to the normal of the surface, as shown in Figure 4.5(b).

To study the convergence accuracy of the system, 25 experiments were performed. These experiments are grouped in 5 series with different control gain settings. The settings for each series are summarized in Table 4.1. In all experiments, the desired contact force was set to $s_f = 4$ N. The initial configuration was such that

- the initial contact force ranged from 3.73 N to 4.55 N,
- the initial target feature error (distance between the target and the central scan line) ranged from 4.29 cm to 5.58 cm,

Table 4.1 – Control gain settings used in the target-specific confidence-driven control convergence experiment series.

series	$\frac{\lambda_f}{k}$ (m N ⁻¹ s ⁻¹)	λ_t (s ⁻¹)	λ_c (s ⁻¹)
A	1.0×10^{-3}	0.2	0.3
B	2.0×10^{-3}	0.4	0.4
C	2.0×10^{-3}	0.6	0.6
D	2.0×10^{-3}	0.8	0.8
E	2.0×10^{-3}	1.0	1.0

- and the initial angular confidence feature error ranged from 1.0° to 7.8°.

For each experiment, we computed the mean feature errors at convergence. In order to have a common reference to compare the different experiments, we used the characteristic time τ of the system to define the convergence time. Specifically, τ was defined as

$$\tau = \frac{1}{\min(\lambda_t, \lambda_c)}, \quad (4.1)$$

which is the characteristic time of the slowest control component. Then, we defined the final feature errors for each feature $s \in \{s_f, s_t, s_c\}$ as the average of $|s - s^*|$ in the time slot $5\tau < t < 10\tau$, where t is the time elapsed from the activation of the control law.

The final feature errors for each experiment are reported in Figure 4.6. The maximum observed final feature errors were 0.3 N for the force control task, 0.16 mm for the target centering task, and 1.0° for the confidence control task. Thus, the force control gain impacts the convergence accuracy of the contact force feature, which was in average 0.06 N for $\frac{\lambda_f}{k} = 1.0 \times 10^{-3}$ m N⁻¹ s⁻¹, and 0.23 N for $\frac{\lambda_f}{k} = 2.0 \times 10^{-3}$ m N⁻¹ s⁻¹. On the other hand, the control gains for the target centering task and the confidence control task did not appear to impact significantly the convergence accuracy of the associated features.

The evolution of the system features during one of the convergence experiments of series E is presented in Figure 4.7. In this particular example, the mean final feature errors were 0.19 N for the force control task, 0.13 mm for the target centering task, and 1.0° for the confidence control task. These results indicate that our system can reach an excellent precision simultaneously for target centering and confidence control.

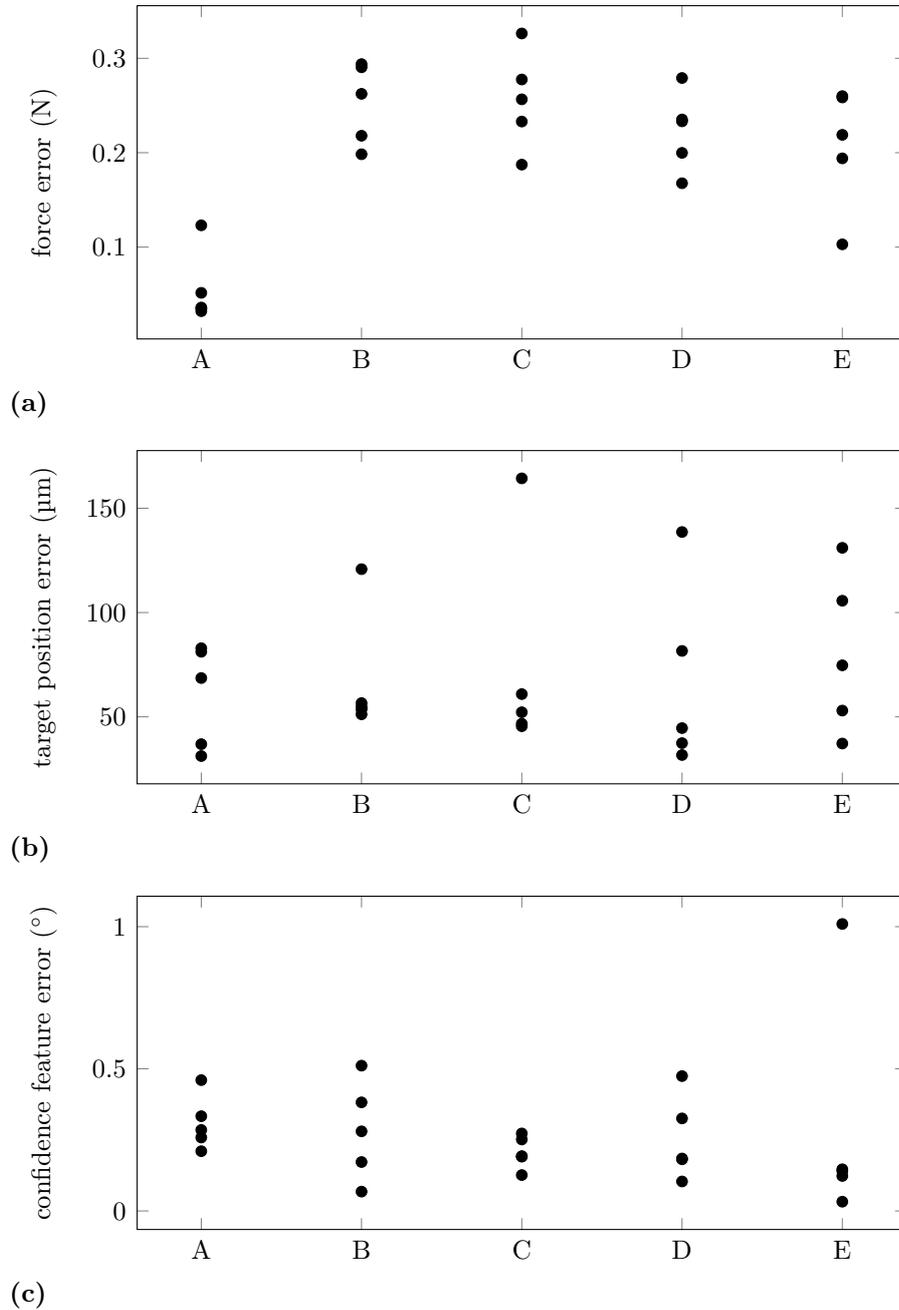


Figure 4.6 – Convergence accuracy of the target-specific confidence-driven control. Each dot corresponds to the final feature error for one experiment. (a) Force error e_f in Newtons. (b) Target centering error e_t in micrometers. (c) Confidence feature error e_c in degrees.

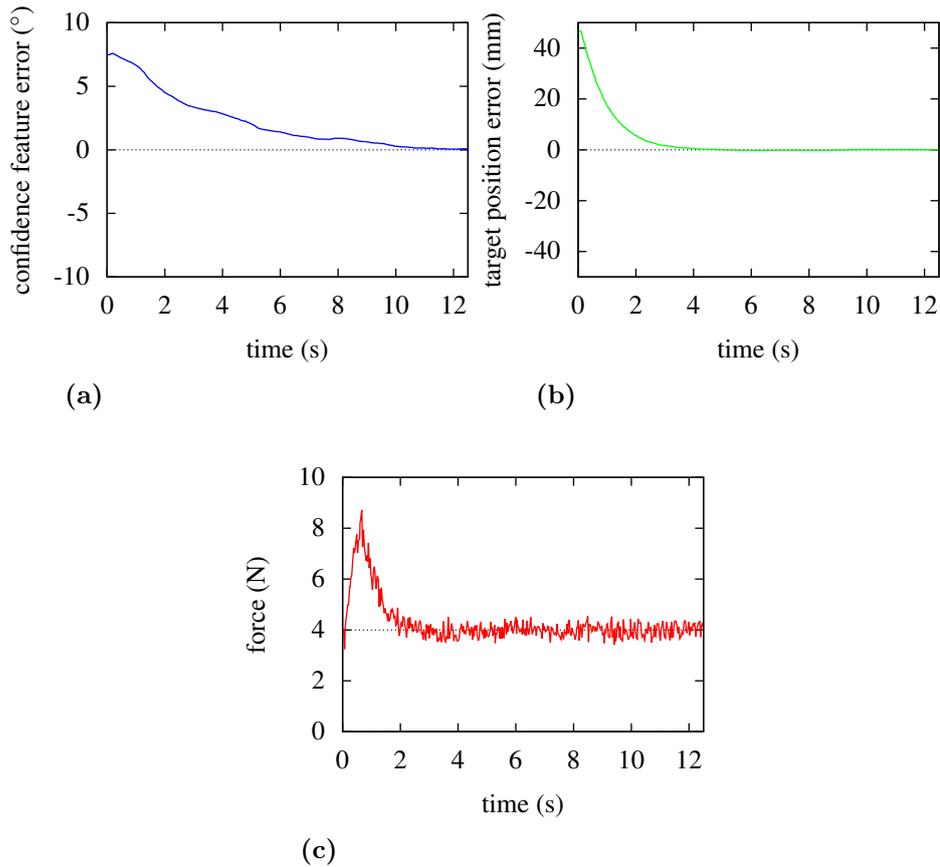


Figure 4.7 – Convergence of the target-specific confidence-driven control in a static environment. (a) Confidence feature error e_c in degrees. (b) Target centering error e_t in millimeters. (c) Contact force s_f in Newtons.

4.3 Reaction to disturbances

In the previous section, we have reported experimental results on the convergence of our confidence-driven control approaches in a static environment. We will now present the results of experiments aimed at testing the reaction of the system to disturbances such as shadowing and motion.

4.3.1 Global confidence-driven control

Experimental setup			
robot	scanner	probe	phantom
KUKA iiwa	SonixRP	4DC7 (3D mode)	CIRS

We start by describing an experiment that illustrates the behavior of the global confidence-driven controller in the presence of shadows. To this end, a band of paper was inserted manually between the ultrasound probe and the phantom, thus generating a strong shadow in the ultrasound image, as illustrated in Figure 4.8. During the experiment, the band of paper was inserted and withdrawn on different sides of the ultrasound probe. Consequently, our system automatically reoriented the probe in order to avoid the shadowed area. The control gains were set to $\frac{\lambda_f}{k} = 5.0 \times 10^{-4} \text{ m N}^{-1} \text{ s}^{-1}$ for the force control task, $\lambda_t = 0.8 \text{ s}^{-1}$ for the target centering task, and $\lambda_c = 0.8 \text{ s}^{-1}$ for the confidence control task.

The evolution of the confidence features during a shadowing experiment is represented in Figure 4.9, where the 4 gray areas correspond to the periods during which a shadow was generated:

- At $t = 21 \text{ s}$, the band of paper was inserted on the right¹ [Figure 4.8(c)], resulting in a negative shift of θ_c , which triggered an in-plane rotation of the probe. Similarly, when the band of paper was withdrawn at $t = 27 \text{ s}$, we observed a positive shift of θ_c , and the probe moved back to its original position. During this perturbation, the largest absolute deviation of the features was 3.12° for θ_c and 0.84° for ϕ_c .
- From $t = 33 \text{ s}$ to $t = 37 \text{ s}$, the band of paper was inserted at the rear of the probe [Figure 4.8(d)], inducing a perturbation of ϕ_c .
- From $t = 45 \text{ s}$ to $t = 50 \text{ s}$, the band of paper was inserted on the left of the probe, inducing a perturbation of θ_c .

¹The terms *right*, *left*, *rear* and *front* refer to the viewpoint of Figure 4.8.

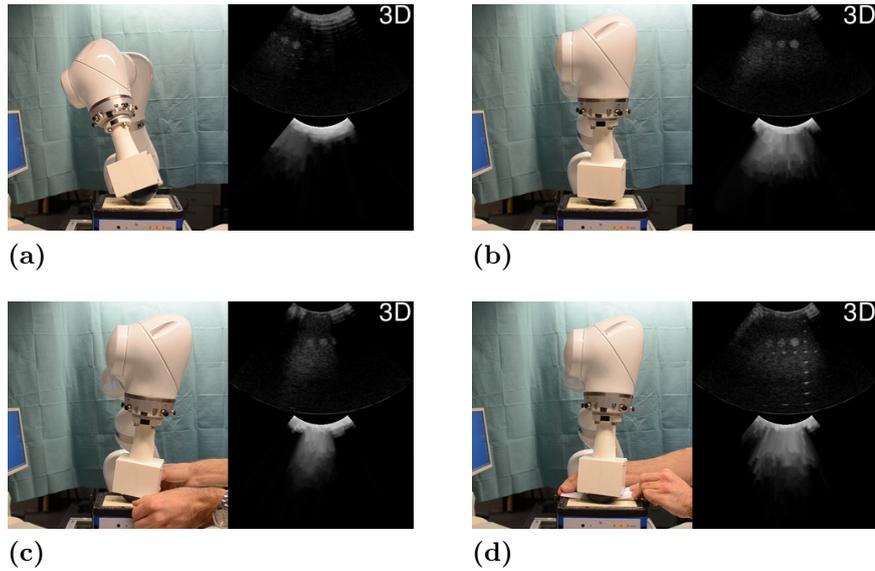


Figure 4.8 – 3D global confidence-driven control. (a) Initial configuration – the probe is not fully in contact with the phantom. (b) The system has converged, and the image quality is higher. (c) A shadow is generated using a paper band, which is inserted between the probe and the phantom by the experimenter. The system reacts automatically by reorienting the probe.

- From $t = 57$ s to $t = 61$ s, the band of paper was inserted at the front of the probe, inducing a perturbation of ϕ_c .

The results of these 4 perturbations are summarized in Table 4.2. We observe that the absolute deviation is significantly higher for the feature which corresponds to the direction of the perturbation (highlighted in bold). Thus, this experiment validates the discriminative power of the chosen features.

Table 4.2 – Maximum absolute deviation of the angular confidence features during 4 different perturbations of the system. Values in bold correspond to the main direction of the perturbation

direction	start	end	max $ e_\theta $	max $ e_\phi $
right	21 s	27 s	3.12°	0.84°
rear	33 s	37 s	1.42°	6.55°
left	45 s	50 s	4.51°	1.60°
front	57 s	61 s	0.74°	4.02°

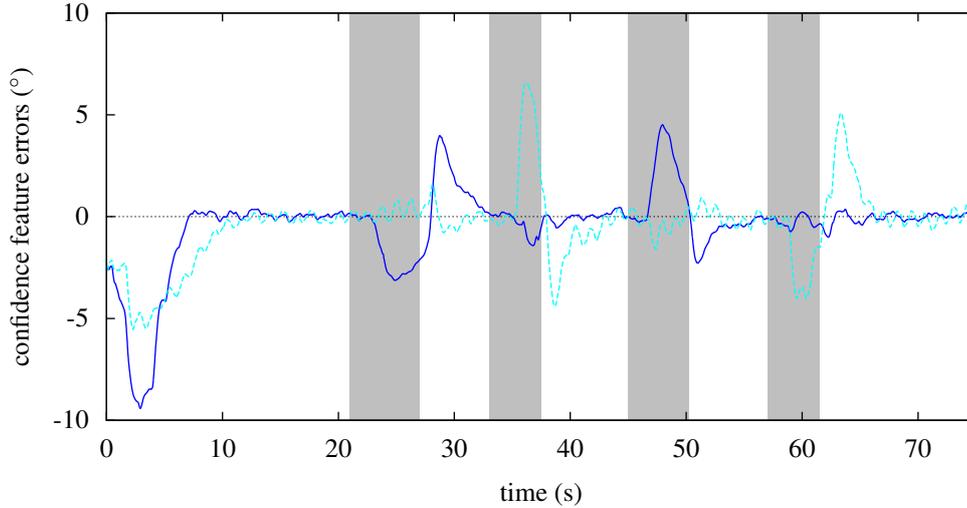


Figure 4.9 – Evolution of the confidence feature errors θ_c (—) and ϕ_c (---), in degrees, during a shadow avoidance experiment under global confidence-driven control. The four gray areas corresponds to the periods of disturbance, during which a band of paper was manually inserted between the probe and the phantom’s surface in order to generate a shadow.

4.3.2 Target-specific confidence-driven control

Experimental setup			
robot	scanner	probe	phantom
KUKA iiwa	SonixRP	4DC7 (2D mode)	CIRS

Let us now consider the target-specific confidence-driven control. In order to test the tracking capability of the system, we disturbed the equilibrium of the system by manually moving the phantom, as illustrated in Figure 4.10(c). In addition, we used the same method as previously, with a paper band, to generate a shadow in the ultrasound image [Figure 4.10(d)]. The control gains were set to $\frac{\lambda_f}{k} = 5.0 \times 10^{-4} \text{ m N}^{-1} \text{ s}^{-1}$ for the force control task, $\lambda_t = 0.8 \text{ s}^{-1}$ for the target centering task, and $\lambda_c = 0.8 \text{ s}^{-1}$ for the confidence control task.

The evolution of the confidence and target centering errors during an experiment is represented in Figure 4.11, where the gray areas correspond to periods of disturbance:

- Motion: The phantom was manually moved [Figure 4.10(c)] forward ($24 \text{ s} < t < 33 \text{ s}$) and backward ($34 \text{ s} < t < 47 \text{ s}$). The robot

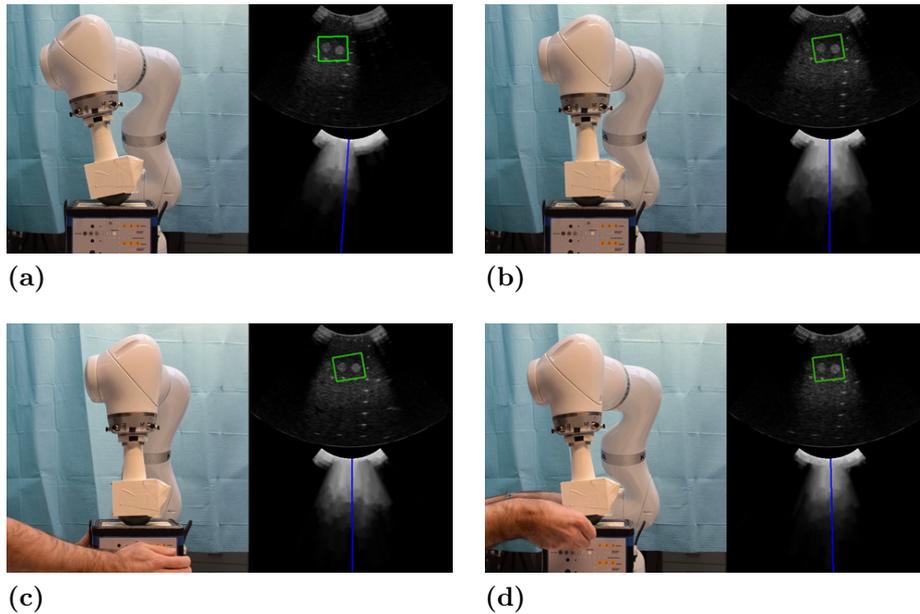
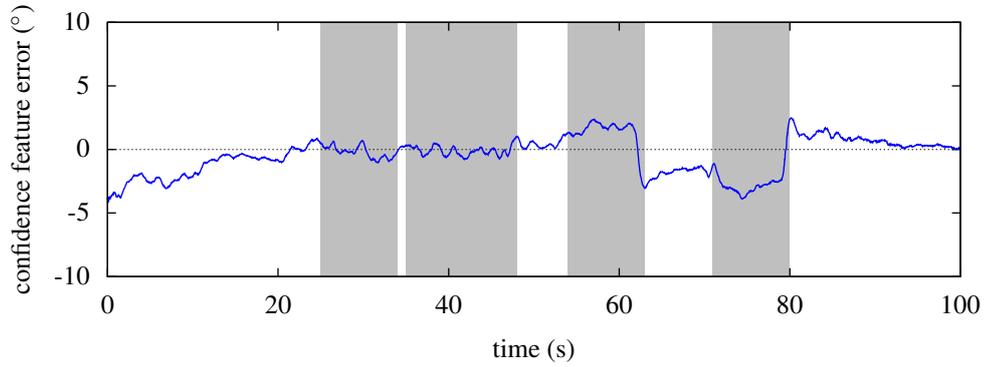


Figure 4.10 – Target-specific confidence-driven control. The green rectangle displayed on the ultrasound image represents the target. The blue line displayed on the confidence map represents the angular confidence feature. (a) Initial configuration. (b) The system has converged. The target is centered, and the image quality is higher. (c) The system tracks the moving target, while maintaining a constant image quality. (d) The system adapts to a manually generated shadow, while keeping the target centered.

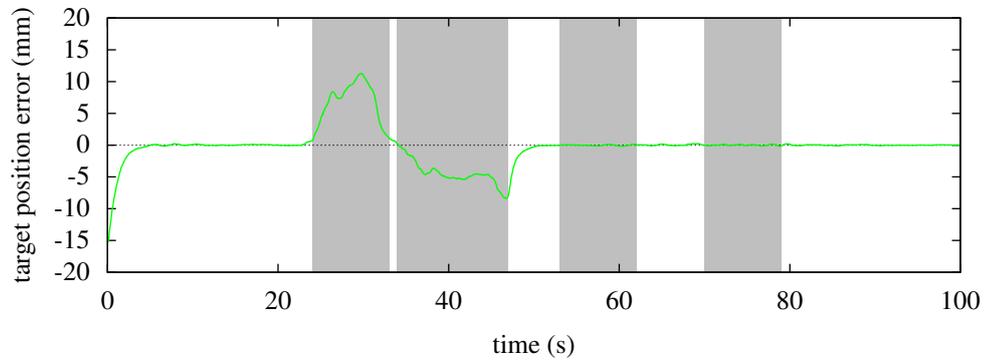
automatically compensated by moving the probe along, while keeping an optimal orientation. The maximal target centering error was 11.25 mm, and the maximal confidence feature error was 1.04° .

- Shadowing ($53\text{ s} < t < 62\text{ s}$) and ($70\text{ s} < t < 79\text{ s}$): Similar to section 4.3.1, a band of paper was manually inserted between the probe and the phantom [Figure 4.10(d)] in order to generate an important shadow. The robot automatically compensated by changing the orientation of the probe, while keeping the region of interest centered. The maximal confidence feature error was 3.89° , and the maximal target centering error was 0.18 mm.

The maximal feature errors of the 3 tasks for each perturbation are reported in Table 4.3. We observe a good decoupling of the different tasks. In particular, the effect of shadow compensation on the target positioning task is negligible, with a maximum displacement of 0.18 mm.



(a)



(b)

Figure 4.11 – Evolution of the features during target-specific confidence-driven control, in a dynamic environment. The gray areas correspond to periods of disturbance (moving phantom, shadowing). (a) Confidence feature error θ_c in degrees. (b) Target centering error e_t in millimeters.

Table 4.3 – Maximum absolute feature errors during 4 different perturbations of the system. Values in bold correspond to the feature which was specifically excited by the perturbation

perturbation	start	end	max $ e_c $	max $ e_t $	max $ e_f $
motion right	24 s	33 s	1.04°	11.26 mm	1.47 N
motion left	34 s	47 s	0.77°	8.34 mm	0.95 N
shadow left	53 s	62 s	2.35°	0.18 mm	0.95 N
shadow right	70 s	79 s	3.89°	0.18 mm	0.86 N

4.3.2.1 Evaluation of target tracking quality

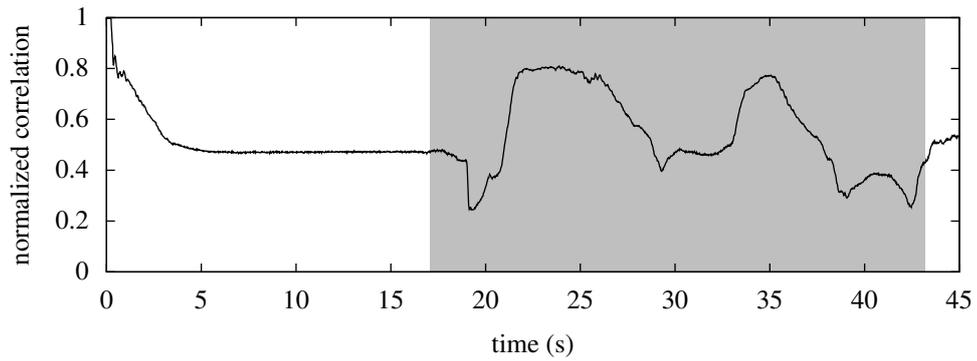
In order to evaluate the benefit of performing confidence-driven control during target tracking, we propose an investigation of the impact of our confidence-driven control scheme on the tracking quality. To this end, we use the normalized correlation between the intensities $\mathbf{s}_{\text{track}}$ in the tracked region of interest and those of the initial template, $\mathbf{s}_{\text{track}}^*$. Let us recall the expression of the normalized correlation, which was already introduced in section 1.2.1.1:

$$\rho(\mathbf{s}_{\text{track}}, \mathbf{s}_{\text{track}}^*) = \sum_i \frac{(\mathbf{s}_{\text{track}}(i) - \bar{\mathbf{s}}_{\text{track}})(\mathbf{s}_{\text{track}}^*(i) - \bar{\mathbf{s}}_{\text{track}}^*)}{\sigma(\mathbf{s}_{\text{track}})\sigma(\mathbf{s}_{\text{track}}^*)}, \quad (4.2)$$

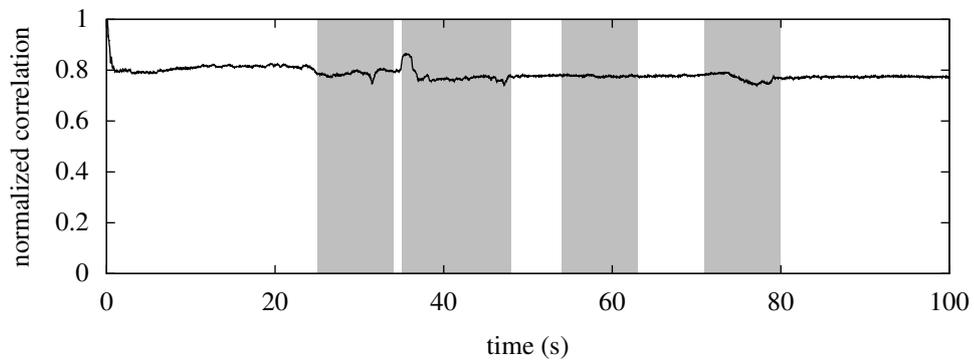
where $\bar{\mathbf{s}}_{\text{track}}$ and $\sigma(\mathbf{s}_{\text{track}})$ are the mean and standard deviation of $\mathbf{s}_{\text{track}}$, respectively. The normalized correlation ρ takes values in $[-1, 1]$, where 1 denotes a perfect correlation. Note that the normalized correlation can account both for tracking errors and for a decrease in the signal quality. Therefore, it is an interesting tool to measure the tracking quality.

To assess the impact of confidence control on the quality of target tracking, we compared the evolution of the normalized correlation during two region tracking experiments, first without confidence control, and then with confidence control. In both experiments, the target was initialized manually in a region with high confidence. Results are displayed in Figure 4.12, where the gray areas correspond to external disturbances. We observe that, in the experiment with confidence control, the normalized correlation was high and stable, at 0.78 ± 0.02 . On the other hand, without confidence control, the normalized correlation took significantly lower values and varied greatly (0.53 ± 0.15) when the phantom was moved. This is because nothing prevented the target region from being shadowed.

To better visualize the evolution of the normalized correlation, we colored the target region according to the value of ρ in the live ultrasound display. The color of the rectangle is set, in RGB format, as $(1 - \max(\rho, 0), \max(\rho, 0), 0)$, so that the rectangle is red when $\rho \leq 0$, green when $\rho = 1$, and its color varies continuously in between. Figure 4.13 illustrates the target tracking experiment without confidence control. As nothing prevents the target from being shadowed, the template correlation drops significantly in Figure 4.13(b), which leads to a failure of the tracking algorithm.



(a) Without confidence control



(b) With confidence control

Figure 4.12 – Evolution of the normalized correlation between the tracked target and the initial template, (a) without and (b) with confidence control.

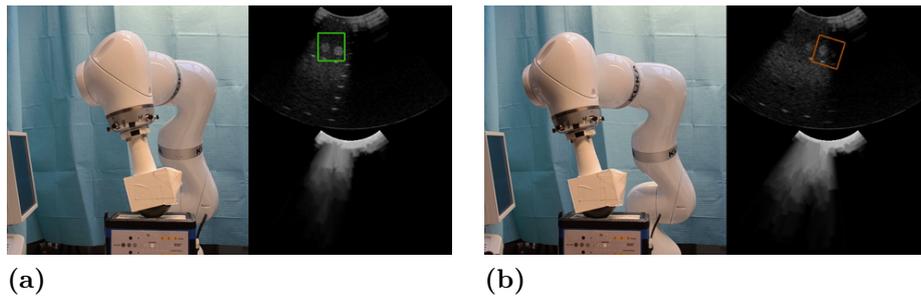


Figure 4.13 – Target tracking without confidence control. The rectangle displayed on the ultrasound image represents the target. Its color indicates the tracking quality, with green corresponding to a high correlation with the initial template, and red a low correlation with the template. (a) Initial configuration. (b) Without confidence control, the visibility of the target is not guaranteed, which can result in a tracking failure.

4.4 Confidence-optimized tele-echography

In the previous sections, we have reported experimental results which illustrate the behavior of the proposed control laws in a well-controlled environment. In particular, we have investigated the convergence accuracy of the different features, and we have tested the reaction of the system to specific disturbances, such as motion and shadowing.

We now illustrate the use of the global confidence-driven control scheme for teleoperation. The objective of this control scheme is to ease the manipulation of the probe, by optimizing its orientation during the teleoperated examination. The robotic system is controlled according to (3.45), so that the control is shared between the computer and the human operator.

4.4.1 Phantom experiment

Experimental setup			
robot	scanner	probe	phantom
Viper s850	SonixTOUCH	SonoSite	ABDFAN

During teleoperation, the user has the control over the two translation motions of the probe along the surface. Translation along the y-axis (depth direction) is managed by the force controller. The orientation of the probe is automatically adapted by the confidence controller to provide a good global image quality. As a result, high image quality is maintained even when scanning shadowed areas such as behind the ribs (Figure 4.14).

The evolution of the system's features during a teleoperation experiment is represented in Figure 4.15. Between the frames presented in Figure 4.14(c) and (d), the probe was guided by the operator towards the side of the phantom, over the ribs. The peak in feature error observed around $t = 40$ s in Figure 4.15(a) corresponds to the instant when the probe reached the rib cage. A peak in the contact force, due to the salient ribs, occurred at the same instant. The feature error then quickly decreased thanks to our confidence-based control scheme. As a result, although the ribs are strongly shadowing, our confidence-based visual servoing method allowed the probe to be automatically re-oriented in a way that maintained the visibility of inner tissues.

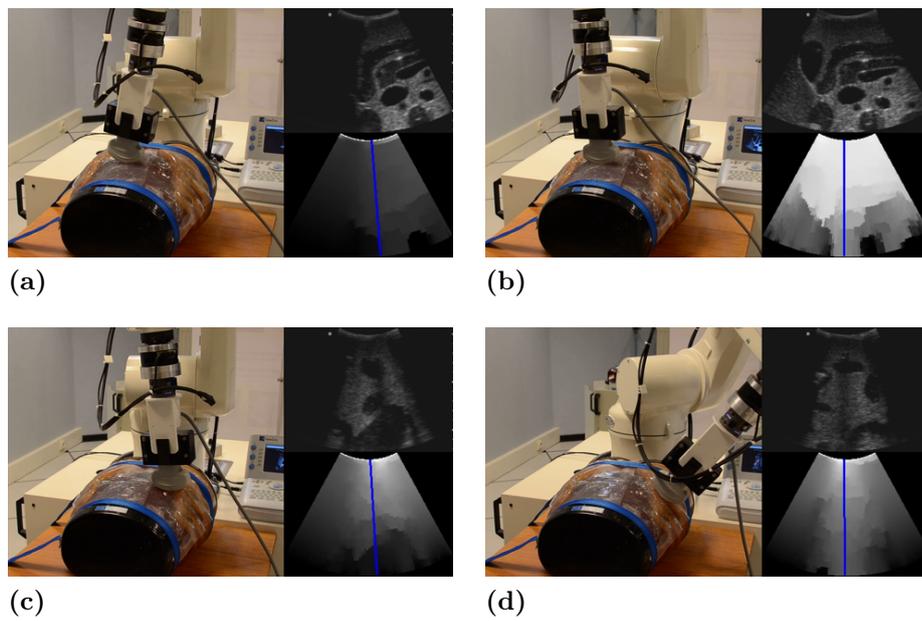
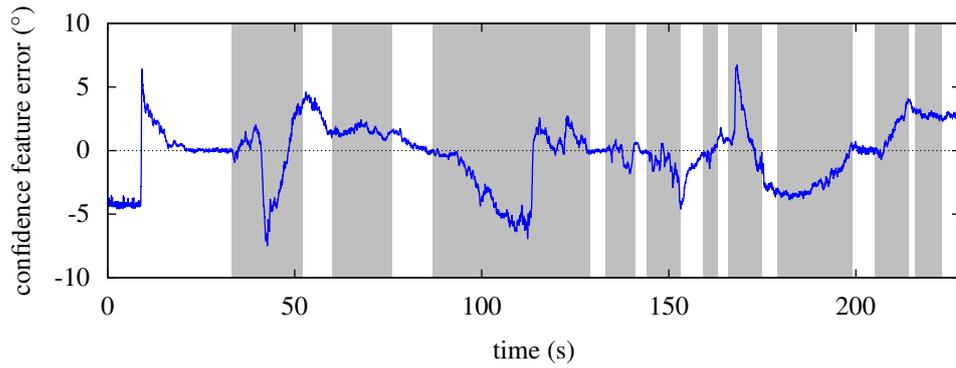
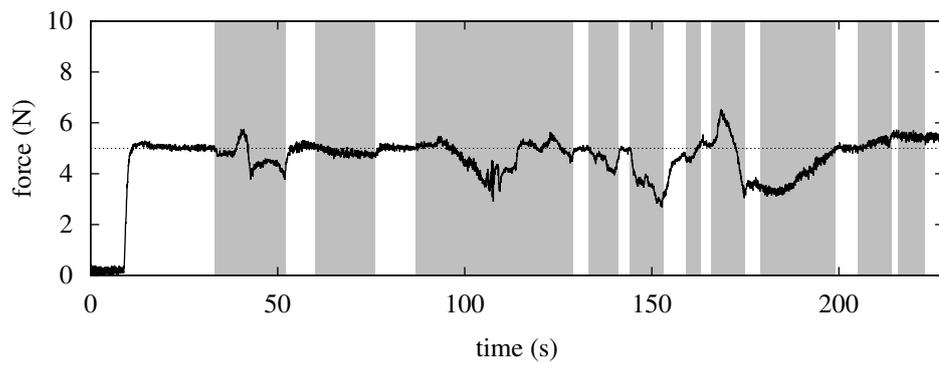


Figure 4.14 – Confidence-optimized tele-echography on an ABDFAN phantom with a 2D ultrasound probe. (a) Initial contact with the phantom. (b) Position at convergence ($t = 23$ s). (c) The operator moves the probe towards the ribs ($t = 171$ s). (d) The probe is placed between the ribs ($t = 209$ s)



(a)



(b)

Figure 4.15 – Tele-echography with the global confidence-driven control. (a) Confidence feature error θ_c in degrees. (b) Contact force s_f in Newtons. The gray parts correspond to the periods during which the teleoperation was active.

4.4.2 In vivo experiment

Experimental setup			
robot	scanner	probe	subject
KUKA iiwa	SonixRP	4DC7 (3D mode)	human

In the previous section, we have presented a teleoperation experiment on phantom. The purpose of this experiment was to validate the shared control framework, with a human operator controlling the translation motion along the x - and z -axes (along the surface). We now present a demonstration of teleoperation on a human volunteer, aimed at validating the proposed control framework in real conditions. In particular, the system now has to cope with breathing and tissue motion.

For these experiments, we used the force controller provided by the KUKA software in order to guarantee the safety of the procedure. The desired contact force was set to 2 N, which was sufficient to obtain a satisfying image quality. The confidence control law was the same as previously, with a force control gain of $\lambda_c = 0.8$.

We describe thereafter the results of three experiments: frontal abdominal scan under normal breathing, frontal abdominal scan under heavy breathing, and lateral abdominal scan under normal breathing.

Normal breathing For our first test on a human subject, the volunteer was asked to breath normally. The probe was placed vertically above the abdomen [Figure 4.16(a)]. Force control was first activated, followed by the global confidence-driven control law. After initial convergence, the feature errors were $0.58^\circ \pm 0.31^\circ$ for θ_c (in-plane orientation), and $-0.33^\circ \pm 0.02^\circ$ for ϕ_c (out-of-plane orientation) [Figure 4.16(b)]. Then, the operator moved the probe towards the right side of the patient. Consequently, the system automatically tilted the probe in order to follow the curvature of the body surface. As a result, a good image quality was maintained [Figure 4.16(c)]. Finally, the operator moved the probe back towards the center of the abdomen, and the robot tilted back in the opposite direction [Figure 4.16(d)]. During the teleoperation, the maximum confidence feature errors were 3.0° for θ_c and 2.6° for ϕ_c . The evolution of the confidence feature errors during the experiment is represented in Figure 4.17. These results show that our confidence-driven controller is able to adapt the orientation of the probe in order to maintain a good global image quality. The maximum confidence feature error (3°) was small compared to the field of view of the ultrasound probe and the rotation applied to the probe.

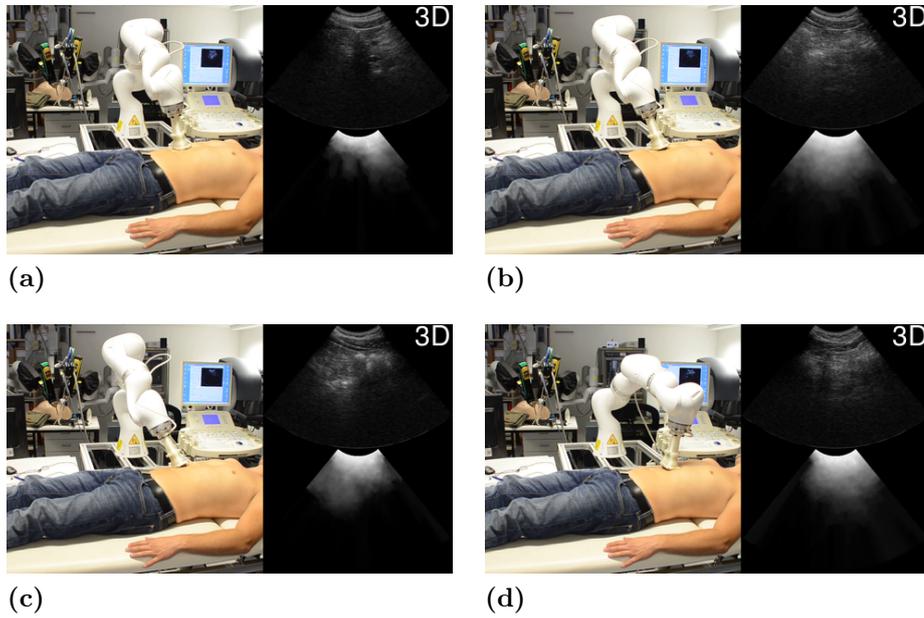


Figure 4.16 – Confidence-optimized tele-echography on a human volunteer – frontal abdominal scan – normal breathing. (a) Initial position. (b) Position at convergence ($t = 11$ s). (c) The operator moves the probe towards the right side of the patient ($t = 21$ s). (d) The operator moves the probe back to the center of the torso ($t = 51$ s)

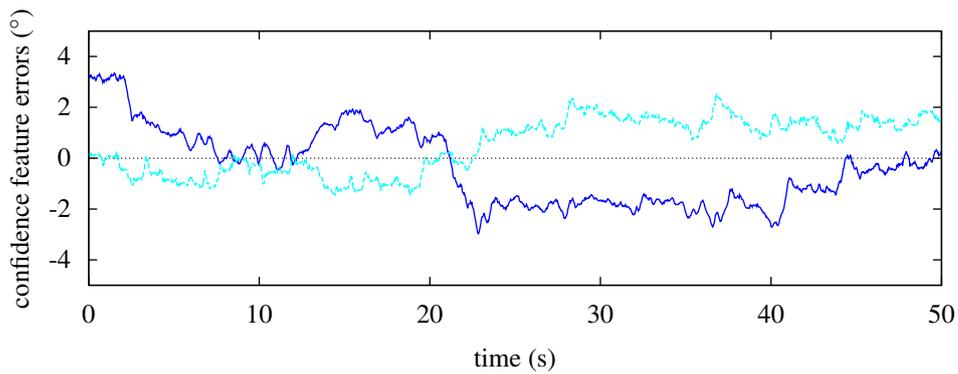


Figure 4.17 – Evolution of the confidence feature errors θ_c (—) and ϕ_c (—), in degrees, during a tele-echography on a human volunteer – frontal abdominal scan – normal breathing.

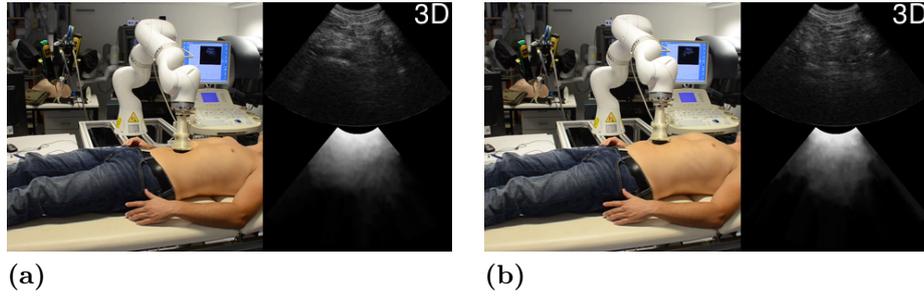


Figure 4.18 – Confidence-optimized ultrasound acquisition on a human volunteer – frontal abdominal scan – heavy breathing. (a) Exhalation. (b) Inhalation.

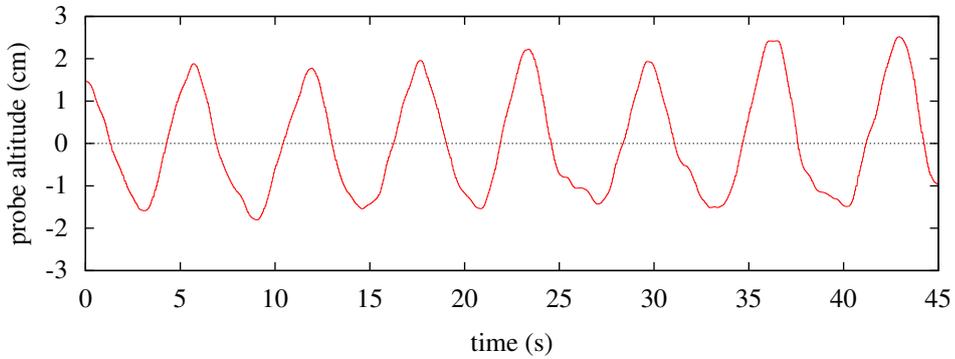


Figure 4.19 – Position of the probe along the z -axis (altitude) during heavy breathing.

Heavy breathing In a second experiment, the volunteer was asked to breath heavily. The purpose of this experiment was to assess the system’s capability to adapt the orientation of the ultrasound probe when large motions occur. The probe was placed vertically above the abdomen (Figure 4.18). Again, force control was activated first, followed by confidence control. The patient had a breathing period of approximately 6 s, with an amplitude of 4 cm along the elevation axis (Figure 4.19). Breathing also impacted the optimal scanning direction, so that the system automatically re-oriented the probe, following the breathing pattern. We report in Figure 4.20 the confidence feature errors during the experiment. Thanks to confidence control, the confidence feature errors were kept below 3.2° for θ_c , and 1.6° for ϕ_c . Therefore, we did not observe any significant increase in the confidence feature error, compared to the experiment performed under normal breathing. This shows that our system is able to compensate for large motions.

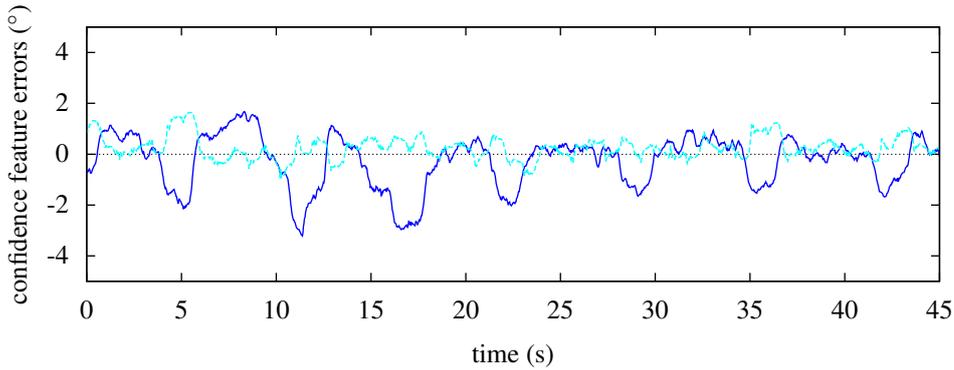


Figure 4.20 – Evolution of the confidence feature errors θ_c (—) and ϕ_c (---), in degrees, during an ultrasound acquisition on a human volunteer – frontal abdominal scan – heavy breathing.

Lateral scan Finally, we present an experiment where the probe is placed on the right side of the patient’s abdomen (Figure 4.21). In this experiment, the volunteer was asked to breath normally. After initial convergence, the confidence feature errors were $-0.45^\circ \pm 0.11^\circ$ for θ_c , and $0.13^\circ \pm 0.27^\circ$ for ϕ_c [Figure 4.21(b)]. Then, the operator moved the probe down along the side of the patient. Consequently, the system automatically tilted the probe in order to follow the curvature of the body surface, maintaining a good image quality [Figure 4.21(c)]. Finally, the operator moved the probe back up, and the robot tilted back in the opposite direction [Figure 4.21(d)].

During teleoperation, the maximum confidence feature errors were 1.4° for θ_c and 4.6° for ϕ_c . The evolution of the confidence feature errors during the experiment is represented in Figure 4.22. As in the previous experiments, the maximum confidence feature errors remained acceptable, and the image contents was clearly visible throughout the experiment.

4.5 Conclusion

In this chapter, we presented the results of a series of experiments that illustrate our framework. We tested the confidence-driven control methods presented in chapter 3 in terms of convergence and reaction to disturbances. The results show that the proposed control schemes effectively yield a simultaneous convergence of the force contact and visual features. In addition, confidence control allows the robotic system to efficiently compensate for disturbances such as shadowing and tissue motion.

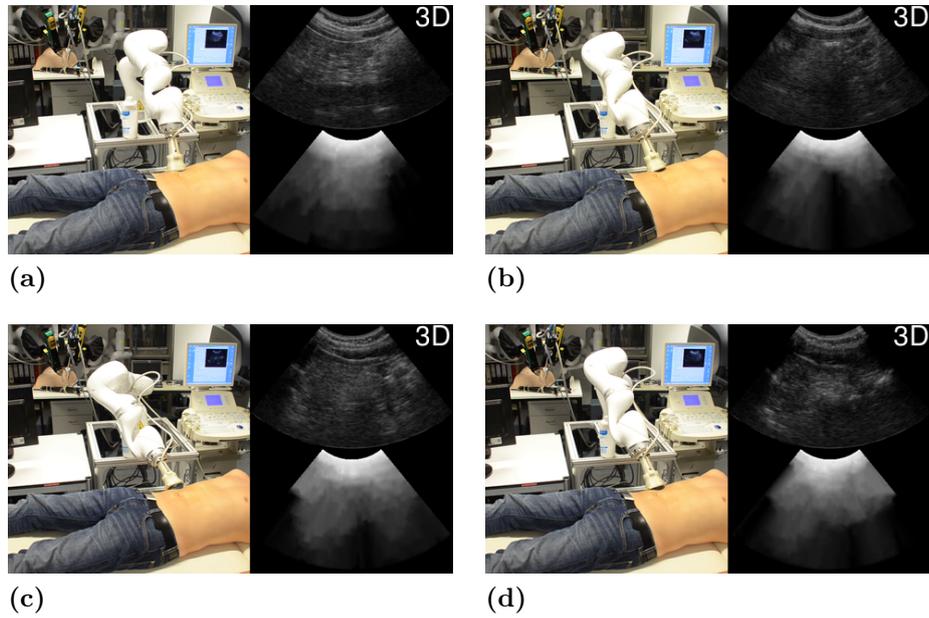


Figure 4.21 – Confidence-optimized tele-echography on a human subject – lateral abdominal scan – normal breathing. (a) Initial position. (b) Position at convergence ($t = 11$ s). (c) The operator moves the probe towards the right side of the patient ($t = 21$ s). (d) The operator moves the probe back to the center of the torso ($t = 51$ s)

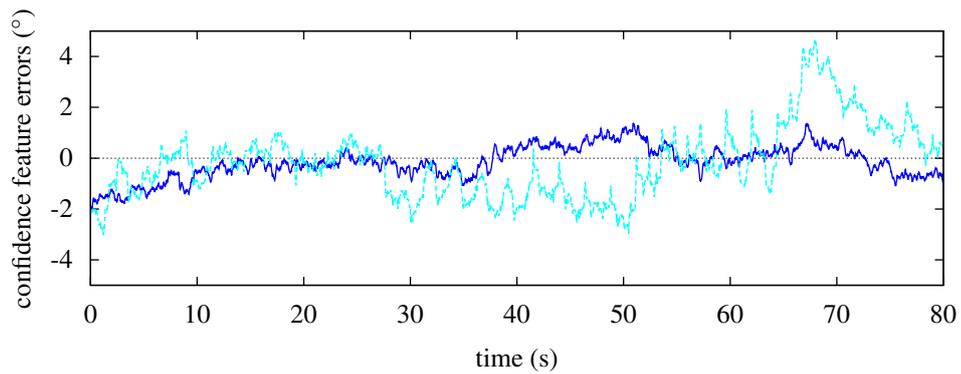


Figure 4.22 – Evolution of the confidence feature errors θ_c (—) and ϕ_c (—), in degrees, during a tele-echography on a human volunteer – lateral abdominal scan – normal breathing.

Then, we presented an illustration of our framework in the context of tele-echography, where the control was shared between the automatic controller and a teleoperator. In particular, we performed a series of experiments on a human volunteer. These experiments validate the use of our methods in realistic conditions. In that regard, they constitute an initial step towards a clinical validation of our framework.

Conclusion

In this dissertation, we discussed several issues in the field of robotized ultrasound imaging. In chapter 1, we provided an introduction to ultrasound imaging, and a review of real-time tracking methods for ultrasound images. Then, we addressed the issue of ultrasound image quality, and we presented different methods for estimating the quality of the ultrasound signal. In chapter 2, we introduced the state-of-the-art visual servoing approaches for the control of an ultrasound probe and ultrasound-guided instrument manipulation. In chapter 3, we proposed a study of the interaction between image quality and probe positioning. We developed a framework for controlling a robot-held ultrasound probe from the confidence map, in order to optimize the image quality. In chapter 4, we presented the results of experiments illustrating the proposed confidence-driven control methods. Thereafter, we draw the conclusions of this thesis on the topics of soft tissue tracking, ultrasound image quality, ultrasound-based visual servoing, and confidence-driven visual servoing. Then, we provide some perspectives for further developments of the work presented herein.

Soft tissue tracking

We started this dissertation with an introduction to ultrasound imaging, and a review of tracking algorithms. Localizing anatomical landmarks is a key prerequisite in the design of ultrasound-guided control strategies. The segmentation, detection, and tracking of structures in ultrasound image are active subjects of research, with a wide corpus of approaches originating from the communities of medical image analysis and computer vision. In our review of the literature, we focused on methods with real-time tracking capability, which is essential when considering image-guided control schemes. In this regard, template tracking and active contour approaches currently provide a good trade-off in terms of robustness, accuracy, and speed. However, template tracking algorithms are limited to rigid motions, and cannot cope with the deformations that

occur in living tissues. Deformable block matching is a promising solution to overcome this limitation, but it is more computationally expensive. Incorporating anatomy-specific shape priors could help designing a faster and more robust tracking.

Ultrasound image quality

Estimating the quality of the ultrasound signal is important for the development of robot-assisted ultrasound acquisition systems, in order to ensure an optimal image quality. A measure of ultrasound signal quality also provides a mean to assess the reliability of the image content, which is useful to subsequent image analysis steps. Among the different approaches that we presented, the concept of confidence map seems to be the most promising. It provides a real-valued measure of signal quality for each pixel in the ultrasound image. The confidence map estimation approach originally presented by [Karamalis et al., 2012], which is based on the random walks algorithm, incorporates a simplified model of ultrasound propagation properties. However, its computational complexity is relatively high, and the results present temporal discontinuities. We proposed a simpler method, based on the integration of ultrasound intensity along the scan lines, which is much faster to compute, and provides smoother results. We are convinced that alternative methods can still be found, that incorporate the advantages of both approaches.

Ultrasound-based visual servoing

In chapter 2, we presented a review of ultrasound-based visual servoing methods. We addressed two main topics: the control of an ultrasound probe, and needle insertion under ultrasound guidance. For the control of an ultrasound probe by visual servoing, the main challenge consists in the modeling of the interaction between the motion of the probe and the image content. The topic has been widely covered in the literature, and consists of geometric approaches and intensity-based approaches. Needle insertion is an other important topic in the field of robot-assisted ultrasound-guided procedures. We provided an overview of needle tracking methods for ultrasound images, which are necessary to the design of robotic assistance systems for needle insertion. Then, we presented a review of robotized needle insertion techniques, and in particular needle steering. We proposed a new 3D ultrasound-guided needle steering method based on visual servoing.

Confidence-driven visual servoing

In chapter 3, we covered the main topic of this thesis. We proposed a visual servoing framework for controlling the orientation of an ultrasound probe based on the confidence map. The issue of controlling the image quality had not been addressed in the literature on ultrasound-based visual servoing. First, we provided an analysis of the dynamics of the confidence maps, and we proposed a model of the interaction between the probe's motion and the confidence values. Then, we introduced two different confidence-driven control schemes, aimed at (i) globally optimizing the image quality, and (ii) optimizing the image quality for a specific anatomical target. In particular, we showed how to obtain a decoupled control of contact force, target positioning, and confidence. We presented experimental results that validate the proposed methods. Through experiments on phantom, we studied the convergence accuracy of the system, and its behavior in the presence of disturbances such as motion and shadowing. Then, we presented an application of our global confidence-driven control method to tele-echography, with a shared (human/computer) control configuration. In particular, we presented a validation of our framework on a human volunteer. The results confirm that the proposed method is applicable in real conditions.

Perspectives

In the following, we discuss possible further developments that could be made in the continuity of the work presented in this thesis, both in terms of theoretical research and clinical integration. We first present short-term perspectives, that could be attained in the next few years. Then, we propose long-term perspectives, which may require longer developments because of technical challenges or clinical constraints.

Short-term

Theoretical perspectives

Confidence estimation We presented a review of quality estimation methods for ultrasound images. From these methods, the concept of confidence maps, providing a pixel-wise measure of ultrasound signal quality, is a promising direction. In our implementation, we used two versions of the confidence map. The random walks confidence map, based on a simplified model of sound propagation in soft tissues, and the scan line integration method. The second one is very simple, and may not

reflect precisely the actual sound attenuation profile, but it is very time-efficient. In addition, it presents a better temporal continuity than the random walks confidence map. Surely, alternative methods could be designed, that better model the physics of ultrasound propagation, while maintaining a reasonable computational cost. In order to obtain a fine estimation of ultrasound signal quality, it would be desirable to go away from B-mode ultrasound, and to work directly on the radio-frequency data. Indeed, this would allow to be closer to the raw physical signal, and thus to avoid the deformations applied to the signal by post-processing steps. The use of the RF signal was recently investigated in [Klein and Wells III, 2015], where a measure of similarity in the distribution of RF intensities is used in the definition of the weights for the random walks algorithms.

Servoing schemes In addition to improvements in the estimation of the ultrasound signal quality, there is still room for research in the modeling of the interaction between the quality and the positioning of the probe. In particular, the methods we proposed are based on the servoing of angular confidence features. It would be interesting to have a direct control from the confidence map values. For instance, the visual servoing control law could be designed in a dense manner, similar to intensity-based visual servoing as proposed in [Nadeau and Krupa, 2013]. This could allow one to have a finer control on the quality of a precise target. On the control aspect, we proposed an hybrid control strategy, to fulfill simultaneously different control tasks. We presented an example with force control, target positioning, and confidence control. Other types of tasks could be considered as well. Indeed, the confidence-driven control method that we have introduced is generic, and it aims at being integrated in other ultrasound-based visual servoing frameworks. In our target-specific confidence-driven control method, we used the redundancy formalism as presented in [Siciliano and Slotine, 1991] to combine the different tasks. This framework was sufficient in our case, where a perfect decoupling of the tasks could be found. Other task redundancy approaches have been proposed, that allow the fusion of tasks that are more strongly coupled [Mansard and Chaumette, 2009].

Contact modeling Moreover, a precise modeling of the acoustic transmission at the surface would be desirable. As is it defined, the confidence map mainly models the sound propagation in the tissue. This model could be coupled with a model of sound transmission between the probe and the skin, to have a precise estimate of acoustic attenuation at the

surface. Such a model would allow, for instance, an automatic detection of a lack of acoustic gel. Indeed, the issue of acoustic gel was not directly addressed in this thesis. In the future, one could imagine having a gel dispenser on the robot, autonomously applying gel when needed to maintain a good acoustic coupling with the patient's skin.

Clinical perspectives

Tele-echography The contributions we presented in this dissertation ranged from the modeling of the system to its practical implementation and validation on a human volunteer. The next step will consist in the integration of the proposed method in a clinical setup. In particular, the main application we considered was robotized tele-echography, where confidence control could assist the teleoperator by maintaining automatically a good image quality. This application could be transferred easily to the clinic, since it does not involve any invasive procedure. The robot is used as a tool to hold the probe, and partial automation is used to ease the task of the physician, who remains in control of the procedure. Clinical integration will require an extensive interaction with physicians, in order to build a solution that best answers their specific needs.

Autonomous screening In addition to tele-echography, our framework could be used in different scenarios. The proposed target-specific confidence-driven control allows the system to automatically maintain a good image quality while tracking an anatomical target. This opens perspectives towards autonomous screening of a region or organ, in order to obtain a high-quality 3D scan of a structure of interest. Initial steps in this direction were recently proposed in [Virga et al., 2016] for screening the aorta. Ideally, a preoperative scan from the patient, or an anatomical atlas, could be used to select a desired region of interest to be scanned autonomously by the robotized ultrasound system. In this scenario, confidence-driven control would be useful to ensure an optimal image quality.

Long-term

Clinical integration The first long-term perspective of the work presented in this thesis is a full clinical integration of the proposed methods. To this end, all aspects of the solution should be carefully considered, from the design of the robot to the user interface. There again, continuous interaction with the clinical staff is primordial. In addition, attention

must be paid to all the requirements for a smooth integration of the system to the operating room, such as described in [Navab et al., 2016]. An important challenge in the translation of a new technology into clinical practice is the acceptance of the system by the medical staff. This necessitates, on one hand, a careful design of the system and the human/machine interface. On the other hand, it also requires the development of education and training solutions to form the medical staff to the use of this technology. In this regard, immersive simulation environments could help prototyping and testing new technologies, and allow a faster translation of experimental systems into clinical applications [Wucherer et al., 2013; Promayon et al., 2013].

Intraoperative imaging In terms of clinical applications, we have mainly discussed the use of our method for diagnostic purposes, in the context of tele-echography. Eventually, our methods could also be of use in robot-assisted surgery, to improve needle placement during biopsy or drug delivery. Indeed, controlling the image quality during ultrasound-guided needle insertion could be a key to improve the reliability and precision of such procedures. We presented a method for controlling simultaneously the position of a target within the image and the quality of the image, using the redundancy algorithm. With this approach, an additional degree of freedom (the rotation around the depth direction) remains available. It could be used, for instance, to automatically align the needle shaft with the longitudinal axis of the probe, as was proposed in [Chatelain et al., 2013]. This way, one would ensure that the needle shaft is visualized in a direction where the image resolution is the highest, thus improving its visibility. While such a method could be developed and tested experimentally in the next years, we classify it in the long-term perspectives, because the translation of intraoperative solutions to the clinic is more challenging than for diagnostic imaging. In addition, it would be desirable to fuse multiple imaging modalities for the guidance of interventional procedures, in order to leverage the advantages of the different modalities. For instance, fusing ultrasound with CT or MRI can help performing a precise needle insertion towards a target that is not directly detectable in the ultrasound images [Crocetti et al., 2008; Mauri et al., 2014]. In the context of MRI or CT imaging, it is also important to design robotic systems that are compatible with the imaging device [Shahriari et al., 2015; Hungr et al., 2016].

Therapeutic ultrasound In the introduction, we have mentioned the use of high-intensity focused ultrasound for thermal ablation of tissues.

Ultrasound-based visual servoing has recently been proposed to improve the accuracy of HIFU by performing motion compensation [Seo et al., 2011; Chanel et al., 2015]. In this context, the confidence-driven control framework we proposed could be use to ensure the visibility of the target.

Clinical workflow optimization Finally, a more general perspective is the closed-loop integration of quality optimization into the clinical workflow. In this thesis, we have focused on the optimization of image quality. In order to obtain optimal results for a clinical task, it would be necessary to consider the procedure in its entirety, and to optimize all components with respect to the expected outcome. For instance, if the aim is to obtain a high-quality 3D ultrasound image of an organ, the acquisition procedure should be designed so as to optimize the quality of the final reconstructed 3D image, and not solely of each individual image. This could imply predicting the quality of the reconstruction from the data that has been acquired, and predicting how to acquire new data to improve the reconstructed image. Similarly, for target tracking, it should be possible to use the confidence map to predict the quality of the tracking, and to take actions to guarantee the success of the tracking. In addition, such an indicator on the confidence of the measures extracted from the image could provide useful information to the medical staff. In this regard, machine learning algorithms could help predicting the long-term outcome of the procedures, and optimizing each step so as to get the best possible outcome.

List of Publications

Published

- **Chatelain, P.**, Krupa, A., and Navab, N. (2016). Confidence-driven control of an ultrasound probe: target-specific acoustic window optimization. In *IEEE Int. Conf. on Robotics and Automation, ICRA'16*, pages 3441–3446, Stockholm, Sweden.
- Peter, L., Pauly, O., **Chatelain, P.**, Mateus, D., and Navab, N. (2015). Scale-Adaptive Forest Training via an Efficient Feature Sampling Scheme. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015*, part I, Lecture Notes in Computer Science 9349, pages 637–644, Munich, Germany.
- **Chatelain, P.**, Krupa, A., and Navab, N. (2015). Optimization of ultrasound image quality via visual servoing. In *IEEE Int. Conf. on Robotics and Automation, ICRA'15*, pages 5997–6002, Seattle, WA.
- **Chatelain, P.**, Krupa, A., and Navab, N. (2015). 3D ultrasound-guided robotic steering of a flexible needle via visual servoing. In *IEEE Int. Conf. on Robotics and Automation, ICRA'15*, pages 2250–2255, Seattle, WA.
- Peter, L., Mateus, D., **Chatelain, P.**, Schworm, N., Stangl, S., Multhoff, G., Navab, N. (2014). Leveraging Random Forests for Interactive Exploration of Large Histological Images. In *Medical Image Computing and Computer Assisted Intervention — MICCAI 2014*, part I, Lecture Notes in Computer Science 8673, pages 1–8, Boston, MA.
- **Chatelain, P.**, Pauly, O., Peter, L., Ahmadi S.-A., Plate A., Boetzel K., Navab, N. (2013). Learning from Multiple Experts with Random Forests: Application to the Segmentation of the Midbrain

in 3D Ultrasound. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2013*, part II, Lecture Notes in Computer Science 8150, pages 230–237, Nagoya, Japan.

- **Chatelain, P.**, Krupa, A., Marchal M. (2013). Real-time needle detection and tracking using a visually servoed 3D ultrasound probe. In *IEEE Int. Conf. on Robotics and Automation, ICRA'13*, pages 1668–1673, Karlsruhe, Germany.

In press

- Peter, L., Mateus, D., **Chatelain, P.**, Declara, D., Schworm, N., Stangle, S., Multhoff, G., and Navab, N. Assisting the Examination of Large Histopathological Slides with Adaptive Forests. *Medical Image Analysis*.

Under review

- **Chatelain, P.**, Krupa, A., and Navab, N. Confidence-driven control of an ultrasound probe. *IEEE Transactions on Robotics*.

Abstracts of Additional Contributions

We provide the abstracts of those publications which were not discussed in the dissertation.

Real-Time Needle Detection and Tracking Using a Visually Servoed 3D Ultrasound Probe

P. Chatelain, A. Krupa, M. Marchal

In this paper, we present a method to localize and track manually inserted needles in real-time using a three-dimensional ultrasound probe mounted on a robotized arm. The system tracks the needle in real-time using image processing. We first propose a new algorithm capable of robustly detecting a needle from the moment it is inserted, without any a priori information on the insertion direction. By combining the random sample consensus (RANSAC) algorithm with Kalman filtering in closed loop, we achieve robust real-time tracking of the needle. In addition, we propose a control scheme to automatically guide the ultrasonic probe in order to keep the needle within the field of view, while aligning its axis with the ultrasound beam. This method will ease the insertion of the needle by the operator, and allow the development of autonomous needle insertion by medical robots.

Learning from Multiple Experts with Random Forests: Application to the Segmentation of the Midbrain in 3D Ultrasound

P. Chatelain, O. Pauly, L. Peter, A. Ahmadi, A. Plate, K. Bötzel, N. Navab

In the field of computer aided medical image analysis, it is often difficult to obtain reliable ground truth for evaluating algorithms or supervising statistical learning procedures. In this paper we present a new method for training a classification forest from images labeled by variably performing experts, while simultaneously evaluating the performance of each expert. Our approach builds upon state-of-the-art randomized classification forest techniques for medical image segmentation and recent methods for the fusion of multiple expert decisions. By incorporating the performance evaluation within the training phase, we obtain a novel forest framework for learning from conflicting expert decisions, accounting for both inter- and intra-expert variability. We demonstrate on a synthetic example that our method allows to retrieve the correct segmentation among other incorrectly labeled images, and we present an application to the automatic segmentation of the midbrain in 3D transcranial ultrasound images.

Leveraging Random Forests for Interactive Exploration of Large Histological Images

L. Peter, D. Mateus, P. Chatelain, N. Schworm, S. Stangl, G. Multhoff, N. Navab

The large size of histological images combined with their very challenging appearance are two main difficulties which considerably complicate their analysis. In this paper, we introduce an interactive strategy leveraging the output of a supervised random forest classifier to guide a user through such large visual data. Starting from a forest-based pixelwise estimate, subregions of the images at hand are automatically ranked and sequentially displayed according to their expected interest. After each region suggestion, the user selects among several options a rough estimate of the true amount of foreground pixels in this region. From these one-click inputs, the region scoring function is updated in real time using an online gradient descent procedure, which corrects on-the-fly the shortcomings of the initial model and adapts future suggestions accordingly. Experimental validation is conducted for extramedullary hematopoiesis localization

and demonstrates both the practical feasibility of the procedure as well as the benefit of the online adaptation strategy.

Scale-Adaptive Forest Training via an Efficient Feature Sampling Scheme

L. Peter, O. Pauly, P. Chatelain, D. Mateus, N. Navab

In the context of forest-based segmentation of medical data, modeling the visual appearance around a voxel requires the choice of the scale at which contextual information is extracted, which is of crucial importance for the final segmentation performance. Building on Haar-like visual features, we introduce a simple yet effective modification of the forest training which automatically infers the most informative scale at each stage of the procedure. Instead of the standard uniform sampling during node split optimization, our approach draws candidate features sequentially in a fine-to-coarse fashion. While being very easy to implement, this alternative is free of additional parameters, has the same computational cost as a standard training and shows consistent improvements on three medical segmentation datasets with very different properties.

Bibliography

- Abayazid, M., Moreira, P., Shahriari, N., Zompas, A., and Misra, S. (2016). Three-Dimensional Needle Steering Using Automated Breast Volume Scanner (ABVS). *Journal of Medical Robotics Research*, 01(01):1640005.
- Abayazid, M., Vrooijink, G. J., Patil, S., Alterovitz, R., and Misra, S. (2014). Experimental Evaluation of Ultrasound-Guided 3D Needle Steering in Biological Tissue. *International Journal of Computer Assisted Radiology and Surgery*, 9(6):931–939.
- Abolmaesumi, P., Salcudean, S. E., Zhu, W.-H., Sirouspour, M. R., and DiMaio, S. P. (2002). Image-guided control of a robot for medical ultrasound. *IEEE Transactions on Robotics and Automation*, 18(1):11–23.
- Aboofazeli, M., Abolmaesumi, P., Mousavi, P., and Fichtinger, G. (2009). A new scheme for curved needle segmentation in three-dimensional ultrasound images. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1067–1070.
- Adebar, T. K., Fletcher, A. E., and Okamura, A. M. (2014). 3-D Ultrasound-Guided Robotic Needle Steering in Biological Tissue. *IEEE Transactions on Biomedical Engineering*, 61(12):2899–2910.
- Alterovitz, R., Goldberg, K., and Okamura, A. (2005). Planning for Steerable Bevel-tip Needle Insertion Through 2D Soft Tissue with Obstacles. In *IEEE International Conference on Robotics and Automation*, pages 1640–1645.
- Arbeille, P., Poisson, G., Vieyres, P., Ayoub, J., Porcher, M., and Boulay, J. L. (2003). Echographic examination in isolated sites controlled from an expert center using a 2-D echograph guided by a teleoperated robotic arm. *Ultrasound in Medicine & Biology*, 29(7):993–1000.
- Arbeille, P., Ruiz, J., Herve, P., Chevillot, M., Poisson, G., and Perrotin, F. (2005). Fetal tele-echography using a robotic arm and a satellite link. *Ultrasound in Obstetrics and Gynecology*, 26(3):221–226.
- Bakthavatchalam, M., Chaumette, F., and Marchand, E. (2013). Photometric moments: New promising candidates for visual servoing. In *2013 IEEE*

- International Conference on Robotics and Automation (ICRA)*, pages 5241–5246.
- Banovac, F., Glossop, N., Lindisch, D., Tanaka, D., Levy, E., and Cleary, K. (2002). Liver Tumor Biopsy in a Respiring Phantom with the Assistance of a Novel Electromagnetic Navigation Device. In Dohi, T. and Kikinis, R., editors, *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002*, number 2488 in Lecture Notes in Computer Science, pages 200–207. Springer Berlin Heidelberg.
- Barva, M., Uherčík, M., Mari, J.-M., Kybic, J., Duhamel, J.-R., Liebgott, H., Hlavac, V., and Cachard, C. (2008). Parallel integral projection transform for straight electrode localization in 3-D ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 55(7):1559 – 1569.
- Basarab, A., Liebgott, H., Morestin, F., Lyshchik, A., Higashi, T., Asato, R., and Delachartre, P. (2008). A method for vector displacement estimation with ultrasound imaging and its application for thyroid nodular disease. *Medical Image Analysis*, 12(3):259–274.
- Beigi, P., Rohling, R., Salcudean, S. E., and Ng, G. C. (2016). Spectral analysis of the tremor motion for needle detection in curvilinear ultrasound via spatiotemporal linear sampling. *International Journal of Computer Assisted Radiology and Surgery*, 11(6):1183–1192.
- Bernardes, M. C., Adorno, B. V., Borges, G. A., and Poignet, P. (2014). 3D Robust Online Motion Planning for Steerable Needles in Dynamic Workspaces Using Duty-Cycled Rotation. *Journal of Control, Automation and Electrical Systems*, 25(2):216–227.
- Bernardes, M. C., Adorno, B. V., Poignet, P., and Borges, G. A. (2013). Robot-assisted automatic insertion of steerable needles with closed-loop imaging feedback and intraoperative trajectory replanning. *Mechatronics*, 23(6):630–645.
- Berton, F., Cheriet, F., Miron, M.-C., and Laporte, C. (2016). Segmentation of the spinous process and its acoustic shadow in vertebral ultrasound images. *Computers in Biology and Medicine*, 72:201–211.
- Bierman, W. (1954). Ultrasound in the treatment of scars. *Archives of Physical Medicine and Rehabilitation*, 35(4):209–214.
- Biot, J.-B. (1802). Sur la théorie du son. *Journal de physique*, 55:173–182.

- Boctor, E. M., Choti, M. A., Burdette, E. C., and Webster III, R. J. (2008). Three-dimensional ultrasound-guided robotic needle placement: an experimental evaluation. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 4(2):180–191.
- Bohs, L. N. and Trahey, G. E. (1991). A novel method for angle independent ultrasonic imaging of blood flow and tissue motion. *IEEE Transactions on Biomedical Engineering*, 38(3):280–286.
- Boman, K., Olofsson, M., Forsberg, J., and Boström, S.-Å. (2009). Remote-Controlled Robotic Arm for Real-Time Echocardiography: The Diagnostic Future for Patients in Rural Areas? *Telemedicine and e-Health*, 15(2):142–147.
- Boudet, S., Gariepy, J., and Mansour, S. (1997). An integrated robotics and medical control device to quantify atheromatous plaques: experiments on the arteries of a patient. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 1533–1538. IEEE.
- Boukerroui, D., Noble, J. A., and Brady, M. (2003). Velocity Estimation in Ultrasound Images: A Block Matching Approach. In Taylor, C. and Noble, J. A., editors, *Information Processing in Medical Imaging*, number 2732 in Lecture Notes in Computer Science, pages 586–598. Springer Berlin Heidelberg.
- Brody, H. (2015). Prostate cancer. *Nature*, 528(7582):S117–S117.
- Cachard, C., Barva, M., Kybic, J., Hlavac, V. C., and Liebgott, H. (2006). P3a-3 Comparison of Methods for Tool Localization in Biological Tissue from 3D Ultrasound Data. In *2006 IEEE Ultrasonics Symposium*, pages 1983–1986.
- Carneiro, G. and Nascimento, J. C. (2013). Combining Multiple Dynamic Models and Deep Learning Architectures for Tracking the Left Ventricle Endocardium in Ultrasound Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2592–2607.
- Chalana, V., Linker, D. T., Haynor, D. R., and Kim, Y. (1996). A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 15(3):290–298.
- Chan, C., Lam, F., and Rohling, R. (2005). A needle tracking device for ultrasound guided percutaneous procedures. *Ultrasound in Medicine & Biology*, 31(11):1469–1483.
- Chan, V. and Perlas, A. (2011). Basics of Ultrasound Imaging. In Narouze, S. N., editor, *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, pages 13–19. Springer New York.

- Chanel, L. A., Nageotte, F., Vappou, J., Luo, J., Cuvillon, L., and de Mathelin, M. (2015). Robotized High Intensity Focused Ultrasound (HIFU) system for treatment of mobile organs using motion tracking by ultrasound imaging: An in vitro study. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2571–2575.
- Chapman, G. A., Johnson, D., and Bodenham, A. R. (2006). Visualisation of needle position using ultrasonography. *Anaesthesia*, 61(2):148–158.
- Chatelain, P., Krupa, A., and Marchal, M. (2013). Real-time needle detection and tracking using a visually servoed 3D ultrasound probe. In *2013 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1676–1681.
- Chatelain, P., Krupa, A., and Navab, N. (2015a). 3D ultrasound-guided robotic steering of a flexible needle via visual servoing. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2250–2255.
- Chatelain, P., Krupa, A., and Navab, N. (2015b). Optimization of ultrasound image quality via visual servoing. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5997–6002.
- Chatelain, P., Krupa, A., and Navab, N. (2016). Confidence-Driven Control of an Ultrasound Probe: Target-Specific Acoustic Window Optimization. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 3441-3446.
- Chaumette, F. and Hutchinson, S. (2006). Visual servo control, Part I: Basic approaches. *IEEE Robotics and Automation Magazine*, 13(4):82–90.
- Chaumette, F. and Hutchinson, S. (2007). Visual servo control, Part II: Advanced approaches. *IEEE Robotics and Automation Magazine*, 14(1):109–118.
- Chevrie, J., Krupa, A., and Babel, M. (2016a). Needle steering fusing direct base manipulation and tip-based control. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4450–4455.
- Chevrie, J., Krupa, A., and Babel, M. (2016b). Online prediction of needle shape deformation in moving soft tissues from visual feedback. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2375–2380.
- Chilowsky, C. and Langevin, P. (1916). Procédés et appareils pour la production de signaux sous-marins dirigés et pour la localisation à distance d’obstacles sous-marins. French patent, 502913.

- Cohen, B. and Dinstein, I. (2002). New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition*, 35(2):455–463.
- Coleman, J., Goettsch, A., Savchenko, A., Kollmann, H., Wang, K., Klement, E., and Bono, P. (1996). TeleInViVoTM: Towards collaborative volume visualization environments. *Computers & Graphics*, 20(6):801–811.
- Comaniciu, D., Zhou, X. S., and Krishnan, S. (2004). Robust real-time myocardial border tracking for echocardiography: an information fusion approach. *IEEE Transactions on Medical Imaging*, 23(7):849–860.
- Cootes, T. F. and Taylor, C. J. (1992). Active Shape Models — ‘Smart Snakes’. In Hogg, D. and Boyle, R., editors, *BMVC92*, pages 266–275. Springer London.
- Courreges, F., Vieyres, P., and Istepanian, R. S. H. (2004). Advances in robotic tele-echography services - the OtelO system. In *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004. IEMBS '04*, volume 2, pages 5371–5374.
- Crocetti, L., Lencioni, R., DeBeni, S., See, T. C., Pina, C. D., and Bartolozzi, C. (2008). Targeting Liver Lesions for Radiofrequency Ablation: An Experimental Feasibility Study Using a CT-US Fusion Imaging System. *Investigative Radiology*, 43(1):33–39.
- Curie, J. and Curie, P. (1880). Développement par compression de l’électricité polaire dans les cristaux hémihédres à faces inclinées. *Bulletin de la Société minéralogique de France*, 3(4):90–93.
- Curie, J. and Curie, P. (1881). Contractions et dilatations produites par des tensions électriques dans les cristaux hémihédres à faces inclinées. *Comptes rendus hebdomadaires des séances de l’Académie des sciences*, 93:1137–1140.
- Davies, B. L., Chauhan, S., and Lowe, M. J. S. (1998). A robotic approach to HIFU based neurosurgery. In Wells, W. M., Colchester, A., and Delp, S., editors, *Medical Image Computing and Computer-Assisted Intervention — MICCAI’98*, number 1496 in Lecture Notes in Computer Science, pages 386–396. Springer Berlin Heidelberg.
- Davis, M. (2007). Audio and Electroacoustics. In Rossing, T. D., editor, *Springer Handbook of Acoustics*, pages 743–781. Springer New York.
- De Cunha, D., Gravez, P., Leroy, C., Maillard, E., Jouan, J., Varley, P., Jones, M., Halliwell, M., Hawkes, D., Wells, P., and others (1998). The MIDSTEP system for ultrasound guided remote telesurgery. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1998*, volume 3, pages 1266–1269. IEEE.

- De Luca, V., Székely, G., and Tanner, C. (2015). Estimation of large-scale organ motion in B-mode ultrasound image sequences: A survey. *Ultrasound in Medicine and Biology*, 41(12):3044 – 3062.
- Delgorge, C., Courrèges, F., Bassit, L. A., Novales, C., Rosenberger, C., Smith-Guerin, N., Brù, C., Gilabert, R., Vannoni, M., Poisson, G., and others (2005). A tele-operated mobile ultrasound scanner using a light-weight robot. *IEEE Transactions on Information Technology in Biomedicine*, 9(1):50–58.
- DiMaio, S. and Salcudean, S. (2005). Needle steering and motion planning in soft tissues. *IEEE Transactions on Biomedical Engineering*, 52(6):965–974.
- Dorileo, É., Albakri, A., Zemiti, N., and Poignet, P. (2015). Simplified adaptive path planning for percutaneous needle insertions. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1782–1788.
- Drukker, K., Giger, M. L., and Mendelson, E. B. (2003). Computerized analysis of shadowing on breast ultrasound for improved lesion detection. *Medical Physics*, 30(7):1833–1842.
- Duflot, L. A., Krupa, A., Tamadazte, B., and Andreff, N. (2016). Towards Ultrasound-based visual servoing using shearlet coefficients. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3420–3425.
- Duindam, V., Xu, J., Alterovitz, R., Sastry, S., and Goldberg, K. (2010). Three-dimensional Motion Planning Algorithms for Steerable Needles Using Inverse Kinematics. *The International Journal of Robotics Research*, 29(7):789–800.
- Dussik, K. T. (1942). Über die Möglichkeit, hochfrequente mechanische Schwingungen als diagnostisches Hilfsmittel zu verwenden. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 174(1):153–168.
- Easley, G., Labate, D., and Lim, W.-Q. (2008). Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis*, 25(1):25–46.
- El-Zehiry, N., Yan, M., Good, S., Fang, T., Zhou, S. K., and Grady, L. (2013). Learning the Manifold of Quality Ultrasound Acquisition. In Mori, K., Sakuma, I., Sato, Y., Barillot, C., and Navab, N., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, number 8149 in Lecture Notes in Computer Science, pages 122–130. Springer Berlin Heidelberg.

- Espiau, B., Chaumette, F., and Rives, P. (1992). A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3):313–326.
- Fichtinger, G., Fiene, J. P., Kennedy, C. W., Kronreif, G., Iordachita, I., Song, D. Y., Burdette, E. C., and Kazanzides, P. (2008). Robotic assistance for ultrasound-guided prostate brachytherapy. *Medical Image Analysis*, 12(5):535–545.
- Foroughi, P., Abolmaesumi, P., and Hashtrudi-Zaad, K. (2006). Towards real-time registration of 4D ultrasound images. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 404–407.
- Frangi, A. F., Niessen, W. J., and Viergever, M. A. (2001). Three-dimensional modeling for functional analysis of cardiac images, a review. *IEEE Transactions on Medical Imaging*, 20(1):2–5.
- Fronheiser, M. P., Idriss, S. F., Wolf, P. D., and Smith, S. W. (2008). Vibrating interventional device detection using real-time 3-D color Doppler. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 55(6):1355–1362.
- Frush, D. P. (2004). Review of radiation issues for computed tomography. *Seminars in Ultrasound, CT and MRI*, 25(1):17–24.
- Gilbertson, M. W. and Anthony, B. W. (2015). Force and Position Control System for Freehand Ultrasound. *IEEE Transactions on Robotics*, 31(4):835–849.
- Gill, S., Abolmaesumi, P., Fichtinger, G., Boisvert, J., Pichora, D., Borshneck, D., and Mousavi, P. (2012). Biomechanically constrained groupwise ultrasound to CT registration of the lumbar spine. *Medical Image Analysis*, 16(3):662–674.
- Glozman, D. and Shoham, M. (2007). Image-Guided Robotic Flexible Needle Steering. *IEEE Transactions on Robotics*, 23(3):459–467.
- Golemati, S., Sassano, A., Lever, M. J., Bharath, A. A., Dhanjil, S., and Nicolaides, A. N. (2003). Carotid artery wall motion estimated from b-mode ultrasound using region tracking and block matching. *Ultrasound in Medicine & Biology*, 29(3):387–399.
- Golemati, S., Stoitsis, J. S., Gastouniotti, A., Dimopoulos, A. C., Koropouli, V., and Nikita, K. S. (2012). Comparison of Block Matching and Differential Methods for Motion Analysis of the Carotid Artery Wall From Ultrasound Images. *IEEE Transactions on Information Technology in Biomedicine*, 16(5):852–858.

- Gordon, N. J., Salmond, D. J., and Smith, A. F. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140, pages 107–113. IET.
- Grady, L. (2006). Random Walks for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783.
- Güttner, W., Fiedler, G., and Pätzold, J. (1952). Über Ultraschallabbildungen am menschlichen Schädel. *Acta Acustica united with Acustica*, 2(4):148–156.
- Hager, G. D. and Belhumeur, P. N. (1998). Efficient region tracking with parametric models of geometry and illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(10):1025–1039.
- Hamper, U. M., Savader, B. L., and Sheth, S. (1991). Improved needle-tip visualization by color Doppler sonography. *American Journal of Roentgenology*, 156(2):401–402.
- Han, M., Chang, D., Kim, C., Lee, B. J., Zuo, Y., Kim, H.-J., Petrisor, D., Trock, B., Partin, A. W., Rodriguez, R., Carter, H. B., Allaf, M., Kim, J., and Stoianovici, D. (2012). Geometric Evaluation of Systematic Transrectal Ultrasound Guided Prostate Biopsy. *The Journal of Urology*, 188(6):2404–2409.
- Hellier, P., Coupé, P., Morandi, X., and Collins, D. L. (2010). An automatic geometrical and statistical method to detect acoustic shadows in intraoperative ultrasound brain images. *Medical Image Analysis*, 14(2):195–204.
- Herlin, I. L. and Ayache, N. (1992). Features extraction and analysis methods for sequences of ultrasound Images. *Image and Vision Computing*, 10(10):673–682.
- Hong, J., Dohi, T., Hashizume, M., Konishi, K., and Hata, N. (2004). An ultrasound-driven needle-insertion robot for percutaneous cholecystostomy. *Physics in Medicine and Biology*, 49(3):441–455.
- Hong, J.-S., Dohi, T., Hasizume, M., Konishi, K., and Hata, N. (2002). A Motion Adaptable Needle Placement Instrument Based on Tumor Specific Ultrasonic Image Segmentation. In Dohi, T. and Kikinis, R., editors, *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002*, number 2488 in Lecture Notes in Computer Science, pages 122–129. Springer Berlin Heidelberg.
- Hough, P. V. C. (1962). Method and means for recognizing complex patterns. US Patent number 3069654.

- Hounsfield, G. N. (1973). Computerized transverse axial scanning (tomography): Part 1. Description of system. *The British Journal of Radiology*, 46(552):1016–1022.
- Howard, M. H., Nelson, R. C., Paulson, E. K., Kliewer, M. A., and Sheafor, D. H. (2001). An Electronic Device for Needle Placement during Sonographically Guided Percutaneous Intervention. *Radiology*, 218(3):905–911.
- Howry, D. H., Holmes, J. H., Cushman, C. R., and Posakony, G. J. (1955). Ultrasonic visualization of living organs and tissues; with observations on some disease processes. *Geriatrics*, 10(3):123–128.
- Hung, A. J., Abreu, A. L. D. C., Shoji, S., Goh, A. C., Berger, A. K., Desai, M. M., Aron, M., Gill, I. S., and Ukimura, O. (2012). Robotic Transrectal Ultrasonography During Robot-Assisted Radical Prostatectomy. *European Urology*, 62(2):341–348.
- Hungr, N., Baumann, M., Long, J. A., and Troccaz, J. (2012). A 3-D Ultrasound Robotic Prostate Brachytherapy System With Prostate Motion Tracking. *IEEE Transactions on Robotics*, 28(6):1382–1397.
- Hungr, N., Bricault, I., Cinquin, P., and Fouard, C. (2016). Design and Validation of a CT- and MRI-Guided Robot for Percutaneous Needle Procedures. *IEEE Transactions on Robotics*, 32(4):973–987.
- Jacob, G., Noble, J. A., Behrenbruch, C., Kelion, A. D., and Banning, A. P. (2002). A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography. *IEEE Transactions on Medical Imaging*, 21(3):226–238.
- Jacob, G., Noble, J. A., Mulet-Parada, M., and Blake, A. (1999). Evaluating a robust contour tracker on echocardiographic sequences. *Medical Image Analysis*, 3(1):63–75.
- Karamalis, A., Wein, W., Klein, T., and Navab, N. (2012). Ultrasound confidence maps using random walks. *Medical Image Analysis*, 16(6):1101–1112.
- Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331.
- Katouzian, A., Karamalis, A., Sheet, D., Konofagou, E., Baseri, B., Carlier, S. G., Eslami, A., König, A., Navab, N., and Laine, A. F. (2012). Iterative Self-Organizing Atherosclerotic Tissue Labeling in Intravascular Ultrasound Images and Comparison With Virtual Histology. *IEEE Transactions on Biomedical Engineering*, 59(11):3039–3049.

- Kaye, D. R., Stoianovici, D., and Han, M. (2014). Robotic Ultrasound and Needle Guidance for Prostate Cancer Management: Review of the Contemporary Literature. *Current opinion in urology*, 24(1):75–80.
- Khallaghi, S., Leung, C. G. M., Hastrudi-Zaad, K., Foroughi, P., Nguan, C., and Abolmaesumi, P. (2012). Experimental validation of an intrasubject elastic registration algorithm for dynamic-3D ultrasound images. *Medical Physics*, 39(9):5488–5497.
- Klein, T. and Wells III, W. M. (2015). RF Ultrasound Distribution-Based Confidence Maps. In Navab, N., Hornegger, J., Wells, W., and Frangi, A., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, volume 9350 of *Lecture Notes in Computer Science*, pages 595–602. Springer International Publishing.
- Knoll, C., Alcañiz, M., Grau, V., Monserrat, C., and Carmen Juan, M. (1999). Outlining of the prostate using snakes with shape restrictions based on the wavelet transform (Doctoral Thesis: Dissertation). *Pattern Recognition*, 32(10):1767–1781.
- Ko, S. Y. and Rodriguez y Baena, F. (2013). Toward a Miniaturized Needle Steering System With Path Planning for Obstacle Avoidance. *IEEE Transactions on Biomedical Engineering*, 60(4):910–917.
- Kobayashi, Y., Suzuki, M., Kato, A., Hatano, M., Konishi, K., Hashizume, M., and Fujie, M. G. (2012). Enhanced Targeting in Breast Tissue Using a Robotic Tissue Preloading-Based Needle Insertion System. *IEEE Transactions on Robotics*, 28(3):710–722.
- Kojcev, R., Fuerst, B., Zettinig, O., Fotouhi, J., Lee, S. C., Frisch, B., Taylor, R., Sinibaldi, E., and Navab, N. (2016). Dual-robot ultrasound-guided needle placement: closing the planning-imaging-action loop. *International Journal of Computer Assisted Radiology and Surgery*, 11(6):1173–1181.
- Kontaxakis, G., Walter, S., and Sakas, G. (2000). EU-TeleInViVo: an integrated portable telemedicine workstation featuring acquisition, processing and transmission over low-bandwidth lines of 3D ultrasound volume images. In *2000 IEEE EMBS International Conference on Information Technology Applications in Biomedicine, 2000. Proceedings*, pages 158–163.
- Krombach, G. A., Mahnken, A., Tacke, J., Staatz, G., Haller, S., Nolte-Ernsting, C. C. A., Meyer, J., Haage, P., and Günther, R. W. (2001). US-guided Nephrostomy with the Aid of a Magnetic Field-based Navigation Device in the Porcine Pelvicaliceal System. *Journal of Vascular and Interventional Radiology*, 12(5):623–628.

- Krupa, A. (2014). 3D steering of a flexible needle by visual servoing. In Golland, P., Hata, N., Barillot, C., Hornegger, J., and Howe, R., editors, *Medical Image Computing and Computer-Assisted Interventions – MICCAI 2014*, volume 8673 of *Lecture Notes in Computer Science*, pages 480–487. Springer International Publishing.
- Krupa, A. and Chaumette, F. (2006). Guidance of an ultrasound probe by visual servoing. *Advanced Robotics*, 20(11):1203–1218.
- Krupa, A., Fichtinger, G., and Hager, G. D. (2007). Full Motion Tracking in Ultrasound Using Image Speckle Information and Visual Servoing. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 2458–2464.
- Krupa, A., Fichtinger, G., and Hager, G. D. (2009). Real-time Motion Stabilization with B-mode Ultrasound Using Image Speckle Information and Visual Servoing. *The International Journal of Robotics Research*, 28(10):1334–1354.
- Krupa, A., Folio, D., Novales, C., Vieyres, P., and Li, T. (2016). Robotized Tele-Echography: An Assisting Visibility Tool to Support Expert Diagnostic. *IEEE Systems Journal*, 10(3):974–983.
- Kuc, R. and Schwartz, M. (1979). Estimating the Acoustic Attenuation Coefficient Slope for Liver from Reflected Ultrasound Signals. *IEEE Transactions on Sonics and Ultrasonics*, 26(5):353–361.
- Kuhlemann, I. (2013). *Force and image adaptive strategies for robotised placement of 4D ultrasound probes*. Master’s Thesis, University of Lübeck.
- Kurohiji, T., Sigel, B., Justin, J., and Machi, J. (1990). Motion Marking in Color Doppler Ultrasound Needle and Catheter Visualization. *Journal of Ultrasound in Medicine*, 9:243–245.
- Kutarnia, J. and Pedersen, P. (2015). A Markov random field approach to group-wise registration/mosaicing with application to ultrasound. *Medical Image Analysis*, 24(1):106–124.
- Lauterbur, P. C. (1973). Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature*, 242:190–191.
- LaValle, S. and Kuffner, J. (1999). Randomized kinodynamic planning. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 473–479.
- Lee, D. and Krupa, A. (2011). Intensity-based visual servoing for non-rigid motion compensation of soft tissue structures due to physiological motion using

- 4D ultrasound. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2831–2836.
- Leroy, A., Mozer, P., Payan, Y., and Troccaz, J. (2004). Rigid Registration of Freehand 3D Ultrasound and CT-Scan Kidney Images. In Barillot, C., Haynor, D. R., and Hellier, P., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*, number 3216 in Lecture Notes in Computer Science, pages 837–844. Springer Berlin Heidelberg.
- Leung, C., Hashtrudi-Zaad, K., Foroughi, P., and Abolmaesumi, P. (2009). A real-time intrasubject elastic registration algorithm for dynamic 2-D ultrasound images. *Ultrasound in Medicine and Biology*, 35(7):1159 – 1176.
- Li, T., Krupa, A., and Collewet, C. (2011). A robust parametric active contour based on fourier descriptors. In *2011 18th IEEE International Conference on Image Processing*, pages 1037–1040.
- Linguraru, M. G., Vasilyev, N. V., Marx, G. R., Tworetzky, W., Del Nido, P. J., and Howe, R. D. (2008). Fast block flow tracking of atrial septal defects in 4D echocardiography. *Medical Image Analysis*, 12(4):397–412.
- Lippmann, G. (1881). Principe de la conservation de l’électricité. *Annales de chimie et de physique*, 5(24):145–178.
- Loizou, C. P., Pattichis, C. S., Pantziaris, M., Tyllis, T., and Nicolaidis, A. (2006). Quality evaluation of ultrasound imaging in the carotid artery based on normalization and speckle reduction filtering. *Medical and Biological Engineering and Computing*, 44(5):414.
- Long, J.-A., Lee, B. H., Guillotreau, J., Autorino, R., Laydner, H., Yakoubi, R., Rizkala, E., Stein, R. J., Kaouk, J. H., and Haber, G.-P. (2012). Real-Time Robotic Transrectal Ultrasound Navigation During Robotic Radical Prostatectomy: Initial Clinical Experience. *Urology*, 80(3):608–613.
- Ludwig, G. D. and Struthers, F. W. (1949). Considerations underlying the use of ultrasound to detect gallstones and foreign bodies in tissue. Technical report, Naval Medical Research Inst. Bethesda, MD.
- Madabhushi, A., Yang, P., Rosen, M., and Weinstein, S. (2006). Distinguishing Lesions from Posterior Acoustic Shadowing in Breast Ultrasound via Non-Linear Dimensionality Reduction. In *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2006. EMBS ’06*, pages 3070–3073.
- Magnavita, N., Bevilacqua, L., Mirk, P., Fileni, A., and Castellino, N. (1999). Work-related musculoskeletal complaints in sonologists. *Journal of Occupational and Environmental Medicine*, 41(11):981–988.

- Mallapragada, V., Sarkar, N., and Podder, T. K. (2011). Toward a Robot-Assisted Breast Intervention System. *IEEE/ASME Transactions on Mechatronics*, 16(6):1011–1020.
- Mansard, N. and Chaumette, F. (2009). Directional Redundancy for Robot Control. *IEEE Transactions on Automatic Control*, 54(6):1179–1192.
- Mansfield, P. (1977). Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics*, 10(3):L55.
- Marchand, E., Spindler, F., and Chaumette, F. (2005). ViSP for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics Automation Magazine*, 12(4):40–52.
- Masuda, K., Kimura, E., Tateishi, N., and Ishihara, K. (2001). Three dimensional motion mechanism of ultrasound probe and its application for tele-echography system. In *Proceedings. 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 1112–1116.
- Mauri, G., Cova, L., Beni, S. D., Ierace, T., Tondolo, T., Cerri, A., Goldberg, S. N., and Solbiati, L. (2014). Real-Time US-CT/MRI Image Fusion for Guidance of Thermal Ablation of Liver Tumors Undetectable with US: Results in 295 Cases. *CardioVascular and Interventional Radiology*, 38(1):143–151.
- Mebarki, R., Krupa, A., and Chaumette, F. (2008). Image moments-based ultrasound visual servoing. In *IEEE International Conference on Robotics and Automation, 2008. ICRA 2008*, pages 113–119.
- Mebarki, R., Krupa, A., and Chaumette, F. (2010). 2-D Ultrasound Probe Complete Guidance by Visual Servoing Using Image Moments. *IEEE Transactions on Robotics*, 26(2):296–306.
- Mignon, P., Poignet, P., and Troccaz, J. (2015). Using rotation for steerable needle detection in 3D color-Doppler ultrasound images. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1544–1547.
- Mignotte, M., Meunier, J., and Tardif, J.-C. (2001). Endocardial Boundary Estimation and Tracking in Echocardiographic Images using Deformable Template and Markov Random Fields. *Pattern Analysis & Applications*, 4(4):256–271.
- Mikic, I., Krucinski, S., and Thomas, J. D. (1998). Segmentation and tracking in echocardiographic sequences: active contours guided by optical flow estimates. *IEEE Transactions on Medical Imaging*, 17(2):274–284.

- Minhas, D. S., Engh, J. A., Fenske, M. M., and Riviere, C. N. (2007). Modeling of needle steering via duty-cycled spinning. In *IEEE International Conference on Engineering in Medicine and Biology Society*, pages 2756–2759.
- Misra, S., Reed, K. B., Schafer, B. W., Ramesh, K. T., and Okamura, A. M. (2010). Mechanics of Flexible Needles Robotically Steered through Soft Tissue. *The International Journal of Robotics Research*.
- Mukundan, R. and Ramakrishnan, K. R. (1998). *Moment Functions in Image Analysis — Theory and Applications*. World Scientific.
- Myronenko, A., Song, X., and Sahn, D. J. (2009). Maximum Likelihood Motion Estimation in 3D Echocardiography through Non-rigid Registration in Spherical Coordinates. In Ayache, N., Delingette, H., and Sermesant, M., editors, *Functional Imaging and Modeling of the Heart*, number 5528 in Lecture Notes in Computer Science, pages 427–436. Springer Berlin Heidelberg.
- Nadeau, C. and Krupa, A. (2013). Intensity-Based Ultrasound Visual Servoing: Modeling and Validation With 2-D and 3-D Probes. *IEEE Transactions on Robotics*, 29(4):1003–1015.
- Nadeau, C., Ren, H., Krupa, A., and Dupont, P. (2015). Intensity-Based Visual Servoing for Instrument and Tissue Tracking in 3D Ultrasound Volumes. *IEEE Transactions on Automation Science and Engineering*, 12(1):367–371.
- Najafi, M., Abolmaesumi, P., and Rohling, R. (2015). Single-Camera Closed-Form Real-Time Needle Tracking for Ultrasound-Guided Needle Insertion. *Ultrasound in Medicine & Biology*, 41(10):2663–2676.
- Nakadate, R., Solis, J., Takanishi, A., Minagawa, E., Sugawara, M., and Niki, K. (2011). Out-of-plane visual servoing method for tracking the carotid artery with a robot-assisted ultrasound diagnostic system. In *2011 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5267–5272.
- Nascimento, J. C. and Marques, J. S. (2008). Robust Shape Tracking With Multiple Models in Ultrasound Images. *IEEE Transactions on Image Processing*, 17(3):392–406.
- Navab, N., Hennersperger, C., Frisch, B., and Fürst, B. (2016). Personalized, Relevance-based Multimodal Robotic Imaging and Augmented Reality for Computer Assisted Interventions. *Medical Image Analysis*, 33:64–71.
- Nelder, J. A. and Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313.

- Nelson, T. R., Pretorius, D. H., Sklansky, M., and Hagen-Ansert, S. (1996). Three-dimensional echocardiographic evaluation of fetal heart anatomy and function: acquisition, analysis, and display. *Journal of Ultrasound in Medicine*, 15(1):1–9 quiz 11.
- Neshat, H. R. S. and Patel, R. V. (2008). Real-time parametric curved needle segmentation in 3D ultrasound images. In *2008 2nd IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 670–675.
- Neubach, Z. and Shoham, M. (2010). Ultrasound-Guided Robot for Flexible Needle Steering. *IEEE Transactions on Biomedical Engineering*, 57(4):799–805.
- Noble, J. A. and Boukerroui, D. (2006). Ultrasound image segmentation: a survey. *IEEE Transactions on Medical Imaging*, 25(8):987–1010.
- Novotny, P. M., Stoll, J. A., Dupont, P. E., and Howe, R. D. (2007a). Real-Time Visual Servoing of a Robot Using Three-Dimensional Ultrasound. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 2655–2660.
- Novotny, P. M., Stoll, J. A., Vasilyev, N. V., Nido, P. J. d., Dupont, P. E., Zickler, T. E., and Howe, R. D. (2007b). GPU based real-time instrument tracking with three-dimensional ultrasound. *Medical Image Analysis*, 11(5):458 – 464. Special Issue on the Ninth International Conference on Medical Image Computing and Computer-Assisted Interventions - MICCAI 2006.
- Okazawa, S., Ebrahimi, R., Chuang, J., Salcudean, S. E., and Rohling, R. (2005). Hand-held steerable needle device. *IEEE/ASME Transactions on Mechatronics*, 10(3):285–296.
- Ortmaier, T., Groger, M., Boehm, D. H., Falk, V., and Hirzinger, G. (2005). Motion estimation in beating heart surgery. *IEEE Transactions on Biomedical Engineering*, 52(10):1729–1740.
- Pappalardo, A., Albakri, A., Liu, C., Bascetta, L., De Momi, E., and Poignet, P. (2016). Hunt–Crossley model based force control for minimally invasive robotic surgery. *Biomedical Signal Processing and Control*, 29:31–43.
- Pascal, E. (1900). Zweiter Theil: Besondere Untersuchungen über Determinanten und Anwendungen. In *Die Determinanten*, pages 36–254. B. G. Teubner, Leipzig.
- Penney, G. P., Barratt, D. C., Chan, C. S. K., Slomczykowski, M., Carter, T. J., Edwards, P. J., and Hawkes, D. J. (2006). Cadaver validation of intensity-based ultrasound to CT registration. *Medical Image Analysis*, 10(3):385–395.

- Penney, G. P., Blackall, J. M., Hamady, M. S., Sabharwal, T., Adam, A., and Hawkes, D. J. (2004). Registration of freehand 3D ultrasound and magnetic resonance liver images. *Medical Image Analysis*, 8(1):81–91.
- Penrose, R. (1955). A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(03):406–413.
- Pierrot, F., Dombre, E., Dégoulange, E., Urbain, L., Caron, P., Boudet, S., Gariépy, J., and Mégnien, J.-L. (1999). Hippocrate: a safe robot arm for medical applications with force feedback. *Medical Image Analysis*, 3(3):285–300.
- Pretorius, D. H. and Nelson, T. R. (1995). Fetal face visualization using three-dimensional ultrasonography. *Journal of Ultrasound in Medicine*, 14(5):349–356.
- Promayon, E., Fouard, C., Bailet, M., Deram, A., Fiard, G., Hungr, N., Luboz, V., Payan, Y., Sarrazin, J., Saubat, N., Selmi, S. Y., Voros, S., Cinquin, P., and Troccaz, J. (2013). Using CamiTK for rapid prototyping of interactive Computer Assisted Medical Intervention applications. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4933–4936.
- Qiu, W., Ding, M., and Yuchi, M. (2008). Needle Segmentation Using 3D Quick Randomized Hough Transform. In *First International Conference on Intelligent Networks and Intelligent Systems, 2008. ICINIS '08*, pages 449–452.
- Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., and Ng, A. (2009). ROS: an open-source Robot Operating System. In *ICRA Workshop on Open Source Software*, volume 3, page 5.
- Radon, J. (1917). Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. In *Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig*, number 69 in Mathematisch-Physische Klasse, pages 262–277. B. G. Teubner.
- Richa, R., Poignet, P., and Liu, C. (2010). Three-dimensional Motion Tracking for Beating Heart Surgery Using a Thin-plate Spline Deformable Model. *The International Journal of Robotics Research*, 29(2-3):218–230.
- Roesthuis, R. J., van de Berg, N. J., van den Dobbelsteen, J. J., and Misra, S. (2015). Modeling and steering of a novel actuated-tip needle through a soft-tissue simulant using Fiber Bragg Grating sensors. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2283–2289.

- Röntgen, W. C. (1898). Ueber eine neue Art von Strahlen. *Annalen der Physik*, 300(1):1–11.
- Royer, L., Dardenne, G., Bras, A. L., Marchal, M., and Krupa, A. (2015). Tracking of Non-rigid Targets in 3D US Images: Results on CLUST 2015. In *Proceedings of MICCAI 2015 Challenge on Liver Ultrasound Tracking*.
- Royer, L., Krupa, A., Dardenne, G., Le Bras, A., Marchand, E., and Marchal, M. (2017). Real-time Target Tracking of Soft Tissues in 3D Ultrasound Images Based on Robust Visual Information and Mechanical Simulation. *Medical Image Analysis*, 35:582–598.
- Salcudean, S., Bell, G., Bachmann, S., Zhu, W.-H., Abolmaesumi, P., and Lawrence, P. D. (1999). Robot-assisted diagnostic ultrasound—design and feasibility experiments. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 1999*, pages 1062–1071. Springer.
- Samson, C., Espiau, B., and Le Borgne, M. (1991). *Robot control: The task function approach*. Oxford University Press.
- Schulte zu Berge, C., Baust, M., Kapoor, A., and Navab, N. (2014). Predicate-Based Focus-and-Context Visualization for 3D Ultrasound. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2379–2387.
- Schulte zu Berge, C., Declara, D., Hennesperger, C., Baust, M., and Navab, N. (2015). Real-time uncertainty visualization for B-mode ultrasound. In *2015 IEEE Scientific Visualization Conference (SciVis)*, pages 33–40.
- Sears, P. and Dupont, P. (2006). A Steerable Needle Technology Using Curved Concentric Tubes. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2850–2856.
- Seferidis, V. and Ghanbari, M. (1994). Generalised block-matching motion estimation using quad-tree structured spatial decomposition. *IEE Proceedings — Vision, Image and Signal Processing*, 141(6):446–452.
- Seo, J., Koizumi, N., Funamoto, T., Sugita, N., Yoshinaka, K., Nomiya, A., Homma, Y., Matsumoto, Y., and Mitsuishi, M. (2011). Visual servoing for a US-guided therapeutic HIFU system by coagulated lesion tracking: a phantom study. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 7(2):237–247.
- Shahriari, N., Hekman, E., Oudkerk, M., and Misra, S. (2015). Design and evaluation of a computed tomography (CT)-compatible needle insertion device using an electromagnetic tracking system and CT images. *International Journal of Computer Assisted Radiology and Surgery*, 10(11):1845–1852.

- Shahriari, N., Roesthuis, R. J., van de Berg, N. J., van den Dobbelsteen, J. J., and Misra, S. (2016). Steering an actuated-tip needle in biological tissue: Fusing FBG-sensor data and ultrasound images. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4443–4449.
- Shannon, C. E. (1949). Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1):10–21.
- Sheet, D., Karamalis, A., Eslami, A., Noël, P., Chatterjee, J., Ray, A. K., Laine, A. F., Carlier, S. G., Navab, N., and Katouzian, A. (2014). Joint learning of ultrasonic backscattering statistical physics and signal confidence primal for characterizing atherosclerotic plaques using intravascular ultrasound. *Medical Image Analysis*, 18(1):103–117.
- Siciliano, B. and Slotine, J. J. E. (1991). A general framework for managing multiple tasks in highly redundant robotic systems. In *'Robots in Unstructured Environments', 91 ICAR., Fifth International Conference on Advanced Robotics, 1991*, pages 1211–1216 vol.2.
- Song, X., Myronenko, A., and Sahn, D. J. (2007). Speckle Tracking in 3D Echocardiography with Motion Coherence. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–7.
- Stewart, B. and Wild, C. P. (2014). *World Cancer Report 2014*. Number 3 in World Cancer Reports. International Agency for Research on Cancer.
- Stolka, P. J., Foroughi, P., Rendina, M., Weiss, C. R., Hager, G. D., and Boctor, E. M. (2014). Needle Guidance Using Handheld Stereo Vision and Projection for Ultrasound-Based Interventions. In Golland, P., Hata, N., Barillot, C., Hornegger, J., and Howe, R., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*, number 8674 in Lecture Notes in Computer Science, pages 684–691. Springer International Publishing.
- Sublett, J. W., Dempsey, B. J., and Weaver, A. C. (1995). Design and implementation of a digital teleultrasound system for real-time remote diagnosis. In *, Proceedings of the Eighth IEEE Symposium on Computer-Based Medical Systems, 1995*, pages 292–298.
- Taylor, R. H. and Stoianovici, D. (2003). Medical robotics in computer-integrated surgery. *IEEE Transactions on Robotics and Automation*, 19(5):765–781.
- Tempany, C. M. C., McDannold, N. J., Hynynen, K., and Jolesz, F. A. (2011). Focused Ultrasound Surgery in Oncology: Overview and Principles. *Radiology*, 259(1):39–56.

- Tong, S., Cardinal, H. N., McLoughlin, R. F., Downey, D. B., and Fenster, A. (1998). Intra- and inter-observer variability and reliability of prostate volume measurement via two-dimensional and three-dimensional ultrasound imaging. *Ultrasound in Medicine & Biology*, 24(5):673–681.
- Touil, B., Basarab, A., Delachartre, P., Bernard, O., and Friboulet, D. (2010). Analysis of motion tracking in echocardiographic image sequences: Influence of system geometry and point-spread function. *Ultrasonics*, 50(3):373–386.
- Trahey, G. E., Allison, J. W., and Ramm, O. T. v. (1987). Angle Independent Ultrasonic Detection of Blood Flow. *IEEE Transactions on Biomedical Engineering*, BME-34(12):965–967.
- Uherčík, M., Kybic, J., Liebgott, H., and Cachard, C. (2008). Multi-resolution parallel integral projection for fast localization of a straight electrode in 3D ultrasound images. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 33–36.
- Uherčík, M., Kybic, J., Liebgott, H., and Cachard, C. (2010). Model Fitting Using RANSAC for Surgical Tool Localization in 3-D Ultrasound Images. *IEEE Transactions on Biomedical Engineering*, 57(8):1907–1916.
- Vanderpool, H. E., Friis, E. A., Smith, B. S., and Harms, K. L. (1993). Prevalence of carpal tunnel syndrome and other work-related musculoskeletal problems in cardiac sonographers. *Journal of occupational and environmental medicine*, 35(6):604–610.
- Vilchis, A., Troccaz, J., Cinquin, P., Masuda, K., and Pellissier, F. (2003). A new robot architecture for tele-echography. *IEEE Transactions on Robotics and Automation*, 19(5):922–926.
- Virga, S., Zettinig, O., Esposito, M., Pfister, K., Frisch, B., Neff, T., Navab, N., and Hennersperger, C. (2016). Automatic Force-Compliant Robotic Ultrasound Screening of Abdominal Aortic Aneurysms. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Vitrani, M.-A., Baumann, M., Reversat, D., Morel, G., Moreau-Gaudry, A., and Mozer, P. (2016). Prostate biopsies assisted by comanipulated probeholder: first in man. *International Journal of Computer Assisted Radiology and Surgery*, 11(6):1153–1161.
- Vrooijink, G. J., Abayazid, M., Patil, S., Alterovitz, R., and Misra, S. (2014). Needle path planning and steering in a three-dimensional non-static environment using two-dimensional ultrasound images. *International Journal of Robotics Research*, 33(10):1361–1374.
- Wachinger, C., Klein, T., and Navab, N. (2012). Locally adaptive Nakagami-based ultrasound similarity measures. *Ultrasonics*, 52(4):547–554.

- Waine, M., Rossa, C., Sloboda, R., Usmani, N., and Tavakoli, M. (2016). Needle Tracking and Deflection Prediction for Robot-Assisted Needle Insertion Using 2D Ultrasound Images. *Journal of Medical Robotics Research*, 01(01):1640001.
- Webster III, R. J., Kim, J. S., Cowan, N. J., Chirikjian, G. S., and Okamura, A. M. (2006). Nonholonomic Modeling of Needle Steering. *The International Journal of Robotics Research*, 25(5-6):509–525.
- Wei, Z., Ding, M., Downey, D., and Fenster, A. (2005). 3D TRUS Guided Robot Assisted Prostate Brachytherapy. In Duncan, J. and Gerig, G., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005*, volume 3750 of *Lecture Notes in Computer Science*, pages 17–24. Springer Berlin / Heidelberg.
- Wein, W., Brunke, S., Khamene, A., Callstrom, M. R., and Navab, N. (2008). Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical Image Analysis*, 12(5):577–585.
- Wein, W., Karamalis, A., Baumgartner, A., and Navab, N. (2015). Automatic bone detection and soft tissue aware ultrasound–CT registration for computer-aided orthopedic surgery. *International Journal of Computer Assisted Radiology and Surgery*, 10(6):971–979.
- Wein, W., Roper, B., and Navab, N. (2007). Integrating Diagnostic B-Mode Ultrasonography Into CT-Based Radiation Treatment Planning. *IEEE Transactions on Medical Imaging*, 26(6):866–879.
- Wells, P. N. T. (1975). Absorption and dispersion of ultrasound in biological tissue. *Ultrasound in Medicine & Biology*, 1(4):369–376.
- Wild, J. J. and Reid, J. M. (1952). Application of Echo-Ranging Techniques to the Determination of Structure of Biological Tissues. *Science*, 115(2983):226–230.
- Wojcinski, S., Farrokh, A., Hille, U., Wiskirchen, J., Gyapong, S., Soliman, A. A., Degenhardt, F., and Hillemanns, P. (2011). The Automated Breast Volume Scanner (ABVS): initial experiences in lesion detection compared with conventional handheld B-mode ultrasound: a pilot study of 50 cases. *International Journal of Women’s Health*, 3:337–346.
- Woo, J. and Roh, Y. (2012). Ultrasonic 2D matrix array transducer for volumetric imaging in real time. In *2012 IEEE International Ultrasonics Symposium*, pages 1568–1571.
- Wood, N., Shahrour, K., Ost, M., and Riviere, C. (2010). Needle steering system using duty-cycled rotation for percutaneous kidney access. In *IEEE*

- International Conference on Engineering in Medicine and Biology Society*, pages 5432–5435.
- Wucherer, P., Stefan, P., Weidert, S., Fallavollita, P., and Navab, N. (2013). Development and Procedural Evaluation of Immersive Medical Simulation Environments. In Barratt, D., Cotin, S., Fichtinger, G., Jannin, P., and Navab, N., editors, *Information Processing in Computer-Assisted Interventions*, number 7915 in Lecture Notes in Computer Science, pages 1–10. Springer Berlin Heidelberg.
- Xu, J., Duindam, V., Alterovitz, R., and Goldberg, K. (2008). Motion planning for steerable needles in 3D environments with obstacles using rapidly-exploring Random Trees and backchaining. In *2008 IEEE International Conference on Automation Science and Engineering*, pages 41–46.
- Yeung, F., Levinson, S. F., Fu, D., and Parker, K. J. (1998). Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh. *IEEE Transactions on Medical Imaging*, 17(6):945–956.
- Zaaroor, M., Bejerano, Y., Weinfeld, Z., and Ben-Haim, S. (2001). Novel magnetic technology for intraoperative intracranial frameless navigation : In vivo and in vitro results. *Neurosurgery*, 48(5):1100–1108.
- Zahnd, G., Boussel, L., Marion, A., Durand, M., Moulin, P., Sérusclat, A., and Vray, D. (2011). Measurement of Two-Dimensional Movement Parameters of the Carotid Artery Wall for Early Detection of Arteriosclerosis: A Preliminary Clinical Study. *Ultrasound in Medicine & Biology*, 37(9):1421–1429.
- Zetting, O., Fuerst, B., Kojcev, R., Esposito, M., Salehi, M., Wein, W., Rackerseder, J., Sinibaldi, E., Frisch, B., and Navab, N. (2016). Toward real-time 3D ultrasound registration-based visual servoing for interventional navigation. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 945–950.
- Zhao, Y., Liebgott, H., and Cachard, C. (2012). Tracking micro tool in a dynamic 3D ultrasound situation using Kalman filter and RANSAC algorithm. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1076 –1079.
- Zhou, H., Qiu, W., Ding, M., and Zhang, S. (2008). Automatic needle segmentation in 3D ultrasound images using 3D improved Hough transform. In *Medical Imaging 2008: Visualization, Image-guided Procedures, and Modeling*, volume 6918 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*.

List of Figures

1.1	Specular reflection and transmission of a sound wave at an interface between two media	12
1.2	Generation of a focused ultrasound beam	13
1.3	Radio frequency ultrasound signal	14
1.4	Envelope-detected ultrasound signal	15
1.5	Geometry of a 2D linear ultrasound probe	17
1.6	Geometry of a 2D convex ultrasound probe	17
1.7	Geometry of a motorized convex 3D ultrasound probe	20
1.8	Sliding window approach for block matching	22
1.9	Gauss-Newton optimization approach for block matching	23
1.10	Deformable tracking with an irregular mesh	25
1.11	Vessel tracking in ultrasound images using an active contour algorithm	27
1.12	Optical needle tracking system	30
1.13	Curve fitting for needle detection in ultrasound images	33
1.14	Polynomial curve model of a flexible needle	34
1.15	Needle detection using the RANSAC algorithm	36
1.16	Needle tip localization by intensity drop detection	39
1.17	Dynamic model of a flexible needle	41
1.18	Particle filtering	43
1.19	Ultrasound images with shadowing	44
1.20	Acoustic shadow detection	47
1.21	Graph layout for the computation of the random walks confidence map	49
1.22	Confidence maps estimated via the random walks method	51
1.23	Confidence maps estimated via the scan line integration method	54
1.24	Computation time for the random walks confidence map and the scan line integration methods	55
1.25	Temporal regularity for the random walks and scan line integration methods	57
2.1	Visual servoing control loop	60
2.2	Configuration of a robotic manipulator	61
2.3	Null-space projection of a secondary task	65

LIST OF FIGURES

2.4	In-plane and out-of-plane directions for a 2D ultrasound probe	67
2.5	Control frame attached to an ultrasound probe	69
2.6	Relation between a target detected in a 2D ultrasound image and a physical point	70
2.7	Intersection between an object and the image plane	71
2.8	Dense visual feature for intensity-based visual servoing	73
2.9	Schematic of a bevel-tip needle inserted in soft tissues	76
2.10	Needle steering for obstacle avoidance	77
2.11	Insertion types for bevel-tip needles	78
2.12	Control velocities in the needle tip frame	82
2.13	Visual features for ultrasound-based needle steering	83
2.14	Needle steering under 3D ultrasound guidance	85
3.1	Acoustic window in ultrasound imaging	88
3.2	Relation between the mean image confidence and the contact force	90
3.3	Relation between the in-plane orientation of a probe and the polar coordinates of a point	92
3.4	Angular confidence feature for three different positions of an ultrasound probe	96
3.5	Estimation of the direction of maximum confidence on a sequence of ultrasound images	97
3.6	Estimation of the confidence-weighted angular feature on a sequence of ultrasound images with the scan line integration method	98
3.7	Estimation of the confidence-weighted angular feature on a sequence of ultrasound images with the random walks method	98
3.8	Imaging geometry of a motorized 3D convex ultrasound probe	100
3.9	Overview of the global confidence-driven control framework	102
3.10	Configuration of the probe frame, contact frame, sensor frame, and gravity frame	103
3.11	Elastic compression model following Hooke's law	105
3.12	2D illustration of the system's behavior with an homogeneous medium	107
3.13	2D illustration of the system's behavior with the presence of a strong acoustic attenuator	108
3.14	Overview of the target-specific confidence-driven control framework	109
3.15	2D illustration of the constraints induced by the force control and target centering tasks	111
4.1	Experimental setup	116
4.2	Multi-threaded architecture of the control application	120
4.3	3D global confidence-driven control	121
4.4	Convergence of the global confidence-driven control in a static environment	121

4.5	Target-specific confidence-driven control	122
4.6	Convergence accuracy of the target-specific confidence-driven control	124
4.7	Convergence of the target-specific confidence-driven control in a static environment	125
4.8	3D global confidence-driven control	127
4.9	Evolution of the confidence feature errors during a shadow avoidance experiment under global confidence-driven control	128
4.10	Target-specific confidence-driven control	129
4.11	Feature evolution during target-specific confidence-driven control	130
4.12	Evolution of the normalized correlation between the tracked target and the initial template, (a) without and (b) with confidence control.	132
4.13	Target tracking without confidence control	132
4.14	Confidence-optimized tele-echography on an ABDFAN phantom with a 2D ultrasound probe	134
4.15	Tele-echography with the global confidence-driven control . . .	135
4.16	Confidence-optimized tele-echography on a human volunteer – frontal abdominal scan – normal breathing	137
4.17	Evolution of the confidence feature errors during a tele-echography on a human volunteer – frontal abdominal scan – normal breathing	137
4.18	Confidence-optimized ultrasound acquisition on a human volunteer – frontal abdominal scan – heavy breathing	138
4.19	Ultrasound probe elevation during heavy breathing	138
4.20	Evolution of the confidence feature errors during an ultrasound acquisition on a human volunteer – frontal abdominal scan – heavy breathing	139
4.21	Confidence-optimized tele-echography on a human subject – lateral abdominal scan – normal breathing	140
4.22	Evolution of the confidence feature errors during a tele-echography on a human volunteer – lateral abdominal scan – normal breathing	140

Acronyms

- ABVS** Automated Breast Volume Scanner. 81
- CIS** Computer-Integrated Surgery. 2
- CT** Computerized Tomography. 1, 46, 48, 79, 80
- DOF** Degrees of Freedom. 61, 66, 71
- GPU** Graphics Processing Unit. 32
- HIFU** High Intensity Focused Ultrasound. 4, 149
- MRI** Magnetic Resonance imaging. 1, 46, 79, 80
- MSS** Minimal Sampling Set. 36
- PIP** Parallel Integral Projection. 32
- RANSAC** Random Sample Consensus. 32, 35–39
- RF** Radio Frequency. 13–15, 45, 48, 52, 146
- ROI** Region of Interest. 23
- RRT** Rapidly-Exploring Random Tree. 79, 81
- RW** Random Walks. 48, 96, 97
- SIR** Sequential Importance Re-Sampling. 42, 43, 82
- SLI** Scan Line Integration. 48, 96
- SSD** Sum of Squared Differences. 22, 67
- TRUS** Transrectal Ultrasound. 3, 4
- US** Ultrasound. 1, 46, 48

Résumé

La manipulation robotique d'une sonde échographique a été un important sujet de recherche depuis plusieurs années. Plus particulièrement, des méthodes d'asservissement visuel guidé par échographie ont été développées pour accomplir différentes tâches, telles que la compensation de mouvement, le maintien de la visibilité d'une structure pendant la téléopération, ou le suivi d'un instrument chirurgical. Cependant, en raison de la nature des images échographiques, garantir une bonne qualité d'image durant l'acquisition est un problème difficile, qui a jusqu'ici été très peu abordé.

Cette thèse traite du contrôle de la qualité des images échographiques acquises par une sonde robotisée. La qualité du signal acoustique au sein de l'image est représentée par une carte de confiance, qui est ensuite utilisée comme signal d'entrée d'une loi de commande permettant d'optimiser le positionnement de la sonde échographique. Une commande hybride est également proposée pour optimiser la fenêtre acoustique pour une cible anatomique qui est détectée dans l'image. L'approche proposée est illustrée dans le cas d'un scénario de télé-échographie, où le contrôle de la sonde est partagé entre la machine et le téléopérateur.

Mots-clés — Robotique médicale, échographie, asservissement visuel, qualité de l'image

Abstract

The robotic guidance of an ultrasound probe has been extensively studied as a way to assist sonographers in performing an exam. In particular, ultrasound-based visual servoing methods have been developed to fulfill various tasks, such as compensating for physiological motion, maintaining the visibility of an anatomic target during teleoperation, or tracking a surgical instrument. However, due to the specific nature of ultrasound images, guaranteeing a good image quality during the procedure remains an unaddressed challenge.

This thesis deals with the control of ultrasound image quality for a robot-held ultrasound probe. The ultrasound signal quality within the image is represented by a confidence map, which is used to design a servo control law for optimizing the placement of the ultrasound probe. A control fusion is also proposed to optimize the acoustic window for a specific anatomical target which is tracked in the ultrasound images. The method is illustrated in a teleoperation scenario, where the control is shared between the automatic controller and a human operator.

Keywords — Medical robotics, ultrasonography, visual servoing, image quality