

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE LA
MAÎTRISE EN GÉNIE ÉLECTRIQUE
M. Ing.

PAR
Juliette CONRATH

RECALAGE D'IMAGES ÉCHOGRAPHIQUES PAR APPRENTISSAGE POUR
L'ÉCHOGRAPHIE 3D MAIN LIBRE SANS CAPTEUR DE POSITION

MONTRÉAL, LE 24 AOÛT 2012

© Tous droits réservés

Cette licence signifie qu'il est interdit de reproduire, d'enregistrer ou de diffuser en tout ou en partie, le présent document. Le lecteur qui désire imprimer ou conserver sur un autre media une partie importante de ce document, doit obligatoirement en demander l'autorisation à l'auteur.

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE:

Mme Catherine Laporte, directrice de mémoire
Département de génie électrique à l'École de technologie supérieure

M. Christian Desrosiers, président du jury
Département de génie logiciel et des TI à l'École de technologie supérieure

M. Jean-Marc Lina, membre du jury
Département de génie électrique à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 14 AOÛT 2012

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

En premier lieu, je tiens à exprimer ma plus sincère et profonde reconnaissance envers ma directrice de recherche, Catherine Laporte. Elle a su, par son soutien, ses conseils et ses encouragements, m'offrir les meilleures conditions possibles pour réaliser le travail qu'elle m'a confié, et c'est bien grâce à elle que ma maîtrise à l'ÉTS s'est révélée une expérience des plus enrichissantes. C'est à ses côtés que j'ai découvert et pris goût à la recherche, et je lui suis infiniment redevable des perspectives d'avenir qui se sont ainsi ouvertes à moi. J'espère que le présent mémoire et mes travaux futurs seront à la hauteur de ses enseignements.

Je remercie également les autres étudiants de notre groupe de travail, François-Pierre Chartrand-Dufresne, Hacène Chebrek et Housseem Gueziri pour leur écoute attentive et leur regard critique sur mes différents travaux, ainsi que les professeurs et étudiants du laboratoire LATIS.

Merci également à Sylvain Brisebois pour son travail minutieux sur le système de positionnement motorisé, je regrette de ne pas avoir eu le temps de réellement m'en servir, mais je suis certaine qu'il sera un atout précieux pour la suite des travaux.

Je tiens également à remercier les personnes de l'INSA de Lyon m'ayant offert l'opportunité de ce double diplôme, notamment Tanneguy Redarce, Guillaume Villemaud et Claudine Gehin, ainsi que les responsables des relations internationales côté ÉTS, Éric Doré et Perrine Wirth, pour leur soutien précieux dans les démarches administratives parfois complexes. Je remercie également mes collègues insaliens et amis qui m'ont accompagnée dans cette aventure, Simon Noirot et Cyril Daran-Daneau, avec qui j'ai découvert la culture québécoise et appris, avec enthousiasme, à m'adapter au milieu universitaire québécois.

Un merci tout particulier à Faten Mhiri pour son écoute, ses conseils et suggestions, mais aussi et surtout pour son amitié et nos bons moments détentes pendant les périodes difficiles. Je lui souhaite une grande réussite dans la poursuite de son doctorat et le reste de sa carrière, sans aucun doute de son succès. Je remercie également Julien Couët et Jonathan Hadida qui ont su m'inspirer par leur passion, et que je suis heureuse de compter parmi mes amis.

Je tiens à exprimer ma gratitude envers mes parents Juana et Emmanuel Conrath, ainsi que mon frère Pierre et mes soeurs Alice et Susan, qui m'ont toujours soutenue et appuyée dans la poursuite de mes études, et qui ont su rester très présents malgré la distance. Je remercie également mes amis lointains, de France et d'ailleurs, mais toujours proches dans mon coeur, ainsi que les amis québécois avec qui j'ai eu le plaisir de partager mes moments de loisir, et que je regrette de quitter. Le dernier mot revient à mon compagnon, Kotaro Fukasaku, qui a traversé cette expérience avec moi et a su m'épauler, m'écouter, et faire preuve d'une infinie patience pendant mes longues heures de travail nocturne. Je n'aurais pas pu venir à bout de ces travaux sans son soutien sans faille, et je ne peux le remercier assez pour la force qu'il me donne jour après jour.

Je remercie finalement le CRSNG qui a financé ce projet.

RECALAGE D'IMAGES ÉCHOGRAPHIQUES PAR APPRENTISSAGE POUR L'ÉCHOGRAPHIE 3D MAIN LIBRE SANS CAPTEUR DE POSITION

Juliette CONRATH

RÉSUMÉ

L'échographie 3D main libre sans capteur de position est une approche permettant la reconstruction de volumes échographiques par l'analyse du contenu des images d'une séquence acquise sans contrainte avec le matériel conventionnel 2D. La corrélation entre les images permet en effet de déduire la distance les séparant grâce à un modèle de décorrélation du speckle, tandis que les méthodes conventionnelles de recalage permettent d'estimer le mouvement dans le plan des images. La complexité du mouvement main libre engendre cependant des erreurs dans l'estimation du déplacement de la sonde, du fait des sources multiples de décorrélation. Ce mémoire présente d'abord une étude de ces erreurs grâce à des expérimentations réalisées sur des paires d'images synthétiques. Cette étude permet de confirmer l'hypothèse qu'une grande partie des erreurs sont systématiques et qu'il existe une relation entre ces erreurs et les valeurs réelles du déplacement. Par la suite, une méthode permettant de modéliser cette relation est présentée, dans le but de corriger les estimations. La modélisation est réalisée par apprentissage bayésien de régression par Machine à Vecteurs de Pertinence. Les expérimentations sur imagerie synthétique ont permis de démontrer la validité de cette nouvelle méthode : les résultats montrent en effet une réduction significative des erreurs par rapport à la méthode conventionnelle.

Mot-clés: Recalage d'images échographiques, Décorrélation du speckle, Correction d'estimation par apprentissage, Échographie 3D main libre

RECALAGE D'IMAGES ÉCHOGRAPHIQUES PAR APPRENTISSAGE POUR L'ÉCHOGRAPHIE 3D MAIN LIBRE SANS CAPTEUR DE POSITION

Juliette CONRATH

ABSTRACT

Sensorless freehand 3D ultrasound imaging is an approach which allows the reconstruction of ultrasound volumes by analyzing the contents of the 2D scans in a freely acquired sequence to estimate their relative positions, instead of relying on a position sensor. Indeed, the distance between scans can be inferred using a speckle decorrelation model which describes the relationship between image correlation and elevational distance, while in-plane motion can be estimated by conventional registration methods. However, each type of motion (in-plane and out-of-plane) occurring in freehand motion is a source of decorrelation, and these multiple sources cause motion estimation errors. A study of these errors based on experiments on synthetic imagery is first presented, confirming the hypothesis that the errors are largely systematic and that a relationship between these errors and the ground truth displacement exists. A novel method to correct the estimations is thus proposed, based on the learning of this relationship using the Relevance Vector Machine algorithm. Experiments on synthetic imagery demonstrate the validity and reliability of our method : the results show a significant error reduction in comparison with the conventional method.

Keywords: Ultrasound registration, Speckle decorrelation, Estimation correction by learning, Freehand 3D ultrasound

TABLE DES MATIÈRES

	Page
INTRODUCTION.....	1
CHAPITRE 1 PRINCIPES FONDAMENTAUX ET REVUE DE LITTÉRATURE	7
1.1 Les signaux échographiques	7
1.1.1 Fonctionnement de l'échographie	7
1.1.2 Le speckle	11
1.1.3 Modélisation et statistiques des signaux échographiques	12
1.2 Recalage.....	16
1.2.1 Recalage hors plan et échographie 3D main libre sans capteur de position	18
1.2.2 Recalage dans le plan	21
1.2.3 Recalage global et erreurs engendrées	24
1.3 Problème étudié.....	26
1.3.1 Définition du contexte d'étude	26
1.3.2 Formalisation du problème	27
1.4 Conclusion	29
CHAPITRE 2 ALGORITHME DE RECALAGE NAIF	31
2.1 Recalage hors plan.....	33
2.2 Calculs de corrélation et localisation sous-pixel	34
2.2.1 Matrice de corrélation.....	35
2.2.2 Précision sous-pixel et interpolation	39
2.2.3 Calcul du vecteur de translation à trois dimensions	40
2.3 Calcul de la transformation globale par alignement de Procruste	42
2.4 Expérimentations sur l'algorithme naïf	47
2.4.1 Protocole expérimental	47
2.4.2 Mesures d'erreur	49
2.4.3 Expériences sur séquences avec translations pures.....	50
2.4.4 Expériences sur séquences avec mouvements parasites.....	54
2.5 Conclusion	64
CHAPITRE 3 ALGORITHME DE CORRECTION PAR APPRENTISSAGE.....	67
3.1 Méthodes d'apprentissage	68
3.1.1 Apprentissage par RVM monovarié	70
3.1.2 RVM à sortie multivariée	76
3.2 Expérimentation sur la correction de l'erreur	78
3.2.1 Comparaison des modèles	79
3.2.2 Tests sur séquences avec mouvement restreint	82
3.3 Conclusion	91
CONCLUSION.....	95

BIBLIOGRAPHIE	100
---------------------	-----

LISTE DES FIGURES

	Page
Figure 0.1	Méthodologie proposée 5
Figure 1.1	Image échographique avec speckle - Tirée de Karim et Nancy Nice (2000).. 8
Figure 1.2	Organisation des transducteurs et focalisation..... 9
Figure 1.3	Réflexion et transmission des ondes ultrasonores 10
Figure 1.4	Interférences dues à deux particules (gauche) et à de multiples particules (droite) - Tirées de Støylen (2010) 11
Figure 1.5	Superposition des cellules de résolution 14
Figure 1.6	Exemple de courbe de décorrélation - Adaptée de Laporte (2011) 17
Figure 1.7	Description des différents types de mouvement 19
Figure 1.8	Définition du repère 28
Figure 2.1	Schéma fonctionnel de l'algorithme de recalage naif 32
Figure 2.2	Construction de la matrice de corrélation pour une fenêtre de taille 3×3 et la zone de référence correspondante de taille 7×7 36
Figure 2.3	Exemple de matrice de corrélation représentée sous forme de surface 39
Figure 2.4	Estimation des trois dimensions de translation pour le recalage d'une fenêtre 41
Figure 2.5	Représentation schématique des translations dans le plan estimées pour chaque fenêtre 41
Figure 2.6	Représentation schématique des trois translations de recalage estimées pour chaque fenêtre 43
Figure 2.7	Disposition des points dans le repère 44
Figure 2.8	Étapes de calcul de la transformation rigide optimale 45
Figure 2.9	Exemple d'image synthétique et division en fenêtres 48
Figure 2.10	Résultats sur séquences de mouvement pur selon chaque composante 52

Figure 2.11	Représentation des vecteurs de translation estimés pour chaque fenêtre, dans le cas d'une translation pure de 1mm selon x	53
Figure 2.12	Représentation des vecteurs de translation estimés pour chaque fenêtre, dans le cas d'une translation pure de 1mm selon y	54
Figure 2.13	Résultats sur séquences avec combinaison de translations dans le plan.....	56
Figure 2.14	Résultats sur séquences avec combinaison de translations et rotations dans le plan	57
Figure 2.15	Résultats sur séquences de mouvement hors plan	58
Figure 2.16	Résultats sur séquences de translations dans le plan constantes associées à des translations hors plan variables	60
Figure 2.17	Résultats sur séquences de translations dans le plan constantes associées à des rotations hors plan variables	61
Figure 2.18	Résultats sur séquences de translations hors plan associées à du mouvement dans le plan	63
Figure 3.1	Construction des modèles selon les différentes méthodes proposées	69
Figure 3.2	Comparaison de la distribution des erreurs de validation obtenues pour chaque méthode sur la base de données de 400 séquences libres.....	80
Figure 3.3	Comparaison des distributions des erreurs de validation sur chaque composante pour l'estimation naive et RVM	81
Figure 3.4	Comparaison des estimés naifs et RVM de z pour les 290 séquences libres, pour des valeurs réelles croissantes de z	82
Figure 3.5	Résultats sur séquences de mouvement pur dans le plan	84
Figure 3.6	Résultats sur séquences de mouvement pur hors plan	85
Figure 3.7	Résultats sur séquences de combinaisons de translations dans le plan	86
Figure 3.8	Résultats sur séquences de combinaisons de translations et rotations dans le plan	87
Figure 3.9	Résultats sur séquences de mouvement hors plan	88
Figure 3.10	Résultats sur séquences de translations dans le plan constantes associées à des translations hors plan variables	90

Figure 3.11	Résultats sur séquences de translations dans le plan constantes associées à des rotations hors plan variables	91
Figure 3.12	Résultats sur séquences complexes	92

INTRODUCTION

L'imagerie par ultrasons, ou échographie, est une technique d'imagerie médicale qui s'est développée durant la seconde moitié du 20^{ème} siècle. Elle est basée sur les échos d'ondes ultrasonores émises en direction du sujet et permet en particulier une bonne visualisation des tissus mous. Elle est de nos jours très bien maîtrisée et très utilisée, notamment en obstétrique mais aussi pour diverses autres applications, comme le guidage lors d'opérations chirurgicales. En effet, elle présente de nombreux avantages qui en font la modalité de choix dans plusieurs situations : le matériel nécessaire est peu encombrant, facilement transportable, et peu coûteux. Ces arguments permettent une utilisation beaucoup moins limitée que d'autres modalités comme l'imagerie à résonance magnétique ou la tomodensitométrie. Un autre avantage primordial par rapport aux autres techniques est l'absence de rayonnements ionisants, qui présentent des risques pour la santé des patients.

Actuellement, la visualisation en deux dimensions (2D) constitue la norme pour les applications médicales de l'échographie. Dans ces conditions, le praticien acquiert de multiples images sous différents angles, afin de reconstruire mentalement le volume observé. Ceci implique une subjectivité potentiellement problématique. Une visualisation directe en trois dimensions (3D) permet quant à elle d'observer les objets sous n'importe quel angle et de faciliter l'observation de structures complexes. Les mesures deviennent alors objectives, précises et répétables (Prager *et al.*, 2010). L'acquisition de volumes permet également de faciliter la visualisation multimodale, c'est-à-dire le recalage avec des images provenant d'autres modalités d'imagerie.

Les travaux présentés dans ce mémoire concernent une méthode d'acquisition de volumes échographiques : *l'échographie 3D main libre sans capteur de position*. Cette méthode a l'avantage important de ne requérir aucun ajout de matériel par rapport à l'acquisition 2D : elle repose sur la reconstruction de volumes à partir des échographies 2D, leurs positions relatives étant obtenues grâce à une analyse de leur contenu. Cependant, cette analyse est complexe et présente de nombreuses difficultés (Li *et al.*, 2002; Housden *et al.*, 2007; Laporte et Arbel, 2008). Dans cette optique, il n'est pas surprenant que d'autres solutions, très différentes, aient

été développées afin d'aboutir à des images échographiques en 3D (Prager *et al.*, 2010). Elles diffèrent entre autres en termes de matériel, de précision, de complexité des algorithmes.

Une de ces approches consiste à concevoir des sondes spécifiques pour l'échographie 3D (Smith *et al.*, 1991; Austeng et Holm, 2002). Les sondes échographiques, dont le fonctionnement est expliqué en détail au chapitre 1, sont constituées d'un réseau de transducteurs piézo-électriques. Dans le cas des sondes traditionnelles 2D, ces transducteurs sont positionnés le long d'une ligne (droite pour les sondes linéaires ou courbe pour les sondes convexes). La technique la plus courante de conception de sonde dédiée 3D consiste à placer une telle ligne de transducteurs dans un dispositif permettant de la déplacer mécaniquement dans un mouvement de balayage. Une autre technique consiste à organiser les transducteurs en quadrillage à deux dimensions, afin d'acquérir directement des volumes. La densité des transducteurs nécessaires pose cependant des problèmes importants au niveau électronique, et par ailleurs la surface de la sonde est rendue très large, ce qui rend le maintien permanent du contact de tous les transducteurs avec la peau difficile. De plus, pour les deux techniques décrites, le champ de vision offert reste limité (Prager *et al.*, 2010). Ce type d'approche présente des inconvénients qui nuisent à l'intérêt initial de l'échographie par rapport aux autres modalités. En effet, elles sont plus encombrantes, coûtent plus cher et nécessitent un investissement important pour les hôpitaux, qui disposent tous déjà d'équipement 2D. Par ailleurs, certains examens nécessitent des sondes spécialisées (forme, fréquence, etc.), et la conception de sondes 3D adaptées à toutes les applications semble peu réaliste.

Pour ces raisons, certaines études s'orientent plutôt vers une construction de volumes à partir d'images acquises avec une sonde conventionnelle 2D, déplacée autour de la zone d'intérêt. Il s'agit d'acquérir des tranches de l'objet, puis de reconstruire le volume par interpolation (Solberg *et al.*, 2007). La reconstruction requiert la connaissance de la position relative des différentes images acquises. Une solution pour cela consiste à fixer la sonde sur un bras mécanique dont le déplacement est programmé, et donc parfaitement connu (Janvier *et al.*, 2008). Cependant ce type de dispositif rend le processus d'acquisition peu pratique et est peu adapté à l'imagerie d'objets irréguliers, non rigides et mobiles comme le corps humain.

L'*échographie 3D main libre*, qui est l'approche abordée dans ce mémoire, consiste à acquérir des tranches de l'objet en déplaçant la sonde manuellement, et à déterminer leurs positions relatives afin d'effectuer la reconstruction. Un ensemble de méthodes d'échographie 3D main libre s'appuie sur un capteur de position fixé sur la sonde, permettant d'enregistrer sa position absolue associée à chaque acquisition. Les deux principales méthodes développées pour cela utilisent soit des capteurs optiques (par exemple Blackall *et al.* (2005)), soit des capteurs électromagnétiques (par exemple Cool *et al.* (2008)). Cependant, ces capteurs ajoutent un encombrement à la sonde, et imposent des restrictions supplémentaires. Dans le cas d'un capteur optique, il est indispensable que la sonde reste en tout temps visible par une caméra, ce qui est extrêmement contraignant en pratique dans le cadre d'une salle d'opération où les médecins ont besoin de pouvoir se déplacer librement. Le capteur optique est pourtant reconnu comme ayant des résultats plus fiables (Birkfellner *et al.*, 1998). Le capteur électromagnétique est, quant à lui, légèrement moins précis et, s'il ne pose pas de contrainte géométrique, implique la prise en compte de la présence de métaux dans la salle et des distorsions générées par ceux-ci (Lang *et al.*, 2009). De plus, dans tous les cas, l'utilisation d'un capteur implique un calibrage préalable (Lindseth *et al.*, 2003; Mercier *et al.*, 2005), qui demande du temps et un savoir-faire, donc une formation supplémentaire pour les praticiens.

La méthode d'échographie 3D main libre qui fait l'objet de ce mémoire, qui pourrait à terme être utilisée en combinaison avec une approche par capteur, consiste à utiliser le contenu des images pour estimer leurs positions relatives. Ce concept n'est pas intuitif car pour reconstruire un volume, la technique usuelle consiste à acquérir des coupes approximativement parallèles de l'objet, et c'est principalement la distance entre ces coupes qui doit être déterminée par l'analyse du contenu des images. Avec la plupart des techniques d'imagerie, ceci est difficile, voire impossible sans connaissance préalable de l'anatomie imagée. Cependant, les signaux échographiques présentent des particularités, détaillées au chapitre 1, qui ont pour effet l'existence de structures corrélées pour des tranches suffisamment proches. L'importance de la corrélation est directement liée à la distance séparant les tranches, et cette relation peut donc être utilisée pour déterminer leurs positions relatives.

Ce concept, développé et exploité dans de multiples études (Chen *et al.*, 1997; Tuthill *et al.*, 1998; Prager *et al.*, 2003; Housden *et al.*, 2007; Laporte et Arbel, 2011), fonctionne très bien lorsque les tranches sont parfaitement parallèles et alignées. Cependant, la problématique abordée dans ce mémoire émane du constat suivant : la sonde étant déplacée librement par le praticien, l'hypothèse d'un mouvement parfaitement rectiligne est irréaliste. Les tremblements probables de la main lors de l'acquisition d'une séquence ont pour conséquence le désalignement des plans, ainsi que leurs rotations relatives. Le déplacement de la sonde résultant comporte ainsi six degrés de liberté : deux translations et une rotation dans le plan des images, et une translation perpendiculaire à ce plan (nommée translation élévationnelle, il s'agit de la direction principale et intentionnelle), et deux rotations hors plan. Le problème primaire qui en découle est la nécessité, afin d'obtenir un positionnement correct, d'estimer les valeurs de toutes les composantes du mouvement, et pas seulement de la translation élévationnelle. La solution proposée par Housden *et al.* (2007) consiste à utiliser une méthode conventionnelle de recalage d'images pour estimer le mouvement dans le plan des images, puis de combiner cette estimation à celle de la translation élévationnelle. Le problème secondaire est le suivant : les composantes du mouvement n'ont pas des effets indépendants, et la qualité de l'estimation de l'une est nécessairement affectée par la présence de l'autre. Dans le cas de la translation élévationnelle, la relation entre la corrélation et la distance séparant les tranches est basée sur l'hypothèse que la décorrélation est uniquement due à l'augmentation de la distance. Or chaque composante du mouvement constitue une source de décorrélation. La présence des autres composantes fausse ainsi les estimations de distance élévationnelle. De manière similaire, la composante de translation élévationnelle fausse les résultats des méthodes d'estimation des autres composantes. La solution proposée, qui constitue la contribution principale de ce mémoire, est basée sur l'hypothèse que les erreurs d'estimation sont en grande partie systématiques : une même combinaison de composantes de mouvement de la sonde produit des erreurs similaires d'estimation des composantes. Nous proposons alors d'utiliser une approche d'apprentissage automatique pour construire un modèle de la régression liant les estimations faussées aux valeurs réelles des composantes du mouvement réel correspondant. Ceci implique de disposer d'un très grand nombre d'exemples de séquences pour lesquelles le mouvement réel est connu (éventuellement des séquences obtenues par simulation), afin de pouvoir construire une base

de données d'apprentissage constituée des estimations et des valeurs réelles. Le modèle de régression pourra ainsi être utilisé pour corriger de nouvelles estimations pour des séquences de mouvement inconnu.

La méthodologie proposée, détaillée sur le schéma de la figure 0.1 est ainsi composée de deux étapes :

- a. estimer les composantes du mouvement par les méthodes conventionnelles, et en déduire une estimation du mouvement global. Cette estimation globale est nommée «estimation naïve» au cours de ce mémoire, puisqu'il est attendu qu'elle comporte des erreurs. Cette étape, et les erreurs qu'elle implique, sont décrites au chapitre 2 ;
- b. corriger les estimations grâce au modèle de correction, préalablement construit par apprentissage statistique.

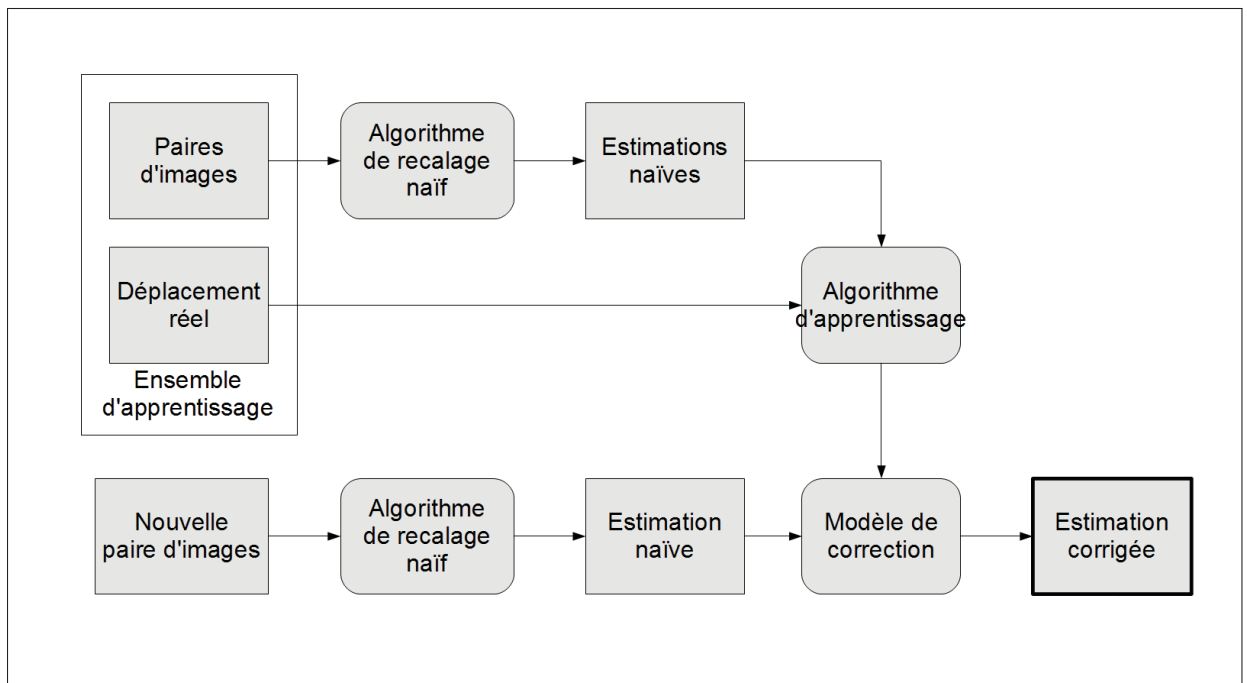


Figure 0.1 Méthodologie proposée

La construction préalable du modèle de correction est réalisée selon les étapes suivantes :

- a. obtenir un grand nombre de séquences synthétiques de mouvement connu ;
- b. obtenir des estimés naïfs du mouvement de ces séquences pour construire la base de données d'apprentissage ;
- c. construire le modèle de correction des estimations grâce à un apprentissage par machine à vecteurs de pertinence. Cette étape fait l'objet du chapitre 3.

Les autres approches existantes pour résoudre le problème des erreurs d'estimation sont toutes basées sur des données externes obtenues grâce à un capteur (Housden *et al.*, 2008; Lang *et al.*, 2009), ou grâce à des informations provenant d'autres modalités (Lang *et al.*, 2012). À la différence de ces solutions, l'approche proposée s'appuie uniquement sur le contenu des images, et est novatrice pour cette raison.

Ce mémoire est structuré comme suit. Tout d'abord, la revue de la littérature et des connaissances est présentée dans le chapitre 1. Celui-ci rassemble les connaissances nécessaires à la compréhension de la méthode proposée, notamment concernant le fonctionnement de l'échographie et les statistiques des signaux engendrés, ainsi qu'une revue de la littérature pertinente, présentant les travaux existants à ce jour entrant dans le cadre de la problématique. Ensuite, le chapitre 2 concerne la première partie de la méthodologie proposée. Chaque étape de l'algorithme naïf y est expliquée, puis des expérimentations sont menées afin d'étudier son fonctionnement. Les résultats obtenus permettent de vérifier les attentes théoriques : en présence de mouvement complexe, des erreurs systématiques apparaissent. Le chapitre 3 présente la seconde partie de la méthodologie : la correction des erreurs par apprentissage. Le modèle de correction est construit grâce à deux implémentations des machines à vecteurs de pertinence : l'une à sortie monovariée, l'autre à sortie multivariée. Ces deux méthodes sont d'abord présentées, puis une étude expérimentale est réalisée, qui permet de démontrer de bonnes performances de l'algorithme proposé dans la plupart des situations, tout en notant que certaines restent problématiques. Finalement, la conclusion de ce mémoire présente un résumé et une discussion des perspectives futures.

CHAPITRE 1

PRINCIPES FONDAMENTAUX ET REVUE DE LITTÉRATURE

Ce chapitre a pour but d'introduire les notions nécessaires à la compréhension de notre démarche, et de présenter les solutions existantes à ce jour pour réaliser l'échographie 3D main libre sans capteur de position, afin de préciser le contexte de notre étude. Tout d'abord, la partie 1.1 présente le fonctionnement de l'échographie, et les particularités du signal échographique exploitées pour le recalage des images permettant la reconstruction du volume échographique. Les différentes méthodes de recalage existantes dans la littérature, ainsi que les difficultés qu'elles présentent, sont ensuite présentées dans la partie 1.2. Finalement, la partie 1.3 définit précisément le problème abordé dans ce mémoire et sa formalisation.

1.1 Les signaux échographiques

L'échographie produit des images très particulières par rapport aux autres modalités d'imagerie. En effet, les interférences entre les signaux échographiques produisent des structures spécifiques dans le plan image, comportant notamment une structure particulière nommée *speckle*, apparaissant sur les images comme une texture granuleuse, caractéristique des images échographiques (figure 1.1).

Ces structures peuvent nuire à l'interprétation visuelle, mais sont un atout pour l'application considérée ici car elles peuvent aider à l'identification des positions relatives. Il est nécessaire d'étudier l'origine de ces structures et la nature des signaux produits afin de comprendre les méthodes présentées par la suite, qui exploitent pour la plupart les particularités de cette modalité.

1.1.1 Fonctionnement de l'échographie

Lors de l'acquisition d'images échographiques, des impulsions sonores de haute fréquence (de 1 à 10 MHz) sont produites par des éléments piézoélectriques qui vibrent à leur fréquence de résonance, excités par un potentiel électrique. Ces éléments (nommés transducteurs) sont

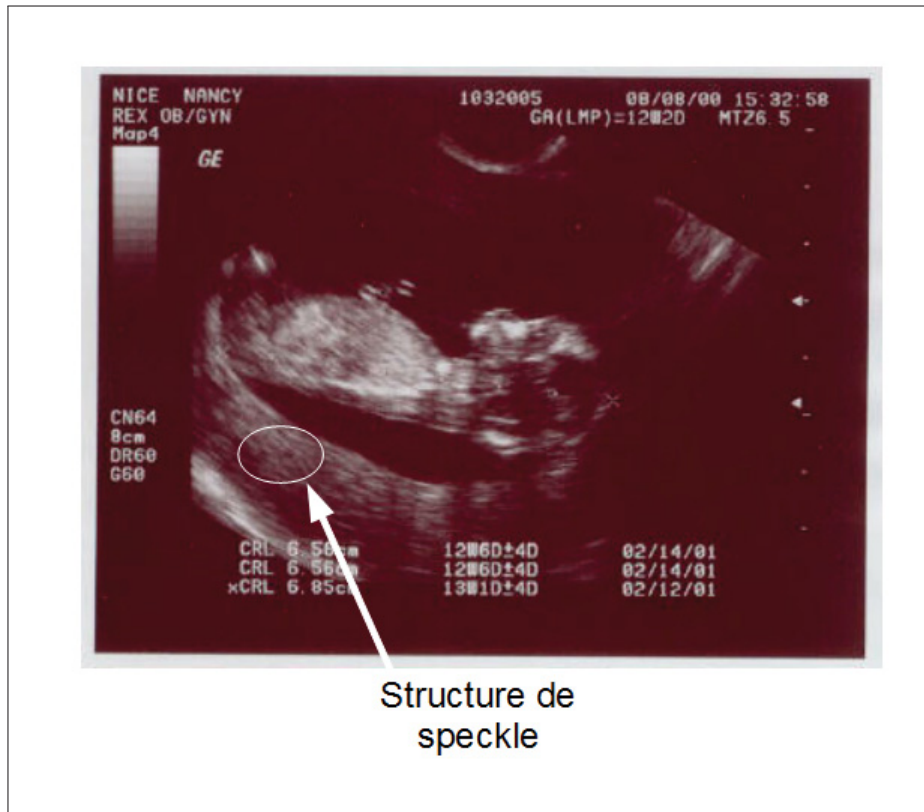


Figure 1.1 Image échographique avec speckle
Tirée de Karim et Nancy Nice (2000)

alignés sur une sonde, qui transmet ces impulsions ultrasonores à l'intérieur du corps du patient (figure 1.2). Afin d'améliorer la résolution latérale, une focalisation électronique est réalisée : un jeu de retards symétriques est appliqué aux impulsions électriques afin de créer un front d'onde de forme concave convergeant vers la zone focale.

Les ondes sonores voyagent alors dans le milieu et finissent par rencontrer une interface entre des tissus. L'impédance acoustique, qui représente la difficulté qu'a le signal à se propager dans un tissu donné et à causer des déplacements de particules, est caractéristique du tissu. Ainsi, lorsque les ondes atteignent une telle interface, la variation d'impédance provoque la réflexion d'une partie des ondes en direction de la sonde, alors que le reste est transmis plus loin jusqu'à rencontrer une autre interface et se réfléchir, ou jusqu'à absorption totale dans le milieu (figure 1.3).

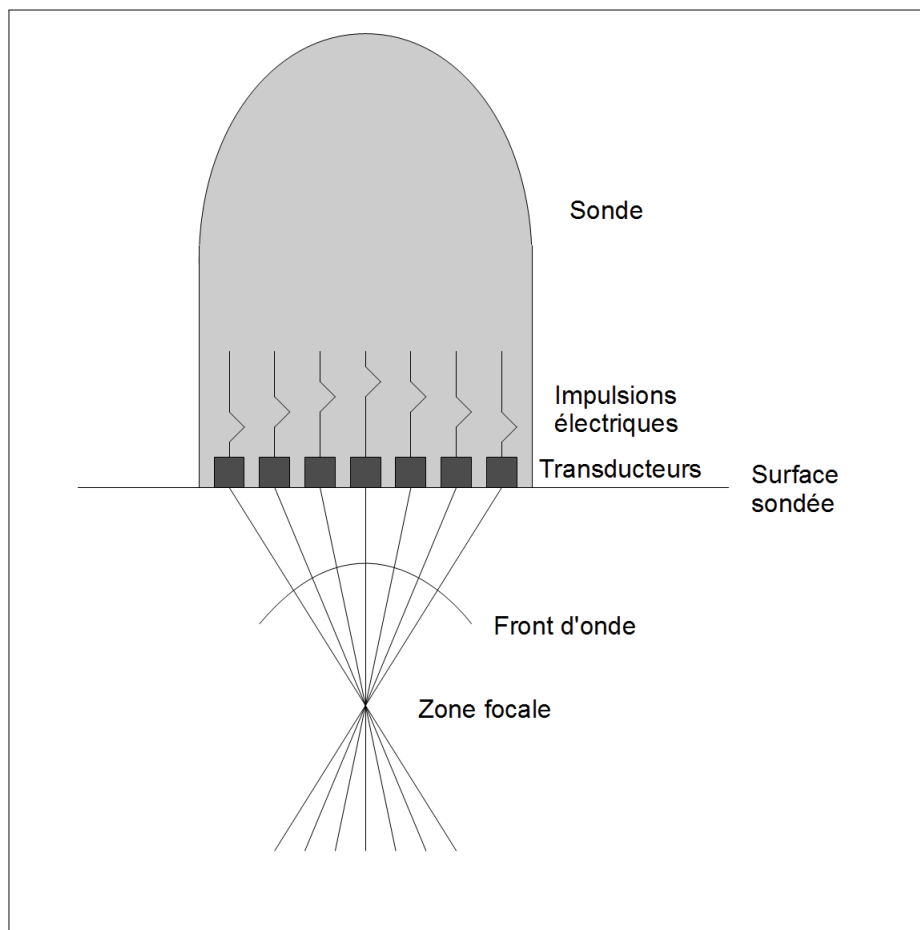


Figure 1.2 Organisation des transducteurs et focalisation

Les ondes réfléchies sont ensuite reçues par la sonde et transformées en signaux électriques par le processus inverse de leur émission. Ainsi, on peut calculer la distance entre la sonde et le tissu ou l'organe (interface) en utilisant la vitesse du son dans les tissus biologiques (environ 1540 m.s^{-1}), qui est admise constante dans le milieu, et le temps entre l'émission d'une impulsion et la réception de son écho. Pour cela, ce signal électrique est transmis à l'ordinateur de l'échographe pour être traité. On parle alors de signal RF, à cause de la similitude avec les ondes radio-fréquence : il s'agit d'une onde porteuse à la fréquence des impulsions, modulée par un signal d'amplitude représentant l'importance de l'écho. Un tel signal RF est associé à chaque élément piézoélectrique, indiquant la position et l'importance des interfaces rencontrées par les ondes transmises par cet élément.

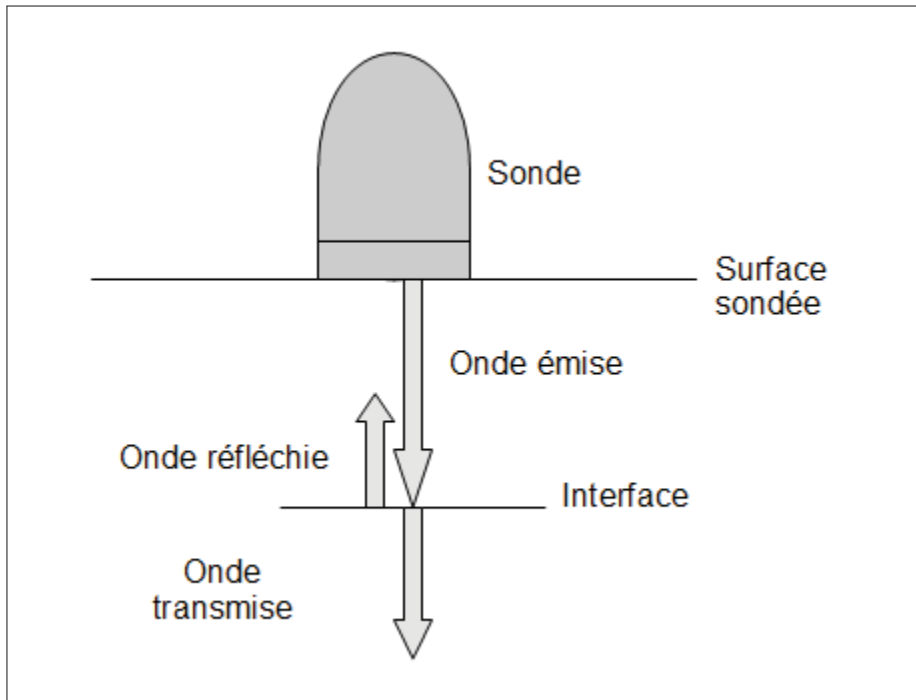


Figure 1.3 Réflexion et transmission des ondes ultrasonores

On procède ensuite, pour chacun des signaux RF, à une série d'opérations permettant de visualiser les signaux sous forme d'image. Tout d'abord, une détection d'enveloppe est réalisée afin d'obtenir des signaux d'amplitude : chaque signal d'amplitude représente l'échogénicité (i.e. la capacité à rétrodiffuser les ondes ultrasonores) en fonction de la profondeur dans la structure imagée, détectée par un élément piézoélectrique. Puis une compression logarithmique de la plage dynamique du signal, ainsi que diverses autres opérations particulières au matériel employé, sont réalisées pour obtenir une meilleure visualisation. Chaque signal représente alors une ligne dans le plan image, et la brillance de chaque point de la ligne est fonction de l'échogénicité correspondante en ce point du signal. Leur combinaison permet de former les images dites de mode B, qui représentent ainsi une vue en coupe de la structure.

Il est important de noter que les appareils d'échographie effectuent automatiquement la conversion en mode B, et que l'accès aux signaux RF est généralement restreint et disponible à l'utilisateur uniquement via certaines interfaces de recherche très coûteuses fournies seulement par quelques fabricants.

1.1.2 Le speckle

Comme rappelé dans la partie 1.1.1, les ondes ultrasonores sont en partie réfléchies lorsqu'elles rencontrent un changement de milieu. Cependant, les tissus sont rarement des milieux parfaitement homogènes : ils contiennent des particules de densités différentes, qui créent des inhomogénéités acoustiques. Lorsque ces particules sont de taille inférieure à la longueur d'onde (il s'agit alors de micro-particules), l'onde est réfléchiée par chacune d'entre elles dans toutes les directions : ces particules sont des diffuseurs. Des interférences constructives et destructives apparaissent alors entre les multiples ondes rétrodiffusées (figure 1.4), créant ainsi des structures particulières dans le signal : le speckle.

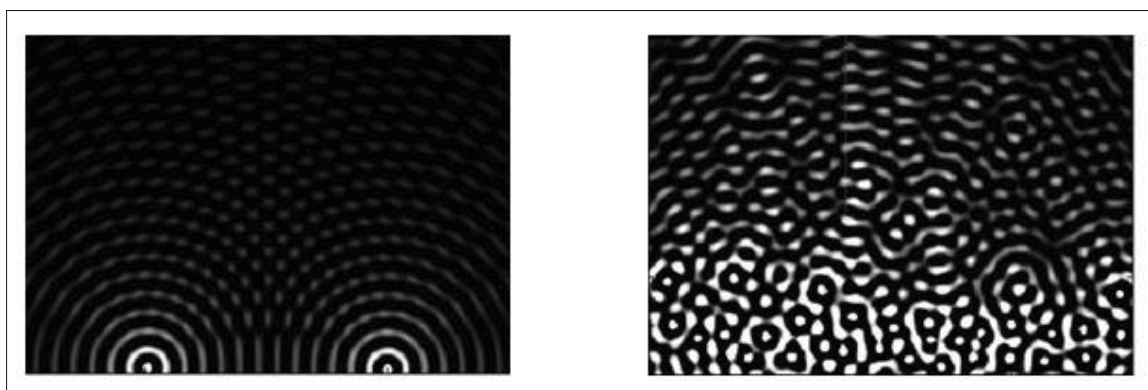


Figure 1.4 Interférences dues à deux particules (gauche)
et à de multiples particules (droite)
Tirées de Støylen (2010)

Le speckle n'est pas représentatif des structures macroscopiques généralement intéressantes pour le clinicien, et est donc souvent considéré comme un bruit. En effet, il induit une diminution importante de la résolution de contraste, qui rend les images échographiques très difficiles à « lire » par rapport, par exemple, aux images obtenues par IRM. De nombreux chercheurs se sont donc dédiés à la recherche d'un moyen d'éliminer ce « bruit » (Evans et Nixon, 1995; Coupé *et al.*, 2009; Tay *et al.*, 2010).

Pourtant, si le speckle rend l'interprétation des images difficile pour les non-initiés, il ne constitue pas réellement un bruit. En effet, les structures ajoutées à l'image ne sont pas aléatoires,

mais peuvent être reproduites pourvu que l'on image le sujet toujours de la même manière (en termes de sonde, de position...). Les statistiques du speckle dépendent de paramètres liés à la structure des tissus (densité, taille, distribution des diffuseurs), et de la sonde employée. Par conséquent, la texture du speckle varie en fonction des tissus. Ainsi, plutôt que de tenter de s'en débarrasser, il peut être intéressant de l'exploiter comme une source d'information à part entière. Pour cela, il est bien nécessaire de prendre en compte que le speckle n'est pas une représentation directe de la configuration spatiale des particules, puisqu'il s'agit du résultat des interférences dues à celles-ci. Il existe cependant un lien statistique entre la configuration des particules et la structure résultante du speckle. Il peut donc être utilisé pour reconnaître un même objet sur deux images différentes, acquises avec une même sonde.

Ainsi, le speckle est souvent utilisé pour suivre le mouvement des tissus, en observant des images d'une même zone à des instants différents (Bohs et Trahey, 1991; Yu *et al.*, 2006). De manière similaire, il peut être utilisé pour suivre le mouvement de la sonde sur un même objet immobile. C'est cette application dont il sera question dans ce mémoire. Les méthodes existant pour y parvenir seront présentées dans la section 1.2. Une analyse statistique du phénomène de formation du speckle fait l'objet de la section suivante.

1.1.3 Modélisation et statistiques des signaux échographiques

Rappelons tout d'abord que la réponse impulsionnelle d'un système, (ou PSF, pour *Point Spread Function*), est la réponse de ce système lorsque l'entrée est une impulsion. Dans le cas de l'échographie, la PSF représente donc la répartition spatiale de l'intensité acoustique lorsque la structure imagée est un réflecteur ponctuel idéal, et son enveloppe peut être modélisée par une fonction gaussienne (Anderson et Trahey, 2000).

La formation des images par ultrasons peut être modélisée par un système linéaire (Ng *et al.*, 2006). Si l'on considère des régions de petite taille dans le plan image, on peut poser l'hypothèse que la PSF est indépendante de la position. Le signal RF peut alors être exprimé comme la convolution de la PSF par la fonction des diffuseurs, qui décrit l'échogénicité des diffuseurs

en fonction de leur position.

$$U(x, y, z) = g(x, y, z) \otimes t(x, y, z), \quad (1.1)$$

avec :

$U(x, y, z)$ le signal RF,

$g(x, y, z)$ la PSF du système,

$t(x, y, z)$ la fonction des diffuseurs,

$\otimes$ le produit de convolution.

L'enveloppe de la PSF correspond à ce que l'on appelle la cellule de résolution. Celle-ci décrit le volume fini auquel correspond chaque pixel dans le plan image, étant donné que le faisceau ultrasonore n'est pas parfaitement ponctuel mais possède une taille finie. Les cellules sont allongées dans le sens de la direction élévationnelle, ce qui implique une superposition importante des cellules sur des plans avoisinants (figure 1.5). Il y a donc une corrélation entre les signaux des plans suffisamment proches, qui dépend du degré de superposition et qui peut donc être utilisée pour estimer la séparation entre deux plans (Chen *et al.*, 1997; Tuthill *et al.*, 1998; Prager *et al.*, 2003; Housden *et al.*, 2007; Laporte et Arbel, 2011).

Dans le but de modéliser avec précision le processus complexe de formation des images échographiques, les statistiques du speckle ont été étudiées en profondeur dans la littérature (Burckhardt, 1978; Wagner, 1983; Wear et Popp, 1987; Ng *et al.*, 2006). Afin de modéliser la superposition des cellules de résolution, nous nous intéressons ici aux statistiques de second ordre. Celles-ci permettent de décrire et de quantifier les relations entre pixels voisins, par le biais de la fonction d'autocorrélation du signal.

Dans le cas où les diffuseurs sont répartis aléatoirement et qu'ils sont présents en très grand nombre, l'enveloppe du signal est distribuée selon le modèle de Rayleigh (Wagner, 1983). Dans ces conditions, dénommées *conditions de Rayleigh*, le speckle est dit « pleinement développé ». D'autre part, les facteurs physiques affectant la formation des images peuvent être distingués

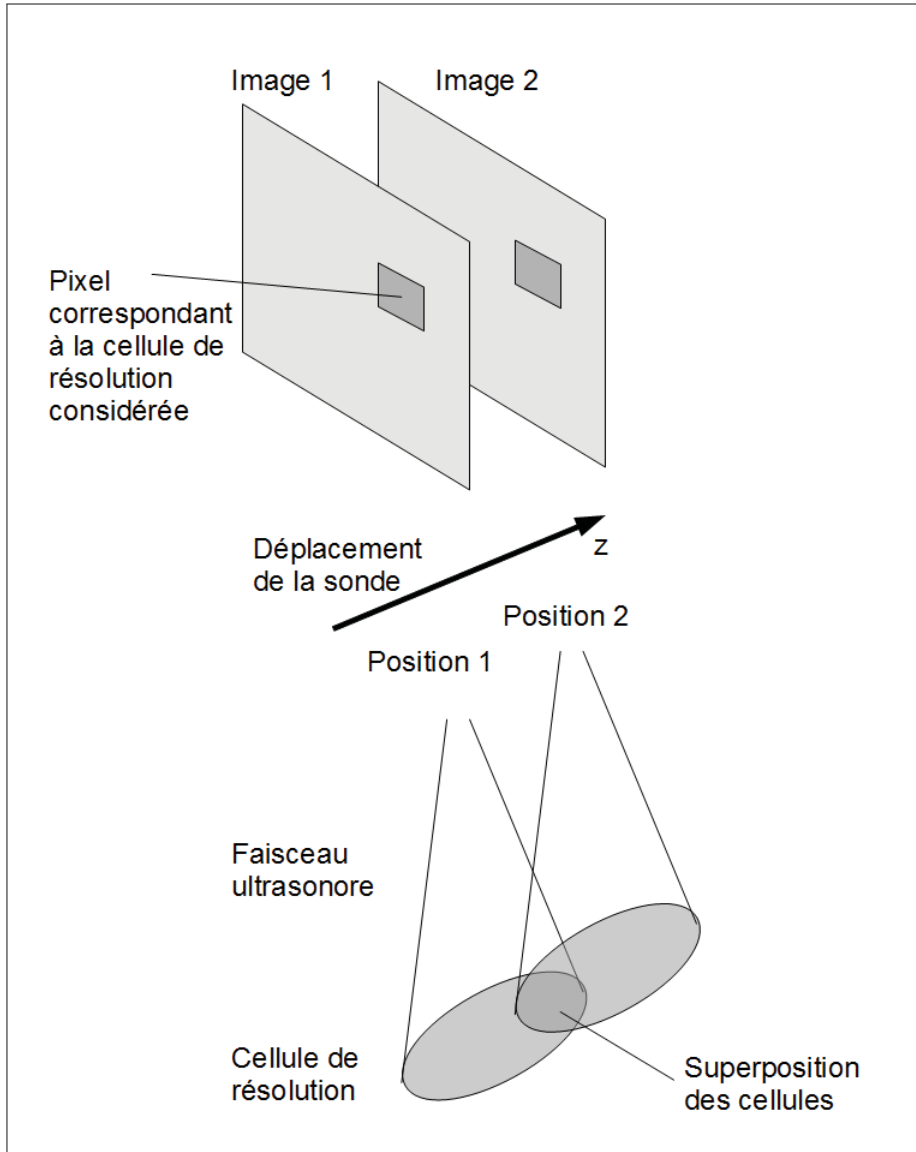


Figure 1.5 La superposition des cellules de résolution lors de l'acquisition de deux images consécutives conduit à une corrélation entre ces images

selon chaque direction, ce qui permet de supposer la PSF séparable (Wagner, 1983) :

$$g(x, y, z) = g_x(x)g_y(y)g_z(z).$$

Le problème peut donc être ramené à une seule dimension, notée r (désignant x , y ou z). La fonction d'autocorrélation du signal s'exprime ainsi :

$$R_U(r, r + \Delta r) = E[U(r)U(r + \Delta r)],$$

avec : $R_U(r, r + \Delta r)$ la fonction d'autocorrélation du signal aux positions r et $r + \Delta r$, et $E[\cdot]$ l'espérance mathématique. D'après l'expression du signal RF (1.1), on peut alors développer de la manière suivante :

$$R_U(r, r + \Delta r) = R_t(r, r + \Delta r) \otimes g(r) \otimes g(r + \Delta r),$$

avec $R_t(r, r + \Delta r) = E[t(r)t(r + \Delta r)]$, la fonction d'autocorrélation de la fonction des diffuseurs t aux positions r et $r + \Delta r$.

Considérant que la force des diffuseurs est peu variante au niveau macroscopique par comparaison au niveau microscopique, on peut écrire :

$$R_t(r, r + \Delta r) = |t(r)|^2 \mu_t(\Delta r),$$

où μ_t représente la fonction d'autocorrélation de la micro-structure. Si la distribution des diffuseurs est considérée homogène, $t(r)$ peut être considérée constante, notée t_0 . Par ailleurs on peut montrer que la fonction d'autocorrélation du speckle est un processus localement stationnaire, c'est-à-dire que $R_U(r, r + \Delta r) = R_U(\Delta r)$, pour de petits Δr (Ng *et al.*, 2006). Après simplifications, il vient alors :

$$R_U(\Delta r) = t_0^2 g(-\Delta r) \otimes \mu_t(\Delta r) \otimes g(\Delta r).$$

La fonction d'autocorrélation du speckle n'est donc pas fonction de la position, mais seulement de la différence des positions, soit du déplacement.

Finalement, dans les conditions de Rayleigh, la microstructure n'est pas corrélée, et μ_t peut être décrite comme une impulsion de Dirac : $\mu_t(\Delta r) = \delta(\Delta r)$ (Wagner, 1983). Alors,

$$R_U(\Delta r) = t_0^2 g(-\Delta r) \otimes g(\Delta r).$$

Ainsi, la fonction d'autocorrélation du speckle ne dépend que de la PSF du système.

La corrélation normalisée est définie comme suit :

$$\rho(\Delta r) = \frac{R_U(\Delta r)}{R_U(0)}.$$

La PSF étant considérée approximativement gaussienne (et séparable), les fonctions de corrélation normalisée dans les trois directions (axiale, latérale et élévationnelle) peuvent être modélisées par des fonctions Gaussiennes dépendant uniquement des caractéristiques du système (Wagner, 1983).

On peut ainsi définir la notion de courbe de décorrélation, lorsque les déplacements axial et latéral sont nuls : elle correspond à la fonction de corrélation normalisée élévationnelle, et dépend uniquement du déplacement élévationnel et des caractéristiques du système (soit des caractéristiques de la sonde employée). Cette spécificité de l'échographie peut ainsi être utilisée pour retrouver le déplacement élévationnel. Un exemple de courbe de décorrélation est présenté sur la figure 1.6. La démarche permettant d'obtenir une telle courbe est décrite dans la partie 1.2.1.

1.2 Recalage

Les particularités du speckle, et notamment les statistiques du second ordre qui lui sont associées, décrites à la partie 1.1.3, en font un phénomène très utile lorsqu'il s'agit de déterminer le déplacement hors plan. Si les conditions de Rayleigh sont respectées et que le déplacement est une translation élévationnelle pure, il existe en effet une correspondance directe entre le coefficient de corrélation et la distance élévationnelle séparant deux images. Cette correspondance

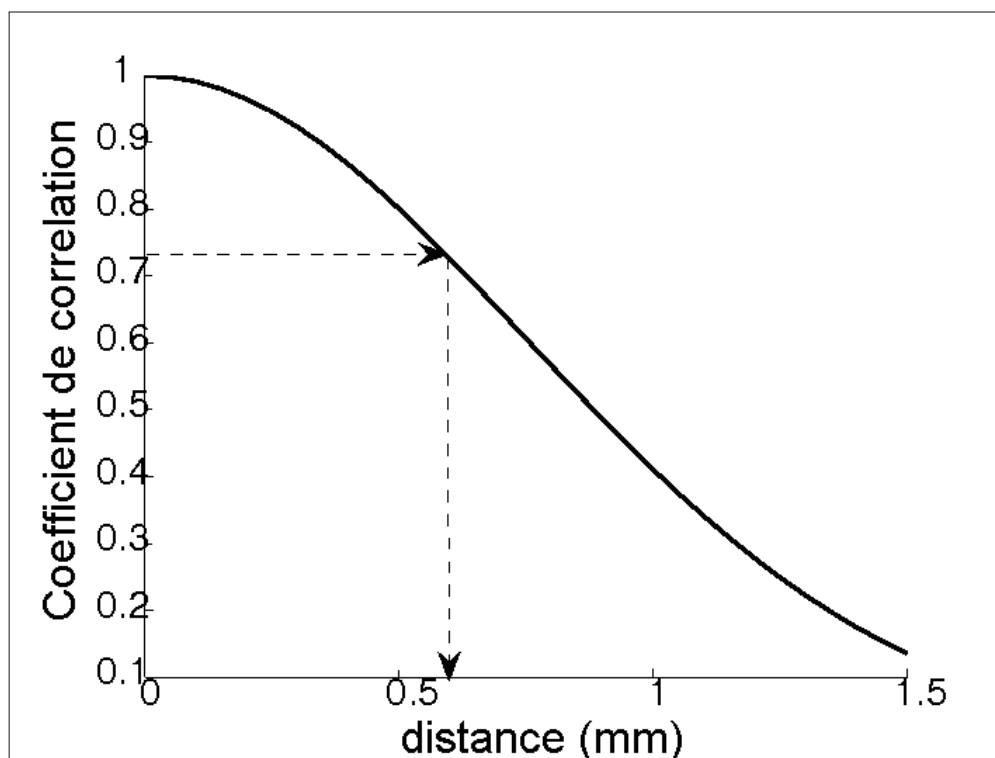


Figure 1.6 Exemple de courbe de décorrélation : coefficient de corrélation en fonction de la distance élévationnelle
Adaptée de Laporte (2011)

est décrite par la courbe de décorrélation, représentant le coefficient de corrélation en fonction de la distance de séparation entre les plans.

Cependant, la méthode étudiée ici considère un mouvement non contraint de la sonde, ce qui implique qu'un déplacement consistant uniquement d'une translation élévationnelle pure est peu probable : d'autres composantes sont présentes, et il est nécessaire de les déterminer également. Il s'agit alors d'un problème de *recalage*. En traitement d'image, le recalage consiste à déterminer l'alignement spatial entre plusieurs images représentant le même objet. En imagerie médicale, il peut s'agir d'images provenant d'un seul sujet à des temps différents, par exemple avant et après un traitement, ou à des localisations légèrement différentes, ou bien d'une même région de plusieurs sujets à des fins de comparaison. Ces images peuvent également être acquises avec différentes modalités d'imagerie (on parle alors de recalage multimodal), et il peut s'agir de recalage 2D-2D, 2D-3D ou encore 3D-3D. Par ailleurs les transformations à déter-

miner peuvent être rigides (six degrés de liberté), ou non rigides (dus au mouvement physiologique, au mouvement du patient, aux distorsions causées par les appareils, etc.). Il s'agit d'un problème central dans le domaine de l'imagerie médicale, et diverses méthodes ont été développées afin de traiter les différentes situations (Hill *et al.*, 2001).

Ici, l'objectif est de recalcr des images en deux dimensions, provenant d'un même sujet, lorsque la sonde subit un déplacement en trois dimensions. La sonde subit une transformation rigide, et l'on cherche en fait à estimer cette transformation en analysant son effet sur les images. Nous considérons ici qu'il n'y a pas de mouvement physiologique ni de variation de la compression des tissus, donc le seul mouvement correspond à la transformation rigide de la sonde. Il s'agit d'une transformation composée de six degrés de liberté dans l'espace : trois translations et trois rotations, plus précisément définies dans la partie 1.3.

Ce problème complexe de recalage est communément abordé en séparant le mouvement de la sonde en deux parties (Housden *et al.*, 2007) : d'une part, le mouvement hors plan, composé d'une translation hors plan (dans notre cadre de travail, ce mouvement est considéré comme le mouvement principal appliqué à la sonde) et deux rotations hors plan (figure 1.7 (1)), et d'autre part le mouvement dans le plan, composé de deux translations (axiale et latérale) et d'une rotation (figure 1.7 (2)). Notons que le recalage dans le plan correspond en fait à la notion conventionnelle de recalage en imagerie, tandis que le recalage hors plan est plus particulier à l'échographie. La littérature abordant le recalage de ces deux types de mouvement sont abordées dans les parties 1.2.1 et 1.2.2.

1.2.1 Recalage hors plan et échographie 3D main libre sans capteur de position

La courbe de décorrélation (figure 1.6), calibrée pour la sonde employée, permet de réaliser l'estimation de la distance élévationnelle entre deux images échographiques. Ce concept, développé relativement récemment, a été exploité par plusieurs équipes de chercheurs afin de rendre possible l'échographie 3D main libre sans capteur de position. Tout d'abord, Chen *et al.* (1997) ont réalisé une étude théorique démontrant qu'il était possible d'approximer la fonction d'autocorrélation normalisée du signal d'intensité dans la zone focale par une gaussienne. Ils ont ensuite utilisé ce concept pour estimer la distance élévationnelle entre des images de

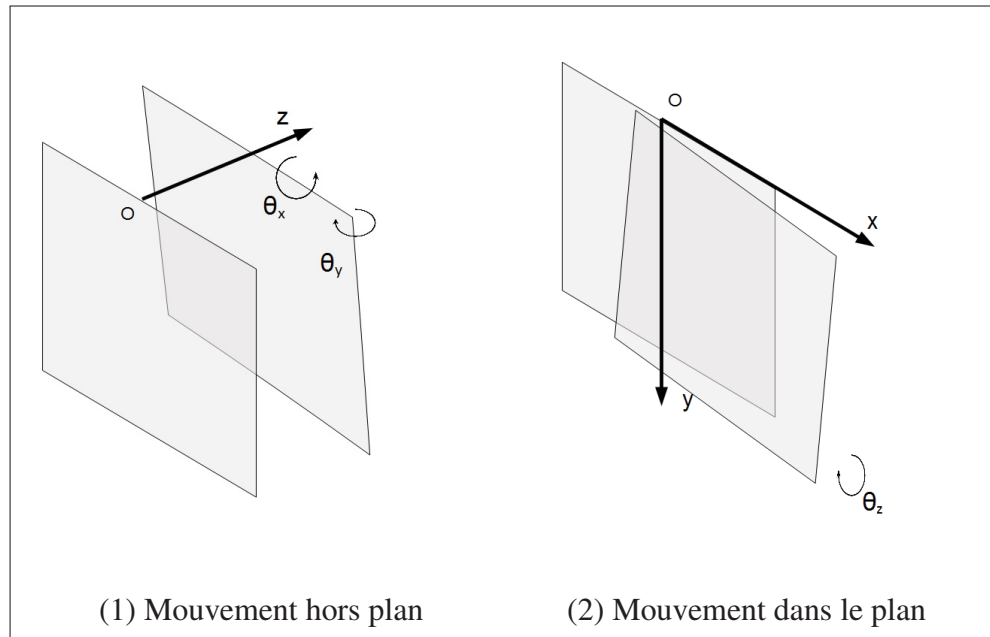


Figure 1.7 Description des différents types de mouvement

mode B dans le cas de speckle pleinement développé et pour un mouvement purement élévational. Leur méthode pour l'estimation de la distance entre deux images consistait à acquérir 10 images et à mesurer la corrélation dans la zone focale entre les images, puis à ajuster une gaussienne aux valeurs obtenues, et ainsi déterminer la distance entre les deux images centrales de l'ensemble des 10 images.

Par la suite, Tuthill *et al.* (1998) se sont basés sur cette approche et l'ont développée en adaptant le modèle de décorrélation en fonction de la profondeur, pour permettre l'utilisation de mesures hors zone focale afin d'estimer également les rotations hors plan grâce à des mesures de corrélation en plusieurs endroits des images. Par ailleurs, ils ont également élaboré une méthode permettant de localiser automatiquement les zones de speckle pleinement développé dans les images. Le modèle ne tient en effet que dans les conditions de Rayleigh, qui induisent un speckle pleinement développé. Or, dans l'imagerie réelle il existe des zones de speckle partiellement développé où les estimations seraient incorrectes. Ainsi, la localisation automatique des zones de speckle pleinement développé permet de choisir les meilleures régions pour effectuer les mesures.

Chang *et al.* (2003) ont proposé une méthode alternative d'estimation des distances ne nécessitant pas de calibrer la courbe de décorrélation, en calculant uniquement des rapports de distance entre chaque image de la séquence, puis en utilisant une image perpendiculaire à toutes les images de la séquence comme référence pour obtenir les valeurs d'estimation. Cette méthode utilise un détecteur de speckle tel que celui proposé par Tuthill *et al.* (1998). Elle a l'inconvénient d'imposer que toutes les images de la séquence soient parallèles, et que l'image de référence soit parfaitement perpendiculaire à la séquence. Le processus d'acquisition de la séquence est rendu difficile, et ne permet pas suffisamment de liberté de mouvement. Cette méthode n'a donc pas été poursuivie.

Revenant à l'approche de Chen *et al.* (1997) et Tuthill *et al.* (1998) avec calibrage de la courbe, Prager *et al.* (2003) se sont intéressés à développer encore la liberté du mouvement entre les images en proposant d'estimer également le mouvement dans le plan grâce aux méthodes conventionnelles de recalage dans le plan (présentées dans la partie 1.2.2). Une analyse principalement qualitative des résultats expérimentaux ainsi obtenus a permis de confirmer la validité de l'approche et d'encourager la poursuite des recherches dans cette direction.

Gee *et al.* (2006) ont ensuite travaillé sur la problématique du degré de développement du speckle. Les zones de speckle pleinement développées sont en réalité rares et peu étendues dans les tissus biologiques réels (ceci a été démontré par Hassenpflug *et al.* (2005)), et la sélection de suffisamment de zones appropriées, comme proposé par Tuthill *et al.* (1998), n'est donc pas toujours réalisable. Les tissus réels contiennent en effet des zones de diffusion cohérente, qui sont à l'origine de speckle partiellement développé dont la décorrélation est plus lente. Gee *et al.* (2006) ont donc proposé d'adapter la courbe de décorrélation élévationnelle, préalablement calibrée pour du speckle pleinement développé, à chaque zone considérée, en calculant un facteur heuristique de diffusion cohérente. Cette méthode a permis d'obtenir des estimations fiables du mouvement hors plan. Plus tard, Laporte et Arbel (2011) ont également travaillé sur le problème de speckle partiellement développé, en proposant une méthode basée sur les statistiques du speckle pour adapter les estimations hors plan par apprentissage, obtenant ainsi des

résultats très satisfaisants sur des séquences d'images de mode B avec mouvement élévational pur.

Housden *et al.* (2007) ont abordé le problème des mouvements non monotones : la mesure de corrélation donne une indication de distance absolue entre deux images, mais ne permet pas de déterminer s'il y a un changement de sens dans une même séquence, ou si les images s'intersectent (en présence de rotations hors plan). Pour répondre à ce problème, Housden et al. ont élaboré une méthode imposant une contrainte de planarité des images permettant de détecter les intersections, et estimant le déplacement sur des sous-ensembles de la séquence afin de repérer les changements de sens éventuels dans la séquence. Laporte et Arbel (2008) ont proposé une méthode alternative réalisant d'abord une analyse permettant la détection des mouvements non monotones sur différentes zones des images, avant d'imposer la contrainte de planarité. Cette méthode a permis d'obtenir des résultats plus précis sur des séquences purement élevationnelles. Housden *et al.* (2007) ont également rassemblé des solutions aux diverses difficultés de l'échographie 3D main libre pour proposer un système global de reconstruction, intégrant le recalage dans le plan (Prager *et al.*, 2003; Housden *et al.*, 2006), la détection de mouvements non monotones et l'adaptation de la courbe de décorrélation aux zones hors conditions de Rayleigh (Gee *et al.*, 2006). Les résultats ainsi obtenus sont intéressants, mais des erreurs ont été notées dans les estimations en présence de mouvement dans le plan et de rotations, malgré l'utilisation du recalage dans le plan. Ces erreurs seront plus précisément étudiées à la partie 1.2.3 de ce chapitre.

1.2.2 Recalage dans le plan

Dans le cadre de l'échographie 3D main libre, il est nécessaire d'estimer le mouvement dans le plan car celui-ci est également à l'origine de décorrélation, ce qui fausse les estimations de déplacement élévational car les causes sont confondues. Le recalage des images dans le plan permet d'atténuer la décorrélation causée par ce mouvement afin de retrouver une valeur de corrélation permettant une estimation du déplacement élévational plus proche de la réalité (Housden *et al.*, 2006). Le recalage dans le plan consiste spécifiquement à estimer les translations et la rotation subies par une image, dans le cadre d'une transformation rigide 2D.

Parmi les méthodes possibles, on distingue les méthodes de recalage par caractéristiques et les méthodes de recalage par intensité (Hill *et al.*, 2001). Le recalage par caractéristiques s'appuie sur des caractéristiques géométriques dans les images (points, lignes, surfaces). La reconnaissance de ces structures peut être réalisée par segmentation ou extraction de points saillants, de manière manuelle ou automatique. La segmentation automatique dans les images échographiques est très difficile à mettre en oeuvre, principalement à cause de la présence de speckle. Malgré les études visant le « de-speckling » citées dans la partie 1.1.2, le recalage par caractéristiques est moins utilisé en échographie par rapport aux modalités d'imagerie présentant des structures mieux définies (IRM, CT...), avec quelques exceptions comme les travaux de Leung *et al.* (2009). En effet, ce type d'approche est surtout utile pour des déplacements importants alors que dans notre application, les mouvements sont de très faible amplitude, et le recalage par caractéristiques ne fournit pas une précision assez bonne pour être fiable.

Le recalage par intensité utilise, quant à lui, une mesure de similitude entre les pixels des images, sans prendre en compte directement l'information géométrique. Ce type de méthode est par conséquent très bien adapté à l'imagerie échographique où les structures géométriques macroscopiques sont difficiles à reconnaître. De nombreuses mesures de similitude ont été développées à ce jour, supposant chacune une relation différente entre les intensités des pixels des deux images lorsque les images sont recalées. Parmi elles, les plus utilisées sont notamment la SAD (« sum of absolute differences », soit somme des différences absolues), qui suppose que les intensités des deux images sont les mêmes, ou des mesures issues de la théorie de l'information comme par exemple l'information mutuelle (relation statistique) ou encore le coefficient de corrélation de Pearson (Penney *et al.*, 1998). On trouve ensuite la transformation qui optimise le critère choisi.

Les méthodes de recalage des échographies basées sur l'intensité ont d'abord été développées pour des applications de suivi de mouvement physiologique, par exemple pour la mesure du débit sanguin. Pour cela, le *speckle tracking*, ou suivi de speckle, est réalisé grâce à des mesures de similitude permettant d'apparier des blocs d'image. La mesure de similitude la plus couramment utilisée pour ceci est la SAD (Bohs et Trahey, 1991; Kontogeorgakis *et al.*, 1994;

Yeung *et al.*, 1998), car celle-ci représente un temps de calcul moins élevé que la corrélation normalisée par exemple (Bohs *et al.*, 2000), et peut donc être calculée plus rapidement, ce qui permettait à l'époque des mesures en temps réel.

Une autre application est le recalage élastique d'échographies acquises à des instants différents (par exemple avant et pendant l'opération). La plupart des techniques sont basées sur le recalage par intensité (Krucker *et al.*, 2002), en particulier par mesure de corrélation (Xiao *et al.*, 2002; Chen *et al.*, 2003), mais nous pouvons noter que Foroughi et Abolmaesumi (2005) et Leung *et al.* (2009) ont proposé des méthodes de recalage par caractéristiques, où des attributs sont calculés pour chaque pixel, tels que l'intensité, le module du gradient et le Laplacien de gaussienne, pour ensuite effectuer une reconnaissance de points sélectionnés grâce à une fonction d'importance basée sur les vecteurs d'attributs.

L'avancement des recherches concernant l'échographie 3D main libre sans capteur de position a progressivement révélé la nécessité de prendre en compte également le mouvement dans le plan, et de procéder à un recalage de ce mouvement avant de s'intéresser au mouvement hors plan. Prager *et al.* (2003), se basant sur une idée de Treece *et al.* (2002) (initialement utilisée pour la compensation des variations de pression), ont proposé d'estimer le mouvement dans le plan en divisant les images en bandes horizontales et en effectuant une recherche du maximum de corrélation pour chaque bande, permettant ainsi d'obtenir une estimation des translations axiale et latérale, mais aussi de la rotation dans le plan, grâce à une optimisation au sens des moindres carrés. Housden *et al.* (2007) ont par la suite intégré cette étape à leur système de reconstruction de volume échographique, en utilisant des fenêtres au lieu des bandes horizontales. Les translations dans le plan étant inférieures à la taille d'un pixel, il est important de pouvoir obtenir une précision plus raffinée que la taille d'un pixel. Ils ont ainsi proposé dans un précédent article (Housden *et al.*, 2006) d'utiliser une technique d'interpolation. Leur étude présente une comparaison de différentes méthodes d'interpolation regroupées en deux types : des méthodes interpolant directement les pixels de l'image pour rechercher le maximum de corrélation ensuite, et d'autres interpolant les valeurs de corrélation. Ce dernier type de méthode a permis d'obtenir les meilleurs résultats.

1.2.3 Recalage global et erreurs engendrées

Il existe ainsi diverses méthodes pour estimer le mouvement dans le plan d'une part, et le mouvement hors plan d'autre part. Ces méthodes amènent des résultats satisfaisants dans le cadre de translations pures selon la ou les composantes auxquelles elles sont dédiées. Cependant, il serait erroné de considérer que dans le cas d'une transformation rigide à six degrés de liberté, il suffit de combiner les deux estimations pour obtenir une estimation du mouvement global. En effet, chaque degré de liberté influence la corrélation et de ce fait influence également l'estimation des autres degrés de liberté. Il est donc nécessaire de s'intéresser à l'influence des différentes composantes du mouvement sur les différentes estimations. C'est l'objet de la discussion qui suit.

Comme précisé précédemment, en présence d'une simple translation élévotionnelle, et dans les conditions de Rayleigh, le recalage hors plan peut s'effectuer grâce à la connaissance de la courbe de décorrélation. Mais le lien direct entre la décorrélation du speckle et le déplacement élévotionnel n'existe que dans ces conditions. Dès lors que d'autres mouvements entrent en jeu, la décorrélation est affectée, et il est indispensable de tenir compte de ces effets pour effectuer un recalage fiable.

Afin d'étudier avec précision les erreurs commises en fonction des types de mouvement, Li *et al.* (2002) se sont appuyés sur des simulations présentant différentes combinaisons de composantes de mouvement. Ils ont ainsi réalisé une étude des effets de la rotation dans le plan. Leur conclusion a été que même de très faible amplitude, une rotation (dans le plan ou hors plan) cause des erreurs importantes dans les estimations de translation élévotionnelle, et que ces rotations ne peuvent pas être elles-mêmes estimées efficacement par la décorrélation. L'étude menée par Housden *et al.* (2007) parvient à la même conclusion qu'une des principales sources d'erreurs d'estimation de déplacement élévotionnel est la présence de rotations, qui semble résulter en une surestimation systématique du déplacement. Les travaux de Laporte et Arbel (2008) montrent que les rotations hors plan autour de l'axe latéral sont plus particulièrement source d'erreurs, ce qui est lié au fait que la forme de la cellule de résolution varie le plus dans la direction axiale. La raison pour cela a été expliquée par Housden *et al.* (2008). Une rota-

tion de la sonde dans le plan cause une modification des caractéristiques du speckle. Les effets sur le signal sont multiples, mais le phénomène le plus important concerne le changement de phases relatives des différentes composantes du signal. En, effet, si l'angle d'observation des particules est modifié, les interférences entre les signaux individuels le sont également, et le signal résultant est donc lui aussi modifié.

Il est également intéressant de considérer l'effet de rotations de tout type sur l'estimation du mouvement dans le plan. Une étude approfondie de la formation des images ultrasonores, menée par Kallel *et al.* (1994), a permis de démontrer que la PSF d'un échographe présente une courbure dans la direction axiale, ce qui implique une variance spatiale du système, causant des artefacts lors de rotations de la sonde autour des axes latéral et élévational. La conséquence de cette courbure est un décalage apparaissant dans la fonction de corrélation : le pic de corrélation ne correspond alors pas au meilleur alignement entre les images. Dans le cas de signaux RF, Kallel *et al.* proposent de remédier à ce problème en compensant le décalage grâce à une estimation de celui-ci obtenue par la connaissance théorique de la courbure, afin d'obtenir une estimation plus fiable. Maurice et Bertrand (1999) ont par ailleurs proposé une méthode de restauration du speckle basée sur une description Lagrangienne du mouvement, dans le cas de transformations linéaires. Cette restauration permet de compenser les changements dans la morphologie du speckle dus aux mouvements dans le plan, et ainsi de corriger les estimations de transformation. Les mouvements hors plan ne sont cependant pas traités dans cette étude.

Housden *et al.* (2008) et Lang *et al.* (2009) ont proposé de combiner l'information contenue dans l'image avec l'information obtenue par l'intermédiaire d'un capteur de position secondaire pour améliorer la précision du recalage. Ceux-ci permettent un suivi du mouvement très efficace à grande échelle, mais manquent de précision à faible échelle. A l'inverse, le suivi basé sur le speckle présente de forts décalages à grande échelle, mais permet une précision très satisfaisante au niveau des détails. Les deux techniques semblent donc très complémentaires, et les résultats des expériences montrent une amélioration notable. Lang *et al.* (2012) ont développé une approche s'appuyant sur des volumes obtenus par tomodensitométrie : à partir du volume, des images échographiques sont simulées, et une méthode de recalage 3D similaire à

celle proposée par Housden *et al.* (2007) est utilisée pour recalibrer les images échographiques réelles aux images simulées. Le volume tomographique est ainsi utilisé comme référence. Les résultats expérimentaux ont démontré que ceci permet de compenser la propagation des erreurs de reconstruction.

1.3 Problème étudié

Dans le cadre de l'échographie 3D main libre sans capteur de position, il n'existe ainsi pas encore de solution entièrement satisfaisante au problème de l'estimation de la transformation rigide appliquée à la sonde lors de l'acquisition d'une séquence d'images avec un mouvement complexe. Les travaux présentés dans ce mémoire s'intéressent donc à la recherche d'une solution fiable et robuste qui n'exige pas de matériel additionnel.

1.3.1 Définition du contexte d'étude

Le contexte de l'étude est précisé par les hypothèses suivantes, et pourra être élargi dans des travaux ultérieurs par les méthodes citées :

- les images à traiter sont des échographies de mode B, qui sont accessibles sans restriction sur tous les appareils échographiques, contrairement aux images RF. Cela permet d'élargir de façon importante le champ d'application de la méthode proposée, ainsi que sa reproductibilité ;
- les images sont acquises avec une sonde linéaire. Cette hypothèse permet de faciliter les estimations de mouvement dans le plan : une translation latérale de la sonde produit une translation latérale dans l'image, ce qui n'est pas directement le cas avec une sonde convexe ;
- les séquences considérées sont constituées uniquement d'une paire d'images. La reconstruction de volumes nécessite des séquences constituées d'un grand nombre d'images, ce qui implique un problème de propagation des erreurs. Housden *et al.* (2008) proposent de corriger la trajectoire globale grâce à un capteur de position ;

- l'objet imagé respecte les conditions de Rayleigh et génère donc uniquement du speckle pleinement développé. Les courbes de décorrélation peuvent ainsi être utilisées sans procéder à une adaptation préalable au milieu telle que proposée par Laporte et Arbel (2011), nécessaire dans les cas de speckle partiellement développé ;
- l'objet imagé est immobile, et ne subit ni mouvement physiologique, ni variation dans la compression des tissus par la sonde. Cette hypothèse nous permet de considérer que le mouvement appliqué à la sonde entre chaque acquisition se traduit par un mouvement rigide sur les images, et que les erreurs observées sont uniquement dues aux mesures et aux méthodes d'estimation. Treece *et al.* (2002) ont proposé une méthode de compensation des variations de compression des tissus ;
- le mouvement appliqué à la sonde est tel qu'il n'y a pas d'intersection entre les images successives. En effet les intersections sont problématiques car l'estimation de distance élévationnelle ne permet pas de distinguer le sens de déplacement, et une analyse complexe telle que celles proposées par Housden *et al.* (2007) et Laporte et Arbel (2008) doivent être réalisées afin de repérer les intersections causées par les rotations hors plan.

1.3.2 Formalisation du problème

Le problème est formalisé de la manière suivante. La première image de la paire, nommée « image de référence », est désignée par le symbole P , tandis que la seconde, nommée « image à recaler » est désignée par le symbole Q . L'espace est rapporté à un repère cartésien de centre O et d'axes x , y et z tels que définis sur la figure 1.8. Ce repère est lié à l'image P , et l'objectif est de positionner l'image Q dans ce repère. La direction de l'axe y correspond à la direction de propagation des ultrasons (direction axiale), la direction de l'axe x correspond à la direction latérale, et la direction de l'axe z correspond à la direction élévationnelle.

Pour définir et nommer les rotations, le vocabulaire de l'aviation est emprunté. La rotation dans le plan, autour de l'axe z , est nommée roulis, et est désignée par le symbole θ_z . La rotation autour de l'axe y est nommée lacet, de symbole θ_y , celle autour de l'axe x est nommée tangage,

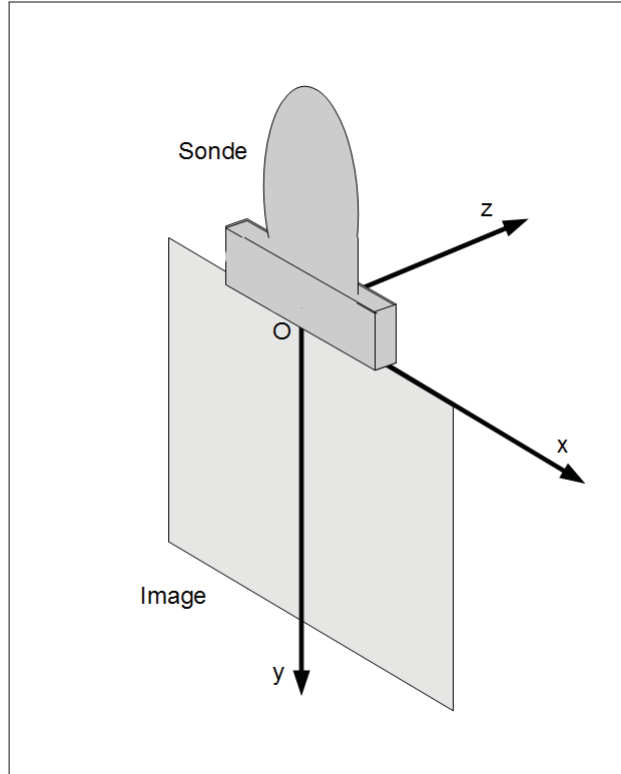


Figure 1.8 Définition du repère

de symbole θ_x . Le sens antihoraire est choisi comme sens positif pour toutes les rotations. Par ailleurs, il est également utile de définir la matrice de rotation R , de dimension 3×3 :

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

Les différentes rotations sont définies par les formules suivantes :

$$\theta_z = \tan^{-1} \left(\frac{r_{21}}{r_{22}} \right),$$

$$\theta_y = \tan^{-1} \left(-\frac{r_{31}}{\sqrt{r_{32}^2 + r_{33}^2}} \right),$$

$$\theta_x = \tan^{-1} \left(\frac{r_{32}}{r_{33}} \right).$$

Ainsi, le mouvement dans le plan est composé de la translation axiale y , de la translation latérale x et du roulis θ_z . Le mouvement hors plan est composé de la translation élévationnelle z , du lacet θ_y et du tangage θ_x (figure 1.7). Notons que les valeurs des déplacements en translation sont notées x , y et z , à ne pas confondre avec la position des images.

1.4 Conclusion

Ce chapitre a permis d'introduire les notions pertinentes à la compréhension des travaux présentés dans ce mémoire, concernant les signaux échographiques et le recalage des images échographiques. En particulier, l'étude des statistiques du second ordre a permis de définir le modèle de décorrélation du speckle, utilisé pour l'estimation de la distance élévationnelle entre les images. De plus, l'étude des méthodes conventionnelles de recalage dans le plan a permis d'identifier les méthodes utiles à l'application que nous considérons. Enfin, les approches existantes d'échographie 3D main libre ont été présentées, ainsi que les problèmes persistants liés à la complexité du mouvement. La suite de ce mémoire est ainsi dédiée à l'élaboration d'une méthode permettant de corriger les erreurs d'estimation dues aux causes multiples de décorrélation.

CHAPITRE 2

ALGORITHME DE RECALAGE NAIF

Dans l'idéal, la reconstruction de volumes échographiques est faite à partir de séquences d'images parfaitement alignées, acquises par un mouvement de sonde en translation élevationnelle. Dans le cas d'une telle séquence idéale, il suffit ainsi d'estimer l'écart entre chaque paire d'images consécutives de la séquence pour reconstruire le volume. La section 1.2.1 décrit les algorithmes existants permettant de réaliser cette estimation. L'algorithme de recalage proposé ici est construit autour d'un tel algorithme de recalage hors plan, adapté de manière à tenir compte de l'imperfection du mouvement de la sonde, introduisant du mouvement dans le plan et des rotations hors plan. Cet algorithme est basé sur celui proposé par Housden *et al.* (2007). Le schéma de la figure 2.1 présente son flux de données.

L'image à recaler est d'abord divisée en k fenêtres. Chaque fenêtre est ensuite traitée individuellement dans le but d'obtenir pour chacune un vecteur de translation à trois dimensions dans les directions x , y et z décrivant le recalage de la fenêtre dans l'image de référence. Pour cela, la matrice de corrélation entre la fenêtre et une zone correspondante dans l'image de référence est construite, et le maximum, ou pic de corrélation, est recherché dans cette matrice. La position de ce maximum permet conventionnellement d'obtenir les points de recalage dans le plan (selon x et y). La valeur de ce maximum permet quant à elle d'estimer la distance élevationnelle (selon z). L'ensemble des k vecteurs à trois dimensions ainsi obtenu permet d'obtenir une estimation de la transformation rigide définie par les six composantes $\{x, y, z, \theta_z, \theta_y, \theta_x\}$, grâce à une méthode d'alignement de Procruste.

Le présent chapitre est composé comme suit. Tout d'abord, le principe de fonctionnement de la méthode de recalage hors plan utilisée, et les étapes qui la composent, sont présentés dans la section 2.1. Ensuite, la section 2.2 décrit le calcul de la matrice de corrélation, ainsi que la méthode de localisation de précision sous-pixel du pic de corrélation, permettant d'obtenir, pour chaque fenêtre, la valeur de corrélation à utiliser pour l'estimation hors plan, ainsi que les estimations de translations dans le plan. Il est ensuite expliqué à la section 2.3 comment estimer

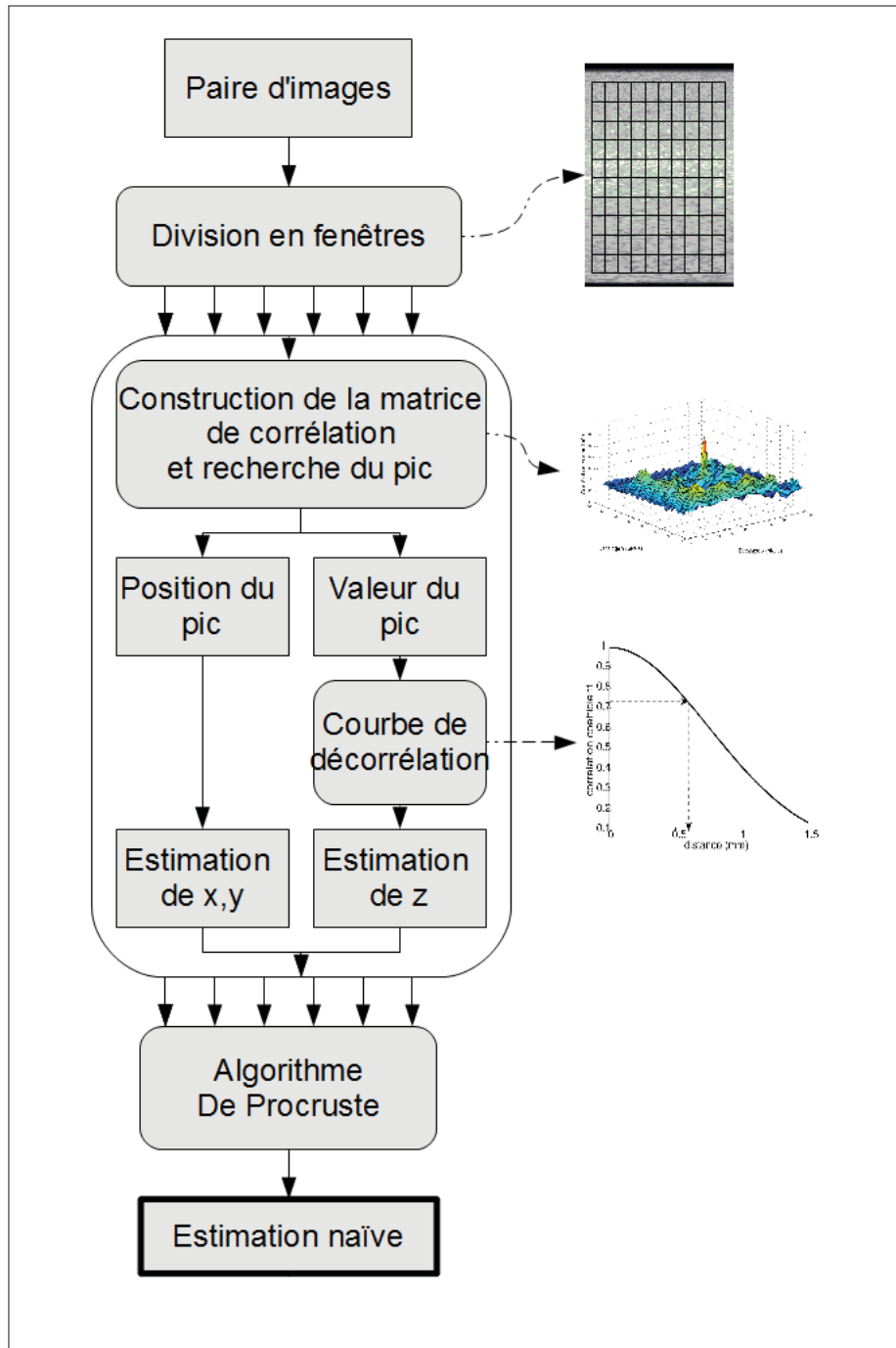


Figure 2.1 Schéma fonctionnel de l'algorithme de recalage naïf

la transformation rigide liant l'image à recaler à l'image de référence, à partir des estimations par fenêtre. Finalement, la section 2.4 présente l'étude expérimentale menée sur l'algorithme naïf afin d'évaluer sa performance et sa validité.

2.1 Recalage hors plan

L'estimation du déplacement hors plan dans les images échographiques est réalisable grâce aux spécificités du signal échographique décrites dans le chapitre 1. En effet, le phénomène de décorrélation du speckle permet de construire un modèle liant la corrélation entre deux images et la distance élévationnelle les séparant. Nous utilisons ici partiellement l'implémentation de Laporte (2011) (les hypothèses posées dans ce mémoire ne permettent en effet pas d'en exploiter la plupart des contributions). Cette implémentation modélise la décorrélation du speckle par une courbe linéaire par morceaux.

Dans les conditions de Rayleigh, ce modèle est spécifique à la sonde employée, mais aussi à la profondeur de la zone imagée. En effet, la largeur des faisceaux d'ultrasons émis par les transducteurs varie avec la profondeur, ce qui influence directement la corrélation du speckle, et donc le modèle de son évolution. Ceci explique la nécessité de diviser l'image en fenêtres et d'associer un modèle spécifique à chaque fenêtre, comme proposé par Housden *et al.* (2007). Par la suite, une mesure de corrélation devra donc être réalisée pour chaque fenêtre, et ainsi une estimation de distance élévationnelle par fenêtre pourra être obtenue. Ces multiples estimations non colinéaires seront notamment utiles pour l'estimation des rotations hors plan lors de la détermination de la transformation globale (partie 2.3).

La construction de ces modèles se fait par une étape de calibrage, au cours de laquelle les paramètres de la courbe sont ajustés pour correspondre au mieux à des valeurs mesurées. Ces mesures sont obtenues à partir d'un fantôme de speckle pleinement développé, c'est-à-dire un objet approximant les conditions de Rayleigh. Celui-ci est balayé à intervalles réguliers, dans un mouvement de translation élévationnelle pure et monotone. Les images obtenues sont divisées en fenêtres de la même manière que pour l'étape d'estimation. Les coefficients de corrélation entre fenêtres correspondantes sont mesurés, et une fonction linéaire par morceaux est ajustée aux valeurs obtenues. Un modèle de décorrélation du speckle peut ainsi être construit pour chaque fenêtre.

Une fois ces modèles construits pour la sonde choisie, l'estimation de la distance élévationnelle pour chaque fenêtre de la paire d'images considérée peut être réalisée. Dans la méthode originale, le coefficient de corrélation est calculé entre chaque fenêtre de l'image à recaler et la même fenêtre de l'image de référence, puis la distance élévationnelle correspondant à ce coefficient est lue sur la courbe. Afin de tenir compte du mouvement dans le plan, une recherche des valeurs optimales de corrélation à utiliser est implémentée dans le présent algorithme, présentée dans la section 2.2. Ces valeurs de corrélation sont ensuite utilisées pour lire l'estimation de la distance hors plan sur la courbe correspondant à la fenêtre considérée. Il est important de noter que l'estimation obtenue dans la direction élévationnelle est une valeur absolue représentant la distance entre les fenêtres : la direction de la translation hors plan ne peut pas être obtenue de cette manière.

2.2 Calculs de corrélation et localisation sous-pixel

Cette section concerne la recherche du pic de corrélation pour chaque fenêtre de l'image à recaler. Chaque fenêtre constitue un problème indépendant dont la résolution est identique. Nous nous intéressons donc ici au cas d'une seule fenêtre.

L'estimation du déplacement élévationnel, décrite à la section 2.1, est basée sur une mesure du coefficient de corrélation. Pour effectuer le recalage dans le plan sur de petits déplacements, l'utilisation d'une méthode par intensité est préconisée dans le cadre de l'échographie. Rappelons que ce type de méthode consiste à rechercher la transformation relative entre les deux images donnant le maximum de similitude. Ainsi, il semble opportun de choisir le coefficient de corrélation comme mesure de similitude, afin de mieux optimiser la méthode et les temps de calcul associés.

L'objectif de cette étape est d'obtenir la position de la fenêtre à recaler dans l'image de référence donnant la valeur de corrélation maximale, ainsi que cette valeur. La position permet en effet de déduire une estimation des translations dans le plan, tandis que la valeur de corrélation à cette position est nécessaire à l'estimation de la distance élévationnelle (section 2.1).

Compte tenu de l'ampleur des mouvements dans le plan considérés (des translations d'ordre cent fois inférieur à la taille de l'image), il n'est pas utile de rechercher le maximum de similitude dans l'ensemble de l'image de référence, et il est ainsi possible de réduire significativement le temps de calcul. Chaque fenêtre de l'image à recaler est donc couplée avec une zone dans l'image de référence qui doit être définie. Dans l'algorithme présenté, nous définissons cette zone telle que la fenêtre à recaler ait toujours au minimum un pixel superposé avec la fenêtre à laquelle elle est comparée dans l'image de référence. Pour une fenêtre de taille $m \times n$ pixels, la zone de référence correspondante a donc la taille $(3m - 2) \times (3n - 2)$ pixels.

La partie suivante décrit la méthode permettant de calculer pour chaque fenêtre les coefficients de corrélation dans la zone de référence. La partie 2.2.2 s'intéresse ensuite au moyen de localiser et mesurer le pic de corrélation avec précision.

2.2.1 Matrice de corrélation

La recherche du pic de corrélation nécessite de construire la matrice de corrélation croisée normalisée, qui est constituée des coefficients de corrélation normalisée. La signification de ces coefficients est d'abord expliquée ci-après, puis la méthode de calcul de la matrice employée dans notre algorithme est définie.

Chaque élément de la matrice de corrélation représente la valeur de la corrélation entre la fenêtre à recaler (identique pour tous les éléments) et une portion de même taille dans la zone de référence. La position de l'élément dans la matrice est dépendante de la position de cette portion dans l'image de référence. Le calcul de la matrice de corrélation est réalisé d'après la méthode rapide proposée par Lewis (1995). Dans ce qui suit, F représente la fenêtre à recaler, telle que $F(x, y)$ représente l'intensité du pixel positionné en (x, y) de F . Similairement, Z représente la matrice d'intensités de la zone de référence. F est de taille $m \times n$, Z est de taille $(3m - 2) \times (3n - 2)$. Les schémas de la figure 2.2 représentent la démarche de construction de la matrice de corrélation pour un exemple simplifié d'une fenêtre de taille 3×3 .

Chaque coefficient de corrélation est calculé entre F et une portion de Z de taille $m \times n$. L'expression du coefficient de corrélation tire son origine d'une mesure de distance entre F

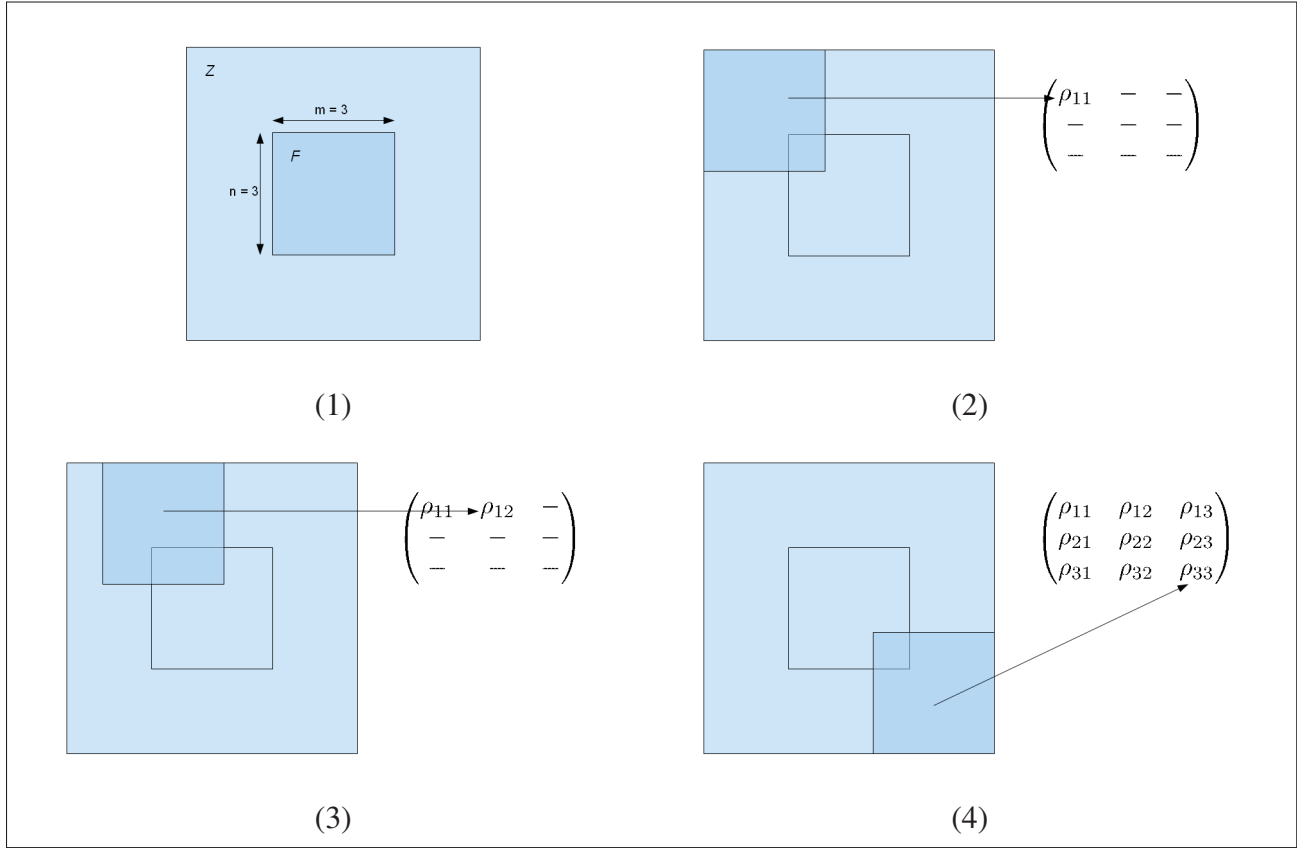


Figure 2.2 Construction de la matrice de corrélation pour une fenêtre de taille 3×3 et la zone de référence correspondante de taille 7×7

et la portion considérée (la distance Euclidienne au carré), dont l'expression pour la portion située en (u, v) est :

$$d_{Z,F}^2(u, v) = \sum_{x,y} (Z(x, y) - F(x - u, y - v))^2,$$

où (x, y) varie de (u, v) à $(u + m, v + n)$.

Cette expression peut être développée de la manière suivante :

$$d_{Z,F}^2(u, v) = \sum_{x,y} (Z^2(x, y) - 2Z(x, y)F(x - u, y - v) + F^2(x - u, y - v)).$$

Le terme $\sum_{x,y} F^2(x-u, y-v)$ est constant (il s'agit de la somme des intensités de F au carré). Si le terme $\sum_{x,y} Z^2(x, y)$ est approximativement constant, ce qui est raisonnable si la taille de la fenêtre est suffisamment grande par rapport à la taille de la zone considérée (il y a alors des superpositions importantes entre les différentes portions $m \times n$ de Z pour toutes valeurs de (u, v)), alors le terme restant, qui correspond à la corrélation croisée, constitue une mesure de similitude entre F et la portion considérée de Z :

$$c(u, v) = \sum_{x,y} Z(x, y) F(x-u, y-v).$$

Cependant, le facteur $\sum_{x,y} Z(x, y)$ dans cette mesure peut être source d'erreur : la corrélation entre la fenêtre et une portion identique de la zone de référence pourrait être inférieure à la corrélation entre la fenêtre et une portion de la zone de référence comportant un point très brillant. Pour cette raison, il est préférable d'utiliser le coefficient de corrélation normalisée :

$$\rho(u, v) = \frac{\sum_{x,y} [Z(x, y) - \bar{Z}_{u,v}] [F(x-u, y-v) - \bar{F}]}{\sqrt{(\sum_{x,y} [Z(x, y) - \bar{Z}_{u,v}]^2 \sum_{x,y} [F(x-u, y-v) - \bar{F}]^2)}},$$

où $\bar{F}$ représente la moyenne de F et $\bar{Z}_{u,v}$ représente la moyenne de la portion $m \times n$ de Z située en (u, v) .

Il n'est pas nécessaire de calculer individuellement chaque coefficient pour toutes les portions possibles de la zone de référence. Il est en effet possible d'effectuer le calcul de la matrice de corrélation dans son ensemble, dans le domaine fréquentiel. Cette méthode, proposée par Lewis (1995), est plus rapide dans le cas où la différence de taille entre F et Z n'est pas trop importante, ce qui est notre cas ici.

Soit $F'(x, y) = F(x, y) - \bar{F}$. Le numérateur de $\rho(u, v)$ s'exprime alors :

$$\begin{aligned} \rho_{num}(u, v) &= \sum_{x,y} (Z(x, y) - \bar{Z}_{u,v}) F'(x-u, y-v) \\ &= \sum_{x,y} Z(x, y) F'(x-u, y-v) - \bar{Z}_{u,v} \sum_{x,y} F'(x-u, y-v). \end{aligned}$$

D'après l'expression de F' , sa somme est nulle, donc le terme $\bar{Z}_{u,v} \sum F'(x-u, y-v)$ est nul. Ainsi,

$$\rho_{num}(u, v) = \sum_{x,y} Z(x, y) F'(x-u, y-v).$$

Cette expression correspond à une convolution : $\rho_{num}(u, v) = Z(x, y) \otimes F'(x-u, y-v)$. Rappelons que l'opération de convolution dans le domaine spatial correspond à la multiplication dans le domaine fréquentiel. Ainsi, la matrice des numérateurs peut être calculée en effectuant le produit des transformées de Fourier (le conjugué doit être utilisé pour F' afin de tenir compte des coordonnées négatives), puis en calculant la transformée inverse du résultat de ce produit :

$$\mathcal{F}^{-1}\{\mathcal{F}(Z)\mathcal{F}^*(F')\},$$

où $\mathcal{F}$ représente la transformée de Fourier, $\mathcal{F}^{-1}$ la transformée inverse, et $\mathcal{F}^*$ le conjugué de la transformée.

Le calcul du dénominateur du coefficient de corrélation est cependant plus difficile à calculer rapidement. Rappelons son expression :

$$\rho_{den}(u, v) = \sqrt{\left(\sum_{x,y} [Z(x, y) - \bar{Z}_{u,v}]^2 \sum_{x,y} F'(x-u, y-v)^2\right)}.$$

Le terme $\sum_{x,y} F'(x-u, y-v)^2$ est constant pour tout (u, v) et peut être pré-calculé. Le second terme est plus problématique car il doit être calculé pour chaque (u, v) : $\sum_{x,y} [Z(x, y) - \bar{Z}_{u,v}]^2$. Lewis (1995) propose de construire au préalable des tables contenant la somme et la somme au carré de chaque portion de Z , permettant ainsi ensuite de calculer le dénominateur très rapidement.

Ainsi, nous pouvons finalement obtenir la matrice de corrélation de la fenêtre considérée. Un exemple d'une telle matrice est représenté sur la figure 2.3, sous forme de surface.

Cependant, telle que cette matrice est définie, le pic de corrélation ne peut être déterminé qu'au pixel près. Or les déplacements recherchés peuvent être du même ordre de grandeur que la

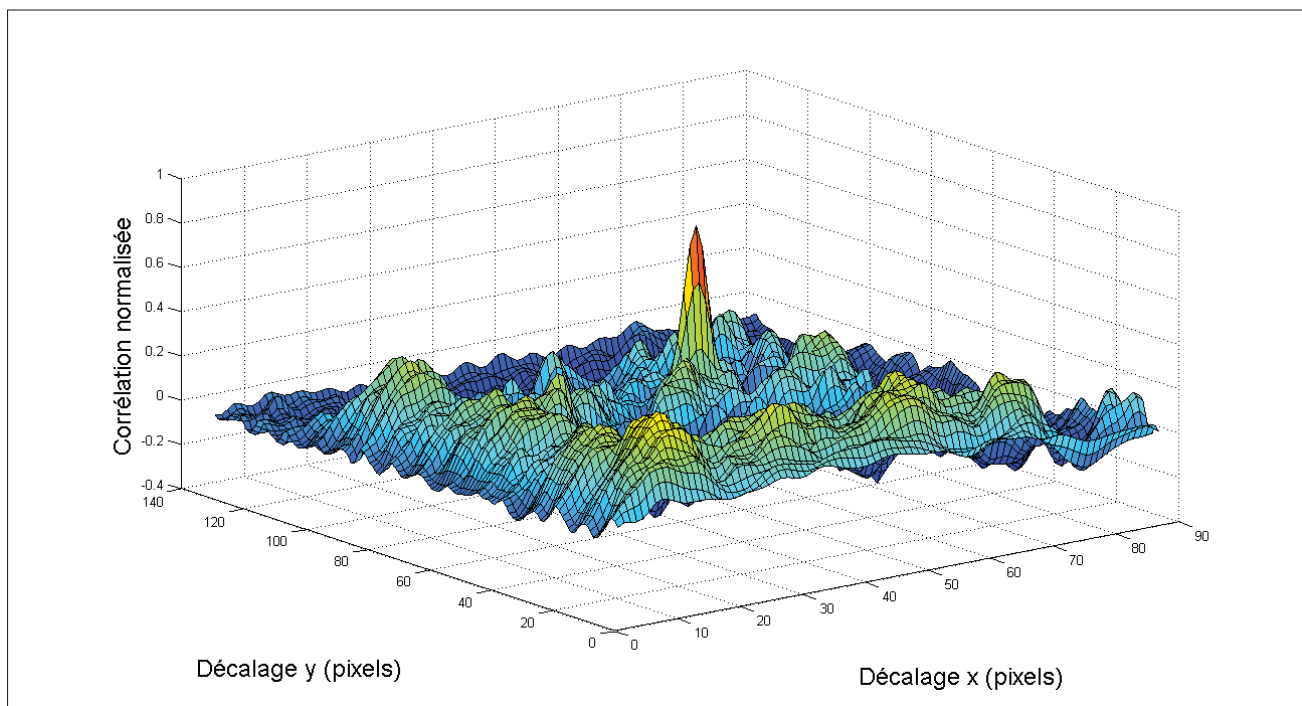


Figure 2.3 Exemple de matrice de corrélation représentée sous forme de surface

taille des pixels (10^{-4} m). De ce fait, la précision au pixel près n'est pas suffisante. La partie suivante décrit la méthode employée pour obtenir une meilleure précision.

2.2.2 Précision sous-pixel et interpolation

Un bon moyen d'augmenter la précision de la localisation du maximum est d'utiliser l'interpolation pour ajouter des échantillons entre les valeurs calculées. Par exemple, si nous voulons améliorer la précision par un facteur s , il nous faut ajouter $s - 1$ échantillons entre les échantillons existants.

L'interpolation correspond à la convolution avec un noyau d'interpolation. Un signal à bande limitée peut être parfaitement reconstruit, sous certaines conditions, avec une fonction noyau de type sinus cardinal. Il s'agit alors d'une «interpolation idéale». La convolution par une fonction à support infini (comme le sinus cardinal) dans le domaine spatial implique néanmoins une approximation à une fonction à support borné qui dégrade fortement le résultat de l'interpolation.

Il est utile de rappeler que la convolution d'un signal avec un sinus cardinal dans le domaine spatial est équivalente à la multiplication de la transformée de Fourier de ce signal par une fonction rectangle, ce qui revient à couper parfaitement les hautes fréquences. La technique d'interpolation choisie exploite précisément ce concept. Elle consiste à effectuer une opération de «zero padding» dans le domaine fréquentiel, c'est-à-dire d'ajouter des zéros avant et après le spectre, dans les deux dimensions. Plus précisément, pour augmenter la précision par un facteur s , la matrice spectrale est plongée au centre d'une matrice de taille s fois plus grande dans chaque dimension, dont toutes les autres valeurs sont nulles. Ceci simule la multiplication par une fonction rectangle, en ajoutant des hautes fréquences et en les fixant à la valeur nulle. Ces hautes fréquences supplémentaires se traduisent dans le domaine spatial par des échantillons supplémentaires intercalés entre les échantillons préexistants ($s - 1$ entre chaque échantillon) (Schafer et Rabiner, 1973). Ainsi, après calcul de la transformée de Fourier inverse, le pic peut être recherché s fois plus précisément.

2.2.3 Calcul du vecteur de translation à trois dimensions

Le schéma de la figure 2.4 résume la méthode de calcul du vecteur à trois dimensions recherché pour chaque fenêtre.

Le calcul de la fonction de corrélation croisée normalisée, puis l'application d'une méthode de suréchantillonnage par interpolation de la matrice obtenue, permettent d'obtenir une surface décrivant le comportement de la corrélation en fonction de la position de la fenêtre par rapport à la zone de référence.

Une fois localisée par simple recherche du maximum, la position du pic doit être remise à l'échelle des pixels. Ainsi, les translations dans le plan peuvent être estimées au $1/s$ pixel près, et ce pour chaque fenêtre. La figure 2.5 représente un exemple de résultat, où les translations dans le plan estimées pour chaque fenêtre sont représentées par un vecteur ayant pour origine le centre de la fenêtre.

Par définition, la valeur du coefficient de corrélation à la position du pic correspond au maximum de corrélation. Cette valeur s'approche de la valeur qui serait obtenue dans le cas d'un mouvement purement élévational, sans mouvement dans le plan. C'est donc cette valeur qui

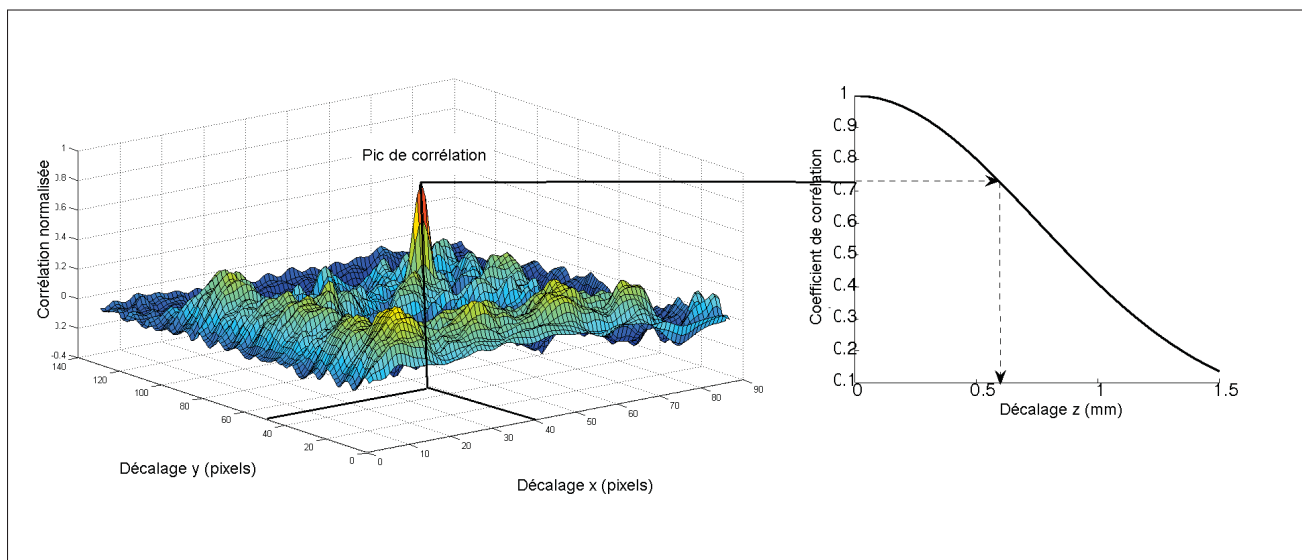


Figure 2.4 Estimation des trois dimensions de translation pour le recalage d'une fenêtre

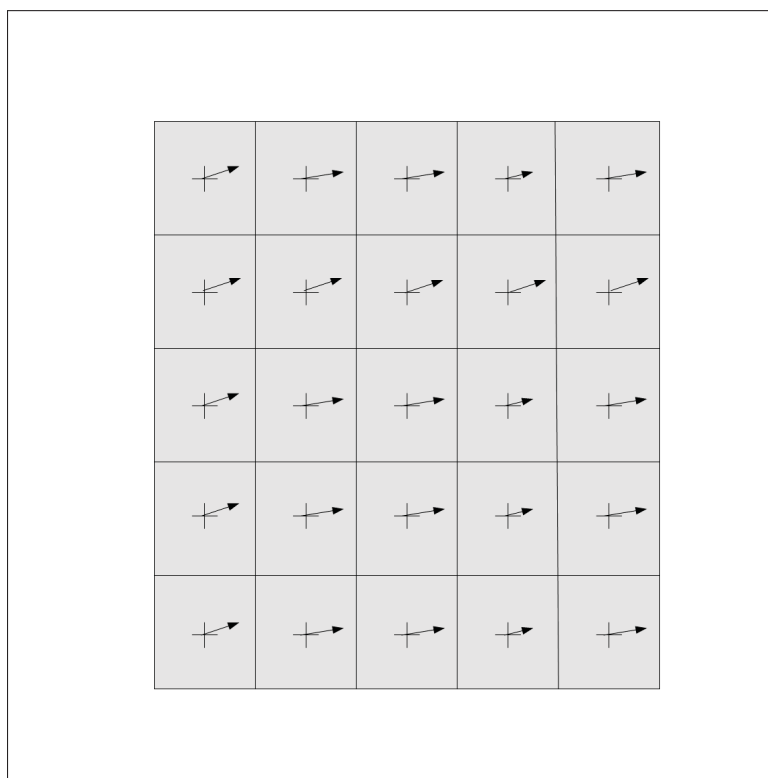


Figure 2.5 Représentation schématique des translations dans le plan estimées pour chaque fenêtre

est utilisée dans l'algorithme d'estimation de distance élévationnelle (section 2.1). Néanmoins, la présence éventuelle de roulis (rotation dans le plan) peut faire diminuer significativement la valeur maximale de corrélation par rapport à la valeur pour un déplacement hors plan pur. Ainsi, étant donné que la courbe de décorrélation est modélisée et calibrée pour des translations élévationnelles pures, les distances hors plan devraient être sur-estimées en présence d'autres mouvements. Ceci est étudié en détail par expérimentation (section 2.4).

Finalement, trois estimations sont obtenues pour chaque fenêtre, correspondant aux trois translations dans l'espace à effectuer à partir de la position d'origine dans le repère de référence, pour recaler au mieux chaque fenêtre de l'image à recaler dans l'image de référence (figure 2.6).

2.3 Calcul de la transformation globale par alignement de Procruste

Le mouvement de la sonde est exprimé par une transformation rigide à six composantes : $\{x, y, z, \theta_z, \theta_y, \theta_x\}$. La contrainte de planarité des images et l'hypothèse d'absence de mouvement physiologique impliquent que le mouvement dans l'image correspond à cette même transformation rigide. L'objectif est donc de retrouver la transformation rigide la plus représentative des k vecteurs à trois dimensions obtenus par la méthode décrite à la section précédente appliquée aux k fenêtres.

Les vecteurs obtenus permettent de définir les positions des centres des fenêtres de l'image à recaler dans le repère de référence (figure 2.7). Ainsi, les positions des centres des fenêtres de l'image de référence sont désignées par $C_P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_k\}$, et les positions estimées des centres des fenêtres de l'image à recaler dans le repère de référence sont définies par $C_Q =$

$\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$, telles que $\mathbf{p}_i + \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \mathbf{q}_i$, où $\mathbf{p}_i$ et $\mathbf{q}_i$ sont des points décrits par trois coordonnées selon les axes x, y et z , et x_i, y_i et z_i désignent les translations estimées selon ces axes pour la i ème fenêtre.

Une méthode de recalage par points de contrôle peut ainsi être appliquée afin de trouver la transformation rigide en trois dimensions entre les deux images, en mettant en correspondance

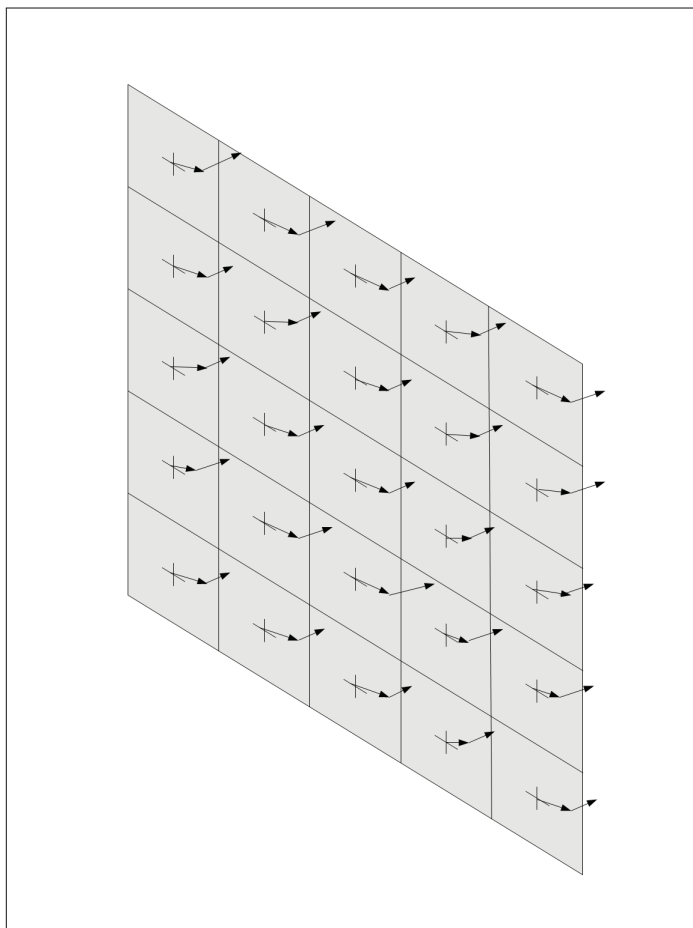


Figure 2.6 Représentation schématique des trois translations de recalage estimées pour chaque fenêtre

les positions des centres à recaler avec les positions des centres de référence. La méthode optimale (au sens des moindres carrés) pour résoudre un tel problème est connue sous le nom de méthode de Procruste¹. Cette méthode permet de déterminer une transformation de similitude entre deux ensembles de points, en minimisant un critère d'erreur quadratique. La validité de cette méthode a été démontrée à plusieurs reprises (Schönemann, 1966; Farrell *et al.*, 1966; Arun *et al.*, 1987; Umeyama, 1991).

1. Ce nom fait référence à un personnage de la mythologie grecque qui étirait ou coupait les membres de ses invités pour les ajuster à la taille de son lit. Il s'agissait initialement d'une référence péjorative critiquant la déformation d'un ensemble de points pour l'ajuster à une référence dans le cadre d'analyses statistiques. Le terme est aujourd'hui utilisé sans cette connotation négative.

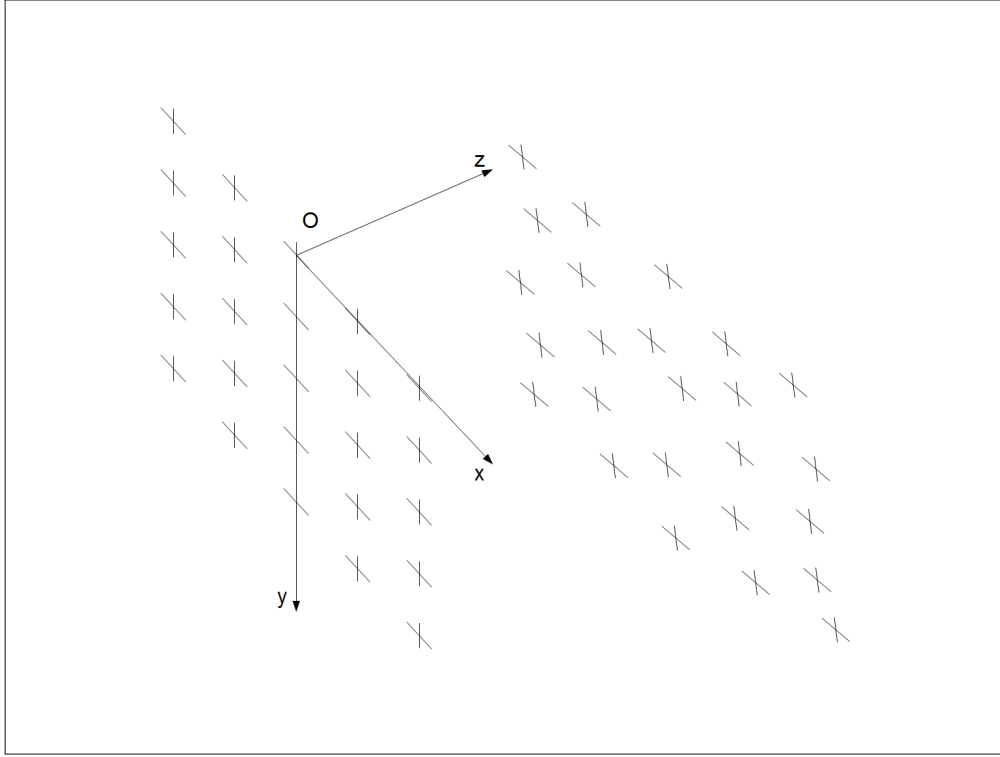


Figure 2.7 Disposition des points dans le repère

L'objectif est de trouver la matrice de rotation $\mathbf{R}$ et le vecteur de translation $\mathbf{t}$ tels que :

$$(\mathbf{R}, \mathbf{t}) = \underset{\mathbf{R}, \mathbf{t}}{\operatorname{argmin}} \sum_{i=1}^k \|(\mathbf{R}\mathbf{p}_i + \mathbf{t}) - \mathbf{q}_i\|^2. \quad (2.1)$$

Le vecteur de translation contient directement les trois composantes recherchées : les translations latérale, axiale et élévationnelle :

$$\mathbf{t} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Les trois composantes rotationnelles que nous exprimons sous forme d'angles d'Euler, peuvent être extraites de la matrice $\mathbf{R}$ selon la méthode présentée à la partie 1.3.2.

Les étapes de calcul de la transformation rigide optimale, présentées sur la figure 2.8, sont les suivantes :

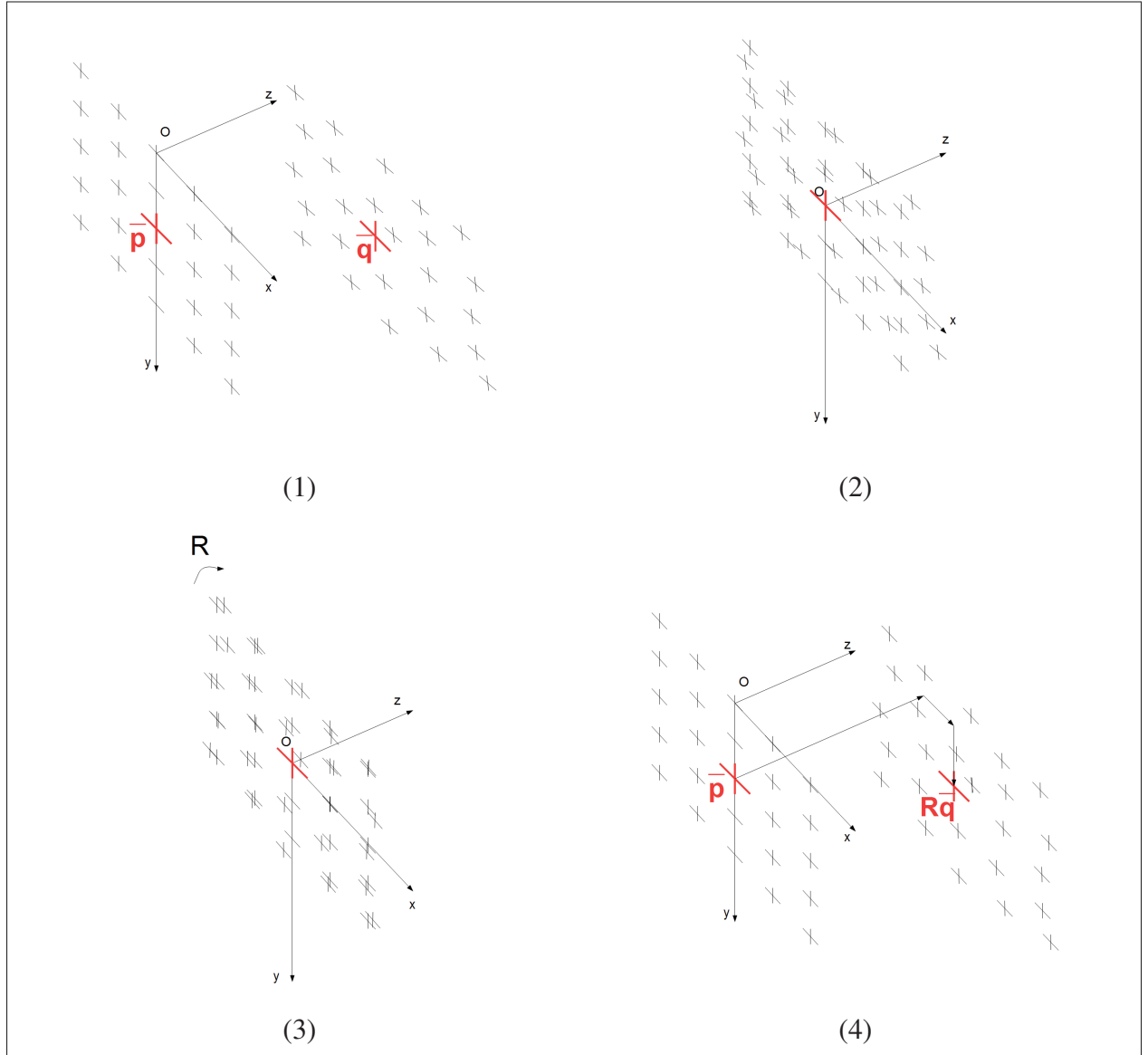


Figure 2.8 Étapes de calcul de la transformation rigide optimale

- calculer les centroïdes des deux ensembles de points (figure 2.8 (1)) :

$$\bar{\mathbf{p}} = \frac{\sum_{i=1}^k \mathbf{p}_i}{k},$$

$$\bar{\mathbf{q}} = \frac{\sum_{i=1}^k \mathbf{q}_i}{k};$$

- calculer les positions centrées (figure 2.8 (2)) :

$$\mathbf{p}'_i = \mathbf{p}_i - \bar{\mathbf{p}},$$

$$\mathbf{q}'_i = \mathbf{q}_i - \bar{\mathbf{q}};$$

- calculer la matrice de covariance de taille 3×3 :

$$\mathbf{S} = \mathbf{P}'\mathbf{Q}'^T,$$

où $\mathbf{P}'$ et $\mathbf{Q}'$ sont les matrices $3 \times k$ contenant respectivement les vecteurs colonne $\mathbf{p}'_i$ et $\mathbf{q}'_i$;

- calculer la décomposition en valeurs singulières :

$$\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{V}^T;$$

- calculer la matrice de rotation $\mathbf{R}$ (figure 2.8 (3)) :

$$\mathbf{R} = \mathbf{V}\mathbf{\Delta}\mathbf{U}^T,$$

avec $\mathbf{\Delta} = \text{diag}(1, 1, \det(\mathbf{V}\mathbf{U}^T))$;

- extraire les composantes de rotation θ_z , θ_y et θ_x à partir de $\mathbf{R}$ (formules en 1.3.2) ;
- calculer le vecteur de translation $\mathbf{t}$ (figure 2.8 (4)) :

$$\mathbf{t} = \bar{\mathbf{q}} - \mathbf{R}\bar{\mathbf{p}}.$$

Une estimation de la transformation rigide, exprimée par ces six composantes, est ainsi finalement obtenue.

2.4 Expérimentations sur l'algorithme naïf

2.4.1 Protocole expérimental

Afin d'évaluer les performances de l'algorithme présenté dans ce chapitre, il est nécessaire de le tester sur différents types de données. Les données en question sont ici des séquences de deux images échographiques de mode B, la sonde étant déplacée suivant une trajectoire connue entre les deux. Notre méthode a été testée sur des images synthétiques, et donc idéales, avec des trajectoires parfaitement connues.

De telles séquences ont été générées artificiellement grâce au logiciel Field II, qui permet de simuler le processus d'acquisition échographique. Ce logiciel, développé par Jensen et Svendsen (1992) puis Jensen (1996), est basé sur une méthode développée par Tupholme (1969) et Stepanishen (1971), qui consiste à calculer la convolution d'une fonction d'excitation (par exemple une sinusoïde à enveloppe gaussienne) avec la PSF du transducteur, qui est constituée par superposition des PSF de chaque élément du transducteur. Grâce à ce logiciel, il est ainsi possible de simuler un fantôme bien défini et d'en acquérir différentes images.

Les simulations ont été faites à partir d'un fantôme de speckle pleinement développé, obtenu grâce à une concentration de 16 diffuseurs ponctuels par mm^3 . Celui-ci a été balayé à l'aide d'une sonde linéaire de 3.5 MHz, à une profondeur de 6cm et avec une zone focale à 3cm, avec une fréquence d'échantillonnage de 50MHz. Des paires d'images RF ont ainsi pu être synthétisées, avec différentes trajectoires de sonde, définies par les six composantes de transformation rigide. Les trajectoires choisies seront décrites pour chaque expérience.

Ces images ont ensuite été transformées en mode B par un processus de détection d'enveloppe, compression logarithmique et interpolation des lignes, afin de simuler les images obtenues par les appareils conventionnels (un exemple d'image synthétique obtenue après ce processus, puis divisée en fenêtres, est présenté sur la figure 2.9). Ces images ont ensuite été décompressées selon l'algorithme proposé par Prager *et al.* (2003) afin d'obtenir une approximation du signal d'enveloppe, nécessaire pour l'application du modèle de décorrélation.

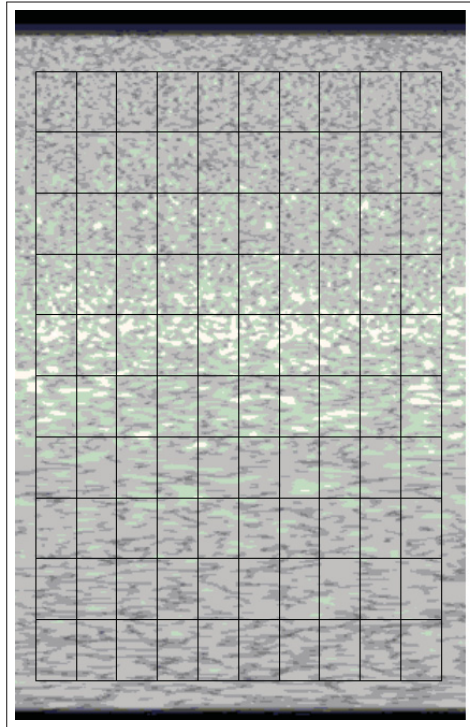


Figure 2.9 Exemple d'image synthétique et division en fenêtres

Une étude préliminaire a permis de montrer que le nombre de fenêtres a peu d'importance, mais une division en 10×10 fenêtres donne des résultats marginalement meilleurs et a donc été adoptée. Par ailleurs, une précision au quart de pixel près a été choisie pour la localisation du pic de corrélation ($s = 4$).

Avant de pouvoir mener les expériences, une étape préalable de calibrage doit être réalisée, afin de construire les courbes de décorrélation de chaque fenêtre. Dans ce but, une séquence de 100 images a d'abord été générée, avec un mouvement de translation purement élévacionnelle de la sonde, et une acquisition tous les 0.1mm. Ces images ont ensuite été divisées en 10×10 fenêtres. Puis pour chaque fenêtre, une courbe de décorrélation a été ajustée aux mesures réalisées. Par la suite, chaque paire d'images testée est divisée de la même manière, et les estimations de distance élévacionnelle sont réalisées à partir de la courbe correspondante.

2.4.2 Mesures d'erreur

Afin d'évaluer les résultats obtenus, différentes mesures d'erreur ont été choisies. Tout d'abord, lorsque l'évaluation de la précision pour une seule composante de la transformation est nécessaire, une simple différence entre la valeur estimée et la valeur réelle de cette composante est utilisée. Ensuite, lorsque l'évaluation de la précision de l'estimation de la transformation globale est recherchée, la mesure utilisée est «l'erreur moyenne de recalage des cibles», notée *mTRE* (*mean Target Registration Error*). Cette mesure, prônée par van de Kraats *et al.* (2005) est parmi les plus répandues pour évaluer la performance des algorithmes de recalage parce qu'elle ne dépend pas du choix d'un centre de rotation. Ici, les cibles considérées sont les centres des fenêtres, tels que définis à la section 2.3. Ce type de mesure peut en fait être utilisé pour évaluer différentes étapes de notre algorithme.

Pour évaluer l'étape d'estimation de déplacement de chaque fenêtre (translations dans le plan et hors plan), décrite aux sections 2.1 et 2.2, il s'agit de calculer la moyenne des distances entre les centres $\mathbf{p}_i$ transformés selon la transformation réelle T_r , ($T_r(\mathbf{p}_i)$) et les points $\mathbf{q}_i$, correspondant aux positions de recalage obtenues pour chaque fenêtre. Ceci permet de rendre compte de la précision des valeurs utilisées ensuite par l'algorithme de Procruste. Cette mesure, notée $mTRE_{init}$, est donc définie de la manière suivante, pour une paire d'images divisées en k fenêtres :

$$mTRE_{init} = \frac{1}{k} \sum_{i=1}^k ||T_r(\mathbf{p}_i) - \mathbf{q}_i||.$$

Pour évaluer l'étape de calcul de la transformation de Procruste à partir des estimations de translations individuelles des fenêtres, il faut calculer la moyenne des distances entre les centres $\mathbf{p}_i$ transformés selon la transformation de Procruste T_{proc} et les points $\mathbf{q}_i$ (ceci correspond à la valeur minimisée par l'algorithme de Procruste, voir l'équation 2.1). Ceci permet d'évaluer à quel point les positions estimées correspondent à une transformation rigide. Cette mesure est notée $mTRE_{proc}$:

$$mTRE_{proc} = \frac{1}{k} \sum_{i=1}^k ||T_{proc}(\mathbf{p}_i) - \mathbf{q}_i||.$$

Finalement, pour évaluer l'ensemble de cette première partie de l'algorithme, la mesure $mTRE_{naive}$ est définie comme la moyenne des distances entre les centres $\mathbf{p}_i$ transformés selon la transformation de Procruste et ces centres transformés par la transformation réelle :

$$mTRE_{naive} = \frac{1}{k} \sum_{i=1}^k ||T_{proc}(\mathbf{p}_i) - T_r(\mathbf{p}_i)||.$$

2.4.3 Expériences sur séquences avec translations pures

L'objectif de l'expérimentation sur l'algorithme naïf est d'étudier son comportement pour des séquences présentant différents types de mouvements restreints à certaines composantes, afin d'identifier l'origine des erreurs observées sur des séquences avec mouvement libre. Grâce aux différentes mesures d'erreur définies précédemment, une étude du fonctionnement de l'algorithme au cours de ses différentes étapes peut également être réalisée.

Ainsi, des mouvements composés de diverses combinaisons des composantes ont été définis pour l'acquisition des paires d'images de test. Pour chaque type de mouvement, les résultats sont présentés sous la forme de trois graphiques. Les deux premiers présentent les erreurs sur les trois translations et les trois rotations respectivement (la différence entre la composante estimée par l'algorithme et la composante réelle), en fonction de la valeur réelle de la composante variable. Le troisième présente l'évolution des trois $mTRE$ ($mTRE_{init}$, $mTRE_{proc}$ et $mTRE_{naive}$), également en fonction de la valeur réelle de la composante variable.

Tout d'abord, des séquences présentant des translations "pures", c'est-à-dire dans une seule direction, ont été générées, afin d'étudier le comportement de l'algorithme pour chaque composante :

- 10 paires d'images avec un mouvement en translation selon x , variant de 0.1 à 1 mm avec un pas de 0.1 mm ;
- 10 paires d'images avec un mouvement en translation selon y , variant de 0.1 à 1 mm avec un pas de 0.1 mm ;

- 25 paires d'images avec un mouvement en translation selon z , variant de 0.02 à 0.5 mm avec un pas de 0.02 mm (la composante en z étant la composante principale du mouvement dans l'application qui nous intéresse, celle-ci été examinée plus précisément).

Les résultats pour l'estimation du mouvement de ces séquences sont présentés sur la figure 2.10. Dans le cas des mouvements de translation pure selon x (figure 2.10 (1)), les graphiques d'erreurs sur chaque composante montrent que les trois composantes dans le plan, ainsi que la composante de lacet (rotation selon θ_y) sont estimées avec des erreurs négligeables. Par contre, pour ces mouvements, des erreurs importantes sont observées sur les deux autres composantes hors plan, qui ne sont pas reconnues comme étant nulles. z est sur-estimée et θ_x est sous-estimée.

La sur-estimation de z provient de la décorrélation impliquée par la translation dans le plan : même après recalage dans le plan, les valeurs de corrélation obtenues pour chaque fenêtre sont diminuées par rapport à la valeur attendue pour une translation élévationnelle nulle, donc l'algorithme de recalage hors plan estime des valeurs non nulles de translation élévationnelle pour chaque fenêtre. Pour expliquer les erreurs d'estimation de tangage (rotation selon θ_x), rappelons que la longueur de la cellule de résolution dans la direction élévationnelle varie significativement avec la profondeur (à cause de la divergence des faisceaux ultrasonores). Ainsi, une même valeur de corrélation est interprétée par l'algorithme comme une distance plus faible pour une fenêtre plus basse dans l'image. Ce phénomène est illustré sur la figure 2.11. L'algorithme de Procruste interprète à son tour ce phénomène comme une rotation négative selon θ_x associée à une translation selon z .

Les valeurs prises par les différentes $mTRE$ témoignent également de ce phénomène : la perturbation causée par l'amplitude des pics de corrélation provoque lors de l'alignement de Procruste une disposition des points de contrôle très différente de celle attendue, ce qui cause des erreurs importantes de $mTRE_{init}$. L'algorithme de Procruste en déduit une estimation de transformation erronée, reflétée par les valeurs de $mTRE_{naive}$ très proches de celles de $mTRE_{init}$. Cependant ces erreurs n'affectent pas la rigidité apparente de la transformation indiquée par la disposition des points de contrôle, d'où la bonne performance démontrée par $mTRE_{proc}$.

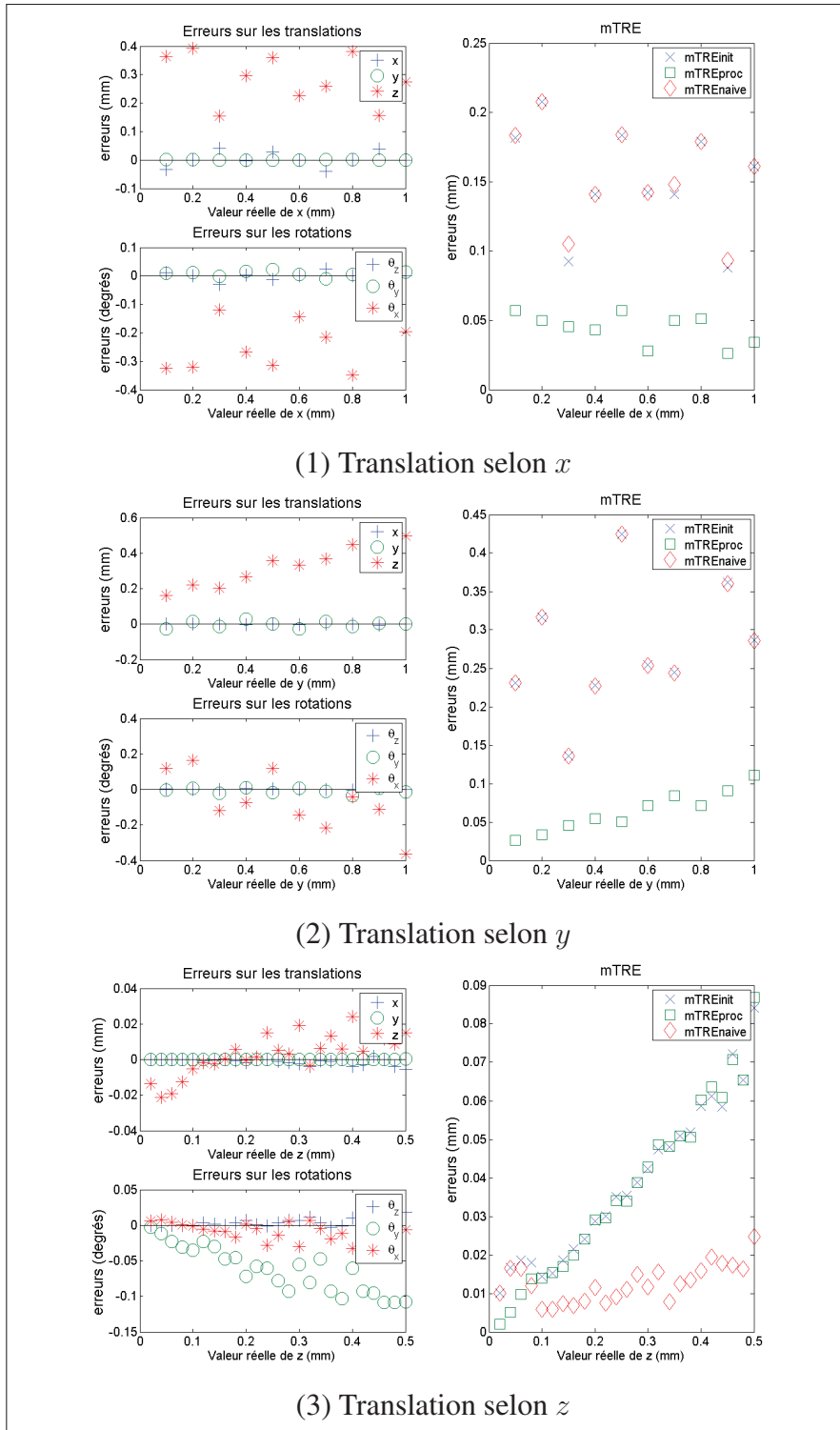


Figure 2.10 Résultats sur séquences de mouvement pur selon chaque composante. À gauche : Erreurs sur les composantes en translation (haut) et en rotation (bas). À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante en mouvement

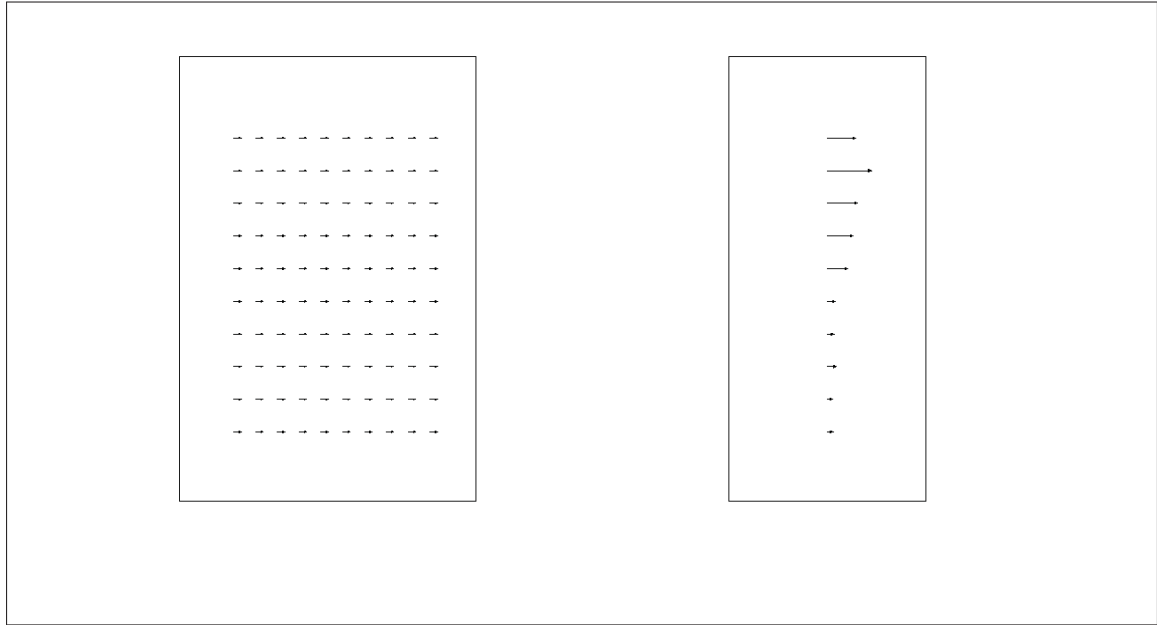


Figure 2.11 Représentation des vecteurs de translation estimés pour chaque fenêtre, dans le cas d'une translation pure de 1mm selon x . À gauche : Vecteurs de translation de recalage dans le plan par rapport aux centres des fenêtres. À droite : Vecteurs de translation élévationnelle estimés pour chaque fenêtre de la colonne centrale

Les résultats pour les mouvements de translation pure selon y (figure 2.10 (2)) sont similaires à ceux obtenus pour x , avec quelques différences cependant. L'erreur sur z augmente presque proportionnellement avec la valeur réelle de y , et l'erreur sur θ_x est beaucoup plus irrégulière que pour les séquences de déplacement selon x . De plus, les $mTRE$ sont environ deux fois plus importantes. Ceci est dû à l'épaisseur moins importante de la cellule de résolution dans la direction axiale : la décorrélation est plus importante pour une translation axiale que par une translation latérale de valeur égale. Par conséquent les erreurs sur z sont directement fonction de la valeur de y . Les erreurs sur θ_x sont plus mitigées car la décorrélation étant plus importante, les valeurs de distance élévationnelle estimées pour chaque fenêtre (figure 2.12) ne coïncident pas nécessairement avec un mouvement de tangage, comme c'est le cas pour les translations selon x .

Dans le cas des mouvements de translation élévationnelle pure (figure 2.10 (3)), les graphiques des erreurs sur les composantes montrent une erreur faible sur l'estimation de z (inférieure à 0.03mm pour toutes les valeurs). Il est intéressant de noter que l'estimation de z est optimale

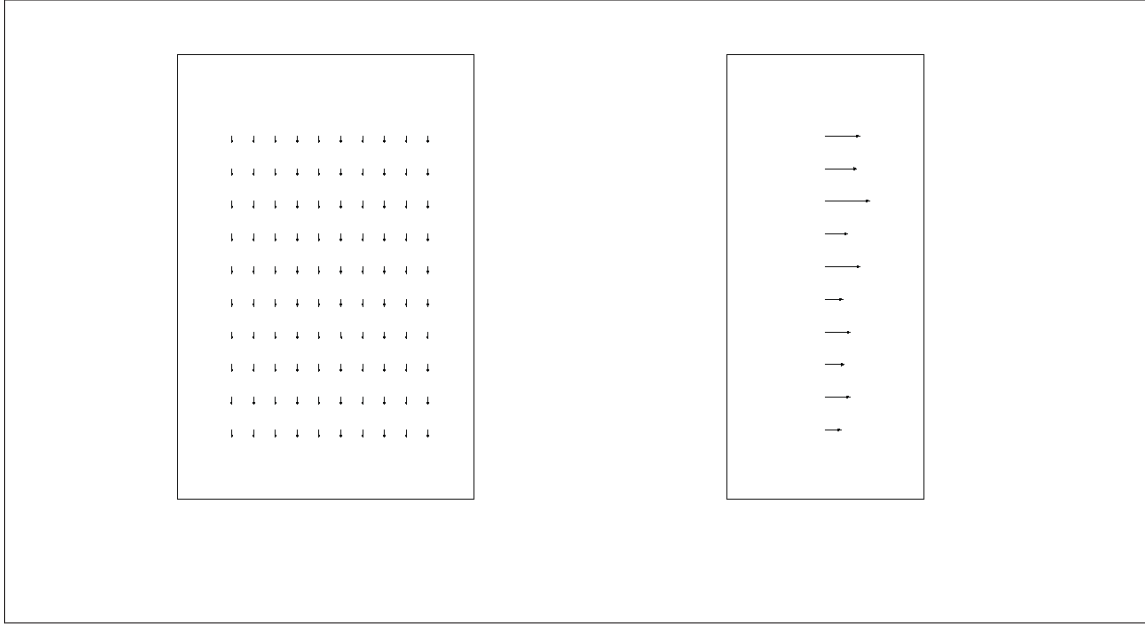


Figure 2.12 Représentation des vecteurs de translation estimés pour chaque fenêtre, dans le cas d’une translation pure de 1mm selon y . À gauche : Vecteurs de translation de recalage dans le plan par rapport aux centres des fenêtres. À droite : Vecteurs de translation élévationnelle estimés pour chaque fenêtre de la colonne centrale

pour des valeurs réelles comprises entre 0.1 et 0.3mm, et que z est notamment sous-estimée pour des translations inférieures à 0.1mm. Ce phénomène, observé auparavant par Hassenpflug *et al.* (2004), est dû à la forme de la courbe de décorrélation : pour des valeurs de distance élévationnelle faibles, la pente de la courbe n’est pas assez importante pour permettre une estimation précise. Les erreurs sur les rotations hors plan augmentent avec la valeur de z . Par ailleurs, l’estimation du mouvement dans le plan est bonne. Le graphique des $mTRE$ montre que des erreurs croissantes sont faites sur l’estimation de la disposition des points de contrôle ($mTRE_{init}$). Cependant, cette disposition ne correspond pas à une transformation rigide (erreur également croissante de $mTRE_{proc}$), et la transformation rigide estimée permet de corriger en grande partie les erreurs initiales d’estimation : $mTRE_{naive}$ est faible et relativement constante.

2.4.4 Expériences sur séquences avec mouvements parasites

Par la suite, nous nous sommes intéressés à l’impact de l’introduction de composantes de mouvement additionnelles, afin d’évaluer la performance de l’algorithme pour chaque combinaison

et d'identifier les situations problématiques. Tout d'abord, nous nous sommes concentrés sur l'estimation des translations dans le plan avec uniquement du mouvement dans le plan, puis sur l'estimation de la translation élévationnelle avec uniquement d'autres composantes hors plan. Ensuite, nous avons observé l'effet de l'introduction de mouvement hors plan sur l'estimation des translations dans le plan, puis de la même manière l'effet de l'introduction de mouvement dans le plan sur l'estimation de la translation élévationnelle.

Les séquences suivantes présentant du mouvement soit uniquement dans le plan soit uniquement hors plan ont d'abord été générées :

- deux groupes de 10 paires d'images combinant les deux types de translations dans le plan : une translation constante de 0.4mm selon x (respectivement y), et une translation variant de 0.1 à 1mm avec un pas de 0.1mm selon y (respectivement x) ;
- deux groupes de 20 paires d'images combinant une translation dans le plan constante de 0.4mm (selon x pour un groupe et y pour l'autre groupe), et une rotation dans le plan (roulis) variant de 0.1 à 2° avec un pas de 0.1° ;
- deux groupes de 10 paires d'images combinant une translation constante de 0.4mm selon z et une rotation hors plan selon θ_x (tangage) pour un groupe et θ_y (lacet) pour l'autre groupe variant de 0.1 à 1° avec un pas de 0.1°.

Les graphiques des figures 2.13 et 2.14 présentent les résultats obtenus pour les translations dans le plan constantes combinées à d'autres composantes dans le plan variables. La combinaison des deux translations dans le plan (figure 2.13 (1) et (2)) montre des résultats similaires aux translations pures, indiquant que les estimations ne sont pas affectées par ce type de combinaison.

L'effet de l'introduction du roulis est similaire pour les translations selon x et y (figure 2.14 (1) et (2)). Plus la valeur réelle de θ_z augmente, plus cette composante est sous-estimée, et plus la translation selon x est sur-estimée : la décorrélation induite par la rotation est interprétée comme une translation latérale. Ce résultat confirme les observations de Morrison *et al.* (1983)

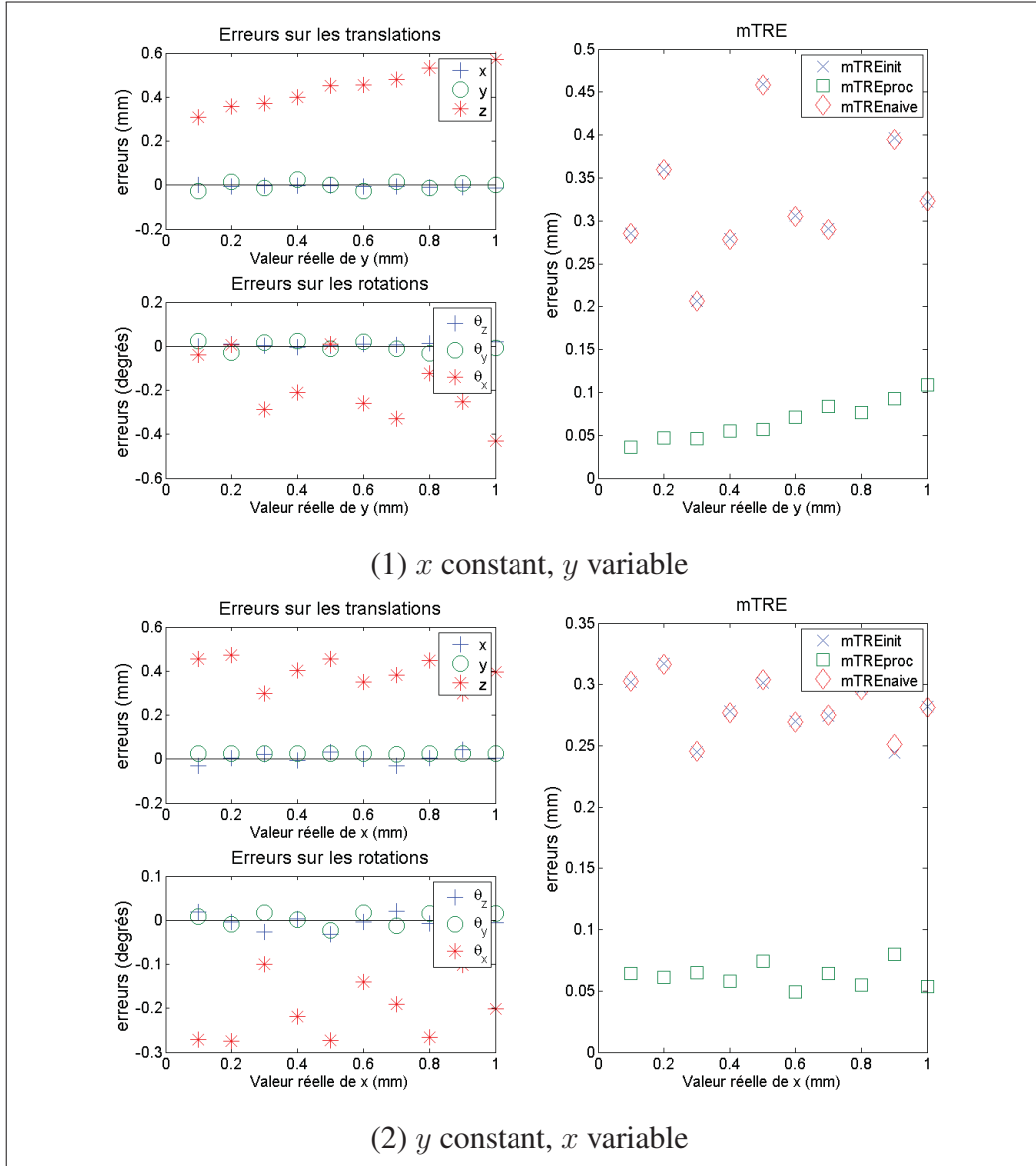


Figure 2.13 Résultats sur séquences avec combinaison de translations dans le plan. À gauche : Erreurs sur les composantes en translation (haut) et en rotation (bas). À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

et Kallel *et al.* (1994), qui ont expliqué ce phénomène par la courbure de la PSF (qui déplace le pic de corrélation). Des erreurs apparaissent également sur les composantes z et θ_x , qui sont toutes les deux progressivement sur-estimées. En effet, les rotations dans le plan causent une décorrélation importante du speckle car les interférences entre les ondes rétro-diffusées sont modifiées. Ceci cause donc des erreurs importantes lors de l'estimation du mouvement

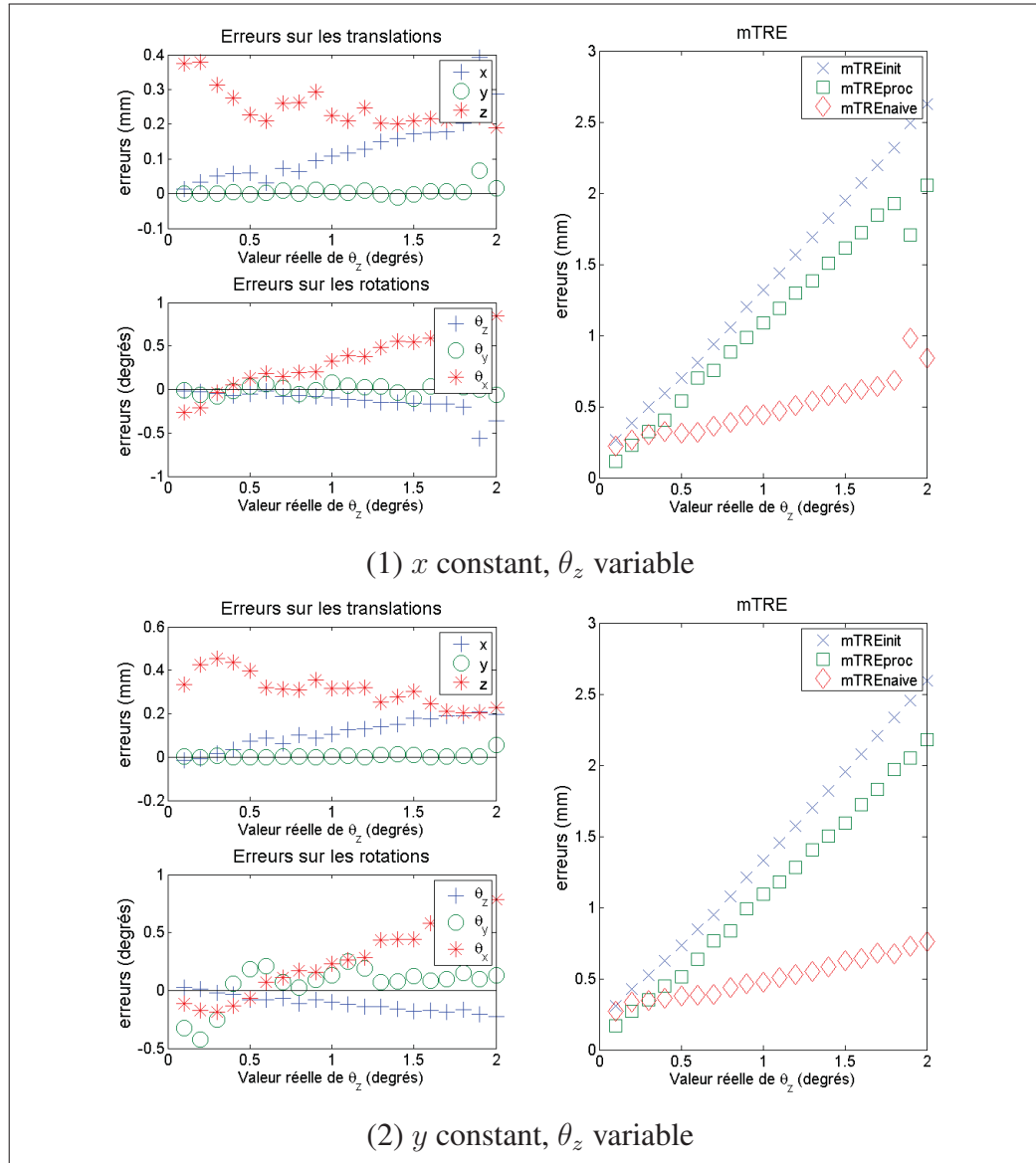


Figure 2.14 Résultats sur séquences avec combinaison de translations et rotations dans le plan. À gauche : Erreurs sur les composantes en translation (haut) et en rotation (bas). À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

hors plan. Le graphique des $mTRE$ (figure 2.14 (1) et (2)) montre des valeurs de $mTRE_{init}$ et $mTRE_{proc}$ beaucoup plus élevées que lorsque seules des translations sont effectuées. Ces valeurs de $mTRE$ sont proportionnelles à la valeur de θ_z . Ceci indique que des erreurs importantes sont faites lors de l'estimation de la disposition des points de contrôle, et que cette disposition correspond de moins en moins à une transformation rigide lorsque le roulis aug-

mente. Le calcul de la transformation rigide permet cependant de corriger l'estimation pour la rendre plus proche de la transformation réelle ($mTRE_{naive}$ est beaucoup moins élevée que les autres $mTRE$).

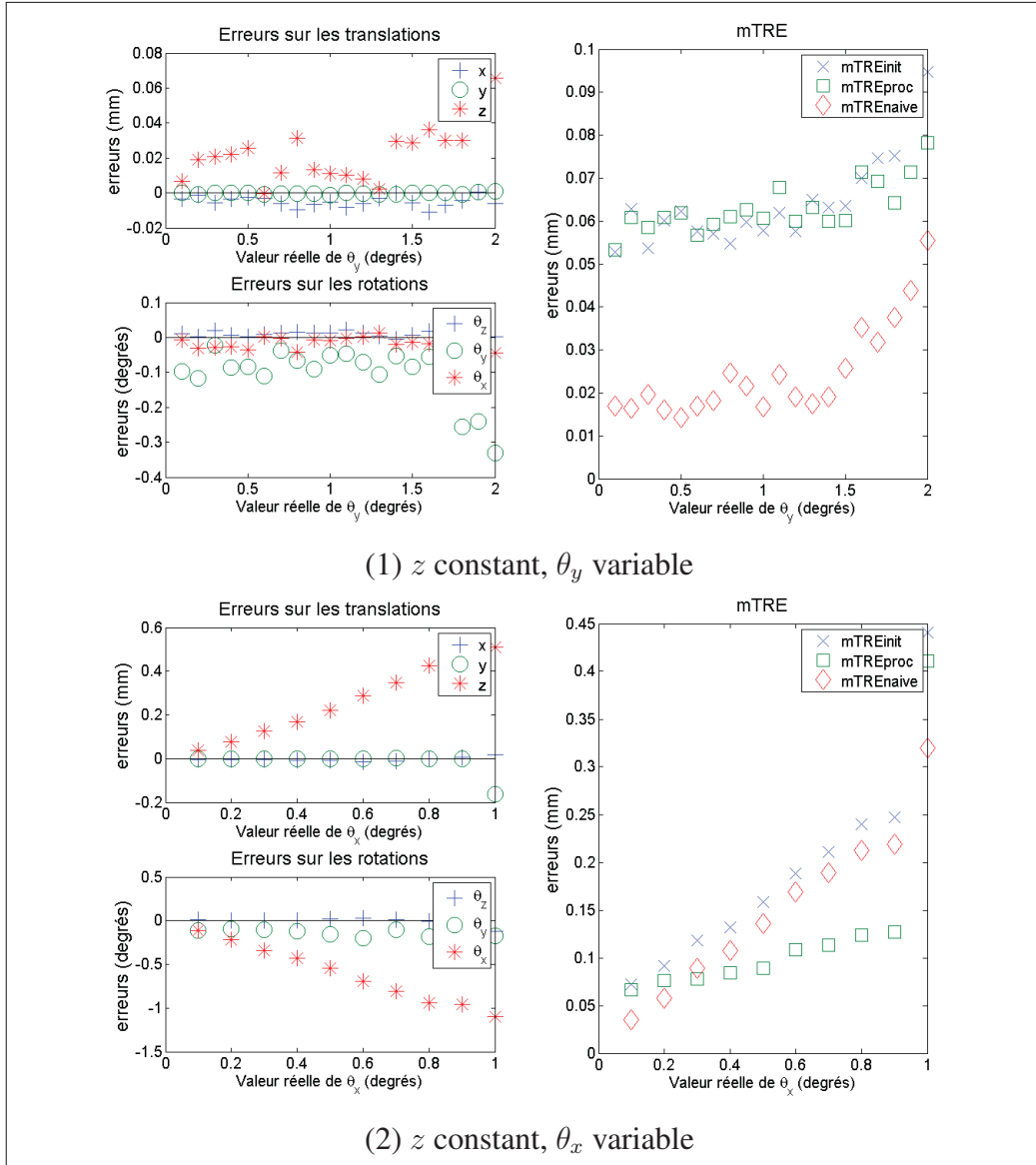


Figure 2.15 Résultats sur séquences de mouvement hors plan. À gauche : Erreurs sur les composantes en translation (haut) et en rotation (bas). À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

Les graphiques de la figure 2.15 présentent les résultats obtenus pour une translation hors plan constante combinée à des rotations hors plan variables. Les résultats produits pour θ_y et θ_x sont très différents. L'introduction du lacet (figure 2.15 (1)) ne change pas beaucoup les résultats par rapport à la translation élévationnelle pure (figure 2.10 (3)). Notons que même θ_y est relativement bien estimée quelle que soit sa valeur. L'introduction du tangage (figure 2.15 (2)) est beaucoup plus problématique. L'erreur effectuée sur θ_x correspond presque parfaitement à sa propre valeur : θ_x est constamment estimée nulle. Toute la décorrélation engendrée semble être attribuée à z qui est progressivement sur-estimée lorsque θ_x augmente. Il s'agit en fait ici encore une fois du problème lié à la courbure de la PSF, observé précédemment pour une combinaison de x et θ_z . La courbure causant des distortions dans les directions latérales et élévationnelles, le même phénomène observé par Morrison *et al.* (1983) et Kallel *et al.* (1994) se produit dans les deux situations.

Afin d'étudier l'effet de l'introduction de mouvement hors plan sur l'estimation dans le plan et inversement, les séquences suivantes ont été générées :

- deux groupes de 10 paires d'images combinant une translation constante de 0.4mm selon x pour un groupe et y pour l'autre groupe, et une translation variant de 0.1 à 1mm avec un pas de 0.1mm selon z ;
- deux groupes de 10 paires d'images combinant une translation constante de 0.4mm selon x pour un groupe et y pour l'autre groupe, et une rotation variant de 0.1 à 2° avec un pas de 0.1° selon θ_x (tangage) (les mouvements de lacet avec seulement du mouvement dans le plan causeraient des intersections et sont donc hors du contexte d'étude) ;
- deux groupes de 10 paires d'images combinant une translation constante de 0.4mm selon z , et une translation variant de 0.1 à 1.9mm avec un pas de 0.1mm selon x pour un groupe et y pour l'autre groupe ;
- 20 paires d'images combinant une translation constante de 0.4mm selon z , et une rotation variant de 0.1 à 2° avec un pas de 0.1° selon θ_z (roulis).

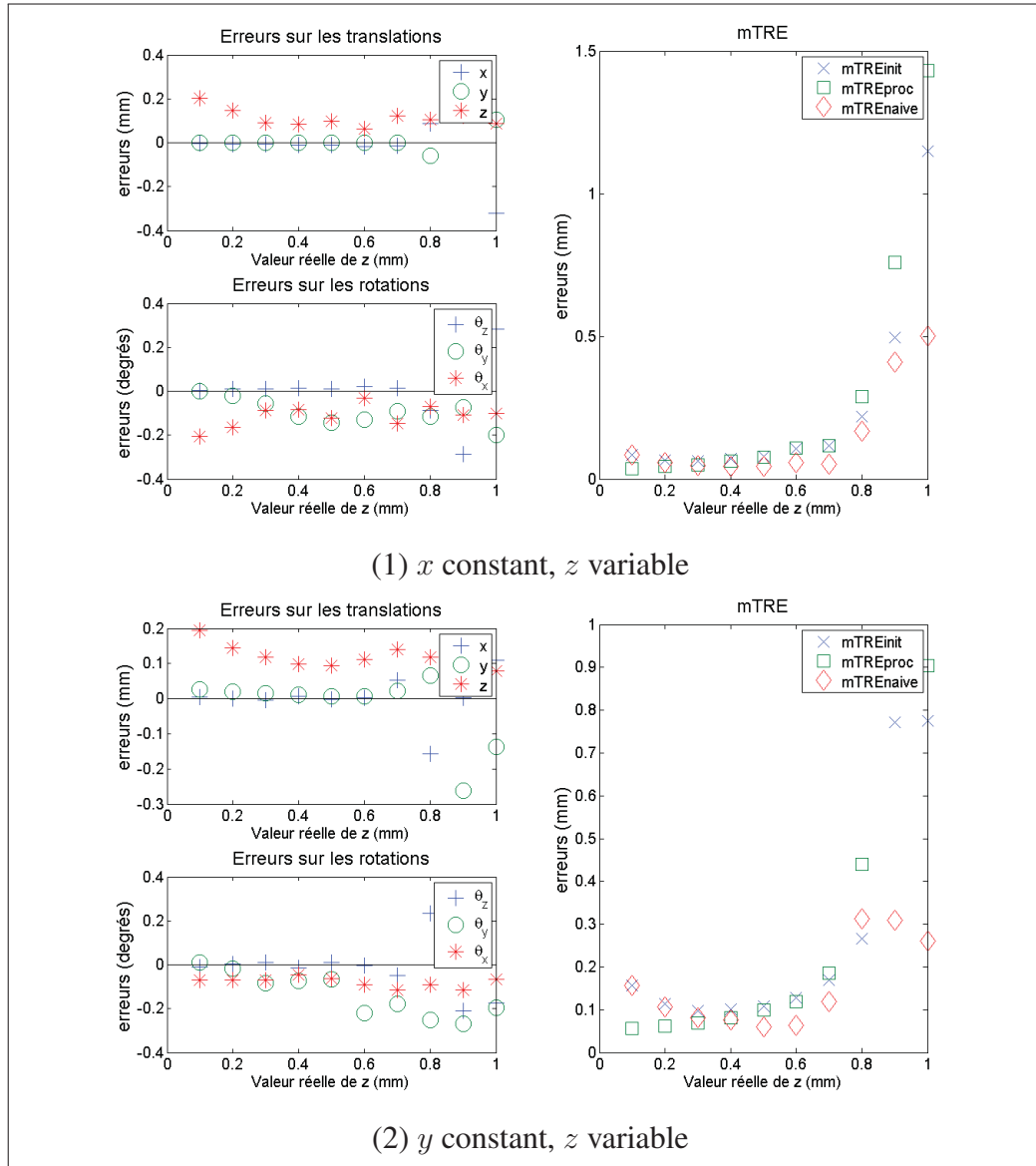


Figure 2.16 Résultats sur séquences de translations dans le plan constantes associées à des translations hors plan variables. À gauche : Erreurs sur les composantes en translation (haut) et en rotation (bas). À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

Les graphiques des figure 2.16 et 2.17 présentent les résultats obtenus pour le mouvement dans le plan constant et le mouvement hors plan variable, sous la même forme que précédemment. L'effet de l'introduction de z à une translation constante selon x et y (figure 2.16 (1) et (2)) est très similaire. L'erreur d'estimation de z , élevée lorsque z est nulle ou très faible, diminue progressivement pour atteindre un minimum lorsque la valeur réelle de z atteint 0.3mm.

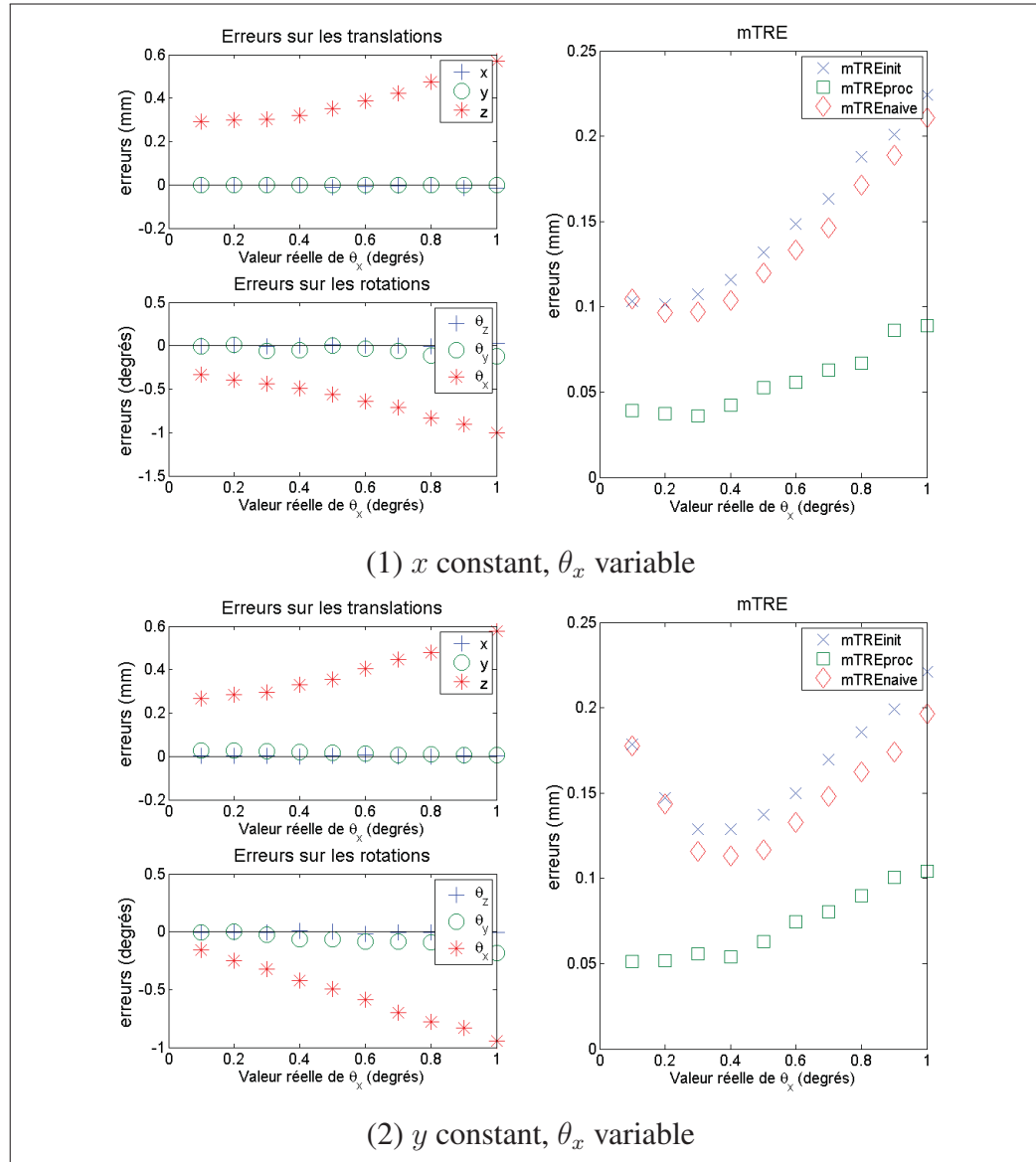


Figure 2.17 Résultats sur séquences de translations dans le plan constantes associées à des rotations hors plan variables. À gauche : Erreurs sur les composantes en translation (haut) et en rotation (bas). À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

Comme précédemment, ceci est dû à la pente de courbe de décorrélation. Notons cependant que l'erreur sur z reste plus élevée en présence de translations dans le plan, que lors de mouvements en translation élévationnelle pure. En effet, la décorrélation due au mouvement dans le plan s'ajoute indistinctivement à la décorrélation due à la translation élévationnelle, et z est sur-estimée. Pour des valeurs de z inférieures à 0.6mm, la précision des estimations des autres

composantes n'est pas affectée. Cependant, au delà de 0.6mm, on observe une dégradation significative de toutes les estimations. En effet, le recalage dans le plan est basé sur la similitude des images, et celle-ci diminue nécessairement progressivement lorsque la distance séparant les images augmente. Il est cependant intéressant de noter que l'algorithme de recalage dans le plan n'est pas affecté pour des distances inférieures à 0.6mm (l'estimation des composantes dans le plan est bonne). Les valeurs obtenues pour les différentes $mTRE$ sont proches, et elles sont faibles pour $z < 0.6\text{mm}$. Ceci indique que l'estimation de la position des points de contrôle correspond bien aux positions réelles, et à une transformation rigide également.

En présence de translations constantes selon x ou y , l'introduction de tangage est très mal interprétée par l'algorithme (figure 2.17 (1) et (2)). θ_x est estimée continuellement nulle, et en contrepartie z est progressivement sur-estimée proportionnellement à la valeur de rotation. L'observation des résultats de $mTRE$ confirme l'origine du problème : la disposition des points de contrôle est erronée ($mTRE_{init}$ augmente progressivement avec la rotation), mais cette disposition est cohérente avec une transformation rigide comportant une translation élévatoire, donc $mTRE_{proc}$ reste faible et presque constante. Comme la transformation est erronée, $mTRE_{naive}$ augmente également progressivement et prend des valeurs très proches de $mTRE_{init}$.

Les graphiques de la figure 2.18 présentent les résultats obtenus pour une translation hors plan constante et du mouvement dans le plan variable. L'augmentation progressive de la translation selon x ou y en présence d'une translation constante selon z (figure 2.18 (1) et (2)) amène les mêmes observations que les translations pures selon x ou y , c'est-à-dire une sur-estimation de z accompagnée d'une sous-estimation de θ_x , causées par la décorrélation supplémentaire due à la translation dans le plan, de valeur approximativement constante dans la profondeur. Remarquons également que les valeurs de $mTRE$ sont beaucoup plus importantes lors de l'introduction de translations selon y que selon x , ce qui est expliqué comme précédemment par la faible épaisseur de la cellule de résolution dans la direction axiale.

L'introduction de roulis à la translation élévatoire constante (figure 2.18 (3)) amène de nouvelles observations. Lorsque θ_z augmente, l'estimation de z semble d'une part être légè-

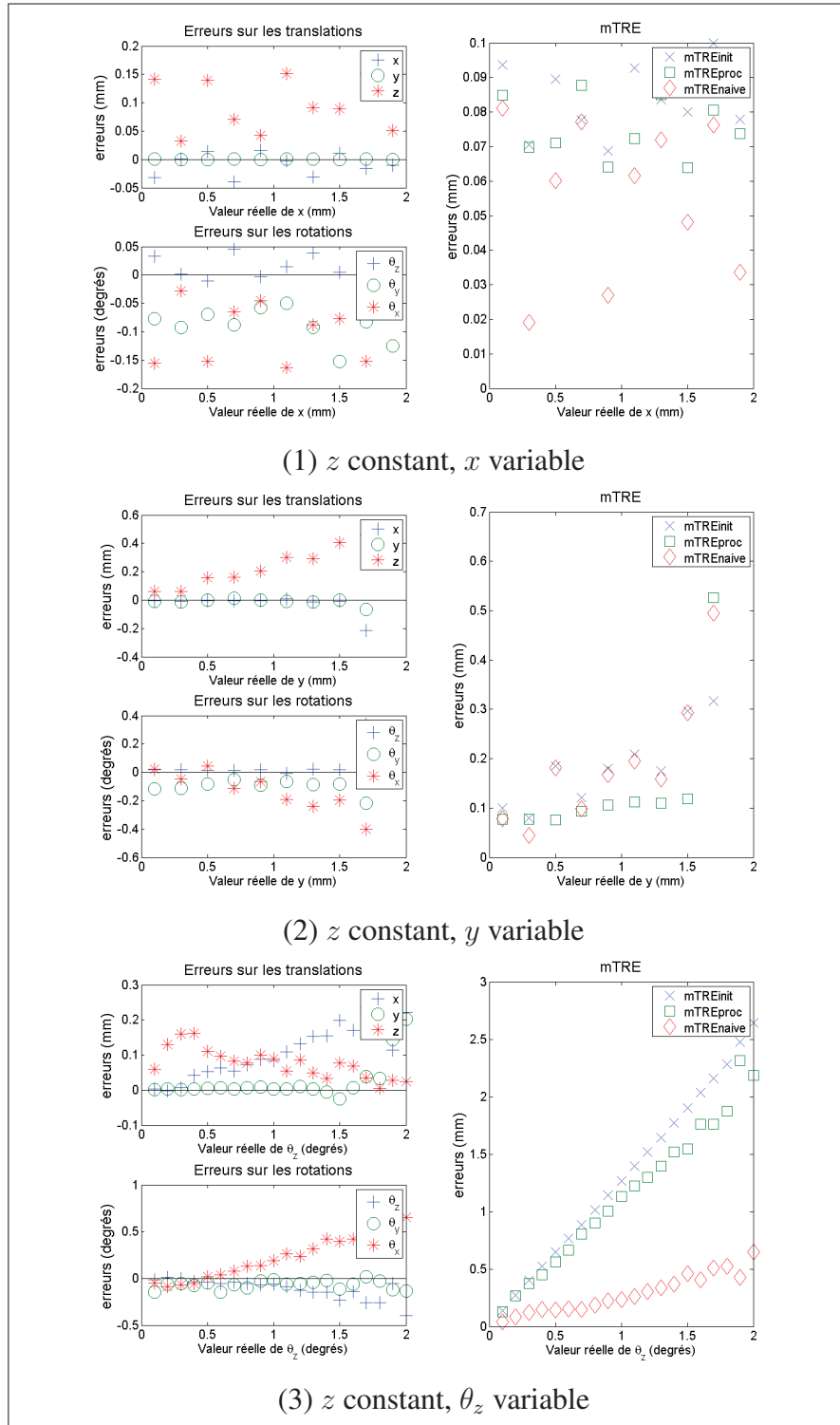


Figure 2.18 Résultats sur séquences de translations hors plan associées à du mouvement dans le plan. À gauche : Erreurs sur les composantes en translation (haut) et en rotation (bas). À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

rement améliorée. D'autre part, l'estimation de θ_z se dégrade progressivement, tout en restant raisonnable, tandis que x et θ_x sont progressivement sur-estimées. Ainsi, la réorganisation des interférences est incorrectement attribuée à l'association d'une translation latérale et d'une rotation selon θ_x . Les résultats de $mTRE$ montrent que $mTRE_{init}$ est de plus en plus élevée car la disposition des points est de plus en plus éloignée de ce qui est attendu, et $mTRE_{proc}$ augmente de la même manière car l'algorithme ne parvient pas à fournir une transformation rigide cohérente avec la disposition des points de contrôle ainsi obtenue. Cependant, $mTRE_{naive}$ augmente bien moins rapidement, ce qui suggère que la transformation estimée est finalement assez proche de la transformation réelle.

2.5 Conclusion

Le présent chapitre a permis de présenter l'implémentation de l'algorithme de recalage d'images échographiques le plus couramment utilisé pour l'application de l'échographie 3D main libre sans capteur de position. Celui-ci permet d'associer les méthodes d'estimation du mouvement dans le plan et hors plan afin d'obtenir une estimation de la transformation rigide correspondant au mouvement appliqué lors de l'acquisition d'une paire d'images.

Les expérimentations effectuées sur cet algorithme ont permis d'étudier son comportement pour différents types de mouvements complexes, au travers de l'analyse des erreurs d'estimation de chaque composante ainsi que des $mTRE$ représentant les résultats de différentes étapes de l'algorithme. Dans beaucoup de cas, la courbure de la PSF cause des erreurs sur les estimations de rotations autour de l'axe latéral et élévational (tangage, roulis). Par ailleurs, la translation selon z est très souvent sur-estimée, du fait de la présence de décorrélation causée par d'autres composantes, qui est indistinctement attribuée à la translation élévationalle. Les translations dans le plan sont quant à elles bien estimées dans la plupart des cas, tant que la valeur de la translation élévationalle n'est pas trop élevée (si elle l'est, la perte de corrélation devient trop importantes pour pouvoir mesurer la similitude). Par ailleurs, l'imposition de la contrainte de planarité des images par l'alignement de Procruste permet dans certains cas de réduire les erreurs initiales d'estimation de la position des points de contrôle ($mTRE_{naive}$ est alors inférieure à $mTRE_{init}$). Globalement, ces expérimentations ont permis de démontrer la

validité de l'hypothèse selon laquelle des erreurs systématiques sont causées par la présence de composantes autres que la composante de translation élévationnelle. En effet, des tendances ont pu être identifiées dans les évolutions d'erreurs pour chacune des combinaisons spécifiques de composantes. L'origine de ces erreurs peut souvent être expliquée par les caractéristiques des signaux échographiques. Cependant, tenter de compenser ces erreurs grâce à l'analyse de ces signaux (comme proposé par Maurice et Bertrand (1999) grâce à un filtre de restauration du speckle) s'avère peu réaliste pour des mouvements présentant six composantes, donc six sources potentielles de décorrélation. De plus, ce type d'approche nécessite les données RF (sans détection d'enveloppe). Ainsi, une approche de correction par apprentissage automatique est proposée au chapitre 3.

CHAPITRE 3

ALGORITHME DE CORRECTION PAR APPRENTISSAGE

Ce chapitre traite de la méthode choisie pour corriger les erreurs d'estimation observées et analysées dans le chapitre 2. Comme le montrent ces analyses, chaque composante du mouvement de la sonde échographique a sa propre influence sur la corrélation des images. Ainsi, quand plusieurs types de mouvements sont présents, ils influencent la corrélation simultanément et il devient difficile d'établir l'origine des décorrélations. Compte tenu de la complexité des signaux, due aux interférences multiples et à la non-linéarité introduite par la détection d'enveloppe en mode B, tenter de décrire analytiquement les effets de chaque mouvement sur la corrélation serait probablement peu efficace. Cependant, les expérimentations sur l'algorithme naïf ont permis d'observer des tendances claires dans les erreurs effectuées par cet algorithme pour certains types de mouvements : des mouvements similaires ont une influence similaire sur la corrélation, donc l'erreur causée sur l'estimation globale est similaire également. Autrement dit, «les mêmes causes ont les mêmes effets», et les erreurs dues aux multiples causes de décorrélation sont systématiques (non aléatoires), dans une certaine mesure. Ainsi, l'hypothèse pourrait être posée que la relation entre l'estimation de la transformation globale (désignée ci-après par le terme «estimation naïve»), et la transformation globale réellement appliquée à la sonde (nommée «transformation réelle») est une fonction bijective : à chaque transformation réelle (ensemble d'arrivée) correspond une seule estimation de Procruste (ensemble de départ). Dans ce cas, la connaissance de cette relation permettrait de corriger les erreurs d'estimation : pour chaque estimation naïve, la ou les transformations réelles potentiellement correspondantes pourraient être retrouvées.

Cependant, les mesures de corrélation sont nécessairement bruitées, ce qui fausse les estimations (Laporte et Arbel, 2008). Par ailleurs, même en considérant des mesures non bruitées et dans le cas où un seul type de mouvement (dans le plan ou hors plan) est présent, l'expérimentation a montré que des erreurs subsistent, inhérentes aux méthodes d'estimation. De plus, dans certains cas, par exemple pour des rotations trop importantes, les tendances observées ne sont

plus valables car la corrélation n'est pas assez importante. Ainsi il est probable que les erreurs ne soient pas entièrement systématiques et que le comportement de l'algorithme naïf ne soit pas entièrement prévisible, mais les résultats du chapitre 2 montrent que dans beaucoup de cas, les erreurs correspondant à la tendance sont bien supérieures aux erreurs aléatoires.

Afin de corriger ces erreurs, l'objectif est donc ici de parvenir à prédire la transformation réelle la plus probable étant donnée une estimation naïve, pour ainsi obtenir une estimation corrigée. Pour cela, nous proposons d'utiliser une méthode d'apprentissage statistique, présentée à la section 3.1. Deux types d'implémentations de cette méthode sont utilisées : la première est une implémentation à sortie monovariée, qui conduit à construire un modèle par composante de sortie et à rassembler les résultats de ces six modèles ensuite. La seconde est une implémentation à sortie multivariée, qui permet de construire un modèle unique et d'obtenir directement les six composantes de la transformation rigide. Ces deux implémentations sont également testées avec différents ensembles de variables d'entrée. Les expérimentations et les résultats obtenus présentés à la section 3.2 permettent de sélectionner la meilleure méthode parmi celles-ci et d'évaluer ses performances en détail.

3.1 Méthodes d'apprentissage

Acquérir la connaissance de la relation liant les estimations naïves aux transformations réelles est un problème de régression qui peut être résolu à partir d'observations. Dans notre cas, ces observations proviennent d'un ensemble de paires d'images échographiques dont les transformations réelles $\{x, y, z, \theta_z, \theta_y, \theta_x\}$ sont connues. L'algorithme naïf présenté au chapitre 2 est appliqué à ces paires d'images, afin d'obtenir les estimations naïves de la transformation rigide, dont les composantes sont notées $\{\hat{x}, \hat{y}, \hat{z}, \hat{\theta}_z, \hat{\theta}_y, \hat{\theta}_x\}$, ainsi que la valeur d'erreur $mTRE_{proc}$ telle que définie lors des expérimentations. Le schéma de la figure 3.1 présente comment ces données sont ensuite utilisées dans la construction des modèles selon les différentes méthodes proposées.

Nous disposons de N observations constituées de couples d'entrée-sortie. Les estimations naïves sont ainsi utilisées comme variables d'entrées, auxquelles est éventuellement ajoutée la valeur de $mTRE_{proc}$. Cette valeur permet effectivement d'évaluer à quel point la transforma-

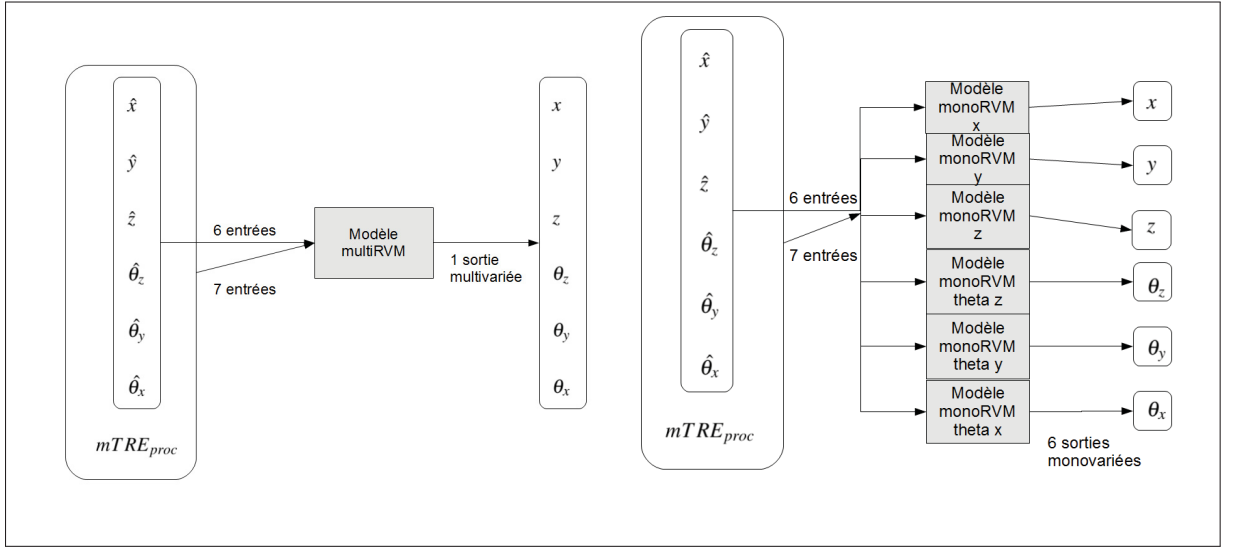


Figure 3.1 Construction des modèles selon les différentes méthodes proposées. À gauche : Modèle multivarié unique, avec six ou sept entrées. À droite : Six modèles monovariés, avec six ou sept entrées

tion rigide estimée reflète les positions de recalage de chaque fenêtre obtenues par les méthodes de recalage dans le plan et hors plan (sections 2.1 et 2.2), et pourrait potentiellement apporter des connaissances supplémentaires pour construire le modèle. L'ensemble des vecteurs de variables d'entrées pour les N observations est noté $\mathbf{A}_N = \{\mathbf{a}_n\}_{n=1}^N$, où $\mathbf{a}_i$ est composé des six composantes de l'estimation naïve pour la i ème observation $\{\hat{x}^{(i)}, \hat{y}^{(i)}, \hat{z}^{(i)}, \hat{\theta}_z^{(i)}, \hat{\theta}_y^{(i)}, \hat{\theta}_x^{(i)}\}$, et $mTRE_{proc}^{(i)}$ pour la méthode à sept variables d'entrées. Dans le cas de l'implémentation à sortie multivariée, les vecteurs de variables de sortie, constituant les «cibles» du modèle, sont les six composantes de la transformation réelle. L'ensemble des vecteurs de variables de sortie est alors noté $\mathbf{b}_N = \{\mathbf{b}_n\}_{n=1}^N$, avec $\mathbf{b}_i = \{x^{(i)}, y^{(i)}, z^{(i)}, \theta_z^{(i)}, \theta_y^{(i)}, \theta_x^{(i)}\}$. Dans le cas de l'implémentation à sortie monovariée, la variable de sortie correspond à l'une des six composantes, et un modèle est ainsi construit pour chaque composante. L'ensemble des variables de sortie est alors noté $\mathbf{b}_N = \{b_n\}_{n=1}^N$, b_i prenant alternativement la valeur $x^{(i)}, y^{(i)}, z^{(i)}, \theta_z^{(i)}, \theta_y^{(i)}$ ou $\theta_x^{(i)}$.

La fonction recherchée, supposée représenter la relation entrée-sortie, est notée respectivement f_{multi6} pour le modèle à six variables d'entrée et sortie multivariée, f_{multi7} pour le modèle à sept variables d'entrée et sortie multivariée, f_{mono6} pour le modèle à six variables d'entrée et sortie monovariée, où b représente l'une des six composantes de la transformation rigide,

et finalement f_{mono7} pour le modèle à sept variables d'entrée et sortie monovariée. Pour plus de simplicité, la fonction nommée f dans le développement mathématique des modèles des sections suivantes désigne l'une des quatre fonctions décrites précédemment, dépendamment du contexte.

La méthode d'apprentissage choisie est la machine à vecteurs de pertinence (RVM, pour *Relevance Vector Machine*). Dans un premier temps, elle est présentée dans sa forme monovariée, qui constitue la forme la plus simple, dans la partie 3.1.1. Son extension à des sorties multivariées est présentée dans la partie 3.1.2.

3.1.1 Apprentissage par RVM monovarié

La régression par apprentissage supervisé consiste à utiliser un ensemble de couples entrée-sortie connus pour «entraîner» un modèle représentant la dépendance entre les entrées et les sorties. Le problème réside dans la modélisation de la fonction $f(\mathbf{a})$, supposée représenter la relation entre $\mathbf{a}$ et b . On considère cependant la présence de bruit, de sorte que chaque cible b_i est représentative du vrai modèle $f(\mathbf{a}_i)$, avec l'ajout de bruit ϵ :

$$b = f(\mathbf{a}) + \epsilon. \quad (3.1)$$

La composante de bruit est modélisée par une distribution gaussienne de moyenne zéro et de variance σ^2 :

$$p(\epsilon|\sigma^2) = N(0, \sigma^2). \quad (3.2)$$

La forme la plus simple de régression est la régression linéaire, qui consiste à exprimer le modèle sous la forme $f(\mathbf{a}, \mathbf{w}) = \mathbf{w}\mathbf{a}$, où $\mathbf{w} = (w_0, w_1, \dots, w_N)$ est le vecteur des poids ajustables du modèle. Nous considérons ici des modèles linéaires généralisés, c'est-à-dire que f est une somme pondérée linéairement de N fonctions de base Φ_n (qui elles peuvent être non linéaires). Le modèle prend ainsi la forme :

$$f(\mathbf{a}, \mathbf{w}) = w_0 + \sum_{i=1}^N w_i \Phi_i(\mathbf{a}). \quad (3.3)$$

La résolution passe alors par la recherche des valeurs optimales pour $\mathbf{w}$. La méthode la plus commune pour y parvenir est la méthode des moindres carrés. Il s'agit d'une approche fréquentiste, qui considère que les informations obtenues par les données sont suffisantes pour estimer correctement les poids. Cependant, compte tenu des erreurs dues à l'inexactitude des mesures, et des erreurs d'estimations initiales sur lesquelles est basée l'estimation naïve, il est plus adapté d'employer une approche bayésienne. Dans ce type d'approche, des informations supplémentaires sont ajoutées aux données sous la forme d'une probabilité *a priori* sur les poids : $p(\mathbf{w})$. Celle-ci est combinée avec la *vraisemblance* $p(\mathbf{b}_N | \mathbf{w}, \mathbf{A}_N)$ selon la loi de Bayes, pour obtenir une probabilité *a posteriori* sur les poids :

$$p(\mathbf{w} | \mathbf{A}_N, \mathbf{b}_N) = \frac{p(\mathbf{w})p(\mathbf{b}_N | \mathbf{w}, \mathbf{A}_N)}{p(\mathbf{b}_N | \mathbf{A}_N)}.$$

Une fois cette expression obtenue, la prédiction pour une nouvelle valeur cible b' , $p(b' | \mathbf{a}', \mathbf{A}_N, \mathbf{b}_N)$ peut finalement être calculée.

Une caractéristique importante à prendre en compte lors de la résolution de ce problème est la capacité de la méthode à générer des modèles parcimonieux. En effet, si l'on choisit un modèle trop complexe (c'est-à-dire comportant un trop grand nombre de paramètres), il y a un risque important de «sur-spécifier» le modèle pour les valeurs connues sur lesquelles il est construit, et qu'il ne soit plus du tout représentatif de la relation réelle qui existe entre les variables. Il faut que le modèle soit le plus précis possible, non pas pour les valeurs d'observation, mais pour les valeurs prédites sans connaissance préalable. Par ailleurs, le caractère parcimonieux assure également que le modèle analytique soit calculable (Bishop, 2006).

Nous avons choisi la méthode stochastique d'apprentissage par RVM. Celle-ci, développée par Tipping (2001), est une méthode d'inférence bayésienne apparentée aux méthodes par Processus Gaussiens (Quiñonero-Candela et Rasmussen, 2005), incluant directement la contrainte de parcimonie dans son mode de construction. Elle permet ainsi très bien de contrôler la complexité du modèle et donc d'éviter la sur-spécification. Cette méthode est très similaire aux machines à vecteurs de support (SVM pour *Support Vector Machine*), une méthode d'appren-

tissage très populaire (Vapnik, 1999). Cependant, les SVM constituent une méthode non probabiliste et fournissent un résultat «ferme», contrairement aux RVM qui permettent d'obtenir une mesure d'incertitude sur le résultat obtenu. C'est en particulier cette différence qui a motivé la préférence pour les RVM dans ce projet. Le fonctionnement de cette méthode est décrit ci-après, suivant de près le développement et la notation exposés dans Tipping (2001).

D'après les équations 3.1 et 3.2, la distribution de chaque cible est de la forme suivante :

$$p(b_i|\mathbf{a}_i, \mathbf{w}, \sigma^2) = N(f(\mathbf{a}_i, \mathbf{w}), \sigma^2).$$

En supposant que tous les exemples sont statistiquement indépendants, la vraisemblance de l'ensemble des données est la suivante :

$$\begin{aligned} p(\mathbf{b}_N|\mathbf{A}_N, \mathbf{w}, \sigma^2) &= \prod p(b_i|\mathbf{a}_i, \mathbf{w}, \sigma^2) \\ &= \prod (2\pi\sigma^2)^{1/2} \exp \left[-\frac{(b_i - f(\mathbf{a}_i, \mathbf{w}))^2}{2\sigma^2} \right]. \end{aligned}$$

Rappelons que nous choisissons un modèle linéaire (équation 3.3). Les fonctions de base Φ_n peuvent être de tout type, mais il est usuel de choisir des fonctions noyau centrées sur chaque observation.

Une fonction noyau est une fonction non-négative, intégrable et à valeurs réelles, souvent notée K , qui doit vérifier les deux conditions suivantes :

- $\int_{-\infty}^{+\infty} K(u)du = 1$;
- $K(-u) = K(u)$ pour tout u .

Les fonctions de base sont rassemblées dans la matrice Φ de taille $N \times (N + 1)$, avec $\Phi_{nm} = \Phi_m(\mathbf{a}_n)$. La vraisemblance s'exprime alors :

$$p(\mathbf{b}_N|\mathbf{A}_N, \mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp \left[-\frac{1}{2\sigma^2} \|\mathbf{b} - \Phi\mathbf{w}\|^2 \right].$$

La distribution *a priori* sur les valeurs prises par $\mathbf{w}$ est choisie de manière à restreindre la complexité du modèle. Un choix conventionnel est une distribution gaussienne centrée en zéro, ce qui permet d'exprimer une préférence pour des modèles plus lisses. Par ailleurs, afin d'assurer le caractère parcimonieux du modèle, une variance spécifique est associée à chaque poids, pour pouvoir contrôler la force de l'*a priori* de chacun indépendamment :

$$p(\mathbf{w}|\alpha_n) = N(\mathbf{0}, \alpha_n^{-1}\mathbf{I}),$$

$$\text{donc } p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=0}^N N(\mathbf{0}, \alpha_i^{-1}\mathbf{I}).$$

Le vecteur $\boldsymbol{\alpha}$ contient ainsi les $N + 1$ hyperparamètres α_n représentant l'inverse de la variance pour chaque poids.

La distribution *a posteriori* sur les poids peut ensuite être exprimée grâce à la loi de Bayes. Le développement, disponible dans les travaux de Tipping (2001), permet d'obtenir une distribution gaussienne multivariée :

$$p(\mathbf{w}|\mathbf{b}_N, \mathbf{A}_N, \boldsymbol{\alpha}, \sigma^2) = N(\boldsymbol{\mu}, \Sigma),$$

dont la moyenne et la matrice de covariance sont données par :

$$\boldsymbol{\mu} = \sigma^{-2}\Sigma\boldsymbol{\Phi}^T\mathbf{b}_N,$$

$$\Sigma = (\sigma^{-2}\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \mathbf{H})^{-1},$$

et où les hyperparamètres $\boldsymbol{\alpha}$ ont été rassemblés dans la matrice diagonale $\mathbf{H}$:

$$\mathbf{H} = \begin{pmatrix} \alpha_0 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_N \end{pmatrix}.$$

Un résultat très important à noter concernant ces équations est que si α_n tend vers l'infini, alors la moyenne μ_n de la distribution du poids correspondant tend vers zéro, ce qui a pour conséquence de faire tendre le poids lui-même vers zéro. Le vecteur correspondant n'est ainsi pas pris en compte, et la complexité du problème peut être réduite, comme il sera explicité plus loin.

Finalement, la *vraisemblance marginale* $p(\mathbf{b}_N|\boldsymbol{\alpha}, \sigma^2)$ des cibles observées peut être obtenue par intégration sur les poids. Ainsi, la probabilité d'une valeur cible donnée sera moyennée sur toutes les valeurs possibles de $\mathbf{w}$. Ceci permet de prendre en compte l'incertitude sur les poids.

$$\begin{aligned} p(\mathbf{b}_N|\boldsymbol{\alpha}, \sigma^2) &= \int p(\mathbf{b}|\mathbf{w}, \sigma^2)p(\mathbf{w}|\boldsymbol{\alpha})d\mathbf{w} \\ &= (2\pi)^{-N/2}|\sigma^2\mathbf{I} + \boldsymbol{\Phi}\mathbf{H}^{-1}\boldsymbol{\Phi}^T|^{-1/2} \exp[-\frac{1}{2}\mathbf{b}^T(\sigma^2\mathbf{I} + \boldsymbol{\Phi}\mathbf{H}^{-1}\boldsymbol{\Phi}^T)^{-1}\mathbf{b}]. \end{aligned} \quad (3.4)$$

Afin de trouver les valeurs optimales des hyperparamètres, il faut alors maximiser l'expression de l'équation 3.4 par rapport à $\boldsymbol{\alpha}$ et σ^2 . Cette étape d'optimisation peut être réalisée itérativement à l'aide des assignations suivantes :

$$\begin{aligned} \alpha_i &\leftarrow \frac{1 - \alpha_i \Sigma_{ii}}{\mu_i^2} = \frac{\gamma_i}{\mu_i^2}, \text{ avec } \gamma_i = 1 - \alpha_i \Sigma_{ii}, \\ \text{et } \sigma^2 &\leftarrow \frac{\|\mathbf{b} - \boldsymbol{\Phi}\boldsymbol{\mu}\|^2}{N - \sum_i \gamma_i}. \end{aligned}$$

Il est pour cela nécessaire de fixer des valeurs initiales pour les $\{\alpha_i\}_{i=0}^N$ et σ^2 , puis de calculer $\boldsymbol{\mu}$ et Σ , et ensuite de recalculer $\boldsymbol{\alpha}$ et σ^2 à partir des expressions précédentes. Ce processus doit être répété jusqu'à atteindre un critère de convergence fixé.

Les valeurs finales peuvent ensuite être utilisées pour évaluer la distribution prédictive de b' pour une nouvelle valeur d'entrée $\mathbf{a}'$:

$$\begin{aligned} p(b'|\mathbf{a}', \mathbf{b}_N, \mathbf{A}_N, \boldsymbol{\alpha}, \sigma^2) &= \int p(b|\mathbf{w}, \sigma^2)p(\mathbf{w}|\mathbf{a}'\boldsymbol{\alpha}, \sigma^2)d\mathbf{w} \\ &= N(\mu^T\boldsymbol{\Phi}(\mathbf{a}'), \boldsymbol{\Phi}(\mathbf{a}')^T\Sigma\boldsymbol{\Phi}(\mathbf{a}')). \end{aligned} \quad (3.5)$$

Ainsi, l'estimation de b' est la moyenne de cette distribution, $\boldsymbol{\mu}^T \Phi(\mathbf{a}')$. La confiance dans cette estimation est déterminée par la variance de la distribution : $\Phi(\mathbf{a}')^T \Sigma \Phi(\mathbf{a}')$.

Lors de cette procédure, il a été démontré par Tipping (2001) qu'un grand nombre de paramètres α_i vont tendre vers l'infini. Comme remarqué plus haut, cela a pour conséquence de faire tendre la valeur de μ_i correspondante vers zéro. Or, comme $p(\mathbf{w}|\mathbf{t}, \boldsymbol{\alpha}, \sigma^2) = N(\boldsymbol{\mu}, \Sigma)$, cela signifie que les poids correspondants tendent également vers zéro. Les fonctions de base associées $\phi(\mathbf{a}_i)$ peuvent donc être retirées de la matrice Φ à chaque itération, la rendant ainsi progressivement parcimonieuse.

Le principe des RVM réside donc ici : seuls les vecteurs *pertinents*, correspondant aux observations pertinentes, sont conservés, les autres sont écartés au fur et à mesure des itérations.

Concrètement, la construction des six modèles monovariés peut ainsi être réalisée par les étapes suivantes, où (k) correspond à la composante utilisée comme variable de sortie :

- construire la matrice Φ à partir d'une fonction noyau gaussienne appliquée sur les données $\mathbf{A}_N$;
- établir un critère de convergence pour $\{\alpha_i\}$ et σ^2 , ainsi qu'un critère d'«infinité» déterminant la valeur à partir de laquelle α_i est considéré tendre vers l'infini ;
- choisir les valeurs initiales de $\{\alpha_i^{(k)}\}_{i=0}^N$ et $\sigma^{2(k)}$;
- calculer $\boldsymbol{\mu}^{(k)} = \sigma^{-2} \Sigma^{(k)} \Phi^T \mathbf{b}$ et $\Sigma^{(k)} = (\sigma^{(k)-2} \Phi^T \Phi + \mathbf{H}^{(k)})^{-1}$;
- mettre à jour $\alpha_i^{(k)} = \frac{1 - \alpha_i^{(k)} \Sigma_{ii}^{(k)}}{\mu_i^{(k)2}} = \frac{\gamma_i^{(k)}}{\mu_i^{(k)2}}$ et $\sigma^{(k)2} = \frac{\|\mathbf{b} - \Phi \boldsymbol{\mu}^{(k)}\|^2}{N - \sum_i \gamma_i^{(k)}}$;
- retirer de Φ les fonctions $\Phi(\mathbf{a}_i)$ pour lesquelles $\alpha_i^{(k)}$ tend vers l'infini ;
- réitérer les étapes (4) à (6) jusqu'à convergence.

Pour une nouvelle estimation naïve $\mathbf{a}'$, l'estimation corrigée de chaque composante est ensuite obtenue en calculant la moyenne de la distribution 3.5 :

$$k^* = \boldsymbol{\mu}_{(k)}^T \Phi_{(k)}(\mathbf{a}').$$

3.1.2 RVM à sortie multivariée

La construction d'un modèle par composante ne permet pas de tenir compte de la dépendance entre les composantes de la transformation rigide. Pour ce faire, il faudrait construire un modèle unique à six variables de sortie.

Thayananthan *et al.* (2006) ont proposé une extension des RVM pour les sorties multivariées. Leur solution a l'avantage d'utiliser un seul ensemble de vecteurs de pertinence pour toutes les dimensions de sortie. Pour cela, la vraisemblance est exprimée comme un produit de gaussiennes représentant la distribution des poids pour chaque dimension de sortie, permettant ainsi de choisir les hyperparamètres pour toutes les dimensions lors d'une seule étape d'optimisation.

Dans ce cadre, notons maintenant l'ensemble des vecteurs de sortie $\mathbf{B}_N = \{\mathbf{b}_n\}_{n=1}^N$. Chaque vecteur $\mathbf{b}$ comporte P dimensions (dans notre cas $P = 6$). La fonction recherchée est paramétrée par la matrice $\mathbf{W}$ de taille $P \times M$. Le modèle linéaire est alors défini comme suit :

$$\mathbf{b} = \sum_{m=1}^M \mathbf{w}_m \Phi_m(\mathbf{a}) + \boldsymbol{\epsilon},$$

où $\mathbf{w}_m$ est le m ième vecteur de poids de $\mathbf{W}$, et $\boldsymbol{\epsilon}$ est le vecteur de bruit gaussien de dimension P , de moyenne zéro et de matrice de covariance $\mathbf{S}$, qui est une matrice diagonale de taille $P \times P$, contenant les variances σ_r^2 associés à chaque variable de sortie.

Comme définie précédemment, $\mathbf{H}$ est la matrice diagonale contenant les hyperparamètres α : $\mathbf{H} = \text{diag}(\alpha_0, \alpha_2, \dots, \alpha_N)$. La distribution *a priori* des poids est :

$$p(\mathbf{W}|\mathbf{H}^2) = \prod_{r=1}^P \prod_{i=1}^N p(w_{ri}|\alpha_i) = \prod_{r=1}^P p(\mathbf{w}_r|\mathbf{H}^2) = \prod_{r=1}^P N(0, \mathbf{H}^2),$$

où w_{ri} est l'élément (r, i) de $\mathbf{W}$, et $\mathbf{w}_r$ est le vecteur de poids pour la r^{ime} variable de sortie tel que $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_r, \dots, \mathbf{w}_P]^T$.

La forme diagonale de $\mathbf{S}$ permet d'écrire la vraisemblance comme un produit de Gaussiennes séparées pour les vecteurs de poids de chaque dimension de sortie :

$$p(\mathbf{B}_N|\mathbf{A}_N, \mathbf{W}, \mathbf{S}) = \prod_{r=1}^P p(\mathbf{B}_N(r)|\mathbf{A}_N, \mathbf{w}_r, \sigma_r^2) = \prod_{r=1}^P N(\mathbf{w}_r \Phi, \sigma_r^2).$$

Ainsi, la distribution *a posteriori* sur les poids peut être exprimée comme un produit de Gaussiennes séparées pour les vecteurs de poids de chaque dimension de sortie, grâce à la loi de Bayes :

$$\begin{aligned} p(\mathbf{W}|\mathbf{B}_N, S, H^2) &\propto p(\mathbf{B}_N|\mathbf{W}, \mathbf{S})p(\mathbf{W}|\mathbf{H}^2) \\ &\propto \prod_{r=1}^P N(\mathbf{w}_r|\boldsymbol{\mu}_r, \Sigma_r), \end{aligned}$$

où $\boldsymbol{\mu}_r$ et Σ_r s'expriment pour chaque dimension de sortie de manière identique à la formulation monovariée :

$$\begin{aligned} \boldsymbol{\mu}_r &= \sigma_r^{-2} \Sigma_r \Phi^T \mathbf{B}_N(r), \\ \Sigma_r &= (\sigma_r^{-2} \Phi^T \Phi + \mathbf{H})^{-1}. \end{aligned}$$

Par la suite, il suffit d'intégrer la vraisemblance sur les poids pour obtenir la vraisemblance marginale comme pour le cas monovarié, puis les hyperparamètres peuvent être obtenus par optimisation.

3.2 Expérimentation sur la correction de l'erreur

Les expérimentations sur l'algorithme avec apprentissage ont été réalisées sur des données synthétiques obtenues grâce au logiciel Field II. Notamment, les séquences présentées précédemment pour l'expérimentation sur l'algorithme naïf ont été réutilisées dans le but d'effectuer des comparaisons et de tester l'efficacité de la correction proposée dans ce chapitre.

Pour la construction des modèles, l'implémentation de Tipping (2001)¹ a été utilisée pour le modèle à sortie monovariée, et l'implémentation de Thayanathan *et al.* (2006)² a été utilisée pour le modèle à sortie multivariée.

Tout d'abord, 400 paires d'images ont été synthétisées avec un mouvement rigide entièrement libre et aléatoire, à six degrés de liberté. La translation selon z suit une distribution exponentielle de moyenne 0.15mm, tandis que les translations selon x et y suivent une distribution gaussienne centrée en zéro et d'écart type 0.15mm. Les rotations selon θ_z , θ_y et θ_x suivent chacune une distribution gaussienne centrée en zéro et d'écart type 0.10° . Ces distributions permettent de représenter un mouvement typique dans le cadre d'une acquisition pour la reconstruction 3D. Parmi ces séquences, celles présentant des intersections ont été détectées et éliminées afin de conserver uniquement des séquences respectant les hypothèses du contexte d'étude. 290 paires d'images ont ainsi été conservées pour la construction des modèles.

Les mesures d'erreurs utilisées ici sont de même type que la mesure $mTRE_{naive}$, et sont définies comme la moyenne des distances entre les centres p_i transformés selon la transformation estimée par l'algorithme correspondant (T_{multi6} , T_{multi7} , T_{mono6} et T_{mono7} respectivement) et ces centres transformés par la transformation réelle :

1. Disponible en ligne : <http://www.miketipping.com>

2. Disponible en ligne : <http://mi.eng.cam.ac.uk/at315/MVRVM.htm>

$$\begin{aligned}
mTRE_{multi6} &= \frac{1}{n} \sum_{i=1}^n ||T_{multi6}(p_i) - T_r(p_i)||, \\
mTRE_{multi7} &= \frac{1}{n} \sum_{i=1}^n ||T_{multi7}(p_i) - T_r(p_i)||, \\
mTRE_{mono6} &= \frac{1}{n} \sum_{i=1}^n ||T_{mono6}(p_i) - T_r(p_i)||, \\
mTRE_{mono7} &= \frac{1}{n} \sum_{i=1}^n ||T_{mono7}(p_i) - T_r(p_i)||.
\end{aligned}$$

3.2.1 Comparaison des modèles

Dans le but de comparer les différents modèles d'apprentissage présentés à la section 3.1, une validation croisée a été réalisée pour chacun d'entre eux. La validation croisée est une méthode d'estimation de la fiabilité d'un modèle. La base de données est divisée en c sous-ensembles, puis l'un d'entre eux est utilisé comme ensemble de validation, et les $c - 1$ autres sous-ensembles constituent la base de données d'apprentissage sur laquelle le modèle est entraîné. Une fois le modèle construit, il est testé sur l'ensemble de validation, puis l'erreur de validation peut être estimée pour chaque estimation en calculant la $mTRE$. Cette opération est répétée en sélectionnant un autre sous-ensemble parmi les $c - 1$ qui n'ont pas encore été utilisés pour la validation du modèle. L'opération est donc répétée c fois pour qu'en fin de compte chaque sous-ensemble ait été utilisé exactement une fois comme ensemble de validation. La distribution des c $mTRE$ moyennes permet finalement de représenter les résultats de validation de la base de données choisie. Une validation croisée à cinq sous-ensembles a ainsi été réalisée sur la base de données de 290 séquences libres. Les distributions des $mTRE$ obtenues pour chaque méthode (y compris l'algorithme naïf), sont présentées au diagramme de la figure 3.2.

D'après ces résultats, la correction réalisée par les modèles à sortie monovariée n'amène que très peu d'améliorations, que ce soit en termes de moyenne ou de variabilité de l'erreur. Ces modèles ne permettent donc pas de réaliser des corrections notables. Les résultats obtenus

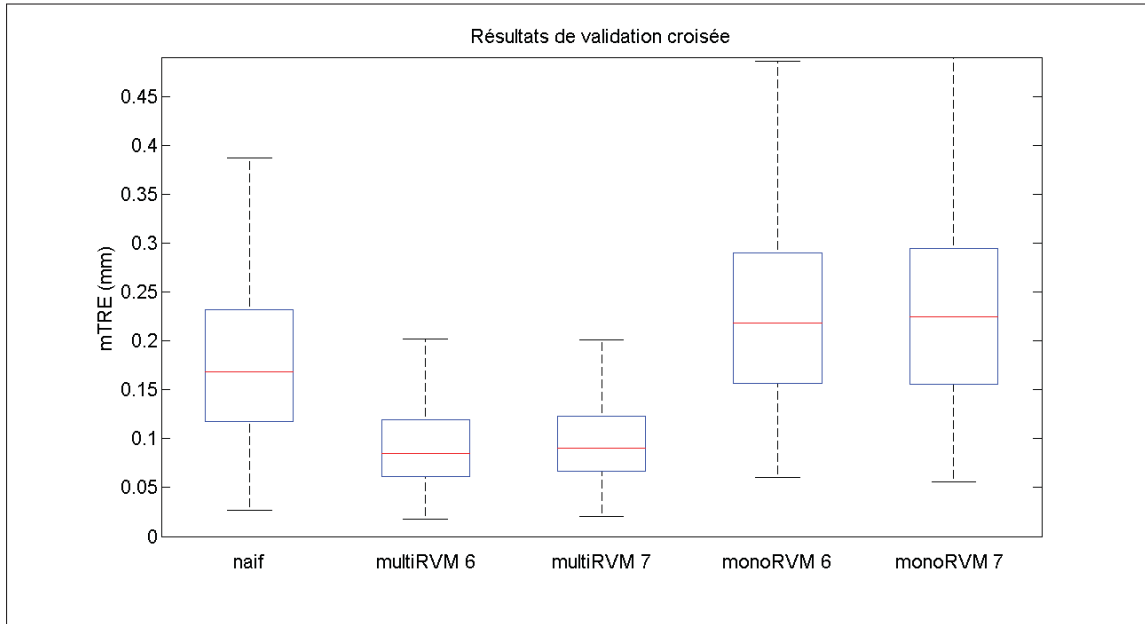


Figure 3.2 Comparaison de la distribution des erreurs de validation obtenues pour chaque méthode sur la base de données de 400 séquences libres

avec les modèles à sortie multivariée montrent quant à eux une amélioration importante. La moyenne, l'écart type, et les valeurs maximales d'erreurs sont visiblement réduites. Un test t pour différence paire a permis de démontrer l'amélioration statistiquement significative de la moyenne ($p < 0.001$) des résultats du modèle multivarié à six composantes d'entrée par rapport aux résultats de l'algorithme naïf. Contrairement à ce que suggère l'intuition, l'ajout de la variable supplémentaire ne rend pas le modèle plus performant. Ceci signifie que l'information fournie par la valeur de $mTRE_{proc}$ ne permet pas de reconnaître la provenance des erreurs.

Dans ce qui suit nous nous intéressons donc plus en détail au comportement et à la performance du modèle obtenu par RVM à sortie multivariée à six composantes d'entrée. Pour plus de simplicité, la méthode est simplement dénommée «RVM» dans la suite du chapitre. Afin d'étudier plus précisément les résultats obtenus sur les séquences aléatoires, la distribution des erreurs sur chaque composante, avec et sans correction par RVM, est représentée sur la figure 3.3.

Les résultats montrent que l'amélioration observée sur la $mTRE$ est principalement due à une amélioration de l'estimation de la translation selon z . Comme observé lors des expéri-

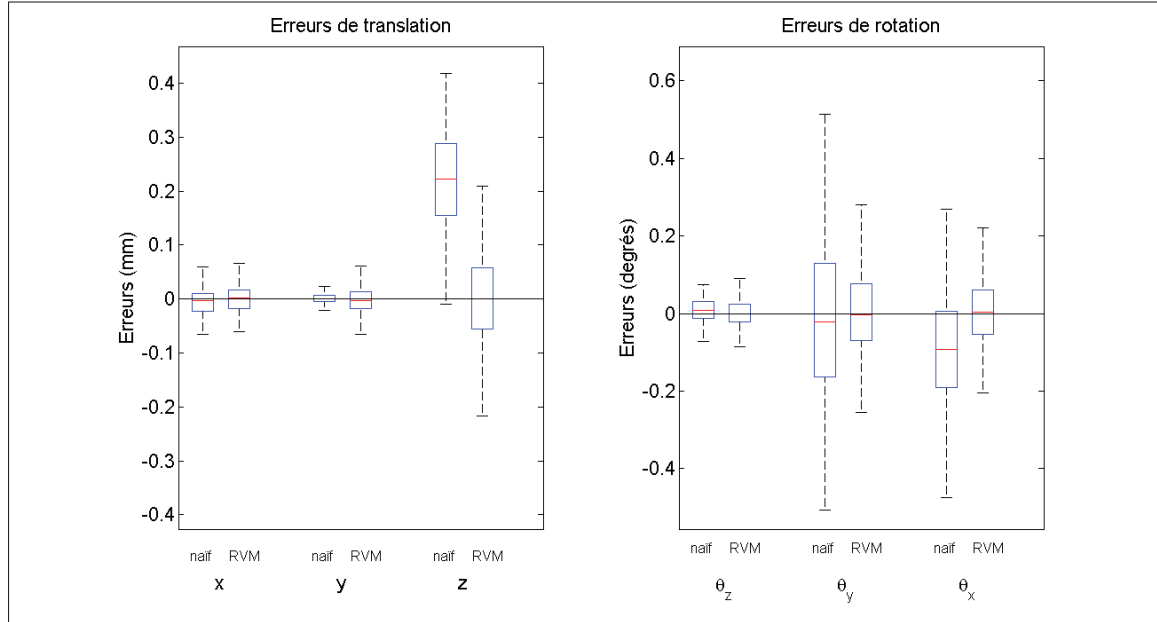


Figure 3.3 Comparaison des distributions des erreurs de validation sur chaque composante pour l'estimation naïve et RVM

mentations du chapitre 2, l'estimation naïve tend à sur-estimer z car la décorrélation due aux autres composantes est attribuée au déplacement élévational. La moyenne des erreurs sur z avec cette méthode est ainsi supérieure à 0.2mm. Après correction, cette moyenne est rendue presque nulle. La variance reste cependant élevée et de nombreux cas où la translation est sous-estimée apparaissent. Ce phénomène sera approfondi par la suite. Par ailleurs, l'estimation de θ_y et θ_x est également améliorée : la moyenne des erreurs par l'estimation naïve est nulle pour θ_y , mais la variance est importante, et celle-ci est visiblement diminuée après correction. Pour θ_x , la moyenne des erreurs, relativement faible mais non nulle lors de l'estimation naïve, est ramenée à zéro, et la variance est également réduite.

Afin d'observer plus précisément le comportement de l'algorithme de correction pour l'estimation de z , qui semble être le facteur principal de la diminution des erreurs, le graphique de la figure 3.4 compare les estimés de z obtenus avec chaque méthode, par rapport à la valeur réelle de z . Les séquences ont été triées selon z croissant, dans le but d'étudier l'évolution des performances lorsque z augmente.

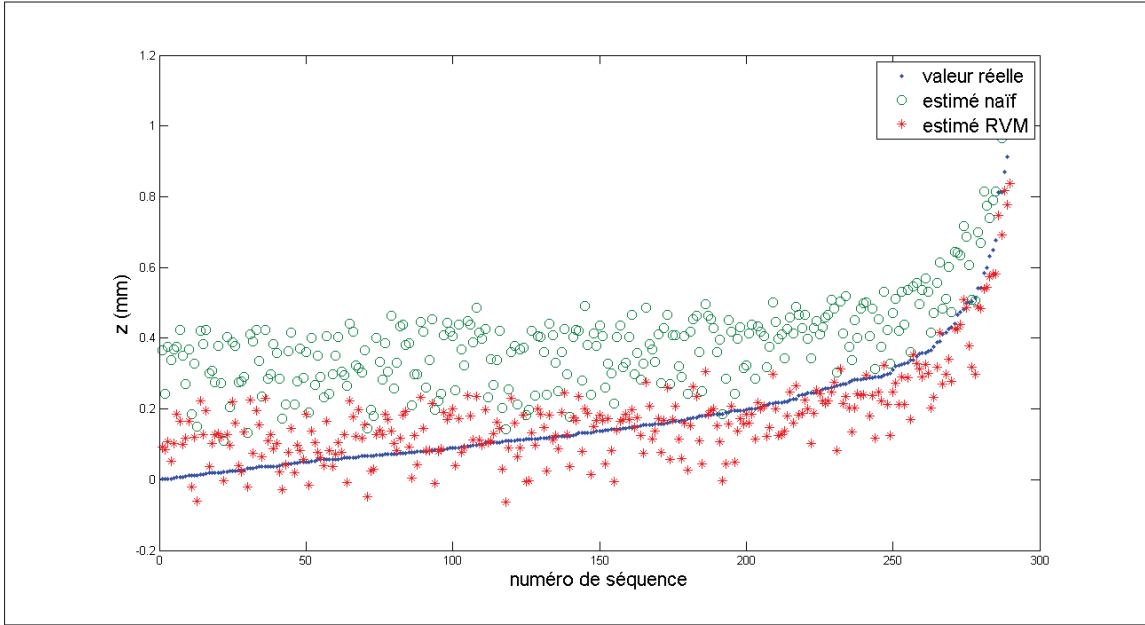


Figure 3.4 Comparaison des estimés naïfs et RVM de z pour les 290 séquences libres, pour des valeurs réelles croissantes de z

D'après ce graphique, l'amélioration est plus importante pour des valeurs faibles de z . Lorsque z augmente, l'erreur moyenne de l'estimation naïve apparaît diminuer, tout en restant toujours positive. L'estimation par RVM devient quant à elle généralement inférieure à la valeur réelle : z est sous-estimée par l'algorithme de correction, pour des valeurs importantes de z .

Pour une étude plus précise, les séquences à mouvement restreint présentées pour l'expérimentation sur l'algorithme naïf sont testées dans la partie suivante.

3.2.2 Tests sur séquences avec mouvement restreint

Afin de tester l'algorithme de correction sur diverses séquences, le modèle est d'abord entraîné sur les 290 séquences à mouvement aléatoire. Les tests, effectués sur des séquences à mouvement restreint, sont évalués par comparaison des valeurs de $mTRE_{naïve}$ et $mTRE_{RVM}$, en fonction des valeurs réelles des composantes. Il est cependant important de noter qu'étant donnée la base de données d'apprentissage, le modèle n'est pas entraîné pour corriger des cas de mouvements restreints, et des erreurs sont attendues pour ce type de cas. Il n'est pas souhaitable de réaliser l'apprentissage sur ces séquences particulières puisqu'elles ne sont pas

réalistes lors d'un mouvement libre, et l'expérimentation sur celles-ci est seulement menée à des fins d'observations.

Tout d'abord, les graphiques des figures 3.5 et 3.6 montrent les résultats sur des séquences de mouvement pur. Les résultats pour les mouvements dans le plan (translations selon x et y et rotation selon θ_z , figure 3.5 (1), (2) et (3)) sont excellents puisque la $mTRE$ est grandement réduite pour toutes valeurs de y et pour des valeurs inférieures à 0.6mm pour x . Notons cependant que la correction augmente les erreurs effectuées sur la composante selon laquelle le mouvement est effectuée. Par exemple, dans le cas de translations pures selon x , l'erreur d'estimation sur x est plus importante après correction. Ceci est probablement dû au fait que la présence de mouvements purs dans la base de données d'apprentissage est faible, donc la correction a tendance à ramener également ces estimations vers zéro. Les résultats sur les estimations de z et θ_x sont par ailleurs très positifs. Rappelons que les erreurs importantes sur z et θ_x apparaissant lors de l'estimation naïve sont dues à la sur-estimation de la translation élévationnelle, augmentant en profondeur (car la taille de la cellule de résolution augmente avec la profondeur, voir le chapitre 2). Ces erreurs sont très bien corrigées pour les mouvements purs dans le plan.

Les résultats pour la translation élévationnelle pure (figure 3.6 (1)) montrent cependant que le modèle tend à «sur-corriger» l'estimation de z lorsque z augmente, ce qui conduit à une augmentation conséquente de la $mTRE$ (qui reste tout de même assez faible) pour de grandes valeurs de z , alors que la performance de l'estimation naïve est très bonne et constante. En effet, lorsque toutes les autres composantes sont nulles, l'algorithme naïf attribue avec raison toute la décorrélation à la translation élévationnelle. Cependant, le modèle de correction est entraîné sur des cas où les autres composantes sont rarement nulles puisque le mouvement est libre, et où z est donc presque toujours sur-estimée. Le modèle diminue donc systématiquement la valeur estimée de z . Lorsqu'il est testé sur des séquences où l'estimation naïve de z est correcte, une sous-estimation est donc systématiquement observée. Des erreurs relativement importantes sur l'estimation naïve de θ_y avaient également été notées, et nous constatons que ces erreurs sont quant à elles très bien corrigées.

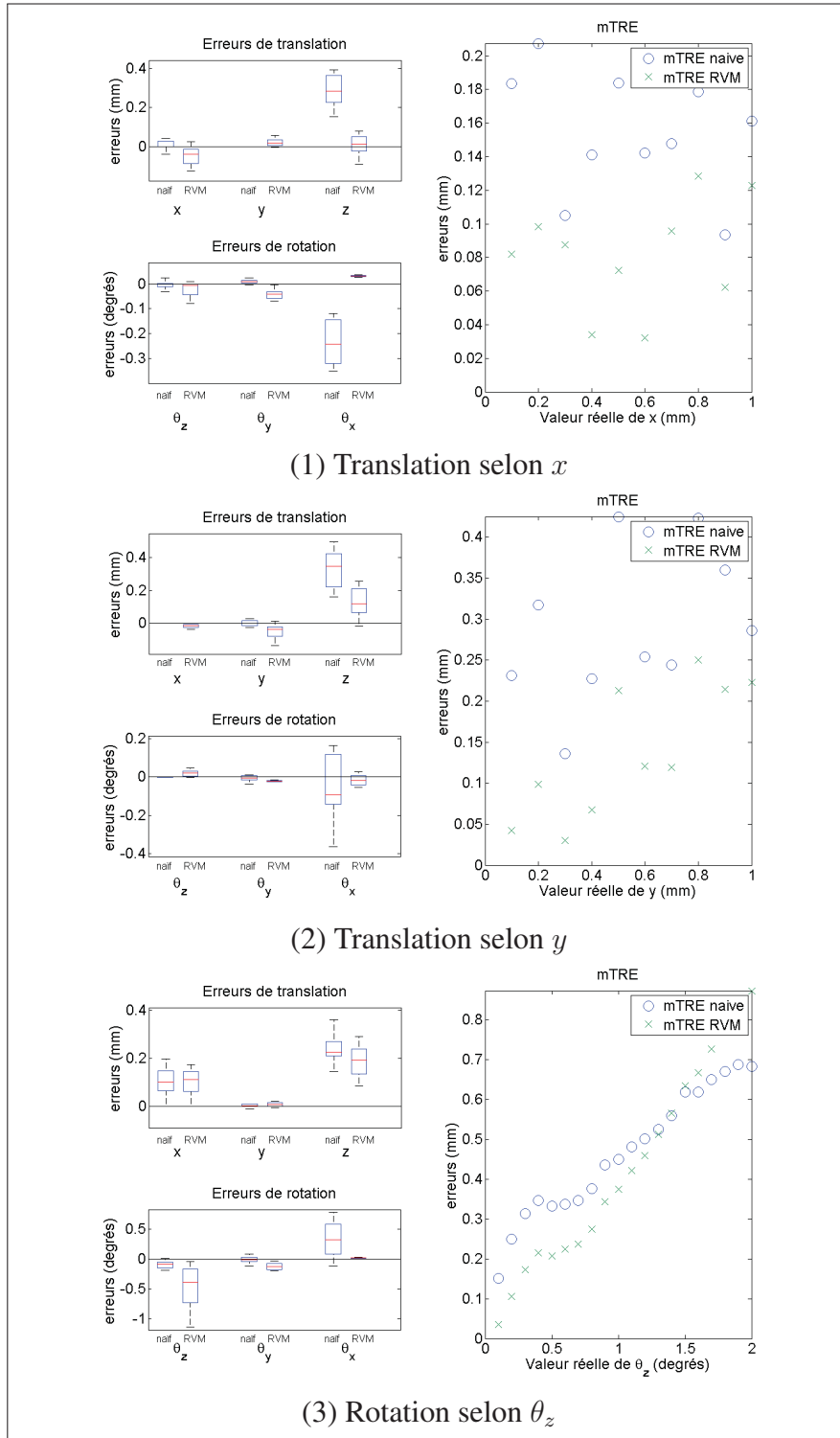


Figure 3.5 Résultats sur séquences de mouvement pur dans le plan. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante en mouvement

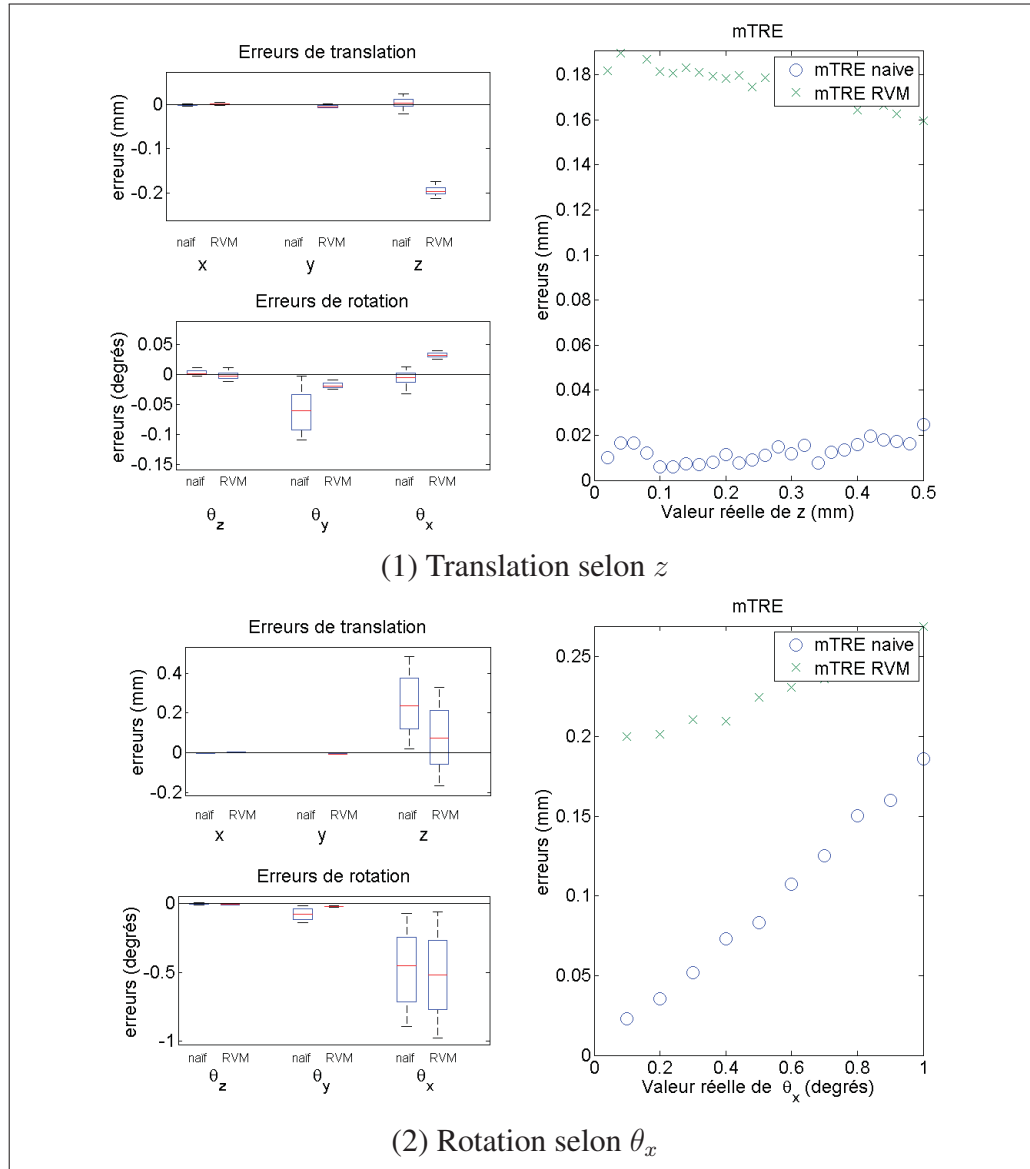


Figure 3.6 Résultats sur séquences de mouvement pur hors plan. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante en mouvement

Les séquences de rotations pures hors plan autour de l'axe latéral (tangage) (figure 3.6 (2)) produisent quant à elles des résultats mitigés. L'évolution des $mTRE$ montre que la correction détériore l'estimation pour des petites rotations, mais produit des résultats comparables à ceux de la méthode naïve pour des rotations suffisamment importantes. Il faut cependant noter que la correction de z est en fait efficace. L'apprentissage ne vise en effet pas à améliorer la $mTRE$,

mais à obtenir les «vraies» valeurs de chaque composante. Du point de vue de l'algorithme d'apprentissage, la correction est donc un succès.

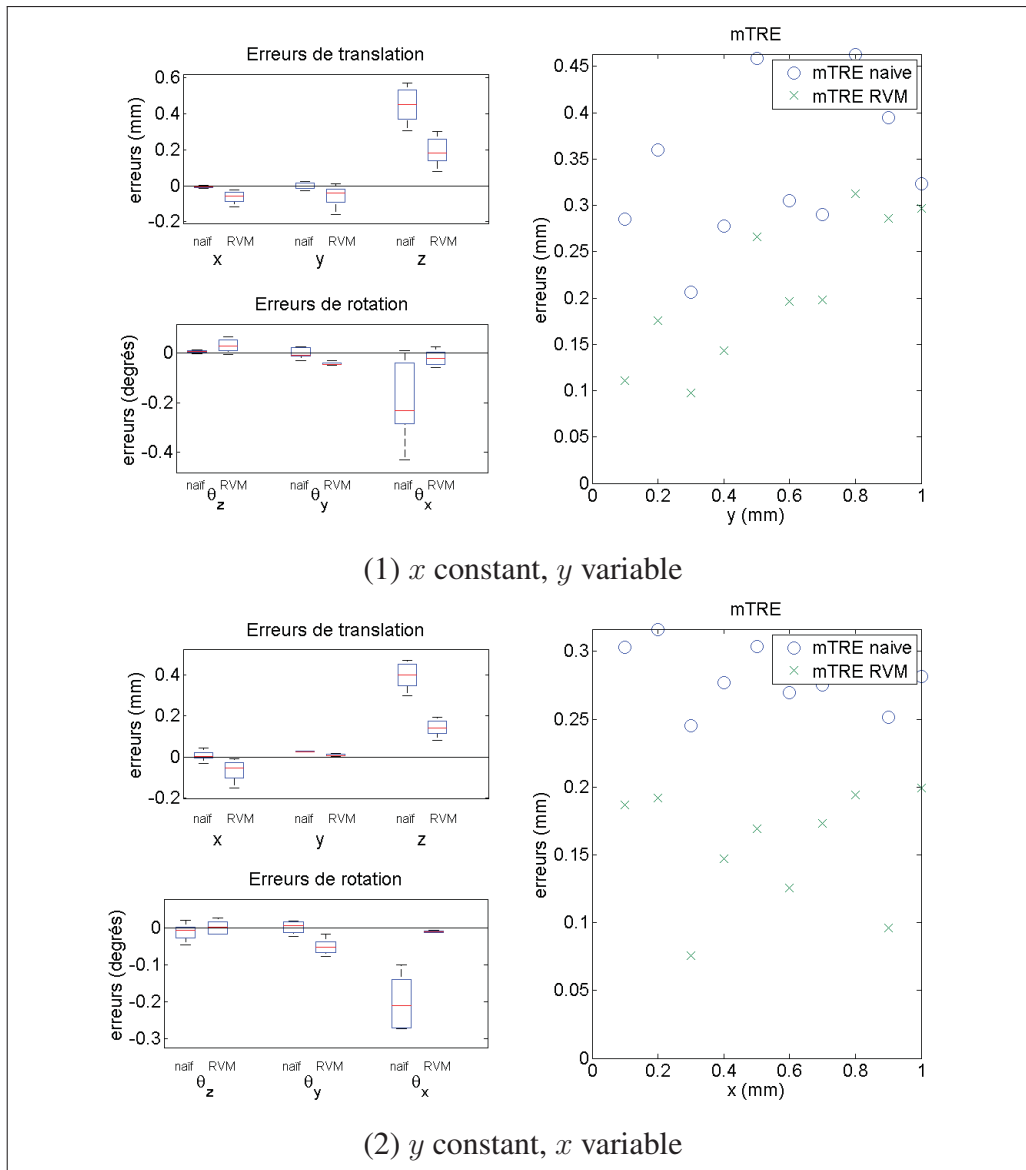


Figure 3.7 Résultats sur séquences de combinaisons de translations dans le plan. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

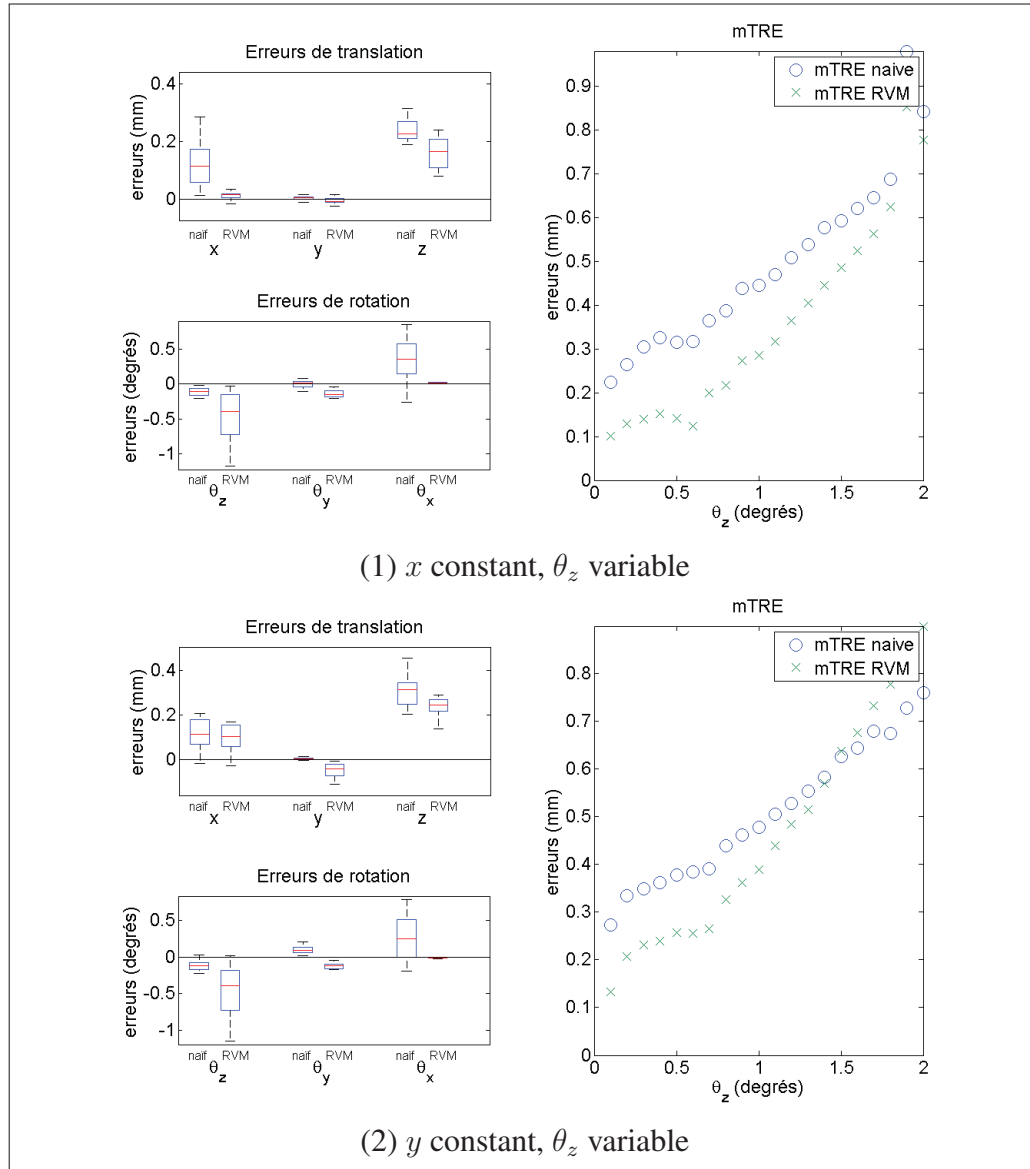


Figure 3.8 Résultats sur séquences de combinaisons de translations et rotations dans le plan. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

Les graphiques des figures 3.7 et 3.8 présentent les résultats pour les séquences combinant des composantes dans le plan (les valeurs de translation constante, selon x ou y , sont de 0.4mm). La correction de l'erreur est démontrée efficace dans tous les cas de mouvement dans le plan. La combinaison de plusieurs composantes produit des résultats très similaires aux mouvements purs dans le plan.

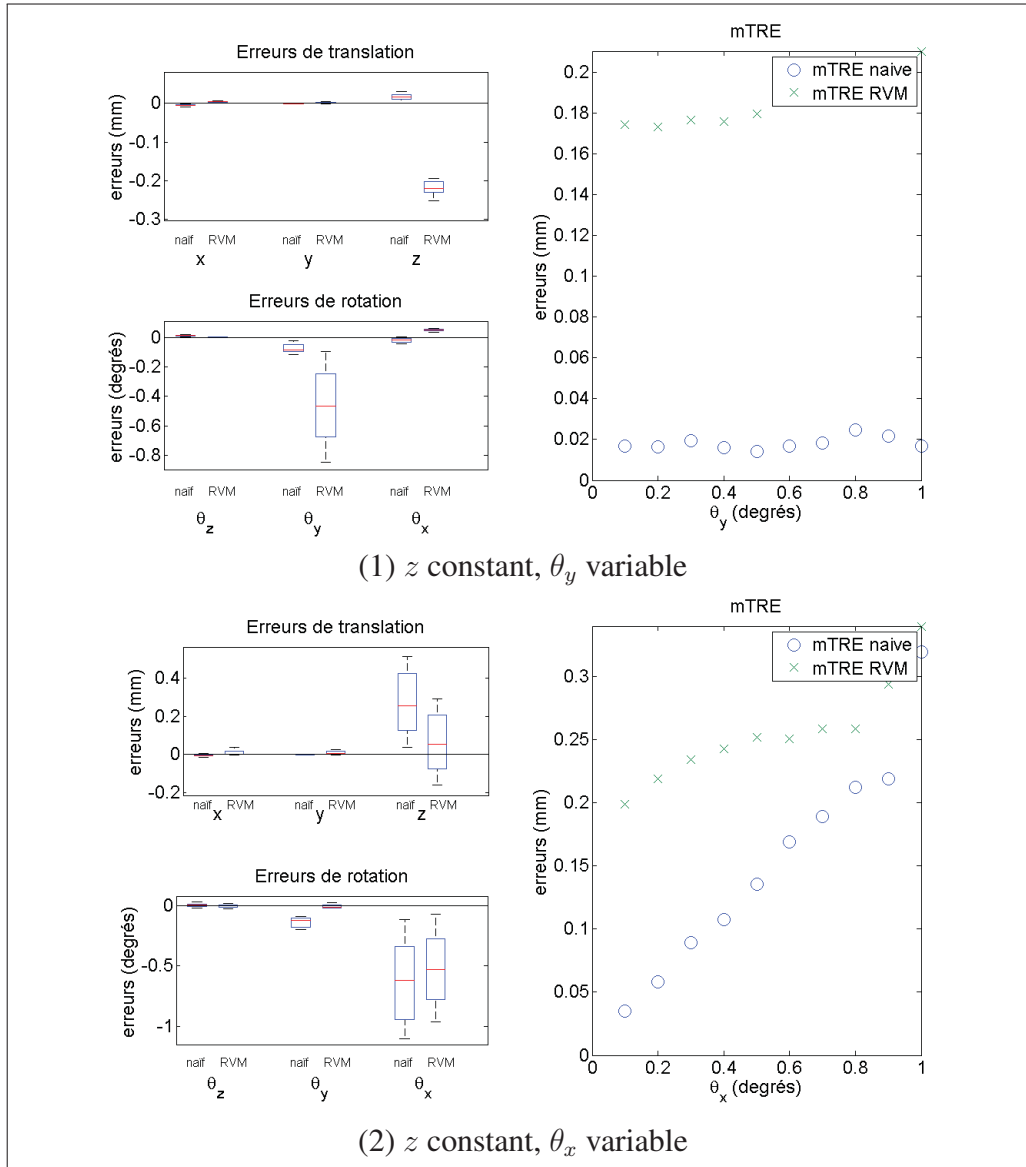


Figure 3.9 Résultats sur séquences de mouvement hors plan. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Évolution des *mTRE* en fonction de la valeur réelle de la composante variable

Les graphiques de la figure 3.9 présentent les résultats pour les séquences combinant des translations élévationnelles constantes (de 0.4mm) avec des rotations hors plan. Comme il a été observé dans les séquences de mouvement pur, l'algorithme répond mal aux cas où une décorrélation supplémentaire n'est pas causée par du mouvement dans le plan. Les résultats de *mTRE* pour les séquences comportant une rotation selon θ_y (lacet) ajoutée à la translation élévationnelle (figure 3.9 (1)) sont très similaires aux résultats de translation élévationnelle pure

(figure 3.6 (1)). Le même problème de sur-correction de z est observé, et survient également pour l'estimation de θ_y , car des valeurs importantes de lacet sont peu présentes dans la base de données d'apprentissage, et des estimés élevés ont donc tendance à être corrigés, même s'ils sont corrects. Les résultats pour les séquences comportant une rotation selon θ_x (tangage) ajoutée à la translation élévationnelle (figure 3.9 (2)) sont quant à eux très similaires aux résultats de tangage pur (figure 3.6 (2)). La correction des estimations de chaque composante est bonne, mais les résultats de $mTRE$ restent meilleurs pour la méthode naïve.

Les graphiques des figures 3.10 et 3.11 présentent les résultats pour les séquences combinant des translations constantes dans le plan avec du mouvement hors plan variable. L'introduction de translation selon z à des translations constantes dans le plan (figure 3.10 (1) et (2)) produit les mêmes effets que la translation élévationnelle pure, c'est-à-dire une sur-correction de z résultant en une sous-estimation, et donc des valeurs croissantes de $mTRE$ lorsque z augmente. Lorsque z est suffisamment faible (inférieure à 0.2mm), la correction est cependant efficace. Ces valeurs, qui représentent selon nos hypothèses les cas les plus probables lors d'une acquisition réelle, correspondent à la vaste majorité des cas représentés dans la base de données d'apprentissage (figure 3.4), et le modèle est donc apte à corriger correctement ces estimations.

L'introduction de rotation selon θ_x aux translations constantes dans le plan (figure 3.11 (1) et (2)) montre que la correction n'est efficace que pour de faibles valeurs de θ_x .

Finalement, l'algorithme a été testé sur deux types de séquences plus complexes :

- 104 paires d'images acquises avec un mouvement combinant les trois composantes translationnelles, variant chacune entre 0.1 et 0.5 mm ;
- 125 paires d'images acquises avec un mouvement uniquement dans le plan : les translations selon x et y varient de 0.1 à 0.5 mm, et la rotation selon θ_z varie de 0.1 à 1.0°.

Les résultats sont présentés sur la figure 3.12. Ces résultats démontrent une très bonne performance de la correction. Les sur-estimations de z sont très bien corrigées, et les estimations

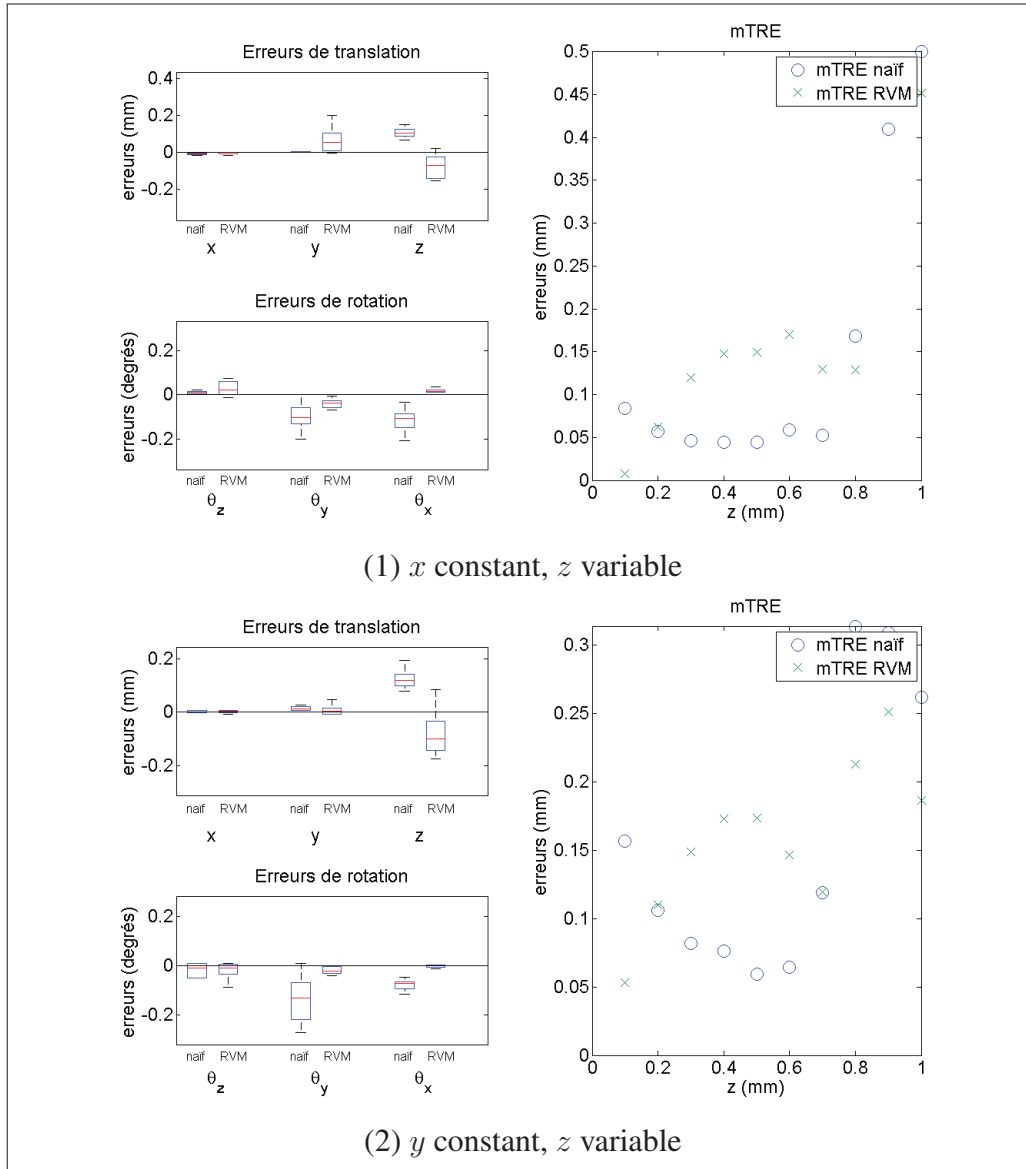


Figure 3.10 Résultats sur séquences de translations dans le plan constantes associées à des translations hors plan variables. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

naïves non nulles des rotations hors plan sont correctement ramenées vers des valeurs nulles. Globalement, les valeurs de $mTRE$ sont ainsi réduites de moitié.

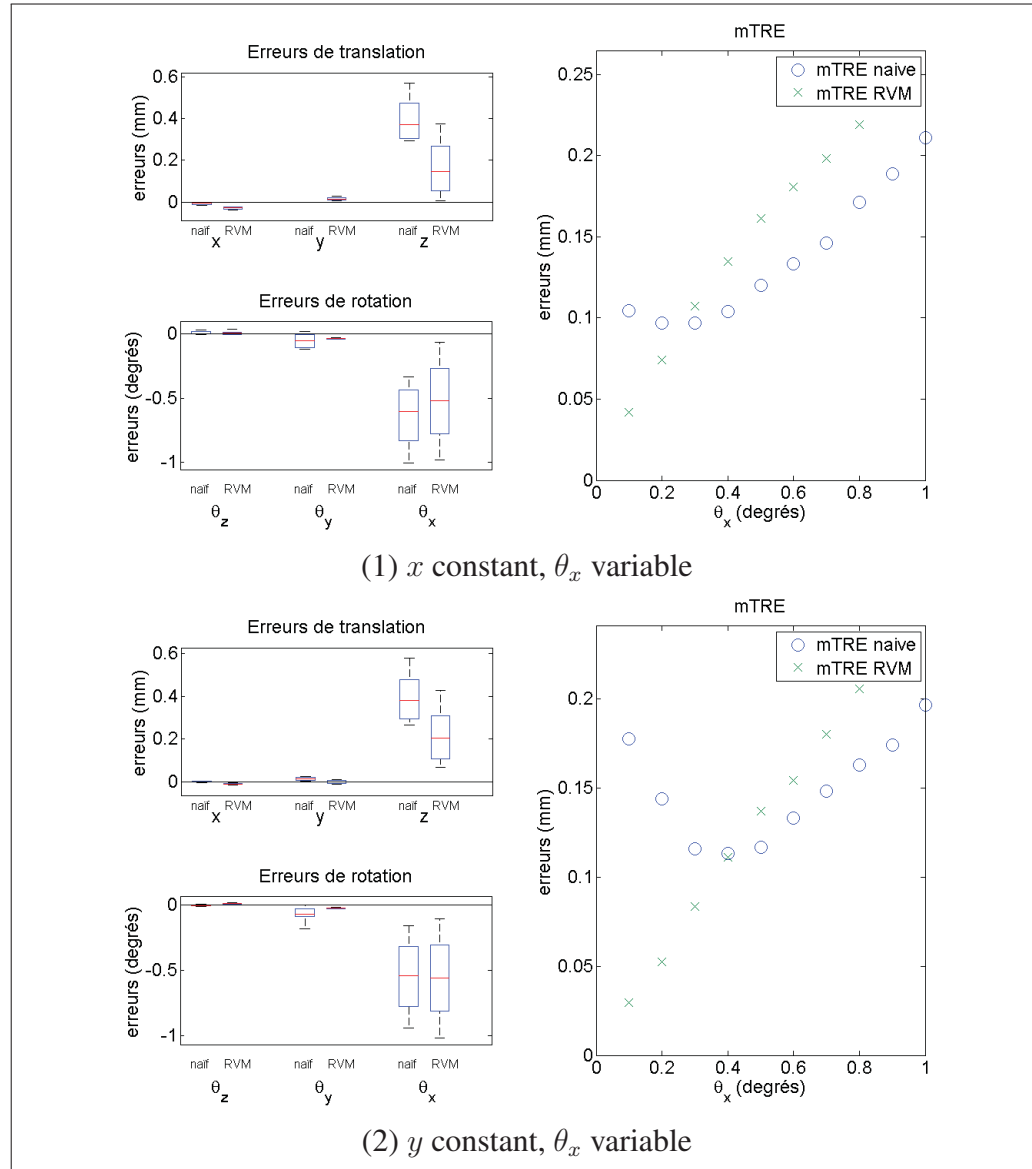


Figure 3.11 Résultats sur séquences de translations dans le plan constantes associées à des rotations hors plan variables. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Évolution des $mTRE$ en fonction de la valeur réelle de la composante variable

3.3 Conclusion

Dans ce chapitre, un nouvel algorithme de correction des estimations de transformation rigide a été présenté. Cette méthode, basée sur l'apprentissage automatique d'un modèle de correction par RVM, a pour objectif de corriger les erreurs d'estimation naïve détaillées au chapitre

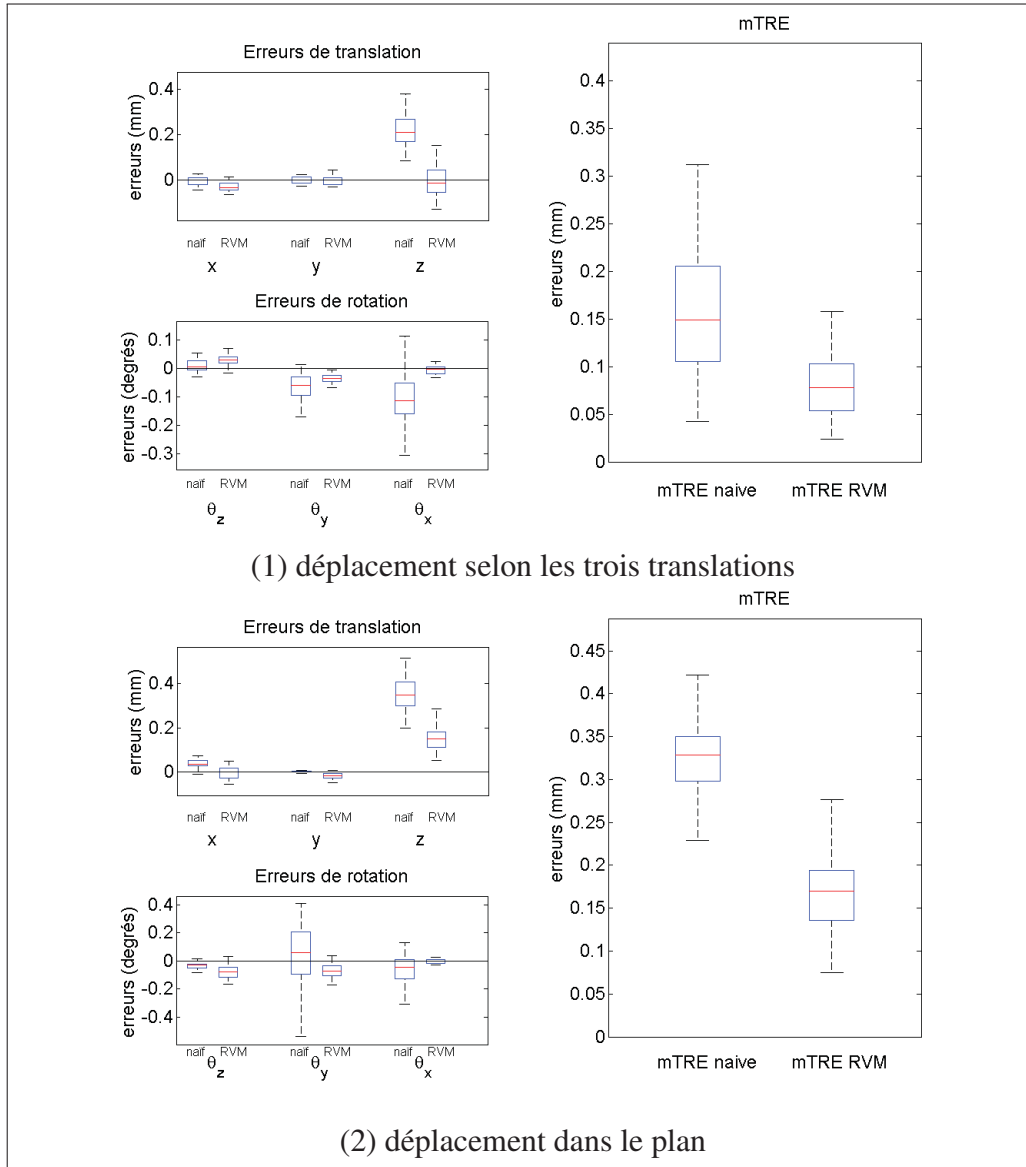


Figure 3.12 Résultats sur séquences complexes. À gauche : Distribution des erreurs d'estimation sur chaque composante pour les deux méthodes. À droite : Distribution des *mTRE* avec et sans apprentissage

2. Il a été observé que ces erreurs, dues au mouvement complexe réalisé lors d'un déplacement libre de la sonde échographique, sont en grande partie systématiques. Cette observation a mené à l'hypothèse qu'un modèle de régression pourrait être construit afin de représenter la relation liant les estimations naïves aux transformations réelles. Une approche d'inférence bayésienne a été choisie, et plus précisément l'apprentissage par RVM. Deux implémentations

de cette méthode ont été utilisées, l'une à sortie monovariée, nécessitant la construction d'un modèle par composante de la transformation rigide, et l'autre à sortie multivariée, permettant de construire un seul modèle pour les six composantes de sortie. Les expérimentations ont permis de comparer les résultats de ces implémentations, et de démontrer une efficacité de correction nettement supérieure pour le modèle à sortie multivariée, comme l'illustre la figure 3.3, notamment. Par la suite, les expérimentations sur mouvements restreints ont permis d'observer les effets de la présence de chaque composante, et d'identifier les cas pour lesquels l'estimation corrigée demeure erronée, et les cas pour lesquels la correction permet d'obtenir des estimations satisfaisantes. Lorsque le mouvement est principalement dans le plan, avec peu ou pas de mouvement hors plan, les résultats sont excellents, et les erreurs importantes dans le plan sont corrigées. Lorsque le mouvement est principalement hors plan, la correction produit le plus souvent un effet négatif sur l'estimation : la $mTRE$ augmente de manière conséquente. La raison principale de ce résultat est liée à l'ensemble d'apprentissage, constitué de séquences de mouvement libre, sur lequel est entraîné le modèle. Cependant, ceci n'est pas nécessairement problématique pour notre application. En effet, les séquences d'entraînement ont été choisies pour être représentatives des séquences obtenues par acquisition réelle, et il est très peu probable que des déplacements uniquement hors plan soient réalisés lors d'un mouvement libre : même s'il s'agit du mouvement principal que l'on retrouvera lors d'une longue séquence (car il s'agit du mouvement intentionnel pour la réalisation d'une séquence adaptée à la reconstruction 3D), il y aura toujours du mouvement plus ou moins aléatoire dans les autres directions. S'il s'avérait cependant que des mouvements purement élévationnels (ou d'autres mouvements mal corrigés actuellement) soient souvent retrouvés dans les séquences réelles, il suffirait d'adapter la base de données d'apprentissage en ajoutant ce type de cas. Une étude future de séquences réelles pourra ainsi permettre de spécifier plus précisément une base de données d'apprentissage cohérente avec la réalité. Par ailleurs, les résultats révèlent un autre type de difficulté qui demeure malgré la correction : le modèle ne parvient pas à traiter les transformations contenant des mouvements de tangage (figure 3.6 (2), figure 3.9 (2), figure 3.11 (1) et (2)). Les perspectives futures présentées dans la conclusion discutent des solutions envisagées. Finalement, les excellents résultats obtenus pour les expérimentations sur des mouvements plus complexes, et plus réalistes, incluant du mouvement dans le plan (figure 3.12) sont très encourageants et

témoignent du bon fonctionnement de l'algorithme de correction, permettant ainsi de valider l'approche proposée.

CONCLUSION

Les travaux présentés dans ce mémoire traitent de l'échographie 3D main libre sans capteur de position, et plus particulièrement d'une méthode visant à corriger les erreurs d'estimation dues à la complexité des mouvements impliqués. En effet, les spécificités du signal échographique détaillées au chapitre 1 permettent un recalage efficace des images dans une séquence limitée à un mouvement purement élévational, et les méthodes de recalage conventionnel permettent d'obtenir des estimations précises pour des images situées dans un même plan. Cependant, comme expliqué dans la revue de littérature du même chapitre, les méthodes existantes combinant ces deux types de recalages produisent des estimations erronées lorsque le mouvement est effectivement libre et donc complexe, c'est-à-dire lorsque le mouvement correspond à une transformation rigide à six degrés de libertés.

Le chapitre 2 a porté sur l'implémentation d'un algorithme dit de «recalage naïf», correspondant à l'état de l'art du recalage pour l'échographie 3D main libre sans capteur de position, pour des conditions simplifiées (une revue de ces conditions et des perspectives d'élargissement de la méthode est présentée ci-après). Cet algorithme utilise ainsi les principes du recalage dans le plan et du recalage hors plan pour des sous-divisions (fenêtres) de l'image, puis impose une contrainte de planarité des images pour parvenir à une estimation des six composantes de la transformation existant dans une paire d'images. Des expérimentations ont ensuite été réalisées sur cet algorithme, afin de mieux comprendre les circonstances menant aux erreurs. Les résultats ainsi obtenus confirment les prédictions des erreurs liées à la forme de la PSF, étudiées notamment par Morrison *et al.* (1983) et Kallel *et al.* (1994), et les résultats obtenus entre autres par Li *et al.* (2002), Housden *et al.* (2007) et Laporte et Arbel (2008), et prolongent ces études grâce à l'analyse de différents types de combinaison de composantes. Il en ressort que chaque composante a un effet systématique sur les estimations, ce qui justifie l'hypothèse posée dans la suite des travaux de ce mémoire : les erreurs observées sont en grande partie liées directement aux valeurs des composantes de la transformation rigide réelle.

La méthode de correction des estimations, constituant la contribution principale de ce mémoire, a donc été présentée au chapitre 3. Cette méthode consiste à effectuer une régression

afin de modéliser la relation liant les estimations naïves de transformation rigide à la transformation réelle. Pour cela, une méthode d'apprentissage bayésien, la RVM, a été choisie. Des expériences préliminaires sur différentes formes d'implémentation ont permis de retenir l'implémentation à sortie multivariée. Celle-ci permet en effet de tenir compte des dépendances entre les composantes de la transformation rigide en construisant un seul modèle liant l'ensemble de la transformation estimée par la méthode naïve à l'ensemble de la transformation réelle. Les résultats de validation croisée démontrent que ce modèle apporte une nette amélioration par rapport à l'estimation naïve. Ce modèle a donc ensuite été testé sur les séquences de mouvement contraint utilisées précédemment pour l'analyse de l'estimation naïve. L'amélioration par rapport aux résultats naïfs a ainsi pu être évaluée en fonction de la composition du mouvement. Cette étude a permis de vérifier la validité de la méthode de correction proposée. Dans la majorité des cas, une réduction importante de l'erreur a en effet pu être observée. Certains cas restent cependant encore problématiques. Les deux types de cas suivants ont ainsi pu être identifiés :

- les mouvements purement hors plan, c'est-à-dire uniquement composé de translation élévationnelle et de rotations hors plan (lacet et tangage) et les mouvements contenant une translation élévationnelle trop élevée, c'est-à-dire supérieure à 0.2mm ;
- les mouvements contenant du tangage.

Le manque d'efficacité de la correction pour le premier type de cas s'explique principalement par le contenu de la base de données d'apprentissage, qui est élaborée pour refléter les mouvements qui seraient effectués lors d'une acquisition main libre réelle. Celle-ci contient ainsi des transformations complexes combinant les six composantes d'une transformation rigide. Le modèle ainsi réalisé n'est donc pas préparé pour traiter des cas restreints à certaines composantes, mais pour corriger les erreurs apparaissant lorsque toutes les composantes sont non nulles. De la même manière, des transformations contenant une translation élévationnelle supérieure à 0.2mm sont peu représentées dans la base de données d'apprentissage. Nous considérons en effet qu'il est peu probable que des translations aussi importantes entre deux acquisitions soient

réalisées. Ainsi, il n'est pas souhaitable d'inclure ce type de transformations pour l'apprentissage, même si la généralisation du modèle n'est pas performante pour ce type de cas. Notons cependant que la base de données pourrait être adaptée si les mouvements problématiques mentionnés s'avéraient plus présents que nous le supposons.

Par ailleurs, des solutions dans la continuité de ces travaux sont envisageables afin de rendre la méthode robuste à ces cas extrêmes, et également de permettre de traiter les problèmes liés au tangage. Ces solutions, ainsi que les travaux futurs envisagés pour élargir les conditions d'applications de la méthode proposée, sont présentées ici. Dans le contexte d'étude choisi, les résultats obtenus demeurent prometteurs, malgré des erreurs subsistantes dans certains cas. Le problème principal est lié à la présence de tangage dans les mouvements. En effet, dès lors que ce type de rotation est contenu dans la transformation réelle, des erreurs importantes dans l'estimation naïve de cette composante apparaissent (elle est presque toujours estimée nulle), et ces erreurs sont peu corrigées par le modèle de correction, menant ainsi à une estimation finale peu représentative de la réalité. Notons cependant que dans le cas où la valeur de tangage est nulle et que l'estimation naïve de cette composante est incorrecte (ceci arrive dans certaines configurations à cause de la forme de la PSF), la correction est cette fois efficace. Housden *et al.* (2008) se sont penchés sur ce problème spécifique, et ont proposé de le résoudre en construisant deux courbes de décorrélation. L'une correspond à celle utilisée ici, calibrée sur une séquence purement élévationnelle à intervalles fixes. L'autre est calibrée sur une séquences présentant une valeur de tangage constante, en plus du déplacement élévationnel entre chaque acquisition. Leur méthode s'appuie par la suite sur un capteur d'orientation permettant de connaître la valeur réelle de tangage entre deux images. Une interpolation des deux modèles permet ainsi d'obtenir une courbe ajustée à la valeur de tangage de la paire d'images considérée, et d'obtenir une estimation plus précise de la distance élévationnelle. Sans l'ajout de capteur, l'idée de construire deux courbes de corrélation pourrait être implémentée dans notre algorithme naïf, permettant d'obtenir deux ensembles de points de contrôle, puis d'utiliser la valeur de $mTRE_{proc}$ définie dans la section 2.4 comme critère de rigidité afin de sélectionner l'ensemble correspondant le plus à une transformation rigide comme estimation optimale.

Par ailleurs, l'ensemble des cas problématiques rencontrés lors de notre étude pourraient être résolus lors de l'estimation de déplacement d'une séquence entière contenant une centaine d'images ou plus. En effet, des estimations erronées pour une paire d'images particulière produiraient des résultats incompatibles avec les autres estimations, et pourraient donc être repérées et éliminées. De plus, le mouvement peut aussi être estimé entre des paires d'images non consécutives, ce qui permet ainsi d'obtenir de la redondance et de repérer les incompatibilités. Deux méthodes différentes d'estimation de déplacement pour une séquence entière ont été proposées dans la littérature, permettant également de repérer les changements de direction de la sonde et les intersections. Housden *et al.* (2007) ont ainsi proposé d'estimer la distance entre toutes les paires d'images potentiellement corrélées, puis de sélectionner des ensembles de trois images parmi les estimations disponibles. Une détection des intersections est ensuite réalisée grâce à une recherche de la configuration optimale, puis une recherche de changement de direction est effectuée par optimisation des quatre estimations disponibles pour chaque ensemble de trois images. La méthode de Laporte et Arbel (2008) consiste quant à elle à ordonner tout d'abord les fenêtres correspondantes suivant la direction élévationnelle afin de minimiser la distance parcourue, ce qui permet d'identifier les changements de sens, puis d'imposer la contrainte de planarité des images afin d'obtenir les estimations des composantes hors plan de la transformation (ces travaux ont été réalisés sous l'hypothèse de mouvement dans le plan nul). L'une de ces solutions pourrait être adaptée et implémentée dans notre méthode, et permettrait ainsi non seulement d'identifier les estimations trop éloignées du déplacement global possible, mais également de lever les hypothèses de monotonie du mouvement et de non-intersections.

Une autre hypothèse posée dans ce mémoire concerne le speckle : la méthode proposée est applicable telle quelle uniquement dans les conditions de Rayleigh, c'est-à-dire pour des échographies présentant du speckle pleinement développé. Certains auteurs (Tuthill *et al.*, 1998; Prager *et al.*, 2003; Hassenpflug *et al.*, 2005) ont proposé de détecter les zones respectant ces conditions, et d'utiliser uniquement celles-ci pour effectuer les estimations. Cependant, ces conditions sont rarement rencontrées dans la réalité, et il est nécessaire de parvenir à traiter les cas de speckle partiellement développé. L'approche de Laporte et Arbel (2011) est envisagée pour cela. Celle-ci consiste à adapter la courbe de décorrélation au degré de développement du

speckle dans chaque fenêtre, grâce à un facteur multiplicatif obtenu par l'apprentissage de la relation entre celui-ci et les statistiques des images. Cette méthode pourrait être implémentée directement dans notre algorithme naïf.

Finalement, il est indispensable de noter que toute l'étude a été menée sur des séquences synthétiques, et que le modèle lui-même a été construit sur de telles séquences. Il est difficile de prédire comment ce modèle se comporterait pour des séquences réelles. Il est probable qu'il soit nécessaire de l'adapter en fonction de la sonde employée, même si cela n'est pas une évidence car la courbe de décorrélation n'entre pas directement en jeu dans la relation modélisée. Des expériences sur des acquisitions réelles permettront de répondre à ces questions et de guider la suite des travaux.

Lorsque toutes ces idées auront été implémentées et que la meilleure précision d'estimation possible aura été obtenue pour les séquences réelles, il pourra être envisagé de combiner notre méthode avec un capteur de position à six degrés de liberté (comme proposé par Lang *et al.* (2009)), ou un capteur d'orientation à trois degrés de liberté (comme proposé par Housden *et al.* (2008)). Notre méthode sans capteur étant plus précise que les méthodes de base (sans capteur) utilisées dans ces références, nous pouvons espérer qu'une méthode hybride basée sur notre méthode résulterait en une amélioration par rapport aux résultats obtenus par les méthodes hybrides citées.

BIBLIOGRAPHIE

- M.E. Anderson et G.E. Trahey. 2000. « A seminar on k-space applied to medical ultrasound ». *Course Notes for use in BME*.
- K.S. Arun, T.S. Huang, et S.D. Blostein. 1987. « Least-squares fitting of two 3-D point sets ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, n° 5, p. 698–700.
- A. Austeng et S. Holm. 2002. « Sparse 2-D arrays for 3-D phased array imaging-design methods ». *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 49, n° 8, p. 1073–1086.
- W. Birkfellner, F. Watzinger, F. Wanschitz, R. Ewers, et H. Bergmann. 1998. « Calibration of tracking systems in a surgical environment ». *IEEE Transactions on Medical Imaging*, vol. 17, n° 5, p. 737–742.
- C.M. Bishop, 2006. *Pattern recognition and machine learning*, volume 4. Springer New York.
- J.M. Blackall, G.P. Penney, A.P. King, et D.J. Hawkes. 2005. « Alignment of sparse freehand 3-D ultrasound with preoperative images of the liver using models of respiratory motion and deformation ». *IEEE Transactions on Medical Imaging*, vol. 24, n° 11, p. 1405–1416.
- L.N. Bohs et G.E. Trahey. 1991. « A novel method for angle independent ultrasonic imaging of blood flow and tissue motion ». *IEEE Transactions on Biomedical Engineering*, vol. 38, n° 3, p. 280–286.
- L.N. Bohs, B.J. Geiman, M.E. Anderson, S.C. Gebhart, et G.E. Trahey. 2000. « Speckle tracking for multi-dimensional flow estimation ». *Ultrasonics*, vol. 38, n° 1–8, p. 369–375.
- C.B. Burckhardt. 1978. « Speckle in ultrasound B-mode scans ». *IEEE Transactions on Sonics and Ultrasonics*, vol. 25, n° 1, p. 1–6.
- R.F. Chang, W.J. Wu, D.R. Chen, W.M. Chen, W. Shu, J.H. Lee, et L.B. Jeng. 2003. « 3-D US frame positioning using speckle decorrelation and image registration ». *Ultrasound in medicine & biology*, vol. 29, n° 6, p. 801–812.
- J.F. Chen, J.B. Fowlkes, P.L. Carson, et J.M. Rubin. 1997. « Determination of scan-plane motion using speckle decorrelation : theoretical considerations and initial test ». *International Journal of Imaging Systems and Technology*, vol. 8, n° 1, p. 38–44.
- X. Chen, X. Liu, D. Sahn, K. Kim, H. Xie, et M. O'Donnell. 2003. « Temporal and spatial registration for cardiac strain rate imaging ». *Proceedings of the IEEE Ultrasonics Symposium*, vol. 2, p. 2134–2137.
- D. Cool, S. Sherebrin, J. Izawa, J. Chin, et A. Fenster. 2008. « Design and evaluation of a 3D transrectal ultrasound prostate biopsy system ». *Medical physics*.

- P. Coupé, P. Hellier, C. Kervrann, et C. Barillot. 2009. « Nonlocal means-based speckle filtering for ultrasound images ». *IEEE Transactions on Image Processing*, vol. 18, n° 10, p. 2221–2229.
- A.N. Evans et MS Nixon. 1995. « Mode filtering to reduce ultrasound speckle for feature extraction ». *IEE Proceedings-Vision Image and Signal Processing*, vol. 142, n° 2, p. 87–94.
- JL Farrell, JC Stuelpnagel, RH Wessner, JR Velman, et JE Brook. 1966. « A Least Squares Estimate of Satellite Attitude (Grace Wahba) ». *SIAM Review*, vol. 8, n° 3, p. 384–386.
- P. Foroughi et P. Abolmaesumi. 2005. « Elastic registration of 3d ultrasound images ». *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, p. 83–90.
- A.H. Gee, R. James Housden, P. Hassenpflug, G.M. Treece, et R.W. Prager. 2006. « Sensorless freehand 3D ultrasound in real tissue : speckle decorrelation without fully developed speckle ». *Medical Image Analysis*, vol. 10, n° 2, p. 137–149.
- P. Hassenpflug, R. Prager, G. Treece, et A. Gee. 2004. « Distance measurement for sensorless 3D US ». *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*, p. 1087–1088.
- P. Hassenpflug, R.W. Prager, G.M. Treece, et A.H. Gee. 2005. « Speckle classification for sensorless freehand 3-D ultrasound ». *Ultrasound in medicine & biology*, vol. 31, n° 11, p. 1499–1508.
- D.L.G. Hill, P.G. Batchelor, M. Holden, et D.J. Hawkes. 2001. « Medical image registration ». *Physics in Medicine and Biology*, vol. 46, n° 3, p. R1.
- R.J. Housden, A.H. Gee, G.M. Treece, et R.W. Prager. 2006. « Subsample interpolation strategies for sensorless freehand 3D ultrasound ». *Ultrasound in medicine & biology*, vol. 32, n° 12, p. 1897–1904.
- R.J. Housden, A.H. Gee, G.M. Treece, et R.W. Prager. 2007. « Sensorless Reconstruction of Unconstrained Freehand 3D Ultrasound Data ». *Ultrasound in Medicine & Biology*, vol. 33, n° 3, p. 408 - 419.
- R.J. Housden, A.H. Gee, R.W. Prager, et G.M. Treece. 2008. « Rotational motion in sensorless freehand three-dimensional ultrasound ». *Ultrasonics*, vol. 48, n° 5, p. 412 - 422.
- M.A. Janvier, L.G. Durand, M.H.R. Cardinal, I. Renaud, B. Chayer, P. Bigras, J. de Guise, G. Soulez, et G. Cloutier. 2008. « Performance evaluation of a medical robotic 3D-ultrasound imaging system ». *Medical Image Analysis*, vol. 12, n° 3, p. 275–290.
- JA Jensen. 1996. « Field : A program for simulating ultrasound systems ». *Nordic-Baltic Conference on Biomedical Imaging*, p. 351–353.

- J.A. Jensen et N.B. Svendsen. 1992. « Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers ». *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 39, n° 2, p. 262–267.
- F. Kallel, M. Bertrand, et J. Meunier. Janvier 1994. « Speckle motion artifact under tissue rotation ». *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 41, n° 1, p. 105–122.
- C. Kontogeorgakis, MG Strintzis, N. Maglaveras, et I. Kokkinidis. 1994. « Tumor detection in ultrasound B-mode images through motion estimation using a texture detection algorithm ». In *Computers in Cardiology 1994*. p. 117–120. IEEE.
- J.F. Krucker, G.L. LeCarpentier, J.B. Fowlkes, et P.L. Carson. 2002. « Rapid elastic image registration for 3-D ultrasound ». *IEEE Transactions on Medical Imaging*, vol. 21, n° 11, p. 1384–1394.
- A. Lang, P. Mousavi, G. Fichtinger, et P. Abolmaesumi. 2009. « Fusion of electromagnetic tracking with speckle-tracked 3D freehand ultrasound using an unscented Kalman filter ». *SPIE Medical Imaging*, vol. 7265.
- A. Lang, P. Mousavi, S. Gill, G. Fichtinger, et P. Abolmaesumi. 2012. « Multi-modal Registration of Speckle-tracked Freehand 3-D Ultrasound to CT in the Lumbar Spine ». *Medical Image Analysis*, vol. 16, n° 3, p. 675–686.
- C. Laporte. 2011. « Statistical methods for out-of-plane ultrasound transducer motion estimation, Thèse de doctorat ».
- C. Laporte et T. Arbel. 2008. « Combinatorial and probabilistic fusion of noisy correlation measurements for untracked freehand 3-D ultrasound ». *IEEE Transactions on Medical Imaging*, vol. 27, n° 7, p. 984–994.
- C. Laporte et T. Arbel. 2011. « Learning to estimate out-of-plane motion in ultrasound imagery of real tissue ». *Medical Image Analysis*, vol. 15, n° 2, p. 202–213.
- C. Leung, K. Hashtrudi-Zaad, P. Foroughi, et P. Abolmaesumi. 2009. « A real-time intra-subject elastic registration algorithm for dynamic 2-D ultrasound images ». *Ultrasound in Medicine and Biology*, vol. 35, n° 7, p. 1159–1176.
- JP Lewis. 1995. « Fast template matching ». In *Vision Interface*. p. 120–123.
- C. Li, C.Y. Li, et W.C. Yeh. 2002. « Tissue motion and elevational speckle decorrelation in freehand 3D ultrasound ». *Ultrasonic Imaging*, vol. 22, n° 1.
- F. Lindseth, G.A. Tangen, T. Langø, et J. Bang. 2003. « Probe calibration for freehand 3-D ultrasound ». *Ultrasound in Medicine & Biology*, vol. 29, n° 11, p. 1607–1623.
- R.L. Maurice et M. Bertrand. 1999. « Lagrangian speckle model and tissue-motion estimation-theory ». *IEEE Transactions on Medical Imaging*, vol. 18, n° 7, p. 593–603.

- L. Mercier, T. Langø, F. Lindseth, et L.D. Collins. 2005. « A review of calibration techniques for freehand 3-D ultrasound systems ». *Ultrasound in Medicine & Biology*, vol. 31, n° 2, p. 143–165.
- DC Morrison, WN McDicken, et DSA Smith. 1983. « A motion artefact in real-time ultrasound scanners ». *Ultrasound in Medicine & Biology*, vol. 9, n° 2, p. 201–203.
- J. Ng, R.W. Prager, N. Kingsbury, G. Treece, et A. Gee. 2006. « Modeling ultrasound imaging as a linear, shift-variant system ». *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 53, n° 3, p. 549–563.
- G.P. Penney, J. Weese, J.A. Little, P. Desmedt, D.L.G. Hill, et al. 1998. « A comparison of similarity measures for use in 2-D-3-D medical image registration ». *IEEE Transactions on Medical Imaging*, vol. 17, n° 4, p. 586–595.
- R.W. Prager, A.H. Gee, G.M. Treece, C.J.C. Cash, et L.H. Berman. 2003. « Sensorless freehand 3-D ultrasound using regression of the echo intensity ». *Ultrasound in Medicine & Biology*, vol. 29, n° 3, p. 437–446.
- R.W. Prager, U.Z. Ijaz, A.H. Gee, et G.M. Treece. 2010. « Three dimensional ultrasound Imaging ». *Proceedings of the Institution of Mechanical Engineers, Part H : Journal of Engineering in Medicine*, vol. 24, n° 2, p. 193–223.
- J. Quiñonero-Candela et C.E. Rasmussen. 2005. « A unifying view of sparse approximate Gaussian process regression ». *The Journal of Machine Learning Research*, vol. 6, p. 1939–1959.
- R.W. Schafer et L.R. Rabiner. 1973. « A digital signal processing approach to interpolation ». *Proceedings of the IEEE*, vol. 61, n° 6, p. 692–702.
- P.H. Schönemann. 1966. « A generalized solution of the orthogonal Procrustes problem ». *Psychometrika*, vol. 31, n° 1, p. 1–10.
- S.W. Smith, H.G. Pavy Jr, et O.T. von Ramm. 1991. « High-speed ultrasound volumetric imaging system. I. Transducer design and beam steering ». *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 38, n° 2, p. 100–108.
- O.V. Solberg, F. Lindseth, H. Torp, R.E. Blake, et T.A. Nagelhus Hernes. 2007. « Freehand 3D ultrasound reconstruction algorithms—a review ». *Ultrasound in Medicine & Biology*, vol. 33, n° 7, p. 991–1009.
- PR Stepanishen. 1971. « The Time-Dependent Force and Radiation Impedance on a Piston in a Rigid Infinite Planar Baffle ». *The Journal of the Acoustical Society of America*, vol. 49, p. 841–862.
- A. Støylen. 2010. « Basic ultrasound, echocardiography and Doppler for clinicians ». <http://folk.ntnu.no/stoylen/strainrate/Ultrasound/>. [En ligne ; consulté 01-Février-2011].

- C. Tay, C.D. Garson, S.T. Acton, et J.A. Hossack. 2010. « Ultrasound Despeckling for Contrast Enhancement ». *IEEE Transactions on Image Processing*, vol. 19, n° 7, p. 1847 -1860.
- A. Thayananthan, R. Navaratnam, B. Stenger, P. Torr, et R. Cipolla. 2006. « Multivariate Relevance Vector Machines for Tracking ». *European Conference on Computer Vision*, p. 124–138.
- M.E. Tipping. 2001. « Sparse Bayesian learning and the relevance vector machine ». *The Journal of Machine Learning Research*, vol. 1, p. 211–244.
- GM Treece, RW Prager, AH Gee, et L. Berman. 2002. « Correction of probe pressure artifacts in freehand 3D ultrasound ». *Medical Image Analysis*, vol. 6, n° 3, p. 199–214.
- GE Tupholme. 1969. « Generation of acoustic pulses by baffled plane pistons ». *Mathematika*, vol. 16, n° 02, p. 209–224.
- T.A. Tuthill, JF Krücker, J.B. Fowlkes, et P.L. Carson. 1998. « Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. ». *Radiology*, vol. 209, n° 2, p. 575–582.
- S. Umeyama. 1991. « Least-squares estimation of transformation parameters between two point patterns ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, n° 4, p. 376–380.
- E.B. van de Kraats, G.P. Penney, D. Tomazevic, T. van Walsum, et W.J. Niessen. 2005. « Standardized evaluation methodology for 2-D-3-D registration ». *IEEE Transactions on Medical Imaging*, vol. 24, n° 9, p. 1177–1189.
- V.N. Vapnik. 1999. « An overview of statistical learning theory ». *IEEE Transactions on Neural Networks*, vol. 10, n° 5, p. 988–999.
- R.F. Wagner. 1983. « Statistics of speckle in ultrasound B-scans ». *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, n° 3, p. 156-163.
- K.A. Wear et R.L. Popp. 1987. « Methods for estimation of statistical properties of envelopes of ultrasonic echoes from myocardium ». *IEEE Transactions on Medical Imaging*, vol. 6, n° 4, p. 281–291.
- G. Xiao, J.M. Brady, J.A. Noble, M. Burcher, et R. English. 2002. « Nonrigid registration of 3-D free-hand ultrasound images of the breast ». *IEEE Transactions on Medical Imaging*, vol. 21, n° 4, p. 405–412.
- F. Yeung, S.F. Levinson, D. Fu, et K.J. Parker. 1998. « Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh ». *IEEE Transactions on Medical Imaging*, vol. 17, n° 6, p. 945–956.
- W. Yu, P. Yan, A.J. Sinusas, K. Thiele, et J.S. Duncan. 2006. « Towards pointwise motion tracking in echocardiographic image sequences - Comparing the reliability of different features for speckle tracking ». *Medical Image Analysis*, vol. 10, n° 4, p. 495 - 508.