Statistical methods for out-of-plane ultrasound transducer motion estimation

Catherine Laporte

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

November 2009

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Ph.D.

Abstract

Freehand 3D ultrasound imaging usually involves moving a conventional tracked 2D ultrasound probe over a subject and combining the images into a volume to be interpreted for medical purposes. Tracking devices can be cumbersome; thus, there is interest in inferring the trajectory of the transducer based on the images themselves. This thesis focuses on new methods for the recovery of the out-of-plane component of the transducer trajectory using the predictive relationship between the elevational decorrelation of ultrasound speckle patterns and transducer displacement. To resolve the directional ambiguities associated with this approach, combinatorial optimisation techniques and robust statistics are combined to recover non-monotonic motion and frame intersections. In order to account for the variability of the sample correlation coefficient between corresponding image patches of fully developed speckle, a new probabilistic speckle decorrelation model is developed. This model can be used to quantify the uncertainty of any displacement estimate, thereby facilitating the use of a new maximum likelihood out-of-plane trajectory estimation approach which fully exploits the information available from multiple redundant and noisy correlation measurements collected in imagery of fully developed speckle. To generalise the applicability of these methods to the case of imagery of real tissue, a new data-driven method is proposed for locally estimating elevational correlation length based on statistical features collected within the image plane. In this approach, the relationship between the image features and local elevational correlation length is learned by sparse Gaussian process regression using a training set of synthetic ultrasound image sequences. The synthetic imagery used for learning is created via a new statistical model for the spatial distribution of ultrasound scatterers which maps realisations of a 1D generalised Poisson point process to a 3D Hilbert space-filling curve. In experiments with imagery of animal tissue, the learning-based approach is shown to give distance estimates more accurate than those obtained using a speckle detection filter and comparable to the state-of-the-art heuristic method. Remaining modelling imperfections are accounted for by a new iterative algorithm which extends the proposed maximum likelihood measurement fusion approach. In this algorithm, probabilistic measurement fusion and measurement selection steps based on statistical hypothesis testing alternate in order to establish a trajectory estimate based on measurements which agree with each other. This approach succeeds in avoiding distance under-estimates arising from image structures exhibiting significant but uninformative correlation over long distances.

Sommaire

L'échographie 3D main-libre consiste habituellement à déplacer et à mesurer le déplacement d'une sonde échographique 2D conventionnelle au-dessus d'un sujet et à créer un volume à partir des images qui sera ensuite interprété dans un but médical. Puisque les capteurs de position externes peuvent être encombrants, il y a un intérêt à calculer la trajectoire de la sonde à partir des images elles-mêmes. Cette thèse se penche sur de nouvelles méthodes pour le calcul de la composante hors-plan de la trajectoire de la sonde utilisant la relation prédictive entre la décorrélation hors-plan du speckle échographique et le déplacement de la sonde. Afin de résoudre les ambiguïtés directionnelles associées à cette approche, un nouveau cadre d'opérations est proposé. Ce cadre combine des techniques d'optimisation combinatoire et des techniques statistiques robustes pour détecter les mouvements non-monotones et les intersections entre les images. Pour tenir compte de la variabilité du coefficient de corrélation échantillonnaire entre deux portions d'images de speckle pleinement développé correspondantes, un nouveau modèle probabiliste de la décorrélation du speckle est développé. Ce modèle permet de quantifier l'incertitude associée à l'estimé d'un déplacement, facilitant ainsi l'utilisation d'une nouvelle approche de maximisation de la vraisemblance pour l'estimation de la trajectoire hors-plan qui exploite pleinement l'information rendue disponible par des mesures de corrélation multiples et redondantes acquises dans des images de speckle pleinement développé. Afin de généraliser l'applicabilité de ces méthodes au cas d'images de tissus véritables, un nouvel algorithme guidé par les données est proposé pour l'estimation de la longueur de corrélation hors-plan locale à partir d'attributs statistiques acquis à même le plan image. Dans cette approche, la relation entre les attributs de l'image et la longueur de corrélation hors-plan locale est établie par régression selon un processus gaussien parcimonieux en utilisant des données d'apprentissage tirées de séquences échographiques synthétiques. L'imagerie synthétique utilisée pour l'apprentissage est créée via un nouveau modèle statistique pour la distribution spatiale de diffuseurs échographiques qui transforme les réalisations d'un processus ponctuel de Poisson généralisé 1D selon une courbe de Hilbert remplissant l'espace en 3D. Lors d'expériences avec de l'imagerie de tissus animaux, il est démontré que le modèle adaptif de décorrélation du speckle basé sur l'apprentissage permet d'estimer les distances de façon plus exacte qu'un filtre de détection de speckle et de façon comparable avec la méthode heuristique qui constitue l'état de l'art. Des imperfections subsistant dans cette

modélisation sont prises en compte par un nouvel algorithme itératif développé à partir de l'approche proposée se basant sur la maximisation de la vraisemblance. Dans cet algorithme, des étapes de fusion probabiliste et de sélection de mesures par test d'hypothèse se succèdent afin d'établir un estimé de la trajectoire fondé sur des mesures qui s'appuient mutuellement. Cette approche permet d'éviter les sous-estimés de distance obtenus en présence de structures d'images exhibant une corrélation significative mais non informative sur de longues distances.

Acknowledgements

First, I wish to thank my thesis supervisor Tal Arbel for giving me a unique opportunity to learn through this ultrasound project. I am grateful for the freedom she gave me in carrying out this research project, as well as the considerable amounts of trust, financial resources, patience, moral and professional support that she invested in my research, as well as much editorial help with this document. I also wish to thank my thesis committee members, Guy Cloutier and Gregory Dudek, for making their respective expertises in ultrasound imaging and probabilistic estimation techniques available to me.

My gratitude also goes to James Clark for introducing me to space-filling curves, and to Denis Larocque for helping me find a way to carry out the statistical analysis of my results on ultrasound phantom data. Donald Pavlasek and his team from the departmental mechanical workshop designed, built and troubleshot the sub-millimetre positioning device that I used for my experiments. My friend and colleague Jérôme Pasquero initially assisted me in measuring its accuracy. Boris Chayer assisted me with the choice of the ultrasound phantom that was used throughout the research presented in this thesis and was very helpful in ensuring its maintenance. Dale Einarson, the former CLUMEQ system administrator, has been of great help in helping me get my ultrasound simulations running properly. I also thank Michael Tipping, who shared with me a recent version of his Relevance Vector Machine code; I used it for a long time and it played a significant role in the development of the material presented in Chapter 6. I was fortunate enough to meet Richard Prager an Graham Treece on a visit to their ultrasound laboratory in Cambridge. The technical conversations that ensued have been instrumental in helping me getting started with my research project. For this, and for continued Stradx support over the years, I am endebted to them and also to their colleague Andrew Gee. I also owe thanks to CIM's system administrator Jan Binder for patiently and diligently resolving the technical problems I have encountered over the years. I also wish to thank Michael Langer and Farida Cheriet, two excellent teachers and scientists whose influences have significantly contributed to my development as a scientist over the years.

In completing this thesis, I am also greatly endebted to Rupert Brooks, my friend and colleague of many years. He was always available for fruitful discussions of my research, and on many occasions, his expertise in optimisation has been very helpful. He also gave me much needed feedback on several chapters of this thesis and several of my papers prior

to that. He and Isabelle Bégin have provided me with tremendous amounts of excellent home cooked food and moral support in the latest stages of this endeavour. I also owe special thanks to Prasun Lala, who proofread two chapters of this thesis and has always been available to discuss issues with thesis structure, and the analysis and presentation of experimental results. I also owe him much moral support. I also wish to thank all my other friends and colleagues at CIM, past and present, particularly Sandra Skaff, Matthew Toews, Rola Harmouche, Frank Riggi, Peter Savadjev, Shufei Fan, Prakash Patel, Jérôme Pasquero, Scott McCloskey, Thierry Bélanger-Roy, Mohak Shah, Marta Kersten, Nils Tilton and Meltem Demirkus. Their presence, wisdom and moral support have been instrumental to my well-being and productivity while at CIM.

I cannot thank my family and friends enough for the support they have shown me over the years it took to take this project to fruition. I particularly want to thank my partner Mark Wallace for his kindness, his affectionate support, his lifts to the lab on weekends, his humour, and above all, his patience from the very beginning to the very end of this endeavour. My parents Gilbert Laporte and Ann Langley have worked all their lives in order to support my education and ambitions. They have been role models to me, both as academics and as human beings. I am also grateful to my father for some helpful technical discussions about traveling salesman problems and stochasticity; I do not know many other graduate students whose parents are still able to help them with their homework. I also owe special thanks to my friends Marie-Claude Levac, Jean-François Simoneau, Jean-Sylvain Sormany, Corinne Potvin and Samia Saouaf for their moral support and for understanding why I could not come out and play during this last summer. Thanks also to the musicians in my life, France Champagne, Martin Dagenais, Mary Jane Puiu and Geneviève Lévesque, for giving me the opportunity to participate in musical projects which truly enriched the time I spent as a graduate student, and for giving me a much broader perspective on what it truly means to understand and process acoustic signals.

The research described in this thesis was supported by a doctoral scholarship from the Natural Science and Engineering Research Council of Canada, the Max E. Binz fellowship from McGill University as well as a doctoral scholarship from the Desjardins Foundation. Computational ressources for the ultrasound image simulations in this thesis were provided by CLUMEQ, which is funded in part by NSERC (MRS), FQRNT and McGill University. Further computational resources for these simulations were provided by RQCHP. I thank these institutions for their financial support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Ultrasound imaging systems have been used for medical purposes since the mid 1950s (see [241] for a fascinating account of the invention, early evolution and impact of medical ultrasound technology). Since then, they have continued to find widespread use in both diagnostic and interventional applications. Ultrasound is, in many respects, a highly practical medical imaging modality. Unlike X-ray imaging or computed tomography, it does not emit harmful ionising radiation. Ultrasound image acquisition protocols are very flexible and quick compared to computed tomography or magnetic resonance imaging. Ultrasound imaging equipment is also highly portable and comparatively cheaper than most imaging modalities. These features make it the imaging modality of choice in a broad variety of clinical situations.

In the last two decades, there has been considerable interest in systems allowing the acquisition of 3D ultrasound data [62]. 3D ultrasound imagery is useful in a variety of clinical applications requiring quantitative assessment of 3D structures and physiological phenomena, such as surgical navigation [233, 216, 23, 239, 40], measurement and modelling of organs [56, 217, 175, 39, 77] as well as diagnosis and assessment of pathologies [25, 144, 187, 58]. By re-slicing a 3D ultrasound data set (i.e. interpolating the pixel values of an arbitrarily oriented hypothetical image plane passing through the acquired volume), it becomes possible to visualise 2D images which could otherwise not have been acquired from the patient because of anatomical constraints. Without 3D imagery, such visualisation could only be achieved mentally, through a substantial intellectual effort on the part of the physician.

Freehand 3D ultrasound is the oldest technique allowing the acquisition of 3D ultrasound imagery. It involves sweeping a conventional 2D ultrasound probe over the area of interest while the scanner acquires and stores images in real-time. This yields a set of irregularly positioned 2D images which must be automatically processed and combined to form a 3D image on a regular lattice or to extract other forms of 3D data. Much of the literature on freehand 3D ultrasound has focused on the development of new image interpolation techniques for this purpose [14, 153, 165, 179, 190, 144, 76, 191, 189, 99, 98, 208, 42]. A crucial assumption underlying the use of any of these image reconstruction techniques is that the position and orientation of the individual 2D ultrasound images with respect to each other are known. Such knowledge is typically gained by tracking the motion of the ultrasound probe using an external sensing device, usually an electromagnetic or optical tracker. As will be discussed in in Section 1.1.2, there are important trade-offs between accuracy and clinical practicality. For example, optical trackers are very accurate but also quite cumbersome to use compared to electromagnetic ones. The necessity of such trade-offs is most unfortunate in a context where both criteria may be of critical importance.

As an alternative or, perhaps more realistically, a complement to such tracking technologies, this thesis considers an approach whereby positional information is inferred from ultrasound image content. This is possible because the ultrasound imagery carries information about the relative motion between the probe and the scanned object. As might be expected with any 2D imaging modality, motion occurring within the ultrasound image plane can be recovered using conventional, widely used 2D image registration techniques [89, 140]. Interestingly, it is also possible to recover out-of-plane motion from conventional 2D ultrasound image data. This is because ultrasound images separated by a small out-of-plane displacement have correlated texture patterns owing to the finite width of the ultrasound beam from which they originate. The extent to which the images are correlated is predictive of the magnitude of this displacement, thereby allowing the estimation of one from the other [31]. A more detailed physical explanation of this remarkable phenomenon, called *speckle decorrelation*, is given in Chapter 2.

This thesis focuses on the use of speckle decorrelation for the recovery of out-of-plane transducer trajectories. The image correlation-based estimation of out-of-plane image displacements required for trajectory recovery is subject to many sources of uncertainty and ambiguity. This thesis is particularly concerned with the following issues:

- Only the magnitude of an out-of-plane displacement can be estimated from a correlation measurement, leaving an ambiguity as to the *direction* (or sign) of this displacement. This causes problems when attempting to recover trajectories that are not monotonic or where image frames intersect.

- A correlation measurement computed over a pair of finitely sized windows of image data is subject to sampling noise, which may in turn add considerable uncertainty to a distance estimate.

- Even when ignoring the variability arising from finite sample size effects, the exact relationship between image correlation and displacement cannot be considered to be one-to-one except in very specific conditions known as *Rayleigh scattering*, which impose strong constraints on the possible structure of the scanned medium. Outside these conditions, the relationship depends strongly on the micro-structure of the scanned medium, whose characteristics are usually not well known. This seriously limits the accuracy of the approach when imaging real tissue (as opposed to so-called *speckle phantoms*).

- As with any tracking system based on differential measures of displacement (as opposed to the absolute measurements of position that might be provided by a tracking device), distance estimation errors arising from the above mentioned sources inevitably accumulate over the course of a long trajectory, resulting in significant drift.

In this thesis, these issues are addressed using new methods which favour the translation of redundant and noisy data extracted from ultrasound images into robust knowledge about the trajectory of the ultrasound probe. The methods developed are generally probabilistic in their formulation, but combinatorial data fusion is explored as well. In the latter case, the application is that of finding an ordering and angulation of the ultrasound images in the sequence which accounts for the set of observed correlation measurements between and within frame pairs. This allows for the detection of non-monotonic probe motion and frame intersections, which could not be done by considering each correlation measurement individually.

In the case of probabilistic methods, these were applied to model the effects of sampling error in the measurement of the correlation between images of a phantom with known micro-structure. By modelling the relationship between correlation and displacement as a

set of probability distributions, it becomes possible to quantify the uncertainty attached to any distance estimate associated with a correlation measurement. This, in turn, leads to the development of a probabilistic measurement fusion approach whereby distance estimates arising from redundant but noisy correlation measurements are combined to obtain a more precise estimate of the out-of-plane probe trajectory.

In order to track the motion of the ultrasound probe in scans of real tissue, the measurement model must be generalised to the case of unknown medium structure. The ultrasound physics literature does not provide any satisfactory model relating medium structure to changes in speckle decorrelation rate. Therefore, a new model which accounts for change in tissue micro-structure is derived from a set of synthetically generated ultrasound images from virtual phantoms of varied micro-structure using a statistical machine learning method. In order to create the synthetic imagery required for training, a new statistical model of medium micro-structure is also proposed. The probabilistic data fusion approach outlined earlier is then extended for use with the learned medium-independent speckle decorrelation model. This extension detects and circumvents the remaining imperfections of the learned speckle decorrelation model through a measurement selection strategy based on statistical hypothesis testing.

The remainder of this chapter is structured as follows. Section 1.1 discusses the relevance of the research presented in this thesis in light of the other approaches that could have been chosen. Section 1.2 lists and details the technical contributions made by this thesis. Finally, Section 1.3 provides a chapter by chapter outline of the document.

1.1 Motivation

In light of the difficulties involved with accurately tracking out-of-plane transducer motion using speckle decorrelation, it is both legitimate and reasonable to ask whether the ultimate goal of quantitative 3D ultrasound imaging might not be better achieved by other means. This section discusses the relevance of the research presented in this thesis by answering three commonly expressed concerns. First, one may question the relevance of the freehand 3D imaging protocol in an age where both corporate and academic research in medical imaging are increasingly focused on the development and use of 3D ultrasound probes. Another legitimate question is whether the study of speckle decorrelation-based transducer tracking is justified in the context of freehand 3D ultrasound given the difficulties associated

with this approach and the availability of other tracking methods. Finally, the relevance of probabilistic methods in the context of image-based tracking may also be questioned. These issues are discussed in turn below.

1.1.1 Why freehand 3D ultrasound?

Freehand 3D ultrasound is an established protocol for acquiring 3D ultrasound data and until recently, it was the only practical protocol for doing so, as the image quality and resolution afforded by the first 3D ultrasound probes were simply not satisfactory. Since their invention [203, 225], however, 3D ultrasound probes based on 2D ultrasound transducer arrays have dramatically improved in terms of achievable image quality and resolution, clinical practicality, frame rate and monetary cost [135, 160]. Almost instantaneous acquisition of volumetric imagery also makes these probes suitable for 4D ultrasound scanning, which is at best very difficult using the traditional freehand technique. In spite of this, freehand 3D ultrasound still offers a number of important advantages over the use of 3D probes:

- It offers a wider, nearly unlimited, field of view;
- It allows more control over the spatial sampling and angulation of the images that are used to create the volume;
- The structure of the acquired data makes it easy to discard or add image data if required;
- The 2D probe it uses is comparatively cheap and ubiquitous.

In addition, innovations which facilitate or enhance the use of 2D ultrasound probes are highly relevant because 2D ultrasound is, and will remain for some time, part of the standard instrumentation for many diagnostic and interventional protocols. Image-based transducer tracking is a likely candidate for such an innovation.

1.1.2 Why is speckle decorrelation-based ultrasound probe tracking relevant?

In order to understand the appeal of this approach, it is important to understand the strengths and limitations of alternative transducer tracking methods, most of which rely on external optical or electromagnetic tracking devices. Other types of devices, such as

mechanical or robotic arms [88, 105, 18], acoustic trackers [80] and vision-based systems [6] have also been used in this context, but because of operating restrictions, cost and accuracy considerations, they are not as widespread as optical or electromagnetic tracking technologies. More elaborate comparisons of the tracking technologies currently in use within the medical context are provided in [201, 21]. The present discussion focuses on the more ubiquitous optical and electromagnetic technologies.

Optical tracking typically involves an infra-red stereo camera which tracks the 3D position of markers attached to the ultrasound probe. These may be active infra-red LEDs or passive markers which reflect infra-red light emitted by sources attached to the stereo-camera. Because of their high accuracy [114], optical trackers have become very popular in the operating room [233, 23]. The most serious limitation of optical trackers is that they require an unobstructed line of sight between the markers and the camera, which can restrict the scanning protocol and be very cumbersome, particularly when the environment is cluttered or crowded [21]. This shortcoming also rules out using the technology for imaging inside the body, as with protocols requiring transrectal or intra-vascular ultrasound probes.

Electromagnetic trackers work by generating a magnetic field through a stationary set of emitting coils. The motion of the ultrasound probe is then inferred from measurements of the electrical current induced in a set of receiving coils rigidly attached to it as it moves within this field. Electromagnetic trackers do not have any line-of-sight requirements and miniaturisation has recently made them suitable for interventions where imaging takes place inside the body [239, 40]. However, readings from such trackers are highly vulnerable to electromagnetic field distortions caused by the presence of metal or interference with other devices in the environment [22, 156]. Even under ideal circumstances, their accuracy is considerably worse than that of optical trackers [21].

The choice between these two tracking technologies then becomes a matter of optimising the trade-off between accuracy and clinical practicality, which inevitably leads to sacrifices one way or the other. In addition, the use of any tracking device for freehand 3D ultrasound acquisition requires that two important calibration problems be solved:

1. **Temporal calibration**: the stream of data output by the base unit of the position sensor must be synchronised with the ultrasound data acquisition device (e.g. a frame-grabber) in order to establish the correct associations between image frames and position data. This synchronisation is generally achieved by finding the time

delay which best maps observed changes in the moving ultrasound image plane to the changes in position recorded by the tracker [169, 165, 78, 181].

2. **Spatial calibration**: the spatial transformation between the tracked beacon and the image-centred reference frame must be estimated in order to track the motion of the image plane (as opposed to that of tracking device). This problem is more complicated than temporal calibration and has been the subject of vast amounts of literature [146, 182, 97, 96]. Typical techniques for solving the problem involve the acquisition of ultrasound imagery of a phantom with a special, perfectly known geometry, with the phantom and scanning protocol designed in such a way that the imagery can be used to infer the complete set of transformation parameters relating the image-centred reference frame to that of the position tracking device.

These calibration procedures are necessary, and often times tedious, time-consuming and error-prone, though much effort has been invested in simplifying them.

From this discussion, it transpires that the use of position tracking devices has many shortcomings. All technologies exhibit noise (mainly high frequency jitter) to some degree, require careful calibration, and each one also exhibits a different, serious set of limitations which can have an impact on the quality of the 3D information extracted from the imagery. In addition, none of these devices can recover motion that occurs within the scanned object, arising from physiological phenomena or from the pressure of the ultrasound probe on the target of interest. For all these reasons, there has been interest in image-based tracking techniques which do not rely on, or do not rely solely on, external devices.

As described earlier, image-based tracking of motion within the image-plane can be achieved using conventional image registration methods. These have often been used for freehand 3D ultrasound in conjunction with position tracker data in order to overcome accuracy limitations arising from noise, physiological motion and probe pressure induced tissue deformations [180, 118, 218, 192]. Hossack et al. [91, 71] also proposed a rather unique system which exploits conventional image registration techniques for both in-plane and out-of-plane motion tracking. The system is based on a modified 2D ultrasound probe which, in addition to the usual imaging array, has two small "tracking arrays" positioned perpendicularly to the central one, resulting in three images being acquired at any given time. Image registration applied to the images acquired by the tracking arrays achieves out-of-plane motion tracking. While this solution requires no external tracking device, it

does require substantial modifications to standard 2D imaging instrumentation.

In contrast, out-of-plane transducer tracking using speckle decorrelation does not impose any special hardware requirements on the imaging system. Used in conjunction with image registration techniques for in-plane motion tracking, it makes possible acquisition of freehand 3D ultrasound data without resorting to an external tracking device [166, 94]. While such image-based tracking approaches are vulnerable to many sources of cumulative error and may not, in practice, share the same large scale accuracy as optical, or even magnetic trackers, they are not vulnerable to the type of high-frequency noise that corrupts measurements from such trackers. Because of this, and because their use is not restricted by the same constraints, image-based tracking and device-based tracking can also be seen as complementary to each other, leading to the development of hybrid systems which exploit their respective strengths [95, 92, 124].

It is with these motivating ideas in mind that the research presented in this thesis was conducted. The focus of the thesis is on the estimation of the out-of-plane component of transducer trajectories using speckle decorrelation. In order to facilitate experimentation with the novel components developed in this research, it was assumed throughout this thesis that all imagery was aligned in-plane and that no source of positional data other than the ultrasound imagery was available. While these assumptions may seem somewhat restrictive in the context of the applications envisioned for this research, it was felt that they were necessary in the context of this work. The first assumption (in-plane image alignment) allowed the proposed out-of-plane displacement estimation techniques to be evaluated under well controlled conditions, independently of an error-prone image registration process, which was essential for understanding their behaviour and comparing them with alternatives. The second assumption (no outside source of positional data) was made because it was felt that the limits of the image-based approach should be explored in detail before attempting to broaden them. For these reasons, loosening these assumptions falls outside the scope of this thesis, but can most certainly be envisaged in the context of developing a more complete tracking system.

1.1.3 Why is statistical modelling a good idea in this context?

While some of the difficulties concerning the speckle decorrelation approach have been the subject of scientific research over the last five years or so [133, 204, 86, 72, 94, 121], this

thesis is unique in exploring and addressing them from a statistical standpoint. This particular perspective was heavily inspired by the field of mobile robotics, where trajectory estimation problems of a similar nature to that explored in this thesis are common and statistical modelling is ubiquitous. In that context, it has been shown that the effects of motion estimation errors arising from measurement errors may be alleviated, or at least evaluated, by careful statistical modelling, paying particular attention to the distribution of errors arising from the measurement process and the transformations leading from sensory measurements to trajectory recovery [202]. This thesis does just this by proposing a probabilistic model for the relationship between image correlation and out-of-plane transducer displacement.

Over the years, the mobile robotics community has developed a thorough methodology for leveraging the power of such probabilistic models, allowing the data fusion process which combines the (possibly redundant) measurements to place more weight on accurate, precisely acquired measurements and less weight on error-prone ones, thereby minimising the risk and magnitude of error [202, 139, 49, 150]. This thesis borrows a proven measurement fusion methodology proposed by Lu and Milios [139] in the context of mobile robot localisation in order to integrate redundant noisy correlation measurements into a single, precise trajectory estimate.

One difficulty with the speckle decorrelation approach is that ultrasound physics do not provide any satisfactory model relating image correlation to displacement for the case of arbitrary tissue structure. In similar scenarios, the mobile robotics literature has turned with great success towards statistical learning-based approaches which automatically infer an empirical measurement model from examples [200, 55]. Similarly, this thesis proposes to use statistical machine learning to build a new measurement model which accounts for change in tissue micro-structure from a set of synthetically generated ultrasound images of virtual phantoms with varied micro-structure.

Another advantage of statistical modelling is that it allows for the application of principled hypothesis testing strategies in order detect incompatibilities between two existing models of the same quantity and reject candidate measurements when this occurs. In mobile robotics, this idea is frequently applied to the problem of data association, where the correctness of a correspondence between a recently measured landmark and landmarks in an existing environmental map must be established [28]. Similarly, this thesis uses hypothesis testing in order to assess whether a distance estimate derived from an imperfect

correlation model is compatible with an existing uncertain estimate of the probe trajectory.

1.2 Contributions

As a result of the design choices outlined in the above section, this thesis (and related peer-reviewed publications) makes the following contributions (in order of appearance in this document):

1. **An approach based on combinatorial optimisation and robust statistics for recovering non-monotonic motion and frame intersections in an ultrasound scan [126].** Correlation measurements between localised image patches in an ultrasound image sequence are subject to two constraints: (i) the set of pairwise distance estimates obtained among any triplet of image frames must agree with each other geometrically and (ii) image planarity implies that the set of distances between all pairs of corresponding image patches of two frames must define a rigid transformation. This thesis proposes a new frame ordering algorithm based on the solution to a travelling salesman problem combined with a robust least-median-of-squares rigid transformation fitting method to enforce these constraints. This allows the detection of direction reversals and frame intersections in the ultrasound scan under consideration.

2. **A speckle decorrelation model which explicitly accounts for the stochasticity of the relationship between the sample correlation coefficient and elevational separation [125, 126].** Because correlation measurements are collected over finitely sized image windows, they are inevitably subject to some variability caused by sampling error. This error, in turn, propagates to the distance estimates derived from the measurements. While the existence of such errors has been recognised by others [133, 204], this thesis proposes the first speckle decorrelation model to explicitly account for them. For a given transducer, this new model can be estimated from ultrasound imagery of a speckle phantom using a maximum entropy approach. It does not require any more data than its deterministic counterparts but provides a richer representation of the relationship between image correlation and transducer displacement.

3. **A probabilistic measurement fusion approach for out-of-plane trajectory estimation [125, 126].** Given the wealth of pairwise image correlation measurements that can be computed from an ultrasound image sequence, finding an optimal way to combine them into a trajectory estimate is an important problem. For this purpose, previous work has mostly focused on selecting a minimal number measurements which are least error-prone [204, 85, 94]. Instead, this thesis proposes that the uncertainty associated with each available measurement be quantified by a probabilistic speckle decorrelation model and that a trajectory then be computed by weighting the contribution made by each measurement according to its associated uncertainty. In addition to potentially improving the accuracy of the tracking process, this probabilistic approach can be used to quantify the uncertainty associated with the end result (the estimated trajectory) and therefore lends itself well to extensions involving fusion with other sources of data or decision making processes.

4. **A fractal multi-dimensional point distribution model for simulating ultrasound scatterer distributions representative of the micro-structure associated with a variety of media [128, 129].** Many point processes used to model the spatial organisation of ultrasound scatterers in the context of ultrasound simulation work are 1D [123, 43, 195] whereas ultrasound image processing tasks require a 2D or 3D signal. The few existing models which provide multi-dimensional modelling capacity tend to be limited in the breadth of spatial organisations they can represent or else are difficult to tune or sample [220, 152, 193, 194]. This thesis builds on an existing 1D model which has an intuitive and flexible parameterisation [123, 43] by mapping its output to multiple dimensions using a Hilbert space-filling curve. It is shown that this preserves the desired spatial organisation characteristics of the sample with no changes to model parameterisation. The approach is used in the simulation of 3D ultrasound image sequences of varied media whose characteristics are reliably predictable from the model parameters and agree with predictions from ultrasound physics. Such synthetic imagery can be used for the validation of algorithms for various ultrasound image processing tasks, including segmentation and motion estimation. A further application of the model is demonstrated in the context of the next contribution.

5. **A learning-based approach for the discovery of a tissue invariant speckle**

decorrelation model driven by synthetic imagery [127]. Most speckle decorrelation models in the literature are only valid under Rayleigh scattering conditions, i.e. when the micro-structure of the medium under consideration contains a large number of randomly positioned ultrasound scatterers. Outside these conditions, the correlation length of the imagery (the rate at which imagery decorrelates over distance) changes. Because of this, estimation of out-of-plane transducer displacement from imagery of real tissue has relied strongly on filters which discard image data where the Rayleigh scattering assumption is not met [219, 86, 177, 178]. This thesis proposes instead that the speckle decorrelation model be dynamically adapted to the local characteristics of the medium, resulting in a much denser and more accurate set of displacement estimates. Adaptation to local medium characteristics is achieved through a learning process whereby synthetic image sequences of a variety of media with known out-of-plane transducer trajectory are used to establish a regressive relationship between observable statistical image features and local out-of-plane correlation length. The resulting regressor can then be applied to statistical features measured in new imagery from a real ultrasound transducer. This approach is one of only two existing methods for dynamically adapting a speckle decorrelation model to the local characteristics of the medium; the other is a heuristic method proposed by Gee et al. [72]. The approach proposed in this thesis offers two unique advantages over the state-of-the-art: (i) the resulting regressor provides an estimate of the variance associated with the predicted correlation length and (ii) it is data-driven rather than heuristic, making its behaviour easier to analyse and eventually improve.

6. **An iterative measurement selection and fusion approach for the recovery of out-of-plane transducer trajectories from scans of real tissue.** The probabilistic measurement fusion approach labelled as contribution 3 is extended to allow for the detection and principled rejection of erroneous distance estimates arising from unavoidable imperfections in the proposed tissue invariant speckle decorrelation model. The proposed strategy begins with a highly uncertain trajectory estimate derived from a small number of correlation measurements. Subsets of the remaining measurements are iteratively tested for compatibility with the current trajectory estimate via a statistical hypothesis test and either integrated to the trajectory estimate or discarded. Though measurement selection has been discussed in the context of

decorrelation based transducer tracking [94], the case of model failure has never been explicitly addressed before.

1.3 Thesis structure

While the application domain of this thesis is biomedical engineering, some of its readers will likely be electrical engineers and computer scientists. Much of the work presented in this thesis relies on phenomena related to ultrasound image formation process which may not be familiar to such readers. Therefore, Chapter 2 begins this thesis by providing the necessary technical and contextual background. In this chapter, the fundamental principles governing the formation of ultrasound images are described, with emphasis on the phenomenon of backscattering which leads to the formation of speckle patterns. A simple mathematical model of speckle pattern formation, as well as its consequences on the statistics of such patterns, are described. This, in turn, leads to the reasoning behind the use of speckle decorrelation for the estimation of out-of-plane transducer motion. Following the technical discussion, this chapter also provides a short literature review describing the challenges and advances resulting from the study of this approach. More detailed and focused literature reviews will be presented as part of every other chapter.

Chapter 3 introduces the global framework that will be used throughout this thesis for the image-based recovery of out-of-plane transducer motion. In particular, it explains how image correlation measurements are combined within each pair of image planes and across different pairs of frames in order to recover intersections between frames and changes in the direction of transducer motion using methods from combinatorial optimisation and robust statistics. A short simulation experiment is presented, showing that the proposed combinatorial frame ordering approach is highly robust to noise. More extensive demonstrations of the validity of these techniques are implicitly given in Chapters 4 and 7 where they are used as part of the experimental framework for the contributions presented therein.

Chapter 4 presents a novel probabilistic speckle decorrelation model which accounts for the variability of correlation measurements associated with a given elevational displacement. It also describes an approximate maximum likelihood data fusion approach for estimating the relative poses of a set of corresponding image patches from redundant and uncertain distance estimates obtained from the correlations measured between an arbitrary number of pairs in the sequence. Combined with the global framework presented in Chapter 3, this

allows the estimation of the out-of-plane trajectory of the ultrasound probe from speckle phantom imagery. Trajectory estimation experiments conducted on synthetic and real imagery of speckle phantoms are described where the probabilistic approach is compared to two deterministic base-line approaches and found to give good results in terms of accuracy. Furthermore, the effects of variations in the type and speed of the trajectory undergone by the probe on the accuracy of the recovered trajectory are examined in detail.

The rest of the thesis focuses on generalising the ideas presented in Chapter 4 to imagery of arbitrary tissue. In this context, the relationship between image correlation and probe displacement can no longer be expected to be static. The approach taken is to automatically discover a relationship between a set of image statistics computed within the image-plane and variations in elevational decorrelation rate by observing a pool of synthetic imagery representing a variety of different tissue types with associated ground truth trajectories. In order to enable the creation of this pool of synthetic imagery, Chapter 5 describes a new approach for simulating the micro-structure of a 3D virtual imaging target whereby the density and spatial organisation of ultrasound scatterers can be tuned in an intuitive and highly flexible manner. The synthetic imagery generated from these virtual targets is shown to display a broad range of first and second order statistics which are predictable from the model parameters and in agreement with ultrasound physics.

Chapter 6 then describes a machine learning approach for establishing a probabilistic transducer-invariant regression model of the relationship between out-of-plane correlation length and locally observable image statistics using the pool of varied synthetic data generated using the method outlined above. Through distance estimation experiments on synthetic and real animal tissue imagery acquired from different types of transducers, the approach is shown to generalise well across changes in transducer specifications and to yield results of accuracy comparable to the state-of-the-art.

In Chapter 7, correlation length estimation is integrated within the probabilistic speckle decorrelation model of Chapter 4. Modelling inaccuracies arising from persistent image features such as specular reflections and boundaries between different tissue types are addressed through a new trajectory estimation algorithm which begins with a highly uncertain trajectory estimate based on a minimal number of distance estimates and progressively integrates new measurements into its maximum likelihood estimate as they are approved by a measurement selection process based on statistical hypothesis testing. The effectiveness of the measurement selection approach is illustrated by experiments on imagery of animal

tissue in the context of monotonic and non-monotonic transducer trajectories. Chapter 8 concludes the thesis with a summary and discussion of the aforementioned research and discusses numerous directions for future work.

Chapter 2

Background and literature review

This thesis is about the estimation of out-of-plane ultrasound transducer motion through the exploitation of an imaging phenomenon known as *speckle decorrelation*. A basic understanding of the physics governing this phenomenon is necessary in order for the reader to follow the discussions presented in further chapters of this thesis, and for this, the reader should also be initiated to the fundamental principles of ultrasound image formation and interpretation. This chapter is meant to provide the technical and contextual backgrounds necessary for understanding the contributions made through this thesis.

The first three sections of this chapter briefly describe how ultrasound images like that shown in Figure 2.1 are formed and how their textural appearance relates statistically to the underlying structure of the target. They are not intended to provide an exhaustive treatment of ultrasound physics or ultrasound image interpretation (for this, the interested reader is referred to textbooks such as [117, 36]), but rather to provide an introduction to ultrasound imaging for the uninitiated reader. The emphasis is on the principles of ultrasound imaging physics and interpretation which are directly relevant to this thesis. Specifically, Section 2.1 describes the flow of data within a modern ultrasound machine, from ultrasound transmission to display. Section 2.2 describes how acoustic waves interact with the scanned medium and how these interactions are manifested in the received echo signals, paying particular attention to backscattering phenomena and the formation of speckle patterns that results. Section 2.3 discusses the mathematics of backscattering, the resulting first order statistics of ultrasound speckle and speckle decorrelation phenomenon of interest in this thesis. Following this technical discussion, Section 2.4 explains how

speckle decorrelation has been exploited in the context of freehand 3D ultrasound research, outlines the current state of the art and provides a context for the work presented in this thesis through a short literature review. Aspects of related work which are of particular relevance to the contributions made in this thesis are discussed in more detail in each chapter.

2.1 Ultrasound imaging: instrumentation and data flow

Consider the ultrasound image shown in Figure 2.1. This particular image depicts a cross section of a 12 week human embryo. It was produced by sending short ultrasound signals from different starting points and/or in different directions through the body and collecting the ultrasound signals echoed by the different structures in the organs. Assuming that the speed of sound is known and constant through the imaged tissues, the time delay between the initial transmission of an ultrasound pulse and a given echo determines the location of the structure which caused it. This information is displayed to the user as the location of a pixel in the image whose brightness is indicative of the strength (amplitude) of the echo. The type of ultrasound image shown in Figure 2.1 is sometimes called a B-scan, where the "B" refers to the fact that "brightness" is used to visually quantify echo strength.

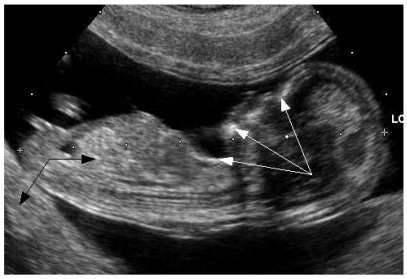


Figure 2.1 Example ultrasound image depicting a 12 week old human embryo [37]. White arrows point to areas of specular reflection. Black arrows point to areas of speckle. The significance of these different types of ultrasound echo patterns is discussed in Section 2.2

The diagram shown in Figure 2.2 illustrates the data flow which leads to the formation

of such a B-scan. Modern ultrasound transducers (Figure 2.2(a)) consist of an array of piezoelectric crystals (called *elements*) which can be electrically stimulated to produce an acoustic wave or, conversely, mechanically stimulated to produce an electrical signal. The transmit beam former (Figure 2.2(b)) produces a focused ultrasound beam by stimulating a group of adjacent transducer elements to produce short ultrasound pulses at slightly different times, allowing them to meet at a single point. Similarly, the echoes received by groups of adjacent transducer elements can be delayed by slightly different amounts of time and summed by the receive beam former (Figure 2.2(d)) to obtain a strong signal representative of the interaction between the pulse and the medium at a particular point in space. As such, ultrasound imaging systems belong to the family of *coherent imaging* systems, along with laser and synthetic aperture radar imaging systems. Linear transducer arrays produce focused beams using groups of transducer elements centred about different locations, for an image geometry that is rectangular. Linear phased arrays use the same group of transducer elements with different pulse delay patterns in order to produce ultrasound beams that travel in different directions, yielding a sector shaped image (or a rectangular shaped image on a polar coordinate grid).

The echo signals thus created are known as *radio-frequency* (RF) vectors. Here, the term "radio-frequency" is used in analogy with short wave radio communications to denote the fact that these signals are essentially an amplitude signal modulated by a carrier wave of frequency similar to the emitted ultrasound pulse. The RF vectors are subsequently envelope detected (Figure 2.2(e)) (i.e. demodulated), yielding signals called *A-lines* which can be interpreted in terms of the echogenicity (or the propensity to reflect sound waves) of structures in the scanned medium, at the cost of losing phase information. A post-processing module (Figure 2.2(f)) compresses the dynamic range of the envelope detected signal for display purposes. This is done by taking the logarithm of the envelope and is sometimes preceded by other non-linear mappings selected by the user ((Figure 2.2(h)). The scan converter (Figure 2.2(g)) then produces a single image by combining the log-compressed A-lines according to their different starting points and directions of acquisition. Essentially, the scan converter carries out the spatial shifting, sub-sampling and interpolation operations necessary to form a geometrically meaningful image on the rectangular pixel grid of the ultrasound machine's display monitor (Figure 2.2(i)). In the case of linear phased array, for example, the scan-converter is reponsible for mapping the image originally acquired in polar coordinates to the Cartesian pixel grid of the display.

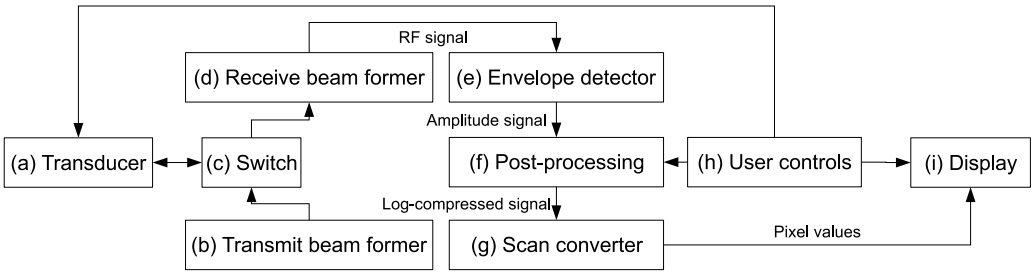


Figure 2.2 Data flow for B-scan formation and display in a modern ultrasound machine. This diagram is an adaptation of similar diagrams found in textbooks such as [117, 159].

2.2 Reflection, backscattering and speckle

Echoes are produced when the transmitted acoustic pulse encounters inhomogeneities in the medium across which it is travelling. In this context, inhomogeneities, or *scatterers*, refer to structures whose compressibility or mass density differ significantly from those of the surrounding medium. When the acoustic wave hits such a structure, the latter responds to the oscillatory pressure by compressing and expanding and/or by moving back and forth with respect to the homogeneous medium [7, 101]. In this process, some acoustic energy is lost (or absorbed) by the scatterer, some is transmitted further along the original path of the incident wave to hit further scatterers, and some is reflected in the form of echoes.

When the size of the scatterer is much larger than the ultrasound pulse wavelength, sound is reflected back to the transducer in a single direction, producing specular reflection. This often occurs at organ boundaries, such as illustrated by the embryo head and neck structures pointed to by the white arrows in Figure 2.1, or when the sound hits a man-made structure such as the bottom of a water bath or a biopsy needle. When the scatterer is smaller than the pulse wavelength, the echo is backscattered in all directions. Because the ultrasound beam has a finite spatial extent, many such scatterers can be imaged in a given location. The echoes backscattered from all these scatterers interfere positively and negatively with one another, creating a pattern called speckle, which becomes the grainy texture characteristic of ultrasound B-scans (see the areas indicated by black arrows in Figure 2.1).

The geometry of ultrasound speckle does not correspond to that of any existing macro-structure in the scanned object. As such, speckle is often considered to be a type of noise, as evidenced by the continued abundance and variety of methods proposed for "speckle reduction" and "de-speckling" of ultrasound imagery in the last two or three decades [12, 214, 138, 115, 53, 250, 83, 2, 244, 147]. De-speckling is often considered to be a necessary step towards the improvement of image contrast for visualisation [44, 137] or in carrying out other image processing tasks which rely on macroscopic geometry, such as multi-modality image registration [131], delineation of organ boundaries [209] or interpolation of freehand 3D ultrasound data [190].

Though it can make image interpretation difficult, speckle is not random noise in the conventional sense: if a given transducer scans a given target in a given spatial configuration twice, the speckle patterns in the two images should be the same. In that sense, speckle is a repeatable signal which conveys information about the micro-structure of the target medium and the physical characteristics of the transducer. Although speckle patterns cannot be used to recover the exact spatial configuration of the underlying scatterers, their textural appearance and statistics are useful cues with respect to the general spatial organisation of the scatterers. Since the spatial organisation of scatterers varies across different types of tissue, speckle patterns actually have diagnostic value. Textural analysis of speckle was first exploited by physicians [116, 34, 20, 172] and, subsequently, using computer algorithms for diagnostic tissue characterisation tasks [16, 198, 212, 29, 211, 238] and image segmentation tasks [185, 245]. Likewise, in this thesis, speckle is treated as a meaningful signal, not as noise. Statistical analysis of speckle will be used for both tissue and transducer characterisation. The mathematical justification and fundamental assumptions underlying the particular type of analysis used are provided in the next section.

2.3 The statistics of ultrasound speckle

This section discusses a simple mathematical model of ultrasound image formation and its theoretical consequences on the first and second order statistics of ultrasound speckle. Since the formation of speckle patterns is common to all coherent imaging modalities, it comes as no surprise that many of these considerations were first addressed outside the realm of ultrasonics. Therefore, many of the works cited in this section concerning the statistics of speckle patterns originate from the optical systems literature. In what follows, let x, y and

z denote the three dimensions of Cartesian space, as illustrated in Figure 2.3. The *axial* direction is defined as the direction of ultrasound wave propagation. For images acquired with a linear transducer array, this corresponds to the y direction. For other geometries (such as the sector images produced with linear phased arrays), the axial direction varies as a function of the location within the image. The *lateral* direction is defined as the direction perpendicular to the axial direction, within the image plane. For images acquired with a linear transducer array, this corresponds to the x direction. The z direction is called the *elevational* direction and is perpendicular to the image plane. In the remainder of this thesis, out-of-plane rotations about the x axis will be referred to as *tilt* and out-of-plane rotations about the y axis will be referred to as *yaw*.

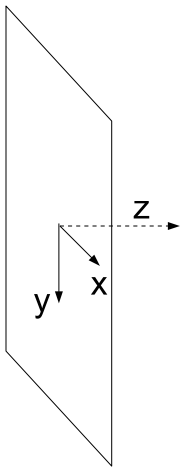


Figure 2.3 The chosen coordinate system illustrated with respect to an ultrasound image frame.

In ultrasound imaging, echoes generated by a given scatterer sometimes hit other scatterers before returning to the transducer. When the reflections are strong, this multiple reflection phenomenon is responsible for reverberation artefacts sometimes observed in B-mode imagery. Throughout this thesis, it is assumed that the amplitude of such multiply reflected signals is negligible. This assumption corresponds to *weak scattering* conditions. Under the assumption of weak scattering conditions, the ultrasound imaging process can be modelled as a linear system. If, without loss of generality, it is assumed that the probe

is linear, the wave travels in the y direction and the RF signal is given by

$$\varrho(x, y, z) = (h * \varsigma)(x, y, z), \tag{2.1}$$

where h is the transducer's point spread function (PSF) and ς is a map of acoustic inhomogeneities, or *scatter function* describing the medium. Thorough physical justification for the validity of this linear systems model, as well as in-depth treatment of the computation of the PSF h for given transducer specifications are provided in [107, 155]. For the purpose of the current discussion, h is a pulse which can be qualitatively approximated as a sinusoidal wave with zero mean modulated by a Gaussian envelope. The envelope of the PSF is often referred to as the *resolution cell*. The scatter function ς is defined as a collection of N_s point scatterers such that

$$\varsigma(x, y, z) = \sum_{i=0}^{N_s} e_i \text{Dirac}(x - x^{(i)}, y - y^{(i)}, z - z^{(i)}), \tag{2.2}$$

where Dirac denotes the Dirac impulse function and e_i, $x^{(i)}$, $y^{(i)}$ and $z^{(i)}$ denote the strength (or *echogenicity*) and x, y, z location of scatterer i, respectively.

The first order statistics of ϱ are determined by how densely the scatterers occupy the resolution cell and how they are organised spatially with respect to the incident wave. If the scatterers are located randomly (according to a uniform distribution) in space with strengths independent of location, ϱ will, on average, have zero mean. This is because the ultrasound pulse is zero mean and the echoes emitted by the individual scatterers have uniformly distributed phase as a result of their random spatial locations, resulting in random positive and negative interferences. If, in addition, the number of scatterers located within the resolution cell is large, ϱ will have a (zero-mean) Gaussian distribution as a consequence of the central limit theorem [79]. These asymptotic conditions are known as *Rayleigh scattering* conditions and the resulting speckle pattern is called *fully developed speckle*.

Let us now consider the envelope detected (or *echo amplitude*) signal $A(x, y, z) = |\varrho(x, y, z) + \sqrt{-1}\breve{\varrho}(x, y, z)|$, where $\breve{\varrho}$ denotes the Hilbert transform of ϱ along the direction of wave propagation (y). Under Rayleigh scattering conditions, A can be shown to be

Rayleigh distributed [79] with probability density function

$$p(A) = \frac{A}{\psi^2} \exp\left(-\frac{A^2}{2\psi^2}\right), A \geq 0, \tag{2.3}$$

where ψ^2 depends on the average signal energy determined by the distribution of scatterer strengths. The *echo intensity signal* $I = A^2$, which will sometimes be used in this thesis, has an exponential distribution [79] with probability density function

$$p(I) = \frac{1}{2\psi^2} \exp\left(-\frac{I}{2\psi^2}\right), I \geq 0. \tag{2.4}$$

Notice how the distributions of A and I depend only on ψ^2, which is a function of the echogenicity of the scatterers. An important consequence of this is that first order statistics of these distributions, when normalised to account for different signal energy levels, are constants. Estimates of such quantities obtained from a patch of ultrasound imagery thus lend themselves well as features for tissue characterisation tasks. Examples of first order statistical identities commonly used to detect fully developed speckle (or departures from it) are [27, 229, 167]:

$$\frac{E\{A\}}{\sqrt{VAR\{A\}}} = \sqrt{\frac{\pi}{4-\pi}} \approx 1.91, \tag{2.5}$$

$$\frac{E\{I\}}{\sqrt{VAR\{I\}}} = 1, \tag{2.6}$$

$$\frac{E\{I - E\{I\}\}^3}{VAR\{I\}^{\frac{3}{2}}} = 2. \tag{2.7}$$

The distribution of echo envelope samples changes outside Rayleigh scattering conditions. Many models have been suggested to describe the statistical distribution of the echo envelope in more general settings [102, 52, 195, 196, 197, 57, 112, 19], most of which can be described as mathematical generalisations of the Rayleigh probability density function of (2.3). Rather than describing those models in mathematical detail, Figure 2.4 illustrates the qualitative behaviour of the echo envelope distribution in different physical conditions. The Rayleigh probability density function corresponding to fully developed speckle is shown along with the three commonly used generalisations discussed in [52], all using the same signal energy related parameter ψ: the K distribution, the Rician distribution and the

homodyned K distribution. The so-called K distribution [103] models the decrease in the distribution's mode and the fattening of its tail with decreasing numbers of scatterers per resolution cell, or their tendency to cluster in space. The Rician distribution [102], with its larger mode and decreased skewness, corresponds to the case of non-random, positive interference between backscattered signals caused by periodic arrangements of scatterers along the direction of wave propagation. Combining the two types of non-Rayleigh phenomena accounted for by the K and Rician distributions leads to the homodyned K distribution model [103, 104].

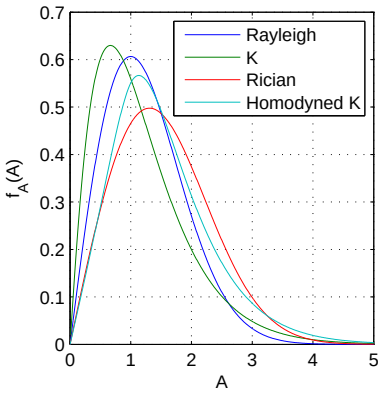


Figure 2.4 Probability density functions which are members of the homodyned K distribution family and have been used to model the first order statistics of envelope detected ultrasound signals.

So far, this discussion has been limited to first order statistics, i.e. the distribution of the individual samples of ultrasound signals. In the context of ultrasound speckle, second order statistics, which describe the distribution of the differences between ultrasound signal samples as a function of their respective locations, are very important because these samples cannot be treated as statistically independent. Indeed, the finitely sized resolution cells centred about two neighbouring points overlap in space and therefore image much of the same scatterers, causing the signals sampled at either point to be mutually correlated. Correlation decreases with the amount of overlap as a function of distance; this phenomenon is known as *speckle decorrelation*.

The second order statistics of ultrasound speckle are described in terms of autocorrelation functions. For simplification, the imaging system's PSF h is assumed to be separable,

i.e.

$$h(x, y, z) = h_x(x)h_y(y)h_z(z), \tag{2.8}$$

an assumption which is very reasonable for ultrasound imaging [229]. Since this thesis focuses precisely on motion estimation in the out-of-plane (elevational) direction, further analysis in this section will be carried out using the z axis as the direction of interest, but similar results apply to the lateral (in this case, x) and, assuming slow changes in beam shape, to the axial (in this case, y) directions. Following classical linear systems theory, the autocorrelation function of the RF signal ϱ along the elevational direction is given by [161]

$$\begin{aligned} \mathfrak{R}_\varrho(z_1, z_2) &= E\{\varrho(z_1)\varrho(z_2)\} & (2.9)\\ &= \mathfrak{R}_\varsigma(z_1, z_2) * h_z(z_1) * h_z(z_2), & (2.10) \end{aligned}$$

where $\mathfrak{R}_\varsigma$ is the autocorrelation function of the random point process which generated the scatter function ς. If it is assumed that the PSF is shift invariant (or slowly shift variant if the analysis is done in the axial direction [155]), that the strength of scatterers is independent of their spatial position, and that the average scatterer strength is given by $\bar{e}$, then this point process is shift invariant and [229]

$$\begin{aligned} \mathfrak{R}_\varrho(z_1, z_2) &= \mathfrak{R}_\varrho(\delta) \\ &= \bar{e}^2 h(-\delta) * \mathfrak{R}'_\varsigma(\delta) * h(\delta), & (2.11) \end{aligned}$$

where δ represents a displacement along the z direction, $\mathfrak{R}'_\varsigma(\delta)$ describes the autocorrelation of the micro-structure of the medium independently of scatterer strength. Under Rayleigh scattering conditions, the positions of the scatterers are uncorrelated with each other, and $\mathfrak{R}'_\varsigma$ becomes a Dirac impulse so that the autocorrelation function of the RF signal becomes

$$\mathfrak{R}_\varrho(\delta) = \bar{e}^2 h(-\delta) * h(\delta). \tag{2.12}$$

Noting that under these conditions ϱ is zero-mean, the normalised correlation coefficient ρ_ϱ of fully developed speckle patterns as a function of elevational displacement is given by

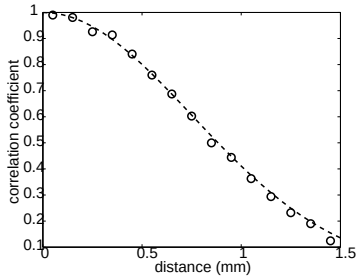
$$\rho_\varrho(\delta) = \frac{\mathfrak{R}_\varrho(\delta)}{\mathfrak{R}_\varrho(0)}. \tag{2.13}$$

(2.12) and (2.13) show that under Rayleigh scattering conditions, the second order statistics of speckle patterns depend exclusively on the transducer's characteristics, i.e. the PSF h [229]. Clearly, this remains true of the normalised autocorrelation functions of the envelope and echo intensity signals deterministically derived from the RF signal. Furthermore, it can be shown that in these conditions, if the resolution cell is approximately Gaussian in shape, the normalised autocorrelation function, henceforth also to be referred to as the *decorrelation curve*, of the echo intensity signal I is a Gaussian function of distance [166, 155]. An example of such a decorrelation curve is shown in Figure 2.5. Chen et al. [31] have shown that this is a very good approximation even for more realistic PSFs. The decorrelation curve associated with the envelope signal A takes a more complicated mathematical form, but for practical purposes, is very similar in shape to that of the echo intensity signal [227, 72].

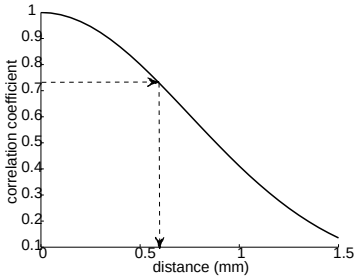
2.4 Transducer motion estimation using speckle decorrelation

The independence of the decorrelation curves from the medium under Rayleigh scattering conditions suggests that empirical second order statistics of fully developed speckle patterns can be used to estimate the unknown elevational separation between two different ultrasound images acquired in close proximity by the same transducer. A typical scheme for doing this is illustrated in Figure 2.5. In a calibration step, a *speckle phantom*, i.e. an object which approximates Rayleigh scattering conditions and contains a large number of randomly located ultrasound scatterers, is scanned at known regular elevational intervals and a transducer-specific decorrelation curve is fitted to the set of resulting correlation measurements. This curve can then be used to estimate distances from correlation measurements.

This idea is the foundation block of the research presented in this thesis and was initially developed by Chen et al. [31]. In their seminal paper, the authors derive the normalised elevational autocorrelation function of the echo intensity signal in the focal zone of a typical clinical transducer from first principles and show that it is well approximated by a Gaussian curve. Failing the availability of exact transducer specifications, they propose that a speckle phantom be scanned at regular elevational intervals and a Gaussian decorrelation curve be computed from the resulting ultrasound imagery. Chen et al. then propose to estimate the spacing between two frames by fitting this transducer dependent Gaussian curve to a



(a) Calibration step. The circles represent the average of correlation measurements collected among images separated by known distance intervals. A decorrelation model, shown by the dashed line, is fitted to this collection of measurements. In this case, the model is a Gaussian curve, but other types of curves, including piecewise linear models, can also be used.



(b) Distance estimation step. Correlation between two images is measured and the corresponding elevational separation is read off the decorrelation curve estimated during calibration.

Figure 2.5 A typical scheme for estimating elevational displacement from speckle decorrelation.

local estimate of the normalised autocorrelation function based on a neighbourhood of 10 frames, assuming locally constant probe velocity.

Following Chen et al.'s study, several researchers proposed useful generalisations of their work. Tuthill et al. [219] generalised their approach for application outside the focal zone of the transducer by allowing the standard deviation of the Gaussian decorrelation model to vary as a function of depth. By computing local image correlations at different locations in the image plane, they also allowed for the estimation of transducer yaw (rotation about the y axis) and tilt (rotation about the x axis). Their paper describes the first experiment using speckle decorrelation to estimate out-of-plane transducer motion *in vivo*, using a rudimentary filter to find regions of fully developed speckle. Prager et al. [166] use the Pearson correlation coefficient between pairs of image patches (which they reformulate as a linear regression parameter) in order to estimate distance using a single correlation measurement (as in the example shown in Figure 2.5(b)). Unlike Chen et al.'s original strategy which fits the decorrelation model to the local autocorrelation function of the data over multiple frames, this strategy does not make any assumption about the local regularity of frame spacing and allows trajectory estimation in a broader range of conditions. They

combined this approach with a conventional image registration technique to obtain good qualitative 3D reconstructions of ultrasound imagery acquired using the freehand technique, demonstrating the potential of image-based tracking as an alternative to external position sensors. Chang et al. [29] proposed a unique, calibration-free, transducer independent approach where speckle decorrelation is used to infer the ratios of different displacements with respect to each other in a linear sweep of the transducer. In this approach, the global scale of the scan is estimated by registering the images to an additional ultrasound frame acquired perpendicularly to the linear sweep. This approach has the disadvantage that it significantly complicates the image acquisition process.

In spite of the promising results obtained in these early works, a paper by Li et al. [133] criticises the suggestion that quantitatively accurate transducer tracking may be achieved using image content only. The paper argues that speckle decorrelation arises from causes other than elevational translation, such as transducer rotation, and that it is difficult, if not mathematically impossible, to express the observed decorrelation as a combination of the decorrelations caused by the different types of motion. Through simulation experiments, the authors show that even very small rotations, which are very likely to occur in freehand scans, can cause substantial errors in the estimation of elevational motion. They also show that even when there is no transducer rotation, sample correlation coefficients can exhibit substantial variability, limiting the accuracy of the speckle decorrelation approach. Work by Smith et al. [204] and Hassenpflug et al. [85] further studied the accuracy of distance prediction made from correlation measurements and found that accuracy is not uniformly obtained over all distances; very short distances, corresponding to high correlation, in particular, were found to give rise to large errors. Chapter 4 takes the results of these studies one step further by explicitly modelling the stochasticity of the relationship between the sample correlation coefficient and elevational separation, leading to a new probabilistic approach for combining such measurements in the context of trajectory estimation.

Additional difficulties arise from the fact that the transducer-specific decorrelation model obtained during calibration is only accurate for imagery of fully developed speckle, i.e. when Rayleigh scattering conditions hold. Outside the context of Rayleigh scattering, the width of the decorrelation curve increases depending on the micro-structure of the medium, as described by $\mathfrak{R}'_\varsigma$ in (2.11). Though there have been attempts to model such variations in the decorrelation curves of the echo envelope and intensity signals as a function of changes in the micro-structure of the medium, the results have been asymptotic [228] or

empirical [158] rather than algebraic and also lack generality.

This lack of directly applicable models for the spatial decorrelation of the ultrasound echo envelope outside Rayleigh scattering conditions poses serious challenges when it comes to estimating the elevational separation between two images of real tissue based on their correlation. This issue has most commonly been tackled by filtering the imagery to detect regions of fully developed speckle where the model applies and ignoring the rest [219, 166, 177, 178]. However, Hassenpflug et al. [86] demonstrated that Rayleigh scattering conditions occur rarely in real tissue, limiting the use and accuracy of this approach. A more effective approach which heuristically estimates the shape of the local elevational decorrelation curve based on the axial and lateral decorrelation curves was proposed by Gee et al. [72]. These methods and other related work will be discussed in a more detailed fashion in Chapter 6, which proposes an alternative, data-driven approach to the problem.

A further difficulty with the speckle decorrelation approach is that a single correlation measurement can only provide an absolute estimate of distance. This creates ambiguities when the trajectory of the probe is not entirely monotonic, or when there are intersections between the frames. In the absence of external sensors, these ambiguities must be resolved using correlations between three or more pairs of frames, exploiting the physical constraint that images must be planar. Housden et al. [94] were the first to propose a method for doing this (their approach as well as alternate ones [121, 126], are discussed in detail in Chapter 3). Combining this with the adaptive speckle decorrelation method of [72] and in-plane image registration [93], their work produced the first (and currently, the only) system capable of entirely image-based transducer tracking in a truly unconstrained setting. Though this is a most remarkable achievement, the lack of constraints on the possible probe trajectories implies that the system is vulnerable to the sources of ambiguity predicted by Li et al. [133]. In particular, the authors note that the use of speckle decorrelation techniques for the estimation of elevational translation loses some of its accuracy in the presence of in-plane probe motion, even when the latter is corrected for by registration [93]; this is the result of speckle decorrelation arising from the in-plane motions themselves as well as inevitable interpolation errors in the registration process. Furthermore, the results presented in [94] clearly show that the accuracy of trajectory estimates is diminished by rotations with an axial component (tilt, in the case of a linear array transducer) because the rotations themselves introduce speckle decorrelation which is not well modelled or accounted for by a translational speckle decorrelation model.

Recent research has thus investigated methods for combining the information provided by position tracking devices and that provided by the imagery, exploiting not only image registration techniques as was done in earlier work [180, 118, 218, 192], but out-of-plane speckle decorrelation as well [95, 124], leading to both small and large scale accuracy. A further development, which builds on the strengths of such hybrid systems, is the introduction of a rotational speckle decorrelation model [92]. Combined with orientation information from a position tracking device, such a model allows the speckle decorrelation arising from translations and rotations to be disambiguated and exploited for more accurate trajectory recovery, as originally predicted by Li et al. [133]. Hybrid systems such as these are a promising context for the exploitation of the new statistical models and techniques proposed in this thesis, as discussed in Chapter 8.

2.5 Summary

This chapter summarised the fundamental physical principles of ultrasound B-mode image formation, with an emphasis on the phenomena studied in this thesis. In particular, the formation of speckle patterns was described mathematically and some important first and second order characteristics of speckle in the limit of Rayleigh scattering conditions were derived. This leads to the important idea that the correlation between neighbouring speckle images provides an indication of spatial separation. A brief overview of research exploiting this idea within the context of freehand 3D ultrasound was given, providing a context for the contributions made in the next chapters of this thesis. The literature on speckle decorrelation-based out-of-plane transducer trajectory estimation is relatively scarce and represents the work of only a small number of individuals. As a result, is much room for innovation in this domain. For instance, while the variability of correlation measurements associated with a given elevational separation under Rayleigh scattering condition has been recognised, there have been no attempts to model it explicitly. Also, while the inadequacy of the Rayleigh scattering assumption in real tissue has also been recognised, the field definitely lacks methods which address the problem in a manner that is effective, principled and tied, if only indirectly, to ultrasound physics. Methods for resolving the directional ambiguity of correlation measurements in the case of non-monotonic probe motion and image sequences with intersecting frames has only begun to be addressed by others through the use of multiple redundant measurements and the full range of possibilities for achieving

this has not yet been explored. These are all questions which will be addressed in this thesis, beginning with the resolution of directional ambiguities in the next chapter.

Chapter 3

Discovering structure from redundant correlation measurements

The previous chapter explained how, under Rayleigh scattering conditions, the correlation between speckle patterns in two ultrasound images is predictive of the magnitude of the elevational displacement undergone by the probe. This is because the ultrasound beam has a finite extent and images much of the same micro-structure when moved between neighbouring locations. An elevational speckle decorrelation model thus provides a means of estimating the absolute elevational separation between two parallel ultrasound image windows from the correlation between them. However, in general, a single correlation measurement between two arbitrarily positioned ultrasound images is of very little use by itself, as there are many different types of motion which could give rise to this measurement. In freehand ultrasound data sequences, there can be in-plane motion and out-of-plane rotations of the transducer, non-rigid deformations of the target, intersections between frames as well as changes in the direction of the scan. Even if the images are assumed to be parallel, there remains an ambiguity as to the direction of the translational motion because the distance estimate obtained from the speckle decorrelation is not signed. In order to recover the transducer trajectory under general conditions, it is necessary to exploit multiple correlation measurements and to establish useful constraints describing the relationship between the different measurements. This chapter presents a method for recovering rigid out-of-plane probe motion by exploiting such constraints, assuming that there is no in-plane motion or that it has otherwise been accounted for by an image registration process.

Correlation measurements between two images can be sampled at many pairs of corresponding locations in the respective image planes. Because the samples are known to lie on a plane for any given image, the distance estimates resulting from these redundant measurements provide a means of estimating out-of-plane rotations as well as translations between pairs of images. Additionally, one can exploit correlation measurements between any two images in the sequence, not just consecutive pairs. The geometric consistency of distance measurements thus obtained can be used to establish the local directionality of the scan in the presence of non-monotonic motion patterns. Combined, the two types of constraints also allow the detection of intersections between frames.

This chapter presents a framework for recovering the out-of-plane trajectory of an ultrasound transducer based on multiple redundant correlation measurements whose general form will be used throughout this thesis, with small variations which will be discussed in the relevant chapters. The chapter begins with a discussion of existing approaches for the image-based recovery of the structure of ultrasound scans (Section 3.1). Section 3.2 then describes the new framework in its globality, explaining the role of each of its components in recovering the structure of ultrasound scans. Sections 3.3 and 3.4 describe some of these components in detail, explaining how intersections between frames and non-monotonic transducer motion are recovered using methods from combinatorial optimisation and robust statistics. A short validation experiment with artificially generated distance measurement data is described in Section 3.5. More complete experiments involving image data will be presented in later chapters. The majority of the work presented in this chapter was published as part of [126].

3.1 Related work

One of the first attempts at exploiting multiple correlation measurements between pairs of ultrasound images is described by Prager et al. [166]. Their approach is to break each image frame into a number of non-overlapping patches. Distances between corresponding patches of a pair of images are then used to robustly estimate the rigid transformation between the images, exploiting the image planarity constraint. Similar ideas were proposed in patents filed by Siemens and General Electric [132, 67, 68, 32, 149, 74]. These approaches account for non-parallel images but assume that the probe motion is monotonic and that there are no intersections between frames.

Building on the ideas of [166], Housden et al. [94] introduced the first system for the image based recovery of unconstrained transducer trajectories. Their approach robustly averages a small number of independent, coarse reconstitutions of the probe trajectory based on interleaved subsets of image frames. The image planarity constraint is first enforced by robustly fitting a rigid transformation between all pairs of consecutive frames in the coarse data sequences, based on the elevational separations estimated from correlations between corresponding image patches. Outlying distance estimates are treated as an indication that the two frames intersect. Non-monotonic probe motion is then detected in each coarse sequence by treating consecutive entire frames planes in triplets and estimating the motion parameters which best explain the estimated patch-wise elevational separations between the three pairs of frames involved (in the least squares sense). Coarse trajectories which are in agreement with each other regarding the total length of the scan are then averaged to obtain a refined trajectory estimate.

The directional ambiguity of correlation-based estimates of elevational displacement was also addressed by Krupa et al. [119, 120] for a visual servoing application. In this context, a robot is used to perform image acquisition and stabilisation. The goal is to make the robot manipulate the ultrasound probe such that, as the scanned medium undergoes motion, it acquires images that are as close as possible to a certain target image (previously tagged as such). For this task, the relative position of the robot held probe to the target is estimated from the correlation between the current image and the target, combined with information from the robot's joint encoders. The application requires that sign ambiguities in the elevational distance estimates be resolved so that the image-based servoing guides the robot in the correct direction. To this end, the authors' approach first copes with the possibility of intersections between successive frames using a random sampling algorithm. Starting from randomly assigned displacement signs, this iterative algorithm tentatively reverses the sign of one randomly chosen patch and greedily accepts sign changes that reduce the least squared error criterion arising from the image planarity constraint until a convergence criterion is reached. This yields two solutions (with the target lying on one or the other side of the current frame). These are disambiguated by tracking the sign of the displacements, starting from a controlled negative displacement from the target image at the beginning of the procedure.

In more recent work, Krupa et al. [121] also developed a different algorithm for estimating the sign of elevational displacements which is robust to target motion exceeding

the range of ultrasound elevational speckle correlation. This algorithm also depends on an initial controlled elevational displacement of the ultrasound probe in the negative direction with respect to the target image. Over the course of this displacement, a number of reference ultrasound image frames are acquired and the pairwise displacements (of known sign) between these frames are measured using speckle decorrelation. After this control step, the scanned object is allowed to move and the visual servoing algorithm begins its search for the target image. The signs of the patch-wise displacements required to obtain the target image are obtained by estimating the absolute patch-wise distances of the currently observed image to each of the reference frames acquired during the controlled displacement step and finding the set of signs which minimises the squared error between the inter-reference distances stored during the control step and those computed using the current observation. The search is rendered tractable by the strong constraints imposed by the known order of the reference frames. Once the optimal sign sequence is found, a plane is fitted to the resulting distance offset and used to control the motion of the probe-bearing robotic manipulator.

Though the methods proposed by Krupa et al. [119, 120, 121] do tackle the issue of directional ambiguity in speckle decorrelation based distance measurements, their approach requires that strong prior constraints be established through controlled data acquisition steps. In contrast, the approach suggested by Housden et al. [94] and that proposed in this thesis do not require any modifications to the scanning protocol, though such modifications, if helpful, can always be incorporated. The next section gives an outline of the proposed approach and compares it with the approach of Housden et al [94], its closest relative.

3.2 Overview of the proposed framework

To account for the variation of elevational ultrasound speckle decorrelation length with axial depth, this work adopts the approach used by the group of medical ultrasound researchers at Cambridge University [166, 72, 94]. That is, the ultrasound image is divided into Q equally sized non-overlapping patches. Each patch has its own associated speckle decorrelation model. This subdivision implies that for each frame pair, there are as many correlation measurements as there are patches, allowing yaw and tilt angles to be estimated in addition to elevational translation. A possible such subdivision is illustrated in Figure 3.1.

During a calibration step, a speckle phantom is scanned at regular elevational intervals.

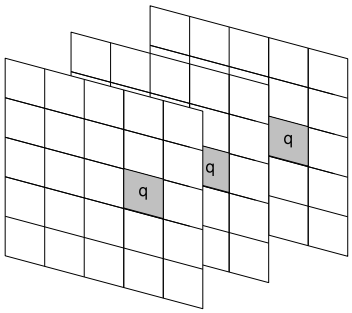


Figure 3.1 The images are divided into non-overlapping patches. Corresponding patches (e.g. grey patches labelled q) in different images are used to build local decorrelation models and are treated as mini-frames when estimating elevational positions.

The correlation coefficients between corresponding image patches in different frames are then measured. A speckle decorrelation model is then estimated for each of the Q image patches. For the moment, let us assume a deterministic, piecewise linear decorrelation curve relating the sampled distance intervals to the average of corresponding correlation coefficients. An alternative probabilistic speckle decorrelation model which describes the *distribution* of these sample coefficients will be introduced in Chapter 4. The speckle decorrelation model obtained through this calibration process can be used to estimate transducer displacements in a new ultrasound image sequence.

Subdividing the $n + 1$ image frames in the ultrasound data set into Q patches corresponding to those used to define the speckle decorrelation model breaks the data set into Q individual smaller ultrasound data sets consisting of $n + 1$ "mini-frames" the size of an individual image patch (see Figure 3.1). The sample correlation coefficients between all the pairs of mini-frames corresponding to the same image location are computed. Using these measurements, the relative positions of the full ultrasound image frames are then recovered in three steps:

1. For each image patch, a consistent ordering of the associated mini-frames along the elevational direction is derived from the redundant correlation measurements using the TSP-sort algorithm, a combinatorial optimisation approach based on the solution of the traveling salesman problem described in Section 3.3. This accounts for reversals in the direction of motion and intersections between frames.

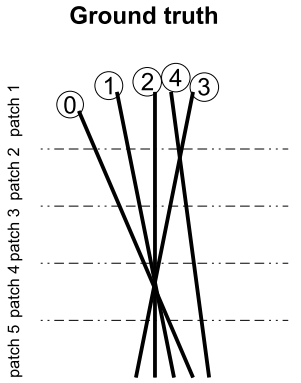
2. For each image patch, a trajectory is estimated based on the ordering found by TSP-sort and the elevational separation estimates provided by the speckle decorrelation model. Different methods for doing this will be investigated in Chapters 4 and 7.

3. A rigid transformation bringing the patch centres of the first frame to each of the other frames in the sequence is robustly estimated, allowing the recovery of elevational translation, yaw and tilt.

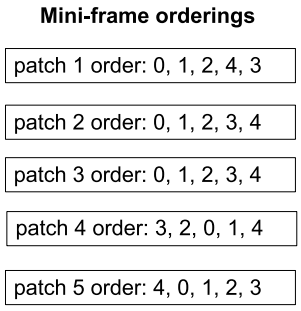
The overall approach is illustrated and explained in more detail by a step-by-step example in Figure 3.2. This general framework most closely resembles that proposed by Housden et al. [94]. At the coarse level, their approach begins with the enforcement of the image planarity constraint (violations of which allow for the detection of frame intersections) and detects non-monotonic motion based on the relative positioning of entire image frames. This thesis proposes an orthogonal approach: non-monotonic motion is first detected at the image patch level using combinatorial optimisation through the TSP-sort algorithm (see Figure 3.2(b)). The image planarity constraint is only enforced at the very end. The result is that frame intersections need not be detected explicitly (according to Housden et al. [94], the most error-prone step of their method): they are already reflected in the mini-frame orderings found by the TSP-sort algorithm. The final rigid transformation fitting stage identifies patches with incorrect mini-frame orderings as outliers and takes care of reversing some of these orderings for the case where the first and second frame of the sequence intersect, as illustrated in Figure 3.2(d). In the next sections, the computational mechanisms required for carrying out the steps illustrated in Figure 3.2 are described in detail, beginning with the combinatorial patch-wise mini-frame ordering algorithm.

3.3 Mini-frame ordering through the TSP-sort algorithm

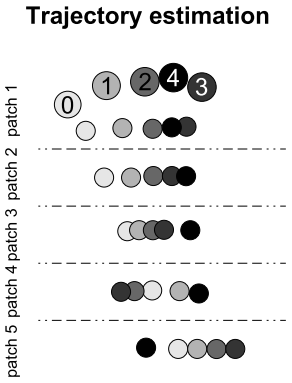
For each each pair of mini-frames i and j in a test ultrasound data set, a correlation measurement ρ_{ij} is used to read an absolute distance estimate δ_{ij} off the average decorrelation curve for the corresponding image patch, as illustrated in Figure 2.5(b) of the previous chapter. Assuming that the images are aligned in-plane, the centres of corresponding mini-frames in any section of the scan approximately lie along a line and the absolute elevational



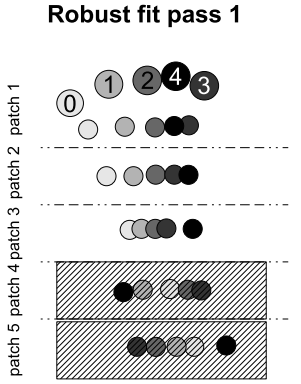
(a) The ground truth geometry of a 5 frame sequence. The circled numbers indicate the frame acquisition order. The frames have been divided into 5 patches. Correlations between all pairs of corresponding image patches are computed to obtain a distance estimate from the speckle decorrelation model.



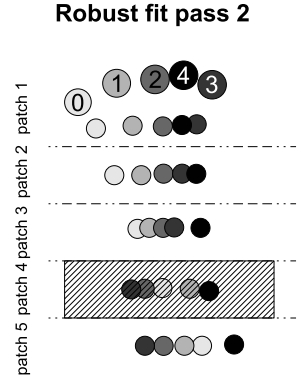
(b) The distance estimates are processed by the TSP-sort algorithm. By convention, the displacement between frames 0 and 1 is treated as positive, leading to a reversed ordering for patch 5. The shuffled ordering found for patch 4 is due to a poor fit of the speckle decorrelation model at the intersection of many frames.



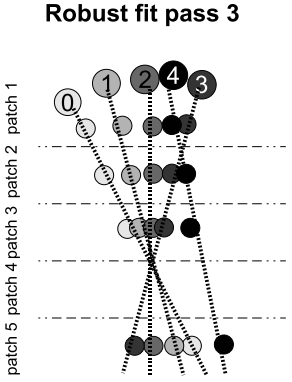
(c) The trajectory associated with each set of mini-frames is calculated using estimated pairwise elevational displacements between mini-frames, respecting the orderings found by TSP-sort. Each dot is the inferred location the patch centre, and its shade indicates which frame it belongs to.



(d) Patches 4 and 5 are labelled as outliers (as indicated by the hatched rectangles) because they do not fit the image planarity constraint. The TSP-sort orderings for these patches are tentatively reversed.



(e) After a second attempt, patch 4 is found to give worse results than before and is set back to its original ordering.



(f) Patch 4 is labelled as a permanent outlier, and all corresponding data are discarded. A least-squares rigid transformation maps each frame to the first, implying the image planes represented by the dotted lines.

Figure 3.2 A step-by-step example depicting how the proposed out-of-plane motion recovery approach works.

separations of the mini-frames should satisfy the constraint

$$\delta_{ik} = \begin{cases} \delta_{ij} + \delta_{jk} & s_{ij} = s_{jk} \\ \delta_{ij} - \delta_{jk} & s_{ij} \neq s_{jk} \wedge \delta_{ij} > \delta_{jk} \\ \delta_{jk} - \delta_{ij} & s_{ij} \neq s_{jk} \wedge \delta_{ij} < \delta_{jk} \end{cases} \tag{3.1}$$

for any three mini-frames i, j and $k \in \{0, \ldots n\}$, where s_{ij} denotes the sign (or direction) of the elevational displacement between mini-frames i and j. The constraint is illustrated in Figure 3.3. This suggests a simple algorithm for determining the signs of elevational displacements between consecutive mini-frames. Initialising $s_{01} = 1$, the signs of all other displacements could be determined as follows: for each set of three consecutive mini-frames i, $j = i + 1$ and $k = i + 2$, consider the measured absolute distances δ_{ij}, δ_{jk} and δ_{ik}. Then, for $i \in 0, \ldots n - 2$,

$$s_{jk} = \begin{cases} s_{ij} & \delta_{ik} \geq \delta_{ij} + \delta_{jk} \\ -s_{ij} & \text{otherwise.} \end{cases} \tag{3.2}$$

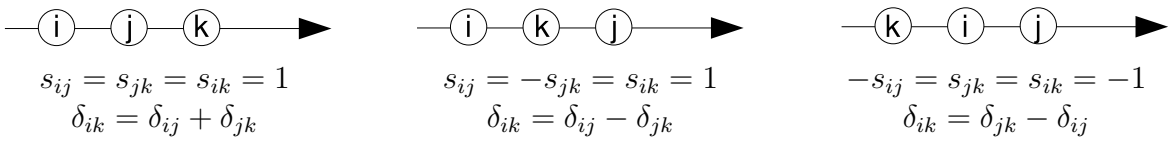


Figure 3.3 An illustration of the collinearity constraint described by the three cases of (3.1).

In reality, the estimated elevational offsets are noisy and most likely do not satisfy (3.1), leading to the failure of this overly simplistic algorithm. In order to achieve the robustness required in the presence of noise, more elevational separation measurements must be taken into account.

Consider again the noiseless case where all correlated mini-frames corresponding to the same image patch lie approximately along a straight line, satisfying (3.1) for all triplets of mini-frames, consecutive or not. This line trivially corresponds to the shortest possible path which includes every mini-frame centre only once. In the noisy case, one can think of the direction assignment problem as that of finding an ordering of the mini-frame centres which minimises the total travelling distance among them, where the distance between

each pair of mini-frames i and j is given by the estimate δ_{ij}, thus choosing elevational displacement signs to make the trajectory as close as possible to the ideal straight line.

Posed this way, the problem is very closely related to the symmetric travelling salesman problem (TSP), a classical problem in combinatorial optimisation. Given a set of cities and the distances between all pairs of cities, the TSP consists in finding the shortest tour which visits each city only once and returns to the originating city.[1]

While the TSP is NP-hard, it is very well studied and clever methods exist which reduce the computational burden of finding an optimal solution in most cases. Such methods are implemented in the free, state-of-the-art Concorde software [8] designed for the solution of the TSP with symmetric distances. Briefly, Concorde performs a branch-and-bound search across possible tours. Increasingly tight lower bounds on the length of the optimal tour are found by solving increasingly tight linear programming relaxations of the TSP. This technique, known as the *cutting plane* method, is rooted in the seminal work of Dantzig, Fulkerson and Jonhson [45], but the modern version of the algorithm uses stronger cutting planes than the original paper, coupled with a highly sophisticated software implementation. Concorde alternates its search for lower bounds with the generation of near-optimal candidate tours, which are iteratively improved using a variety of very effective algorithms (such as variants of the Lin-Kernighan heuristic [136, 87]). The length of these tours provide ever tighter upper bounds on the length of the optimal tour. The search ends when the upper and lower bounds on the length of the optimal tours coincide. The details of algorithmic components of the software, as well as an excellent discussion of the TSP, are available in [8].

In this work, Concorde was used to determine a mini-frame ordering for each patch. By convention, the elevational displacement between the first and second mini-frames is considered to be positive. Solving the TSP then provides a set of mutually consistent signs for the elevational displacements between pairs of mini-frames corresponding to one image patch. The approach, henceforth be referred to as the TSP-sort algorithm, was validated empirically (see Section 3.5 for details).

[1]Note that in the problem at hand, no such round trip is required. To match the TSP description exactly, a fictional point is added whose (non Euclidean) distance to all others is set to zero. A round trip can then be obtained without changing the shortest path among all other points.

3.4 Enforcing the image planarity constraint

The TSP-sort method described above and the subsequent patch-wise trajectory estimation step (illustrated in Figure 3.2(c) and discussed in Chapters 4 and 7) provide consistent elevational pose estimates for individual sets of mini-frames. These must then be combined to obtain an estimate of each frame's elevational translation, yaw and tilt. Context naturally provides an additional constraint to the problem of estimating frame positions: image patches belonging to the same image frame must lie on a plane. The final position of each full frame can be calculated from the rigid transformation parameters $\boldsymbol{\phi}^*$ which best map the centres of the image patches belonging to the first frame to the set of elevational pose estimates obtained for the centres of all image patches belonging to every other frame. In the preliminary work presented in [125], it was assumed that there were no frame intersections and that probe motion was monotonic. Thus, the point-wise registration was carried out according to a standard least-mean-squared error criterion; that is, for frame i,

$$\phi_i^* = \underset{\phi}{\operatorname{argmin}} \sum_{q=1}^{Q} ||\mathcal{R}(\tilde{\mathbf{x}}_0^{[q]}, \phi) - \tilde{\mathbf{x}}_i^{[q]}||^2, \tag{3.3}$$

where Q is the number of image patches in each frame, $\tilde{\mathbf{x}}_i^{[q]}$ denotes the 3D position of the centre of patch q in frame i as inferred from image-based measurements (speckle decorrelation for out-of-plane displacements) and $\mathcal{R}(\mathbf{x}, \phi)$ denotes the 3D position obtained when applying the rigid transformation with parameter vector ϕ to point $\mathbf{x}$. In this case, ϕ_i^* can be computed using the standard methods described in [221, 65]. This technique is problematic when non-monotonic probe motion and frame intersections are present, and is extended here in order to account for such circumstances.

The mini-frame orderings found by the TSP-sort algorithm follow the convention that the displacement between the first and second mini-frames corresponding to a given image patch is always positive. If the first and second image frames intersect, orderings must be reversed for a certain set of image patches on one side of the intersection. Intersections between frames may also cause the TSP-sort algorithm to fail in another way: for image patches where intersections occur, the estimated elevational separation should be close to zero. This will generally not be reflected in the corresponding correlation measurements due to significant speckle decorrelation arising from the rotations which caused the intersections.

This decorrelation will be incorrectly interpreted as resulting from elevational translation. In areas where multiple frames intersect, this may lead to meaningless mini-frame orderings. Image patches with shuffled or reversed TSP-sort orderings create outliers with respect to the image planarity constraint.

Outliers are accounted for by replacing the least-mean-squares fit with a robust fitting approach. This thesis adopts the robust least-median-of-squares (LMedS) fitting paradigm [183, 184], though any of several other robust methodologies could have been used instead [63, 243, 230, 186, 248, 59]. In a LMedS rigid transformation fit for frame i, the optimal rigid transformation parameters ϕ_i^{LMedS} are given by

$$\phi_i^{\text{LMedS}} = \underset{\phi}{\text{argmin}} \underset{q}{\text{med}} ||\mathcal{R}(\tilde{\mathbf{x}}_0^{[q]}, \phi) - \tilde{\mathbf{x}}_i^{[q]}||^2, \tag{3.4}$$

where med is the median operator. Under certain conditions (including the normality of the noise and uniform spatial distribution of outliers), the LMedS definition of optimality can handle up to 50% of the data points being outliers [183, 184]. An exhaustive search for the best fitting rigid transformation parameters ϕ_i^{LMedS} in the LMedS sense is computationally prohibitive. Exhaustive search is avoided by a Monte-Carlo sampling strategy similar to that used in the well known Random Sample Consensus (RANSAC) algorithm [63], as described in [184, 145]. The sampling procedure consists in sampling minimal subsets of data points (samples of 3 pairs of matching points in the case of a rigid transformation fit) and fitting the required primitive (in this case, a rigid transformation) to those. The quality of the fit is then evaluated based on the entire point set (the Q image patch centres belonging to the frame under consideration). In the case of RANSAC, the quality of the fit is usually quantified in terms of the number of data points which satisfy the hypothesis according to some outlier detection threshold. In the LMedS case, the quality of a fit is instead quantified using the median of the squared residual error, as described by (3.4). The fit implied by the new sample is temporarily adopted as the correct one if its quality is better than that of a previously adopted fit. This sampling procedure is repeated enough times to guarantee a high probability of selecting an outlier free minimal subset of points. The required number of such trials is determined based on simple probabilistic arguments [183]. The LMedS approach was preferred over RANSAC because it does not require any parameter tuning.

A robust standard deviation is defined as [184]

$$\chi_i = 1.4826 \left(1 + \frac{5}{Q-3}\right) \underset{q}{\text{med}} \, ||\mathcal{R}(\tilde{\mathbf{x}}_0^{[q]}, \phi_i^{\text{LMedS}}) - \tilde{\mathbf{x}}_i^{[q]}||, \tag{3.5}$$

where the factor $\left(1 + \frac{5}{Q-3}\right)$ is used to correct for small Q (number of image patches). This robust standard deviation is used to determine a set of outlier points which do not agree well with the estimated rigid transformation parameters. Patch q of frame i is deemed an outlier if

$$||\mathcal{R}(\tilde{\mathbf{x}}_0^{[q]}, \phi_i^{\text{LMedS}}) - \tilde{\mathbf{x}}_i^{[q]}|| > 2.5\chi_i. \tag{3.6}$$

Three passes of the robust fitting algorithm are carried out. Every pass consists in finding the rigid transformation from the first frame to all other image frames based on their patch-wise elevational pose estimates. After the first pass, a number of image patches (and their set of corresponding mini-frames) are identified as candidates for having their order (and elevational pose signs) reversed. Image patches which are considered to be outliers by (3.6) for 5% or more of the frames are sign reversal candidates. During the second pass, the LMedS fitting algorithm is run again for every frame, this time with tentatively reversed orderings for the previously flagged image patches. The effect of this is evaluated after the second robust fitting pass: image patches which are found to be outlying more often than before the sign reversal are returned to their original positions. A final pass identifies permanent outliers as image patches which are considered to be outliers by the LMedS fitting procedure for more than than 20% of the frames. Final least-mean-squares rigid transformation fits (described by (3.3)) are then carried out on the inlier patches using the standard methods given in, e.g. [221, 65], for all frames in the sequence. This yields a set of rigid transformation fits $\boldsymbol{\phi}^*$ relating each frame to the first which accounts for a possible intersection between the first two frames as well as possible large-scale failures of the TSP-sort algorithm.

3.5 Experiments with the TSP-sort algorithm

In order to assess the validity and robustness to noise of the TSP-sort algorithm, a short simulation study was conducted. In this study, 30 sequences (trajectories) of 500 1D points from the real line were artificially generated, with consecutive points drawn randomly within

2.0 units of each other in the direction of travel, with a 5% probability of direction reversal at every point. The true distances between all pairs of 1D points were computed by simple subtraction and corrupted with varying levels of Gaussian noise. The standard deviation of the noise varied from 0 to 50% of the true distances measured. Distances larger than 5.0 units were assigned a large constant value (to reflect the fact that elevational distances between ultrasound images cannot be estimated below a certain level of correlation). The TSP-sort algorithm was applied to the noisy distance values in order to recover the signs of the measured distances. The vast majority of runs took under 10 seconds on a regular PC using the Concorde software.[2] The TSP-sort algorithm was compared with the simplistic algorithm introduced at the beginning of Section 3.3, which only exploits constraints between triplets of consecutive points and whose operation is described by (3.2).

The accuracy of the two algorithms was evaluated by comparing the recovered order of the points on the real axis with the ground truth using the Kendall rank correlation defined as [38]

$$\tau = \frac{2(N_c - N_d)}{n(n+1)}, \tag{3.7}$$

where $n+1$ is the number of points and N_c and N_d denote the number of concordant and discordant pairs of points in the ground truth and automatically ordered points. A pair of points i and j is called concordant if the sign of the difference between the rank of point j and the rank of point i in the inferred ordering is the same as that found for the ground truth ordering. τ takes the value 1 when the and inferred point orderings are identical, 0 when they are uncorrelated and -1 when they are the exact opposite of each other. There remains an overall, unavoidable sign ambiguitiy in the assignment of signs to unsigned distances, which was resolved by adopting the convention that the displacement between the first two points is positive. Therefore, Table 3.1 reports $|\tau|$ as a performance measure since both the true ordering and its opposite are considered correct.

In the noiseless case, both algorithms exhibit ideal behaviour. However, whereas the smallest amount of noise causes the base-line ordering algorithm to fail, the TSP-sort

[2]It should be noted that the Concorde TSP solving functionalities work exclusively in integer arithmetic whereas the simulated data as well as data to be gathered from ultrasound imagery later in this thesis are more naturally expressed in floating point representation. In order to use Concorde, a scale factor must be applied to the distance estimates in order to get a fixed point representation of the distances. An inadequately chosen scale factor causes a loss of precision which in turn can cause inaccurate comparisons and (very occasional) ordering errors even when there is no noise. For the experiments presented in this chapter, the scale factor was tuned to ensure that there were no errors in the noiseless case.

Noise level (%)	$\lvert\tau\rvert$	
	Base-line	TSP-sort
0%	1.000 ± 0.000	1.000 ± 0.000
1%	0.243 ± 0.190	$1.000 - 3 \times 10^{-5} \pm 2 \times 10^{-5}$
2%	0.270 ± 0.217	$1.000 - 6 \times 10^{-5} \pm 2 \times 10^{-5}$
3.5%	0.267 ± 0.193	$1.000 - 10^{-4} \pm 4 \times 10^{-5}$
5%	0.285 ± 0.188	$1.000 - 2 \times 10^{-4} \pm 8 \times 10^{-5}$
10%	0.261 ± 0.174	$1.000 - 4 \times 10^{-4} \pm 8 \times 10^{-5}$
20%	0.336 ± 0.186	0.999 ± 10^{-4}
35%	0.279 ± 0.170	0.986 ± 0.009
50%	0.329 ± 0.193	0.908 ± 0.163

Table 3.1 Kendall rank correlation ($|\tau| \pm$ 1 standard deviation) between the true point ordering and those inferred by the base-line sign assignment algorithm and the TSP-sort algorithm, as a function of measurement noise level.

algorithm remains extremely robust even at high noise levels. Whereas a single sign error early in the data sequence can completely derail the simplistic base-line algorithm, the TSP-sort algorithm is generally able to maintain a globally correct ordering of the points, with noise causing only localised incorrectly ordered segments in short parts of the sequence.

The average performance of the TSP-sort algorithm at the 50% noise level may seem surprisingly good. This is partly because large distances were treated as unmeasurable and were thus not corrupted by noise, allowing the TSP-sort algorithm to operate on the partial information provided by the smaller distances which were inherently less corrupted by noise in the absolute sense.

While the base-line algorithm based on triplets of consecutive points clearly failed in this experiment, it is important to realise that this base-line approach is very naive. Constraints between distance measurements within triplets of coarsely spaced, consecutive image frames (not mini-frames) can, and have been used very effectively and robustly for frame ordering, as exemplified by the work of Housden et al. [94]. The important feature of robust frame ordering algorithms is the regularisation afforded by the simultaneous use of multiple, redundant constraints. To achieve regularisation, the TSP-sort approach proposed in this chapter exploits all available triplets of corresponding patches at once. In Housden et al.'s work [94], regularisation arises from the fact that triplets of frames are treated as a collection of distance measurements which are not only constrained by elevational collinearity, but

also by image planarity. Whether one approach is superior to the other in terms of accuracy remains to be experimentally assessed. Housden et al. implement additional regularisation by independently computing frame orderings for distinct interleaved subsets of the frames and deriving a consensus from these. This robust averaging approach is independent of the frame ordering strategy and could easily be implemented in the context of the methods presented in this thesis for their further improvement.

3.6 Summary and discussion

This chapter described a new framework for discovering the structure of ultrasound scans, focusing on the recovery of directional changes in the trajectory and frame intersections. In particular, a new algorithm for recovering the spatial ordering of aligned image patches was proposed based on the solution to the well known travelling salesman problem. The algorithm was validated through a short simulation study whereby it was shown to be very robust to the corruption of distance estimates by random noise. This robustness stems from the ability of the algorithm to simultaneously exploit elevational collinearity constraints within a large number of redundant triplets of points. In addition, a least-median-of-squares approach was introduced for obtaining the rigid transformation parameters relating every frame of an ultrasound image sequence to the first using the elevational distance offsets computed from speckle decorrelation. This approach allows the identification of outlying patch-wise pose estimates symptomatic of inverted or shuffled patch-wise frame orderings. The framework as a whole has yet to be validated in experiments with ultrasound image data. Such experiments will be presented in Chapters 4 and 7, which describe new methods for the estimation of patch-wise trajectories based the fusion of multiple noisy correlation measurements. These new methods integrate naturally within the framework proposed in this chapter, so the experiments presented later will provide implicit validation for it.

Chapter 4

Probabilistic fusion of noisy correlation measurements

The previous chapter showed how correlation measurements between redundant triplets of image patches, or mini-frames, combined with the image planarity constraint, could be used to recover the general structure of an ultrasound scan in the presence of non-monotonic probe motion and frame intersections. This chapter addresses the problem of obtaining a refined trajectory estimate for each set of corresponding mini-frames in the sequence using the ordering provided by the TSP-sort algorithm of the previous chapter and correlation measurements between pairs of mini-frames, as illustrated in Figure 3.2(c) of the schematic example presented in the previous chapter. There is much choice when it comes to methods for doing this as every single frame can be correlated to multiple other frames in the image sequence; each correlation measurement in turn provides a potentially informative distance estimate, but each of these estimates is also corrupted by error.

This chapter describes a new approach for modelling the relationship between the sample correlation coefficient measured between two image patches and the elevational separation between them. Whereas all previously published work considers this relationship to be deterministic, this thesis proposes instead that it be modelled as a probability density function. This representation is useful because it allows the uncertainty attached with any correlation-based distance estimate to be quantified. This feature is then exploited in a maximum likelihood data fusion algorithm inspired from the mobile robotics literature for estimating the position of each mini-frame relative to the others for a given image location.

Section 4.1 opens the chapter with a discussion of previous research findings which are relevant to the trajectory estimation problem under consideration as well as other attempts to solve it. Section 4.2 then presents the new probabilistic speckle decorrelation model and describes how to estimate its transducer-specific parameters using ultrasound images acquired from a speckle phantom. Section 4.3 explains how this new model can be used to combine multiple redundant correlation measurements in a principled way in order to recover the relative positions of mini-frames corresponding to a single image patch. Experiments on synthetic and real ultrasound imagery of speckle phantoms are presented in Section 4.4. These experiments use the global framework discussed in the previous chapter for the recovery of scan structure, thereby demonstrating its strengths and weaknesses. In these experiments, the proposed probabilistic data fusion strategy is compared to two deterministic methods and the effect of certain trajectory parameters (monotonicity, presence of rotations and speed) on the results is evaluated. The theoretical framework and some of the experimental results presented in this chapter were published in [125, 126].

4.1 Related work

Classical correlation-based estimation of elevational separation between ultrasound images proceeds by first modelling the relationship between correlation and distance as a monotonically decreasing curve and by subsequently measuring correlation between images and reading the corresponding distance off the curve. The accuracy of this process depends on the statistical variability of sample correlation measurements as well as on the true and hypothesised functional forms of the relationship between correlation and distance. Li et al. [133] have shown that the sample correlation coefficient between pairs of ultrasound images of fully developed speckle can exhibit substantial variability. This variability arises from two types of finite sample size effects. The first of these stems from the finite number of pixels from which the correlation coefficients are computed. The second stems from the finite size of the physical portion of the medium represented in the image patches, which may change the local second order statistics of the medium even under Rayleigh scattering conditions.

The shape of the decorrelation curve determines how this variability in the sample correlation coefficient can affect the distance estimate. For a transducer with a Gaussian resolution cell, theory predicts that the normalised correlation of the echo intensity (squared

envelope) signal is a Gaussian function of distance [166]. Using the echo amplitude signal yields a very similar shape, i.e. a curve that is nearly flat but concave and highly non-linear for small distances, nearly linear for moderate distances, and non-linear, convex, and asymptotically flat for large distances. The accuracy of the distance estimate read off a reasonable approximation of such a curve largely depends on the value of the correlation coefficient used as input. This relationship was first studied by Smith and Fenster [204], and subsequently by Hassenpflug et al. [85]. Both studies concluded that distance estimates were most accurately obtained for moderate correlation values lying in the nearly linear portion of the decorrelation curve, and that large relative error was obtained in the high and low correlation portions of the curve.

Several factors explain the findings of [204, 85]. Assuming a perfectly accurate decorrelation curve (obtained from equally accurate and exhaustive transducer specifications, for example), the high and low correlation portions of the curve are error prone because they are nearly flat, implying that a slight change in the measured correlation coefficient can lead to a large change in the distance estimate read off the curve. In practice, decorrelation curves are obtained empirically [31, 219, 166, 72, 94] by fitting some monotonically decreasing function to the points defined by elevational separation intervals and the average of the corresponding correlation coefficients measured in imagery of a speckle phantom. In the high correlation region of the curve, the statistical distribution of the sample correlation coefficient is highly skewed (see Figure 4.1 in Section 4.2) and sample measurements are more likely to be larger than the average than lower, thereby leading to skewed distance estimates as well. If a piecewise linear decorrelation model is chosen [72, 94], then it is also likely that the first few linear segments of the curve very poorly represent the true underlying curve, leading to biased distance estimates.

The naive strategy of using only correlation measurements between consecutive frames, which will form the basis for the "nearest-neighbour" approach used in the experiments at the end of this chapter, is therefore highly error prone. Not only are the relative errors in each distance estimate large, but the bias and uncertainty associated with these estimates also accumulate over a large number of frames. In order to avoid such accumulation of errors, Gee et al. [72] have devised a strategy whereby a given frame is chosen as a reference and a number of subsequent frames are positioned with respect to it rather than their immediate predecessor. The reference frame changes over the course of the process in order to ensure a sufficient level of correlation between the pairs of frames considered.

This approach inspired the "shifting reference" approach used a second base-line in the experiments of this chapter.

Housden et al. [94] have developed a more sophisticated approach to the data fusion problem, which involves computing a robust average of coarsely sampled trajectories computed from independent interleaved subsets of the image frames. The composition of the interleaved subsets of frames is optimised in such a way as to make triplets of consecutive frames in a given subset as widely spaced as possible but correlated enough to recover non-monotonic motion. In other words, the subsets of frames are selected so that all or most correlation measurements actually used for trajectory estimation lie on the linear, well-behaved part of the decorrelation curve.

The above discussion demonstrates that there has been interest in understanding the errors that arise from the use of a speckle decorrelation model to estimate out-of-plane displacements and exploiting what knowledge there is about these errors to avoid tracking inaccuracies. However, there have been very few attempts to explicitly model these errors or derive any principled data fusion approaches based on such formal models. The hybrid tracking system of Lang et al. [124] is perhaps an exception, in that it explicitly accounts for the relative accuracies of image-based and electromagnetic tracker based measurements within the well-known Kalman filtering framework. However, the noise models used in the paper for each data source were tuned by trial and error and are not firmly rooted in physical models nor in the systematic analysis of observations. The work presented here innovates by deriving a probabilistic model of the relationship between correlation measurements and elevational displacements using data acquired during calibration. Furthermore, a data fusion approach is described which can exploit an arbitrary number of correlation measurements for each image patch while quantifying the statistical relationship of these measurements to the distances being inferred.

4.2 Probabilistic speckle decorrelation model

The first necessary step in using speckle decorrelation to estimate out-of-plane transducer trajectory is to characterise the imaging instrument in terms of the relationship that can be expected between correlation measurements and elevational displacement. This is done through a calibration process whereby a phantom with Rayleigh scattering properties is scanned at regular elevational intervals. Recall from the previous chapter that in order

to account for axial variations in transducer beam-width and to be able to recover out-of-plane rotations, correlation measurements will be gathered from pairs of corresponding image patches (or *mini-frames*) lying on a pre-defined grid. During calibration, a speckle decorrelation model is established for each patch in the grid.

For each image patch, the calibration scan provides sample mini-frames separated by known, regular elevational distance intervals. The sample correlation coefficient between corresponding image patches u and v of size N_{pix} is given by

$$\rho(u,v) = \frac{N_{\text{pix}} \sum_{i=1}^{N_{\text{pix}}} u_i v_i - \sum_{i=1}^{N_{\text{pix}}} u_i \sum_{i=1}^{N_{\text{pix}}} v_i}{\sqrt{\left[N_{\text{pix}} \sum_{i=1}^{N_{\text{pix}}} u_i^2 - \left(\sum_{i=1}^{N_{\text{pix}}} u_i\right)^2\right]\left[N_{\text{pix}} \sum_{i=1}^{N_{\text{pix}}} v_i^2 - \left(\sum_{i=1}^{N_{\text{pix}}} v_i\right)^2\right]}}, \tag{4.1}$$

where u_i and v_i denote the value of pixel i in image patches u and v, respectively. An assumption underlying the speckle decorrelation methodology is that the image patches u and v can be treated as finite size realisations of correlated random signals U and V with nominal correlation coefficient ρ_0. For large N_{pix}, the sample correlation coefficient $\rho(u,v)$ given in (4.1) should, on average, tend towards the nominal value ρ_0 which is indicative of elevational separation between the image patches. However, the individual sample correlation coefficients do exhibit some variability due to the finite size N_{pix} of the image patch realisations. This variability depends on ρ_0 and on the statistical distributions of U and V.

Previous approaches build a decorrelation curve by averaging the sample correlation coefficients obtained from the calibration data, in effect estimating ρ_0 for given patch location and elevational separation. This chapter argues that it is possible to extract more information than this from the wealth of available calibration data. The statistical variations intrinsic to correlation measurements can be explicitly represented by a probability density function $p(\rho|\delta,q)$ relating the sample correlation coefficient ρ to elevational separation δ in image location q. Assuming that the relationship between elevational separation and the nominal correlation coefficient ρ_0 is one-to-one (as would theoretically be the case for fully developed speckle), $p(\rho|\delta,q) \equiv p(\rho|\rho_0,q)$. The probability density function $p(\rho|\delta,q)$ captures several statistics of ρ other than the mean (such as variance and skewness) which can be exploited for the estimation of elevational separations based on multiple measurements.

Theory does not provide many indications as to a suitable parametric form for $p(\rho|\delta,q)$,

except in certain contrived cases [64, 148]. One approach which could be used to estimate $p(\rho|\delta, q)$ is to model the density using Parzen windows [50]. One well known difficulty with this approach is that it requires that a window size parameter be tuned. Automatic window size selection algorithms based on cross-validation have been suggested (e.g. [51]) and were experimented with for this work, but it was found that in practice, the approach leads to over-fitting for this particular density estimation problem. Here, $p(\rho|\delta, q)$ is instead estimated using a general purpose maximum entropy density estimation method introduced by Baker [10]. The method has interesting theoretical foundations and is applicable to a broad variety of problems [10, 157, 60, 11], but has been largely ignored by the imaging community. The method is described in detail below, as it is applied to the problem of estimating $p(\rho|\delta, q)$.

4.2.1 Density estimation by Baker's method

Baker's method [10] is a general method for estimating a continuous parametric probability density function from a set of data samples lying on a finite domain. In this work, the method is applied to the set of N_ρ sample correlation coefficients $\{\rho_i\}, i = 1, \ldots, N_\rho$ acquired between frames separated by a distance δ in patch q. These data are defined on the closed interval $[-1, 1]$, and therefore, Baker's method applies.

At the core of Baker's method is the idea that the estimation of a statistical distribution from data should depend on as few assumptions as possible; in other terms, the process should follow the principle of maximum entropy [106]. This principle states that in order to make as few additional assumptions about the statistics of the data as possible, the probability distribution must be chosen so as to maximise entropy given known statistics of the distribution. A common assumption is that the moments of the sought distribution are known exactly; in this case, the maximum entropy probability density function is known to take the form [106]

$$p(\rho|\delta, q; \boldsymbol{\mu}) = p_0(\rho) \exp\left(\zeta_0(\boldsymbol{\mu}) + \sum_{\omega=1}^{\Omega} \zeta_\omega(\boldsymbol{\mu})\rho^\omega\right), \tag{4.2}$$

where p_0 is a uniform probability density function over the range $\mathfrak{D}$ of the data (here, $\mathfrak{D} = [-1, 1]$), Ω is the known number of moments of the distribution and $\boldsymbol{\mu}$ is the known

vector of these moments. The normalisation constant $\zeta_0(\boldsymbol{\mu})$ is given by

$$\zeta_0(\boldsymbol{\mu}) = -\ln\left\{\int_{\mathfrak{D}} p_0(\rho)\exp\left(\sum_{j=1}^{\Omega}\zeta_j(\boldsymbol{\mu})\rho^j\right)d\rho\right\}, \tag{4.3}$$

and the parameters $\boldsymbol{\zeta}(\boldsymbol{\mu})$ verify the following system of non-linear equations:

$$\frac{\int_{\mathfrak{D}}\rho^{\omega}p_0(\rho)\exp\left(\sum_{j=1}^{\Omega}\zeta_j(\boldsymbol{\mu})\rho^j\right)d\rho}{\int_{\mathfrak{D}}p_0(\rho)\exp\left(\sum_{j=1}^{\Omega}\zeta_j(\boldsymbol{\mu})\rho^j\right)d\rho} = \mu_{\omega}, \omega = 1, ..., \Omega. \tag{4.4}$$

This form satisfies the maximum entropy constraint because it minimises the Kullback-Leibler divergence to the uniform density p_0,

$$D(p(\rho|\delta, q; \boldsymbol{\mu})||p_0(\rho)) = \zeta_0(\boldsymbol{\mu}) + \sum_{\omega=1}^{\Omega}\zeta_{\omega}(\boldsymbol{\mu})\mu_{\omega}. \tag{4.5}$$

In reality, the moments of the sought distribution are unknown and must be inferred from data. Baker's intuitive choice is to replace the moment vector $\boldsymbol{\mu}$ with the vector $\hat{\boldsymbol{\mu}}$ of *sample moments* of the data in (4.2):

$$\hat{\boldsymbol{\mu}}_{\omega} = \frac{1}{N_{\rho}}\sum_{j=1}^{N_{\rho}}\rho_j^{\omega}. \tag{4.6}$$

This yields a family of probability density functions

$$p_{\Omega}(\rho|\delta, q; \hat{\boldsymbol{\mu}}) = p_0(\rho)\exp\left(\zeta_0(\hat{\boldsymbol{\mu}}) + \sum_{\omega=1}^{\Omega}\zeta_{\omega}(\hat{\boldsymbol{\mu}})\rho^{\omega}\right) \tag{4.7}$$

whose maximum likelihood parameters $\boldsymbol{\zeta}(\hat{\boldsymbol{\mu}})$ can be obtained either by solving (4.4) or by directly minimising (4.5) (with $\hat{\boldsymbol{\mu}}$ substituted for $\boldsymbol{\mu}$), the implementation chosen in this work.[1] The required multi-dimensional non-linear optimisation was carried out using the Fletcher-Reeves conjugate gradient algorithm [66] and the integral required for the computation of (4.3) was computed by Gaussian quadrature [162].[2]

[1] An alternative approach for finding $\boldsymbol{\zeta}(\hat{\boldsymbol{\mu}})$ based on the method of moments was proposed by Urzúa [222].

[2] All numerical computations were carried out using the GNU Scientific Library [70].

This leaves the matter of choosing the unknown model order parameter Ω. Large Ω permits more flexibility and better fit to the data, but increases model complexity. It is well known that the quality of sample moment estimates depends on the amount of data available and high order sample moment estimates (which will be needed with large Ω) are notoriously unreliable. Therefore, high model complexity will likely lead to poor model generalisation. Baker's suggestion for selecting Ω is to minimise the Akaike Information Criterion (AIC) [5],

$$AIC = 2\Omega - 2\sum_{j=1}^{N_\rho} \ln p_\Omega(\rho_j|\delta, q; \hat{\boldsymbol{\mu}}). \tag{4.8}$$

Using (4.7) and dividing (4.8) by $2N_\rho$, notice that this is equivalent to minimising

$$\frac{AIC}{2N_\rho} = \frac{\Omega}{N_\rho} - \left(\zeta_0(\hat{\boldsymbol{\mu}}) + \sum_{\omega=1}^{\Omega} \zeta_\omega(\hat{\boldsymbol{\mu}})\hat{\mu}_\omega\right), \tag{4.9}$$

where the second term is already available for different Ω from the minimisation of (4.5) required for finding $\boldsymbol{\zeta}(\hat{\boldsymbol{\mu}})$. Interestingly, minimising (4.9) (and therefore, minimising the AIC) also has a maximum entropy interpretation because it turns out that (4.9) is an unbiased estimate of the Kullback-Leibler divergence between the true (unknown) distribution of the data and that predicted by the model [10]. The end result is that a good compromise between goodness of fit and model complexity is achieved without tuning any parameters.

Using Baker's method, conditional probability density functions $p(\rho|\delta_j, q)$ are obtained from sample correlation coefficients measured at patch q at elevational separation $\delta_j = j\delta_0,\ j \in \mathbb{N}^*$. Here, δ_0 is the elevational separation between consecutive frames of the calibration data set. Figure 4.1 shows an example set of these distributions obtained from ultrasound phantom data.

4.2.2 Computing the log-likelihood for arbitrary δ

Given a correlation measurement ρ, it is desirable to obtain an estimate of the likelihood $p(\rho|\delta, q)$, or, for practical purposes, of the log-likelihood $L(\rho|\delta, q) = \ln p(\rho|\delta, q)$. The calibration data, together with Baker's density estimation method, provide samples of the (log-)likelihood function at discrete values of δ corresponding to the regular elevational spacings where images were sampled during calibration. Some form of interpolation is thus

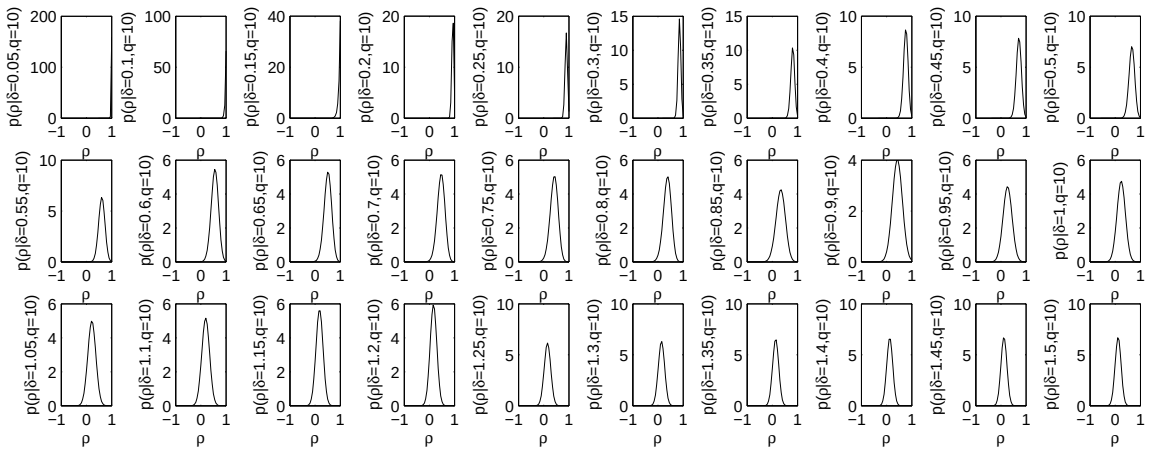


Figure 4.1 The probabilistic speckle decorrelation model for patch location q, obtained from ultrasound speckle phantom data. Each plot is a probability density function $p(\rho|\delta_j, q)$, estimated using Baker's method, of the sample correlation coefficient ρ corresponding to elevational separation δ_j, with δ_j starting at 0.05 mm and increasing at a rate of 0.05 mm per plot from left to right and top to bottom.

needed to evaluate $p(\rho|\delta, q)$ or $L(\rho|\delta, q)$ at arbitrary δ.

First, the number of samples of δ at which $p(\rho|\delta, q)$ can be evaluated is increased $l-$fold (for all experiments in this thesis, $l = 10$). For some elevational separation $\delta_j^{[k]} = (1 - \frac{k}{l})\delta_j + \frac{k}{l}\delta_{j+1}, k = 1, \ldots, l$, corresponding to no calibration data, a set of synthetic sample correlation coefficients are interpolated linearly between sample correlation coefficient data obtained during calibration at δ_j and δ_{j+1}. Baker's method is then applied to the artificial samples, yielding a finely sampled likelihood function $p(\rho|\delta_j^{[k]}, q)$. Increasing the resolution of the discretised likelihood function in this manner allows $L(\rho|\delta, q)$ to be subsequently interpolated at arbitrary δ using cubic splines without much undue oscillation.

4.3 Maximum likelihood mini-frame pose estimation

Once a speckle decorrelation model has been obtained through calibration, it can be used to estimate the elevational displacements between pairs of correlated image patches in a new ultrasound scan. In the trajectory recovery framework proposed in the previous chapter, an elevational trajectory must be estimated for the set of mini-frames corresponding to each image patch within the pre-defined calibration grid. Thus, the noisy, redundant cor-

relation measurements between mini-frames and the signs corresponding to the directions of elevational displacements between them (as found by the TSP-sort algorithm) must now be merged as accurately as possible into a set of consistent elevational pose estimates for each of the mini-frame centres. This problem is very similar to those of automatic robot localisation [202, 139, 173] or sensor network calibration [100, 141], as it may be thought of inferring the maximum likelihood positions of points related by a network of uncertain pose relationships (in this case, provided by correlation measurements). The solution proposed here draws inspiration from these domains. Mathematically, the objective is to estimate the ML positions $z_i^{[q]}, i = 1, \ldots, n$ along the elevational direction of n mini-frames corresponding to patch q with respect to the first mini-frame with position $z_0^{[q]}$. For the remainder of this section, it will be assumed that all computations take place for a particular patch q; thus, for convenience, the superscript $[q]$ will temporarily be dropped from all mathematical notation. Without loss of generality, it is assumed that $z_0 = 0$. A correlation measurement ρ_{ij} between mini-frames i and j provides an estimate of their elevational separation $\delta_{ij} = |z_i - z_j|$ via the log-likelihood function $L(\rho_{ij}|\delta_{ij}, q)$. For every mini-frame, there will be many such correlation measurements, and therefore, many available uncertain distance relationships involving its own pose.

Let $\mathbf{z}$ denote the position vector formed by concatenating the $z_i, i = 1, \ldots, n$. The output of the TSP-sort algorithm provides the signs s_{ij} of each elevational separation between mini-frames i and j, so that $z_j - z_i = s_{ij}\delta_{ij}, j > i$. Assuming that the correlation measurements are mutually independent of each other given the relative configuration of the frames, the maximum likelihood position vector $\mathbf{z}$ is given by

$$\mathbf{z}^* = \underset{\mathbf{z}}{\operatorname{argmax}} \prod_{j>i} p(\rho_{ij}|\ |z_i - z_j|, q) = \underset{\mathbf{z}}{\operatorname{argmax}} \sum_{j>i} L(\rho_{ij}|\ |z_i - z_j|, q). \tag{4.10}$$

Finding $\mathbf{z}^*$ is a difficult multi-dimensional non-linear optimisation problem for arbitrary L, as the objective function is bound to have many local minima. For the problem studied here, it was observed empirically that L often has a strong maximum in the vicinity of (though not necessarily co-located with) the δ corresponding to the correlation measurement input on the average decorrelation curve, suggesting that a quadratic approximation to L is both meaningful and useful. This quadratic approximation amounts to treating $\delta_{ij} = |z_i - z_j|$

as a Gaussian random variable with mean $\bar{\delta}_{ij}$ given by

$$\bar{\delta}_{ij} = \underset{\delta_{ij}}{\operatorname{argmax}}\ L(\rho_{ij}|\delta_{ij}, q) \tag{4.11}$$

and variance σ_{ij}^2 given by

$$\sigma_{ij}^2 = \frac{-1}{\frac{d^2}{d\delta_{ij}^2} L(\rho_{ij}|\bar{\delta}_{ij}, q)}. \tag{4.12}$$

(4.11) is solved easily by a standard 1D optimisation algorithm (Golden Section search was used in the experiments) using, for example, the δ read off an average decorrelation curve as an initial guess. The denominator of (4.12) is simply the second derivative of the cubic spline used to interpolate L (see Section 4.2), evaluated at the previously found value of $\bar{\delta}_{ij}$. Occasionally, the 1D optimiser fails to solve (4.11) (this can happen if the approximate L has no local maximum, for example) and σ_{ij}^2 is calculated as negative. In such cases, the measurement is discarded.

Using this Gaussian approximation, the original multi-dimensional optimisation problem amounts to that of minimising a sum of Mahalanobis distances. That is,

$$\tilde{\mathbf{z}}^* = \underset{\mathbf{z}}{\operatorname{argmin}} \sum_{j>i} \frac{(z_j - z_i - s_{ij}\bar{\delta}_{ij})^2}{\sigma_{ij}^2}, \tag{4.13}$$

where the sign s_{ij} of the displacement between mini-frames i and j is provided by the TSP-sort algorithm of Chapter 3. This is a much simpler optimisation problem than (4.10), and was solved analytically for the context of mobile robot localisation by Lu and Milios [139]. The following steps are direct adaptations from their work. Re-expressing (4.13) in matrix form yields

$$\tilde{\mathbf{z}}^* = \underset{\mathbf{z}}{\operatorname{argmin}}\ (\mathbf{S}^{TSP}\bar{\boldsymbol{\delta}} - \mathbf{H}\mathbf{z})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{S}^{TSP}\bar{\boldsymbol{\delta}} - \mathbf{H}\mathbf{z}), \tag{4.14}$$

where $\mathbf{S}^{TSP}$ is a diagonal matrix containing the s_{ij}, the signs found by the TSP-sort, algorithm, $\bar{\boldsymbol{\delta}}$ is the vector of distance estimates obtained from (4.11), $\mathbf{H}$ is a matrix whose entries consist exclusively of the values 0, 1 and -1, expressing the linear relationships between distances and absolute positions (of the form $s_{ij}\delta_{ij} = z_j - z_i$), and $\boldsymbol{\Sigma}$ is a diagonal matrix of the variances computed using (4.12). The analytical solution to the problem of (4.14) is [139]:

$$\tilde{\mathbf{z}}^* = (\mathbf{H}^\top \boldsymbol{\Sigma}^{-1} \mathbf{H})^{-1} \mathbf{H}^\top \boldsymbol{\Sigma}^{-1} \mathbf{S}^{TSP} \bar{\boldsymbol{\delta}}. \tag{4.15}$$

Because $\mathbf{\Sigma}$ is diagonal, the large matrix inversions involved in (4.15) can be avoided and the problem reduced to the solution of n linear equations, independently of the number of correlation measurements involved. Let $\mathbf{G}$ denote the $n \times n$ matrix $\mathbf{H}^\top \mathbf{\Sigma}^{-1} \mathbf{H}$ and $\mathbf{b}$ denote the n-dimensional vector $\mathbf{H}^\top \mathbf{\Sigma}^{-1} \mathbf{S} \bar{\boldsymbol{\delta}}$. Then, the non-zero entries of $\mathbf{G}$ are given by [139]

$$G_{ij} = \begin{cases} \sum_{k=0}^{n} \frac{1}{\sigma_{ik}^2} & i = j \\ -\frac{1}{\sigma_{ij}^2} & \text{otherwise} \end{cases} \tag{4.16}$$

and the entries of $\mathbf{b}$ are given by

$$b_i = \sum_{j=0, j \neq i}^{n} \frac{s_{ij} \bar{\delta}_{ij}}{\sigma_{ij}^2}. \tag{4.17}$$

The minimum variance solution $\tilde{\mathbf{z}}^*$ is then found by solving the linear system of equations

$$\mathbf{G} \tilde{\mathbf{z}}^* = \mathbf{b}. \tag{4.18}$$

In addition, this weighted least-squares methodology allows one to obtain a covariance matrix for $\mathbf{z}$. This covariance matrix, denoted by $\mathbf{C}$, is simply given by [139]

$$\mathbf{C} = \mathbf{G}^{-1}. \tag{4.19}$$

This covariance matrix captures the relative uncertainties in the inferred positions of the mini-frames; in particular, the main diagonal gives the variance in the position of each mini-frame with respect to the reference point and in this respect, it shows how uncertainty accumulates over frame acquisitions. Though it will not be exploited in this chapter, the covariance matrix $\mathbf{C}$ can provide useful information in the context of tasks which use the output of the mini-frame pose estimation framework or make decisions based on it. This idea will be explored in Chapter 7, where the uncertainty information will condition the inclusion or exclusion of new distance estimates via hypothesis tests in the context of possible speckle decorrelation model failures.

The maximum likelihood estimation framework described above theoretically allows for the fusion of any number of correlation measurements of arbitrary quality. On the one hand, the use of multiple measurements should increase precision by reducing the variance in $\mathbf{z}$

as measurement noise is averaged out. This is particularly appealing where non-monotonic probe motion is present as it allows for later measurements of a given area to confirm earlier measurements of the same area and forces all measurements in between to be consistent with both (a phenomenon known as "loop closing" in the robot localisation literature).

It may not, however, be desirable to use *all* possible measurements. Even in the context of Rayleigh scattering, slight local variations in speckle correlation length can be observed in finitely sized windows, as the random distribution of the scatterers in the target leads to slightly different local statistical properties for finite sample sizes. This is illustrated by Figure 4.2, which shows three different local decorrelation curves from synthetic imagery of fully developed speckle. While one of the curves follows the average curve well, the others show noticeable, though slight, differences in behaviour. While variations in the value of individual sample correlation coefficients are well modelled by the probabilistic model of Section 4.2, the maximum likelihood estimation framework presented above makes the useful, but not entirely valid assumption that these variations are spatially independent of each other. In fact, if the correlation ρ_{ik} between mini-frames i and k is higher than average given their elevational separation and mini-frame j lies between them, then ρ_{ij} and ρ_{jk} will also likely be higher than average given the configuration of the mini-frames. In regions where the correlation length increases, more frames correlate to each other, creating an over-abundance of measurements suggesting shorter inter-frame distances than is accurate. These measurements tend to dominate the data fusion process due to their number.

A suitable compromise is obtained by introducing three adjustable parameters into the proposed scheme, $\varphi_{\min}, \rho_{\min}$ and ρ_{new}. The first adjustable parameter, $\varphi_{\min}$, determines the minimum number of consecutive correlation measurements per mini-frame. $\rho_{\min}$ determines the minimum correlation required between a pair of mini-frames for the measurement to be used. A derived parameter $\varphi_{\max}$ determines the maximum number of consecutive correlation measurements per mini-frame. $\varphi_{\max}$ is simply calculated as the average number of consecutive mini-frames needed to reach $\rho_{\min}$ from a reference mini-frame in the sequence at hand. The second adjustable parameter, ρ_{new} is a minimum correlation value for which new measurements should be taken into account even if the $\rho_{\min}$ or $\varphi_{\max}$ thresholds have previously been violated and is useful for exploiting meaningful correlation measurements in sequences with non-monotonic probe motion. In the experiments presented here, $\varphi_{\min} = 2$ and $\rho_{\text{new}} = 0.7$. This value of ρ_{new} was selected empirically on the basis of performance in monotonic and non-monotonic synthetic ultrasound scans. In practice, it was found that

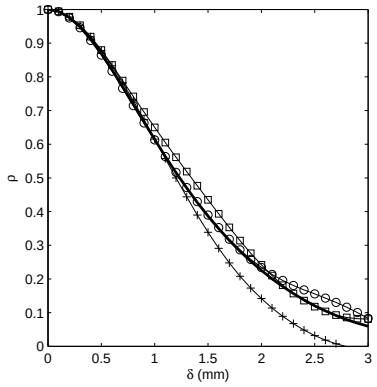


Figure 4.2 The average decorrelation curve constructed from 150 mini-frames of synthetic ultrasound imagery of a speckle phantom (thick solid line) and three local decorrelation curves estimated from 10 frames each in three non-overlapping 1.0 mm windows of the same phantom (squares, circles and crosses).

allowing $\rho_{\text{new}} < 1$ marginally improves performance in non-monotonic scans, but that the algorithm is not overall very sensitive to this parameter. ρ_{min} was defined for each patch by cutting off the lowest 20% of the range of correlation values observed during calibration.

4.4 Experiments

The trajectory recovery approach proposed in this chapter was tested on synthetic and real ultrasound imagery of speckle phantoms. The ML data fusion algorithm was compared to two alternative algorithms based on a deterministic piecewise linear model of speckle decorrelation. Section 4.4.1 describes these base-line approaches and the error measure that will be used for comparisons. Sections 4.4.2, 4.4.3 and 4.4.4 describe the image acquisition protocols and discuss the results of the experiments for different kinds of synthetic and real ultrasound image sequences.

Experimental results for the techniques described in this chapter were originally presented in [126]. Since the publication of this article, the experiments were substantially revised. Some minor software bugs in the implementation of the methods were corrected; others, which primarily arose from third-party libraries, were circumvented. The methodology was also tested on additional synthetic ultrasound data sequences which were still being generated at the time of article submission. More importantly, the experiments with

real imagery were conducted anew in their entirety.

4.4.1 Base-line algorithms and accuracy assessment

The ML elevational pose estimation component of the proposed approach (Section 4.3) was experimentally compared to two base-line algorithms which have the same function: given correlation measurements and the signs of the displacements between pairs of mini-frames corresponding to the same image patch, estimate the elevational position of each mini-frame centre with respect to the first. Both these base-line algorithms use a single correlation measurement per mini-frame in order to achieve this.

The first of these base-line algorithms positions each mini-frame using only its correlation measurement to its immediate predecessor. With some abuse of terminology, this approach is dubbed the "nearest-neighbour" (NN) algorithm and is illustrated in Figure 4.3(a). This approach is not expected to do very well because when the elevational motion is slow, it infers distances based only on error-prone high correlation measurements lying in a flat section of the average decorrelation curve, leading to the accumulation of a large number of inaccurate distance estimates and causing drift. The method is also expected to show strong variability in the accuracy of its reconstructions because each pose estimate is based on the accumulation of many uncertain measurements, leading to steadily increasing uncertainty along the chain.

The second base-line strategy used for comparison positions each mini-frame using its correlation to a certain reference mini-frame until the correlation coefficient falls below a minimal value $\rho_{\min}$ or a maximum number of frames $\varphi_{\max}$ is exceeded, at which point the last mini-frame positioned becomes the reference. $\rho_{\min}$ and $\varphi_{\max}$ were chosen as described in Section 4.3 for the ML approach. The method, henceforth referred to as the shifting reference (SR) method, is illustrated in Figure 4.3(b). A similar method (applied to entire frames) was used in [72] to avoid the accumulation of errors intrinsic to the NN method. The SR method reduces both drift and increases in uncertainty over time by using fewer highly biased distance measurements and by preferring direct distance measurements over the accumulation of several uncertain distances. Unlike the ML approach (illustrated in Figure 4.3(c)), the NN and SR methods use the deterministic average decorrelation curves and a single measurement per mini-frame.

After computing the deterministic and probabilistic elevational speckle decorrelation

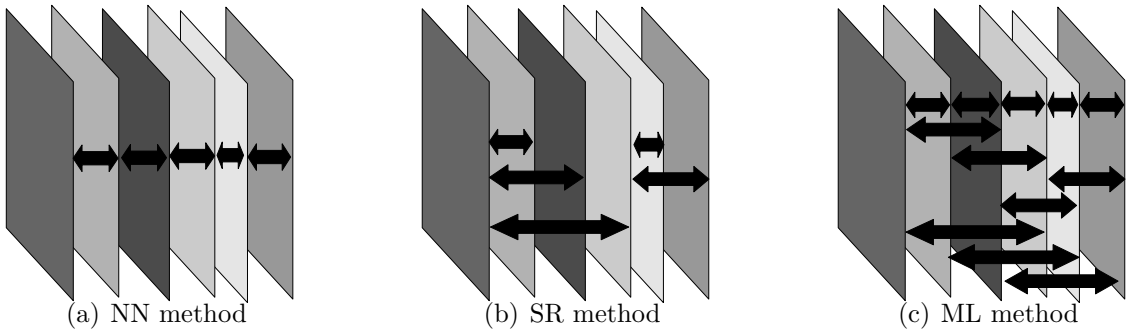


Figure 4.3 Schematic illustration of the three competing measurement fusion approaches in the case of a monotonically acquired scan (so that the geometric order of the frames is the same as their acquisition order in this example). The arrows depict the correlation measurements between pairs of mini-frames which are used by each method. The ML method is the only one to use multiple measurements referring to every mini-frame.

models from a calibration scan of a speckle phantom, each experimental trial amounts to computing the trajectory of the probe from an ultrasound image sequence using the global framework introduced in chapter 3. That is,

1. Correlations between pairs of corresponding mini-frames are calculated;

2. The signs of the elevational pose estimates are found using the TSP-sort algorithm of the previous chapter;

3. For each image patch location, the z positions of all mini-frames are estimated using the ML estimation approach described in Section 4.3 as well as the NN and SR algorithms;

4. The image planarity constraint is enforced using the three step robust rigid transformation fitting scheme described in the previous chapter, thereby correcting inverted mini-frame orderings that may have been given by the TSP-sort algorithm.

Since there remains an unresolvable direction ambiguity as to the overall scanning direction (which is chosen by convention only), a trajectory with reversed z positions is also computed. For all experiments, the reported results correspond to the estimated trajectory that best matches the ground truth.

For most experiments in this thesis, the accuracy of the recovered transducer trajectory is evaluated in terms of the mean target registration error (mTRE). Let ϕ_i^* denote the rigid transformation parameters obtained for frame i after the final rigid transformation fitting step described in Chapter 3 and $\mathbf{x}_i^{[q]}$ denote the ground truth coordinates of the centre of patch q in frame i. Then, the mTRE at frame i is given by

$$mTRE_i = \frac{1}{Q}\sum_{q=1}^{Q} ||\mathbf{x}_i^{[q]} - \mathcal{R}(\mathbf{x}_0^{[q]}, \phi_i^*)|| \tag{4.20}$$

This method of evaluating the accuracy of inferred geometrical transformations was advocated by van de Kraats et al. [224]. It has several advantages over other approaches (such as directly comparing the parameters of the recovered trajectory to those of the ground truth):

- The mTRE does not depend on the choice of a particular centre of rotation;
- It is easy to compute;
- It reflects the spatial distribution of locations where accuracy is desirable. For example, in a sector shaped ultrasound scan, there are more image data at the bottom of the image than at the top. This is reflected in the chosen grid of image patches for these scans and implies that rotational errors that mainly affect the position of the bottom of the image will be penalised more heavily than those which affect mainly the position of the top of the image.

In the context of this research, the mTRE was useful in quantifying the comparative accuracy of the different data fusion algorithms. For the purpose of statistical analysis, the accuracy of an entire trajectory will also often be summarised by the mTRE averaged over all its $n+1$ frames. Where accuracy is being compared between trajectories of significantly different lengths, the average mTRE will also be normalised by the total length of the trajectory (along the z direction) in order to allow for more meaningful comparisons. One disadvantage of the mTRE is that it does not provide an indication of the *type* of error which exists between geometrical transformations. As this is sometimes useful in order to understand the behaviour of a particular algorithm, the inferred out-of-plane transducer trajectory, expressed as displacement, yaw and tilt (calculated using the centroid of the

image patches as the centre of rotation), is also provided along with the mTRE in the full experimental results shown in Appendices A and B.

4.4.2 Experiments with synthetic ultrasound imagery

As a first step towards experimental validation, the probabilistic measurement fusion approach proposed in this chapter was first tested on synthetic ultrasound imagery obtained using the Field II ultrasound simulator [110, 108]. These experiments allowed verification of the theory in ideal imaging conditions with complete control and flexibility in the choice of transducer trajectories, with uncorrupted ground truth. This flexibility allowed for exploratory experimentation with controlled out-of-plane rotational motion, which was not possible with the setup used in the experiments with the real ultrasound scanner.

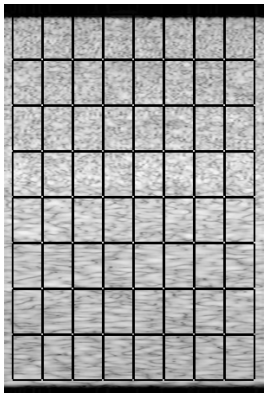
4.4.2.1 Generation of synthetic ultrasound data and trajectories

The synthetic imagery was generated using the Field II ultrasound simulator with a virtual 3.5 MHz linear array probe and a virtual speckle phantom scanned at a depth of 6 cm with a focus at 3 cm and a sampling frequency of 50 MHz. The virtual speckle phantom was created by randomly positioning a very large number of point scatterers within the scanning volume. The resulting RF vectors were envelope detected using the Hilbert transform. In order to demonstrate the feasibility of the proposed approach with RF scans, the synthetic data were not scan-converted or log-compressed.[3] The synthetic images were divided into 64 non-overlapping patches of 14 A-lines laterally by 490 samples axially (Figure 4.4(a)), for each of which a probabilistic speckle decorrelation model and an average speckle decorrelation curve were obtained by scanning the virtual phantom at 0.1 mm intervals.

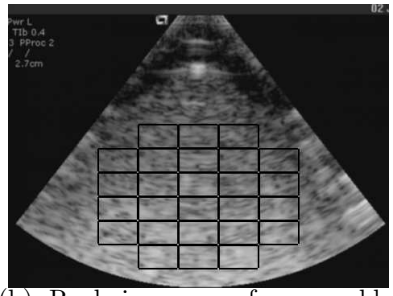
In addition to the calibration scan, 38 synthetic image sequences of fully developed speckle were generated using the Field II simulator with random probe trajectories in new virtual speckle phantoms. All these trajectories were constrained to remain purely out-of-plane, reflecting the assumption made throughout this thesis that any in-plane motion has already been accounted for. Different types of trajectories were considered:

1. Pure monotonic translations by exponentially distributed random steps with a mean of 0.15 mm (8 data sets);

[3] Scan-converted log-compressed synthetic data were used in earlier versions of this work [125] with very similar results.



(a) Synthetic imagery (interpolated laterally, sub-sampled axially to correct scale and log-compressed for illustration)



(b) Real imagery of a speckle phantom

Figure 4.4 The chosen subdivision of ultrasound images into non-overlapping patches.

2. Pure non-monotonic translations, where parallel frames were generated in the same manner, with the direction of the probe reversed with a probability of 2% at any step (10 data sets);

3. Monotonic translations (randomly generated as before) with rotational noise. At every step, the rotational noise consists of additional yaw and tilt motions drawn from zero mean normal distributions with a standard deviation of 0.1 degree (6 data sets).

4. Non-monotonic translations with rotational noise, with changes of direction occurring with a probability of 2% at any step (7 data sets);

5. Rotations about the vertical axis by exponentially distributed random steps with a mean of 0.2 (slow) or 0.3 (fast) degree and translational noise. At each step, an elevational translation is drawn from a zero mean normal distribution with a standard deviation of 0.15 mm (7 data sets: 2 slow, 5 fast). The first two frames in all these sequences were intersecting. In order to ensure that this was the case (so as to test the behaviour of the algorithms in such circumstances), a few non-intersecting frames

at the beginning of each sequence were sometimes deleted manually, as needed.

The rotations were applied using the centre of the transducer surface as the centre of rotation. Except in the case of the rotational sequences, all data sequences consisted of 150 frames.

Experiments on these different types of scans will evaluate several aspects of the methods proposed in this and the preceding chapter. The pure translational sequences will enable comparison of the different measurement fusion strategies under ideal circumstances. In noisy translational sequences, the addition of small amounts of rotational noise to otherwise translational sequences introduces a source of speckle decorrelation which is not captured by the speckle decorrelation models (deterministic or probabilistic) during calibration. Rotational noise, can thus be confused for translational motion and decrease the accuracy of the trajectory recovery process. These noisy translational sequences will also test how well small amounts of yaw and tilt can be recovered implicitly from the elevational translations measured at different locations. Experiments on non-monotonic trajectories will show how well the TSP-sort mini-frame ordering approach can perform under the ideal and less than ideal cases corresponding to pure and noisy translational motion. Experiments with non-monotonic noisy translational sequences and rotational sequences will test the ability of the framework proposed in the previous chapter to correctly recover intersections between frames, as well as the ability of the entire framework to recover large yaw motions.

4.4.2.2 Preliminary assessment of results and choice of statistical analysis

To avoid clutter, the complete set of results for these experiments on synthetic data were placed in Appendix A rather in the body of this chapter. The appendix provides plots comparing the progression of the mTRE over time as well as plots comparing the displacement, yaw and tilt components of the estimated trajectories to those of the true trajectory for the ML, NN and SR approaches. When appropriate, selected examples from this complete set will be discussed for illustrative purposes.

In order to guide the interpretation and statistical analysis of these results, the accuracy of each trajectory estimate was first summarised and quantified by the average mTRE calculated over all its frames. These results are presented in boxplot form, grouped by algorithm and trajectory type, in Figure 4.5.[4] From this boxplot, it is evident that all al-

[4]The boxplots presented in this thesis were generated using the SPSS statistical package, which explicitly

gorithms yielded smaller average mTREs in the (both pure and noisy) translational image sequences than in the primarily rotational ones. However, this observation is not particularly meaningful because the overall displacements undergone by any part of the image in the rotational sequences are very different, leading to a very different mode of error accumulation over time. The average mTREs obtained for the pure and noisy translational cases lend themselves much better to quantitative comparisons as the total displacement undergone by any part of the ultrasound image over an entire trajectory, and hence the potential for accumulating error, is, by design, very similar from one image sequence to the next.

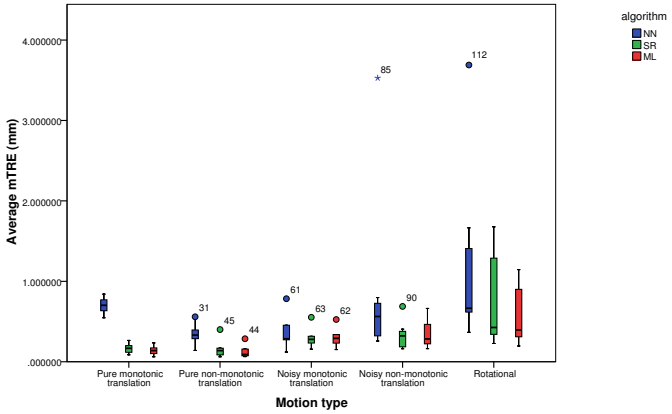


Figure 4.5 Boxplot of the average mTRE between the true and estimated trajectories for the NN, SR and ML algorithms, grouped by trajectory type.

A more meaningful observation is that the mTREs recorded for non-monotonic noisy translations appear to be more variable across cases than those recorded for the other types of translational trajectories, and that one of these cases produced an extreme outlier when the NN algorithm was used. It was found that the increased variability in the case of noisy non-monotonic translations was the result of frame ordering errors committed by the TSP-sort algorithm which arose because of misinterpretations of the unmodelled speckle decorrelation introduced by the rotational noise. Such frame ordering errors were not observed in the case of pure translations.

Based on these observations, a decision was made to carry out a formal statistical anal-

displays outlying data points as round dots or stars (for extreme outliers). The numbers shown next to these outliers are simply the unique identifier associated with the data in the software.

ysis based on a restrained portion of the results. The goal of this statistical analysis is to compare the performance of the NN, SR and ML measurement fusion algorithms and the direct effects of trajectory type on each of them under the assumption of correct input data from the frame-ordering portion of the framework. Because the average mTREs obtained with the rotational sequences could not be meaningfully compared to those of the pure and noisy translational ones, they were left out of this analysis. Among the translational sequences, three cases were also excluded from the analysis because they showed evidence of errors at the level of the TSP-sort algorithm. A case was labelled as showing evidence of frame ordering errors if at least one of the three algorithms tested reported a different number of direction reversals in its displacement estimation plot than existed in the ground truth. The remaining cases were grouped by motion type (pure or noisy translation) and by directionality (monotonic or non-monotonic) and analysed using a mixed between/within-subject factor analysis of variance (ANOVA) [199]. The effects of algorithm choice for any given sequence, motion type and directionality were evaluated, as well as all combinations of these. Where statistically significant effects were detected, pairwise comparisons between algorithms or combinations of algorithms and motion types were carried out, using the Bonferroni correction to adjust the results for multiple comparisons [199]. Statistical significance was evaluated at the 5% level. In the following sections, the main findings of the study in terms of the effects of algorithm choice, motion type and motion directionality will be interpreted in light of this quantitative analysis. Where relevant, cases which were excluded from the analyis will also be discussed as they provide important insights about the interaction between the different measurement fusion algorithms, the TSP-sort algorithm and the robust enforcement of the image planarity constraint discussed in Chapter 3.

4.4.2.3 Translational sequences and choice of measurement fusion algorithm

Overall, for the set of (pure, noisy, monotonic and non-monotonic) translational sequences considered in the statistical analysis, the SR and ML methods achieved the best results and significantly outperformed the NN method ($p < 0.001$ in both cases). An overall significant difference was also found between the SR and ML approaches, with the latter outperforming the former ($p = 0.006$), confirming that the probabilistic decorrelation model and measurement fusion approach proposed in this chapter are effective at quantifying and reducing error under controlled conditions.

The NN method generally did poorly because it uses distance estimates arising from high correlation measurements, which correspond to a region of the decorrelation curve which was poorly represented by its deterministic piecewise linear model. In these simulations, this meant that distances were generally underestimated by the NN method, as illustrated by an example in Figure 4.6.

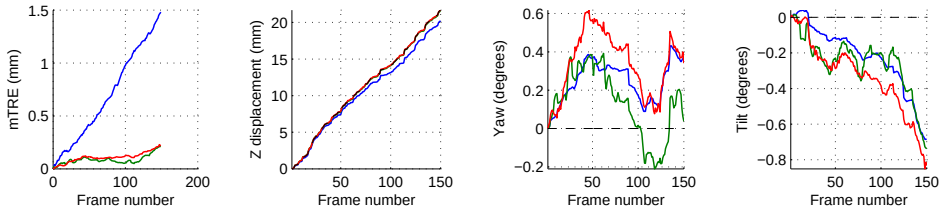


Figure 4.6 An example illustrating the tendency of the NN method to underestimate distances in the simulation study. In this example, the trajectory of the probe was purely translational and monotonic. From left to right: mean target registration error, estimated displacement, estimated yaw, estimated tilt using the centroid of the grid of image patches as the centre of rotation. Dashed-line: ground truth, blue: NN method, green: SR method, red: ML method. Note how the NN method's estimate of displacement lies beneath the ground truth.

Another noteworthy finding concerns the variability of the error across trials. Figure 4.7 shows the average, minimum and maximum mTRE as they evolve over 150 frames in the 8 monotonic, purely translational sequences, chosen for their comparable underlying ground truth trajectories. Essentially, the plot is an average of the 8 individual mTRE plots given in Appendix A. The plot shows that, for such cases, the ML method reduces the average error, but also how it limits the variability of this error, leading to generally more accurate and more repeatable results compared to the NN and SR methods.

4.4.2.4 Monotonic versus non-monotonic translations

Another interesting finding is that in the purely translational case (i.e. in the absence of rotational noise), the average mTRE is significantly reduced when the trajectory is non-monotonic ($p = 0.01$). This is particularly dramatic in the case of the NN method ($p < 0.001$). This improvement in overall accuracy is readily apparent when comparing the progression of the mTRE over time for monotonic and non-monotonic trajectories. Illus-

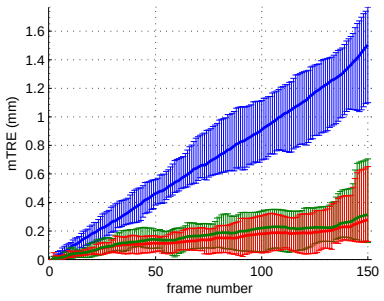
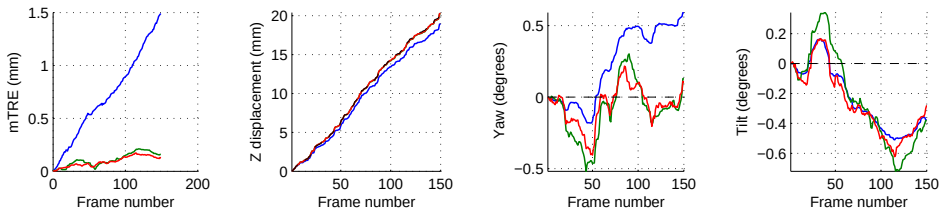


Figure 4.7 The average mTRE (in mm) incurred at each frame by the NN (blue), SR (green) and ML (red) approaches for pure monotonic translational sequences Synthetic 2-9. The top and bottom of the error bars indicate the maximum and minimum mTRE.

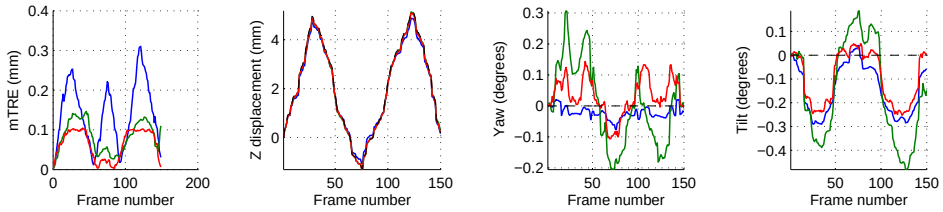
trative cases are shown in Figure 4.8. Whereas in the monotonic case, the mTRE incurred by the NN method continually increases over time, this error starts decreasing when there is a change in scanning direction for the non-monotonic case. The reason for this is that when the probe trajectory changes direction, any systematic drift intrinsic to the inference method also changes direction, thereby cancelling, to an extent, the drift accumulated prior to the direction change. From results such as those shown in Figure 4.8(b), it seems likely that direction changes might similarly affect the SR and ML methods, but that because of the slow drifting behaviour of these methods, more samples would be required to detect a statistically significant effect in these specific cases. Should this be true, it would suggest the possibility of devising new scanning protocols including such direction reversals in order to optimise accuracy.

4.4.2.5 Pure versus noisy translations

As already mentioned, rotational noise introduces speckle decorrelation in the data which is not accounted for by the elevational speckle decorrelation model obtained during calibration. For the SR and ML approaches, this resulted in significant decreases in accuracy ($p = 0.003$ in the SR case, $p < 0.001$ in the ML case), even when excluding cases associated with frame ordering errors. All algorithms had difficulty correctly estimating the small amounts of rotational motion present in these sequences, more so with tilt than with yaw. This is not very surprising since only translation is modelled at calibration time: rotations are only inferred implicitly through the calculation of elevational translation at



(a) Pure monotonic translation example

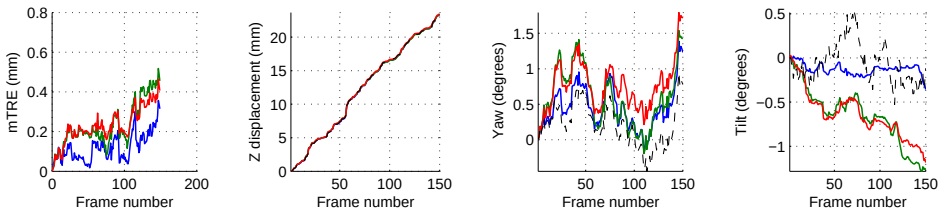


(b) Pure non-monotonic translation example

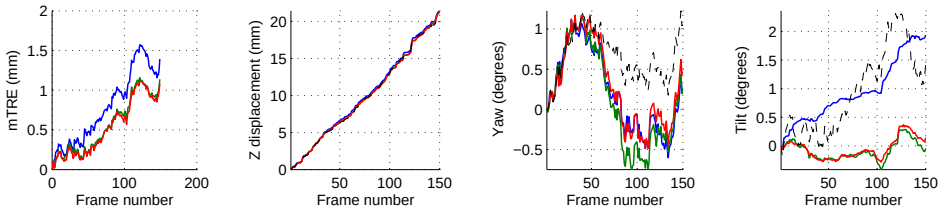
Figure 4.8 Comparison of results for monotonic and non-monotonic trajectories in the purely translational case. From left to right: mean target registration error, estimated displacement, estimated yaw, estimated tilt. Dashed line: ground truth, blue: NN method, green: SR method, red: ML method. For all methods, but especially NN, the mTRE undergoes a decrease when there is a change in direction.

different image locations. The scheme does not account for speckle decorrelation caused by rotational motion itself. Tilt estimates are particularly sensitive to this because the shape of the resolution cell varies significantly along the axial (in this case, vertical) direction, causing fast changes in the spatial distribution of the scatterers actually imaged by the transducer as it rotates, thereby substantially changing the way speckle decorrelates over distance [151, 111, 92]. Speckle decorrelation caused by small amounts of tilting is largely misinterpreted by the proposed model as part of the translational displacement, yielding poor tilt estimates as a result. This is evident in the examples shown in Figure 4.9.

While the SR and ML algorithms saw their respective accuracies decreased in the presence of rotational noise, statistical analysis revealed that, to the contrary, the unmodelled rotational speckle decorrelation significantly *improved* the accuracy of the NN algorithm in comparison to the purely translational case when the probe trajectory was monotonic ($p = 0.001$). This surprising result could be explained by the fact that the additional speckle decorrelation (implying larger distances) introduced by the rotations compensates



(a) A noisy translation example with poorly estimated tilt



(b) Another noisy translation example with poorly estimated tilt

Figure 4.9 Samples showing how tilt is poorly accounted for by a translational speckle decorrelation model. Left to right: mean target registration error, estimate of displacement, estimate of yaw, estimate of tilt using the centroid of the image patches as the rotation centre. Dashed line: ground truth, blue: NN method, green: SR method, red: ML method. In assessing the results presented in this figure and all others, the reader should keep in mind that different examples may have their axes scaled differently.

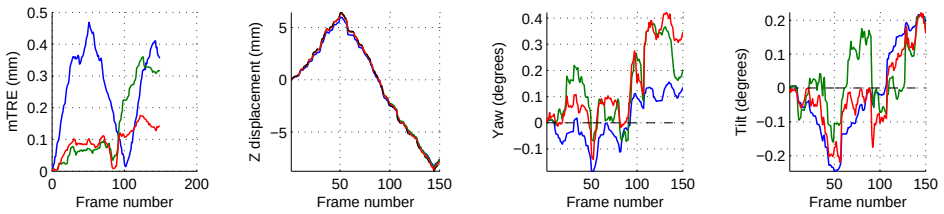
for the tendency of the NN algorithm to underestimate distances in the purely translational case. There was no such significant effect in the case of non-monotonic trajectories, possibly because non-monotonicity already compensates for drift, as discussed earlier. The case shown in Figure 4.9(a) is a particularly interesting one because for this case, the NN method did unusually well in terms of average mTRE and outperformed both the SR and ML methods. The plot of estimated tilt angles for this case provides an explanation for the atypical competitiveness of the NN method: on average, its estimated tilt component is closest to the ground truth, whereas the tilt curves for the SR and ML methods show strong evidence of drift error. However, the shape of the NN tilt curve is not at all similar to the ground truth, whereas the SR and ML curves do share features with it (hills and valleys in similar locations).

Another, indirect, effect of rotational noise is that the misinterpretation of tilt induced speckle decorrelation can lead the TSP-sort algorithm to commit mini-frame ordering errors in non-monotonic trajectories. As mentioned before, cases where this occurred were not

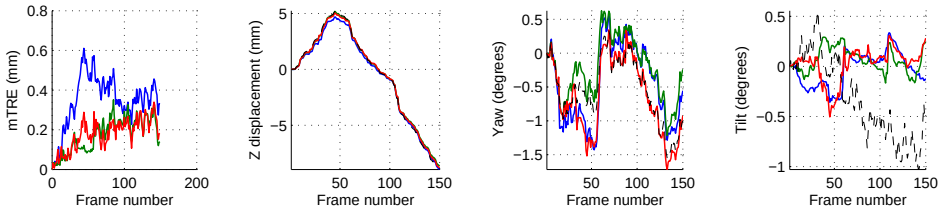
included in the formal statistical analysis, but they are still worthy of interest. Figure 4.10 shows a few examples of successful and unsuccessful frame ordering for pure and noisy non-monotonic translational sequences.

The behaviour of the different measurement fusion algorithms in the presence of frame ordering errors is interesting. The example shown in Figure 4.10(c) corresponds to the extreme outlier recorded for the NN case of the non-monotonic, noisy translational trajectories in the boxplot of Figure 4.5. From the displacement estimates shown in the second plot of Figure 4.10(c), it transpires that the source of the large average mTRE obtained for the NN method is that it completely missed the change in scanning direction near frame 20, implying that the mini-frame orderings found by the TSP-sort algorithm were incorrect, probably due to the considerable tilting motion occurring at that time. The SR and ML methods, in spite of not perfectly following the ground truth near the direction change, did not altogether miss it, even though they used the exact same TSP-sort output. The likely cause of this difference is that the sorting error probably involved a very small number of closely located frames in the vicinity of frame 20. Displacement sign errors therefore mostly occurred between frames close in acquisition order (those likely to be used by the NN approach), leaving the signs of larger displacements intact for use by the more robust SR and ML algorithms.

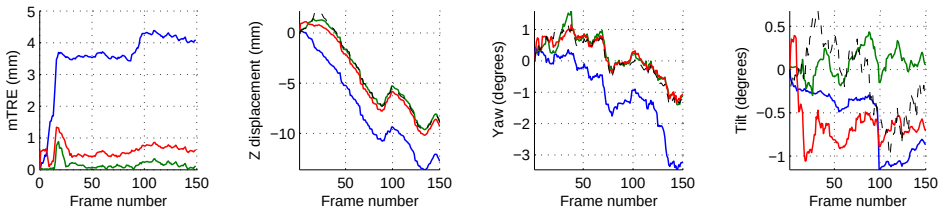
Figure 4.10(d) shows a more subtle case of frame ordering error. The displacement estimates obtained in the case of the NN algorithm in the second plot show that there was some shuffling of mini-frames for at least some of the image patches near frame 80. Once again, the SR and ML algorithms were more robust to these errors than the NN algorithms. However, note how between frames 80 and 140, all three data fusion algorithms show tilt estimates which are the reverse of the ground truth, with substantial error in the yaw component of the trajectory for exactly the same frames. This suggests that locally incorrect but consistent mini-frame orderings were obtained for laterally clustered patches in that part of the scan. In all but the NN case shown in Figure 4.10(c), all three algorithms were able to recover well enough from localised sorting errors to begin following the translational component of the trajectory reasonably consistently. This illustrates the strength of the TSP-sort algorithm: the errors it does make are generally local ones and do not completely change the global structure of the inferred trajectory.



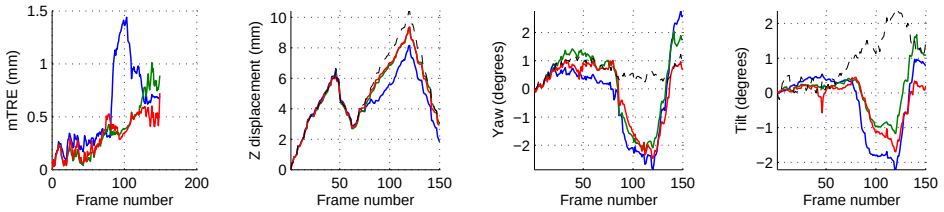
(a) Pure translation case with successful frame ordering



(b) Noisy translation case with successful frame ordering



(c) Noisy translation with a frame ordering error

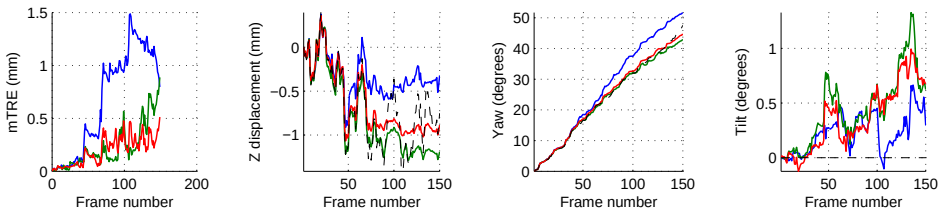


(d) Noisy translation with a frame ordering error

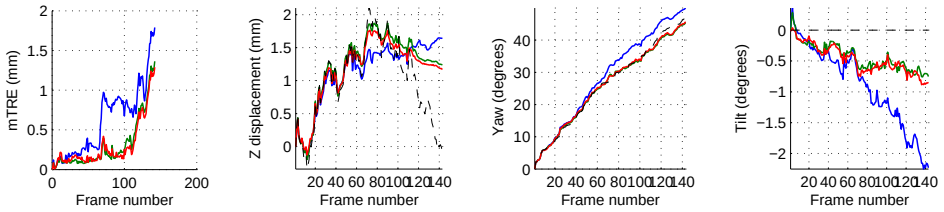
Figure 4.10 Sample non-monotonic translational trajectory results in the absence and presence of frame ordering errors. From left to right: mean target registration error, estimated displacement, estimated yaw, estimated tilt using the centroid of the grid of image patches as the centre of rotation. Dashed-line: ground truth, blue: NN method, green: SR method, red: ML method.

4.4.2.6 Rotational trajectories

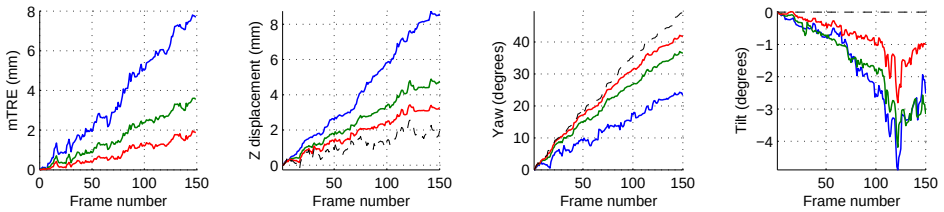
The results obtained for rotational image sequences were left out of the statistical analysis because they cannot be compared to those obtained for the translational results in any meaningful way. Trajectory recovery results for these sequences are thus interpreted separately in this section. Sample results are shown in Figure 4.11. The rotational image sequences presented a particular challenge to the approach proposed in this thesis because they violate its implicit assumption that displacements occur along a single direction. That is, the yaw motion which dominates these image sequences causes individual image patch displacements in both the lateral and elevational directions, the former of which are not explicitly accounted for by the fusion algorithms, but rather left for inference by the robust rigid transformation fitting step described in the previous chapter. As the magnitude of the rotation between a given frame and the first frame in the sequence increases, the fitting process becomes more difficult because of the increase in the size of the residuals caused by incorrect distribution of the estimated displacement over the elevational and lateral directions. In the case depicted in Figure 4.11(a), all methods managed to recover the true trajectory in spite of this, though this was achieved with varying levels of accuracy — the NN method once again yielded considerably poorer performance than its competitors, with ML generally outperforming SR slightly. More typically, errors become significant for very large yaw angles, as in the example shown in Figure 4.11(b), which displays relatively low error for most of the trajectory but a steep increase in the mTRE towards the end. While this leaves much room for conceptual improvement of the proposed approach, reasonably good results were still obtained for non-trivial rotation angles in most cases. The case shown in Figure 4.11(c) was particularly challenging because its first two frames were positioned so that there were nearly the same number of patches on either side of their intersection line. Thus, the LMedS fitting component described in Chapter 3 had to cope with a large number of inverted mini-frame orderings at relatively high noise levels. The robust standard deviation of (3.5) was too large to distinguish inliers from outliers, leading to highly inaccurate positioning for all methods. The ML approach was the one to fail most gracefully, giving results substantially closer to the ground truth than its competitors.



(a) Excellent trajectory recovery for fast yaw motion with noisy translation



(b) Typical trajectory recovery for fast yaw motion with noisy translation



(c) Poor trajectory recovery for fast yaw motion with noisy translation

Figure 4.11 The results obtained for three rotational data sequences generated using the same random trajectory parameters. From left to right: mean target registration error, estimate of displacement, estimate of yaw, estimate of tilt. Dashed line: ground truth, blue: NN method, green: SR method, red: ML method. Note the extreme differences in accuracy between the top and bottom cases for all methods.

4.4.3 Experiments with real imagery: controlled monotonic trajectories

In order to assess the validity of the proposed probabilistic speckle decorrelation model and measurement fusion approach in a more realistic scenario, experiments were also performed on real ultrasound imagery of a speckle phantom. This section presents the results of controlled experiments with monotonic translational trajectories for which the ground truth was well-known. Qualitative results obtained in the less constrained scenario of non-monotonic motion will be presented in Section 4.4.4.

4.4.3.1 Controlled phantom imagery acquisition

Real speckle phantom imagery was acquired one frame at a time using the Stradx software [165, 164] on a PC equipped with a video frame grabber connected to an Acuson Cypress clinical ultrasound machine with a 7V3C 5 MHz linear phased array probe at a depth setting of 2.7 cm. The phantom was a multi-modality vascular imaging phantom [35] made at and purchased from the Laboratory of Biorheology and Medical Ultrasonics, Montreal. The phantom contains a tubular structure which mimics a blood vessel and is surrounded by Rayleigh scattering material. For the experiments described in this thesis, the phantom was scanned at such a shallow depth that only speckle, not the tubular structure, appeared in the images. The real images were divided into 26 non-overlapping patches of 50×30 pixels (Figure 4.4(b)). The logarithmic dynamic range compression parameter was recovered by applying the algorithm described in [168] to one of the data sequences and used to obtain the echo intensity signal associated with all imagery.

The motion of the probe was restricted to the elevational direction and controlled to a 0.01 mm resolution using a manual positioning device based on a linear stage (Figure 4.12). The stage was mounted (with adjustable height) on a metallic stand and a metallic arm was attached to it in order to hold the transducer above the target. The device was custom designed and assembled by the departmental mechanical workshop. Several steps were taken in order to ensure the best possible accuracy of the ground truth measurements acquired using this device:

- The table where the experiments were carried out was purposely reinforced with a heavy, rigid, redundant steel frame.
- The positioning device was solidly secured to the table by way of a clamp.
- To avoid friction between the probe and the phantom, the latter was scanned in a water bath.
- Brass plates were affixed to the casing of the phantom with double sided tape in order to counter its natural buoyancy and ensure that it always remained submerged.
- Motion of the phantom within the water bath was further prevented by placing a sheet of water resistant sandpaper at the bottom of the bath and avoiding contact with the phantom, the bath or the table during the acquisition of data.

In similar conditions, single displacement measurements were found to repeatably match independent displacement measurements of the tip of the transducer bearing arm reported by a solidly anchored dial indicator with a 0.0005" (approximately 0.01 mm) resolution, supporting the use of the measurements reported by the device as ground truth.

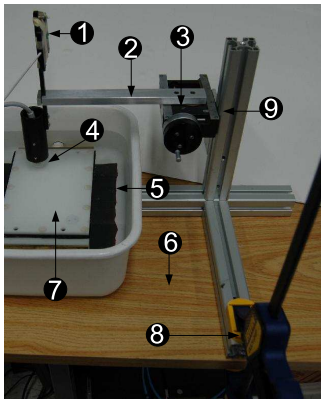


Figure 4.12 The real imagery acquisition setup. (1) Polaris tracker beacon (2) Probe bearing arm (3) Sub-millimetre linear stage (4) Tip of ultrasound transducer (5) Water resistant sandpaper (6) Reinforced table (7) Phantom (8) Clamp (9) At the time the data for the experiments of Section 4.4.4 were acquired, a mechanical wedge could be inserted or removed in this location to create chaotic or controlled monotonic trajectories, respectively.

The probe was attached to the beacon of a Polaris optical tracking device spatially calibrated using the algorithm presented in [169] and implemented in the Stradx software [165, 164] so that the tracker readings could be compared against when needed. The Polaris data could have been used as ground truth; however, its position measurements, though driftless, are corrupted by noise often more substantial than the smallest distance intervals travelled in the experiments, and it was felt that, in light of the results obtained with synthetic data, a less noisy data source would be required for comparison of the proposed algorithm to the SR algorithm, which can exhibit similar accuracy.

A note on experiment duplication When conducting the original quantitative phantom experiments presented in [126], some data sets were acquired for which the NN method gave more accurate results than the SR or ML methods. In light of the results obtained

with synthetic imagery, this was assumed to be anomalous. Upon visual inspection, irregularities were found in these image sequences which suggested that unwanted mechanical vibrations in the experimental setup had contaminated the ground truth measurements. These image sequences were thus excluded from the study. It was later discovered that several of the image sequences which had been included in the original analysis [126] were also affected by motion artefacts. As a result of this finding, the decision was made to duplicate the experiments in more carefully controlled conditions for the quantitative assessment of the different approaches in this thesis.

For the new experiments, mechanical adjustments were made to the positioning device in order to improve its stiffness and all image sequences were inspected for vibrations before running the trajectory recovery algorithms (few problems were encountered). The new set of experiments also provided an opportunity to carry out a more rigourous analysis and to test the algorithms more evenly over a broader range of transducer motion speeds than was originally done for [126]. Namely, 10 phantom image sequences were acquired for each of 6 different inter-frame spacings (0.03, 0.04, 0.05, 0.06, 0.08 and 0.1 mm), yielding 60 data sequences of at least 161 frames each (three of them had 171). This permits an analysis of the effects of transducer speed on the different algorithms.

4.4.3.2 Preliminary assessment of results and choice of statistical analysis

Since all image sequences consist of regularly spaced frames, any of them could be used as a calibration scan to build the speckle decorrelation model of Section 4.2. Over the course of preliminary evaluations, it was unexpectedly discovered that the choice of calibration scan was a source of substantial variability in the accuracy of the trajectories recovered by the SR and ML algorithms. Figure 4.13 illustrates the cause of this variability: upon analysis of the speckle decorrelation models obtained from different data sequences, it appears that there were sizeable variations in overall correlation length in different parts of the phantom. Given that the decorrelation curves and probabilistic decorrelation model are calculated from averages over large numbers of correlation measurements, these variations across different data sets suggest that the phantom may not have perfectly homogeneous Rayleigh scattering properties, or that the image patches were too small to warrant the repeatability of the average decorrelation curves observed in the case of synthetic data. The choice of image patch size was not investigated heavily in the context of this thesis, but there

most likely exists a trade-off between the repeatability of the observed correlation-distance relationship and the regularisation afforded by a large number of patches in enforcing the image planarity constraint. Preliminary evaluation also suggested that the speed at which the calibration scan was acquired (i.e., the resolution of the speckle decorrelation model) might affect the different algorithms in various ways.

In light of these preliminary observations, a decision was made to evaluate the accuracy of each algorithm on every possible pair of calibration and test image sequences, thereby providing as thorough and detailed an assessment of the different approaches as possible. The complete set of detailed results for these experiments is too large to include in this thesis in its entirety, owing to the thousands of possible pairings between image sequences. Appendix B provides a subset of these results, consisting of one case for each different combination of calibration and test sequence inter-frame spacings, for a total of 36 examples. In order to give a fair impression of the proposed method's performance, the examples were selected by finding, for each combination of calibration and test sequence inter-frame spacings, the image sequence which gave the median mTRE (averaged over the entire sequence) associated with the proposed ML measurement fusion scheme. The results presented in the appendix are plots comparing the mTRE, elevational displacement estimates, yaw and tilt estimates (using the centroid of the image patches as the centre of rotation) for the NN, SR and ML algorithms in the selected median cases. Illustrative examples from this subset of the results will be discussed here when relevant. Figure 4.14 summarises the results by showing, for a variety of calibration and test sequence inter-frame spacings, the progression of the mTRE over time averaged over all pairs of relevant image sequences.

4.4.3.3 Statistical analysis

Figure 4.14 suggests that the average accuracy of the different algorithms, and possibly their respective rankings, can vary substantially depending on the inter-frame spacing of the test and calibration image sequences. In light of this, a formal statistical analysis was conducted to assess and compare the accuracy of the recovered transducer trajectories under the influence of three factors:

1. The choice of measurement fusion algorithm (NN, SR or ML);

2. The inter-frame spacing of the image sequence for which trajectory is recovered;

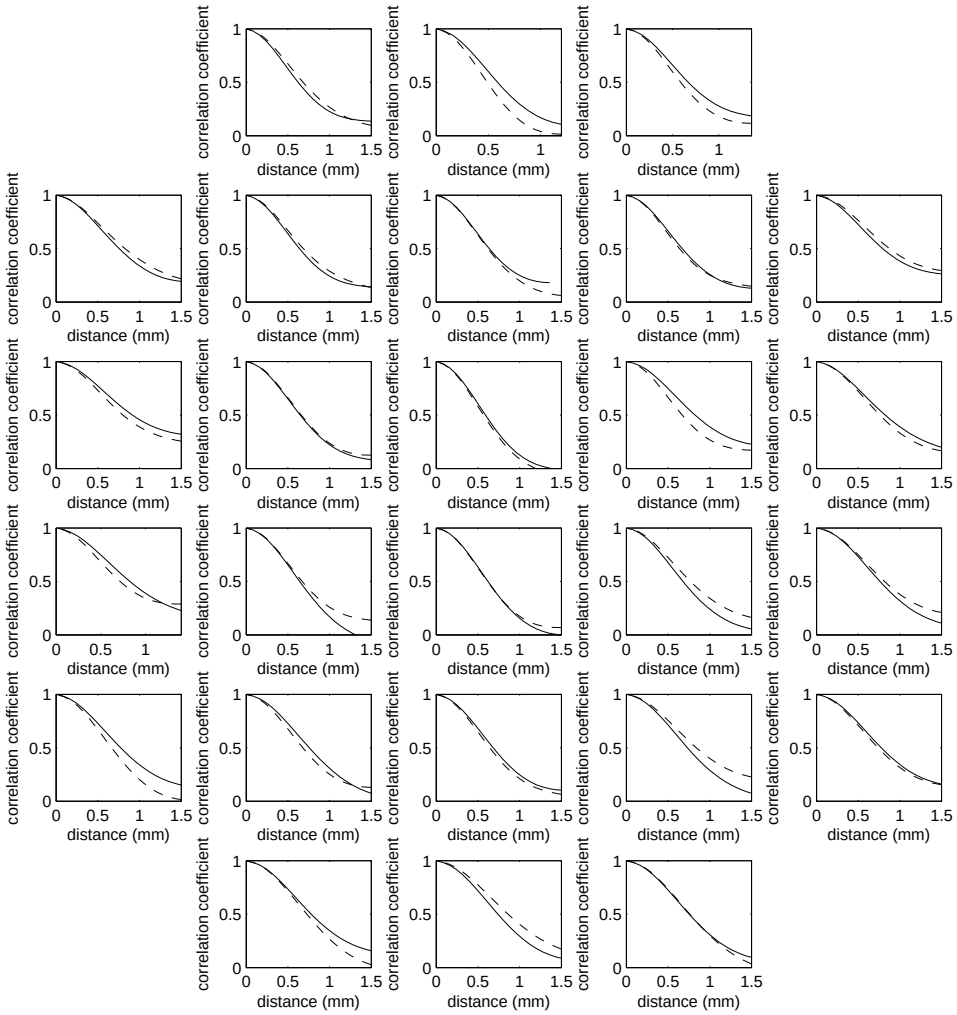


Figure 4.13 The decorrelation curves collected from two different data sets (corresponding to the solid and dashed lines) of speckle phantom imagery over the grid of image patches defined in Figure 4.4. Note how the decorrelation curves from the two data sets are in close agreement for some locations but have sizeable discrepancies in others. Large discrepancies are not expected in the context of pure Rayleigh scattering.

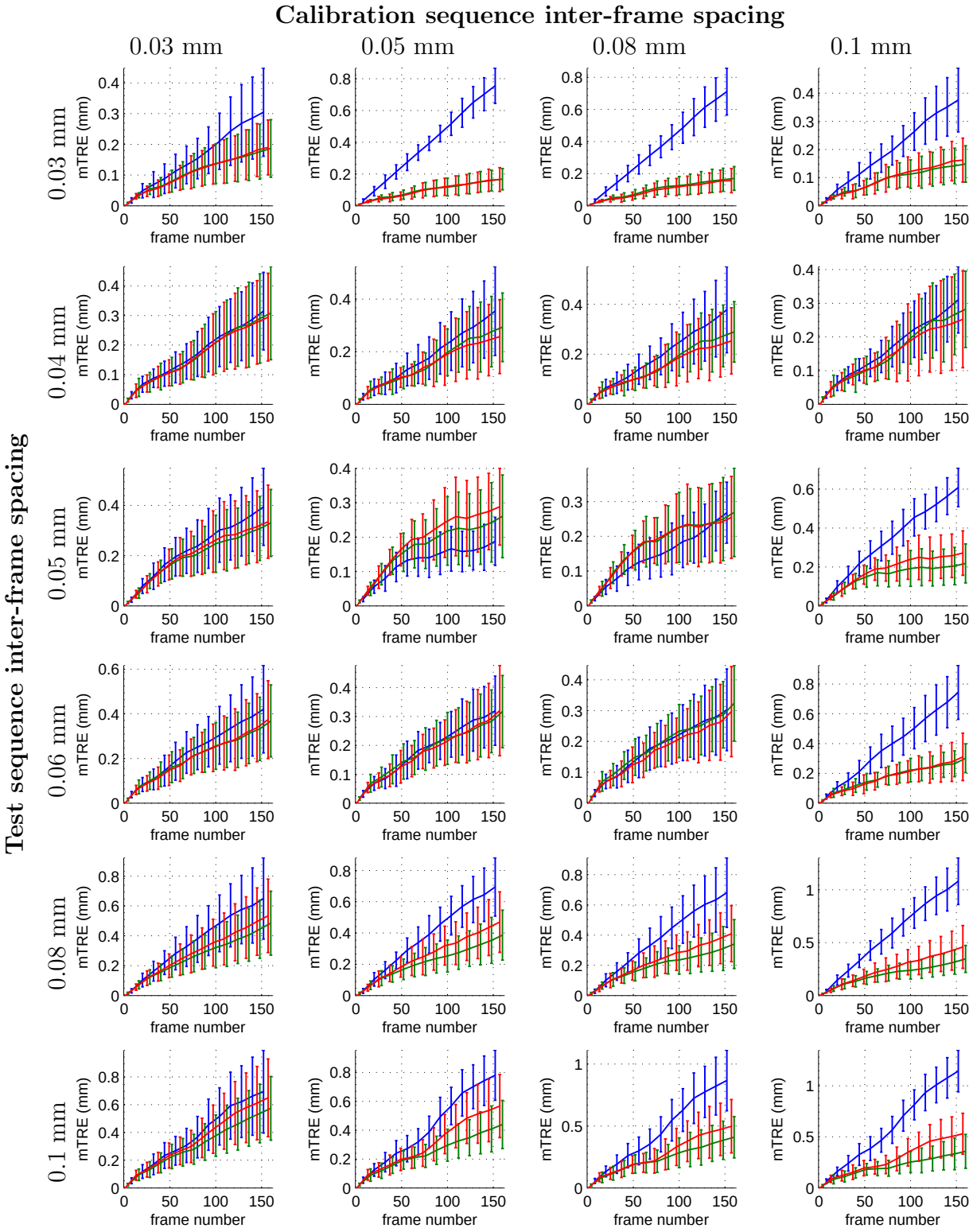


Figure 4.14 The progression of the average mTRE (in mm) over time for the NN (blue), SR (green) and ML (red) approaches in real phantom data sequences. The error bars represent plus and minus one standard deviation. For legibility, results for calibration inter-frame spacings of 0.04 mm and 0.06 mm were omitted and in each plot, the data points and error bars were sub-sampled at interleaved locations for the NN, SR and ML algorithms.

3. The inter-frame spacing of the image sequence used to obtain the speckle decorrelation model at calibration time.

Ideally, this analysis would have taken place over all possible pairings of calibration and test image sequences. This, however, introduces important statistical dependence relationships between the subjects of the analysis and, as a result, the raw data cannot be treated properly using conventional statistical analysis tools (i.e. the various flavours of analysis of variance and other standard linear models).[5] In order to eliminate problematic statistical dependencies between different pairs of subjects while accounting for the variability of the results obtained with different calibration data sequences, the image sequences were divided into a test set and a calibration set and analysed as follows:

1. For each level of inter-frame spacing, 3 sequences were randomly chosen to become part of the calibration data set, leaving 7 for the test data set. There were therefore 18 calibration sequences and 42 test sequences.

2. In order to keep the analysis manageable, the effect of calibration inter-frame spacing was evaluated at 3 levels instead of the original 6. This was done by collapsing pairs of consecutive levels of inter-frame spacing. The resulting three levels were labelled as dense (0.03 or 0.04 mm spacing), moderate (0.05 or 0.06 mm spacing) or sparse (0.08 or 0.1 mm spacing).

3. For each test sequence, measurement fusion algorithm and calibration sequence within a given calibration spacing level, the accuracy of the recovered trajectory was summarised as the mTRE averaged over all the frames of the test sequence and normalised by the true length of the sequence. This allows meaningful comparison of relative accuracy between test sequences acquired at different inter-frame spacings. These normalised average mTRE scores were then averaged over all calibration sequences within each level of calibration spacing, yielding one score for each test sequence, algorithm and calibration spacing combination. This last averaging process is meant to account for some of the observed variability arising from the use of different calibration sequences.

[5]This type of experimental design (involving multiple redundant pairs of subjects) is sometimes encountered in psychological research which considers relationships between and within pairs of individuals. It is known as the *round-robin* experimental design and while statistical methods for analysing data which arise from this type of design exist [113], they are somewhat difficult to use and are not available as part of commercial statistical analysis packages.

4. A mixed effects linear model [235] was fitted to these data, treating the identity of test sequences as a random effect, and algorithm and calibration inter-frame spacing as fixed, repeated effects and test sequence inter-frame spacing as a between-subject fixed effect.

5. The effects of test sequence inter-frame spacing, calibration sequence inter-frame spacing (in terms of the 3 above defined levels), choice of measurement fusion algorithm and combinations thereof were tested for statistical significance at the 5% significance level. Multiple comparisons were then carried out between the different algorithms using the Bonferroni correction.

This statistical analysis revealed that the test and calibration sequence inter-frame spacings, not the just the choice of measurement fusion algorithm, had an important influence on the accuracy of the recovered trajectories. These results are discussed in more detail below.

The effect of test sequence inter-frame spacing The inter-frame spacing of the test sequence was determined to have a statistically significant effect on the accuracy of the NN algorithm ($p < 0.001$). In fact, the behaviour of the NN algorithm varies tremendously with the inter-frame spacing of the test image sequence and this effect is also highly dependent on the spacing of the sequence used for calibration. This is readily apparent from some of the plots shown in Figure 4.14, as well as Figure 4.15. The latter figure shows how the normalised average mTRE score varies with different combinations of test and calibration sequence inter-frame spacing. It is evident from this plot that the NN approach had difficulty accurately recovering trajectories with very tight (0.03 mm) inter-frame spacing. Excluding this extreme case, there were also noticeable variations in accuracy among the other test sequence inter-frame spacings, with a tendency to achieve better accuracy at moderate (0.05, 0.06 mm) rather than extreme spacings; in such cases, it sometimes outperforms the SR and ML approaches, as shown in Figure 4.14. The source of these somewhat strange results is investigated in Section 4.4.3.4.

The test sequence inter-frame spacing also had a statistically significant effect on the accuracy of the SR algorithm ($p = 0.037$). The effect of test sequence inter-frame spacing on the SR algorithm manifests itself through more accurate results in the case of wide inter-frame spacings, irrespective of calibration sequence inter-frame spacing. Recall from Figure 4.3(b) that for a given set of corresponding mini-frames, the SR algorithm uses

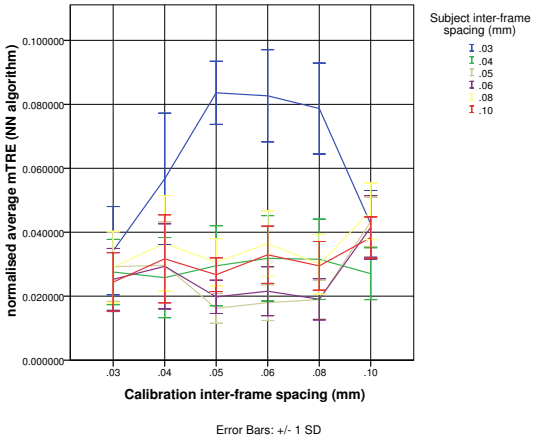


Figure 4.15 Normalised average mTRE obtained with the NN algorithm for different combinations of calibration and test sequence inter-frame spacings.

some of the same distance estimates and commits the same distance estimation errors as the NN approach. In most cases, these errors are worst at small inter-frame spacings (e.g. 0.03 mm). When the inter-frame spacing is wider, these NN-type distance estimates improve and the accuracy of the SR approach improves correspondingly. In contrast, test sequence inter-frame spacing was not found to have a statistically significant effect on the performance of the proposed ML approach ($p = 0.558$), implying that in image sequences with irregularly spaced frames, this approach would exhibit more consistent performance than the others over changes in transducer speed.

The effect of calibration inter-frame spacing The inter-frame spacing of the image sequence used for calibration was found to have a statistically significant effect on the accuracy of all three measurement fusion algorithms ($p < 0.001$ in all cases). In the case of the NN algorithm, this effect is dependent on the inter-frame spacing of the test sequence under consideration, as discussed earlier. For the SR and ML algorithms, this effect manifests itself through smaller normalised average mTRE in the case of widely spaced calibration scans (i.e. sparsely sampled speckle decorrelation models) than in the case of densely spaced calibration scans, irrespective of the inter-frame spacing of the test sequence under consideration. This trend may seem counterintuitive: one might expect displacement readings from the piecewise linear decorrelation curve to become more accurate as the curve

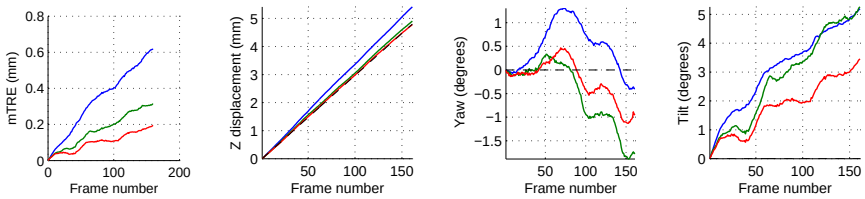
is more densely sampled. However, the opposite was observed. One possible explanation is that in these experiments, widely spaced image sequences sampled a larger portion of the speckle phantom target than densely spaced ones as they all had similar numbers of frames. Because the widely spaced sequences covered a larger distance, the micro-structure imaged in the different frames may have been more varied than in densely spaced sequences and the speckle decorrelation models derived from the imagery may thus have been more representative of scattering conditions encountered elsewhere in the phantom.

The effect of the choice of measurement fusion algorithm Overall, the SR and ML fusion algorithms were found to yield significantly lower normalised average mTRE than the NN algorithm ($p < 0.001$ in both cases). No statistically significant overall normalised average mTRE difference was found between the SR and ML algorithms. However, for some specific combinations of calibration and test sequence inter-frame spacings, differences between the SR and ML algorithms were found which *approached* statistical significance ($p-$values ranging between 0.05 and 0.1). In light of the small sample sizes used (7 cases for each combination of spacing levels) and of the increases in Type II error rates which arise from the use of the Bonferroni correction in multiple comparisons, these results should be considered as ambiguous rather than decisive, and interpreted with caution. These nearly statistical significant results mostly reflected the information conveyed by Figure 4.14, i.e. that for dense spacings in both test and calibration sequences, ML might outperform SR, whereas SR might outperform ML in the opposite situation (sparse inter-frame spacing for both the calibration and test sequences).

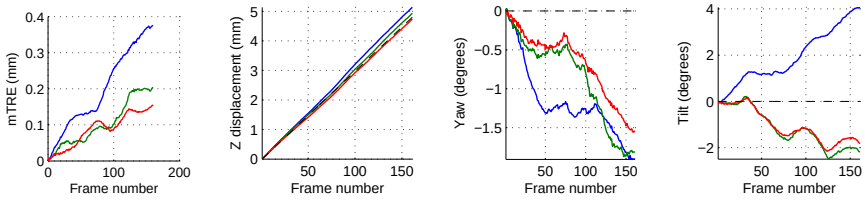
4.4.3.4 Further analysis: source of errors

What neither the statistical analysis nor Figure 4.14 convey is the *types* of errors committed by the NN, SR and ML approaches under different combinations of test and calibration sequence inter-frame spacings. For this information, one must refer to full trajectory recovery data such as the sample results given in Appendix B. For the purpose of discussion, sample results are reproduced in Figure 4.16.

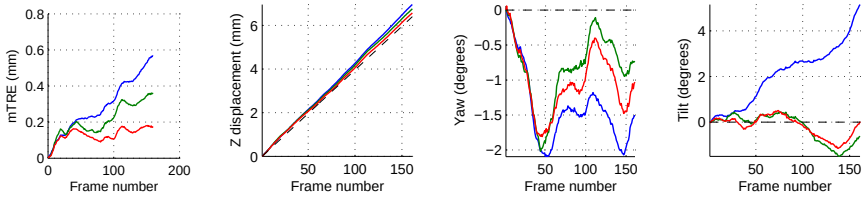
One detail that is immediately apparent from the plots shown in Figure 4.16 (and the more extensive set of results in Appendix B) is that for the real speckle phantom image sequences acquired in these experiments, the NN measurement fusion algorithm generally tended to overestimate transducer displacement in sequences with small (especially



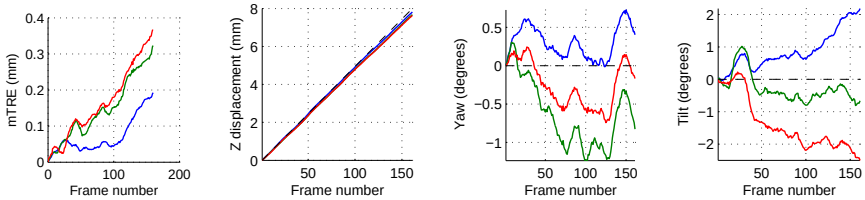
(a) Calibration spacing = 0.06 mm, Subject spacing = 0.03 mm



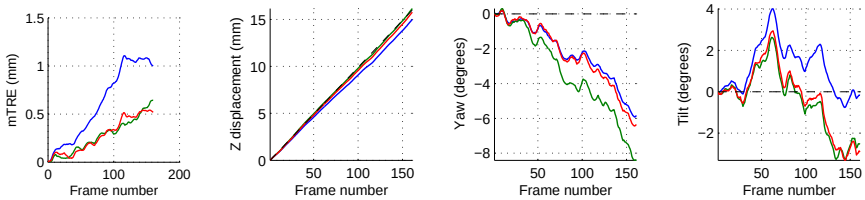
(b) Calibration spacing = 0.1 mm, Subject spacing = 0.03 mm



(c) Calibration spacing = 0.08 mm, Subject spacing = 0.04 mm



(d) Calibration spacing = 0.08 mm, Subject spacing = 0.05 mm

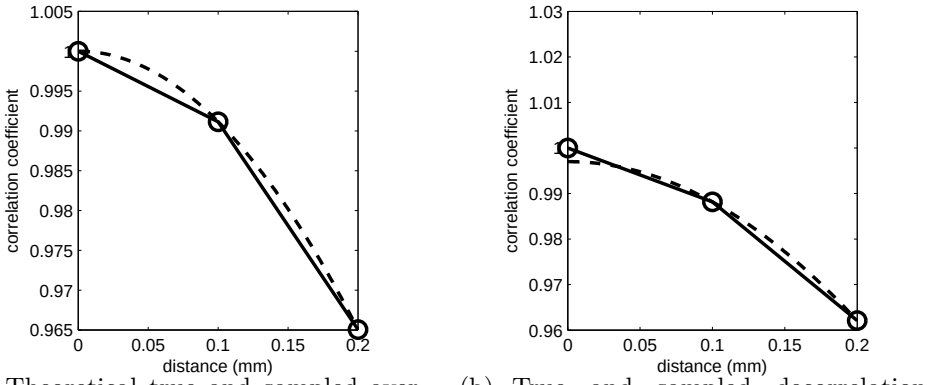


(e) Calibration spacing = 0.1 mm, Subject spacing = 0.1 mm

Figure 4.16 Sample trajectory recovery results for real phantom imagery illustrating the types of errors committed by each algorithm (in particular, the NN algorithm) for different test and calibration sequence inter-frame spacing combinations. From left to right: mean target registration error, estimated displacement, estimated yaw, estimated tilt using the centroid of the grid of image patches as the centre of rotation. Dashed-line: ground truth, blue: NN method, green: SR method, red: ML method.

0.03 mm, but often 0.04 mm) inter-frame spacing and underestimate it for sequences with large (mostly 0.1 mm, but also 0.08 mm) spacing, with continuous variation in between these extremes. This is a somewhat surprising result which deserves an explanation. The theoretical shape of the true speckle decorrelation curve is Gaussian like, with a maximum normalised correlation value of 1, concave but nearly flat for high correlation values. In theory, a piecewise linear representation of this curve would therefore tend to provide *underestimates* of distance for large correlation values rather than overestimates, as illustrated in Figure 4.17(a). In the case of real data (such as the imagery considered in these experiments), however, highly correlated imagery may see its correlation reduced because of uncorrelated noise in the images arising from the data acquisition process (in these experiments, it is believed that interlacing artefacts from the video frame grabber might be at fault). The consequence of this is that the maximum average normalised correlation value may be smaller than 1, which will not be properly reflected by the speckle decorrelation model inferred from the calibration scan. An illustration of this scenario is shown in Figure 4.17(b). In this case, displacement over-estimates may be observed for very small distances, with the displacement estimation process reverting to its theoretical behaviour for larger inter-frame separations. More accurate displacement estimates can be observed where the true decorrelation curve is close to the piecewise linear model, explaining the relatively good performance of the NN method in cases like that shown in Figure 4.16(d). The location of optimal inter-frame spacings depends on the inter-frame spacing of the calibration data sequence. Different results may also have been obtained if the echo amplitude signal had been used instead of the echo intensity signal. In realistic scenarios, freehand ultrasound image sequences are unlikely to have regularly spaced frames. The NN approach is thus unlikely to perform well as the empirical decorrelation model obtained from the calibration phantom is nearly always a very poor representation of the true, non-linear decorrelation behaviour in the high correlation range, as previously noted by others [204, 85].

Whereas the NN approach tends to either over-estimate or under-estimate transducer displacement, the mTREs observed for the SR and ML approach more often tend to arise from errors in the rotational components of the trajectory. This can be seen in several of the examples given in Figure 4.16. It seems likely that these rotational errors arise from local differences in the micro-structure of the phantom which lead to discrepancies between the predictions made by the speckle decorrelation model obtained from the calibration scan



(a) Theoretical true and sampled average decorrelation curves

(b) True and sampled decorrelation curves in the presence of noise

Figure 4.17 Illustration of the discrepancies between the piecewise linear speckle decorrelation model used by the NN approach (solid line, with samples represented by circles) and the "true" average decorrelation curve (dashed line) for short distances. In the absence of noise, the maximum achievable average correlation coefficient is 1 and the NN approach underestimates distance for large correlation values. When noise is present, the true curve can begin with an average correlation lower than 1, leading to distance over-estimates in the case of very short distances.

and the model that would have been obtained if the test scan had been used for calibration instead. Such discrepancies were already illustrated in Figure 4.13. The rotational errors arise when these discrepancies cause displacement estimation errors in image patches at the periphery of the grid (i.e. those which have the strongest influence on the yaw and tilt estimates when enforcing the image planarity constraint).

4.4.4 Experiments with real phantom imagery and non-monotonic trajectories

In order to test the ability of the proposed scheme to recover non-monotonic trajectories in real data (with the help of the TSP-sort algorithm introduced in the previous chapter), another set of experiments was conducted. At the time the experiments were carried out, the probe bearing arm of the sub-millimetre positioning device could be loosened or tightened by inserting or removing a mechanical wedge between the stage and its supporting metallic frame (see Figure 4.12). With the setup thus loosened, 9 data sets of 101 frames of phantom imagery were acquired for which the probe exhibited chaotic elevational dis-

placement patterns with brief direction changes. The speed of displacement varied between approximately 0.03 mm/step and 0.08 mm/step. One or two frames were deleted from the beginning of three of these image sequences because the chaotic probe displacement caused the correlation between these frames to be too low for motion recovery.

In these experiments, the motion of the probe was measured using a Polaris optical tracker spatially calibrated according to the wall-phantom method presented in [169], providing a noisy source of ground truth measurements. Upon comparison of the Polaris data gathered from controlled (i.e. without arbitrary direction changes) experiments with positioning device measurements, it transpired that whether due to a physical imperfection of the setup or errors in the tracker calibration process, the axis of motion of the positioning device was not perfectly perpendicular to what the tracker calibration output determined to be the image plane. The orientation of the positioning device's motion axis within the reference frame inferred by calibration was estimated by fitting a line to the Polaris data associated with the calibration scan. This information was used to adjust the trajectories estimated from image data for comparison with the Polaris measurements.

A single calibration scan with 0.08 mm inter-frame spacing was chosen among the data sets used in the experiments of the previous section, and the probabilistic speckle decorrelation model obtained from it was used in the recovery of the non-monotonic trajectories. Figures 4.18 to 4.19 show the results of these experiments. The plots compare the trajectory inferred using the proposed image-based approach to that recorded by the Polaris in terms of displacement, yaw and tilt (using the centroid of the grid of image patches as the centre of rotation). The displacement plots show that the proposed approach was generally able to recover the structure of the true trajectory quite well, including abrupt direction changes (the results shown in Figure 4.18(b) and 4.18(d) are very illustrative of this). These appear as sharp spikes or deep valleys in both the Polaris and image-based trajectories, among smaller noise-induced fluctuations in the Polaris data for which the image-based trajectory remains smooth.

In some cases, the image-based approach appears to produce overly smooth results. Consider, for example, some of the valleys in the Polaris trajectory in Figure 4.18(c). Such over-smoothing may occur because such deep valleys or sharp spikes in the trajectory force the ML fusion algorithm to rely on very few correlation measurements, which are close to zero (corresponding to large distances) and highly uncertain, in order to estimate the position of the mini-frame corresponding to the spike or valley. The distance estimates for

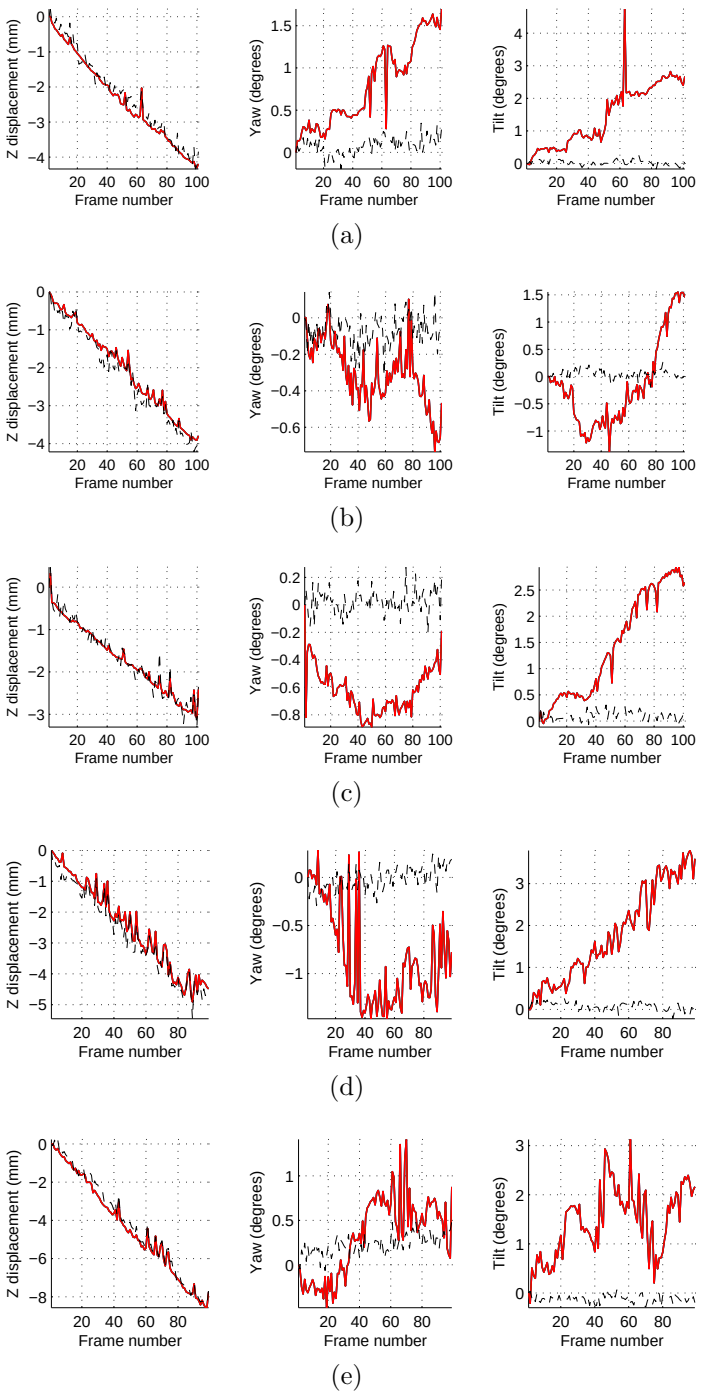


Figure 4.18 ML trajectory recovery results for real phantom imagery and chaotic trajectories, part 1. From left to right: estimated displacement, yaw and tilt angles, measured using the centroid of the grid of image patches as the centre of rotation. Dashed line: Polaris optical tracker measurements, solid line: trajectory recovered by the proposed approach.

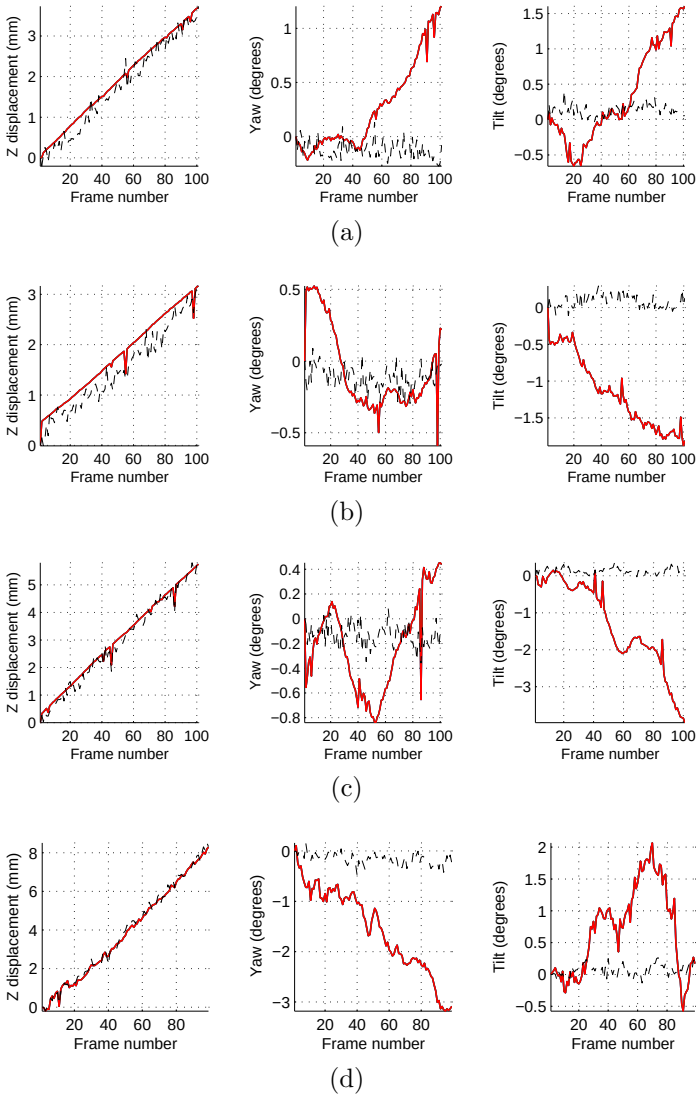


Figure 4.19 ML trajectory recovery results for real phantom imagery and chaotic trajectories, part 2. From left to right: estimated displacement, yaw and tilt angles, measured using the centroid of the grid of image patches as the centre of rotation. Dashed line: Polaris optical tracker measurements, solid line: trajectory recovered by the proposed approach.

pairs of mini-frames surrounding the spike or valley are much more numerous and much more reliable, and may easily dilute the contribution of these uncertain estimates if they disagree with the rest.

In some cases (e.g. Figures 4.18(d) and 4.19(c)), there was significant rotational error, considering the shortness of the trajectories. These errors, as with the previous experiments, may have arisen due to localised discrepancies between the decorrelation model established at calibration time and the local properties of the phantom area under consideration. Another important source of error was the fact that very little effort was made to control the quality of the imagery in these experiments, so that some data sequences were corrupted by substantial blurring and interlacing artefacts arising from motion of probe during the acquisition of data by the video frame-grabber. The image sequence corresponding to Figure 4.18(d) had noticeably poor quality in this respect. These may have decreased inter-frame correlations in localised parts of the image plane, leading to yaw and tilt errors.

4.5 Summary and discussion

This chapter presented a novel speckle decorrelation model which accounts for the variability of the sample correlation coefficients measured between finitely sized image patches. This model captures the uncertain and somewhat skewed relationship that exists between elevational separation and the measured correlation coefficient and allows for the prediction of distance estimates with associated uncertainty information, making it a richer model than the deterministic approaches used in all previous literature. This chapter also showed how the probabilistic speckle decorrelation model can be exploited for the probabilistic fusion of correlation measurements between any number of correlated mini-frames in an ultrasound data set.

Out-of-plane transducer trajectory recovery experiments were carried out to assess the performance of the proposed probabilistic measurement fusion scheme with speckle phantom imagery. The proposed approach was compared to two different deterministic trajectory recovery approaches which use a minimal number of measurements, the naive NN algorithm and the better behaved SR algorithm.

Experiments with synthetic ultrasound data allowed evaluation of the three algorithms under different conditions with respect to the type of trajectory undergone by the trans-

ducer, including monotonic and non-monotonic versions of pure translational trajectories and translational trajectories with mild rotational noise, as well as rotational trajectories. These experiments also allowed implicit validation of the TSP-sort mini-frame ordering algorithm and robust rigid transformation fitting approaches described in the previous chapter. Assuming that there are no frame ordering errors, it was found that, in synthetic imagery, for the types of trajectories analysed, the SR and ML approaches both significantly outperform the NN approach and that the ML approach significantly outperforms the SR approach as well. The presence of rotational motion in the trajectory, in particular rotations with an axial component (tilt), is detrimental to accuracy because neither the deterministic nor the probabilistic speckle decorrelation models considered here account for the speckle decorrelation caused by motion other than elevational translation. Under general conditions, improved results might be obtained by using rotational speckle decorrelation models [92] or prior knowledge about the probe trajectory.

The same three algorithms were also evaluated in experiments on real ultrasound imagery with monotonic translational trajectories of constant speed. The role played by transducer speed (inter-frame spacing) at calibration and trajectory recovery time was also evaluated for each algorithm. It was found that the accuracy of the NN algorithm is highly dependent on the particular combination of calibration and test inter-frame spacings under consideration, making it a poor choice in realistic conditions where transducer speed is not constant. The ML approach was found to be the most stable of the three over changes in test sequence inter-frame spacing, whereas the SR approach appears to have a preference for wide inter-frame spacings. Both the SR and ML approaches were found to be sensitive to the calibration inter-frame spacing, with a preference for widely spaced calibration sequences. It is hypothesised that this preference is due to the broader set of scattering conditions sampled and embedded in the speckle decorrelation model. The ability of the proposed approach to recover non-monotonic transducer motion using the TSP-sort algorithm was also demonstrated.

There remains some ambiguity as to whether there are significant differences in accuracy between the SR and ML algorithm in real speckle phantom imagery and whether some scanning conditions favour one over the other. Qualitatively speaking, the results suggest that the ML approach might be better at slow transducer speeds and that the SR approach might be best with fast transducer speeds. This analysis may have been obscured by the variations observed in the speckle decorrelation rates measured in different scans of the

phantom (as illustrated in Figure 4.13). While these variations were not expected in the context of experiments with speckle phantom imagery, they certainly arise in more realistic scenarios. In imagery of real tissue, variations in speckle decorrelation rate can be far more pronounced and the speckle decorrelation models derived from speckle phantom imagery become insufficient for accurate trajectory recovery.

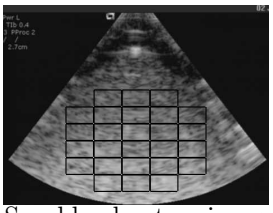
The remainder of the work presented in this thesis aims at generalising the trajectory recovery approach presented in this chapter to the case of real tissue. This begins in the next chapter with the study of a new model for the spatial distribution of ultrasound scatterers. In Chapter 6, this model will be used to generate synthetic imagery from which the characteristics of different tissue types can be learned and exploited for the extension of the proposed probabilistic decorrelation model to arbitrary tissue types. In this context, Chapter 7 will show that the probabilistic approach presented here has a conceptual advantage over deterministic approaches (such as the NN and SR approaches considered here). This advantage stems from the fact that it preserves uncertainty information that can be used to make informed decisions regarding the inaccuracy of distance measurements arising from model failure. This powerful feature will be demonstrated in the context of an iterative measurement selection and fusion strategy.

Chapter 5

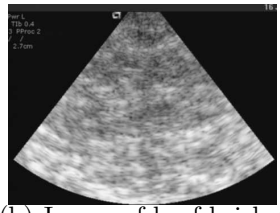
Modelling the spatial distribution of ultrasound scatterers in 3D

The previous chapter showed how the correlation between two image patches could be used to estimate the elevational displacement between them, and how many redundant such distance estimates could be combined to estimate the out-of-plane trajectory of an ultrasound transducer from speckle phantom imagery. In validating this approach, synthetic ultrasound imagery was used profusely. The synthetic validation platform allowed trajectory recovery algorithms to be tested for arbitrarily constrained but perfectly known transducer trajectories, which was not possible using the 1 degree of freedom sub-millimetre positioning device or the noisy data from the Polaris position tracker. The rest of this thesis focuses on the estimation of elevational displacement and the recovery of transducer trajectories from ultrasound imagery of real tissue. This is difficult because the micro-structure of real tissue differs considerably from that of speckle phantoms and also varies considerably, leading to the failure of the elevational speckle decorrelation models derived from speckle phantom imagery, as illustrated in Figure 5.1. The figure compares the average elevational decorrelation curves obtained from imagery of a speckle phantom to those obtained from imagery of a sample of beef brisket. The decorrelation curves associated with the phantom could clearly not be used directly to accurately estimate transducer motion from the beef imagery.

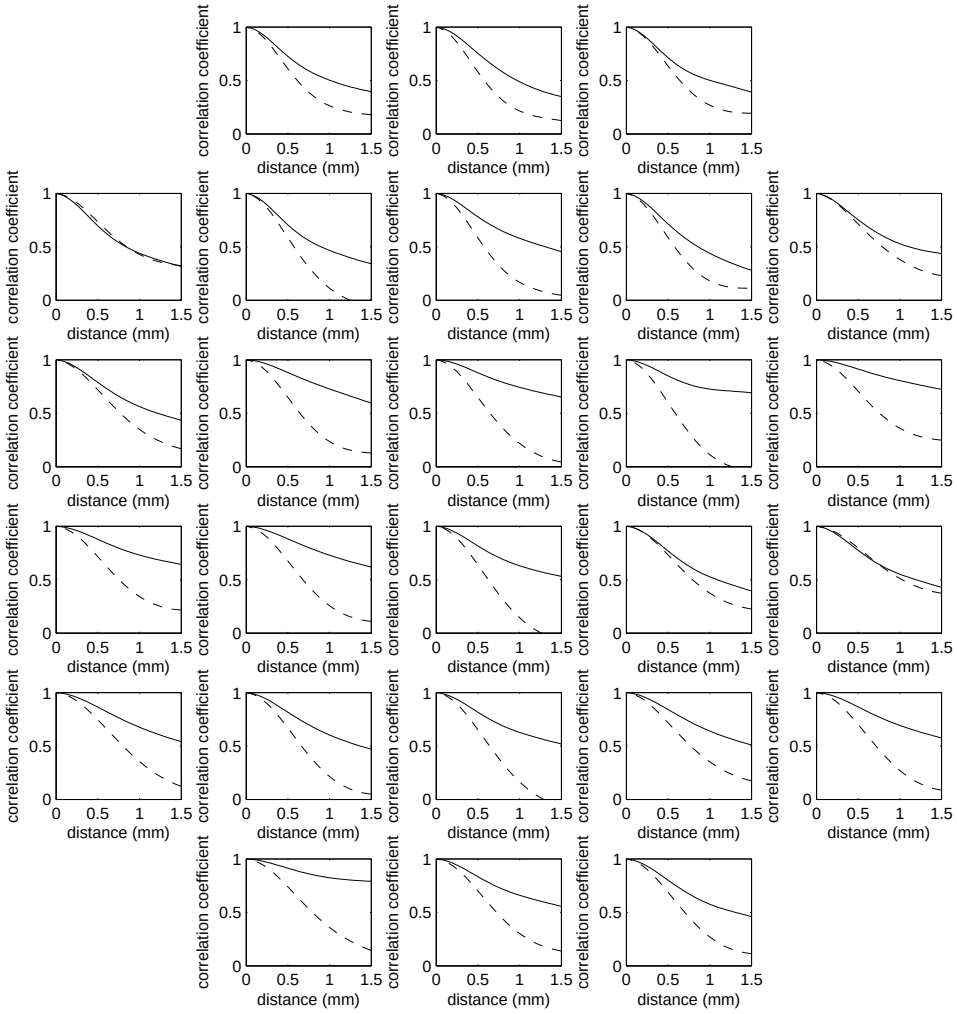
In doing research on the theme of medium invariant speckle decorrelation-based out-of-plane trajectory estimation, synthetic ultrasound imagery has continued to be an invaluable



(a) Speckle phantom image divided into 26 patches for the computation of decorrelation curves



(b) Image of beef brisket



(c) Decorrelation curves for the speckle phantom (dashed lines) and the beef brisket (solid lines)

Figure 5.1 Comparison of the average elevational decorrelation curves obtained from a speckle phantom imagery to those obtained from imagery of beef brisket.

tool, both for validating algorithms and for designing new ones. In this thesis, the synthetic imaging platform allowed fine grained control not only on the parameters of the transducer trajectory, but also on the micro-structure of the virtual medium. This is particulary useful for studying the statistics of ultrasound signals for different hypothetical tissue types, and was used in the development and validation of the tissue-invariant speckle decorreation model of Chapter 6. This chapter describes how the simulation of such varied media was achieved.

The synthetic imagery used in this work was generated using the well-known ultrasound simulator software Field II [110, 108]. Field II takes as input a list of spatial coordinates corresponding to the location of point ultrasound scatterers along with their echogenicity. It is up to the user to specify these parameters in order to reflect the echogenic structure of the types of media to be "scanned", which will in turn determine the texture of the resulting image. The medium structure can be characterised by, among many parameters, the strength of the scatterers, their density and their spatial organisation. In the context of artificial ultrasound imagery used for the validation of image processing algorithms, most studies have focused on the first two of these parameters [109, 134]. Models for the spatial organisation of ultrasound scatterers in multi-dimensional images exist, but have been geared mostly towards the study of specific physiological phenomena, such as the agglomeration of red blood cells [193, 194], rather than towards the development of a general purpose platform for the development and validation of ultrasound image processing algorithms. This context is different because it requires a model which can cover a broader set of spatial configurations and is not subject to special physiological constraints, though such constraints could still be enforced later, as a post-processing step, if needed.

This chapter describes a novel method for generating lists of multi-dimensional scatterer positions with variable strength, density and spatial organisation ranging from tightly clustered to nearly regular in order to mimic a broad range of tissue types. The method extends a previous 1D scatterer distribution model [123, 43] with a simple and intuitive parameterisation by mapping its output to a multi-dimensional space filling curve. The mapping isotropically preserves the characteristics of the original 1D model while providing a fast means of generating multi-dimensional data for the simulations required to validate image processing algorithms such as the transducer motion recovery methods introduced in this thesis. These features are demonstrated by comparing the first and second order statistics of synthetic images produced using the model with theoretical predictions from ultrasound

physics. The majority of the work presented in this chapter was published as [128, 129].

The rest of this chapter is organised as follows. Section 5.1 describes alternative models of spatial organisation in the field of ultrasonics and elsewhere. Section 5.2 presents the chosen adaptation of the 1D model of [123, 43] to multiple dimensions. Section 5.3 describes the pool of synthetic ultrasound data that was generated using this model in conjunction with the Field II ultrasound simulator. The first and second order statistics of this imagery are analysed in Section 5.4 and found to be reliably predictable from the model parameters and also agree with theoretical predictions based on ultrasound physics.

5.1 Other point distribution models

The spatial distribution of scatterers can be thought of abstractly in terms of random point processes. A very simple and ubiquitous point process is one which distributes points randomly in space according to a uniform distribution with a tunable density parameter. Such a process is known in statistics as a Poisson point process. In natural media, the spatial distribution of ultrasound scatterers can exhibit non-random structure, with clustered or periodic arrangements.

In order to generate more structurally varied point patterns, various alternatives and generalisations to the Poisson point process are possible and have been suggested in the literature. For example, clustered data points can be generated in multi-dimensional spaces according to the Neyman-Scott point process [154]. Using this method, an intermediate set of points to be used as cluster centres is generated according to a Poisson process. The final data points are then generated from some other spatial distribution centred about the cluster centres, whose parameters control the extent to which the points are clustered in space. This approach was taken by Savéry and Cloutier to study the characteristics of ultrasound signals in relation to red blood cell aggregation [193].

At the other end of the spatial organisation continuum, the computer graphics literature has been concerned with the generation of random point distributions with high regularity for the creation of random, yet visually pleasing, scene layouts consisting of multiple replicas of the same graphical primitive. In particular, considerable effort has been invested in developing efficient algorithms for generating points according to the so-called Poisson disc distribution [142, 122], which ensures a minimal distance between neighbouring points. Alternately, ultrasound physics research has suggested that quasi-regular data points can

be generated by randomly jittering a lattice of regularly spaced points by varying amounts. This approach was taken by Narayanan et al. [152]. This is similar to the approach taken by Tuthill et al. [220], who used contributions from both a Poisson point process and a regular point lattice to generate scatterer configurations yielding a backscattered ultrasound signal with a mixture of diffuse and coherent components.

None of the approaches described so far is able to model the full continuum of spatial organisations ranging from tightly clustered to nearly regular that might describe the distribution of ultrasound scatterers in different types of tissue. This can be achieved using a Gibbs-Markov pairwise area interaction process. Similar to the Poisson disc model, but more general in scope, this model describes both repulsive and attractive stochastic constraints within pairs of neighbouring points. Current implementations of such models are typically based on random sampling algorithms which begin with a random point distribution (sampled from a Poisson point process) and iteratively perturb the positions of points until the constraints are satisfied. Such an implementation is described by Savéry and Cloutier [194] for the study of the ultrasound signal backscattered by aggregates of red blood cells. In this study, the Gibbs-Markov model was particularly useful because the red blood cells could not be modelled as points scatterers due to the relatively high frequency (and small beam width) of the transducer used. The model was able to account for the fact that such objects with finite spatial extent cannot overlap. This method is computationally demanding and it is difficult to quantitatively assess when and whether it converges towards to the specified model behaviour. The parameters of the model are also much more difficult to tune than the more limited alternatives described earlier.

Interesting alternatives were developed for the field of ecology, where the goal is to model the spatial distribution of plants in an ecosystem [47]. Particularly relevant is the model of Wu et al. [237], a modification of the Poisson point process incorporating an additional parameter which tunes the regularity of points in space such that their spatial organisation varies from clustered to regular. Unfortunately, the parameterisation of the model is somewhat inconsistent in that generating points according to this model involves two different sampling processes, depending on whether the desired effect is clustering or regularity. Wu et al.'s idea of adding a regularity parameter to an ordinary Poisson point process remains intuitively appealing. This idea has led to a flexible and elegantly parameterised 1D point process model, developed independently by Winkelmann the field of economic statistics [236] and by various authors in the field of ultrasonics [123, 43]. The

model, described in more detail in Section 5.2, is a straightforward extension of the 1D Poisson point process whereby the distribution of inter-event interval lengths is modified to reflect the possibility of different spatial organisations. This model embodies most of the characteristics desired for the parametric description of an ultrasound point scatterer spatial distribution model. Namely,

- It can model spatial organisations ranging from tightly clustered to nearly regular;
- Its parameterisation is simple and intuitive;
- It can be sampled efficiently.

The main caveat of this model is that it is one-dimensional and does not have a straightforward analytical extension to the multi-dimensional case of interest for ultrasound image processing. A contribution of this thesis is a simple, yet effective, heuristic extension of this model to multiple dimensions, allowing simulation of 2D and 3D ultrasound data.

5.2 Multidimensional scatterer distribution model

Consider a Poisson point process with rate $\frac{1}{\beta}$. One can mathematically define this random process in terms of statistics on the number of points (or, eventually, scatterers), N_s, counted in non-overlapping areas of unit size. These statistics are Poisson distributed with probability mass function

$$P(N_s) = \frac{\exp(-1/\beta)}{N_s!\beta^{N_s}}, \beta > 0, N_s \geq 0, \tag{5.1}$$

so that there are on average $\frac{1}{\beta}$ points per unit space.

In 1D, there exists an equivalent mathematical definition of the Poisson point process in terms of statistics on the amount of space, d, between two consecutive points, often referred to as the waiting time. These statistics are exponentially distributed [161], with probability density function:

$$p(d) = \frac{\exp(-d/\beta)}{\beta}, \beta, d > 0. \tag{5.2}$$

5.2.1 Tuning regularity

The approach taken in [236, 123, 43] is to generalise (5.2) by making the waiting time d Gamma distributed instead of exponentially distributed.[1] This leads to

$$p(d) = \frac{d^{\alpha-1}\exp(-d/\beta)}{\Gamma(\alpha)\beta^{\alpha}}, \alpha, \beta, d > 0, \tag{5.3}$$

where α is often referred to as a "shape parameter" and Γ denotes the standard Gamma function such that $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\exp(-t)dt$. This distribution can be reparameterised in terms of the mean and variance in waiting time:

$$E\{d\} = \alpha\beta, \tag{5.4}$$

$$VAR\{d\} = \alpha\beta^2 = \frac{E\{d\}^2}{\alpha}, \tag{5.5}$$

where $E\{.\}$ and $VAR\{.\}$ respectively denote the expected value and variance operators. The parameter α may then be viewed as a regularity parameter which, given a certain point density, controlled by $\gamma = \frac{1}{E\{d\}}$, tunes the variance in point spacing [43]. For $\alpha = 1$, (5.3) reduces to (5.2), and the points are randomly distributed in space, according to a Poisson point process. For $\alpha < 1$, the point spacing has high variance and the points tend to group in clusters, whereas for $\alpha > 1$, the point spacing has low variance, yielding more regularly spaced points, as illustrated in Figure 5.2. Should a minimal spacing be required between consecutive points, it is straightforward to include it as a constant term added to the right hand side of (5.3), as proposed in [43].

While no closed form for $p(N_s)$ analogous to (5.1) exists for this generalised Poisson point process, the model described by (5.3) does lead to an asymptotic equivalent in terms of statistics on the point count N_s within an interval of fixed length. Indeed, it has been shown [236] that as the interval length grows toward infinity,

$$\frac{E\{N_s\}}{VAR\{N_s\}} \to \alpha. \tag{5.6}$$

This equivalence will be used to empirically demonstrate the validity of the proposed multi-dimensional extension to the model in the next section.

[1] A similar generalisation based on Weibull distributed waiting times was recently proposed [143].

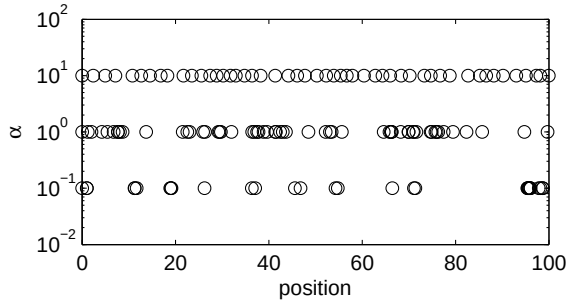


Figure 5.2 Realisations of the 1D generalised Poisson point process with Gamma waiting times. The three realisations were generated using the same point density ($\gamma = 1$) but three different values of the regularity α. Points are randomly placed for $\alpha = 1$, more regularly spaced for $\alpha = 10$, and tightly clustered for $\alpha = 0.1$. In the latter case, the apparently smaller number of points is due to the strong overlap of multiple tightly clustered points, which cannot be visually discriminated from each other at the chosen visualisation scale.

5.2.2 Fractal extension to multiple dimensions

The generalised Poisson point process described above is intrinsically one-dimensional. Therefore, sampling this model becomes difficult in two or more dimensions, except in the case $\alpha = 1$ (Poisson), which amounts to sampling points randomly (according to the uniform distribution) over space. A simple heuristic to generate 2D points with a different spatial organisation would be to sample points according to the 1D generalised Poisson process model along several short segments mapped to a 2D grid of arbitrary precision in a raster scan fashion. This mapping is a poor reflection of what is really desired, as it only preserves the spatial organisation of the original 1D points along one direction. The raster scan mapping is but one of many possible continuous mappings of a line onto a finite multi-dimensional space. Such mappings, which have the property of traversing every cell of a discrete m-dimensional grid of a given precision, are known as discrete space-filling curves. More isotropic preservation of spatial organisation can be achieved by using a different kind of space filling curve. Ideally, distances measured along the original line should correlate well with distances measured in the multi-dimensional space, a property known as *preservation of locality*. The Hilbert curve (illustrated in Figure 5.3) is a recursively defined fractal curve which exhibits excellent preservation of locality and is therefore a good choice [226, 81].

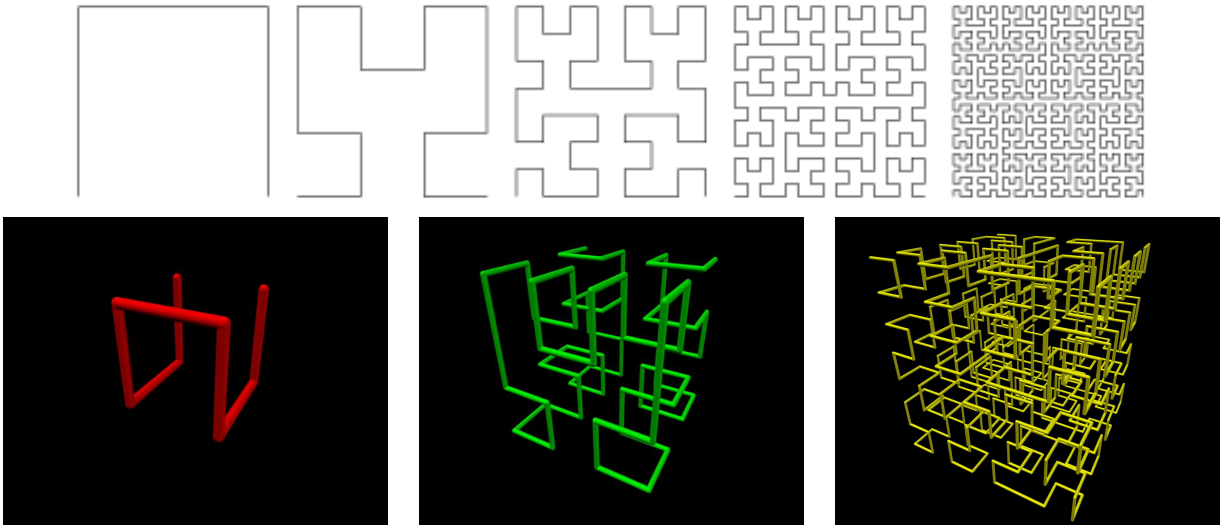


Figure 5.3 Top: discrete approximations of the 2D Hilbert curve [232], with the level of precision k_H of the curves increasing from 1 (left) to 5 (left). Bottom: Discrete approximations of the 3D Hilbert curve [1], with the level of precision k_H of the curves increasing from 1 (left) to 3 (right).

The algorithm introduced in [15] can be used to determine the mapping of 1D points onto a Hilbert curve of precision k_H. This algorithm takes as input a mk_H bit fixed point representation of the position of a 1D point along a line segment. It then iteratively removes m bits off the left of this number and uses them as an index to one of 2^m equally sized sub-cells within the current cell of the m-dimensional space. The algorithm outputs m coordinates in binary fixed point representation with a k_H bit precision. In 2D or 3D, the process is equivalent to decoding the index of cells in a quadtree or octree, respectively. The amount of time required to map a point from 1D to m-D is proportional to the precision k_H.

As a result of applying a Hilbert curve mapping to the output of the 1D point process, a set of 2D or 3D points is obtained whose spatial organisation displays characteristics analogous to that of the original 1D points. A dimensionality independent mathematical description of the desired spatial characteristics is given by (5.6). Thus, in order to verify the ability of the Hilbert curve mapping to preserve spatial organisation, the following steps were taken:

1. Different realisations of the 1D point process (with varying density and regularity parameters) over large intervals were generated.

2. These point process realisations were mapped to 2D and 3D spaces of equivalent size using the Hilbert curve.

3. The 1D, 2D and 3D spaces were divided into an equal number of equally sized non-overlapping bins.

4. The ratio of mean to variance in point count over these bins was calculated and compared to the value of the regularity parameter α used to generate the points.

Predictably, the agreement between the empirical data and (5.6) was excellent in the 1D case. More interestingly, in 2D and 3D, this agreement was also very good for bin sizes corresponding to full sub-cells of the Hilbert curve (which themselves are lower precision Hilbert curves). The agreement between the theory and the data is illustrated in Figure 5.4, where the statistics were collected over a 262,144 unit sized space divided into 64 equally sized bins. Similar results hold irrespective of point density, demonstrating that the proposed multi-dimensional generalisation of the 1D point process preserves its asymptotic statistical properties.

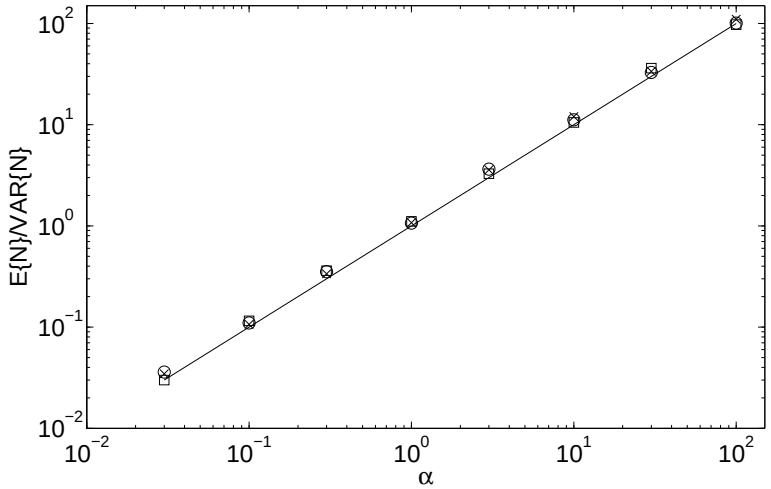


Figure 5.4 Ratio of mean to variance in point count versus regularity parameter α. Solid line: theory, squares: 1D model samples, circles: 2D model samples, crosses: 3D model samples.

The proposed multi-dimensional generalisation approach also preserves the qualitative characteristics of the 1D point process realisations, as shown in Figure 5.5, which shows

example 2D point distributions obtained with the proposed model, compared to those obtained with other point process models used in ultrasonics. Figure 5.5(a) shows the output of a Neyman-Scott point process [193] whose clusters have a Gaussian spatial distribution with variance ε_{NS}^2 controlling their tightness. In these examples, the density of the cluster centres was taken to be the square root of the point density. Figure 5.5(b) shows examples sampled from a jittered lattice model [152], where the lattice points are perturbed by Gaussian noise with variance ε_{JL}^2. Figures 5.5(c) and 5.5(d) show examples from two different Gibbs-Markov pairwise interaction processes. Figure 5.5(c) was obtained using a model very similar to that of Savéry and Cloutier [194] to model aggregates of non-overlapping blood cells: a fixed, strong repulsive potential was used to make pairwise distances below the diameter of a cell virtually impossible, while another more moderate potential was used to draw the scatterers towards a clustered or regular (depending on the sign of the potential) configuration for a range of larger distances within each scatterer's pre-defined zone of influence. The latter moderate potential was proportional to a user-tuned parameter ϵ and to the area of overlap between the influence zones of the two scatterers involved in the relationship. Note how, for clustered configurations, the strong short-range repulsive potential introduces regular structure within the clusters. This same repulsive potential also prevents the formation of truly random point configurations. Figure 5.5(d) shows what happens when the short-range repulsive potential is set to zero: for negative (attractive) long-range potentials, the point configuration abruptly goes from random to extremely clustered, limiting the diversity of available clustered patterns. This shortcoming is well known and documented in the statistics literature [9]. Figure 5.5(e) shows example 2D point distributions obtained using the proposed multi-dimensional generalised Poisson process model for a fixed point density and different values of α. In the Poisson case ($\alpha = 1$), the obtained configuration of points exhibits no particular structure: clusters of points occur by chance and so do areas where points are more uniformly spaced. As α decreases below 1, the 2D point configuration contains more and tighter clusters, following the predictions made by the original 1D model [236, 123, 43]. Similarly, as α grows greater than 1, there are fewer clusters and points more evenly cover the sampling area. Table 5.1 summarises the features of the different models compared in Figure 5.5. Compared to the other models, the proposed generalised multi-dimensional Poisson point process has few parameters (thereby making it easy to tune) and covers a broad range of spatial organisations. It can also be sampled reasonably quickly: the Hilbert curve mapping process is much faster than

the iterative sampling procedure required for Gibbs-Markov models. It is therefore a good candidate for the large scale generation of virtual phantoms representing different media in the context of ultrasound simulation.

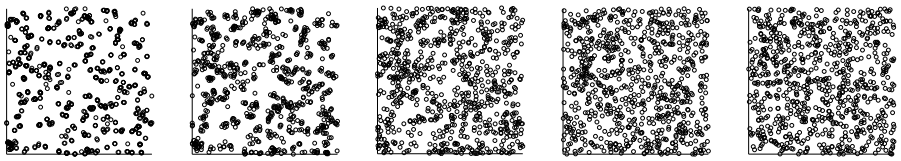
Model	# parameters	Clusters	Regularity	Sampling speed
Poisson	1	No	No	Fast
Neyman-Scott	3	Yes	No	Fast
Jittered lattice	2	No	Yes	Fast
Gibbs-Markov	5	Yes	Yes	Slow
Generalised Poisson	2	Yes	Yes	Moderate

Table 5.1 Comparative summary of different point process models used in ultrasonic simulations.

5.3 Synthetic ultrasound data sets

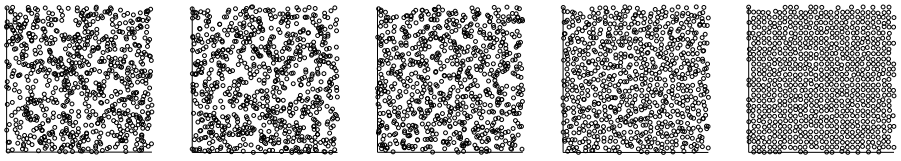
This section describes how the synthetic ultrasound data sets needed for the experiments of Section 5.4 were created. The method described in Section 5.2 was used to generate lists of 3D scatterer positions for all combinations of six scatterer density values ranging from 0.5 to 16 scatterers/mm^3 and ten regularity parameter values ranging from 0.01 to 300, with logarithmic increments, as suggested in [43]. The precision of the 3D Hilbert curve used was 32 bits. The scattering strength of each scatterer was chosen according to a log-normal distribution, as suggested by [43], with mean standard deviation parameters equal to 0.1 and 0.3, respectively.[2] Generation of synthetic ultrasound data from these lists of scatterers was carried out using the realistic ultrasound simulator Field II [110, 108] using a virtual linear array transducer with a central frequency of 3.5 MHz, a depth of 6 cm, a focus at 3 cm and a sampling frequency of 50 MHz. For each combination of scatterer density and regularity parameters, a 3D ultrasound image was obtained by generating 100 2D image slices separated by 0.1 mm increments and consisting of 128 radio-frequency (RF) vectors covering a width of 4 cm. The scatterers occupied a cubic volume of $6 \times 6 \times 6$ cm^3 2 mm below the virtual transducer's surface.

[2] By abuse of terminology, this thesis takes the "mean" and "variance"/"standard deviation" parameters of the log-normal distribution to refer to the mean and variance/standard deviation of the logarithm of the random variable under consideration.



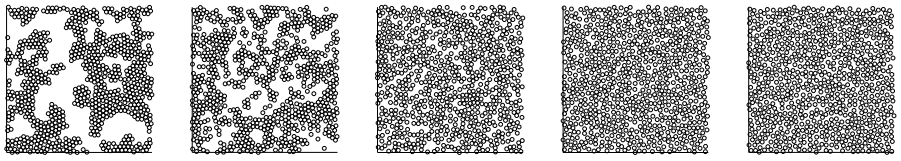
$\varepsilon_{NS} = 0.03$ $\varepsilon_{NS} = 0.1$ $\varepsilon_{NS} = 0.3$ $\varepsilon_{NS} = 1$ $\varepsilon_{NS} = 3$

(a) Neyman-Scott process with varying cluster width ε_{NS}



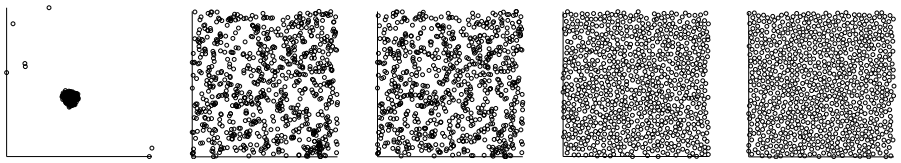
$\varepsilon_{JL} = 3$ $\varepsilon_{JL} = 1$ $\varepsilon_{JL} = 0.3$ $\varepsilon_{JL} = 0.1$ $\varepsilon_{JL} = 0.03$

(b) Regular lattice with varying jitter noise ε_{JL}



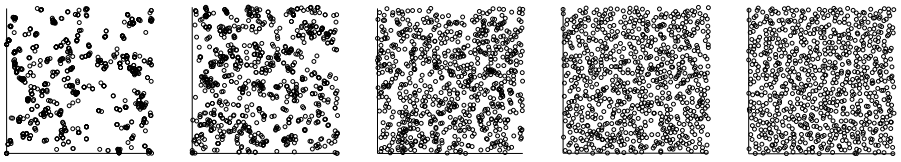
$\epsilon = -30$ $\epsilon = -10$ $\epsilon = 0$ $\epsilon = 10$ $\epsilon = 30$

(c) Gibbs-Markov process with short-range repulsion and varying long-range potential ϵ



$\epsilon = -0.28$ $\epsilon = -0.25$ $\epsilon = 0$ $\epsilon = 10$ $\epsilon = 30$

(d) Gibbs-Markov process with no short-range repulsion and varying long-range potential ϵ.



$\alpha = 0.1$ $\alpha = 0.3$ $\alpha = 1$ $\alpha = 3$ $\alpha = 10$

(e) Proposed generalised Poisson process with varying regularity parameter α. The parameter β was adjusted according to the fixed density $\gamma = \frac{1}{E\{d\}}$ and desired α using (5.4) .

Figure 5.5 Example 2D point distributions obtained using different models with fixed point density. The apparently smaller number of points for some of the more clustered configurations is due to the overlap between multiple tightly clustered points at the visualisation scale presented.

While the generation of scatterer lists is quick, the generation of synthetic 3D ultrasound scans using Field II is not. Thus, the simulations were carried out on parallel processing clusters, thereby exploiting the parallelism-friendly characteristics of Field II. Other less time consuming simulation methods could have been used instead. For example, the preliminary version of this work [128] used a simplistic linear shift invariant model, and recent work by Dillenseger et al. [48] also exploits the current framework within a shift-variant adaptation of Bamber and Dickinson's classical simulation model [13]. The tissue model is completely independent of the type of ultrasound simulator used. The synthetic RF data were envelope detected using the Hilbert transform.

Sample slices from the 3D synthetic data sets, log-compressed for visualisation, are shown in Fig. 5.6, illustrating how the image texture varies from a very grainy to a relatively smooth aspect as scatterer density and regularity parameters increase. It is straightforward to combine different types of micro-structure in the same data set to simulate a non-homogeneous medium such as an area containing different types of tissue. For example, the image in Fig. 5.7 was taken from a 3D data set generated from 20 different combinations of scatterer density, regularity and strength distributions. For this purpose, the volume was segmented into 20 regions by selecting 20 3D points at random and dividing the space according to the corresponding Voronoi diagram. More anatomically realistic arrangements can be achieved using real segmented magnetic resonance or computed tomography images as a starting point, as recently proposed by Dillenseger et al. [48].

5.4 Speckle statistics of synthetic data

In order to be useful in the context of ultrasound image simulation, the proposed scatterer distribution model must be able, in combination with some process which models the propagation of ultrasound waves, to produce speckle images reflective of a broad variety of media. It is also essential that given transducer parameters, the statistics of these images be reliably predictable from the scatterer distribution model parameters and agree with theoretical considerations drawn from previous ultrasound imaging research. This section examines the first and second order statistics of the synthetic imagery obtained using the proposed model.

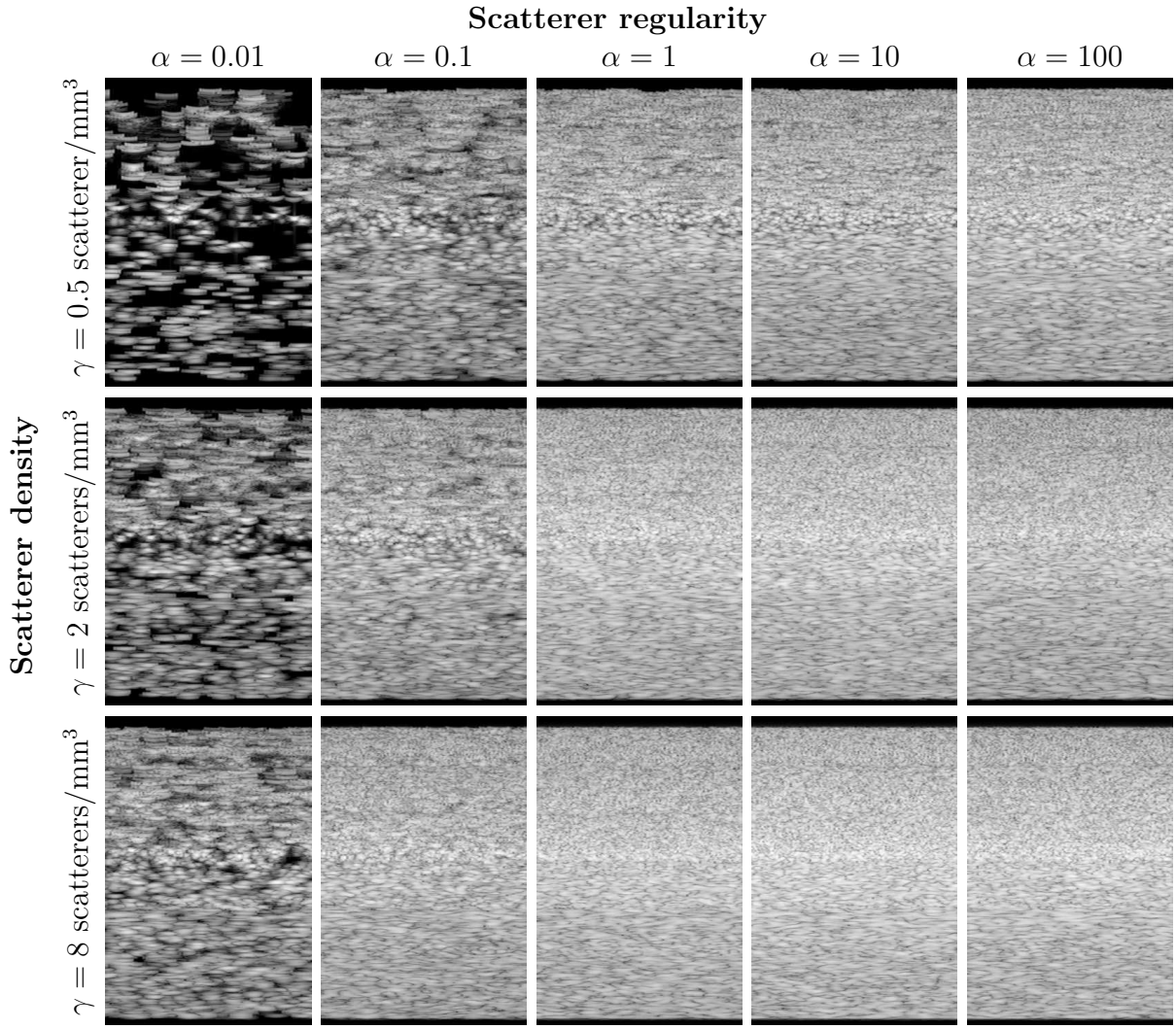


Figure 5.6 Sample synthetic ultrasound images obtained for different combinations of density and regularity.

5.4.1 First order statistics

There is an abundance of parametric models for the first order statistics of ultrasound speckle [229, 102, 52, 195, 196, 197, 57, 112, 19]. The different models account for different sets of assumptions regarding the micro-structure of the scanned object and the imaging process. Older models focused on modelling the local statistics of an ultrasound image patch under different scattering conditions, leading to the well known Rayleigh, K and Rice

Figure 5.7 A slice through a 3D ultrasound data set containing 30 different tissue types.

distribution models [52] briefly discussed in Chapter 2. Recent models (in particular, so-called compound distribution models [197, 57, 112]) try to model the statistics of an entire ultrasound image while taking into account large-scale factors like variations in scattering conditions, signal energy and beam shape which create more or less continuous variations in local speckle statistics across the image.

The present analysis is concerned with two important local speckle statistics, namely the squared signal to noise ratio R^2 and the skewness S of the echo intensity signal, defined as

$$R^2 = \frac{E\{I\}^2}{VAR\{I\}} \tag{5.7}$$

and

$$S = \frac{E\{(I - E\{I\})^3\}}{VAR\{I\}^{\frac{3}{2}}}, \tag{5.8}$$

where I is the echo intensity signal equal to the square of the envelope. These statistics, or transformations thereof, appear as meaningful parameters in many of the probability distributions commonly used to describe local speckle amplitude statistics, including the Rayleigh, K, Rician and homodyned K distributions [52] as well as the Nakagami distribution [196]. They will also be used in the next chapter for tissue characterisation purposes. Theoretically, for Rayleigh scattering (high density and randomly placed scatterers), $R = 1$ and $S = 2$. Lower density or tighter clustering of scatterers lead to lower values for R and higher values for S whereas the opposite may hold true when the scatterers are periodically

placed along the direction of ultrasound wave propagation.

In order to verify whether these predictions were met when using the proposed scatterer distribution model realisations, the synthetic ultrasound data sets described in Section 5.3 were treated as collections of 100 2D image frames each. Each frame was divided into 8×8 equally sized non-overlapping windows measuring 14 A-lines laterally and 490 samples axially. For each of these windows, R^2 and S were estimated.

Fig. 5.8 shows how R^2 and S vary according to scatterer density, regularity and axial depth. As the scattering conditions approach those required for fully developed speckle formation, the value of R^2 approaches 1 and S approaches 2, as predicted by theory. For a given depth and a random arrangement of the scatterers ($\alpha = 1$), these conditions are obtained most closely for high scatterer densities (i.e. closer to 16 scatterers/mm^3). Clustering (low values of α) leads to so-called *sub-Rayleigh* conditions, with $R^2 < 1$ and $S > 2$, with low scatterer densities enhancing the effects of clustering. Also note the changes in R^2 and S with depth; these are especially noticeable for low scatterer densities and small values of α. Near the focus (fourth row), the local resolution cell size is at its smallest, thereby reducing the number of scatterers per resolution cell and leading to sub-Rayleigh conditions.

High values of α, corresponding to more regularly placed scatterers, do not generally appear to lead to conditions very different from Rayleigh scattering, with $R \approx 1$ and $S \approx 2$. This result is different from those presented in [43], where, for certain scatterer densities, the original 1D model with high regularity led to systematic positive interference patterns and Rician statistics ($R > 1$ and $S < 2$). Such statistics were sometimes observed with the model proposed here; indeed, the bottom row of Figure 5.8 shows that for the bottom row of the images and $\alpha > 1$, the average value of R^2 was slightly greater than 1 and the average value of S was slightly below 2. This suggests that positive interference patterns did occur, but not in a systematic fashion. The lack of systematic positive interference effects in the multi-dimensional model is due to the isotropy of the Hilbert curve mapping which was chosen. While the model preserves regularity, there is no guarantee that this regularity will be aligned with a particular direction. That being said, some of the 2D results in [128] show that this *may* occur. For positive interference effects to invariably take place, the scatterers must be placed quasi-periodically *in the direction of wave propagation* in a way that favours constructive interference, which requires a certain degree of anisotropy in the model. Such anisotropy may be achievable through the choice of a different space filling

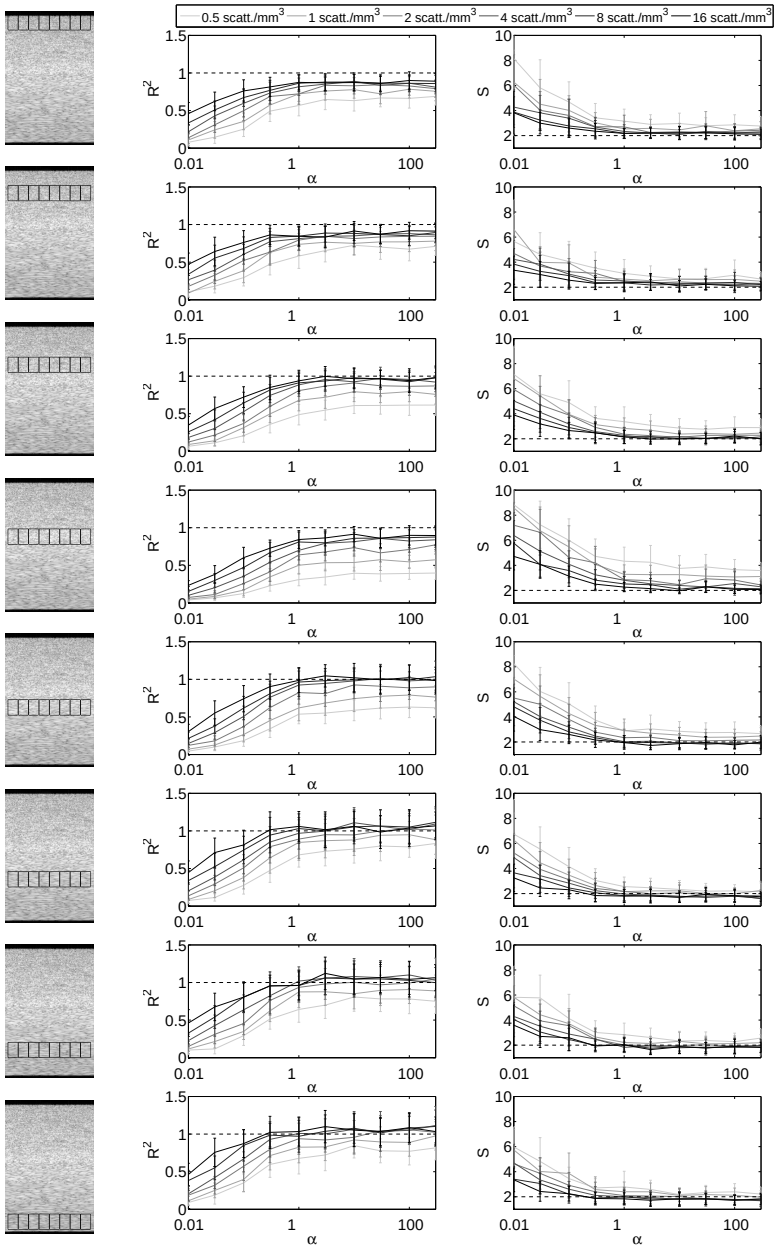


Figure 5.8 Backscattered echo intensity statistics as functions of scatterer density, regularity α and axial depth. Left column: Depth within the ultrasound image. Middle column: squared signal to noise ratio R^2. Right column: skewness S. In these plots, the different shades of grey convey information about scatterer density (see the legend at the top of the figure). The dashed lines in the plots mark the $R^2 = 1$ and $S = 2$ case representative of Rayleigh scattering conditions.

curve to map the 1D point process to 3D. The Hilbert curve mapping nonetheless provides a useful and efficient multi-dimensional extension to the original 1D point process on the basis of which a broad range of synthetic ultrasound image textures can be generated.

5.4.2 Second order statistics

The second order statistics of ultrasound imagery are of paramount relevance to the research presented in the other chapters of this thesis, as elevational image correlation is used as a cue for the estimation of out-of-plane transducer motion. As described in Chapter 2, under Rayleigh scattering conditions, the relationship between image correlation and the transducer's elevational displacement depends entirely on the transducer characteristics and can be described in a simple parametric form for the focal region of the transducer [229], or estimated from imagery of a speckle phantom. Fig. 5.9 shows the average correlation coefficient between pairs of image windows centred near the transducer focus as a function of their elevational separation for selected synthetic data sets. These are compared to the autocorrelation function for a data set consisting of fully developed speckle obtained by placing scatterers at high density according to a plain Poisson process (without the Hilbert curve mapping). Note how the elevational second order statistics of the synthetic data sets of Section 5.3 can be reliably predicted from the parameters of the proposed scatterer distribution model. It was not difficult to find conditions for which the Rayleigh case decorrelation model fails to represent the behaviour of the data, reflecting the results of [86, 72] and the illustrative example shown in Figure 5.1 on imagery of real tissue. A broader range of decorrelation curve shapes can be achieved with virtual phantoms made of a combination of different scatterer distributions, such as the one illustrated in Figure 5.7.

5.5 Summary

This chapter presented a scatterer distribution model for which the strength of individual scatterers as well as their density and spatial organisation can be tuned for the generation of a broad variety of synthetic ultrasound imagery. Spatial organisation of the scatterers can be tuned along the continuum from clustered to regular in an intuitive manner, making the model suitable for the cross-tissue validation of image-based transducer motion estimation. The synthetic platform provides fine grained control over ground truth and target parameters which is at best difficult to achieve in phantom studies, thereby

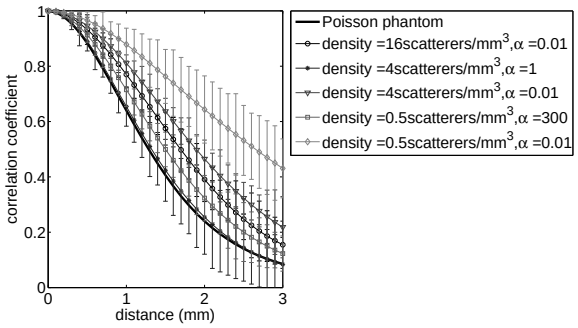


Figure 5.9 Average elevational decorrelation curves at the transducer focus for a variety of combinations of scatterer density and regularity parameters.

providing a useful complementary validation tool. This tool is applicable to a variety of other image processing contexts, a potentially very relevant one being that of texture based image segmentation. It was shown that the model is able to produce imagery with varied first and second order statistics. In particular, it is able to create imagery with elevational decorrelation curves which differ substantially from those obtained under Rayleigh scattering conditions. With this in mind, the framework can also be used to generate training data for automated learning of new ultrasound image processing algorithms which adapt to local tissue structure. This will be illustrated in the next chapter, which describes how a new tissue-invariant speckle decorrelation model was estimated from a pool of synthetic imagery.

Chapter 6

Learning a tissue invariant speckle decorrelation model

Chapter 4 showed how an empirical elevational speckle decorrelation model determined from speckle phantom imagery could be used to recover the elevational motion of a given ultrasound transducer from similar phantom imagery. This worked reasonably well because all scans were acquired under similar scattering conditions, i.e. the phantom was composed of a large number of sub-wavelength scatterers more or less randomly distributed in space, with scattering strengths uncorrelated to spatial location, approximating Rayleigh scattering conditions. Considering the ultrasound imaging system to be a linear system, the autocorrelation function of the output is the convolution of the autocorrelation function associated with the transducer and that associated with the target, as discussed in Chapter 2. Because of the random distribution and large number of scatterers, the autocorrelation function of a speckle phantom is essentially an impulse function, so that the autocorrelation of the output RF ultrasound signal is essentially the transducer's autocorrelation function. In real tissue, scatterers take on spatial configurations that differ from those encountered in a speckle phantom; Chapter 5 showed how to model a variety of such spatial organisations. There are also structured variations in scatterer strength which cause brightness differences in the image, as well as scatterers larger than transducer wavelength, leading to specular reflections. In such conditions, ultrasound signals decorrelate at a slower rate over distance and the transducer specific speckle decorrelation model empirically obtained from the reference phantom fails, as illustrated in Figure 5.1 of the previous chapter.

This chapter presents a novel approach for establishing the local elevational decorrelation rate in imagery of real tissue using the local first and second order statistics of the imagery and a model learned from synthetic ultrasound imagery. This synthetic imagery is first generated from a set of virtual phantoms with varied characteristics using the methods described in Chapter 5 with known elevational trajectories. Sparse Gaussian process regression is then used to establish a relationship between the in-plane first and second order statistics of these images and their elevational correlation lengths. The result is an adaptive speckle decorrelation model which can be applied to new imagery of arbitrary media, synthetic or real. The approach and preliminary experimental results were published as [127].

Section 6.1 discusses other approaches to this problem in the literature. Section 6.2 presents a simplified parametric speckle decorrelation model that is used in this chapter to ease learning, along with a description of the statistical image features used as predictors of elevational correlation length. Section 6.3 then describes the sparse Gaussian process regression method that was chosen for learning the tissue invariant speckle decorrelation model from image features. Section 6.4 describes the experimental methodology adopted in this chapter. Namely, it explains how a training set and test sets of synthetic imagery were generated, how real test imagery of animal tissue samples was acquired and how alternative approaches were implemented for experimental comparison. Experimental results are presented and analysed in Section 6.5, showing the proposed approach to allow accurate distance estimation between image frame pairs. Unexplored ideas which emerged from those results are discussed in Section 6.6.

6.1 Previous work

The literature proposes two main families of approaches for estimating the out-of-plane transducer motion between correlated ultrasound images of real tissue without a position tracker. The first approach is to detect regions of fully developed speckle in the imagery and to obtain distance estimates for those regions exclusively by directly invoking the nominal speckle decorrelation model obtained during calibration. Such speckle detection techniques are discussed in Section 6.1.1. The second approach is to adapt the nominal correlation model to the type of tissue scanned based on the local characteristics of the imagery. Image-based tissue characterisation approaches are discussed in Section 6.1.2. The learning based

approach discussed in this chapter belongs to the latter category.

6.1.1 Speckle detection approaches

Most work on image-based out-of-plane motion estimation assumes that all processed ultrasound image data consist of fully developed speckle. This allows the nominal correlation model (theoretical or empirical) to be used directly for the estimation of distances, but requires a method for reliably identifying regions of fully developed speckle in an image. Many such methods have been developed in the last three decades, often for the purpose of visual image enhancement (speckle removal). For brevity, and because most of these techniques operate on similar principles, only those techniques used in the context of decorrelation based out-of-plane motion estimation will be discussed here.

As discussed in Chapter 2, fully developed speckle is a texture pattern which arises in very narrowly defined theoretical conditions and whose first and second order statistics are simple and very well understood [229]. Namely, the RF signal has a Gaussian probability density function; the envelope of the RF signal is therefore Rayleigh distributed, making its square, the echo intensity signal, exponentially distributed. The exact derivation of the second order statistics (the autocorrelation functions in all three dimensions) of fully developed speckle is substantially more involved and is both transducer and location dependent, but, as demonstrated in this thesis and elsewhere, can be estimated empirically from scans of a speckle phantom. This holds true in the axial and lateral directions as well as in the elevational direction. All the methods described in this section rely on first or second order statistics of ultrasound imagery for the identification of fully developed speckle patterns in imagery.

Tuthill et al. [219] were the first to present experimental image-based motion estimation results on imagery of real tissue. Their approach to speckle detection is based on the fact that the echo intensity signal in fully developed speckle patterns is exponentially distributed [229] and therefore has a theoretically constant mean to standard deviation ratio value of 1. Their paper suggests a simple filter which selects image pixels in whose neighbourhood the ratio of echo intensity signal mean to signal standard deviation is between 0.8 and 1.2. Using this filter in imagery of breast tissue, the authors found that only 4.5% of the data were retained for processing. They also noted that there is an important trade-off between the ability to reliably reject non-speckle data and the availability of enough

image samples for distance estimation using correlation. Similar conclusions were reached by Chang et al. [29].

Georgiou and Cohen [75] proposed a speckle detector based on the Gaussianity of raw RF ultrasound signals. The detector measures how well the Gaussian assumption fits the data in a given region using a non-parametric Kolmogorov-Smirnov test, which compares the empirical cumulative distribution function of the data to that of the normal distribution. The statistical test rejects data samples which have strongly non-Gaussian characteristics as not consisting of fully developed speckle and otherwise accepts the null hypothesis of fully developed speckle. This detector was tested in the context of image-based out-of-plane distance estimation by Hassenpflug et al. [86].

In early work, Dutt and Greenleaf [52] introduced a tissue characterisation methodology based on first order statistics (specifically signal-to-noise ratio R and skewness S defined in Chapter 5) of ultrasound imagery. While these statistics could, in theory, be directly used to carry out a tissue characterisation task, such as determining whether an image patch consists of fully developed speckle or not, the authors found that they require a large sample size to be computed reliably using the envelope detected or echo intensity signal [54]. Based on these findings, they investigated the use of fractional powers of the envelope detected signal in order to reduce the sample size required for stable estimation of these first order statistics. This approach was followed upon by Prager et al. [167], with the specific goal of finding an optimal fractional power to which to raise the data for optimal detection of fully developed speckle patterns, for which they developed a simple elliptical discriminant classifier. Prager et al.'s speckle detector was subsequently tested by Hassenpflug et al. [86] and found to be among the most useful for image-based out-of-plane motion estimation. Rivaz et al. [176] improved upon Prager et al.'s speckle detector [167] by using different, jointly optimal, fractional powers of the ultrasound signal for computation of R and S. Prager et al.'s version of the detector will be used as a base-line approach for the experiments presented in this chapter and is therefore described in more mathematical detail in Section 6.4.3.

Hassenpflug et al. [86] compared a number of speckle detection techniques, including the aforementioned [75] and [167], in the context of their use in obtaining accurate correlation based measurements of elevational separation between two image patches. Their study also included a novel speckle detector based on second order statistics. This detector works by measuring the discrepancy between the axial correlation curve of the data and that

obtained for imagery of a speckle phantom at the same depth; fully developed speckle is detected when the two curves match. The different speckle detectors were applied to non-overlapping patches of a pre-defined grid in images of synthetic data, as well as to images of animal tissue. The findings of the study are mixed: effective and reliable rejection of non fully developed speckle imagery is indeed possible with the detectors studied. However, it must also be noted that effective speckle detectors in fact reject so much image content that obtaining the three corresponding image patches required for estimating full out-of-plane transformations (including translation, yaw and tilt) becomes problematic for realistic subjects.

One limitation of the approach described in [86] is that its application of the various speckle detectors under study was limited to non-overlapping image patches with pre-defined locations, corresponding to the tessellation of the phantom imagery used for the calibration of elevational (and axial) speckle decorrelation curves. Using this approach, substantial areas of fully developed speckle whose extent is smaller than the patch size or does not coincide well with the pre-determined grid are easily missed due to surrounding non-speckle imagery. Rivaz et al. [177] propose an approach which allows tissue classification (using their own detector [176]) in irregularly shaped and located image regions, thereby recovering more regions of usable fully developed speckle from the image data. The intuition behind the approach is that while reliably labelling an image region as fully developed speckle may require a large number of pixels, only a small number of pixels are needed to reject image regions which clearly do not match the model for fully developed speckle. This leads to a two-step robust meshing approach. In the first step, small image regions which are clearly not fully developed speckle are detected and rejected. The second step then looks for larger image patches which contain few rejected pixels and are accepted as fully developed speckle by the classifier. The result is a set of irregularly shaped image patches deemed to contain fully developed speckle. These patches may not be aligned with the calibration grid, which means that the spatially discretised elevational speckle decorrelation model must be interpolated to match the location of detected fully developed speckle regions. The authors demonstrate that the approach is able to recover more usable image data than an approach based on a fixed grid and also yields more accurate correlation based distance estimates. A further improvement of the approach is discussed in [178], where the number of available independent pixels for classification is increased by using two ultrasound images with different beam steering angles at each point in the trajectory.

6.1.2 Tissue characterisation approaches

A major disadvantage of the speckle detection approach to correlation-based out-of-plane motion estimation in real tissue is that it discards a lot of potentially informative image data. Given the very small amounts of fully developed speckle in imagery of animal and human tissue [219, 86, 72], this is a rather serious shortcoming. This section discusses an alternative class of methods which aim at inferring the unknown elevational correlation length of ultrasound signals in the local medium from the associated image data.

As described in Chapter 2, a block of RF ultrasound imagery can be taken to be the output of a (locally) shift-invariant linear system, so that its autocorrelation function in a given direction is the convolution of the transducer specific autocorrelation function (the nominal speckle decorrelation curve) with the autocorrelation function of the scattering medium in the same direction. Under certain conditions, Chen et al. [33] show that it is possible to recover the autocorrelation function of the scattering medium from that of the image by a naive deconvolution process. The authors do not mention the applicability of this approach to the out-of-plane motion estimation problem, but it is easy to see how it could be put to use in predicting the elevational decorrelation behaviour of the RF ultrasound signal. Assuming that the structure of the medium is isotropic, its autocorrelation function could be estimated from axial and lateral decorrelation curves and convolved with the nominal elevational speckle decorrelation curve to obtain the local, medium dependent, elevational decorrelation curve. Such an approach would require the full RF ultrasound signal, which is not available to the user for most clinical ultrasound scanners. It is also likely that important adjustments would be needed in order to overcome the numerical instability which notoriously plagues deconvolution algorithms.

Gee et al. [72] were the first to propose and validate a method for predicting the elevational decorrelation behaviour of the ultrasound echo envelope signal in real tissue. In order to cope with departures from Rayleigh scattering, they augment the traditional calibration procedure by computing nominal axial and lateral speckle decorrelation curves from phantom data in addition to the elevational one. In order to adapt elevational distance estimation to images of arbitrary tissue, their method estimates medium dependent elevational decorrelation curves from the image data by correcting the calibration curve according to observed discrepancies between the in-plane (axial and lateral) calibration curves and the ones observed in the imagery. The correction process is based on a heuristi-

cally defined measure of scattering coherence and by the authors' admission, is not tied to ultrasound physics. In spite of this, this adaptive method was shown to be very effective in yielding accurate and dense out-of-plane transducer motion estimates in high-resolution envelope detected (but not scan-converted) RF scans of animal tissue and as such, represents the state-of-the-art. An adaptation of this approach is used as a base-line in this study; mathematical and implementation details are given in Section 6.4.3.

Rivaz et al. [177] put forward the idea that variations in elevational correlation length of the echo envelope could be predicted from first order image statistics. They recognise the potential of an alternative calibration procedure whereby an object with varied internal structure, such as a piece of meat, is scanned instead of a speckle phantom and a set of elevational decorrelation curves is established, parameterised by the local first order statistics (signal to noise ratio and skewness) of the imagery, presumably by establishing some sort of regression model. The idea is neither fully formalised (no regression model is in fact established) nor tested in the context of distance estimation in the paper. The approach presented here is closely related to this proposal and the first concrete embodiment of it, in that controlled (synthetic) scans of varied media are used as basis to establish a regression model between locally observable image features and the elevational decorrelation behaviour of the ultrasound signal.

6.2 Features and simplified decorrelation model representation

This chapter considers the adaptation of a nominal, deterministic elevational speckle decorrelation model to imagery of arbitrary media. It is assumed for the moment that this nominal model is a set of piecewise linear curves (one for each image patch in the tessellation defined by calibration) obtained by finding the average sample correlation coefficient ρ between pairs of image patches separated by distance δ. The nominal piecewise linear elevational decorrelation curve associated with patch q will be referred to as $\hat{\rho}_q(\delta)$. Within the limitations already discussed in Chapter 4, such a curve can be used to obtain reasonably accurate estimates of the distance between corresponding image patches containing fully developed speckle.

Ultrasound imagery of real tissue typically decorrelates over distance at a slower rate than imagery of a speckle phantom. The approach presented in this chapter is to learn from examples how the nominal decorrelation curve stretches along the δ axis depending on the

characteristics of the medium observable in the image. In order to ease this learning process, a more concise speckle decorrelation model will be used. For most transducers, the nominal decorrelation curve of the echo amplitude (or echo intensity) signal is qualitatively well represented by a Gaussian function of δ [31, 72]. Thus, for learning purposes, decorrelation curves will be summarised by the standard deviation $\hat{w}_e$ (henceforth termed the *correlation length*) of the Gaussian curve which best fits (in the least-squares sense) the (ρ, δ) pairs observed during calibration. The local elevational correlation length of the target medium will be referred to as w_e and is computed as the standard deviation of the Gaussian curve which best fits the local elevational decorrelation curve.

Assuming that the micro-structure of the target is not completely anisotropic, statistical changes in the ultrasound signal should be observed within the image plane as well as out-of-plane. This assumption is central to all adaptive speckle decorrelation methods, including [72, 177]. The proposed approach monitors these changes using four features. The first two are the squared signal to noise ratio R^2 and the skewness S of the echo intensity signal I, used in Chapter 5 for image characterisation, and redefined here for convenience:

$$R^2 = \frac{E\{I\}^2}{VAR\{I\}} \tag{6.1}$$

$$S = \frac{E\{(I - E\{I\})^3\}}{VAR\{I\}^{\frac{3}{2}}}. \tag{6.2}$$

The other two features are the Gaussian approximators of the axial and lateral decorrelation curves, w_a and w_l. The axial and lateral decorrelation curves associated with a given image patch are obtained by calculating the average correlation between (potentially overlapping) sub-windows of the image patch offset by known numbers of pixels. The size of the sub-windows was chosen to be two-thirds the size of the original patch in the direction for which the decorrelation curve is computed. The axial and lateral correlation lengths w_a and w_l are found by fitting Gaussians to these curves. Nominal axial and lateral decorrelation curves are also computed for the calibration scan; in this case, the decorrelation curve for a given patch is the average of the decorrelation curves obtained for that patch in each frame.

The features R^2, S^2, w_a and w_l were chosen because they are conceptually simple and relatively easy to calculate, and because these and similar features (e.g. full piecewise

linear descriptions of axial and lateral decorrelation curves) have previously been used in the context of speckle detection and characterisation for speckle decorrelation based distance estimation [54, 167, 176, 177, 72]. Other textural features, such as Haralick's co-occurrence matrices [84, 16, 223, 247], which are also commonly used for ultrasonic tissue characterisation, could have been used instead or in conjunction with the chosen statistics.

The key hypothesis underlying the work presented in this chapter is that a transducer independent statistical relationship between the measurable statistical features of ultrasound images and the target local w_e can be established in a data-driven manner from a pool of varied synthetic ultrasound imagery with even elevational spacing. In such imagery, w_e can be measured locally for each image patch using a set of elevationally neighbouring image patches by fitting a Gaussian curve to the local average correlation values observed for different elevational separations.

In order to enforce transducer independence of the regression model, the features and target are normalised with respect to the equivalent features as measured from the calibration data set. That is, the learning method uses correlation length ratios defined as

$$r_l = \frac{w_l}{\hat{w}_l}, \tag{6.3}$$

$$r_a = \frac{w_a}{\hat{w}_a}, \tag{6.4}$$

$$r_e = \frac{w_e}{\hat{w}_e}. \tag{6.5}$$

It is assumed that changes in correlation length along any direction depend entirely on the structure of the medium so that these ratios do not depend on the characteristics of the transducer. Ratios of R^2 and S with respect to equivalent features $\hat{R}^2$ and $\hat{S}$ measured in imagery of the calibration phantom,

$$r_{R^2} = \frac{R^2}{\hat{R}^2}, \tag{6.6}$$

$$r_S = \frac{S}{\hat{S}}, \tag{6.7}$$

are also used, even though the features $\hat{R}^2$ and $\hat{S}$ should theoretically exhibit no dependence on transducer characteristics in imagery of fully developed speckle. Doing this allows the density of the scatterers in the calibration phantom to be slightly less than that required for

perfect Rayleigh scattering conditions without compromising the accuracy of the method too much.

6.3 Learning with sparse Gaussian process regression

Each image patch in the pool of synthetic imagery provides an *observation*, i.e. an association between a feature vector $\mathbf{f} = \{r_{R^2}, r_S, r_a, r_l\}$ and the local r_e. Given a training set $\mathcal{D}$ containing N such observations, i.e. a set $\mathbf{F} = \{\mathbf{f}_i\}, i = 1 \ldots N$ of feature vectors and the set $\mathbf{r}_e$ of their associated normalised elevational correlation lengths, the purpose of regression is to establish a relationship between feature vectors and r_e so that r_e can subsequently be predicted for new input feature vectors. There are many methods for doing this. A textbook approach is parametric regression, where a parametric form is fitted to the data using a least-squares error criterion. The main difficulty with this approach is the selection of an appropriate parametric model among an infinity of possibilities.

So-called kernel-based regression methods are more promising. In kernel-based regression, a kernel function with semi-finite support (such as a Gaussian function) is chosen and centred about every data point in the training set. The regressor is simply a linearly weighted sum of the resulting ensemble of kernels, with the weights optimised to minimise some error criterion. Machine learning research has shown that a cleverly designed error measure leads to very few non-zero weights and hence very sparse regression models, which are highly desirable in order to avoid over-fitting and to limit the computational expense of making a prediction. Two famous examples are support vector regression, which attempts to minimise the mean squared error associated with data points outside a regularisation margin close to the regressor [82, 205], and relevance vector regression, a Bayesian approach which maximises the probability of the data under the assumption of uninformative prior distributions on the weights [213].

Support vector regression was briefly considered for this work, chiefly because its better known classifier form, the support vector machine, has had successful computer assisted diagnosis applications based on ultrasound tissue characterisation [30, 245, 246]. Relevance vector regression, which has also been used successfully in (non ultrasound) image interpretation tasks [234, 249, 3, 210], was more seriously considered and initially experimented with as a regression model for this work. It has two important advantages over support vector regression: (1) it generally yields sparser regressors and (2) provides probabilistic outputs,

i.e. the output is given as a mean and variance. The latter feature seemed particularly desirable because the correlation length estimates to be derived from the regressor would eventually be used in higher level inference tasks (e.g. trajectory estimation) involving additional data. In this context, uncertainty information can help in correctly weighting evidence from different sources of data. Unfortunately, it turns out that the behaviour of the predictive variances output by relevance vector regression is counter-intuitive: the predictive uncertainty decreases as the input feature distance itself from the training set, whereas common sense dictates that uncertainty should be lowest when the features match a training datum exactly. This caveat is well documented and is an undesirable effect of excessive sparsity [213, 174, 170, 171].

Gaussian process regression is an approach closely related to kernel-based regression. In fact, a traditional Gaussian process regressor can be described as a kernel-based regressor with an infinite set of weights associated with an infinite set of kernel functions placed at every point in observation space. As such, it is a non-parametric method and does not yield a sparse regressor in its traditional form. In the context of elevational correlation length estimation, the aim of Gaussian process regression is to create a mapping between any feature vector $\mathbf{f}_{\text{test}}$ to a Gaussian distribution with mean $\bar{r}_e(\mathbf{f}_{\text{test}}|\mathcal{D})$ and variance $\xi(\mathbf{f}_{\text{test}}|\mathcal{D})$. Let K denote a 4-dimensional Gaussian kernel function parameterised by hyper-parameter vector $\boldsymbol{\theta}$, i.e.

$$K(\mathbf{f}_1, \mathbf{f}_2) = \theta_0 \exp\left(-\frac{1}{2}\sum_{i=1}^{4}\theta_i(f_1^{(i)} - f_2^{(i)})^2\right). \tag{6.8}$$

The mathematics in the rest of this section were reproduced from [207]. Using their notation, let us define

$$[\mathbf{K}_N]_{ii'} = K(\mathbf{f}_i, \mathbf{f}_{i'}) \tag{6.9}$$

$$[\mathbf{k}_{\mathbf{f}}]_i = K(\mathbf{f}_i, \mathbf{f}) \tag{6.10}$$

$$K_{\mathbf{ff}} = K(\mathbf{f}, \mathbf{f}). \tag{6.11}$$

It is assumed that the target values $\mathbf{r}_e$ are corrupted by Gaussian noise of variance given by additional hyper-parameter η^2. In traditional non-parametric Gaussian process regression, the probabilistic prediction for a new input feature vector $\mathbf{f}_{\text{test}}$, obtained from

the entire training set $\mathcal{D}$, takes the form

$$p(r_e|\mathbf{f}_{\text{test}}, \mathcal{D}, \boldsymbol{\theta}, \eta^2) = \mathcal{N}(\bar{r}_e, \xi), \tag{6.12}$$

where $\mathcal{N}(\bar{r}_e, \xi)$ denotes a normal probability density function with mean $\bar{r}_e$ and variance ξ. For Gaussian process regression,

$$\begin{aligned} \bar{r}_e &= \mathbf{k}_{\mathbf{f}_{\text{test}}}^{\top}(\mathbf{K}_N + \eta^2\mathbf{I})^{-1}\mathbf{r}_e & (6.13)\\ \xi &= K_{\mathbf{f}_{\text{test}}\mathbf{f}_{\text{test}}} - \mathbf{k}_{\mathbf{f}_{\text{test}}}^{\top}(\mathbf{K}_N + \eta^2\mathbf{I})^{-1}\mathbf{k}_{\mathbf{f}_{\text{test}}} + \eta^2, & (6.14) \end{aligned}$$

where $\mathbf{I}$ denotes the identity matrix. In this traditional approach, the hyper-parameters $\boldsymbol{\theta}$ and η^2 are estimated by maximising the likelihood of the training data,

$$p(\mathbf{r}_e|\mathbf{f}, \boldsymbol{\theta}, \eta^2) = \mathcal{N}(\mathbf{0}, \mathbf{K}_N + \eta^2\mathbf{I}). \tag{6.15}$$

Because traditional Gaussian process regression accounts for observed correlations between all pairs of training inputs, it requires $O(N^3)$ operations for training and $O(N^2)$ operations for a single mean and variance prediction, and also requires $O(N^3)$ units of storage. These requirements make the method impractical for large data sets, such as the large pool of statistical features that this chapter proposes to derive from synthetic ultrasound imagery. Not surprisingly, researchers in the fields of statistics and machine learning have developed a number of useful approximations which aim at reducing the computational requirements of training and prediction by sparsely parameterising the regressor. Loosely speaking, sparsification is achieved by representing the training set by a smaller, representative set of data points[1]. The manner in which these data points are selected, and the assumptions that are subsequently made regarding the manner in which they can be used to represent the full data set determine the particular flavour of the approximation. An excellent overview and comparative mathematical analysis of such approximation methods is given in [170].

In order to achieve reasonable computation times, this thesis adopts the sparse Gaussian process regression method introduced by Snelson and Ghahramani [207]. In this approach, it is assumed that the training samples included in $\mathcal{D}$ are mutually independent of each

[1] In fact, the relevance vector regression method discussed earlier can be described as such and its shortcomings understood from this description [170].

other given a set of $M << N$ *pseudo-input* feature vectors $\tilde{\mathbf{F}} = \{\tilde{f}_j\}, j = 1 \dots M$ and associated pseudo targets $\tilde{\mathbf{r}}_e$, whose locations are inspired, but not directly drawn from the set of training samples. Let us define

$$
\begin{aligned}
[\mathbf{K}_M]_{jj'} &= K(\tilde{\mathbf{f}}_j, \tilde{\mathbf{f}}_{j'}) && (6.16)\\
[\mathbf{k}_{\mathbf{f}}]_j &= K(\tilde{\mathbf{f}}_j, \mathbf{f}) && (6.17)\\
[\mathbf{K}_{NM}]_{ij} &= K(\mathbf{f}_i, \tilde{\mathbf{f}}_j) && (6.18)\\
[\mathbf{K}_{MN}]_{ji} &= K(\tilde{\mathbf{f}}_j, \mathbf{f}_i) && (6.19)\\
\mathbf{\Lambda} &= \mathrm{diag}(\boldsymbol{\lambda}) && (6.20)\\
\lambda_i &= K_{\mathbf{f}_i\mathbf{f}_i} - \mathbf{k}_{\mathbf{f}_i}^\top \mathbf{K}_M^{-1} \mathbf{k}_{\mathbf{f}_i}. && (6.21)
\end{aligned}
$$

Under the assumption of a Gaussian prior distribution over the pseudo-targets $\tilde{\mathbf{r}}_e$, Snelson and Ghahramani's sparse Gaussian process regressor takes the form

$$
p(r_e|\mathbf{f}_{\text{test}}, \mathcal{D}, \tilde{\mathbf{F}}, \boldsymbol{\theta}, \eta^2) = \mathcal{N}(\bar{r}_e, \xi), \tag{6.22}
$$

where the pseudo-targets do not appear as a result of clever marginalisation and

$$
\begin{aligned}
\bar{r}_e &= \mathbf{k}_{\mathbf{f}_{\text{test}}}^\top \mathbf{Q}_M^{-1} \mathbf{K}_{MN} (\mathbf{\Lambda} + \eta^2 \mathbf{I})^{-1} \mathbf{r}_e, && (6.23)\\
\xi &= K_{\mathbf{f}_{\text{test}}\mathbf{f}_{\text{test}}} - \mathbf{k}_{\mathbf{f}_{\text{test}}}^\top (\mathbf{K}_M^{-1} - \mathbf{Q}_M^{-1}) \mathbf{k}_{\mathbf{f}_{\text{test}}} + \eta^2, && (6.24)
\end{aligned}
$$

with

$$
\mathbf{Q}_M = \mathbf{K}_M + \mathbf{K}_{MN} (\mathbf{\Lambda} + \eta^2 \mathbf{I})^{-1} \mathbf{K}_{NM}. \tag{6.25}
$$

The set of pseudo-inputs $\tilde{\mathbf{F}}$ as well as the hyper-parameters $\boldsymbol{\theta}$ and η^2 are estimated simultaneously by numerically maximising the marginal likelihood of the training data

$$
p(\mathbf{r}_e|\mathbf{F}, \tilde{\mathbf{F}}, \boldsymbol{\theta}, \eta^2) = \mathcal{N}(\mathbf{0}, \mathbf{K}_{NM} \mathbf{K}_M^{-1} \mathbf{K}_{MN} + \mathbf{\Lambda} + \eta^2 \mathbf{I}) \tag{6.26}
$$

with respect to $\tilde{\mathbf{F}}$, $\boldsymbol{\theta}$ and η. The important mathematical manipulations leading to the absence the pseudo-targets from (6.23) and (6.24) and the computations required for the optimisation of (6.26) are discussed in more detail in [207] and [206].

The computational requirements associated with the approach are far more modest than for full Gaussian process regression: $O(M^2 N)$ for training and $O(M^2)$ for prediction.

The authors have shown that this approach finds very good pseudo-input set and hyper-parameter values when the desired size M of the pseudo-input set and the dimensionality of the data are fairly small. These conditions are well met by the application described here. This particular sparsification approach was found to have several advantages over some of its more naive counterparts [170]: it represents the training set well and sparsely without being prone to over-fitting; furthermore, the possibility of selecting pseudo-inputs that are outside the training set gives it the power to capture non-stationary (i.e. input dependent) noise in the training set, which even full Gaussian process regression cannot generally do.

Given two corresponding image patches whose elevational separation must be estimated, each patch i, $i = 1, 2$, provides a feature vector $\mathbf{f}_i = \{r_{R^2{}_i}, r_{S_i}, r_{a_i}, r_{l_i}\}$, which sparse Gaussian process regression associates with an estimate $\bar{r}_{e_i}$ of the scaling factor r_{e_i} (required to compute the local correlation length w_e from the nominal correlation length $\hat{w}_e$ as $w_{e_i} = r_e \hat{w}_e$) and an associated variance ξ_i. A global estimate r_e^* of r_e for the pair of image patches is obtained by variance-weighted averaging, i.e.

$$r_e^* = \frac{\dfrac{\bar{r}_{e_1}}{\xi_1} + \dfrac{\bar{r}_{e_2}}{\xi_2}}{\dfrac{1}{\xi_1} + \dfrac{1}{\xi_2}}, \tag{6.27}$$

where heavier weight is given to values of r_e associated with small variance. A locally adapted piecewise linear decorrelation curve for patch q is obtained by stretching the nominal transducer specific decorrelation model along the δ axis by the inferred ratio r_e^* of local to nominal correlation lengths (see (6.5)) according to

$$\rho_q(r_e^* \delta) = \hat{\rho}_q(\delta), \tag{6.28}$$

or, equivalently,

$$\rho_q(\delta) = \hat{\rho}_q\left(\frac{\delta}{r_e^*}\right). \tag{6.29}$$

6.4 Experiments

The proposed learning-based approach to elevational correlation length estimation was tested by first creating a pool of varied synthetic imagery from which to learn the rela-

tionship between elevational correlation length and image statistics. The sparse Gaussian regressor obtained from this imagery was then applied to new sequences of both synthetic and real ultrasound imagery of varied media, including animal tissue. The resulting accuracy of out-of-plane motion estimates for pairs of images was then evaluated and compared against those obtained using three different base-line approaches. This section describes how these experiments were carried out. Section 6.4.1 explains how the synthetic training imagery was created and how the sparse Gaussian regressor was learned. Section 6.4.2 describes the acquisition of synthetic and real test imagery. Section 6.4.3 describes the base-line approaches and provides the implementation details for Prager et al.'s speckle detector [167] and Gee et al.'s heuristic adaptive method [72].

6.4.1 Training

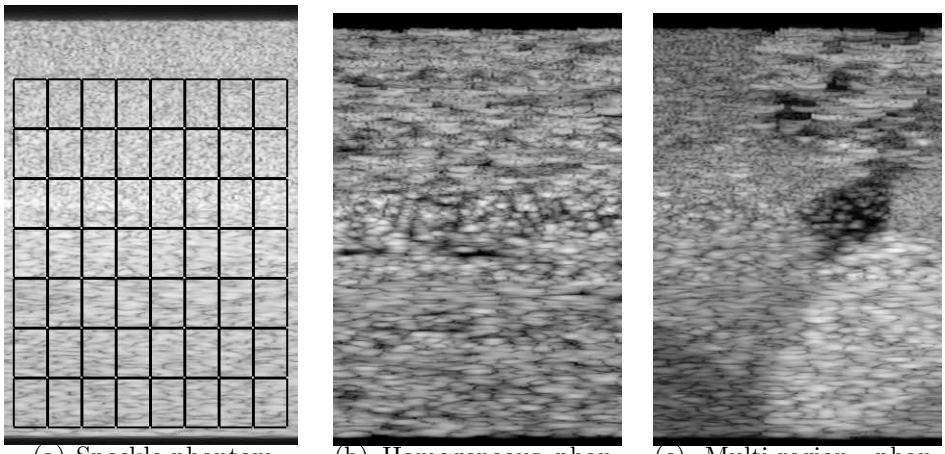
The Field II ultrasound simulation software [108] was used to create a varied set of synthetic training examples of ultrasound data sequences from which to learn a tissue invariant decorrelation model. The synthetic ultrasound data sets were generated with a virtual 3.5 MHz linear transducer with a scanning depth of 6 cm and a focus at 3 cm. The synthetic images were acquired at 0.1 mm elevational intervals. The resulting RF vectors were log-compressed and scan-converted (using a linear interpolator) to mimic the acquisition of data by a clinical ultrasound scanner. The nominal transducer dependent decorrelation curves were estimated from scans of a virtual speckle phantom made of a large number of scatterers distributed randomly according to a Poisson point process model. Additional synthetic ultrasound data sets were generated from 15 virtual phantoms consisting of point scatterers whose echogenicity distribution, density and spatial organisation were tuned differently using the fractal model of Chapter 5. Thirteen of these phantoms had a homogeneous micro-structure, with fixed density and regularity parameters and a log-normal echogenicity distribution. The other two virtual phantoms were multi-region phantoms obtained by generating 20 random seed points in a cube of 60 mm $\times$ 60 mm $\times$ 60 mm and breaking the cube into the corresponding 20 Voronoi regions. The scatterers in each of these regions were then assigned random density, regularity and echogenicity characteristics, as indicated in Table 6.1. The parameters of the simulations were chosen so as to represent a variety of scatterer organisations yielding a broad range of elevational decorrelation lengths. Sample training images are shown in Figure 6.1.

#	Type	Density (scatt./mm^3)	Regularity	Echogenicity	# Frames
1	Homogeneous	$\gamma = 0.5$	$\alpha = 0.01$	$e \sim \ell(0.1, 0.09)$	100
2	Homogeneous	$\gamma = 0.5$	$\alpha = 0.1$	$e \sim \ell(0.1, 0.09)$	100
3	Homogeneous	$\gamma = 0.5$	$\alpha = 10$	$e \sim \ell(0.1, 0.09)$	100
4	Homogeneous	$\gamma = 1$	$\alpha = 0.03$	$e \sim \ell(0.1, 0.09)$	100
5	Homogeneous	$\gamma = 1$	$\alpha = 0.3$	$e \sim \ell(0.1, 0.09)$	100
6	Homogeneous	$\gamma = 2$	$\alpha = 0.01$	$e \sim \ell(0.1, 0.09)$	100
7	Homogeneous	$\gamma = 2$	$\alpha = 0.03$	$e \sim \ell(0.1, 0.09)$	100
8	Homogeneous	$\gamma = 2$	$\alpha = 0.1$	$e \sim \ell(0.1, 0.09)$	100
9	Homogeneous	$\gamma = 2$	$\alpha = 0.3$	$e \sim \ell(0.1, 0.09)$	100
10	Homogeneous	$\gamma = 4$	$\alpha = 0.01$	$e \sim \ell(0.1, 0.09)$	100
11	Homogeneous	$\gamma = 4$	$\alpha = 0.03$	$e \sim \ell(0.1, 0.09)$	100
12	Homogeneous	$\gamma = 4$	$\alpha = 0.1$	$e \sim \ell(0.1, 0.09)$	100
13	Homogeneous	$\gamma = 16$	$\alpha = 100$	$e \sim \ell(0.1, 0.09)$	100
14	Multi-region	$\gamma \sim \lvert\mathcal{N}(2, 4)\rvert$	$\alpha \sim \ell(0, 16)$	$e \sim \ell(\bar{e}, \varepsilon_e^2)$, $\bar{e} \sim \mathcal{U}(0.01, 1)$, $\log \varepsilon_e \sim \mathcal{U}(0.3, 1)$	150
15	Multi-region	$\gamma \sim \lvert\mathcal{N}(2, 4)\rvert$	$\alpha \sim \ell(0, 16)$	$e \sim \ell(\bar{e}, \varepsilon_e^2)$, $\bar{e} \sim \mathcal{U}(0.01, 10)$, $\log \varepsilon_e \sim \mathcal{U}(0.3, 1)$	150

Table 6.1 Simulation parameters for the 15 virtual phantoms used to generate the pool of synthetic training ultrasound imagery. $\ell(a_1, a_2)$ denotes the log-normal distribution with mean parameter a_1 and variance parameter a_2. $\mathcal{U}(a_1, a_2)$ denotes the uniform distribution on the interval $[a_1, a_2]$. The (abusive) notation $\gamma \sim \lvert\mathcal{N}(a_1, a_2)\rvert$ conveys the fact that γ is the absolute value of a normally distributed random variable with mean a_1 and variance a_2.

The training images were divided into 56 windows of 34×52 pixels, as shown in Figure 6.1(a). The statistical features R^2, S, w_l, w_a, and local elevational correlation length w_e were measured for each window and normalised by the transducer dependent values measured in the calibration scan. The local elevational correlation length associated with any image patch was computed from correlations between pairs of corresponding image patches located within 20 frames of it.

The synthetically generated pool of observations was first pruned to remove extreme outliers from the training set. A small number of these arose because of failed Gaussian fits to the decorrelation curves. Pruning was achieved by computing the median and median



(a) Speckle phantom (b) Homogeneous phantom (c) Multi-region phantom

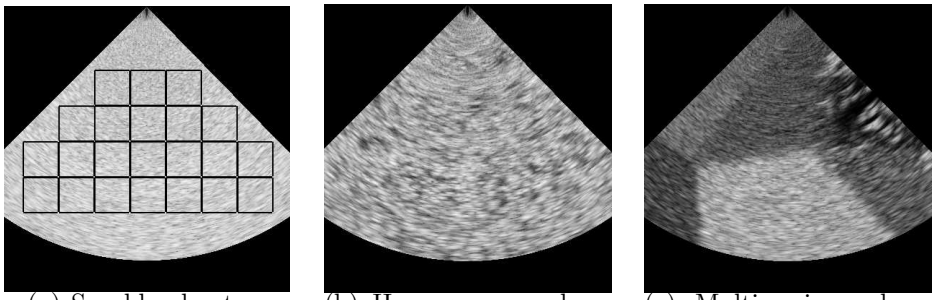
Figure 6.1 Sample synthetic images used for training. The speckle phantom image also shows the subdivision of the synthetic linear transducer images into 56 patches.

absolute deviation of the values obtained for each feature and discarding data points for which any of the features lay more than 7 median absolute deviations from its median. $N = 3000$ $\{r_{R^2}, r_S, r_l, r_a, r_e\}$ tuples were chosen at random from the remaining observations to learn the sparse Gaussian process regressor. The size of the pseudo-input set was $M = 10$ and sparse Gaussian process regression was carried out by the publicly available Gaussian Process Toolbox for Matlab [130]. Although the repeatability of this learning process over different combinations of input data points was not formally evaluated for the sparse Gaussian process model, informal experiments with a related relevance vector regressor proved its estimation be very repeatable, yielding similar regression models regardless of which 3000 data points were used. It is conjectured that this result extends to the sparse Gaussian process model that was ultimately chosen.

6.4.2 Test imagery

The regressor obtained during training was tested on a variety of synthetic and real ultrasound data sets. The synthetic data sets were generated like the training examples using Field II. One group of synthetic data sets was generated using the same virtual linear transducer as for the training images, whereas a second group of data sets was generated using

a virtual 7.5 MHz linear phased array (sector) transducer with a depth setting of 4 cm and a focus at 3 cm. Scan conversion for the envelope detected RF vectors from these simulations was carried out using a Matlab routine provided along with the Field II software.[2] These images were divided into 22 patches of 50×50 pixels, as shown in Figure 6.2(a) and back-converted to polar coordinates using a bi-cubic interpolator. This conversion was done because the axial and lateral directions in the proposed algorithm and Gee et al.'s heuristic (which was used as a base-line) are defined in terms of the direction of ultrasound signal propagation and that perpendicular to it, which do not correspond to the principal directions defined by the pixel grid in the case of scan-converted sector imagery.[3] The targets were virtual phantoms of varied homogeneity and micro-structure and were scanned at intervals of 0.1 mm. Sample synthetic sector scans are shown in Figure 6.2.



(a) Speckle phantom (b) Homogeneous phantom (c) Multi-region phantom

Figure 6.2 Sample synthetic imagery obtained using the virtual linear phased array. The speckle phantom image also shows the subdivision of the synthetic sector imagery into 22 patches.

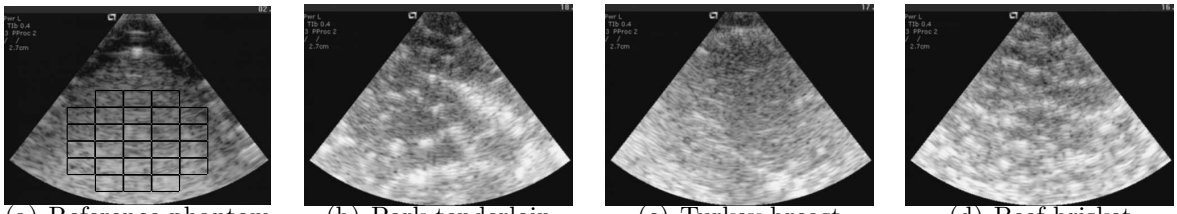
The real data were acquired using an Acuson Cypress ultrasound system with a 7V3C linear phased array transducer at a depth setting of 2.7 cm. The imagery was divided into 26 patches of 50×30 pixels, as illustrated in Figure 6.3(a), back-converted to polar coordinates and the echo amplitude signal was recovered after reversing the log-compression process using the compression parameter estimated for the experiments of Chapter 4.

One of the speckle phantom data sets from the experiments described in Chapter 4 with regular elevational spacing of 0.05 mm was used to obtain a nominal decorrelation

[2] `http://server.oersted.dtu.dk/personal/jaj/field/examples/ftp_files/kidney/`

[3] In practice, it was found that both algorithms still work well without this back-conversion to polar coordinates.

model. Ultrasound imagery of pork tenderloin, turkey breast and beef brisket samples purchased at the local grocery store was acquired at regular intervals of 0.03 mm, 0.04 mm, 0.05 mm, 0.06 mm, 0.08 mm and 0.1 mm using the manual sub-millimetre positioning device described in Chapter 4 (see Figure 4.12). In order to preserve the ground truth quality of the measurements obtained from the positioning device, friction between the probe and the highly deformable targets was prevented by scanning the samples through a thick layer of ultrasound coupling gel. It was found throughout series of similar experiments that poor or uneven coupling of the ultrasound signal, whether due to large air bubbles in the coupling gel or natural air pockets in the target, tends to create fast decorrelating patterns in the imagery. To ensure good and even coupling of the ultrasound signal, the layer of gel was carefully spread onto the target so as to destroy most of the air bubbles present in it. Scans displaying signs of poor coupling were rejected and re-acquired. Three scans of 161 frames each[4] were obtained from each tissue sample for each of the 6 scanning speeds. Exceptionally, no data were obtained for the pork tenderloin at 0.1 mm intervals due to the difficulty encountered in finding a surface that was flat enough with respect to the transducer's fixed height over the required distance. Figure 6.3 shows sample images from the phantom and different animal tissue samples.



(a) Reference phantom (b) Pork tenderloin (c) Turkey breast (d) Beef brisket

Figure 6.3 Sample real ultrasound imagery of the phantom and animal tissue samples used in the experiments. The phantom image shows the sub-division of the imagery into 26 patches. The pork, turkey and beef imagery were acquired using a higher overall gain setting than that used to acquire the phantom imagery in order to enhance visualisation. Note how the signal level is generally lower in the turkey image than in the pork and beef images, reflecting its the lower echogenicity of this medium.

[4]Some had a few extra frames.

6.4.3 Base-line methods

The accuracy of out-of-plane motion estimates obtained with the learning based approach proposed in this chapter was experimentally compared against three base-line approaches. The first approach is the direct use of the nominal elevational speckle decorrelation curves obtained during calibration at all points in the image, without any form of adaptation to the local characteristics of the tissue. The mean target registration errors obtained using this unadapted approach can be thought to quantify how much the imagery of a given data set deviates from fully developed speckle and how much adaptation was needed to accurately recover out-of-plane motion from correlation measurements. The two other base-line approaches were based on Prager et al.'s speckle detection method [167] and on Gee et al.'s heuristic method for adapting elevational decorrelation curves to local image properties [72], respectively. While efforts were made to keep the implementation of these methods as faithful as possible to the intent of the original publications, it is likely that there are implementation differences due to imperfect interpretation of the papers. Moreover, minor adaptations were made, in particular, to Gee et al.'s method, in order to use the methods with the instrumentation available for the experiments. The following sections describe how these base-line methods were implemented and adapted to the proposed experiments.

6.4.3.1 Prager et al.'s speckle detector

Prager et al.'s speckle detector [167] was used to determine whether any given pair of corresponding image patches would be used in the estimation of out-of-plane motion. The implementation of this detector mostly follows that described in [86], where it was numbered as "D3". The classifier uses the signal to noise ratio $R_{1.8}$ and skewness $S_{1.8}$ of the echo envelope A raised to the power of 1.8 as features:

$$R_{1.8} = \frac{E\{A^{1.8}\}}{\sqrt{E\{(A^{1.8})^2\} - E\{A^{1.8}\}^2}} \tag{6.30}$$

$$S_{1.8} = \frac{E\{(A^{1.8} - E\{A^{1.8}\})^3\}}{(E\{(A^{1.8})^2\} - E\{A^{1.8}\}^2)^{\frac{3}{2}}}. \tag{6.31}$$

$R_{1.8}$ and $S_{1.8}$ were computed for each image patch.

An ellipse in $(R_{1.8}, S_{1.8})$ space was used as a discriminant function, with the inside of the

ellipse describing the expected statistics of fully developed speckle. Exactly as described in [86], the ellipse was centred at $R_{1.8} = 1.031, S_{1.8} = 1.837$ with semi-major and semi-minor axes of 0.41 and 0.036, respectively. The major axis of the ellipse was oriented at an angle of 81.9 degrees with respect to the positive $R_{1.8}$ axis. This elliptical discriminant function is shown in Figure 6.4. For each candidate image patch, the distance $\mathfrak{d}$ between the $R_{1.8}, S_{1.8}$ statistics of the image patch to the centre of the ellipse was calculated. The degree of membership ϑ of the image patch to the class of fully developed speckle patterns was determined by calculating the ratio of $\mathfrak{d}$ to the radius $\mathfrak{r}$ of the ellipse in the direction of the line joining its $R_{1.8}, S_{1.8}$ position to the centre of the ellipse. A pair of corresponding image patches with membership values ϑ_1 and ϑ_2 was automatically accepted as fully developed speckle if $\max(\vartheta_1, \vartheta_2) \leq 7.5$.

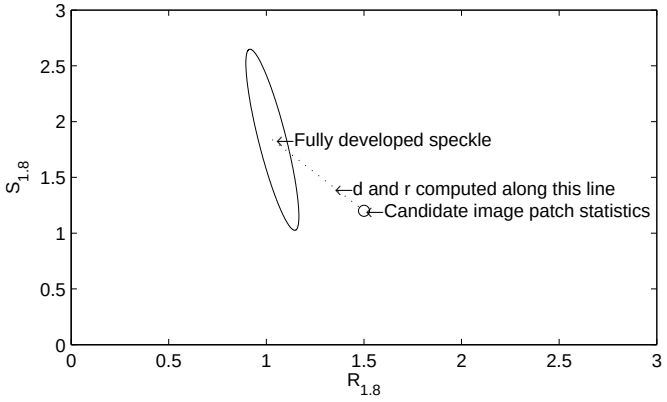


Figure 6.4 Elliptical discriminant function used for speckle classification. The combinations of values of $R_{1.8}$ and $S_{1.8}$ which lie inside the ellipse are considered characteristic of fully developed speckle. The degree of membership of a candidate image patch to the class of fully developed speckle image patches is measured by the ratio of the distance $\mathfrak{d}$ from its position in $R_{1.8}, S_{1.8}$ space to the centre of the ellipse, to the radius $\mathfrak{r}$ of the ellipse measured along the same direction as $\mathfrak{d}$.

This simplistic thresholding procedure often failed to leave enough data to compute the full rigid transformation between two image frames (three matching non-collinear patches are required for this). Therefore, patch pairs with the smallest ϑ values below a second threshold of 9 were greedily added to the usable set until a non-degenerate configuration was obtained for rigid transformation estimation. In a minority of cases this also failed; it should be noted that the statistics presented in Section 6.5 and appendix C do not include

results from these cases.

6.4.3.2 Gee et al.'s adaptive heuristic

Gee et al.'s approach [72] centres around a measure of scattering coherence $0 \leq c \leq 1$ which is used to modify the shape of the elevational speckle decorrelation curve established during calibration using the axial and lateral decorrelation curves measured at a particular location in imagery of real tissue. In this approach, it is assumed that any two patches of echo envelope data can be expressed as A_1+cA_2 and A_2+cA_1. Here, A_1 and A_2 are two echo envelope signal patches consisting of fully developed speckle with correlation coefficient $\hat{\rho}$ predicted by the nominal (Rayleigh) speckle decorrelation model and the elevational separation between the image patches. When the coherence measure $c = 0$, the correlation between the two images is that predicted by the nominal model, $\hat{\rho}$. For $c = 1$, the two image patches are identical and are therefore fully correlated. For intermediate values of c, the correlation coefficient of the two image patches as a function of $\hat{\rho}$ is given by [72]

$$\rho = \frac{(1 + c^2)\hat{\rho} + 2c}{(1 + c^2) + 2c\hat{\rho}}. \tag{6.32}$$

Conversely, c can be determined from the nominal correlation value $\hat{\rho}$ and the measured correlation ρ:

$$c = \begin{cases} 0 & \rho = \hat{\rho} \\ \dfrac{(1 - \rho\hat{\rho})}{(\rho - \hat{\rho})} \pm \dfrac{\sqrt{(1 - \rho^2)(1 - \hat{\rho}^2)}}{(\rho - \hat{\rho})} & \text{otherwise.} \end{cases} \tag{6.33}$$

These equations provide a means of predicting the elevational decorrelation curve from the axial and lateral decorrelation curves measured within the image plane. An image patch thus provides two coherence functions via (6.33), $c_a(\delta_a)$ and $c_l(\delta_l)$, derived from the axial and lateral decorrelation curves, respectively. These are averaged over the two corresponding patches. Here, δ_a and δ_l denote the known axial and lateral distances (counted in pixels, or converted to millimetres or radians, depending on the geometry of the scan) used to calculate the in-plane decorrelation curves. To account for the anisotropy of the nominal decorrelation model (due to the ultrasound beam having different thickness in the elevational, lateral and axial directions) and possible differences in the units used to measure distances in elevational, axial and lateral directions, the coherence c_e of the signal in the

elevational direction, as a function of c_a, is given by

$$c_e^a\left(\frac{\hat{w}_e}{\hat{w}_a}\delta_a\right) = c_a(\delta_a), \tag{6.34}$$

and a similar expression is derived for c_e^l expressed as a function of c_l, with $\hat{w}_e$, $\hat{w}_a$ and $\hat{w}_l$ respectively defined as the nominal elevational, axial and lateral correlation lengths as in the rest of this chapter.[5] The final $c_e(\delta)$ is then given by

$$c_e(\delta) = \frac{1}{2}(c_e^a(\delta) + c_e^l(\delta))(1 - g_1(y - g_2)), \tag{6.35}$$

where the first factor is simply an average of functions c_e^a and c_e^l, and the second factor is an *ad hoc* correction factor which the authors introduced in order to compensate for the tendency of their approach to underestimate distances in the near field of the transducer, and overestimate them in the far field. Here, y is the axial depth of the image patches under consideration, and g_1 and g_2 are adjustable parameters to be tuned empirically for each transducer; the paper suggests that in the case of a linear transducer, a good value for g_2 would be the approximate depth of the transducer's focal point. The local elevational decorrelation curve is then estimated from the new elevational coherence function c_e and the nominal speckle decorrelation model using (6.32).

For the experiments presented here, the values of g_1 and g_2 were tuned empirically for both synthetic transducers and the real transducer. For the synthetic transducers, g_2 was set to the known depth of the transducer's focus and a suitable value for g_1 was determined by trial and error. This process was found to be much easier in the case of the linear transducer, which may be related to the fact that Gee et al.'s algorithm was originally designed and tested for one such linear transducer. Since the location of the focus was unknown for the real transducer, both g_1 and g_2 had to be tuned by trial and error for this case, and it is likely that the selected combination, though it gave reasonable results most of the time, is sub-optimal. The selected values are given in Table 6.2.

[5]The way in which this adjustment is performed constitutes a minor departure from the original paper [72], where the adjustment factor was approximate, was only provided in the case of the axial direction and corresponded to the particular linear transducer used in the experiments.

Transducer	g_1	g_2 (mm)
Virtual linear transducer	-0.04	30.0
Virtual linear phased array (sector) transducer	0.01	30.0
Real 7V3C transducer	0.1	30.0

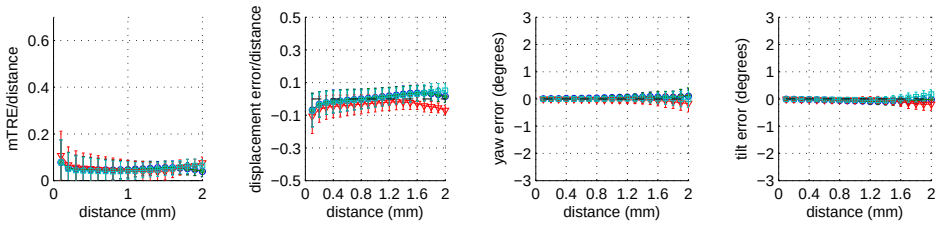
Table 6.2 The values of adjustable parameters g_1 and g_2 of Gee et al.'s adaptive algorithm [72] empirically selected for the two virtual transducers and the real transducer considered in this study.

6.5 Results

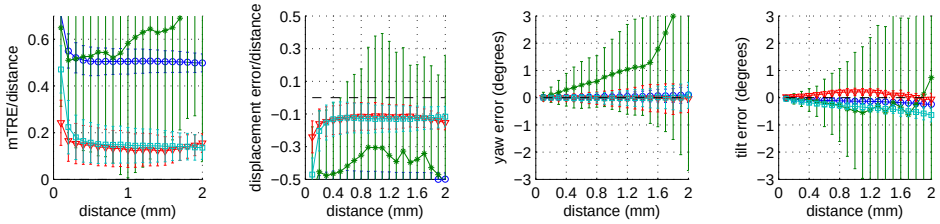
The accuracy of the out-plane-motion estimates resulting from the use of the method proposed in this chapter (or any of the base-line methods) was evaluated by estimating the elevational distances between pairs of corresponding image patches using the chosen speckle decorrelation model and fitting a rigid transformation to these patch-wise distance estimates for each pair of frames separated by less than 20 times the step size used to obtain the calibration imagery (0.1 mm for the synthetic imagery and 0.05 mm for the real imagery). The rigid transformations were fitted using the least-median-of-squares criterion [183, 184] already described in Chapter 3 (see (3.4)). The rest of this section presents and analyses the results.

6.5.1 Results on synthetic imagery

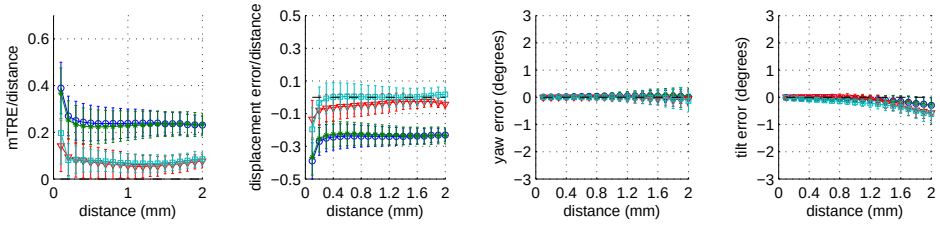
In order to verify that the learning process was effective, out-of-plane motion estimation was first attempted on synthetic ultrasound data sequences obtained using the same virtual transducer as was used to generate the training data, but using different virtual phantoms. Figures 6.5 and 6.6 show the distribution of the pairwise motion estimation error as a function of elevational separation for such cases. Figure 6.5 shows the results obtained for phantoms of homogeneous composition, including a fully developed speckle phantom, whereas Figure 6.6 shows the results obtained for multi-region phantoms. The error is quantified in terms of (1) the ratio of the mean target registration error (mTRE) to the true elevational separation; (2) the ratio of the difference between the estimated separation and the true one to the true elevational separation; (3) the estimated yaw and (4) the estimated tilt, using the centroid of the grid of image patches as the centre of rotation. Both the yaw and tilt should be zero if correctly estimated in these experiments.



(a) Linear scan of homogeneous fully developed speckle phantom: 100 frames. Generated using the fractal model of Chapter 5 with $\gamma = 16$ scatt./mm^3 and $\alpha = 1$.

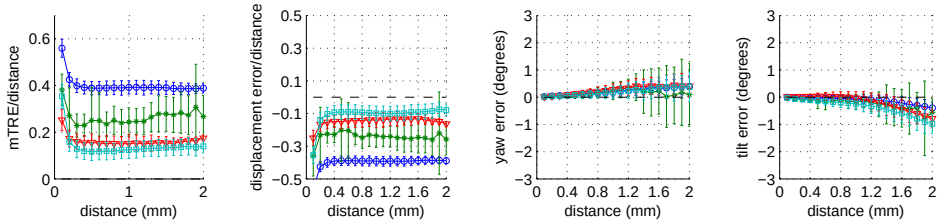


(b) Homogeneous linear 1: $\gamma = 0.5$ scatt./mm^3, $\alpha = 0.03$, 100 frames

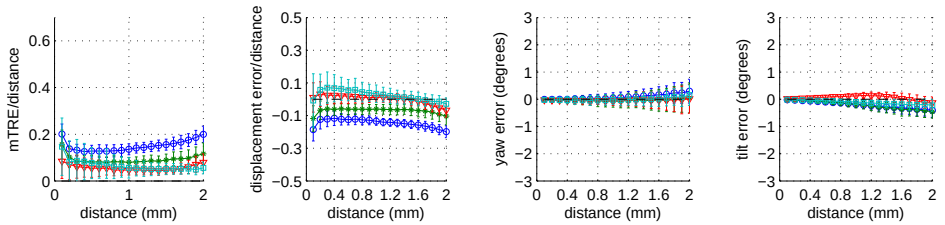


(c) Homogeneous linear 2: $\gamma = 0.5$ scatt./mm^3, $\alpha = 0.3$, 100 frames

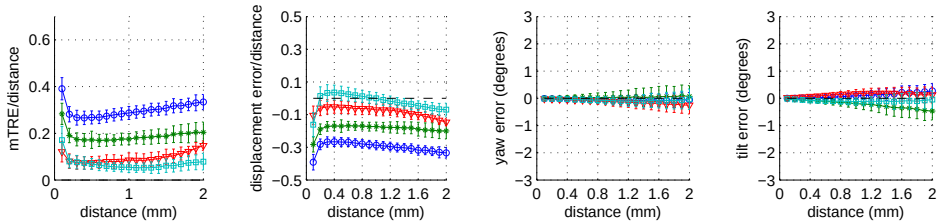
Figure 6.5 Out-of-plane motion estimation error between pairs of synthetic linear array images of homogeneous phantoms. All image sequences were acquired at 0.1 mm intervals in the elevational direction. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation. Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.



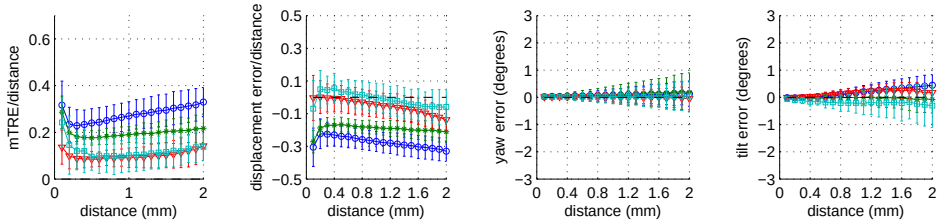
(a) Multi-region linear 1: 50 frames



(b) Multi-region linear 2: 150 frames



(c) Multi-region linear 3: 150 frames



(d) Multi-region linear 4: 150 frames

Figure 6.6 Out-of-plane motion estimation error between pairs of synthetic linear array images of multi-region phantoms. All image sequences were acquired at 0.1 mm intervals in the elevational direction. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation. Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.

From these results, it can be seen that all algorithms succeeded in correctly estimating out-of-plane motion in the case of the homogeneous fully developed speckle phantom (Figure 6.5(a)). This is not surprising since this the nominal decorrelation model was derived from a scan of a very similar phantom and did not require significant adaptation. Nonetheless, it is a good sign for the new approach that distance estimates did not deteriorate.

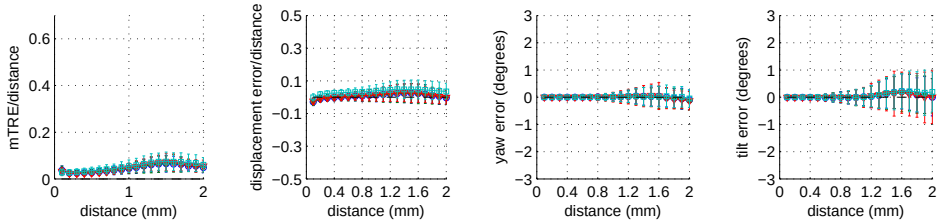
The results obtained for non fully developed speckle virtual phantoms (Figures 6.5 and 6.6) are, predictably, quite different. In these cases, using the nominal speckle decorrelation model often yields large errors. For multi-region phantoms (Figure 6.6), there is often some error in the yaw and tilt components of the estimated rigid transformation. This is explained by the non-uniformity of the error in the individual patch-wise distance estimates caused by the different amounts of adaptation needed in different regions of the phantom.

The speckle detection approach most often succeeded in reducing motion estimation error somewhat, but, as was found by Hassenpflug et al. [86], the trade-off between accurate rejection of non speckle patches and the collection of sufficient data for rigid transformation estimation means that these results are still unsatisfactory. In some cases, such as that shown in Figure 6.5(b) the motion estimate obtained using the speckle detector is even worse than that obtained using the nominal speckle decorrelation model directly. This probably owes to the poor spatial distribution of the greedily selected set of "best" image patches.

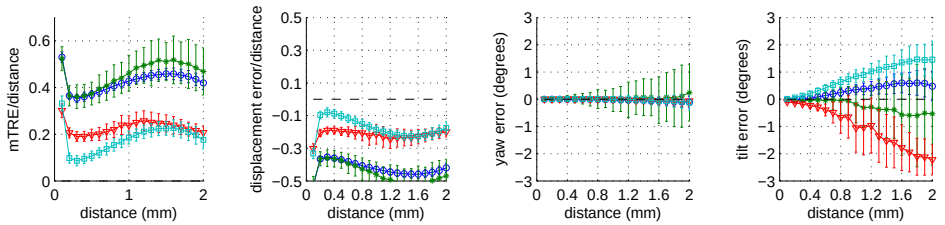
In contrast, both the proposed learning-based approach and Gee et al.'s heuristic approach [72] were generally able to reduce the motion estimation error to levels comparable to that obtained for the calibration phantom. This is a very encouraging result as it shows that the learned relationship between in-plane image statistics and elevational correlation length generalises to imagery outside the training set.

The next set of experiments was carried out on the synthetic imagery generated using the virtual linear phased array transducer. The aim of these experiments was to find out whether the learned model also generalises to imagery acquired using a transducer with characteristics and geometry different from that used to generate the pool of training data. Figures 6.7 and 6.8 show the distribution of motion estimation error over all pairs of frames considered in these experiments. The virtual phantoms were mostly homogeneous ones, but two multi-region phantoms were also used (the last two cases of Figure 6.8).

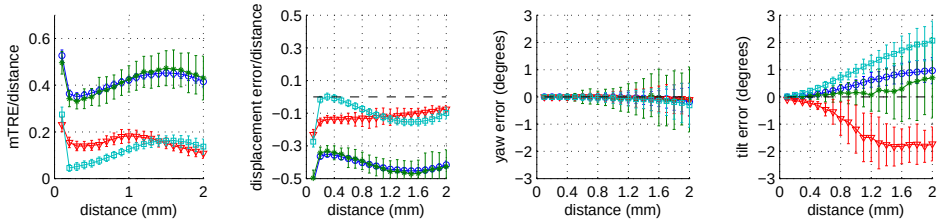
As with the linear transducer results, accurate distance estimation was achieved by



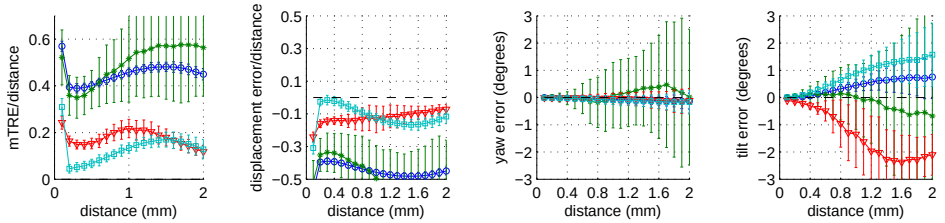
(a) Sector calibration scan: 140 frames



(b) Homogeneous sector 1: $\gamma = 0.5$ scatt./mm^3, $\alpha = 10$, 140 frames

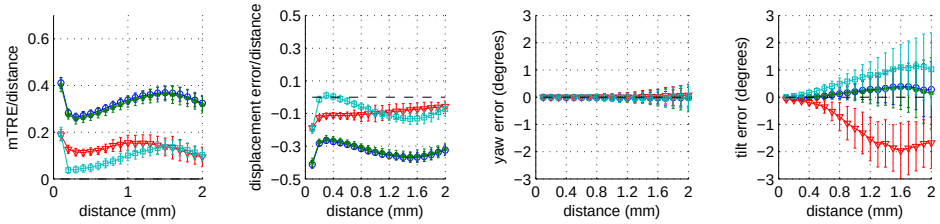


(c) Homogeneous sector 2: $\gamma = 1$ scatt./mm^3, $\alpha = 0.3$, 140 frames

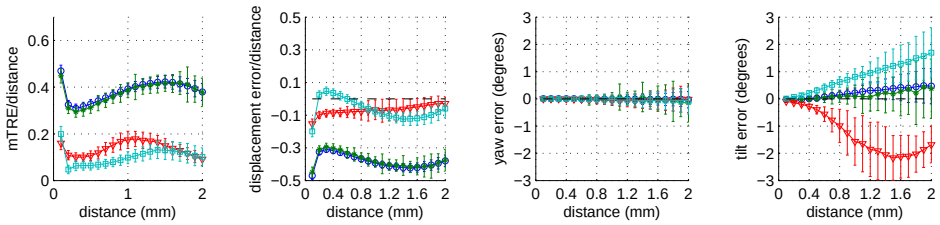


(d) Homogeneous sector 3: $\gamma = 2$ scatt./mm^3, $\alpha = 0.1$, 140 frames

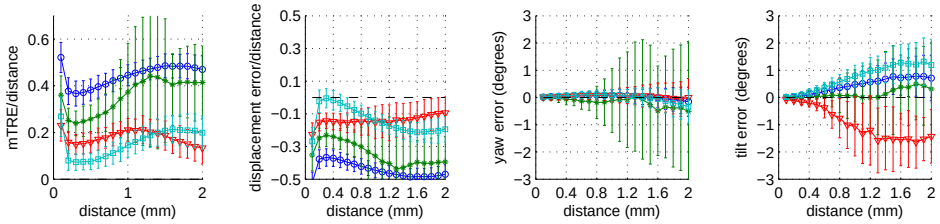
Figure 6.7 Out-of-plane motion estimation error between pairs of synthetic sector scan images of homogeneous phantoms. All image sequences were acquired at 0.1 mm intervals in the elevational direction. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation. Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.



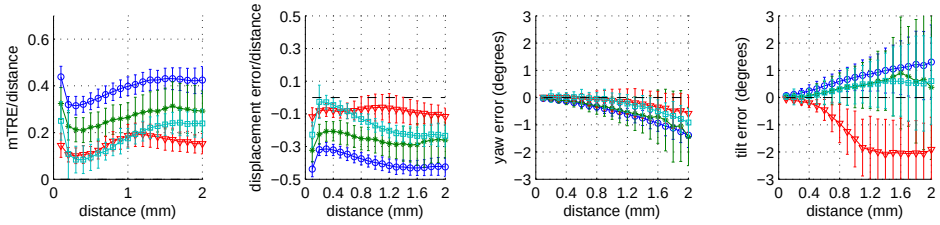
(a) Homogeneous sector 4: $\gamma = 2$ scatt./mm^3, $\alpha = 0.3$, 140 frames



(b) Homogeneous sector 5: $\gamma = 4$ scatt./mm^3, $\alpha = 0.1$, 140 frames



(c) Multi-region sector 1: 140 frames



(d) Multi-region sector 2: 140 frames

Figure 6.8 Out-of-plane motion estimation error between pairs of synthetic sector scan images of homogeneous and multi-region phantoms. All image sequences were acquired at 0.1 mm intervals in the elevational direction. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.

all methods for the calibration phantom (Figure 6.7(a)). Outside the context of Rayleigh scattering, the speckle detection approach was even less successful for the sector scans than for the linear scans. There are two likely reasons for this:

1. The scatterer density parameters (defined in scatterers/mm^3) used to generate the virtual phantoms for the synthetic sector scans were drawn from similar distributions to those used in the case of the virtual linear transducer. However, the size of the resolution cells in the sector imagery was much smaller than for the linear transducer imagery owing to the higher frequency of the ultrasound pulse. This means that the scatterer density in terms of scatterers per resolution cell was also generally lower than for the linear transducer cases, leading to fewer image patches whose content approached fully developed speckle.

2. There was a large difference between the number of patches in the two types of data: the linear scans were divided into 56 patches, whereas the sector imagery only had 22. This means that the speckle detection approach had less choice when collecting image patches suitable for the estimation of a rigid transformation.

The results obtained with the proposed method and with Gee et al.'s method were once again better than for the other alternatives, but not as good as those obtained with the linear transducer data. This is quite possibly a consequence of the smaller numbers of scatterers per resolution cell in those simulations. This hypothesis is supported by the fact that the worst results in the linear transducer case were obtained for the lowest density (Figure 6.5(b)) and the best results in the sector imagery case were obtained for the largest density (Figure 6.8(b)).

Another notable difference concerns the shape of the error curves. For the linear transducer cases, the ratio of the error to the true distance is nearly constant, whereas it appears to be a fast increasing and highly curved function in the case of the linear phased array cases. The reasons for this are not entirely clear, though in some cases, it may have to do with the persistence of image features across the scan, arising from the presence of structured components in the virtual phantom. Nevertheless, it can be observed that the shape of the error curves obtained for the learning based approach tends to follow that obtained using the nominal decorrelation model. This is because the learned model is, by construction, simply a scaled version of the nominal decorrelation curve. The shape of

the error curves obtained for Gee et al.'s method is quite different, reflecting the fact that this approach reshapes the nominal decorrelation curve at every sample, allowing for more complex adaptations.

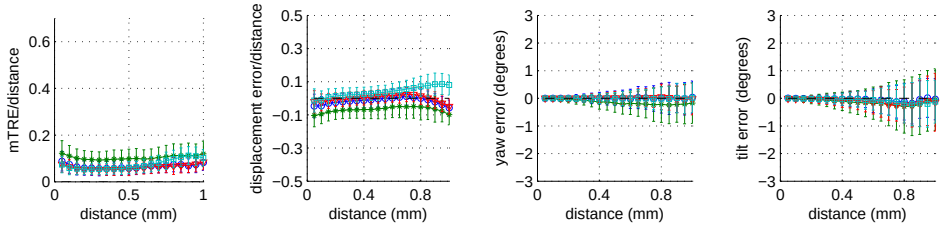
Though they are by no means perfect, the results on synthetic sector scan imagery suggest that the proposed approach generalises reasonably well to data acquired using transducers different from that used to generate the training set. The similarly less than perfect performance of the state-of-the-art method on these cases suggests that the errors encountered are somehow related to the parameters of the artificially generated data sets rather than symptomatic of major flaws in the adaptation algorithms.

6.5.2 Results on real imagery

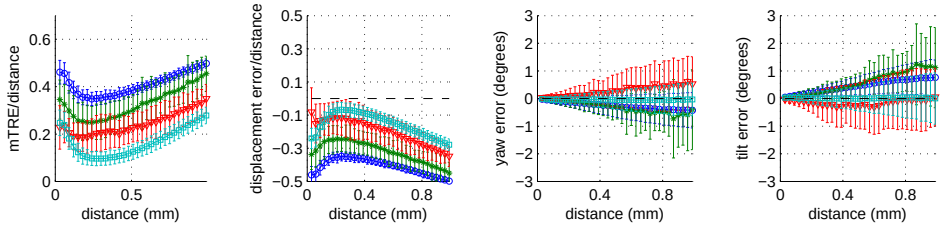
The generality of the learned adaptive decorrelation model was further verified on the real ultrasound scans of animal tissue. Since the experiments with animal tissue involved 51 distinct data sets, it is impractical to include the full set of results in this chapter. Therefore, these are made available in Appendix C. For the purpose of this discussion, Figures 6.9, 6.10 and 6.11 show sample representative results for the reference phantom and pork tenderloin, turkey breast, and beef brisket data sets respectively. The animal tissue data sets selected for discussion were the ones acquired at 0.03 mm intervals, allowing dense performance assessment of the different algorithms over the correlation range of the imagery. As with the results on synthetic data sequences, the plots show (1) the ratio of mean target registration error to true elevational separation; (2) the ratio of the difference between the estimated elevational separation and the true one to true elevational separation, evaluated using the centroid of the grid of image patches as the centre of rotation; (3) yaw evaluated using the centroid of the grid of image patches as the centre of rotation; (4) tilt evaluated using the centroid of the grid of image patches as the centre of rotation.

As with the synthetic scans, the pairwise distance estimation results obtained for the reference phantom scan were good for all methods considered. The speckle detection approach did slightly worse than the others. This implies that the speckle detector did not accept all image patches as fully developed speckle, or the results would have been identical to those obtained for the unadapted case. This may be explained by a combination of three factors:

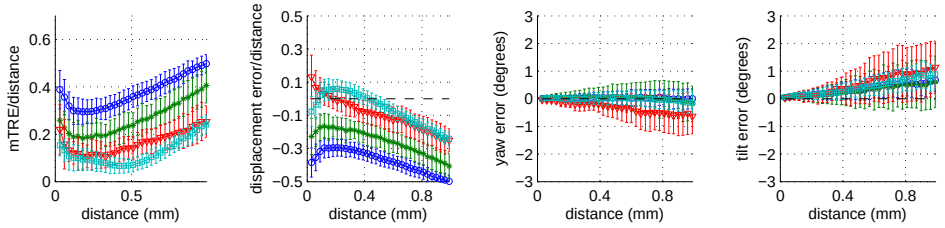
1. The size of the image patches was quite small, making the computation of first order



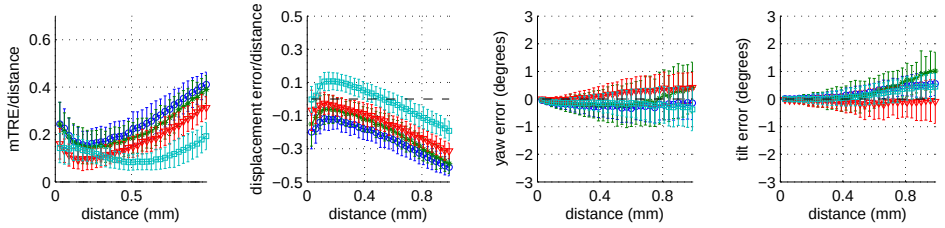
(a) Calibration scan: 161 frames, 0.05 mm intervals



(b) Pork tenderloin 1: 163 frames, 0.03 mm intervals

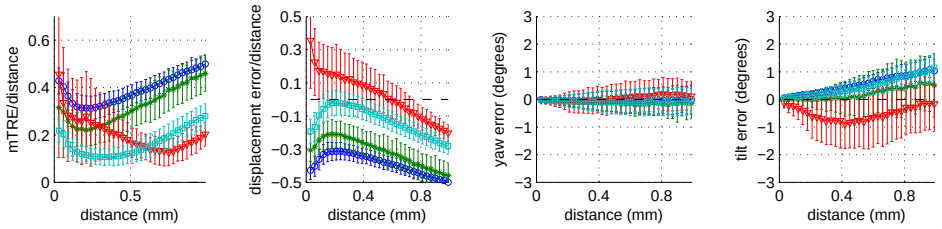


(c) Pork tenderloin 2: 161 frames, 0.03 mm intervals

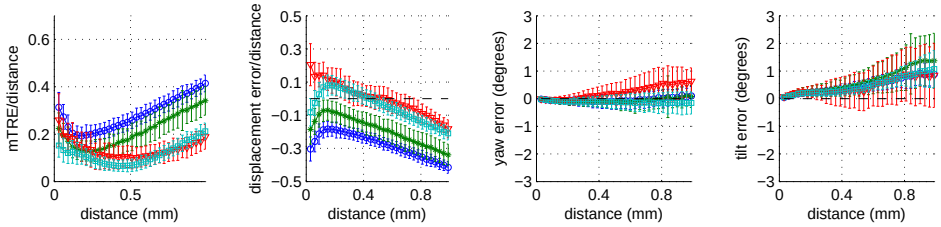


(d) Pork tenderloin 3: 161 frames, 0.03 mm intervals

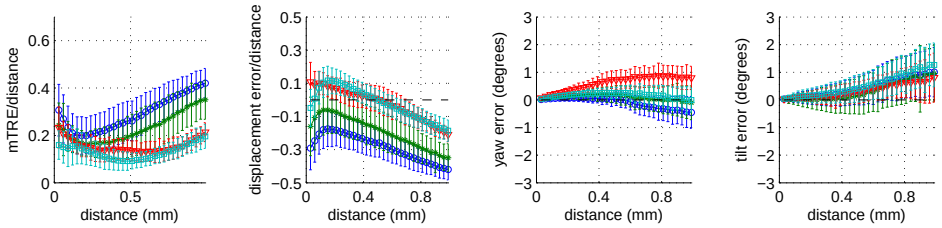
Figure 6.9 Out-of-plane motion estimation error between pairs of speckle phantom and pork tenderloin images. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation. Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.



(a) Turkey breast 1: 161 frames, 0.03 mm intervals

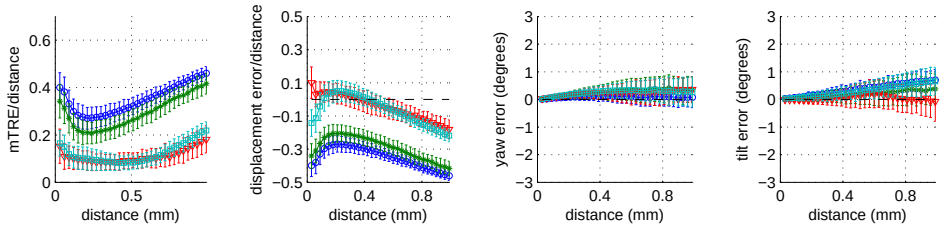


(b) Turkey breast 2: 161 frames, 0.03 mm intervals

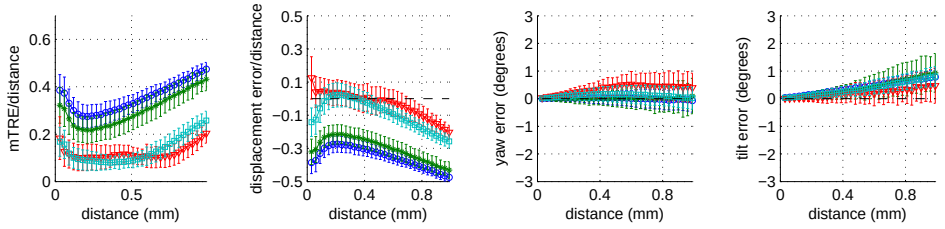


(c) Turkey breast 3: 161 frames, 0.03 mm intervals

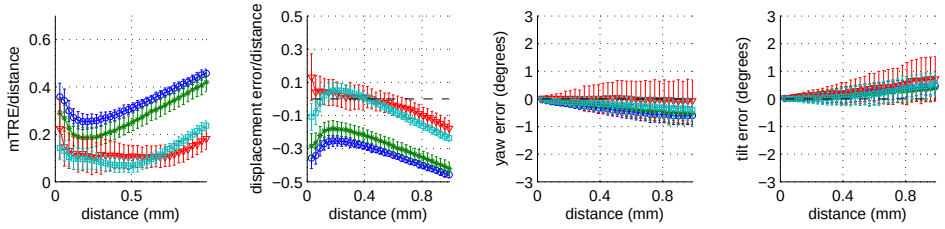
Figure 6.10 Out-of-plane motion estimation error between pairs of speckle phantom and turkey breast images. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation. Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.



(a) Beef brisket 1: 161 frames, 0.03 mm intervals



(b) Beef brisket 2: 161 frames, 0.03 mm intervals



(c) Beef brisket 3: 161 frames, 0.03 mm intervals

Figure 6.11 Out-of-plane motion estimation error between pairs of beef brisket images. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation. Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.

statistics $R_{1.8}$, and particularly $S_{1.8}$, required by the detector susceptible to noise arising from the finite sample size. This may have caused the detector to incorrectly reject fully developed speckle patches.

2. The dynamic range compression parameter may not have been recovered perfectly by Prager et al.'s algorithm [168]. Slightly inaccurate decompression may thus have altered the first order statistics of the decompressed signal compared to those of the original, unavailable envelope detected signal produced by the ultrasound machine.

3. The phantom itself is probably not perfect and exhibits some spatial variations in micro-structure, as already discussed in Chapter 4.

The more interesting results are those obtained for the animal tissue samples. At first glance, the results obtained with the virtual linear transducer imagery appear to have generalised very well to the real data. Without adaptation, results were very poor in most cases, confirming that the tissue samples scanned significantly differed in structure from the reference phantom and justifying the application of adaptive approaches. Generally, the speckle detection method was, again, able to improve on these results somewhat significantly, but did not approach the error levels reached in the case of the phantom. In most cases, Gee et al.'s method and the learning-based approach proposed in this chapter were both much more successful, with mean target registration errors typically neighbouring the 10% mark at their lowest point in the range of elevational separations considered, and errors considerably smaller than for the alternatives elsewhere.[6] The learning-based approach achieved excellent performance, indicating that the regressor learned from synthetic data generalised well across the transition to real ultrasound data and the change in transducer specifications and geometry.

The shape of the error curves (mainly the mTRE and distance estimation error) obtained with the proposed method merits attention, as it generally differs substantially from that of the curves that were obtained with the synthetic linear transducer imagery. In the latter case, the mean target registration error curve was generally flat, except for very short distances corresponding to the region of the decorrelation curve most poorly modelled

[6]The poor behaviour of Gee et al.'s heuristic method on the first turkey data set of Figure 6.10 is the exception rather than the rule, as corroborated by the full set of results presented in appendix C. It is possibly the result of a misunderstood implementation issue, or might be corrected by a different choice for the adjustable parameters g_1 and g_2.

by the piecewise linear approximation. For the real data, however, there is typically a considerable increase in the error for larger elevational separations. The onset and rate of this deterioration in accuracy differ from one data set to the next. Compare, for instance, the error curves obtained for the two beef data sets of Figures 6.11(a) and 6.11(b). The increase in error happens more slowly and for larger distances in the former than in the latter.

A plausible explanation for this error increase for larger distances is illustrated in Figure 6.12 and initially suggested itself in an exploratory experiment on a sample of pork flank. The figure shows an elevational decorrelation curve obtained for imagery of the speckle phantom, which follows a Gaussian-like shape, and the elevational decorrelation curve obtained from a scan of pork flank at the same image location. The adaptive method presented in this chapter essentially re-scales the nominal decorrelation curve along the δ axis according to the local statistics of the image, yielding an adapted decorrelation curve which is still Gaussian-like and tends towards zero correlation. This prediction follows the true decorrelation curve closely for shorter distances; however, as illustrated in Figure 6.12, for this particular image location, the imagery never decorrelates completely, instead reaching a correlation plateau value near 0.5 or 0.6. Notice that the image location corresponding to this curve (Figure 6.12) contains highly specular features due to the presence of a thick layer of fat in the tissue sample. This layer of fat continues to exist over large elevational displacements. Such features also appear in the data sets considered in this thesis, though they are not always as visually blatant.

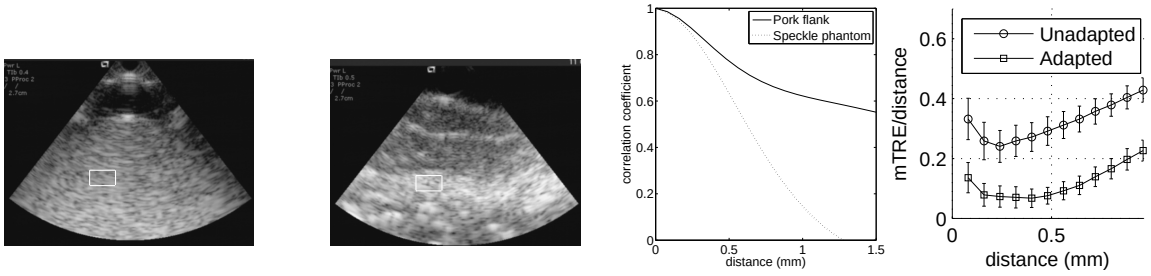


Figure 6.12 Illustration of correlation plateaux caused by persistent structures in the imaged medium. From left to right: speckle phantom image; pork flank image; decorrelation curves associated with the image patches inside the white rectangles; mean target registration error obtained for the pork flank data set without adaptation and using the learning-based approach.

Elevationally persistent structure causing correlation plateaux also arises from boundaries between regions of significantly different echogenicity or micro-structure, such as the random Voronoi regions contained in the virtual multi-region phantoms used in simulation. Upon investigation, it was found that such correlation plateaux did indeed occur for image patches containing region boundaries in the case of the linear transducer simulations presented in this chapter. However, the phenomenon did not occur at enough locations within the image plane to have a sizeable effect on the overall rigid transformation fit between pairs of frames. In fact, image patches where this occurred were often rejected as outliers by the least-median-of-squares fitting algorithm, which was overwhelmed with data from image patches with uniform structure.

Overall, the experimental results suggest that Gee et al.'s heuristic approach [72] is slightly more immune to the effect of structures persisting over large distances than the learning-based approach, with error curves which remain somewhat flatter and increase at slower rates. This is likely because this approach applies a more complex transformation to the nominal decorrelation model than the simple scale factor used by the learning-based method, which means that the shape of the adapted decorrelation curve does not have to follow that of the nominal model. It would seem that the shapes of the axial and lateral decorrelation curves used by the heuristic approach provide some useful information in this respect which the proposed method does not exploit. Nonetheless, neither method can be said to model elevational decorrelation extremely well at larger distance intervals when persistent structures are present.

In a few cases, the learning-based approach, Gee et al.'s approach, or both, had difficulty modelling elevational speckle correlation for short distances as well, leading to overestimates of elevational separation. Such a case, occurring with one of the turkey breast data sets, is illustrated in Figure 6.13. In this case, both of the adaptive approaches substantially over-estimated the elevational separation between pairs of frames. It was observed that such cases tended to occur more commonly when the signal level was comparatively low, i.e. when the scanned tissue sample had low echogenicity properties, as is the case with lean meat samples like turkey breast (see Figure 6.3 for an illustrative example). This suggests the hypothesis that the inaccuracies are caused by noise in the image acquisition process, whose relative contribution to the image content increases when the initial ultrasound signal level is low. Sources of such noise include electronic noise from the ultrasound scanner, quantisation noise arising from the conversion of the analog envelope signal to a

coarsely sampled set of log-compressed 8-bit grey scale pixels as well as potential interlacing artefacts introduced by the video frame grabber.

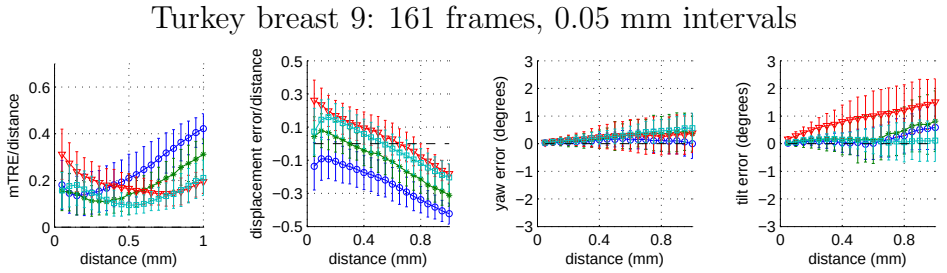


Figure 6.13 Example of out-of-plane motion over-estimation in turkey breast imagery. From left to right: the ratio of the mTRE to the true displacement, the ratio of the estimated displacement error to the true displacement, the yaw error and the tilt error measured using the centroid of the grid of image patches as the centre of rotation. Dark blue circles: using the nominal decorrelation model without adaptation. Green stars: using Prager et al.'s speckle detector [167]. Red triangles: using Gee et al.'s heuristic adaptive method [72]. Cyan squares: using the learning based approach. The error bars represent one standard deviation.

These types of noise were not modelled in the construction of the training set (though they certainly could be in an improved version of the approach) and may have a variety of effects on the accuracy of correlation-based distance estimation:

- The true elevational decorrelation curve may be altered. The maximum achievable correlation value likely decreases, and the global shape of the curve (local slope and curvature) most likely change as well, with the severity of the change depending on the signal-to-noise ratio. It was observed that occasionally, the unadapted nominal speckle decorrelation model itself yielded displacement over-estimates, which is possibly a result of such changes.

- The in-plane image statistics used by the speckle detector, Gee et al.'s algorithm and the learning-based approach may also be significantly altered. In particular, the signal-to-noise ratio statistic R may undergo a systematic decrease due to a noise-induced increase in the variance of the pixel values. Axial and lateral decorrelation curves may be altered in ways more or less similar to the elevational one and become less predictive of its behaviour.

The behaviour of the speckle detection approach does not change much in such conditions, as it only uses the image statistics to determine whether or not an image patch should be used, a binary decision whose tolerance to noise may be quite high. Alteration of the input and output signal statistics most strongly affects adaptive algorithms such as Gee et al.'s or the approach proposed in this chapter, because these approaches are the ones which most strongly rely on the relationship between these statistics. In particular, the accuracy of the learning-based approach depends entirely on the assumption that the training set from which the regression model was established was representative of the data samples to be processed later. Since noise was not modelled at all in the simulations that led to the creation of the synthetic ultrasound training set, it would not be surprising if predictions deteriorated in the presence of noise. This is not a fatal limitation, however, as many sources of noise could easily be modelled as part of the simulation process and an improved elevational correlation length predictor derived from the more realistic training set. In this respect, the proposed learning-based approach is more adaptable than that of Gee et al.

Another potential source of error for the learning-based approach is that the axial and lateral decorrelation curves may not always be well represented by Gaussian curves. The features w_a and w_l computed from the least-squares Gaussian curve fit may thus not always be very representative of the local axial and lateral decorrelation rates, particularly if correlation plateaux similar to that shown in Figure 6.12 appear in the axial and/or lateral directions. These were not observed very often (probably because of the small distance ranges considered in the axial and lateral directions), but in such cases, a least squares Gaussian fit would tend to over-estimate decorrelation rate for small distances, and might lead the regressor to output over-inflated values of w_e, causing distance over-estimates. It is not yet known whether this is in fact a significant problem, but if it is, a potential solution might be to carry out the Gaussian curve fits using a criterion other than the least mean squared error. A fitting approach which assigns less weight to correlation measurements associated with large separations might be more appropriate.

6.6 Summary and discussion

This chapter described a novel approach for adapting a nominal elevational decorrelation model obtained from speckle phantom data to imagery of different media using local image

statistics. In this approach, the relationship between the in-plane first and second order statistics of the image and elevational correlation length is established by applying sparse Gaussian process regression to a large pool of associations between in-plane image statistics and elevational correlation length obtained from varied synthetic ultrasound data sets generated using the method described in Chapter 5.

Experiments on synthetic data and real ultrasound data of animal tissue samples showed that the learned regressor generalises well across transducer and medium changes, achieving significantly lower distance estimation errors than an approach based on the detection and exclusive use of fully developed speckle patches [167, 86]. The approach was also compared against the state-of-the art heuristic method developed by Gee et al. [72]. Quantitative statements regarding this comparison are probably unwise given that the implementation and tuning of Gee et al.'s approach used in this work may differ from what was intended by the authors and are likely sub-optimal. A fair statement would be that, qualitatively, the proposed method exhibited accuracy comparable to the state-of-the-art.

A unique feature of the learning-based method presented in this chapter is that the estimates of correlation length provided by the sparse Gaussian regressor are probabilistic in nature as they are given in terms of a mean and variance. The variance component of the estimates was not fully exploited in this work, but certainly warrants further research interest as it may provide useful quantification of uncertainty or reliability in tasks which require distance measurements. The analysis of the experimental results with imagery of animal tissue also suggests a number of avenues for research. Though it is satisfactory in many respects, the adaptive decorrelation model used in this chapter is perhaps overly simplistic as it assumes that the true decorrelation curve of ultrasound in an arbitrary medium is merely a scaled version of the nominal, transducer specific decorrelation curve. Experimental results show that this is clearly not the case and it seems plausible that the violation of this assumption in the axial and lateral cases might the cause of some of the observed distance estimation inaccuracies. Accuracy might be improved by describing changes in shape of the axial, lateral and elevational decorrelation curves by more than a single scaling parameter. Finally, generalisation of the learned regressor to real ultrasound data can most probably be improved by creating more realistic synthetic training sets. Important unmodelled variables are electronic and frame-grabber noise as well as signal attenuation. The latter would be a very relevant variable to model if data were to be acquired at larger depths. Another modelling parameter of interest would be medium

anisotropy, i.e. the propensity of ultrasound scatterers to have spatially oriented spatial organisations in certain types of tissue (e.g. the muscle fibers imaged in the animal tissue experiments of this chapter). This might be achieved by modifying the scatterer distribution model proposed in Chapter 5. That such variables can eventually be integrated into the construction of the training set to obtain more accurate models is another unique strength of the proposed approach.

In spite of the very simple modelling assumptions it relies on, the learning-based approach presented in this chapter was quite successful in providing accurate distance estimates and its formulation makes it highly adaptable in terms of the imaging conditions that can be accounted for. As such, it is a promising new component within a system for inferring the trajectory of an ultrasound transducer in real-tissue. That is precisely what will be discussed in the next chapter.

Chapter 7

Probabilistic out-of-plane trajectory estimation in ultrasound imagery of real tissue

The goal of the work presented in this thesis was to make possible the accurate estimation of out-of-plane ultrasound transducer trajectories from imagery of arbitrary media. This chapter shows how the different components already presented in this thesis can be modified and integrated within a single framework in order to achieve this and how the limitations of each of these components can be circumvented or overcome when they are used in combination. In what follows, the most important components of the system are reviewed and their limitations identified. A solution which accounts for these limitations through a novel measurement selection process will then be introduced.

Chapter 6 introduced a method for dynamically adapting a nominal, transducer specific elevational speckle decorrelation model to the characteristics of the scanned medium. This is done by learning from images of synthetic phantoms of varied micro-structure created using the scatterer distribution model of Chapter 5. Experimental results showed that this approach was successful, in that relatively accurate elevational separation estimates were generally obtained over a moderate range of true elevational separations. However, it was also found that the accuracy of the dynamically adapted model did not typically extend to the full range of elevational separations originally included in the nominal speckle decorrelation model. The main reason for this is that the locally adapted decorrelation

model is simply a rescaled version of the nominal model and as such, it predicts a Gaussian-like curve which tends towards zero correlations for large distances, thereby failing to model the effect of some elevationally persistent structures in the medium (such as the layer of fat in the pork flank example given in Figure 6.12 of the previous chapter).

Chapter 4 introduced a probabilistic methodology for modelling the uncertainty of distance estimates derived from the sample correlation coefficients between pairs of image patches and obtaining a maximum likelihood trajectory estimate from a set of redundant probabilistic distance estimates. In the development of this approach, it was assumed that the scanned medium exhibited Rayleigh scattering conditions, precluding its direct use in recovering transducer trajectories from imagery of real tissue. Simulations and experiments on speckle phantom data showed that this approach holds promise in terms of reducing possible biases and uncertainty compared with approaches using a single correlation measurement per patch, provided an accurate speckle decorrelation model exists. Chapter 3 also showed how combinatorial optimisation could be used to resolve directional ambiguities in a robust manner given an elevational speckle decorrelation model and multiple distance estimates derived from it. To different extents, these approaches rely on the assumption that the elevational speckle decorrelation model is unbiased over the entire range of elevational separations for which it is defined.

The aim of the research presented in the current chapter was to integrate the ideas proposed in the rest of the thesis in order to perform the estimation of out-of-plane transducer motion from imagery of arbitrary media. The generalisation of the probabilistic data fusion approach of Chapter 4 to imagery of real tissue is non-trivial because currently available adaptive speckle decorrelation models, including the one presented in the previous chapter, present imperfections which lead to biased distance estimates in certain portions of the distance domain. Thus, some sort of automated measurement selection process is needed in order to avoid including these biased measurements in the estimation of trajectory.

Measurement selection in the context of image-based ultrasound trajectory estimation from imagery of real tissue is a problem that has previously been discussed by Housden et al. [94]. They proposed a method for recovering transducer trajectory which calculates a robust average of trajectory estimates obtained from interleaved subsets of the image frames. Considerable effort was invested in optimising the unique assignment of frames to independent interleaved subsets. In particular, consecutive frames in a given subset are required not to lie too close to each other in order to avoid the biases which arise from a

poorly represented area of the decorrelation curve [204, 85] (see Chapter 4). Consecutive frames are also required not to lie so far from each other that they would be uncorrelated and thereby not convey any positional information. This approach does not explicitly account for cases like that illustrated in Figure 6.12 of the previous chapter, where widely separated image patches exhibit significant but uninformative correlation. This issue is never, in fact, discussed by the authors, and so it seems that they might not have encountered the problem in their experimental work, or that their robust averaging approach somehow compensates for it.

In order to address such cases where correlation is uninformative and to compensate for some of the inaccuracies in the adaptive speckle decorrelation model of the previous chapter, this chapter builds on the probabilistic speckle decorrelation model and data fusion scheme of Chapter 4 and proposes an iterative algorithm based on statistical hypothesis testing for incrementally integrating new distance estimates to a conservative initial estimate of the trajectory. The hypothesis testing approach is inspired by the mobile robotics literature on the topic of automated map construction. In that context, it is often the case (e.g. [61, 150]) that updates to positional information about a particular landmark are conditional on the statistical agreement between a model established through prior measurement cycles and newly acquired measurements, which ensures that the new measurements indeed correspond to the hypothesised landmark. In the ultrasound trajectory recovery problem addressed here, it is not the identity of particular landmarks (frame pairs) which is called in question, but rather the accuracy of the potentially biased measurement model.

The remainder of this chapter is structured as follows. Section 7.1 explains how the probabilistic speckle model of Chapter 4 can be extended to exploit the learning-based correlation length estimation approach of Chapter 6 and how to use this extension of the model for data fusion. Section 7.2 shows how this adaptive probabilistic data fusion approach can be further extended to allow for principled measurement selection based on an incremental hypothesis testing scheme. Section 7.3 discusses the order in which measurements are considered for integration in the trajectory estimate and presents three different ways of creating the initial trajectory estimate required by the incremental hypothesis testing scheme. Section 7.4 presents the results of trajectory recovery experiments on the animal tissue data sets introduced in the previous chapter. It also describes additional results on further imagery of animal tissue acquired along non-monotonic probe trajectories, thereby verifying the robustness of the combinatorial frame ordering approach proposed in

Chapter 3. The results of these experiments show that the proposed approach generally succeeds in avoiding biased distance estimates resulting from correlation plateaux, leading to more accurate trajectory estimates than a nearest-neighbour approach or the original probabilistic fusion approach of Chapter 4 without measurement selection.

7.1 Probabilistic correlation measurement fusion in real tissue

Consider the problem of estimating the relative elevational positions of a set of otherwise aligned ultrasound image patches. In the terminology introduced in Chapter 3, these corresponding image patches constitute a set of *mini-frames* whose estimated elevational positions will, at a later stage, be combined to those associated with different image patches by enforcing the image planarity constraint. Chapter 4 introduced a probabilistic elevational speckle decorrelation model which describes the statistical properties of the relationship between the sample correlation coefficient ρ_{ij} measured between mini-frames i and j corresponding to image location q and the distance δ between them. This model assumes that the reference and test targets exhibit similar, near Rayleigh, scattering conditions. The variations in ρ captured by this model are those arising from the finite size of the image patches when computing the value of the sample correlation coefficient. In this chapter, this nominal probabilistic speckle decorrelation model is treated as a special case of a more general model which accounts for variations in correlation length arising from non-Rayleigh scattering conditions.

As in the previous chapter, let $\hat{w}_e$ denote the nominal elevational correlation length of the transducer, and let r_e denote the scale factor which must be applied to $\hat{w}_e$ in order to obtain the local elevational correlation length w_e of the medium under consideration, i.e.

$$w_e = r_e \hat{w}_e. \tag{7.1}$$

Let $p(\rho_{ij}|\delta_{ij}, r_{e_{ij}}, q)$ denote the conditional probability density function of the sample correlation coefficient ρ_{ij} measured between mini-frames i and j separated by distance $\delta_{ij} = |z_i - z_j|$ at image location q for a local elevational correlation length $w_{e_{ij}} = r_{e_{ij}} \hat{w}_e$. As in Chapter 4, z_i denotes the elevational position of mini-frame i and $z_0 = 0$. By construction,

$$p(\rho_{ij}|\delta_{ij}, r_{e_{ij}}, q) = p\left(\rho_{ij} \middle| \frac{\delta_{ij}}{r_{e_{ij}}}, r_{e_{ij}} = 1, q\right), \tag{7.2}$$

where the case $r_{e_{ij}} = 1$ on the right hand side corresponds to the nominal probabilistic speckle decorrelation model originally presented in Chapter 4. This nominal model is estimated using Baker's maximum entropy approach [10] using the correlation measurements obtained from the calibration scan of a speckle phantom.

In the case of real tissue, the elevational correlation length of the medium for any mini-frame i is unknown. It is estimated from the normalised statistical features of the image patch, $\mathbf{f}_i = \{r_{R^2{}_i}, r_{S_i}, r_{a_i}, r_{l_i}\}$ using the sparse Gaussian process regressor introduced in Chapter 6. Recall that this regressor provides a probabilistic estimate for the local normalised elevational correlation length r_{e_i}. This estimate is a normal probability density function with mean $\bar{r}_{e_i}$ and variance ξ_i:

$$p(r_{e_i}|\mathbf{f}_i) = \mathcal{N}(\bar{r}_{e_i}, \xi_i), \tag{7.3}$$

where the dependence of r_{e_i} on the synthetic training set $\mathcal{D}$ used to learn the relationship is henceforth made implicit for the sake of notational simplicity.

Let $n+1$ be the number of frames in the image sequence and $\mathbf{z}$ denote the vector obtained by concatenating the z_i, $i = 1, \ldots, n$. Let $\{\delta(\mathbf{z})\}$ denote the set of pairwise distances between corresponding mini-frames, which are deterministically related to their absolute positions by

$$\delta_{ij} = |z_i - z_j|. \tag{7.4}$$

Finally, let $\{\rho\}$ denote the set of pairwise sample correlation coefficients, $\{\mathbf{f}\}$ denote the set of normalised statistical image feature sets and $\{r_e\}$ denote the set of local normalised elevational correlation lengths, obtained for each mini-frame or pair of corresponding mini-frames.

Very much as in Chapter 4, the problem to be solved is to maximise the likelihood of $\mathbf{z}$ for the correlation measurements obtained at image location q given the observed image features $\{\mathbf{f}\}$:

$$\mathbf{z}^* = \underset{\mathbf{z}}{\operatorname{argmax}}\ p(\{\rho\}|\{\delta(\mathbf{z})\}, q, \{\mathbf{f}\}). \tag{7.5}$$

To solve this problem, one must account for the unknown local elevational correlation lengths which are estimated from $\{\mathbf{f}\}$. This is done by expressing the likelihood function

in the above equation as the result of marginalisation over $\{r_e\}$:

$$\mathbf{z}^* = \underset{\mathbf{z}}{\operatorname{argmax}} \int_0^\infty \dots \int_0^\infty p(\{\rho\}|\{\delta(\mathbf{z})\}, \{r_e\}, q, \{\mathbf{f}\}) p(\{r_e\}|\{\mathbf{f}\}, q, \delta(\mathbf{z}\})) d\{r_e\}. \tag{7.6}$$

It is reasonable to assume that the correlation measurements are independent of the statistical image features $\{\mathbf{f}\}$ given the local normalised elevational correlation lengths $\{r_e\}$ and that the $\{r_e\}$ are independent of the probe trajectory and image location. Therefore,

$$\mathbf{z}^* = \underset{\mathbf{z}}{\operatorname{argmax}} \int_0^\infty \dots \int_0^\infty p(\{\rho\}|\{\delta(\mathbf{z})\}, \{r_e\}, q) p(\{r_e\}|\{\mathbf{f}\}) d\{r_e\}. \tag{7.7}$$

If it is further assumed that the elevational correlation length at any given mini-frame i in the trajectory is independent of the elevational correlation length at all other mini-frames $j \neq i$ given the set of measured statistical image features $\{\mathbf{f}\}$, then

$$\mathbf{z}^* = \underset{\mathbf{z}}{\operatorname{argmax}} \int_0^\infty \dots \int_0^\infty p(\{\rho\}|\{\delta(\mathbf{z})\}, \{r_e\}, q) \prod_{i=1}^{n} p(r_{e_i}|\mathbf{f}_i) d\{r_e\}. \tag{7.8}$$

The marginalisation integral required to account for the uncertainty in the r_e in (7.8) above makes this maximum likelihood estimation problem quite difficult. One possible solution would be to use the expectation-maximisation algorithm [46], an iterative approach which was specifically designed to handle optimisation problems of this kind. However, this approach would require an analytical expression for the expected value of $p(\{\rho\}|\{\delta(\mathbf{z})\}, \{r_e\}, q)$ with respect to $p(\{r_e\}|\{\mathbf{f}\})$, which cannot easily be obtained. For simplicity, the following approximation is instead adopted:

$$\int_0^\infty \dots \int_0^\infty p(\{\rho\}|\{\delta(\mathbf{z})\}, \{r_e\}, q) p(\{r_e\}|\{\mathbf{f}\}) d\{r_e\} \approx p(\{\rho\}|\{\delta(\mathbf{z})\}, \{r_e^*\}, q), \tag{7.9}$$

where

$$r_{e_i}^* = \underset{r_{e_i}}{\operatorname{argmax}}\ p(r_{e_i}|\mathbf{f}_i) = \bar{r}_{e_i}. \tag{7.10}$$

This approximation is valid when the variances in the elevational correlation lengths, ξ_i, are small. The maximum likelihood estimate $\mathbf{z}^*$ for $\mathbf{z}$ can then be approximated as

$$\mathbf{z}^* \approx \underset{\mathbf{z}}{\operatorname{argmax}}\ p(\{\rho\}|\{\delta(\mathbf{z})\}, \{r_e^*\}, q). \tag{7.11}$$

If it is assumed that the correlation measurements are conditionally independent of each other given the spatial configuration of the image frames, the local elevational correlation lengths $\{r_e\}$ and the image patch location q, then

$$\mathbf{z}^* \approx \underset{\mathbf{z}}{\operatorname{argmax}} \prod_{i>j} p(\rho_{ij}|\ |z_i - z_j|, r^*_{e_{ij}}, q). \tag{7.12}$$

The variable $r^*_{e_{ij}}$ was created by slight notational and mathematical abuse: indeed, the learning-based approach of Chapter 6 assigns a correlation length estimate to single mini-frames, not pairs of mini-frames. Thus, for a pair of corresponding mini-frames, there are two such estimates. In practice, $r^*_{e_{ij}}$ is thus taken to be the variance weighted average of the mean elevational correlation lengths estimated for mini-frames i and j, as previously done for the distance estimation experiments of Chapter 6:

$$r^*_{e_{ij}} = \frac{\frac{\bar{r}_{e_i}}{\xi_i} + \frac{\bar{r}_{e_j}}{\xi_j}}{\frac{1}{\xi_i} + \frac{1}{\xi_j}}. \tag{7.13}$$

For the purpose of optimising $\mathbf{z}$, the log-likelihood $L(\rho_{ij}|\ |z_i - z_j|, r^*_{e_{ij}}, q) = \ln p(\rho_{ij}|\ |z_i - z_j|, r^*_{e_{ij}}, q)$ is preferred as the objective function, leading to

$$\mathbf{z}^* \approx \underset{\mathbf{z}}{\operatorname{argmax}} \sum_{i>j} L(\rho_{ij}|\ |z_i - z_j|, r^*_{e_{ij}}, q), \tag{7.14}$$

where

$$L(\rho_{ij}|\ |z_i - z_j|, r^*_{e_{ij}}, q) = L\left(\rho_{ij}\Big|\frac{|z_i - z_j|}{r^*_{e_{ij}}}, r_{e_{ij}} = 1, q\right). \tag{7.15}$$

Here, the elevational correlation length $r^*_{e_{ij}}$ is simply used to stretch the nominal likelihood function obtained for speckle phantoms (see Chapter 4) along the δ dimension.

The methodology used to obtain an optimal estimate for $\mathbf{z}$ then closely follows that described in Chapter 4. Namely, a Gaussian estimate of distance $\delta_{ij} = |z_i - z_j|$ is first obtained for each correlation measurement ρ_{ij} such that its mean $\bar{\delta}_{ij}$ and variance σ^2_{ij} are respectively given by

$$\bar{\delta}_{ij} = \underset{\delta_{ij}}{\operatorname{argmax}}\ L\left(\rho_{ij}\Big|\frac{\delta_{ij}}{r^*_{e_{ij}}}, r_{e_{ij}} = 1, q\right) \tag{7.16}$$

and

$$\sigma_{ij}^2 = \frac{-1}{\frac{d^2}{d\delta_{ij}^2} L\left(\rho_{ij} \middle| \frac{\bar{\delta}_{ij}}{r^*_{e_{ij}}}, r_{e_{ij}} = 1, q\right)}. \tag{7.17}$$

By the weighted-least-squares reasoning adapted from [139] and already detailed in Chapter 4 (see section 4.3 for a reminder), an approximate maximum likelihood estimate $\tilde{\mathbf{z}}^*$ for $\mathbf{z}$ is obtained by solving the system of linear equations

$$\mathbf{G}\tilde{\mathbf{z}}^* = \mathbf{b}, \tag{7.18}$$

where

$$G_{ij} = \begin{cases} \sum_{k=0}^{n} \frac{1}{\sigma_{ik}^2} & i = j \\ -\frac{1}{\sigma_{ij}^2} & \text{otherwise} \end{cases} \tag{7.19}$$

and

$$b_i = \sum_{j=0, j\neq i}^{n} \frac{s_{ij}\bar{\delta}_{ij}}{\sigma_{ij}^2}, \tag{7.20}$$

and s_{ij} denotes the sign of the elevational displacement between mini-frames i and j, obtained from the TSP-sort algorithm described in Chapter 3, such that $z_j - z_i = s_{ij}\delta_{ij}$, $j > i$. Also recall that the methodology provides a covariance matrix $\mathbf{C}$ for $\mathbf{z}$:

$$\mathbf{C} = \mathbf{G}^{-1}. \tag{7.21}$$

This covariance matrix expresses the relative uncertainties in the inferred positions of the mini-frames. The main diagonal of $\mathbf{C}$ captures the amount of uncertainty associated with the position of each individual mini-frame while the (covariance) off-diagonal entries of $\mathbf{C}$ represent the degree to which knowledge about the position of each mini-frame would be helpful in reducing the uncertainty in the position of every other mini-frame. This covariance matrix can be used to determine whether a new uncertain measurement of distance between two mini-frames could reasonably have arisen from the particular distribution of $\mathbf{z}$ established from other measurements, as described in the next section.

7.2 Iterative selection and fusion of measurements

In principle, the maximum likelihood approach described in the previous section could be used to combine any number of correlation measurements into a trajectory estimate at a given image patch location. In practice, using all available correlation measurements was found to be highly error prone. This is because the probabilistic data fusion methodology assumes that the speckle decorrelation model is accurate over its entire range. However, as noted before, the adaptive speckle decorrelation model developed in Chapter 6 to account for variations in elevational correlation length is unable to account for the correlation plateaux which are observed for large distances in the presence of elevationally persistent structures (recall the pork flank example of Figure 6.12 from the previous chapter). Depending on the actual micro-structure of the medium, this model can therefore systematically underestimate large distances. Measurements arising from such biased portions of the model must be excluded from the data fusion process in order to accurately recover the transducer trajectory.

In Chapter 4, small biases introduced by unmodelled variations in the elevational correlation length of ultrasound in speckle phantoms were addressed by explicitly imposing thresholds on the minimal acceptable correlation between two mini-frames and the maximum acceptable number of correlation measurements involving a particular mini-frame. Such thresholds are very difficult to tune and generally quite ineffective in the case of real tissue scans because the displacement range over which the adaptive model is biased and the extent to which it is biased are highly subject dependent. In order to be effective in this context, the measurement selection strategy must dynamically adapt to the data.

Suppose for a moment that an unbiased estimate $\tilde{\mathbf{z}}^*$ of the trajectory for patch q, along with its covariance matrix $\mathbf{C}$ were available. Then, the decision to exploit any correlation-based distance estimate δ_{ij}, normally distributed with mean $\bar{\delta}_{ij}$ and variance σ_{ij}^2 respectively given in (7.16) and (7.17), could be made by carrying out a simple statistical hypothesis test. Namely, let the $\mathcal{H}_0$ denote the null hypothesis that the new distance estimate δ_{ij} is consistent with the estimate of $\mathbf{z}$, i.e. $s_{ij}\delta_{ij} = z_j - z_i,\ j > i$, and let $\mathcal{H}_a$ denote the alternative hypothesis that it is biased, i.e. $s_{ij}\delta_{ij} \neq z_j - z_i$. $\mathcal{H}_0$ should be rejected in favour of $\mathcal{H}_a$ at significance level $\alpha_{\mathcal{H}}$ if

$$|s_{ij}\bar{\delta}_{ij} - (\tilde{z}_j^* - \tilde{z}_i^*)| \geq \Phi^{-1}(1 - 0.5\alpha_{\mathcal{H}}, \nu^2), \tag{7.22}$$

where $\Phi^{-1}(1-0.5\alpha_{\mathcal{H}}, \nu^2)$ is the inverse cumulative distribution function of a Gaussian random variable with variance ν^2 evaluated at probability $1-0.5\alpha_{\mathcal{H}}$, and

$$\nu^2 = \sigma_{ij}^2 + C_{jj} + C_{ii} - 2C_{ij}. \tag{7.23}$$

Should $\mathcal{H}_0$ be accepted, an updated and more precise probabilistic trajectory estimate could then be computed that includes the information contained in the newly accepted distance estimates by solving (7.18) and (7.21). This suggests an incremental, iterative approach to the fusion of correlation measurements. Given an initial trajectory estimate $\mathbf{z}_{[0]}$, iteration t proceeds as follows:

1. A new subset of the available correlation measurements are tested against the current estimate of the trajectory, rejecting candidate distance estimate δ_{ij}, $j > i$, if

$$\begin{aligned} |s_{ij}\bar{\delta}_{ij} - (\tilde{z}^*_{[t-1]_j} - \tilde{z}^*_{[t-1]_i})| &\geq \Phi^{-1}(1-0.5\alpha_{\mathcal{H}}, \nu^2_{[t-1]}), \\ \nu^2_{[t-1]} &= \sigma_{ij}^2 + C_{[t-1]_{jj}} + C_{[t-1]_{ii}} - 2C_{[t-1]_{ij}}. \end{aligned} \tag{7.24}$$

2. The measurements accepted by the above test are fused with all other previously accepted measurements to obtain the trajectory estimate $\mathbf{z}_{[t]}$ by solving (7.18) and (7.21), thereby forming the basis for the null hypothesis to be tested at iteration $t+1$.

Through such an iterative approach, it can be hoped that a robust, increasingly accurate and increasingly precise estimate of the trajectory will be obtained for every set of miniframes. This will depend on the limitations intrinsic to the adaptive speckle decorrelation model, but also on three controllable factors:

1. The order in which the measurements are tested and integrated to the trajectory estimate (hereafter called the *testing schedule*).

2. The initial probabilistic trajectory estimate $\mathbf{z}_{[0]}$.

3. The choice of significance level $\alpha_{\mathcal{H}}$.

The choice of significance level is the most trivial of these parameters. It was found empirically that, for a given initial estimate of trajectory combined with a given testing schedule, $\alpha_{\mathcal{H}} = 0.1$, $\alpha_{\mathcal{H}} = 0.05$ and $\alpha_{\mathcal{H}} = 0.01$ all gave similar results, though these seemed slightly

more accurate for $\alpha_{\mathcal{H}} = 0.01$. Therefore, $\alpha_{\mathcal{H}} = 0.01$ was retained for the experiments presented in this chapter. The other two factors, testing schedule and initial estimate, are of more consequence, and are discussed in more detail in the next section.

7.3 Testing schedule and initial trajectory estimate

At a given iteration t, measurements are tested against the current trajectory estimate using (7.24) and, if accepted as consistent, integrated into the model, thereby creating a new trajectory estimate for iteration $t + 1$. The order in which different subsets of measurements are tested, or testing schedule, influences the outcome of the procedure through the evolution of the hypothesis tests and model updates over time. In this thesis, a very simple testing schedule is adopted: the first iteration of the process considers only those correlation measurements which concern consecutively acquired mini-frames (if they are not already included in the initial trajectory estimate). The second iteration then considers correlation measurements between frames which are two acquisition steps apart, and so on until no new measurements are accepted by the hypothesis test. This testing schedule is illustrated schematically in Figure 7.1. Aside from its ease of implementation, this testing schedule has two advantages:

1. Assuming an adequate initial trajectory estimate, an accurate and reasonably certain estimate of $\mathbf{z}$ has already been formed by the time distance underestimates arising from correlation plateaux are processed, allowing the hypothesis test to reject them with high probability. This is because these frames are quite widely spaced and widely spaced frames can be assumed not to be acquired within a small number of acquisition steps of each other. In contrast, unbiased distance estimates between mini-frames which are far apart in order of acquisition, but close in space (such as may occur after a change in the direction of the trajectory) are more likely to agree with the current estimate of $\mathbf{z}$ and be included in the next probabilistic trajectory update.

2. The variance and covariance terms associated with each mini-frame are updated at a similar rate to those associated with other mini-frames, thereby preventing the outcome of the data fusion process to become entirely dependent on a tiny subset of mini-frame positions that were prematurely constrained by multiple measurements

referring to them. The proposed testing schedule favours even updating across mini-frames because at any given iteration $t > 1$, any particular mini-frame can only be referred to by a single new measurement. For $t = 1$ there can be a maximum of two measurements per mini-frame (one leading to it and one departing from it).

The iterative measurement selection and data fusion scheme described in this chapter was derived assuming that a probabilistic trajectory estimate $\mathbf{z}_{[0]}$ was available as a starting point. In the absence of an external source of data (such as an external position tracking device, or prior knowledge concerning the type and speed of the transducer trajectory), this initial estimate must be obtained from a subset of the correlation measurements derived from the images. An initial trajectory estimate for a set of $n + 1$ mini-frames requires a minimum of n correlation measurements which must relate every mini-frame to at least one other such that there is a connected measurement path from any mini-frame to every other. An example of a minimal initial set of measurements is given in Figure 7.1 (for $t = 0$). More formally, one can think of the mini-frames and available correlation measurements as forming the vertices and edges of a graph $\mathcal{G}$. In this context, a minimal subset of measurements allowing the estimation of a trajectory among the mini-frames is a spanning tree of $\mathcal{G}$. A *minimal* spanning tree of $\mathcal{G}$, obtained easily using any of the classic algorithms (Kruskal's algorithm was used here) [26], thus represents an optimal initial trajectory estimate with respect to some cost function κ defined over the edges (i.e. pairs of connected vertices) of $\mathcal{G}$.

In developing the iterative measurement selection and fusion scheme of Section 7.2, an imaginary, unbiased initial trajectory estimate was assumed. The question of which cost function to use in order to obtain such an unbiased initial trajectory estimate is a difficult one in a context where there is little formal information regarding the correctness of the dynamic speckle decorrelation model over the range of correlation or distance. In this thesis, three alternative cost functions are considered. The first of these is given by

$$\kappa_1(i, j) = |i - j|. \tag{7.25}$$

When a valid Gaussian distance estimate (i.e., with a positive value of σ_{ij}^2 in (7.17)) exists for all ρ_{ij} between consecutively acquired mini-frames, the minimum spanning tree associated with this cost function yields a nearest-neighbour description of the trajectory. The correlation measurements corresponding to this nearest-neighbour reconstitution can

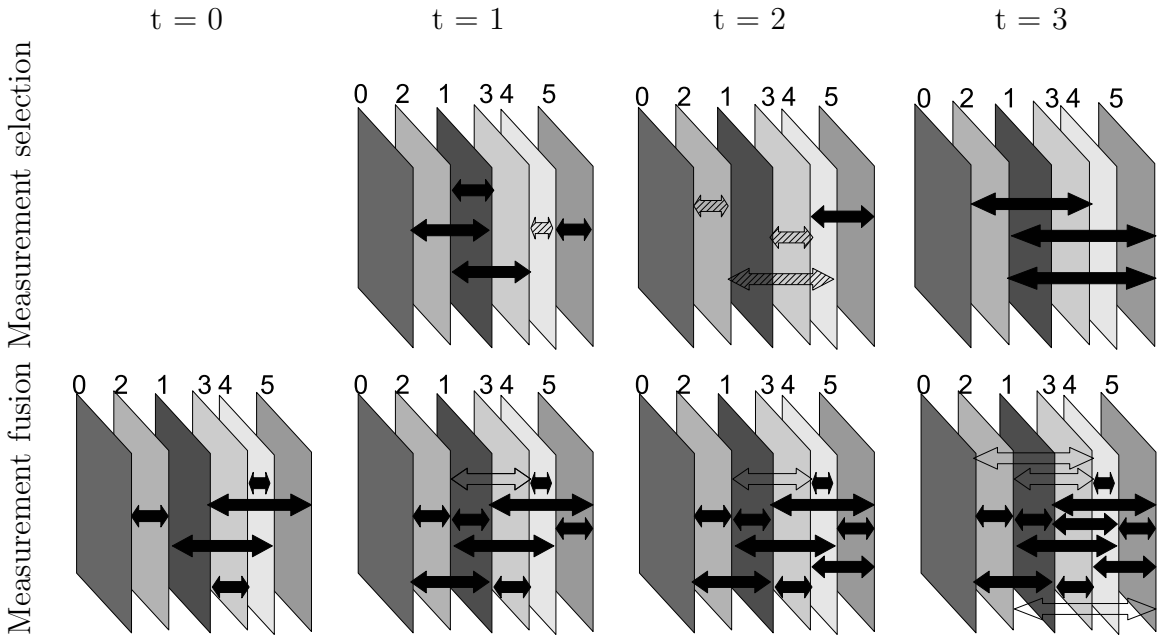


Figure 7.1 Schematic illustration of the chosen testing schedule. The planes represent a set of corresponding mini-frames and the numbers next to each one represent the order in which they were acquired (in this example, the trajectory is not monotonic). The top row shows the distance estimates which are considered as candidates for measurement selection and the bottom row shows the set of distance estimates which are used (solid arrows) or not used (hollow arrows) in each data fusion step. A minimal subset of the available distance estimates is used to initialise the iterative algorithm at iteration $t = 0$. At iteration $t = 1$, distance estimates between mini-frames separated by one acquisition step become candidates for fusion to this trajectory estimate. The hatched distance estimate relating mini-frame 3 to mini-frame 4 is not tested because it was already used to build the initial estimate at $t = 0$. All but one of the candidate distances for $t = 1$ are accepted by the hypothesis test. The estimate that was discarded is shown as an empty arrow in the measurement fusion step. At iteration $t = 2$, distance estimates separated by two acquisition steps are considered. Most of them were already used, but one is tested and accepted. Candidates are selected in the same fashion for iterations $t = 3$ until $t = 5$ (the maximum number of acquisition steps between any two frames) or until no new measurements are accepted.

be probabilistically fused by solving (7.18) and (7.21). This choice of cost function may at first appear to be a poor one because the initial trajectory estimate it provides is, in general, not unbiased owing to the measurements lying on a poorly represented area in the decorrelation model (see Chapter 4). However, it should be noted that in practice, the initial trajectory estimate need not necessarily be unbiased for the proposed iterative scheme to work well. This is because the outcome of the first hypothesis test of (7.22) depends not only on the mean trajectory estimate, but also on its covariance matrix $\mathbf{C}_{[0]}$. In the case of the nearest-neighbour trajectory estimate, this covariance matrix is assigned very large entries by (7.17) owing to the flatness of the log-likelihood function for these high correlation measurements. The initial model then remains heavily biased but is also highly uncertain. This makes the hypothesis test of (7.24) a very conservative one, in the sense that it will not often reject $\mathcal{H}_0$, allowing more certain, less biased measurements to dilute the contributions made by the initial trajectory estimate early on in the data fusion process. As unbiased measurements are added and constrain the range of possible trajectories, the entries on the main diagonal of $\mathbf{C}_{[t]}$ should decrease and the hypothesis test of (7.24) should become more discriminative.

While this simple cost function can work reasonably well in the context of the proposed testing schedule, it can certainly be improved upon. For one thing, the feature which renders it effective, i.e., the small distance to variance ratio of the measurements it uses, can be made explicit in the mathematical definition of the cost function. This would allow measurements arising from frames which are very distant in terms of acquisition time, but very close in terms of location (which may occur in non-monotonic scans), to be advantageously included in the initial trajectory estimate.

More importantly, there can be serious problems with the cost function κ_1 in that it allows overestimates of distance, which sometimes occur for short to moderate distances (whether due to higher noise content in the presence of low signal levels or to otherwise inadequate modelling assumptions), to enter the model too easily, due to the almost invariably large variance of the initial reconstruction. Once such overestimates are integrated into the trajectory estimate, the variance decreases and unbiased or underestimated distance measurements arising from pairs of mini-frames separated by longer distances, considered later in the testing schedule, tend to be rejected by the hypothesis test. As a result, the speed of the transducer may be severely over-estimated in some areas, causing substantial drift. The two other cost functions considered in this thesis attempt to address this issue.

They are defined as:

$$\kappa_2(i,j) = \text{sgn}(r^*_{e_{ij}} - 1)\bar{\delta}_{ij}, \tag{7.26}$$

$$\kappa_3(i,j) = \text{sgn}(r^*_{e_{ij}} - 1)\frac{\bar{\delta}_{ij}}{\sigma^2_{ij}}, \tag{7.27}$$

where

$$\text{sgn}(a) = \begin{cases} -1 & a < 0 \\ 1 & \text{otherwise.} \end{cases} \tag{7.28}$$

For areas of the scan where the estimated local elevational correlation length is larger than the nominal correlation length estimated from the reference phantom, these cost functions act in a very similar way to κ_1, preferring short distance measurements, or measurements with a small distance to variance ratio. This preference changes for areas where the local elevational correlation length is estimated to be *shorter* than the nominal one. In an ideal world, this ought to be a very rare occurrence since theory dictates that the nominal decorrelation model should represent the minimal amount of correlation between neighbouring imagery caused only by the finite elevational beam-width of the transducer. However, as noted in the previous chapter, the lower signal levels obtained when imaging certain types of tissue can imply decreases in the measured correlation, sometimes yielding correlation levels below those predicted by the nominal model, leading to over-estimation of distances. As discussed in Chapter 6, the learning-based adaptive speckle decorrelation model is not entirely able to compensate for such effects because they are not well represented in the training imagery used to build it. Nonetheless, it was observed that when the true correlation length is smaller than the nominal one, it is often able to partially, if not accurately, correct the discrepancy, yielding $r_e < 1$. While such small values of r_e are generally not very accurate, they do provide an indication that distance over-estimates are likely to occur in the corresponding region of the scan. In these cases, the κ_2 and κ_3 cost functions select edges from the graph which correspond to larger distances or larger distance to variance ratios. The effects of this choice are twofold: (1) the very large distance measurements chosen are less likely to be over-estimates than moderate ones and (2) the large distance measurements (or larger distance to variance ratios) chosen are likely to be associated with a smaller variance than the short distance measurements privileged by κ_1, which prevents the hypothesis test from accepting many measurements, especially distance

over-estimates, related to such error-prone areas of the scan.

7.4 Experiments

The proposed iterative measurement selection and fusion scheme was tested on ultrasound imagery of animal tissue. A nominal probabilistic speckle decorrelation model was estimated from one of the phantom ultrasound image sequences described in Chapter 4 with frames regularly spaced at 0.05 mm intervals and the sparse Gaussian regressor obtained in the experiments of the previous chapter was used to determine the mapping between statistical image features and elevational correlation length. For each of the animal tissue sequences, the probe trajectory was estimated following an adaptation of the basic scheme described in Chapter 3. Readers may find useful to remind themselves of this scheme by consulting the schematic example given in Figure 3.2 of Chapter 3. According to this scheme, trajectory recovery proceeded as follows:

1. For every image patch, the local elevational correlation length was estimated using the learning-based model described in Chapter 6.

2. The correlation coefficients between all pairs of corresponding image patches were computed, and a deterministic estimate of distance was obtained for all correlation values greater than 0.1 using the piecewise linear deterministic elevational speckle decorrelation model, scaled to account for the local elevational correlation length.

3. The signs of these distance estimates were estimated using the TSP-sort algorithm described in Chapter 3.

4. An elevational trajectory associated with each image patch location was estimated using the iterative measurement selection and fusion scheme proposed in this chapter, allowing the algorithm to select measurements among the set of all distance estimates deemed shorter than $w_{e_{ij}} = r^*_{e_{ij}} \hat{w}_e$ by the adapted piecewise linear speckle decorrelation model. For comparison, patch-wise trajectories were also estimated using the two base-line algorithms to be described shortly.

5. Because the algorithms described in this and the previous chapters generally gave a noisier cloud of positional data than was the case for the speckle phantom experiments

of Chapter 4, the data corresponding to the estimated positions of each patch centre were smoothed using a median filter (as suggested in [72]) in order to ease the task of fitting rigid transformations from each frame to the first.

6. Rigid transformations relating every frame to the first were computed using the three step robust fitting approach presented in Chapter 3. In order to prevent spurious mini-frame ordering reversals due to the increased level of noise, the number of times that an image patch can be labelled as an outlier without there being an attempt to reverse its ordering was raised to 10% of the number of frames instead of the 5% used in the phantom experiments.

Three versions of the proposed approach were considered in these experiments, corresponding to each of the cost functions defined in (7.25), (7.26) and (7.27) used to build the initial trajectory estimate for each image patch. These three versions of the proposed approach are henceforth respectively labelled HYP-K1, HYP-K2 and HYP-K3. These were compared to two different base-lines. The first is the nearest-neighbour (NN) approach, which uses only deterministic distance estimates from consecutive pairs of mini-frames and has already been used as a base-line algorithm in Chapter 4. The second base-line algorithm is the result of applying the probabilistic fusion approach of Section 7.1 to all distance estimates deemed shorter than $w_{e_{ij}} = r^*_{e_{ij}} \hat{w}_e$ by the adapted piecewise linear speckle decorrelation model, without any further measurement selection. This approach will henceforth be labelled NMS (no measurement selection).

Two distinct sets of experiments were carried out. The first set of experiments, whose results are described in Section 7.4.1, involved the 51 ultrasound image sequences of animal tissue already described in Chapter 6 and used to validate the theory therein. These image sequences were acquired along purely monotonic trajectories. The second set of experiments involved additional ultrasound image sequences of animal tissue samples acquired along chaotic, non-monotonic translational trajectories. The test data and results for this second set of experiments are discussed in Section 7.4.2.

7.4.1 Results for monotonic trajectories

In order to quantitatively evaluate the accuracy of the proposed iterative measurement selection and fusion algorithms, this first set of experiments was carried out on the 51 ultrasound image sequences of pork tenderloin, turkey breast and beef brisket from the

previous chapter. The acquisition of these image sequences was performed using the same sub-millimetre positioning device as was used in the rest of this thesis (see Figure 4.12 in Chapter 4) to control the spacing between the images with high precision; these spacing measurements are used as ground truth for the evaluation of the accuracy associated with the different methods tested. A complete set of results for these experiments can be found in Appendix D, which provides plots comparing the progression of the mean target registration error over time as well as plots comparing the displacement, yaw and tilt components of the estimated trajectories (using the centroid of the grid of image patches as the centre of rotation) to those of the true trajectory for the three versions of the proposed approach as well as the two base-line approaches. Relevant examples from this complete set will be discussed over the course of this section for illustrative purposes.

For the purpose of statistical analysis, the accuracy of the different trajectory estimates was summarised by the mean target registration error measured at the mini-frame centres, averaged over all the frames and normalised by the length of the trajectory. The latter normalisation enabled meaningful comparisons of algorithms across results obtained for trajectories of differing lengths. As a coarse summary of the results, Figure 7.2 shows a boxplot of this error measure, with cases grouped by tissue type (pork, turkey or beef) and by algorithm. From this boxplot, it is immediately apparent that the probabilistic measurement fusion approach did very poorly when no measurement selection (NMS) was applied. This is because the trajectory estimates obtained by this approach were most often overwhelmed with distance underestimates arising from correlation plateaux. Accuracy differences between the other algorithms are not as obvious in the boxplot because accuracy varied substantially across the different image sequences. Note, however, that for the three types of tissue, the median of the normalised average mTRE is consistently smaller for the HYP-K2 and HYP-K3 algorithms than for the NN algorithm, and that it is always smaller than, or else very similar to, the median error for the HYP-K1 algorithm.

In order to properly quantify the effect of the choice of algorithm (or cost function) and tissue type on the accuracy of trajectory recovery, a mixed within/between-subject factor analysis of variance (ANOVA) was carried out. The data were grouped by transducer speed (0.03, 0.04, 0.05, 0.06, 0.08 or 0.1 mm/step) and by tissue type (pork tenderloin, turkey breast or beef). The effects of algorithm choice and tissue type were evaluated, as well as all combinations of these. Where significant effects were detected, pairwise comparisons between algorithms or combinations of algorithms and tissue types were carried out us-

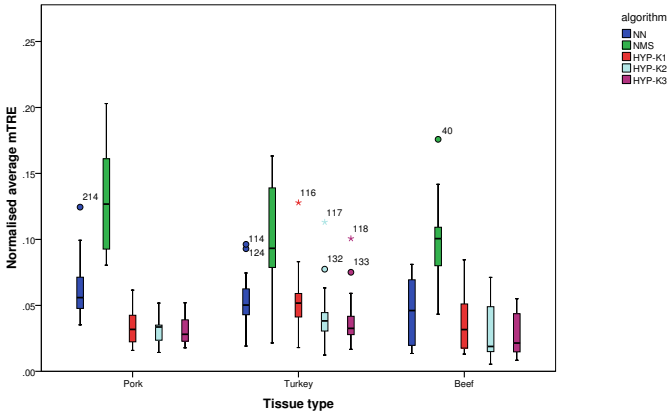


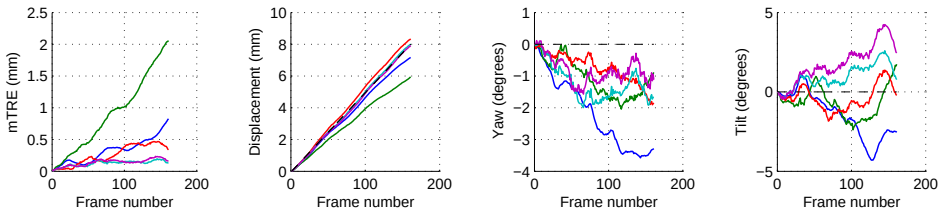
Figure 7.2 Boxplot of the normalised average mTRE between the true and estimated trajectories for imagery of animal tissue for the NN, NMS, HYP-K1, HYP-K2 and HYP-K3 approaches. The statistics were grouped by tissue type.

ing the Bonferroni correction to adjust the analysis for multiple comparisons. Statistical significance was evaluated at the 5% level.

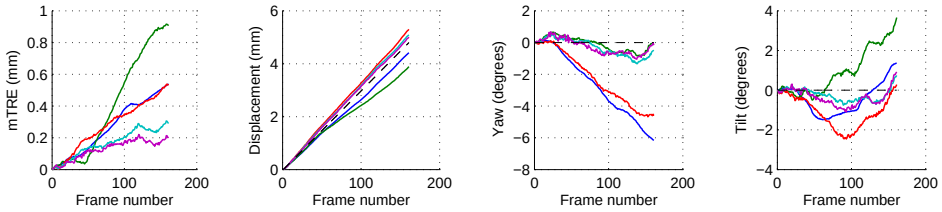
Not surprisingly, the NMS approach was significantly outperformed by all alternatives tested ($p < 0.001$ in all cases). Overall, the HYP-K2 and HYP-K3 versions of the proposed measurement selection approach outperformed both the NN approach and the HYP-K1 approach ($p < 0.001$ in all cases). The overall accuracy of the HYP-K1 approach was slightly better than that of the NN approach, but the difference was not found to be statistically significant ($p = 0.251$).

These results are well illustrated by the examples shown in Figure 7.3, which displays the mean target registration error, displacement, yaw and tilt estimates obtained for scans of beef brisket and pork tenderloin. In both examples, the HYP-K2 and HYP-K3 versions of the proposed algorithm gave excellent results. Though the NN and and HYP-K1 algorithms may exhibit similar overall accuracy in terms of the mTRE, these examples show that they behave quite differently. Indeed, the trajectories they have estimated are qualitatively very different from one another: the NN method underestimated the displacement of the probe whereas the HYP-K1 approach overestimated it. In both cases, this is a general tendency, as can be assessed from the full set of results given in Appendix D.

The tendency of the NN algorithm to underestimate distances stems from the fact that



(a) Beef brisket 7: 161 frames, 0.05 mm intervals



(b) Pork tenderloin 3: 161 frames, 0.03 mm intervals

Figure 7.3 Examples illustrating the typical behaviour of the NN, NMS, HYP-K1, HYP-K2 and HYP-K3 algorithms. From left to right, the mTRE, and estimated displacement, yaw and tilt evaluated using the centroid of the grid of image patches as the centre of rotation. Dashed line: Ground truth; Blue: NN method, green: NMS method, red: HYP-K1 method; cyan: HYP-K2 method; magenta: HYP-K3 method. Note how the HYP-K1 and NN methods have similar mTREs over time, but qualitatively different trajectories associated with them. Also note how the HYP-K2 and HYP-K3 algorithms are able to compensate for the flaws of each.

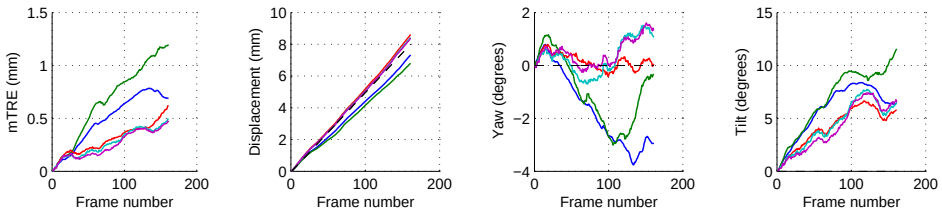
it uses measurements from an area of the decorrelation model which is poorly represented by a piecewise linear curve, as already discussed in Chapter 4. The tendency of the HYP-K1 algorithm to overestimate distances comes from the fact that it easily allows overestimates of distance into its model of the trajectory because its initial estimate of it is under-constrained (large variances). The overestimates of distance are due to limitations in the current version of the adaptive elevational speckle decorrelation model described in Chapter 6, whereby unmodelled noise and variations in the shape of decorrelation curves can affect both the elevational correlation length and the in-plane statistical image features. Algorithms HYP-K2 and HYP-K3, which used more sophisticated cost functions to compensate for this limitation, both gave more accurate results.

Statistically significant differences in accuracy were also found in relation to the type of tissue under observation. Overall, trajectory recovery results were less accurate for turkey

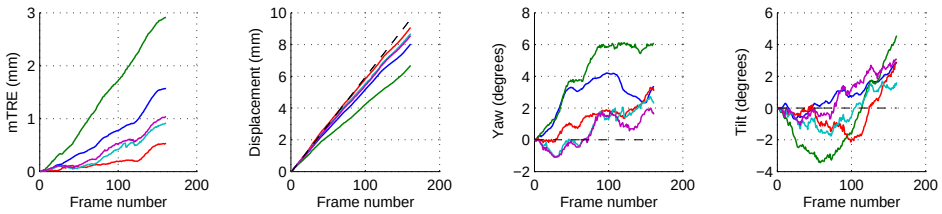
breast data sets than for beef brisket data sets ($p = 0.042$). This is possibly related to the comparatively small amount of echogenic tissue (such as fat) in many of the turkey breast scans, which increased the importance of the contributions made by noise to the image and degraded the accuracy of the learning-based speckle decorrelation model. There was generally more echogenic structure to be seen in the beef and pork scans, possibly because of the higher fat content of the meat.

The HYP-K1 version of the proposed approach was particularly sensitive to differences in tissue type, with significantly worse performance for the turkey data sets than for the beef ($p = 0.044$) or the pork ($p = 0.020$) data sets. Its best performances were achieved with pork data, where it often approached the performance of the HYP-K2 and HYP-K3 algorithms, and occasionally outperformed them, as illustrated by the examples of Figure 7.4. In such cases, the κ_2 and κ_3 cost functions yielded very similar initial trajectory estimates to the κ_1 cost function. The HYP-K2 and HYP-K3 approaches were generally able to maintain competitive performance across different types of tissue, whereas the HYP-K1 approach broke down when the adaptive speckle decorrelation model overestimated correlation length, as often happened with some of the turkey and beef scans. The examples of Figures 7.3 and 7.5 illustrate this. Pork tenderloin was the only one of the three tissue types for which the HYP-K1 approach significantly outperformed the NN approach ($p = 0.013$).

Though they generally gave good results, there were, nonetheless, cases where algorithms HYP-K2 and HYP-K3 were not completely able to overcome the inaccuracies of the adaptive speckle decorrelation model. One such case is illustrated in Figure 7.6. Distances were quite severely over-estimated for this scan. Whether this was due to the low echogenicity properties of the sample or to the lack of fit of the Gaussian model to the axial and lateral decorrelation curves, the κ_2 and κ_3 cost functions probably failed to protect the trajectory estimate against distance overestimates because there were not many image patch pairs for which $r_e < 1$. $r_e < 1$ is a good indicator that the decorrelation model might overestimate distances because it implies that the true elevational correlation length was probably shorter than the nominal one, a case poorly modelled by the training set. However, it is not always the case that $r_e < 1$ in the presence of such modelling inaccuracies, precisely because there were few cases where $r_e < 1$ from which to learn in the training set. When r_e is overestimated and is also greater than 1, the cost functions κ_2 and κ_3 fail to compensate for the problem. In the example of Figure 7.6, this effect is quite extreme.



(a) Pork tenderloin 7: 161 frames, 0.05 mm intervals



(b) Pork tenderloin 11: 161 frames, 0.06 mm intervals

Figure 7.4 Pork tenderloin cases where the HYP-K1 approach gave very good results, as did the HYP-K2 and HYP-K3 approaches. From left to right, the mTRE, and estimated displacement, yaw and tilt using the centroid of the grid of image patches as the centre of rotation. Dashed line: Ground truth; Blue: NN method, green: NMS method, red: HYP-K1 method; cyan: HYP-K2 method; magenta: HYP-K3 method.

Turkey breast 12: 161 frames, 0.06 mm intervals

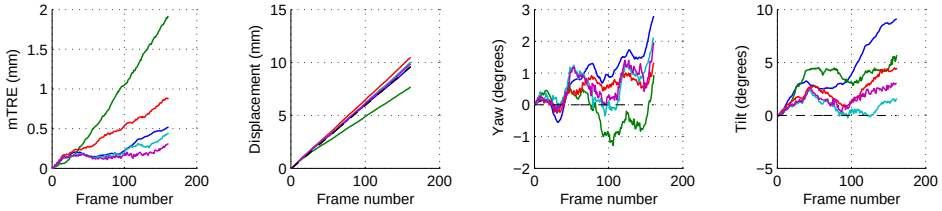


Figure 7.5 Turkey breast case where the HYP-K2 and HYP-K3 approaches gave very good results and outperformed the HYP-K1 approach. From left to right, the mTRE, and estimated displacement, yaw and tilt at the centroid of all image patches. Dashed line: Ground truth; Blue: NN method, green: NMS method, red: HYP-K1 method; cyan: HYP-K2 method; magenta: HYP-K3 method.

All algorithms overestimated displacement except the NMS algorithm, which did uncharacteristically well, compensating for the overestimates of moderate distances by including in its trajectory estimate a large number of underestimates of large distances.

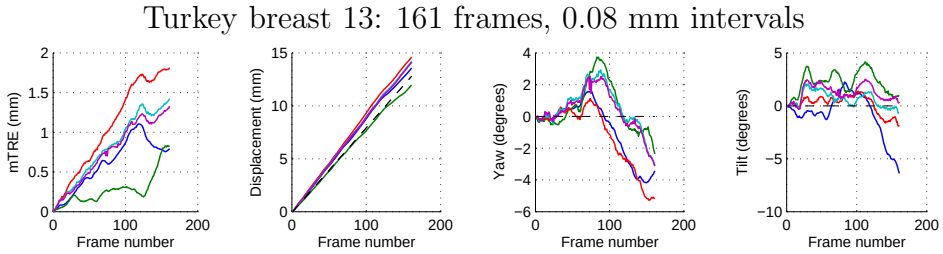


Figure 7.6 Example result where the HYP-K2 and HYP-K3 approaches failed to compensate for distance overestimates and were outperformed by the NMS approach. The example is from a linear scan of turkey breast with 0.08 mm frame intervals. From left to right, the mTRE, and estimated displacement, yaw and tilt measured using the centroid of the grid of image patches as the centre of rotation. Dashed line: Ground truth; Blue: NN method, green: NMS method, red: HYP-K1 method; cyan: HYP-K2 method; magenta: HYP-K3 method.

A less extreme example of displacement over-estimation by algorithms HYP-K2 and HYP-K3 is shown in Figure 7.7. This time, elevational correlation length was overestimated, giving the NN approach, which tends to underestimate distance, an advantage. An interesting observation regarding this example is that although algorithms HYP-K2 and HYP-K3 overestimated transducer speed at the beginning of the image sequence, they eventually adjusted to the correct speed near frame 85. This can be seen by the flatness of the mTRE curves for HYP-K2 and HYP-K3 thereafter, as well as in their displacement curves which finish very nearly parallel to the ground truth. In this case, the κ_2 and κ_3 cost functions were partially able to overcome the limitations of the κ_1 cost function and gave better, though certainly not perfect, results.

A final remark about the results of these experiments concerns the angular errors (yaw and/or tilt) incurred by the various trajectory recovery algorithms. In general, the NN and NMS methods appear to yield the worst rotational errors, with the HYP-K2 and HYP-K3 algorithms being better behaved. There are, however, a few of cases where large angular errors are incurred by all approaches, such as in the example given in Figure 7.8. Such examples illustrate the fact that the accuracy of the mini-frame trajectories corresponding to different image locations is uneven. The accuracy of the proposed algorithms might depend on the local characteristics of the ultrasound signal in the different image patches. The scheme devised in Chapter 3 to enforce the image planarity constraint is somewhat

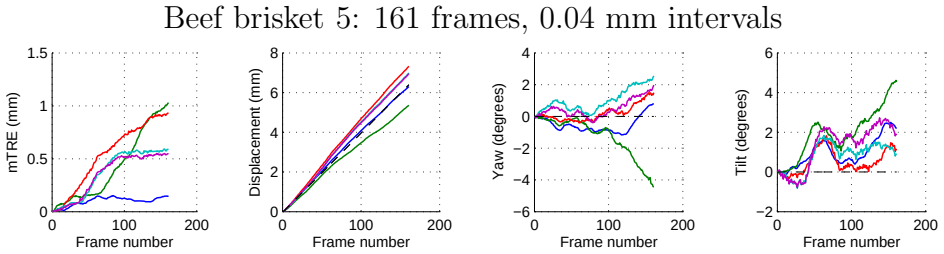


Figure 7.7 Example result from a scan of beef brisket where the NN method outperforms the iterative measurement selection approaches. From left to right, the mTRE, and estimated displacement, yaw and tilt measured using the centroid of the grid of image patches as the centre of rotation. Dashed line: Ground truth; Blue: NN method, green: NMS method, red: HYP-K1 method; cyan: HYP-K2 method; magenta: HYP-K3 method.

vulnerable to non-uniformities in accuracy because it merely fits rigid transformations from a set of patches at $z = 0$ to the z values obtained by patch-wise trajectory estimation over the entire scan. The accumulation of error in a cluster of neighbouring patch-wise trajectory estimates can lead to rotational errors in the estimated rigid transformations. The approach might be improved by re-adjusting the z values for frames $i + 1, \ldots, n$ to account for the z positions implied by the rigid transformation fitted to frame i, thereby more strongly coupling the image planarity constraint to the patch-wise trajectory estimates.

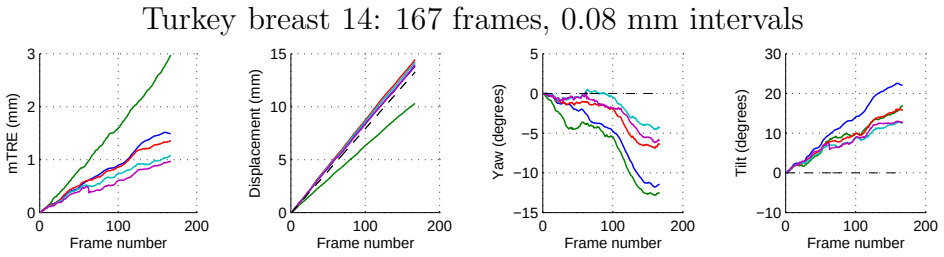


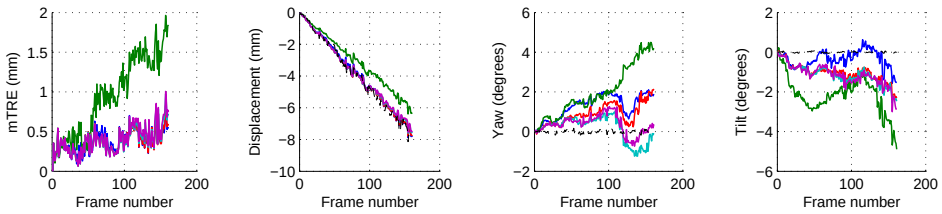
Figure 7.8 Turkey breast case for which all algorithms yielded large angular errors. From left to right, the mTRE, and estimated displacement, yaw and tilt measured using the centroid of the grid of image patches as the centre of rotation. Dashed line: Ground truth; Blue: NN method, green: NMS method, red: HYP-K1 method; cyan: HYP-K2 method; magenta: HYP-K3 method.

7.4.2 Results for non-monotonic image sequences

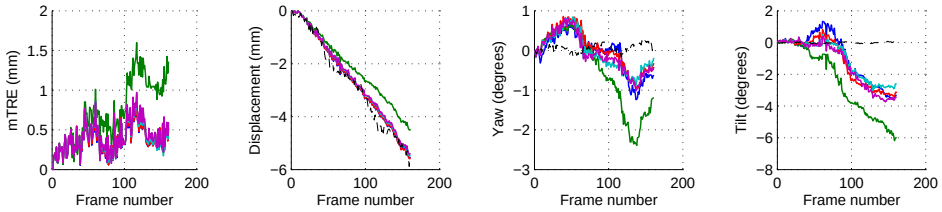
This chapter described a methodology to discard biased distance estimates caused by inaccuracies in the adaptive elevational speckle decorrelation model of Chapter 6 within the probabilistic patch-wise trajectory estimation method described in Chapter 4. The experiments of the previous section showed that this measurement selection methodology was necessary: indeed, the original formulation of the probabilistic trajectory estimation method (i.e., the NMS approach) did not work well at all on image sequences of real tissue. However, these experiments were limited in that the probe motion was always monotonic and so the assumption that the correct signs for displacement estimates would be available from the TSP-sort algorithm was not tested. This assumption is quite a strong one, given that the TSP-sort algorithm itself relies on the speckle decorrelation model to estimate these signs. This section describes the results of further experiments which show that this assumption holds true in the case of chaotic probe trajectories.

At the time these experiments were carried out, the probe bearing arm of the sub-millimetre positioning device could be loosened or tightened by inserting or removing a mechanical wedge between the stage and its supporting metallic frame (see Figure 4.12). With the setup thus loosened, 9 ultrasound image sequences of each of pork and beef tenderloin samples were acquired for which the motion of the transducer was constrained to remain along the elevational axis but was punctuated by chaotic, uncontrolled direction changes. The speed of the transducer varied between approximately 0.03 mm/step and 0.08 mm/step. The trajectory of the probe was estimated for each of the 18 resulting image sequences using the NN, NMS, HYP-K1, HYP-K2 and HYP-K3 approaches and compared to the trajectory measured by a Polaris optical tracker attached to the probe. Appendix E provides the full set of results for these experiments by showing plots of the mean target registration error and of the estimated displacement, yaw and tilt evaluated using the centroid of the image patches as the centre of rotation. Sample results are duplicated in Figure 7.9.

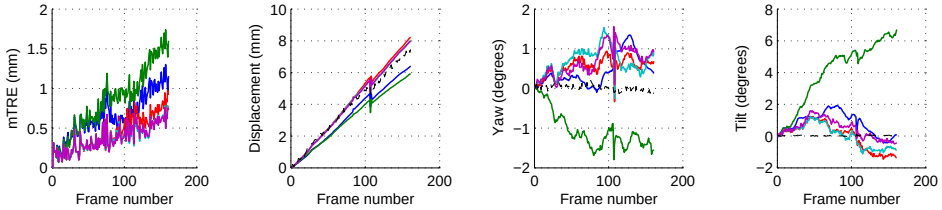
The first observation that can be made about the results shown in Figure 7.9 (and all those of Appendix E) is that the TSP-sort algorithm has worked very well. For all trajectory recovery algorithms, the estimated displacement curves display dips and spikes at the same locations and these most often correspond very well to similar dips and spikes in the measurements made by the Polaris optical tracker. Thus, it appears that the TSP-sort



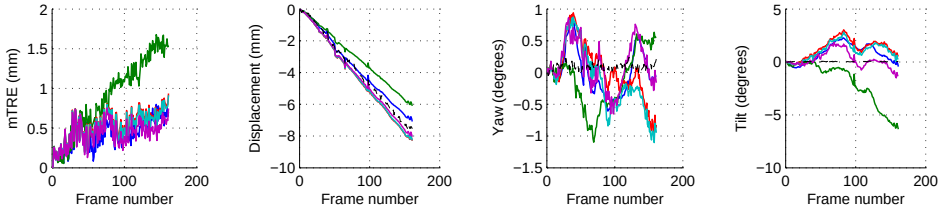
(a) Chaotic pork tenderloin 5



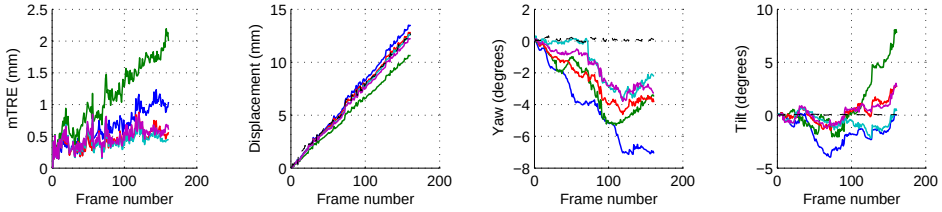
(b) Chaotic pork tenderloin 6



(c) Chaotic beef tenderloin 1



(d) Chaotic beef tenderloin 2



(e) Chaotic beef tenderloin 8

Figure 7.9 Sample trajectory estimation results for non-monotonic image sequences of pork and beef tenderloin. From left to right, the mTRE, estimated displacement, yaw and tilt using the centroid of the grid of image patches as the centre of rotation. Dashed line: Polaris tracker measurements; Blue: NN method, green: NMS method, red: HYP-K1 method; cyan: HYP-K2 method; magenta: HYP-K3 method.

algorithm is very robust to the effects of the unmodelled correlation plateaux which occur at large distances and is able to provide correct mini-frame orderings even in the presence of this modelling inaccuracy. This suggests that the combinatorial optimisation algorithm draws most of the information relevant to ordering from constraints between sets of shorter distance estimates.

It also appears from the results that once again, the NMS approach, which does not implement any form of measurement selection, is much less accurate than the other alternatives tested. In order to quantitatively compare the performance of the different algorithms, the mean target registration errors obtained for each algorithm were averaged over all frames and normalised by the approximate length of each sequence (evaluated by multiplying the number of increments by the nominal increment size). A boxplot of the resulting data, grouped by algorithm, is shown in Figure 7.10. A repeated measures analysis of variance (ANOVA) was carried out on these normalised average mTRE data, revealing that the choice of algorithm had a statistically significant effect on the accuracy of trajectory recovery. Pairwise comparisons between the different algorithms were then performed using the Bonferroni adjustment to correct for the multiple comparisons and statistical significance was evaluated at the 5% level. These comparisons found the NMS approach to be significantly less accurate than the other algorithms tested, but found no statistically significant differences between the normalised average mTREs incurred by the other algorithms.

The lack of statistically significant results is not particularly surprising given that the Polaris tracker measurements were much noisier than those provided by the sub-millimetre positioning device in the experiments of Section 7.4.1. This is evident when comparing the oscillatory shape of the mTRE curves in Figure 7.9 to the much smoother shape of the mTRE curves obtained using the positioning device measurements as ground truth (e.g. Figure 7.4). The noise in the ground truth measurements may thus have obscured subtle accuracy differences between the algorithms. Much of the noise in the optical tracker measurements is due to quantisation effects caused by the finite size of the pixels of the tracker's stereo camera. Such quantisation effects manifest themselves in the form of oscillatory patterns in the measurements. These patterns are particularly evident in the displacement and yaw measurements for the pork tenderloin example of Figure 7.9(b). These effects sometimes make it ambiguous to decide visually which algorithm is most accurate, such as in the beef tenderloin case of Figure 7.9(d).

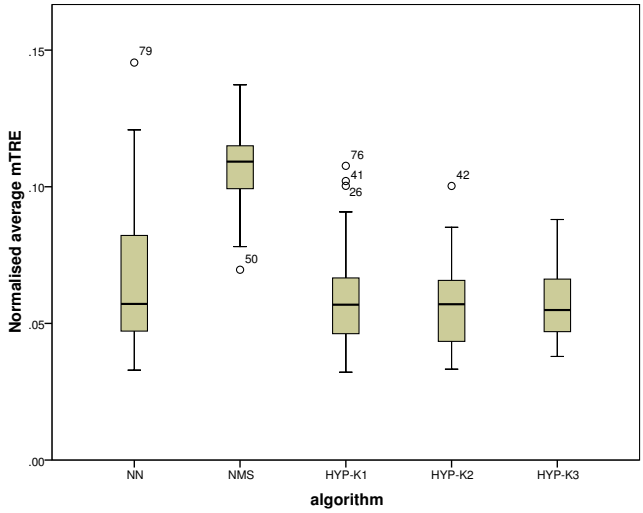


Figure 7.10 Boxplot of the normalised average mTRE between the optically tracked and estimated chaotic trajectories for imagery of animal tissue for the NN, NMS, HYP-K1, HYP-K2 and HYP-K3 approaches.

Noise aside, it seems that accuracy differences between the different algorithms may also have shrunk due to the non-monotonicity of the trajectories considered in the experiments. Indeed, whereas the NN algorithm tended to underestimate displacement in most of the monotonic trajectory experiments, its accuracy appears improved in the case of non-monotonic trajectories and more often approaches that of the iterative measurement selection and fusion approaches HYP-K1, HYP-K2 and HYP-K3. The non-monotonic cases where the NN algorithm still significantly underestimated displacements seem to be ones where the trajectory remained relatively smooth, such as in the beef tenderloin cases shown in Figures 7.9(c) and 7.9(d). The NN approach seems more accurate where the probe undergoes many direction reversals, such as in the two pork tenderloin cases shown in Figure 7.9(a) and 7.9(b), where the NN curves for mTRE and displacement are difficult to see because of their overlap with the other curves. It seems likely that the reason for this relative improvement in performance of the NN algorithm is that the direction reversals also reverse the direction of any biases intrinsic to the method, so that direction changes enable the cancellation of the drift that would otherwise occur for the NN algorithm. The reader may remember that a similar phenomenon was noted in Chapter 4, in results from synthetic ultrasound imagery. Curiously, there are also a few experimental cases where the

NN approach overestimates displacement, such as the beef tenderloin example shown in Figure 7.9(e). These may have resulted partially from the same phenomenon.

A final remark concerns the error variability across experimental trials. Though the error variability associated with the different algorithms was not the object of a formal statistical analysis in this study, further scrutiny of the boxplot of Figure 7.10 suggests that the variants of the proposed iterative measurement selection and fusion approach may exhibit less variability in their ability to accurately recover non-monotonic transducer trajectories than the NN approach. This difference may be explained by the fact that there is less variability in the bias associated with the measurements they use and by the fact that they exploit a broader set of distance estimates whose uncertainties are averaged out by probabilistic fusion.

7.5 Summary and discussion

This chapter proposed a new approach for image-based recovery of out-of-plane ultrasound transducer trajectories from scans of real tissue. The approach generalises the probabilistic data fusion method first introduced in Chapter 4 in two distinct ways: (1) it dynamically adapts to the local characteristics of the scanned medium by exploiting the learning-based correlation length estimation of Chapter 6 and (2) it is able to overcome some limitations of this imperfect, but useful adaptive speckle decorrelation model by incrementally building a reconstitution of the unknown trajectory through the iterative selection and probabilistic fusion of distance measurements which agree with each other. In particular, the approach effectively avoids using distance estimates from uninformative, but significant correlation measurements arising from the correlation plateaux caused by elevationally persistent structures in the scanned medium.

Experimental results showed that, given a well constructed initial estimate of the trajectory, the proposed approach is significantly more accurate than a base-line deterministic approach which uses only distance estimates between consecutive pairs of images and that measurement selection yields significant improvements over the blind use of the probabilistic measurement fusion approach without measurement selection. It was also shown that the approach can be used in the case of non-monotonic (chaotic) probe trajectories without any modification to the TSP-sort frame ordering algorithm of Chapter 3 (beyond scaling the nominal decorrelation curve to account for the local elevational correlation length).

Despite these encouraging results, difficulties remain. Indeed, while distance underestimates resulting from correlation plateaux were generally avoided using the proposed measurement selection approach, distance overestimates resulting from the inaccuracy of the adaptive speckle decorrelation model in the case of moderately low input signal levels were more difficult to control. Though the HYP-K2 and HYP-K3 versions of the proposed algorithm were able to overcome this problem to a certain extent by better controlling the set of measurements which enter the initial trajectory estimate, experimental results show that displacements are still overestimated in some cases. There are many potential solutions to this problem; this chapter has only considered one of them (i.e. using knowledge about the behaviour of the adaptive speckle decorrelation model in the design of cost functions for the initial trajectory estimate). Another solution would be to design better testing schedules. However, these solutions seem somewhat *ad hoc* and perhaps less efficient than solutions that directly address the modelling inaccuracies in the adaptive correlation length estimation scheme. Several ideas for improving the adaptive speckle decorrelation model were already given in the previous chapter.

Further improvement of the proposed probabilistic data fusion approach might be achieved by exploiting the variance information output by the adaptive correlation length estimator of the previous chapter. In the present implementation, this information was, for all intents and purposes, eliminated from the computations through the coarse approximation of (7.9). Though the maximum likelihood problem of (7.8) might not be solvable using the original expectation-maximisation algorithm of [46], it may be solvable through other forms of expectation-maximisation with more realistic mathematical requirements (e.g. [231]).

Another interesting avenue for future research would be to exploit the measurement selection approach presented in this chapter in the context of a hybrid freehand 3D ultrasound imaging system of the kind described by Lang et al. [124], where positional information would be available from an external, potentially highly imprecise, position tracking device. Such a setup would eliminate the need for establishing an initial trajectory estimate from the ultrasound data themselves as the position tracker measurements, along with an appropriate instrument noise model, could be used for this purpose. The initial trajectory estimate thus obtained would stand a much better chance of being unbiased, which ought to improve the achievable accuracy to a great extent.

The results obtained for the chaotic trajectory experiments also suggest another avenue

for research. These results suggest that the accuracy of the NN approach is improved in the presence of non-monotonic motion due to drift cancellation effects which occur with direction reversals. It seems likely that such drift cancellation effects would also occur for other trajectory estimation algorithms (e.g. HYP-K2 or HYP-K3). The development of optimal scanning protocols which exploit similar ideas might help reduce the drift associated with displacement overestimation in the case of media with low echogenicity properties or overcome other modelling inaccuracies.

Chapter 8

Conclusions

The motivating context for the research presented in this thesis was that of freehand 3D ultrasound image acquisition. In this context, image-based transducer tracking can provide an alternative or a complement to measurements obtained from an external position tracking device, thereby eliminating or loosening the constraints associated with their use. This thesis focused on the problem of estimating the out-of-plane trajectory of an ultrasound imaging transducer from the spatial decorrelation of the speckle patterns in the acquired imagery.

Correlation between the speckle patterns of ultrasound images acquired at neighbouring locations arises from the finite beam-width of the ultrasound transducer. Under so-called Rayleigh scattering conditions, the relationship between the absolute elevational displacement of the transducer and the correlation of images theoretically depends only on the characteristics of the imaging instrument. Through a calibration step, a transducer specific elevational speckle decorrelation model can thus be obtained from imagery of a phantom with Rayleigh scattering properties sampled at known elevational displacement intervals. This model can then be used to estimate the elevational transducer displacement between pairs of correlated ultrasound images of objects with similar scattering properties with respect to ultrasound.

In practice, out-of-plane transducer tracking based on speckle decorrelation is much more difficult. One reason is that for any single correlation measurement between a pair of ultrasound image patches, there is an ambiguity as to the direction of the transducer displacement that caused it. When the trajectory of the probe is allowed to be non-monotonic

or lead to intersections between image frames, the trajectory cannot be recovered without somehow resolving this directional ambiguity. Another difficulty is that for a fixed elevational separation, there is significant variability in the sample correlation coefficients that might be measured between finitely sized image windows, even under Rayleigh scattering conditions. These issues become even more problematic when Rayleigh scattering conditions do not hold, which is most of the time when it comes to imagery of real tissue. In imagery of real tissue, the rate at which speckle patterns decorrelate over distance is no longer determined solely by the characteristics of the transducer and the (unknown) properties of the medium must be taken into consideration as well. This thesis examined these difficulties from a unique probabilistic perspective and proposed novel solutions which aimed at minimising the impact of measurement noise and maximising the use of the information present in the data. The resulting technical contributions are reviewed in Section 8.1 below. Section 8.2 discusses the limitations of this work and suggests avenues for future research.

8.1 Summary of research contributions

Chapter 3 presented a framework for recovering the structure of an ultrasound scan using elevational correlation measurements between multiple pairs of corresponding image patches. A novel combinatorial optimisation approach based on the solution to a travelling salesman problem was proposed for ordering corresponding image patches along the elevational direction. This approach enforces consistency in the displacement estimates obtained for every triplet of corresponding correlated image patches in the scan. Simulation experiments showed that the combinatorial approach exhibits considerable robustness to noise. Another important constraint is that of image planarity, which implies that the set of elevational displacements estimated at different locations of two image frames must define a rigid transformation. Combining the results of combinatorial frame ordering with a robust least-median-of-squares rigid transformation fitting approach allows for the implicit discovery of non-monotonic motion and frame intersections. These methods were used successfully in several of the experiments presented in subsequent chapters of this thesis.

Chapter 4 investigated the accuracy of different strategies for the recovery of out-of-plane transducer trajectory under the assumption of Rayleigh scattering. Of particular interest to this study were the variability of the correlation measurements between image

patches which arises from finite sample sizes as well as its consequences on the accuracy and precision of associated displacement estimates. A new probabilistic transducer specific speckle decorrelation model was proposed in order to reflect the variability of sample correlation coefficients measured between finite ultrasound image patches. Like its deterministic counterpart, the probabilistic speckle decorrelation model is obtained through a calibration scan of a reference phantom object with Rayleigh scattering properties. The model takes the form of a probability density function over sample correlation coefficients whose parameters vary with elevational displacement and are estimated through a maximum entropy approach. One advantage of this new model is that it allows quantification of the uncertainty attached to the displacement estimate associated with a correlation measurement. It was shown that this, in turn, allows for the probabilistic fusion of multiple redundant correlation estimates within a maximum likelihood estimation framework. This new approach was compared to two distinct deterministic approaches to trajectory estimation in synthetic and real ultrasound imagery. The experiments on synthetic imagery showed that under ideal conditions, the probabilistic strategy is significantly more accurate than the deterministic ones. These experiments also examined the accuracy of the different measurement fusion strategies as a function of transducer speed, both at calibration and test time. It was found that the inter-frame spacing of the calibration scan had an effect on all algorithms tested and that the proposed approach has a preference for widely spaced (and therefore longer and more diverse) calibration scans. It was also found that while the speed of the transducer motion had a statistically significant effect on the accuracy of both deterministic approaches tested, it did not have a significant impact on the performance of the proposed probabilistic technique, implying that its behaviour is more stable across the range of possible trajectories with varying speed. The experiments on real phantom imagery also showed that even tightly controlled experimental conditions are not ideal in this context, giving rise to unmodelled local variations in speckle correlation length for different locations in the phantom object. As a result, the proposed strategy was able to perform well but did not consistently outperform the best of the two deterministic alternatives.

Key to the practical relevance of speckle decorrelation approaches to out-of-plane transducer motion tracking is the necessity of generalising their application to the case of arbitrary media (e.g. biological tissues) where Rayleigh scattering conditions are not met. Most related work in the literature tackles this issue by filtering the imagery to detect regions of fully developed speckle where the nominal speckle decorrelation model applies; however,

this approach is limited because Rayleigh scattering conditions do not occur frequently in real tissue. Non-Rayleigh scattering conditions arise in a medium as the concentration of ultrasound scatterers in the medium decreases or when these scatterers exhibit non-random spatial organisation. As a first step towards understanding the consequences of these changes in the hypothesised structure of the medium on elevational speckle decorrelation, Chapter 5 introduced a new approach for simulating a broad range of spatial point scatterer distributions. The new method is used for the purpose of creating synthetic ultrasound imagery and studying its properties in relation to known micro-structure parameters. The approach is based on an existing 1D generalised Poisson point process model [123, 236, 43] whose two parameters control the density and spatial organisation (ranging from clustered to regular) of the ultrasound scatterers. The novelty of the proposed approach lies in the generalisation of this 1D model to the multiple dimensions required for the simulation of realistic ultrasound images. This generalisation is achieved by mapping the output of the 1D point process to a Hilbert space-filling curve, thereby preserving its original characteristics in terms of spatial organisation as well as the intuitive parameterisation of the original 1D model. Synthetic ultrasound imagery of virtual phantoms with varying micro-structure was generated using this approach and its first and second order statistics were found to be reliably predictable from the model parameters and in agreement with predictions from ultrasound physics. This makes the proposed scatterer distribution model a useful tool for the validation of ultrasound image processing algorithms at large as well as a means of generating data from which to empirically study phenomena related to ultrasound image formation.

Chapter 6 showed how to leverage the power of the aforementioned simulation method in the context of establishing a new speckle decorrelation model which dynamically adapts to the local characteristics of the imaged medium. The key to this dynamic adaptation is the fact that changes in medium micro-structure affect in-plane first and second order image statistics as well as out-of-plane speckle decorrelation, so that changes in elevational correlation length can be predicted from local image statistics. An important contribution of this thesis is the demonstration that the relationship between the characteristics of the transducer, in-plane and out-of-plane image statistics can be learned by studying a pool of varied synthetic ultrasound imagery with known elevational correlation length. A sparse Gaussian process regressor was established to represent this relationship and was shown to generalise very well to imagery of varied media acquired from virtual and real

transducers with different technical specifications. Displacement estimation experiments carried out on imagery of different types of animal tissue showed that the learning-based adaptive speckle decorrelation model was generally quite accurate, outperforming a speckle detection approach [167] and giving results of accuracy comparable to the state-of-the-art method of Gee et al. [72]. The proposed learning-based approach offers important advantages over the state-of-the-art: it is data driven rather than heuristic and provides a measure of uncertainty along with its output. There are nonetheless circumstances where the proposed approach fails to correctly predict the correlation of images as a function of distance. One such circumstance arises from physical structures whose presence in the medium persists over a long distance in the elevational direction (such as a highly echogenic layer of fat), leading to imagery which remains significantly correlated over long distances, but not in a way that is predictive of transducer motion. Another kind of model failure occurs when the statistics of imagery under consideration fall in an area of the input space that is not well modelled by the training data extracted from the synthetic imagery. This can happen in the presence of image acquisition noise or other artefacts which were left unmodelled in the simulation process.

In Chapter 7, the contributions made in all previous chapters were combined into a single experimental framework for the estimation of out-of-plane ultrasound probe trajectories in imagery of real tissue. The probabilistic speckle decorrelation model of Chapter 4 was generalised to embed the elevational correlation length information provided by the adaptive learning-based method of Chapter 6. In order to circumvent the inevitable inaccuracies intrinsic to the new adaptive speckle decorrelation model, a new iterative algorithm was introduced which robustly recovers increasingly precise trajectory estimates by gradually incorporating the set of available correlation measurements to an initial, highly uncertain trajectory estimate. The chosen approach is to iteratively alternate measurement selection steps with probabilistic measurement fusion steps. During the measurement selection step, candidate distance estimates obtained from the adaptive speckle decorrelation model are evaluated for consistency with the current trajectory estimate using a statistical hypothesis test. Measurements which fail the consistency test are deemed to be biased and are discarded; those which pass the test are integrated to the model of the trajectory. The complete framework, including the frame ordering and robust rigid transformation fitting approaches of Chapter 3, was evaluated in monotonic and non-monotonic trajectory recovery experiments with imagery of animal tissue. While the results varied in terms of

quality, they showed very good promise for the new iterative measurement selection and fusion approach. Indeed, it was found that given a suitable initial trajectory estimate, this method was, overall, significantly more accurate than the original probabilistic measurement fusion approach (which does not perform measurement selection) as well as a base-line deterministic approach.

8.2 Current limitations and future work

Throughout this thesis, an effort has been made to point out the limitations of the proposed technical contributions as well as their strengths. Many possible improvements to the proposed methods have already been discussed as a result and not all of them need repeating in detail here. Among these suggestions, two stand out as most productive. The first is the idea of improving the learning-based adaptive speckle decorrelation approach of Chapter 6 by introducing realistic imaging artefacts into the simulation process used to create the training set. Important such artefacts which were not modelled in this thesis include signal attenuation, electronic noise, spatial and grey level quantisation as well as frame-grabber interlacing artefacts. Another factor which could be interesting to model is anisotropic tissue micro-structure. This could be done by generalising the multi-dimensional scatterer distribution model of Chapter 5 which mostly generates isotropic scatterer configurations because of the Hilbert space-filling curve it uses to achieve the mapping of scatterer distributions from 1D to 3D. Different space-filling curves could be used to generate media with different levels of anisotropy. As a result of modelling these new effects, it is expected that the sparse Gaussian process regressor which currently models the relationship between in-plane statistical image features and elevational correlation length would change substantially and yield improved distance estimation accuracy in situations where these imaging artefacts make significant contributions to image content. This, in turn, would probably improve the robustness of the iterative measurement selection and fusion scheme proposed in Chapter 7.

The other suggestion that was previously made has to do with improving the overall trajectory recovery framework of Chapter 3. As was demonstrated in the experiments presented in Chapters 4 and 7, this framework suffers from a few limitations which were unforeseen at the time of its development. One is that in the presence of modelling inaccuracies, which appear to be almost inevitable in the case of real tissue imagery, elevational

displacement drift error which is localised to a few image patches in the pre-defined grid can cause important rotational errors in the overall trajectory estimate because the image planarity constraint is enforced directly on the complete set of elevational mini-frame trajectories associated with each patch. As already suggested at the end of Chapter 7, trajectories might be recovered more robustly if the image planarity constraint were more tightly coupled to the result of probabilistic measurement fusion and used to incrementally correct the elevational trajectories associated with each mini-frame.

A more fundamental limitation of the framework described in Chapter 3 is its implicit assumption that the trajectory is mostly translational. Indeed, the combinatorial mini-frame ordering and probabilistic mini-frame trajectory estimation approaches are only optimal if the mini-frames associated with one image patch location are aligned along the elevational direction. In situations where they are not, such as in the experiments of Chapter 4 on synthetic rotational image sequences, the burden of correcting the errors resulting from the violation of the assumption falls entirely on the robust rigid transformation fitting approach which enforces the image planarity constraint. While reasonably good results were achieved with synthetic data using this approach in the presence of relatively large rotations, it is probably worthwhile to investigate measurement fusion approaches which do not rely on this implicit assumption. One possibility would be to enforce the image planarity constraint after mini-frame ordering, but before probabilistic data fusion, thereby estimating rigid transformations between every pair of frames. Using a probabilistic methodology similar to that presented in this thesis, one could obtain a covariance matrix describing the uncertainty in the rigid transformation parameters relating pairs of entire frames rather than the single distance values between pairs of mini-frames considered in this thesis. The maximum likelihood data fusion approach described in this thesis and derived from Lu and Milios's work [139] could readily be generalised to such rigid transformation data, as has recently been done in the field of mobile robotics [4].

By far, the most significant limitation of the research presented in this thesis is the very narrow range of trajectory types it focuses on. From the outset, it was assumed that probe motion only occurred out-of-plane, or that any in-plane motion was already accounted for by some external image registration process. While preliminary work was carried out which touched the possibility of out-of-plane transducer rotations (the robust rigid transformation fitting approach of Chapter 3 was designed for this purpose and tested in synthetic ultrasound imagery in Chapter 4), the emphasis throughout most of this thesis has been

on purely translational trajectories. This facilitated the development and evaluation of the medium independent approaches of Chapters 6 and 7, which represent a major contribution of this work.

Generalisation of the methods presented in this thesis to arbitrary trajectories is thus an extremely important direction for future work and would certainly raise the level of clinical applicability of the research presented here. A number of problems must be solved in order to achieve this. First, the recovery of in-plane transducer motion must be made part of the overall trajectory recovery framework. There exist a multitude of methods for in-plane image registration [140, 89], including some ultrasound specific speckle motion tracking techniques [215, 24, 240, 73, 93, 242]. These techniques are all based on the following general idea: the in-plane motion which occurred between the acquisition of one image and that of another is recovered by finding the motion parameters (two translations and one rotation for rigid motion) which maximise a similarity measure (such as normalised correlation) between the first image and the transformed version of the second image. Once this optimal in-plane transformation is determined, the value of the normalised correlation coefficient at the optimum can, in theory, be used to estimate out-of-plane probe translation [93].

One difficulty with this approach pertains to the use of frame-grabbed, scan-converted B-mode ultrasound data (which were the only data available in the context of the present thesis) as opposed to the (envelope detected) high resolution RF signals preferred by several research groups in this line of work [94, 124]. During scan-conversion, image content is interpolated at display locations not covered by the original RF vectors. This interpolation process, combined with the interpolation processes involved in many conventional image registration algorithms, is likely to undermine the accuracy of both in-plane and out-of-plane motion estimates [93]. Using B-mode data also precludes the direct use of theoretical knowledge regarding the properties of envelope detected RF signals to enhance the registration process [93]. A partial solution to this difficulty would be to simply use RF data (or envelope detected RF data), but these are not readily available from most clinical ultrasound scanners. Although this situation is beginning to change within research environments with the recent advent of open ultrasound scanner architectures [69], real-time RF data acquisition is still not entirely trivial. B-mode remains the chief ultrasound imaging modality in use within the clinical setting and it is not yet realistic to pretend otherwise. Furthermore, while the use of RF data eliminates the interpolation artefacts introduced by the scan conversion process, it does not eliminate the interpolation artefacts

caused by the image registration process itself [163, 188]. Therefore, there is still much room for algorithmic improvements in resolving the issue of interpolation artefacts for the in-plane registration of ultrasound imagery; one possible approach is described in [188].

Another major difficulty with image registration in the context of freehand 3D ultrasound imaging is the speckle decorrelation which is expected to arise from out-of-plane motion (and used to estimate it). In fact, elevational speckle decorrelation reduces the accuracy and reliability of in-plane motion tracking and as such, it is considered to be a hindrance in the speckle tracking literature [90]. Conversely, in-plane image motion which has an axial component causes speckle decorrelation which is not accounted for by elevational speckle decorrelation models, leading to out-of-plane motion estimation errors. Similar issues arise when out-of-plane rotations are taken into consideration as well [133, 92]. A rotational speckle decorrelation model has already been proposed by Housden et al. [92]. An interesting challenge would be to statistically model or characterise the interaction between in-plane motion, out-of-plane rotations and out-of-plane translation in terms of the observed speckle decorrelation and develop measurement fusion techniques which minimise the impact of the resulting ambiguities. The work presented in this thesis on the topic of tissue invariant speckle decorrelation models has shown that synthetic ultrasound imagery is a powerful tool for empirically modelling textural phenomena arising from the image formation process. It is believed that this strategy may prove advantageous in studying the effects of complex motions on ultrasound image correlation and motion estimation algorithms that use it.

One way to resolve ambiguities in motion type would be to introduce some form of prior knowledge about the image acquisition process in the trajectory recovery algorithm. This prior knowledge would also help in improving the behaviour of the measurement fusion algorithm proposed in this thesis in the presence of speckle decorrelation model failure. This prior knowledge might be obtained from a position tracking device, leading to hybrid freehand 3D ultrasound acquisition systems like those developed by Housden et al. [95, 92] and Lang et al. [124]. The iterative measurement selection and fusion approach presented in Chapter 7 would lend itself very well to this type of multi-sensor fusion. In this context, a 6 degree-of-freedom position tracker would provide the trajectory estimate necessary to initialise the proposed iterative algorithm, eliminating the need to create such an initial estimate from potentially biased measurements. The tracking device in such a hybrid system need not even provide 6 degree-of-freedom motion tracking; it has been demonstrated that

many ambiguities can be resolved using 3 degree-of-freedom orientation sensors, which offer many practical advantage over optical trackers [133, 95, 92]. Combination of image-based and tracker-based information also opens the door to the proper treatment of cases where the tissue under observation undergoes deformations as the transducer moves. Discrepancies between the motions reported by the two sources of data would provide valuable information about physiological motion. This idea has been investigated in the context of in-plane image deformations [118, 218] but out-of-plane deformations have not yet been considered. A statistical hypothesis testing approach similar to the one presented in Chapter 7 would provide a good starting point for automatically detecting the discrepancies between tracker-based and image-based motion estimates.

Tracking devices are not the only possible source of useful prior knowledge. Indeed, prior knowledge may also arise from constraints which are intrinsic to the clinical context in which freehand 3D ultrasound imagery is acquired. For instance, in some examination protocols, it may be entirely reasonable to assume that the motion of the probe is mostly translational, and a statistical model could be developed which describes the likely magnitude of deviations from this assumption. This could be achieved by systematically studying sets of position tracker measurements associated with multiple scans acquired by single clinicians or groups of clinicians who carry out similar procedures. Prior knowledge may also arise out of anatomical necessity. One remarkable example of this is the case of interventional procedures which use transrectal ultrasound probes, such as prostate biopsy, where 3D ultrasound can be used to establish a reference model of the prostate to guide needle insertion [40]. In this context, the space of possible trajectories taken by the ultrasound probe is heavily limited by anatomical constraints. These anatomical constraints have been used successfully to improve trajectory estimates derived from position tracker measurements [41] and constrain image registration [17]. Similarly, in combination with a good probabilistic model describing the interactions between the different motion parameters, they might prove very useful in reducing or eliminating any ambiguities in the source of the observed speckle decorrelation.

A final note on these suggestions for future research directions concerns the important issue of experimental validation. In the course of the work leading to this thesis, this issue has been the focus of increasing attention and effort, leading to more solid primary research findings as well as some secondary, but interesting research findings which might otherwise have been overlooked. In this context, quantitative experimental validation entails finding

ways of obtaining reliable and accurate ground truth trajectory data and measures for comparing the trajectory recovered algorithmically to the ground truth in a way that allows meaningful statistical comparisons with alternative approaches. Improvements in these areas will be needed should this research be extended in any of the ways described in this chapter.

While ground truth measurements can be obtained from a position tracker, these measurements are noisy and this, as shown by the statistical analysis of the non-monotonic trajectory recovery results of Chapter 7, can obscure some subtle differences between alternative approaches. The need for other sources of ground truth data is even more evident when the evaluation of hybrid image acquisition systems is considered, as the position tracking device is, in this case, part of the framework being tested. Optical trackers have been used for ground truth purposes in the validation of a hybrid system based on less precise electromagnetic tracking [124]. Another way of facilitating validation of such a system would be to design phantoms containing fiducial markers or other artificial structures (along with animal tissue, for instance) whose relative locations and geometry are known with high accuracy and which can be localised reliably and precisely in ultrasound imagery. The accuracy of the recovered trajectory could then be evaluated by comparing the true geometry of the set of markers to that implied by the estimated trajectory.

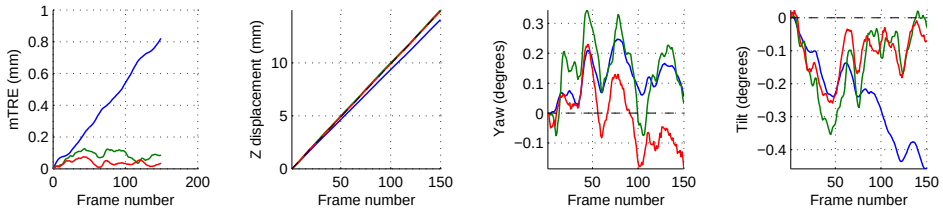
Research in image-based transducer tracking would also benefit from the development of new measures to summarise trajectory accuracy for the purpose of statistical analysis. This thesis used the mean target registration error, averaged over the entire image sequence, as a measure of accuracy for this purpose. While this measure does, in some respect, quantify accuracy, it does so only in a very narrowly-defined way. It quantifies the average deviation between the estimated location of any point of interest in the set of images and its true, absolute position. As a result, a large average error may be reported when significant errors occur in the first few frames of the scan, even if the rest of the trajectory follows the shape of the true one perfectly. An error measure which accounts for relative displacement estimation errors over different scales might provide more refined and highly valuable insights into the strengths and weaknesses of different trajectory recovery strategies and lead to more meaningful comparisons. In clinical applications, relevant error measures could be devised based on the task at hand. Indirect measures, such as the accuracy of needle insertions or other instrument manipulations with respect to specific targets in phantom trials, may in fact provide the most meaningful assessments of the framework

proposed in this thesis, with regards to its usefulness.

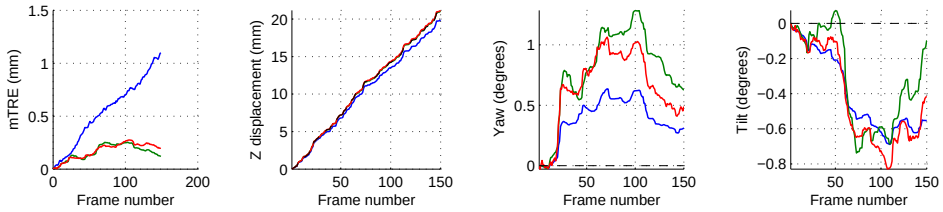
Appendix A

Full set of results for trajectory estimation in synthetic imagery of speckle phantoms

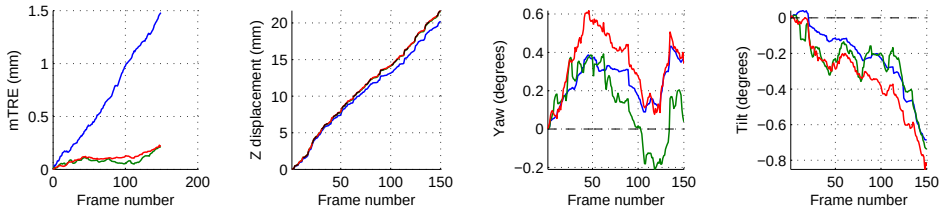
This appendix contains the full set of results obtained for the experiments relating to the recovery of transducer trajectories in synthetic imagery of fully developed speckle described in Chapter 4. Every row of each figure represents the results obtained for a single scan. From left to right, the plots in a given row show (1) the progression of the mTRE over frame acquisition steps, (2) the estimated displacement, (3) the estimated yaw (should be zero) and (4) the estimated tilt (should be zero) evaluated using the centroid of the image patches as the centre of rotation. The dashed line corresponds to the ground truth measurements. The blue line corresponds to the NN method. The green line corresponds the SR method. The red line corresponds to the ML measurement fusion approach of Chapter 4.



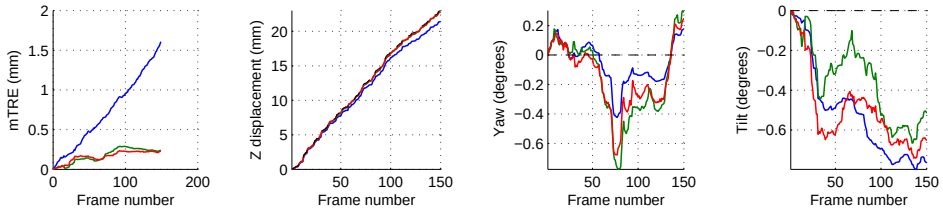
(a) Synthetic 1: Calibration scan



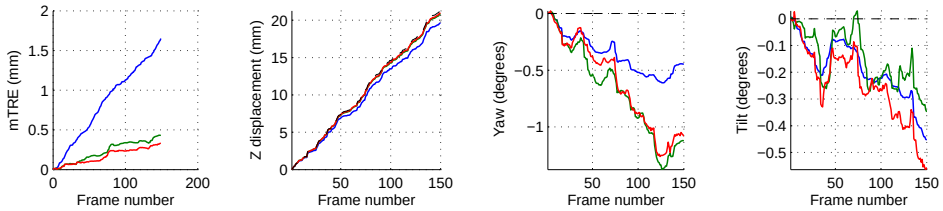
(b) Synthetic 2: Pure monotonic translation



(c) Synthetic 3: Pure monotonic translation

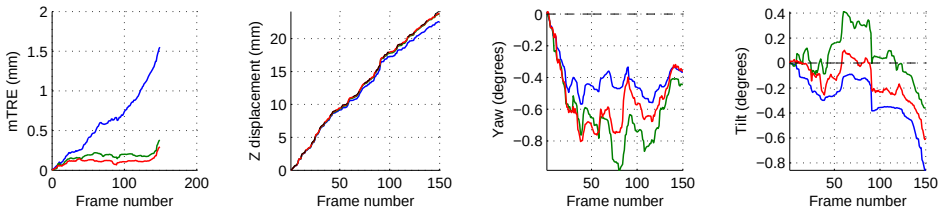


(d) Synthetic 4: Pure monotonic translation

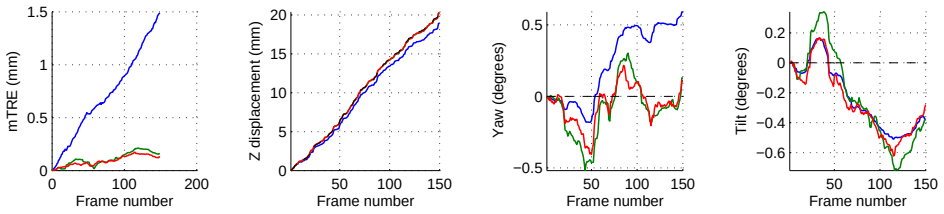


(e) Synthetic 5: Pure monotonic translation

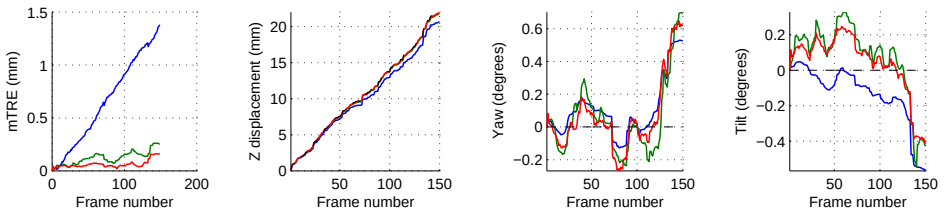
Figure A.1 Results for pure monotonic translational trajectories in synthetic speckle phantom data, part 1



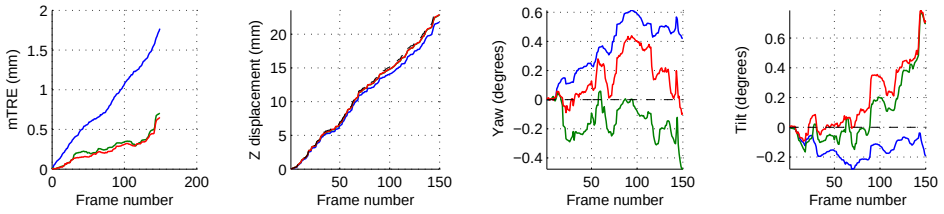
(a) Synthetic 6: Pure monotonic translation



(b) Synthetic 7: Pure monotonic translation

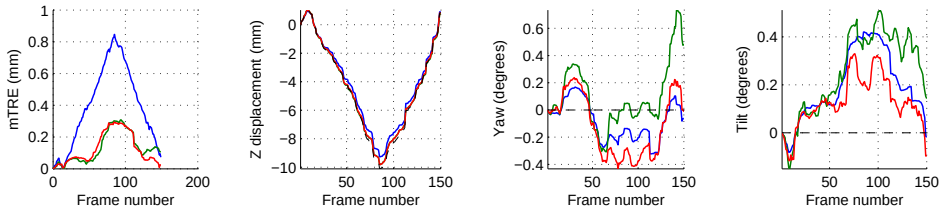


(c) Synthetic 8: Pure monotonic translation

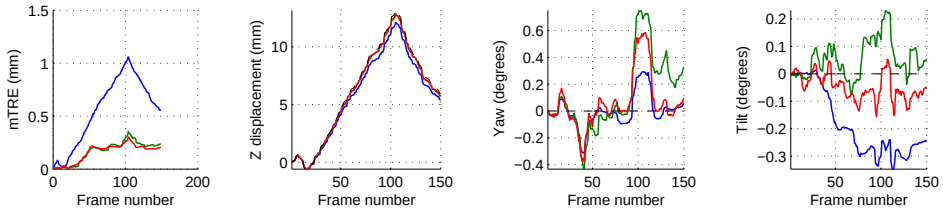


(d) Synthetic 9: Pure monotonic translation

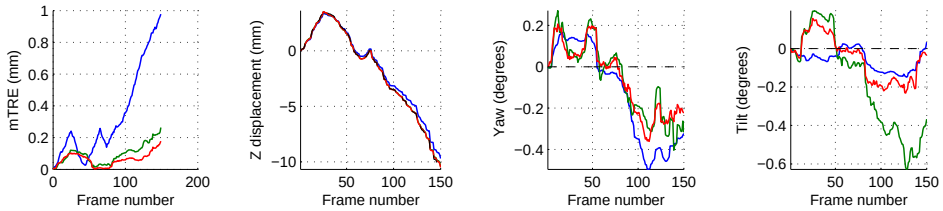
Figure A.2 Results for pure monotonic translational trajectories in synthetic speckle phantom data, part 2



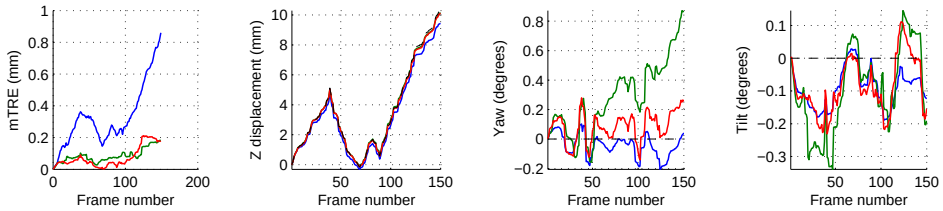
(a) Synthetic 10: Pure non-monotonic translation



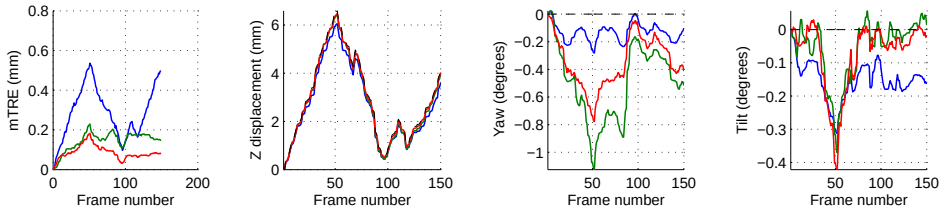
(b) Synthetic 11: Pure non-monotonic translation



(c) Synthetic 12: Pure non-monotonic translation

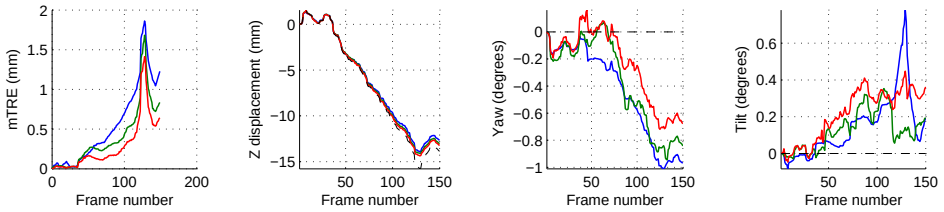


(d) Synthetic 13: Pure non-monotonic translation

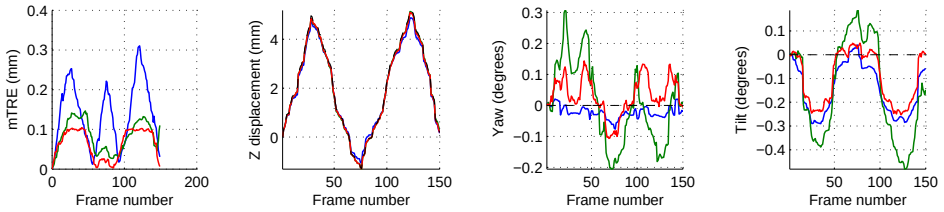


(e) Synthetic 14: Pure non-monotonic translation

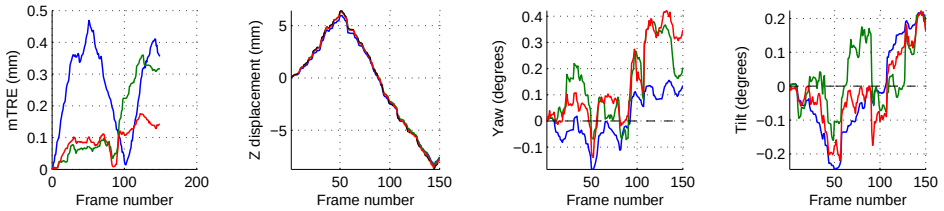
Figure A.3 Results for pure non-monotonic translational trajectories in synthetic speckle phantom data, part 1



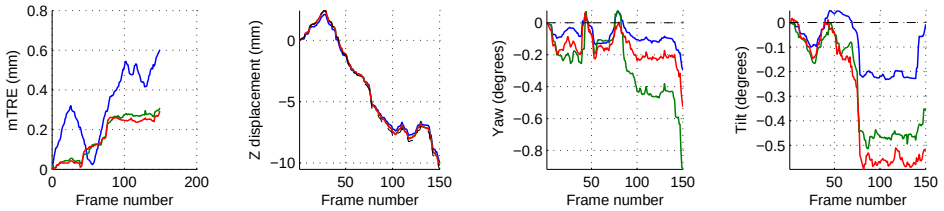
(a) Synthetic 15: Pure non-monotonic translation



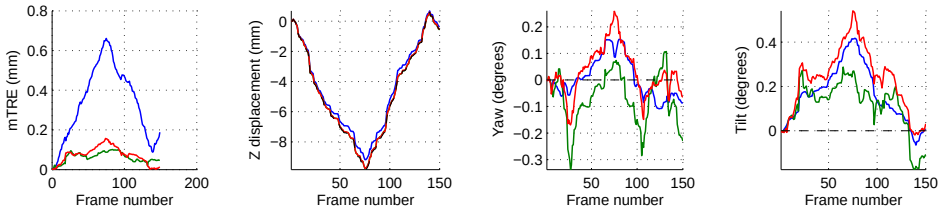
(b) Synthetic 16: Pure non-monotonic translation



(c) Synthetic 17: Pure non-monotonic translation

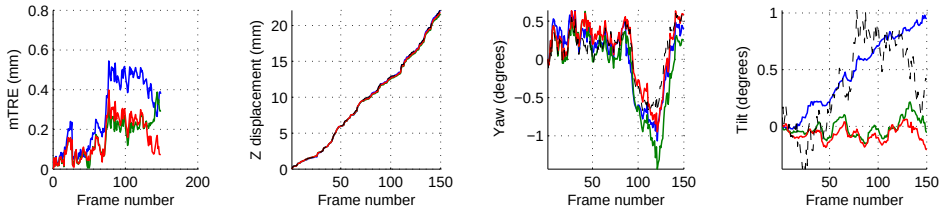


(d) Synthetic 18: Pure non-monotonic translation

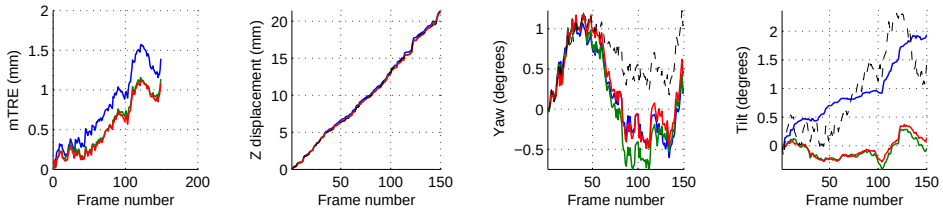


(e) Synthetic 19: Pure non-monotonic translation

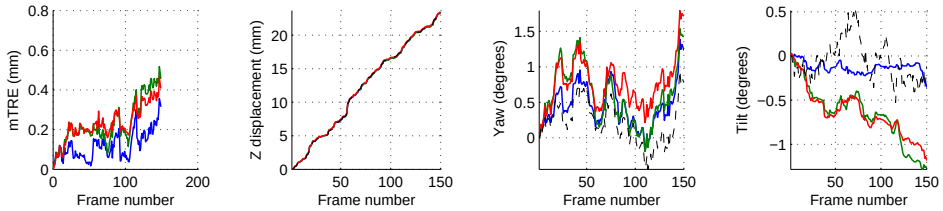
Figure A.4 Results for pure non-monotonic translational trajectories in synthetic speckle phantom data, part 2



(a) Synthetic 20: Monotonic translation with noisy yaw and tilt

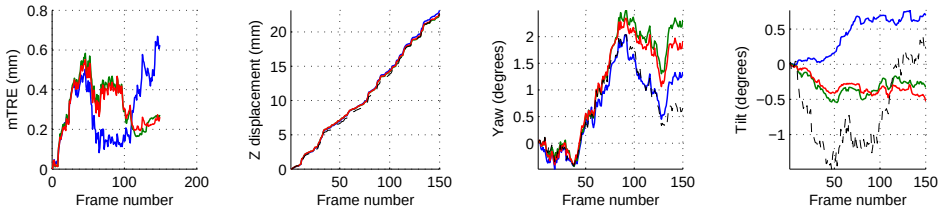


(b) Synthetic 21: Monotonic translation with noisy yaw and tilt

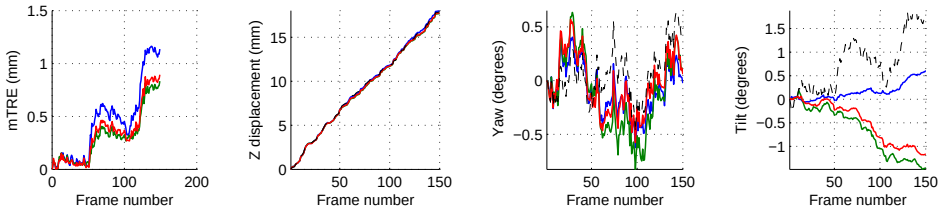


(c) Synthetic 22: Monotonic translation with noisy yaw and tilt

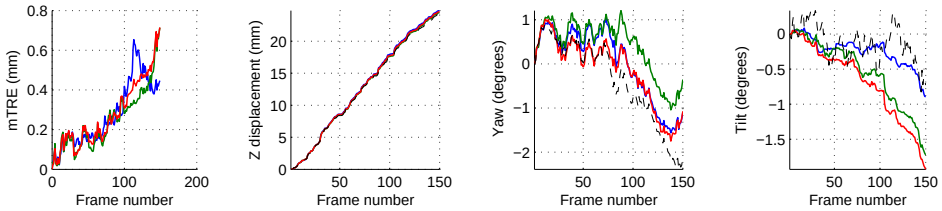
Figure A.5 Results for noisy monotonic translational trajectories in synthetic speckle phantom data, part 1



(a) Synthetic 23: Monotonic translation with noisy yaw and tilt

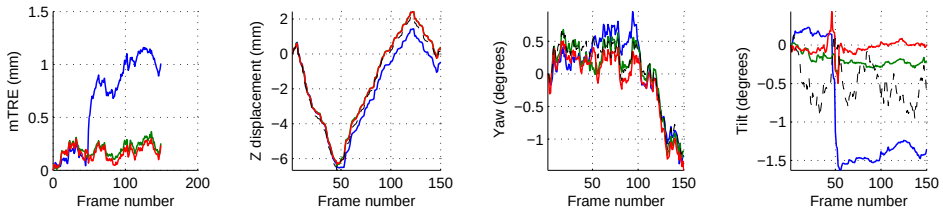


(b) Synthetic 24: Monotonic translation with noisy yaw and tilt

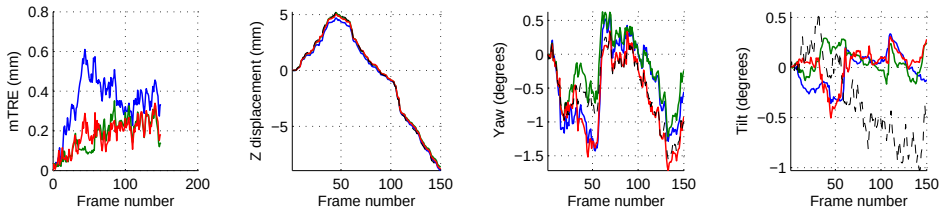


(c) Synthetic 25: Monotonic translation with noisy yaw and tilt

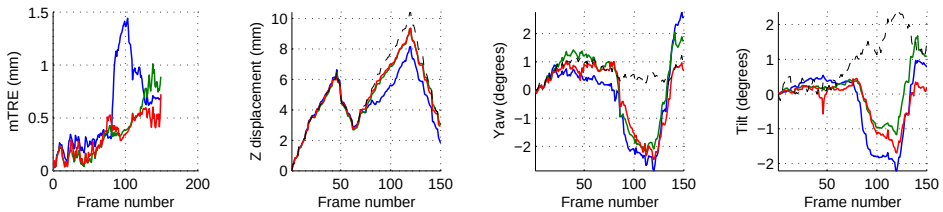
Figure A.6 Results for noisy monotonic translational trajectories in synthetic speckle phantom data, part 2



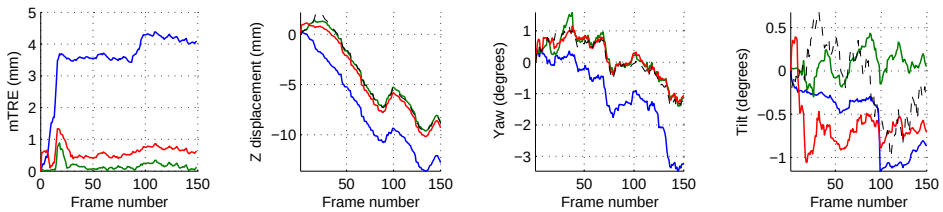
(a) Synthetic 26: Non-monotonic translation with noisy yaw and tilt



(b) Synthetic 27: Non-monotonic translation with noisy yaw and tilt

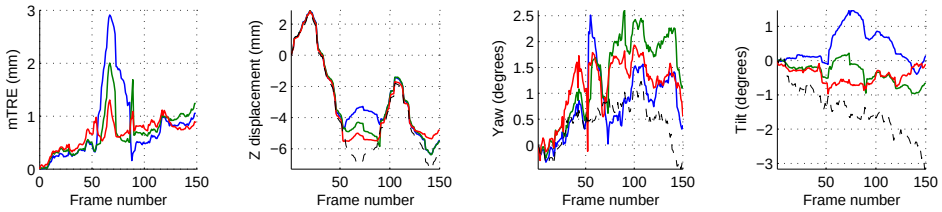


(c) Synthetic 28: Non-monotonic translation with noisy yaw and tilt

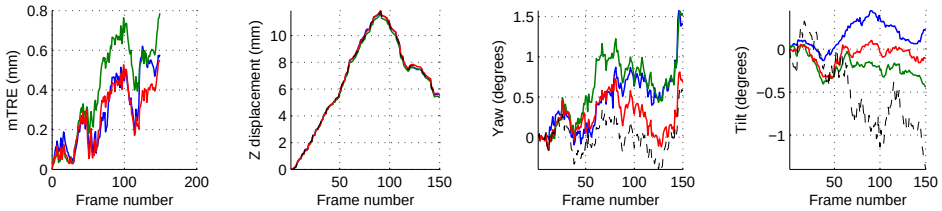


(d) Synthetic 29: Non-monotonic translation with noisy yaw and tilt

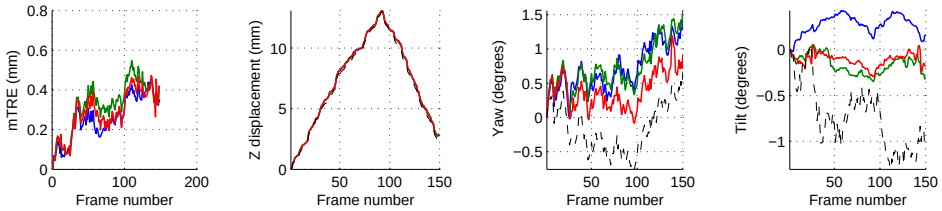
Figure A.7 Results for noisy non-monotonic translational trajectories in synthetic speckle phantom data, part 1



(a) Synthetic 30: Non-monotonic translation with noisy yaw and tilt

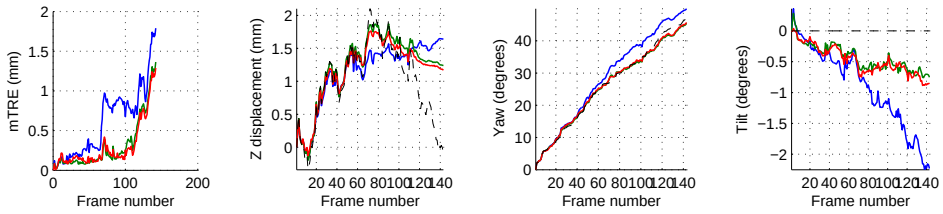


(b) Synthetic 31: Non-monotonic translation with noisy yaw and tilt

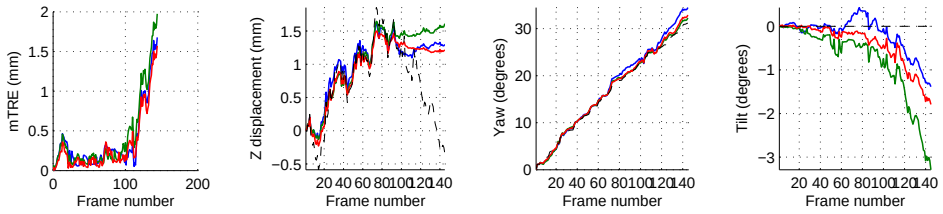


(c) Synthetic 32: Non-monotonic translation with noisy yaw and tilt

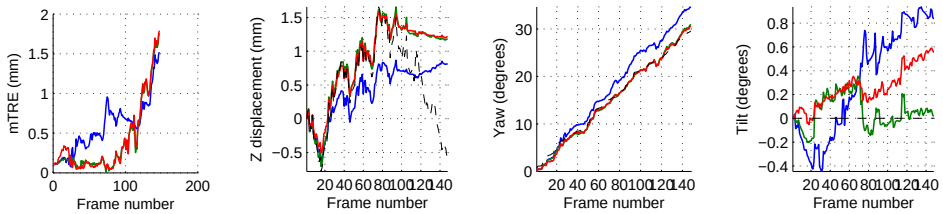
Figure A.8 Results for noisy non-monotonic translational trajectories in synthetic speckle phantom data, part 2



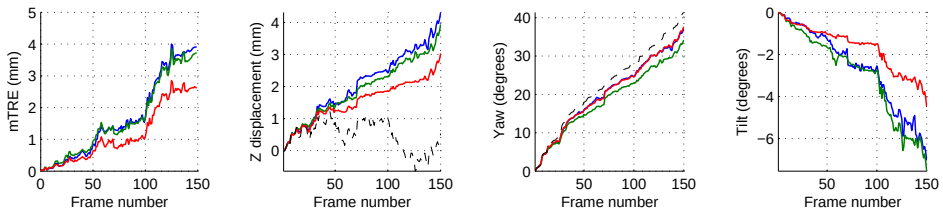
(a) Synthetic 33: Fast yaw motion with noisy translation



(b) Synthetic 34: Slow yaw motion with noisy translation

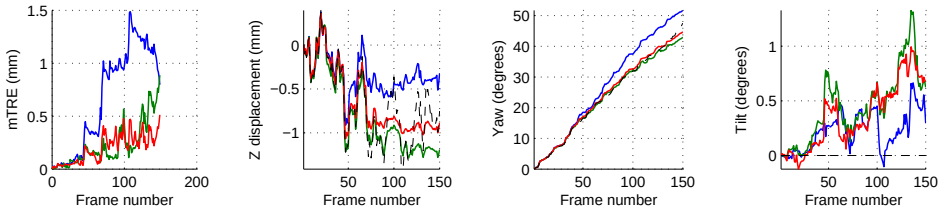


(c) Synthetic 35: Slow yaw motion with noisy translation

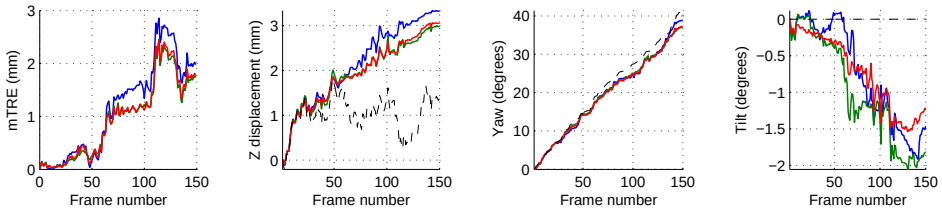


(d) Synthetic 36: Fast yaw motion with noisy translation

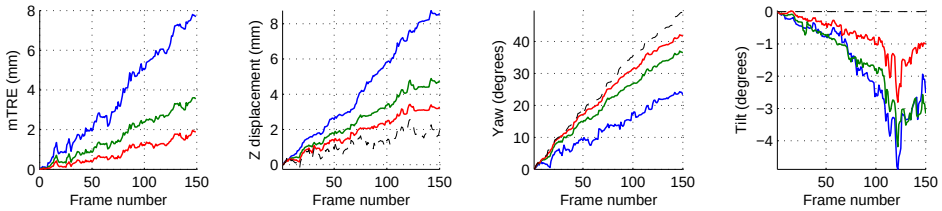
Figure A.9 Results for rotational trajectories in synthetic speckle phantom data, part 1



(a) Synthetic 37: Fast yaw motion with noisy translation



(b) Synthetic 38: Fast yaw motion with noisy translation



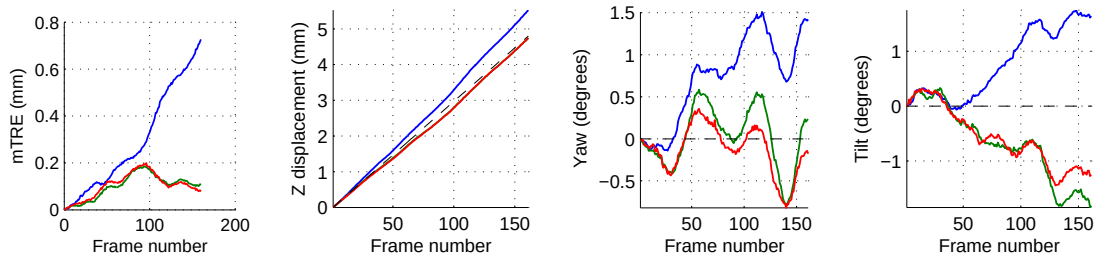
(c) Synthetic 39: Fast yaw motion with noisy translation

Figure A.10 Results for rotational trajectories in synthetic speckle phantom data, part 2

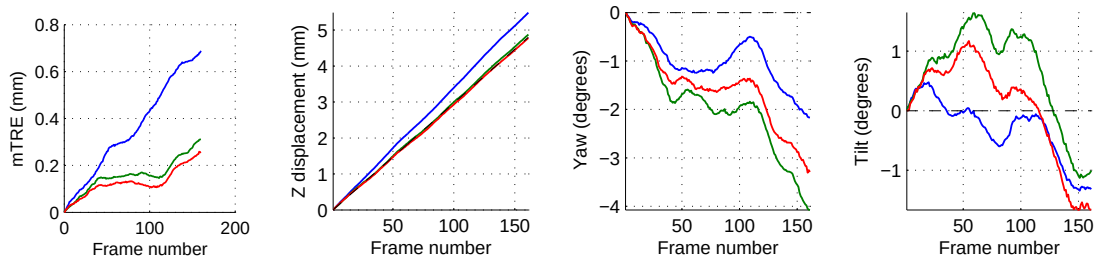
Appendix B

Selected results for monotonic trajectory estimation in real imagery of speckle phantoms

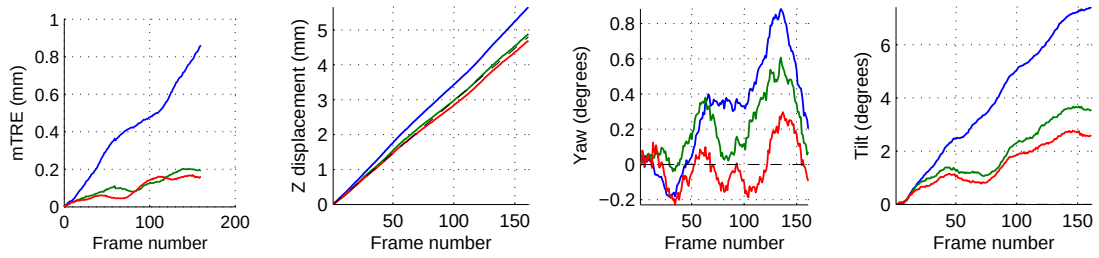
This appendix contains a subset of results obtained for the experiments relating to the recovery of monotonic transducer trajectories in real speckle phantom imagery described in Chapter 4. Since results were obtained for all possible pairings of calibration and test scans, yielding thousands of results, it was impractical to include them all. This appendix shows the results obtained for a single example for each combination of calibration sequence inter-frame spacing test sequence inter-frame spacing. The example was selected as the one which gave the median average mTRE (using the ML measurement fusion approach of Chapter 4. Every row of each figure represents the results obtained for a single scan. From left to right, the plots in a given row show (1) the progression of the mTRE over frame acquisition steps, (2) the estimated displacement, (3) the estimated yaw (should be zero) and (4) the estimated tilt (should be zero) evaluated using the centroid of the image patches as the centre of rotation. The dashed line corresponds to the ground truth measurements. The blue line corresponds to the NN method. The green line corresponds the SR method. The red line corresponds to the ML measurement fusion approach of Chapter 4.



(a) Calibration spacing = 0.03 mm, Subject spacing = 0.03 mm

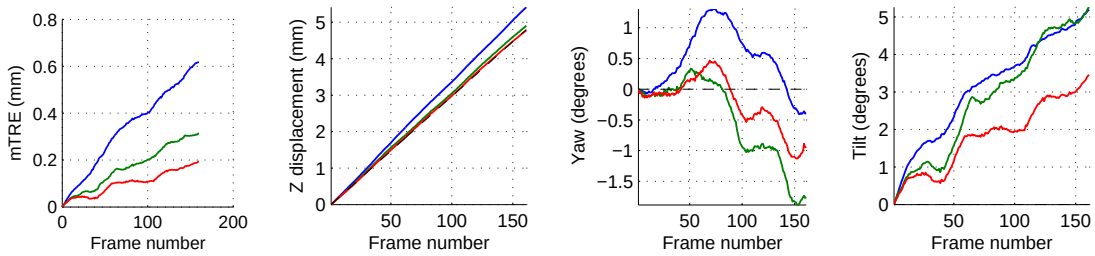


(b) Calibration spacing = 0.04 mm, Subject spacing = 0.03 mm

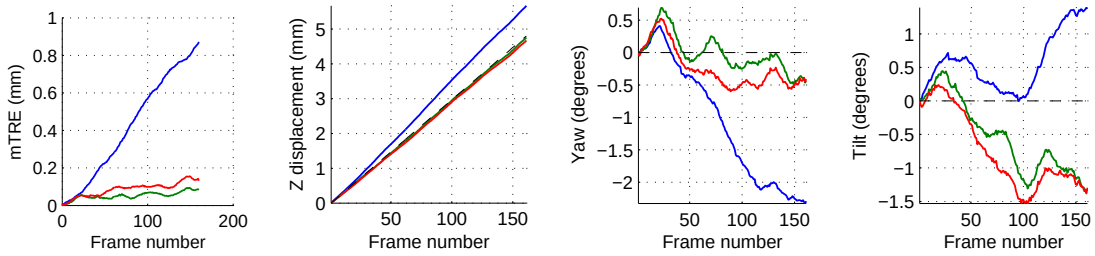


(c) Calibration spacing = 0.05 mm, Subject spacing = 0.03 mm

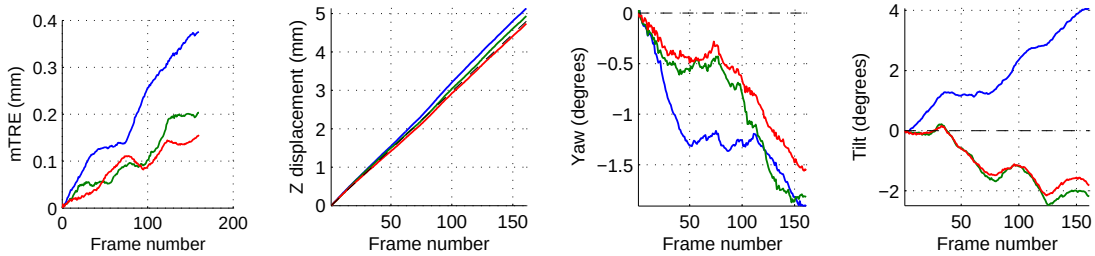
Figure B.1 Speckle phantom trajectory recovery results for 0.03 mm test sequence inter-frame spacing, part 1



(a) Calibration spacing = 0.06 mm, Subject spacing = 0.03 mm

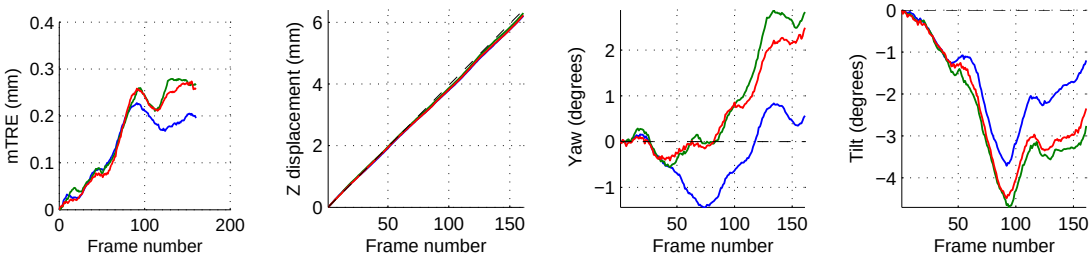


(b) Calibration spacing = 0.08 mm, Subject spacing = 0.03 mm

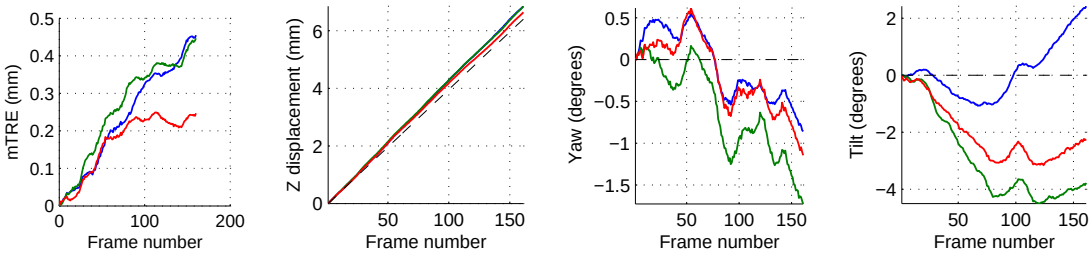


(c) Calibration spacing = 0.1 mm, Subject spacing = 0.03 mm

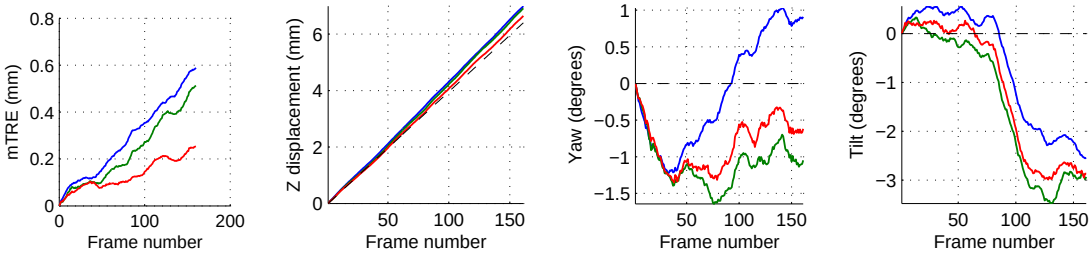
Figure B.2 Speckle phantom trajectory recovery results for 0.03 mm test sequence inter-frame spacing, part 2



(a) Calibration spacing = 0.03 mm, Subject spacing = 0.04 mm

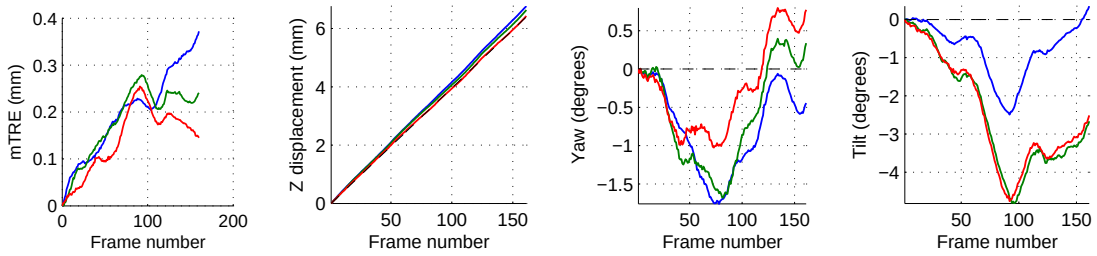


(b) Calibration spacing = 0.04 mm, Subject spacing = 0.04 mm

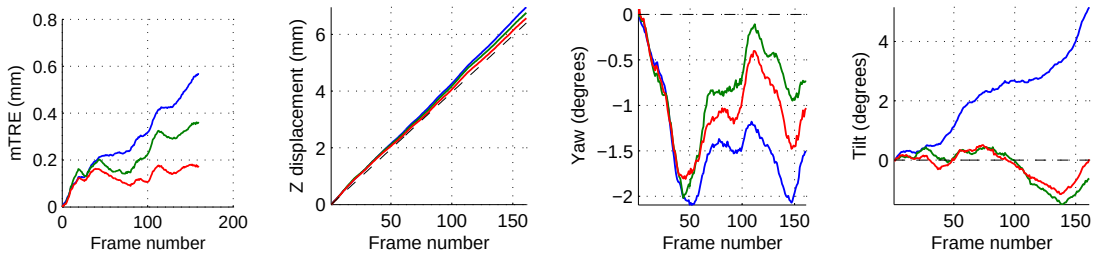


(c) Calibration spacing = 0.05 mm, Subject spacing = 0.04 mm

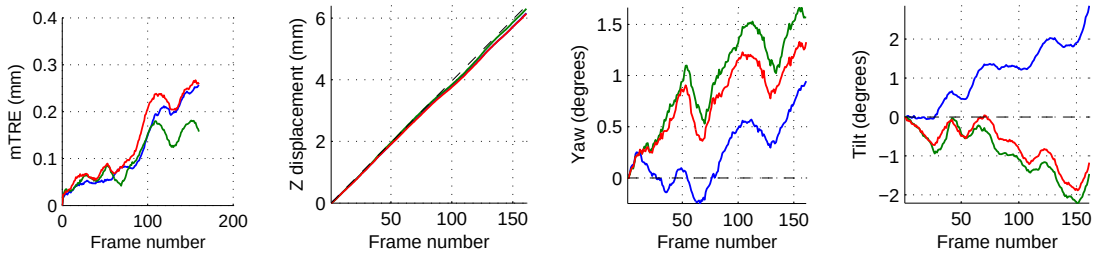
Figure B.3 Speckle phantom trajectory recovery results for 0.04 mm test sequence inter-frame spacing, part 1



(a) Calibration spacing = 0.06 mm, Subject spacing = 0.04 mm

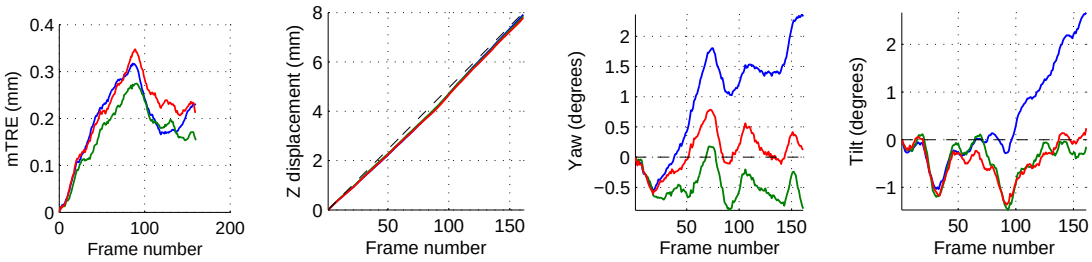


(b) Calibration spacing = 0.08 mm, Subject spacing = 0.04 mm

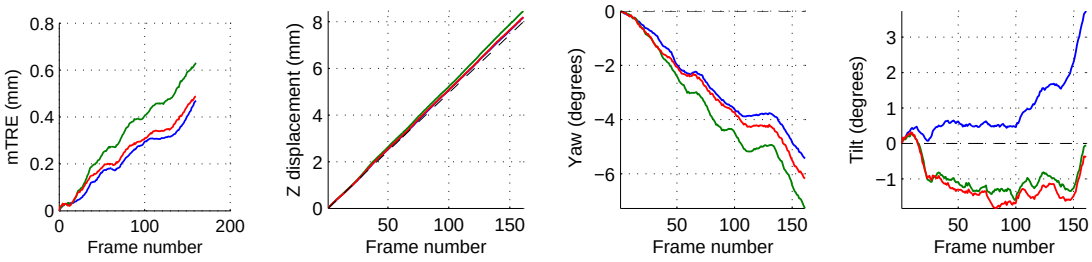


(c) Calibration spacing = 0.1 mm, Subject spacing = 0.04 mm

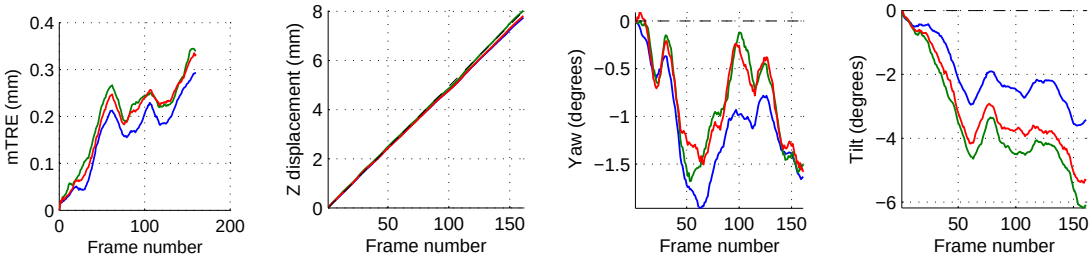
Figure B.4 Speckle phantom trajectory recovery results for 0.03 mm test sequence inter-frame spacing, part 2



(a) Calibration spacing = 0.03 mm, Subject spacing = 0.05 mm

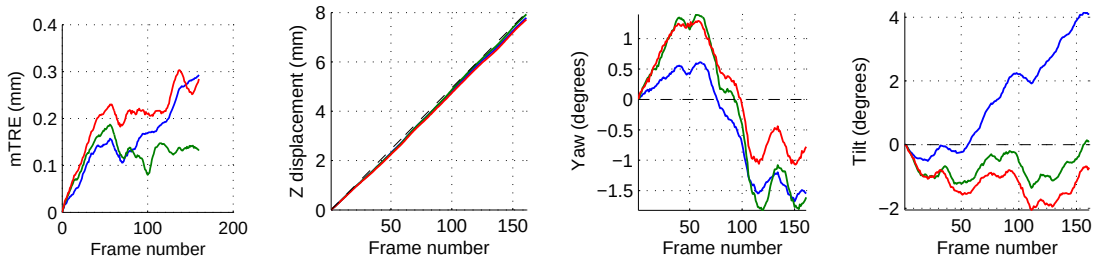


(b) Calibration spacing = 0.04 mm, Subject spacing = 0.05 mm

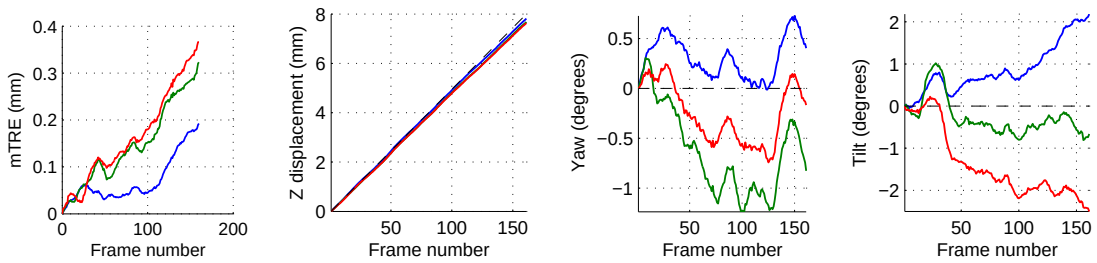


(c) Calibration spacing = 0.05 mm, Subject spacing = 0.05 mm

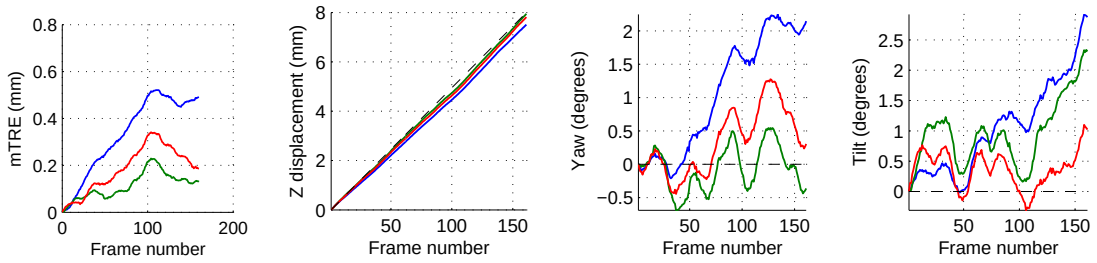
Figure B.5 Speckle phantom trajectory recovery results for 0.05 mm test sequence inter-frame spacing, part 1



(a) Calibration spacing = 0.06 mm, Subject spacing = 0.05 mm

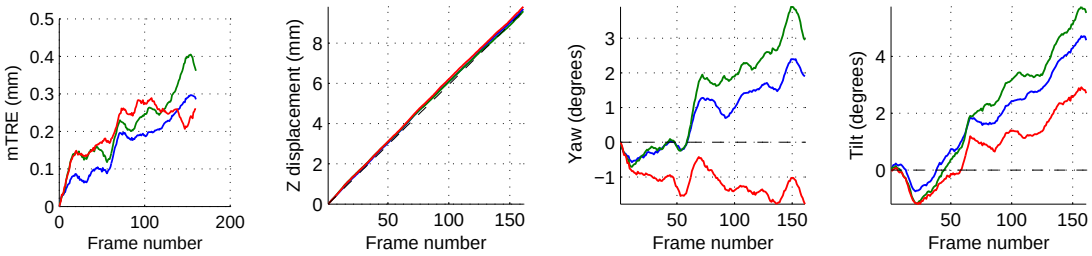


(b) Calibration spacing = 0.08 mm, Subject spacing = 0.05 mm

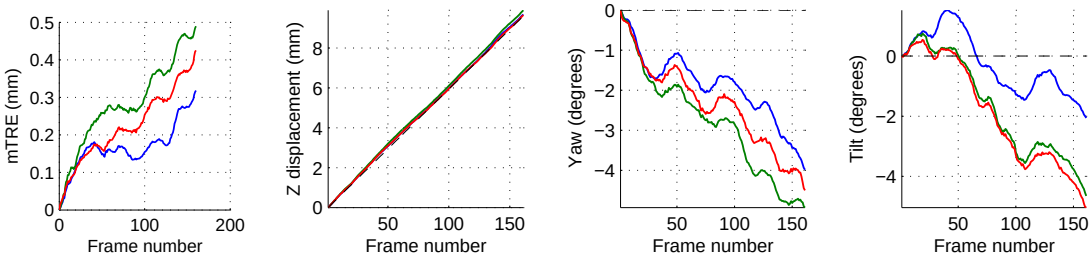


(c) Calibration spacing = 0.1 mm, Subject spacing = 0.05 mm

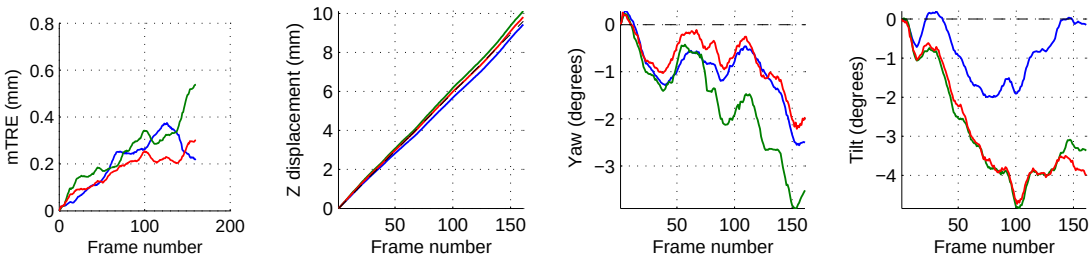
Figure B.6 Speckle phantom trajectory recovery results for 0.05 mm test sequence inter-frame spacing, part 2



(a) Calibration spacing = 0.03 mm, Subject spacing = 0.06 mm

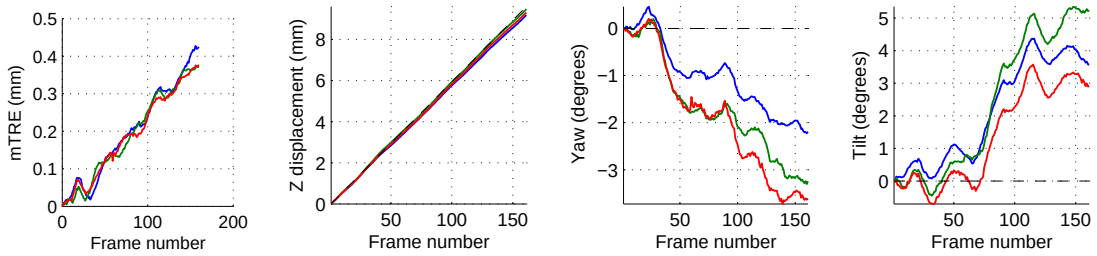


(b) Calibration spacing = 0.04 mm, Subject spacing = 0.06 mm

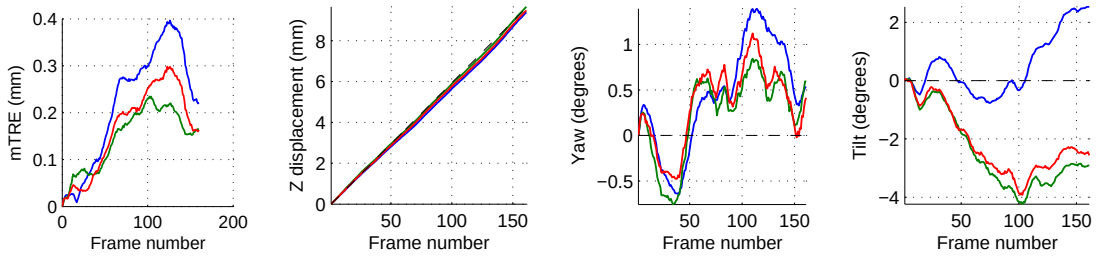


(c) Calibration spacing = 0.05 mm, Subject spacing = 0.06 mm

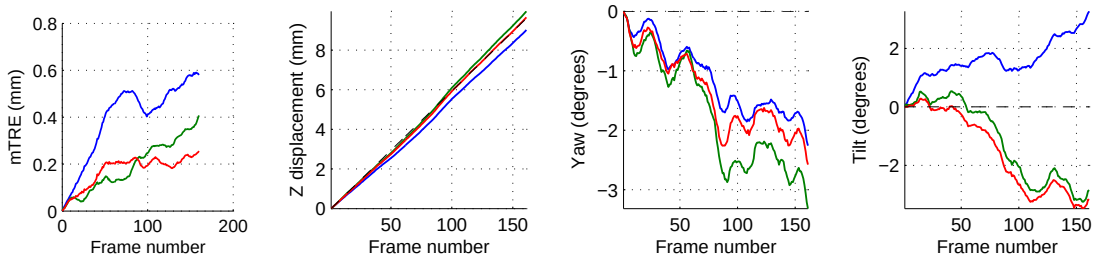
Figure B.7 Speckle phantom trajectory recovery results for 0.06 mm test sequence inter-frame spacing, part 1



(a) Calibration spacing = 0.06 mm, Subject spacing = 0.06 mm

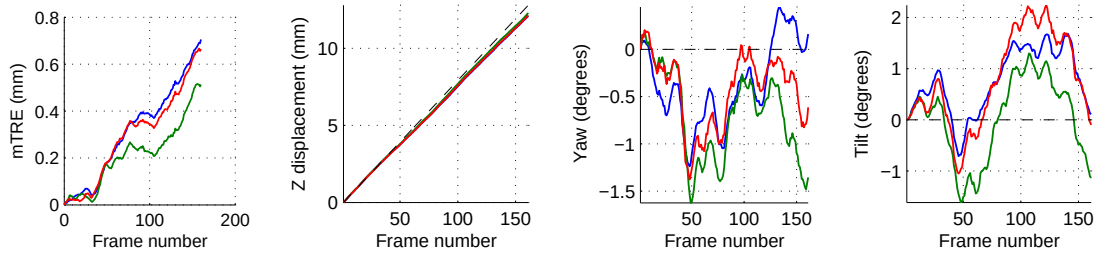


(b) Calibration spacing = 0.08 mm, Subject spacing = 0.06 mm

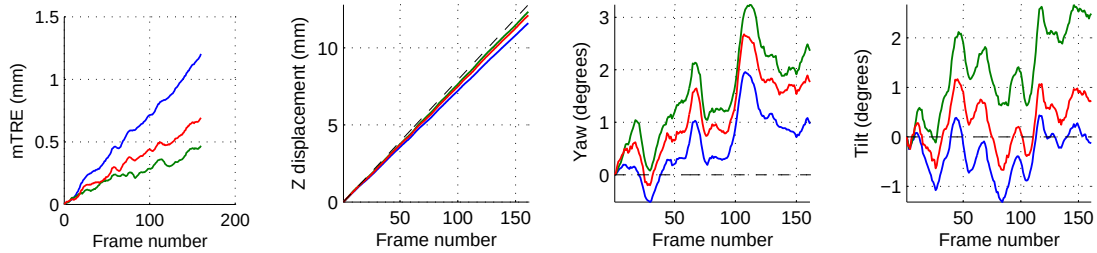


(c) Calibration spacing = 0.1 mm, Subject spacing = 0.06 mm

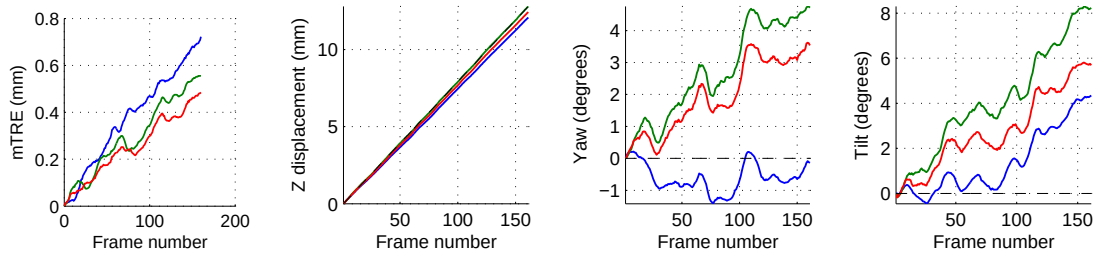
Figure B.8 Speckle phantom trajectory recovery results for 0.03 mm test sequence inter-frame spacing, part 2



(a) Calibration spacing = 0.03 mm, Subject spacing = 0.08 mm

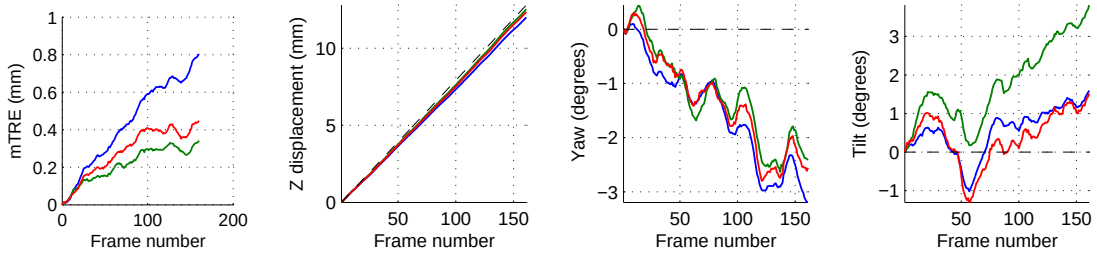


(b) Calibration spacing = 0.04 mm, Subject spacing = 0.08 mm

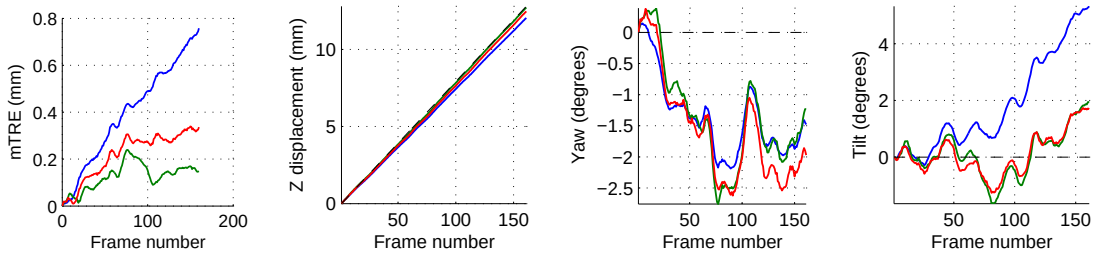


(c) Calibration spacing = 0.05 mm, Subject spacing = 0.08 mm

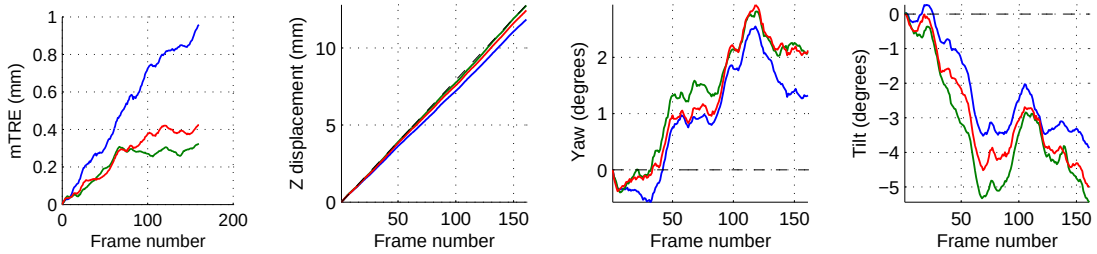
Figure B.9 Speckle phantom trajectory recovery results for 0.08 mm test sequence inter-frame spacing, part 1



(a) Calibration spacing = 0.06 mm, Subject spacing = 0.08 mm

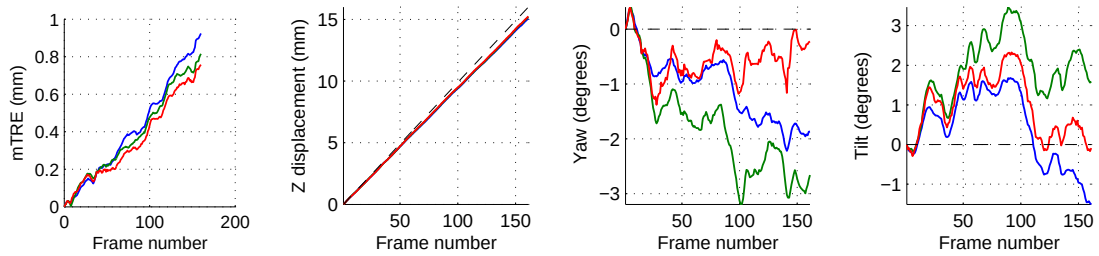


(b) Calibration spacing = 0.08 mm, Subject spacing = 0.08 mm

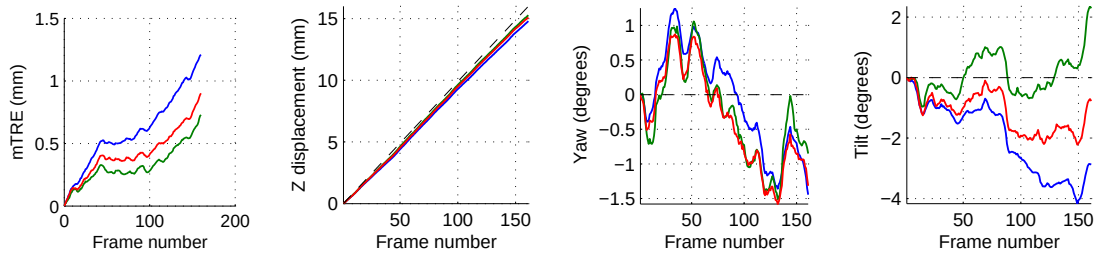


(c) Calibration spacing = 0.1 mm, Subject spacing = 0.08 mm

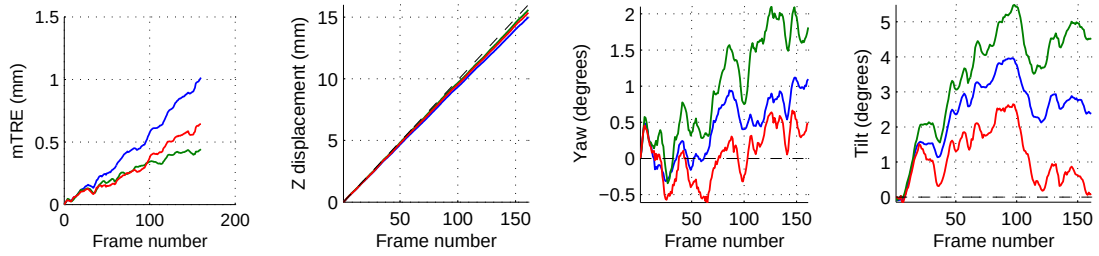
Figure B.10 Speckle phantom trajectory recovery results for 0.08 mm test sequence inter-frame spacing, part 2



(a) Calibration spacing = 0.03 mm, Subject spacing = 0.1 mm

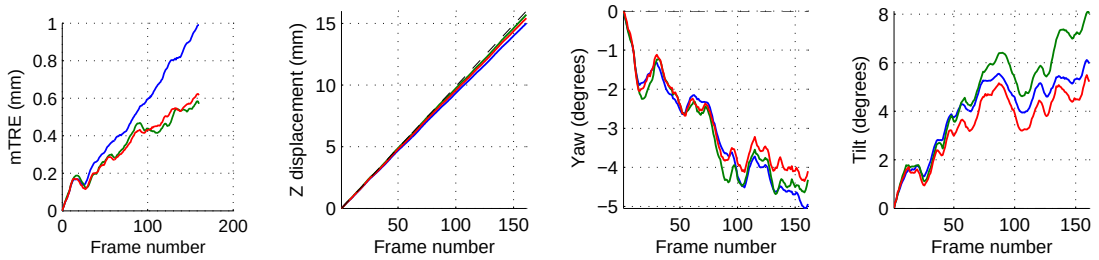


(b) Calibration spacing = 0.04 mm, Subject spacing = 0.1 mm

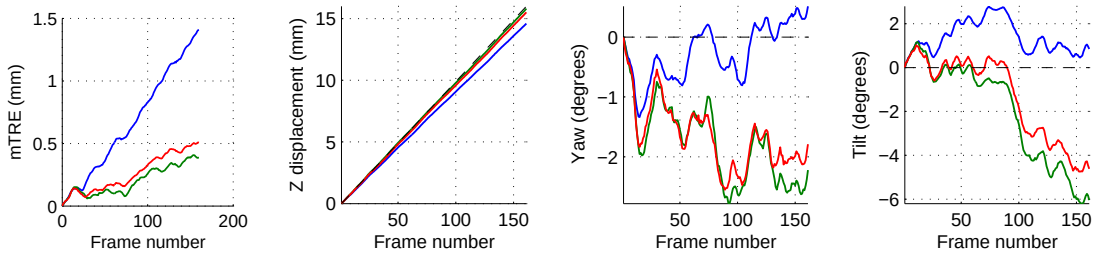


(c) Calibration spacing = 0.05 mm, Subject spacing = 0.1 mm

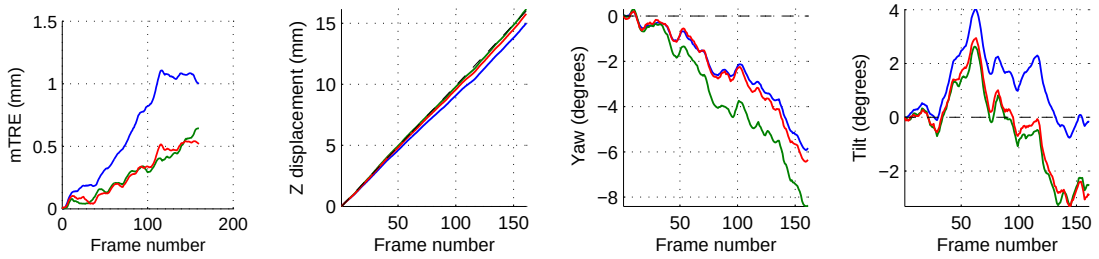
Figure B.11 Speckle phantom trajectory recovery results for 0.1 mm test sequence inter-frame spacing, part 1



(a) Calibration spacing = 0.06 mm, Subject spacing = 0.1 mm



(b) Calibration spacing = 0.08 mm, Subject spacing = 0.1 mm



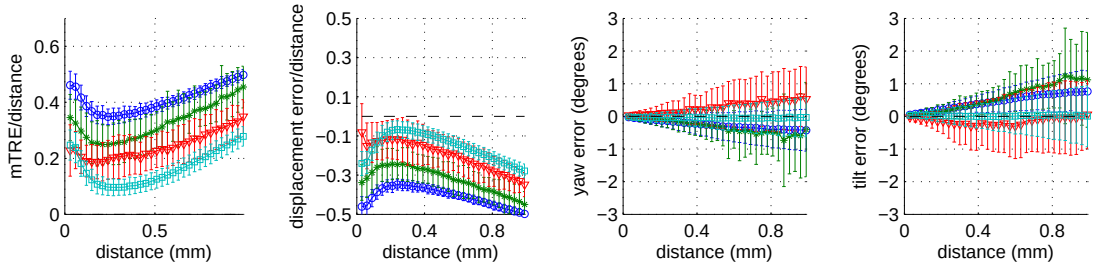
(c) Calibration spacing = 0.1 mm, Subject spacing = 0.1 mm

Figure B.12 Speckle phantom trajectory recovery results for 0.03 mm test sequence inter-frame spacing, part 2

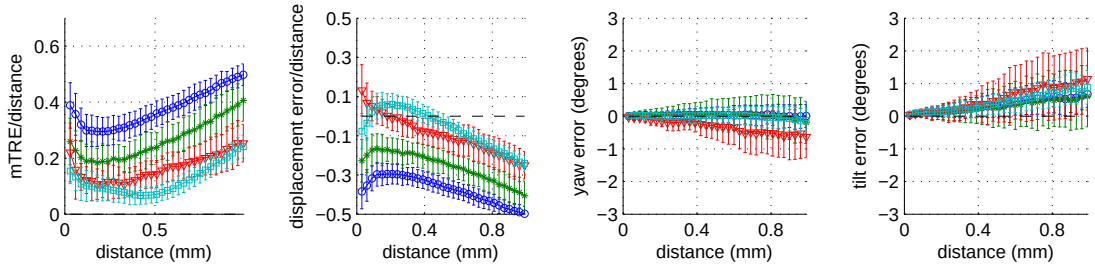
Appendix C

Full set of results for tissue invariant distance estimation

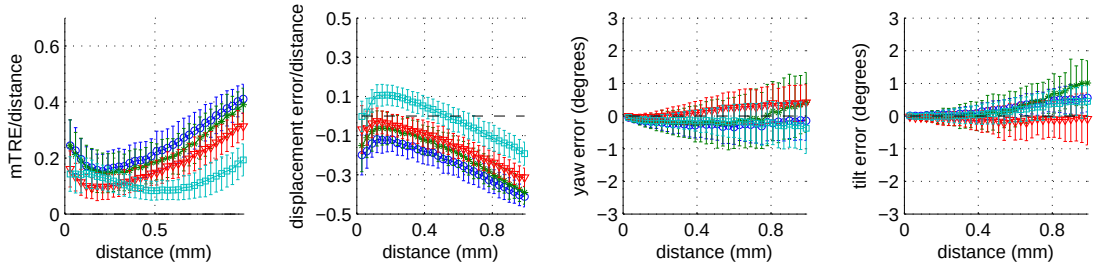
This appendix presents the full set of results obtained for the estimation of elevational separation between pairs of frames in real ultrasound scans of animal tissue during the experiments carried out for chapter 6. Every row of each figure represents the results obtained for a single scan. From left to right, the plots in a given row show (1) the ratio of the mTRE to the true displacement, (2) the ratio of the estimated displacement error to the true displacement, (3) the yaw error and (4) the tilt error measured using the centroid of the grid of image patches as the centre of rotation. The blue circles represent the results obtained using the nominal speckle decorrelation model derived from the speckle phantom. The green stars represent the results obtained using the speckle detector of Prager et al. [164]. The red triangles represent the results obtained using an adaptation of the heuristic approach of Gee et al. [72]. The cyan squares represent the results obtained using the learning based approach of chapter 6. The error bars represent one standard deviation.



(a) Pork tenderloin 1: 163 frames, 0.03 mm intervals

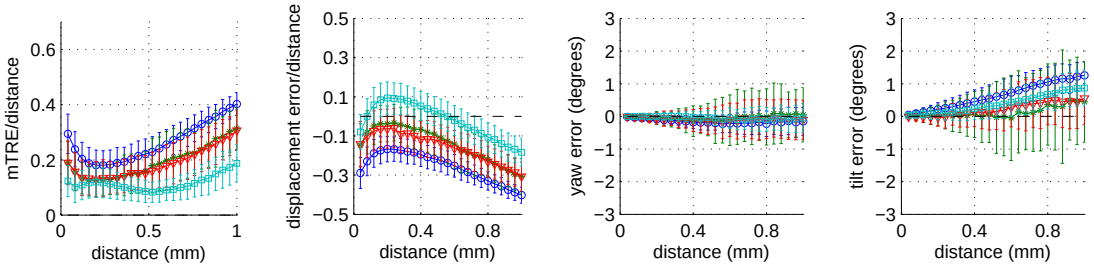


(b) Pork tenderloin 2: 161 frames, 0.03 mm intervals

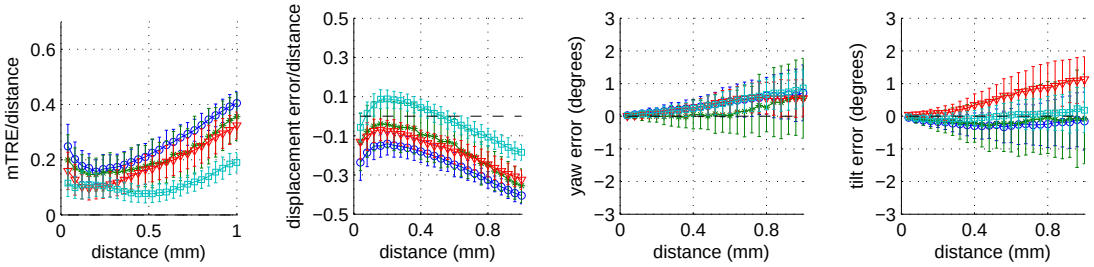


(c) Pork tenderloin 3: 161 frames, 0.03 mm intervals

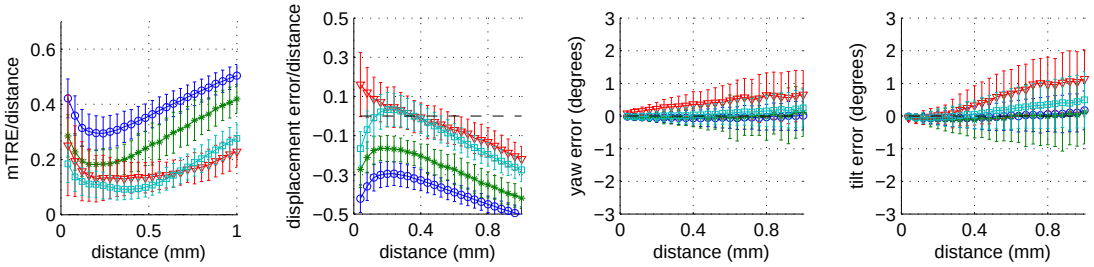
Figure C.1 Distance estimation results for pork tenderloin, 0.03 mm intervals



(a) Pork tenderloin 4: 161 frames, 0.04 mm intervals

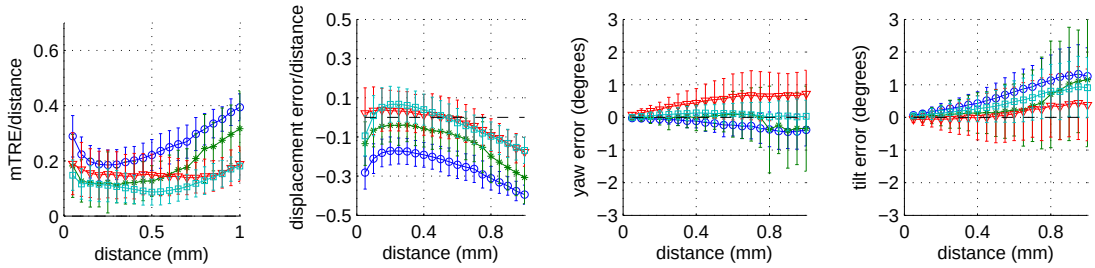


(b) Pork tenderloin 5: 161 frames, 0.04 mm intervals

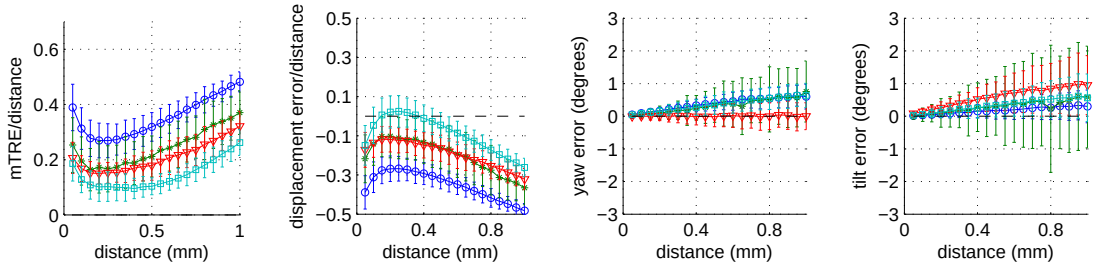


(c) Pork tenderloin 6: 161 frames, 0.04 mm intervals

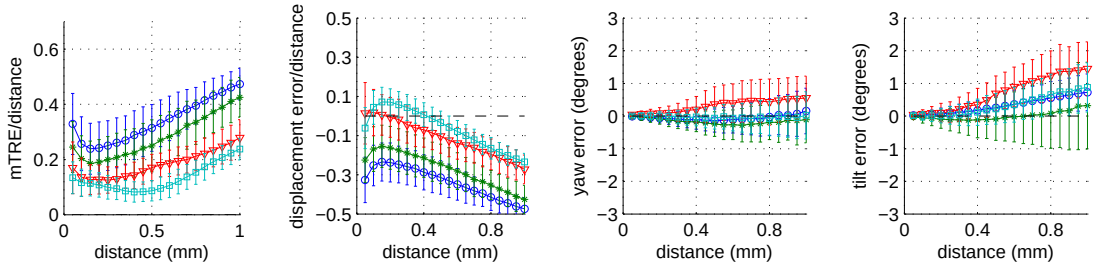
Figure C.2 Distance estimation results for pork tenderloin, 0.04 mm intervals



(a) Pork tenderloin 7: 161 frames, 0.05 mm intervals

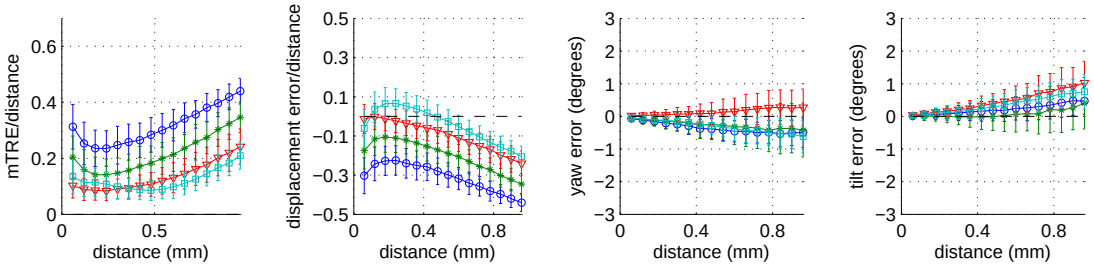


(b) Pork tenderloin 8: 161 frames, 0.05 mm intervals

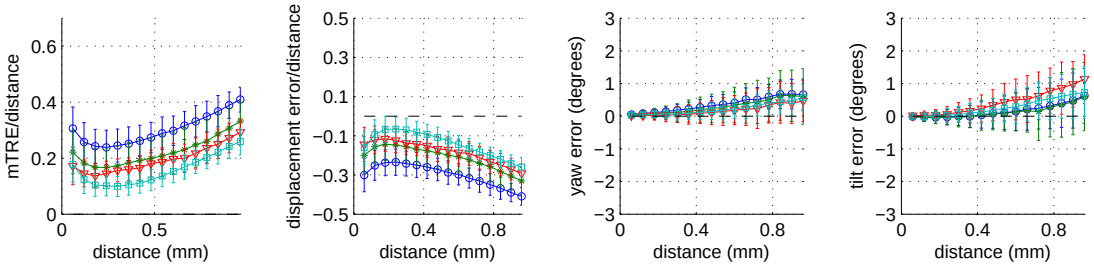


(c) Pork tenderloin 9: 161 frames, 0.05 mm intervals

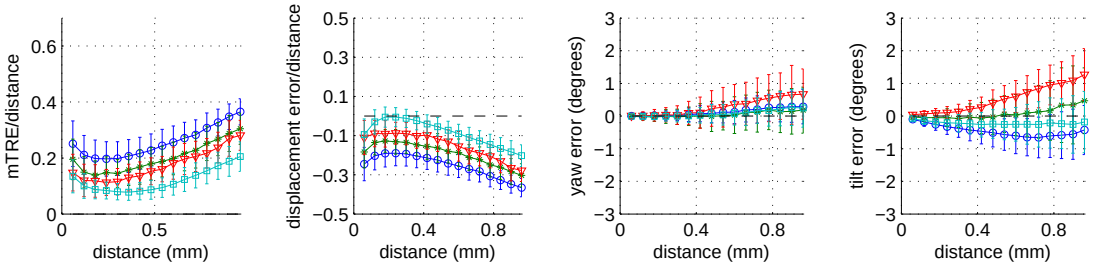
Figure C.3 Distance estimation results for pork tenderloin, 0.05 mm intervals



(a) Pork tenderloin 10: 161 frames, 0.06 mm intervals

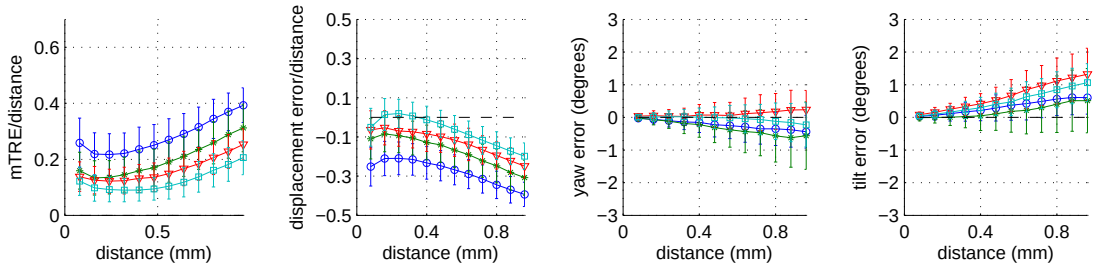


(b) Pork tenderloin 11: 161 frames, 0.06 mm intervals

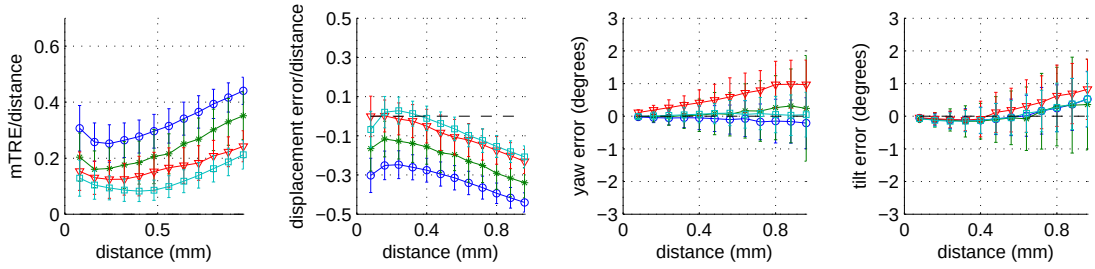


(c) Pork tenderloin 12: 161 frames, 0.06 mm intervals

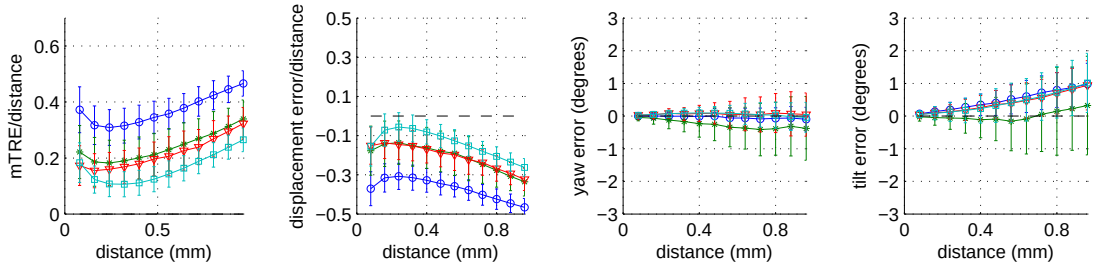
Figure C.4 Distance estimation results for pork tenderloin, 0.06 mm intervals



(a) Pork tenderloin 13: 161 frames, 0.08 mm intervals

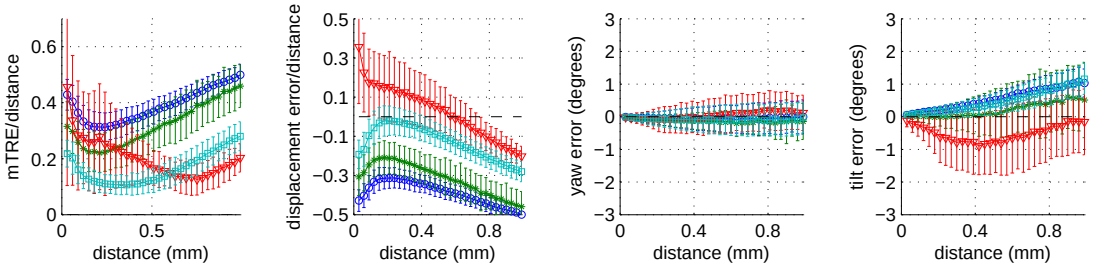


(b) Pork tenderloin 14: 161 frames, 0.08 mm intervals

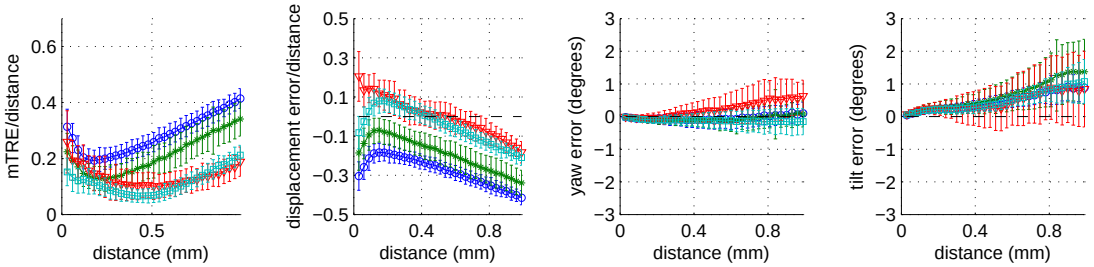


(c) Pork tenderloin 15: 161 frames, 0.08 mm intervals

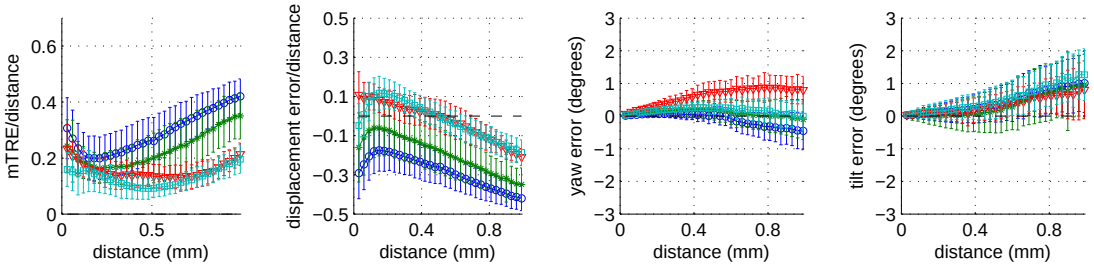
Figure C.5 Distance estimation results for pork tenderloin, 0.08 mm intervals



(a) Turkey breast 1: 161 frames, 0.03 mm intervals

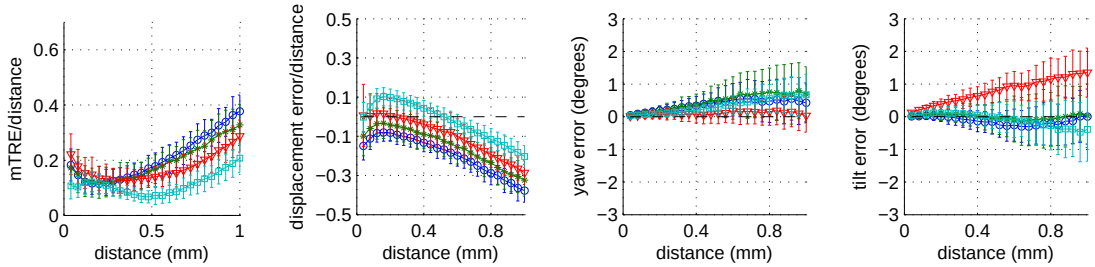


(b) Turkey breast 2: 161 frames, 0.03 mm intervals

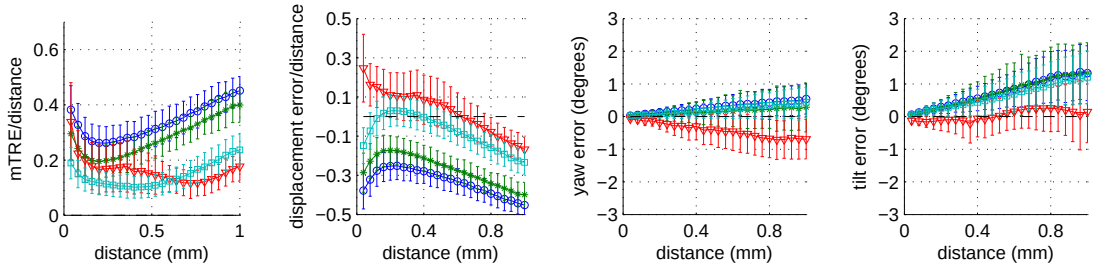


(c) Turkey breast 3: 161 frames, 0.03 mm intervals

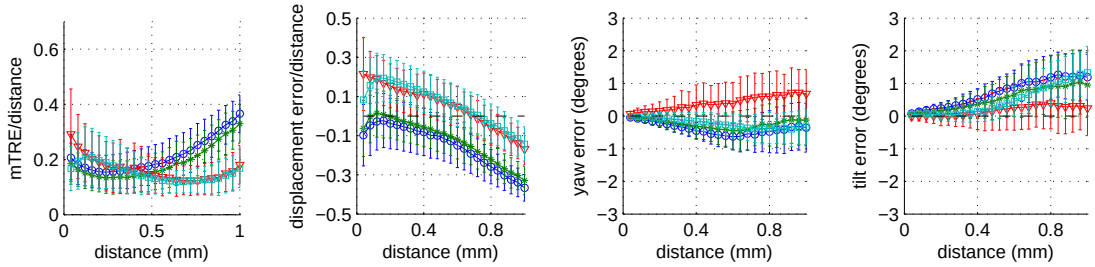
Figure C.6 Distance estimation results for turkey breast, 0.03 mm intervals



(a) Turkey breast 4: 161 frames, 0.04 mm intervals

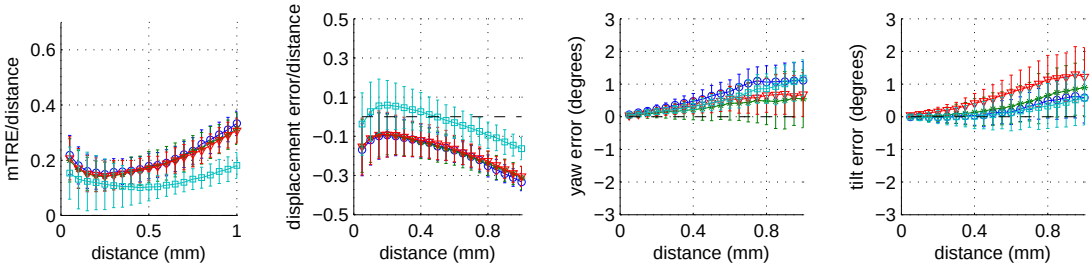


(b) Turkey breast 5: 161 frames, 0.04 mm intervals

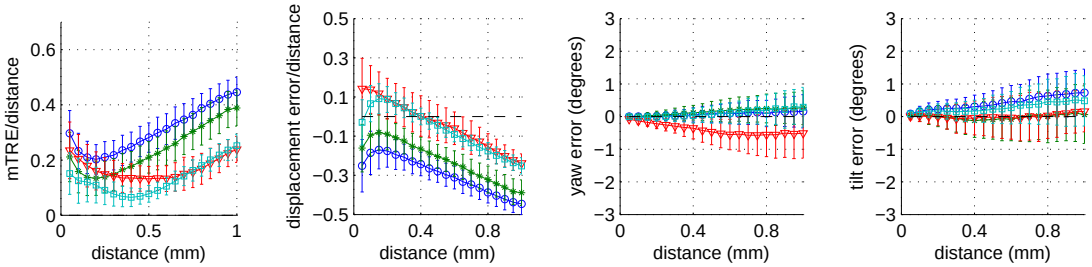


(c) Turkey breast 6: 161 frames, 0.04 mm intervals

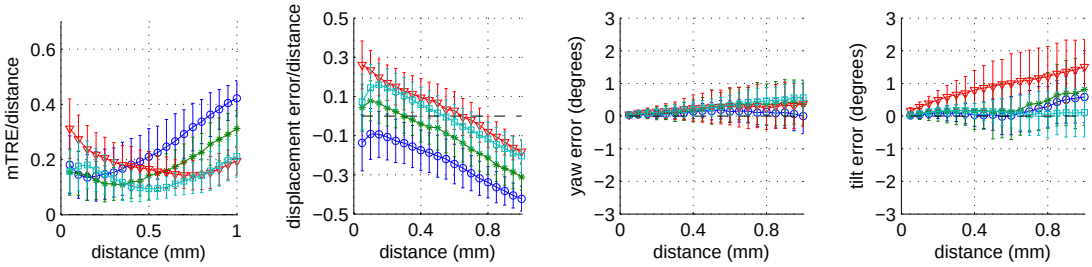
Figure C.7 Distance estimation results for turkey breast, 0.04 mm intervals



(a) Turkey breast 7: 161 frames, 0.05 mm intervals

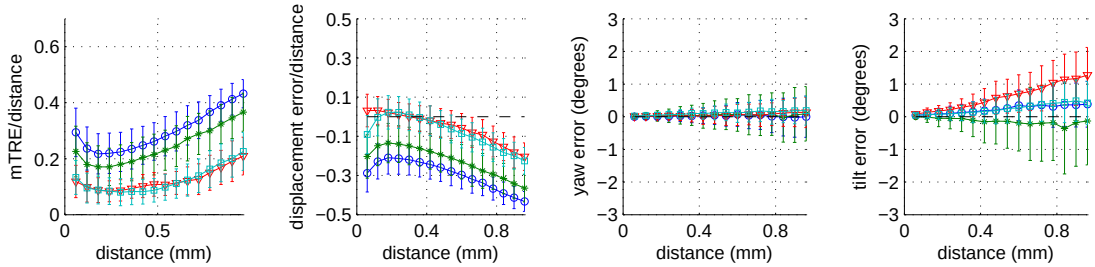


(b) Turkey breast 8: 161 frames, 0.05 mm intervals

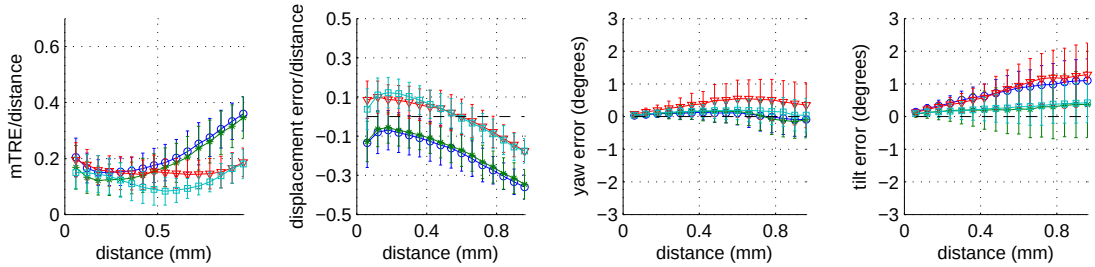


(c) Turkey breast 9: 161 frames, 0.05 mm intervals

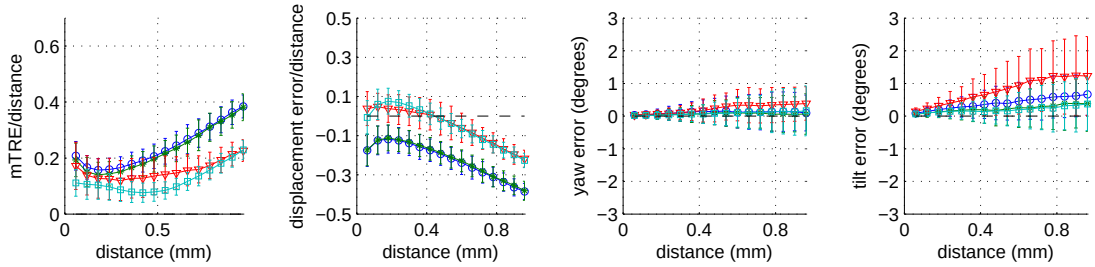
Figure C.8 Distance estimation results for turkey breast, 0.05 mm intervals



(a) Turkey breast 10: 161 frames, 0.06 mm intervals

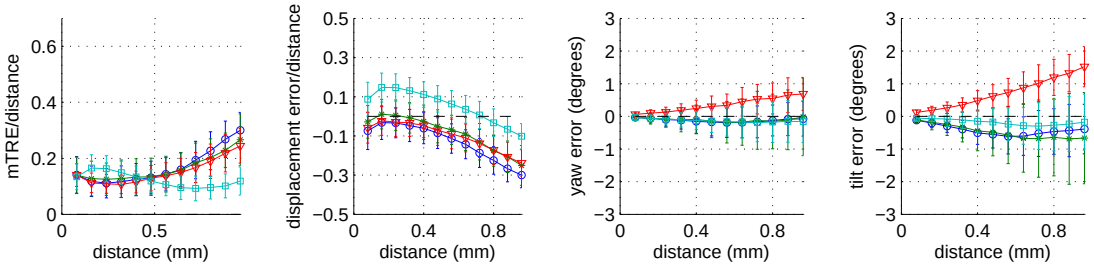


(b) Turkey breast 11: 161 frames, 0.06 mm intervals

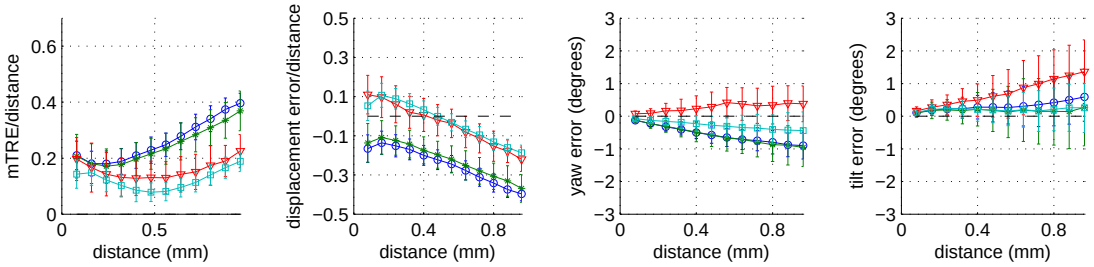


(c) Turkey breast 12: 161 frames, 0.06 mm intervals

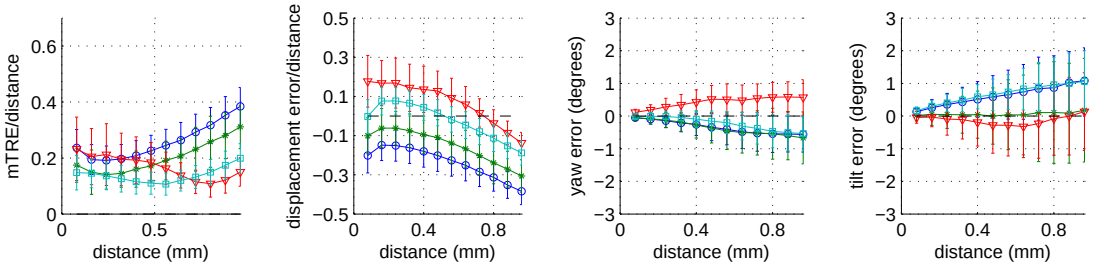
Figure C.9 Distance estimation results for turkey breast, 0.06 mm intervals



(a) Turkey breast 13: 161 frames, 0.08 mm intervals

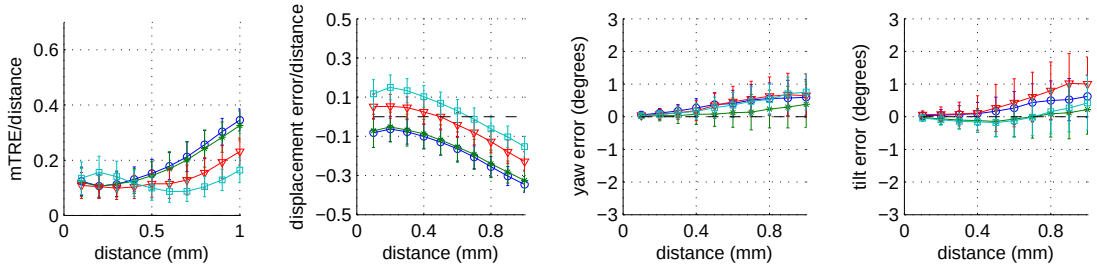


(b) Turkey breast 14: 167 frames, 0.08 mm intervals

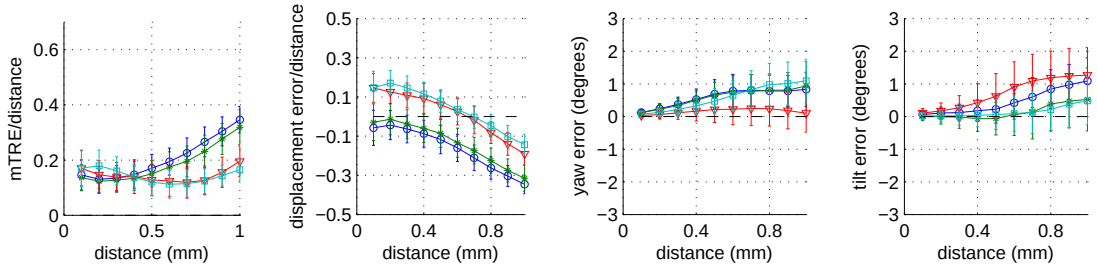


(c) Turkey breast 15: 161 frames, 0.08 mm intervals

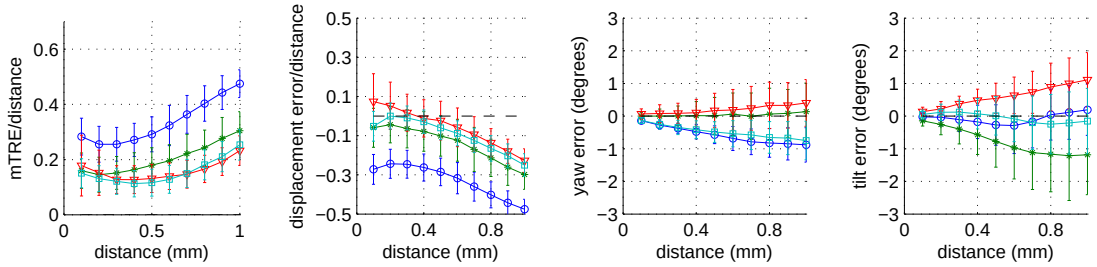
Figure C.10 Distance estimation results for turkey breast, 0.08 mm intervals



(a) Turkey breast 16: 161 frames, 0.1 mm intervals

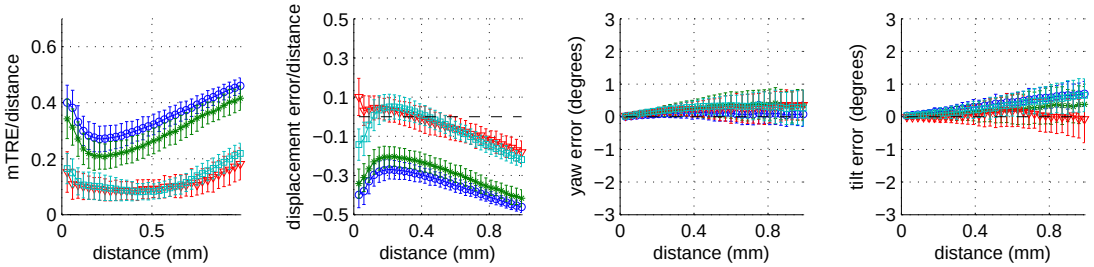


(b) Turkey breast 17: 161 frames, 0.1 mm intervals

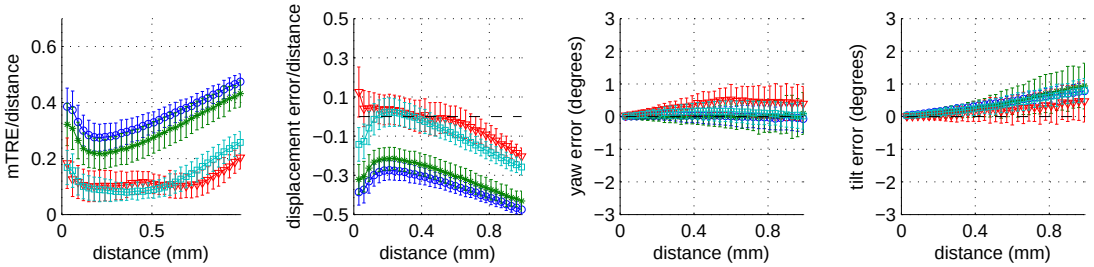


(c) Turkey breast 18: 161 frames, 0.1 mm intervals

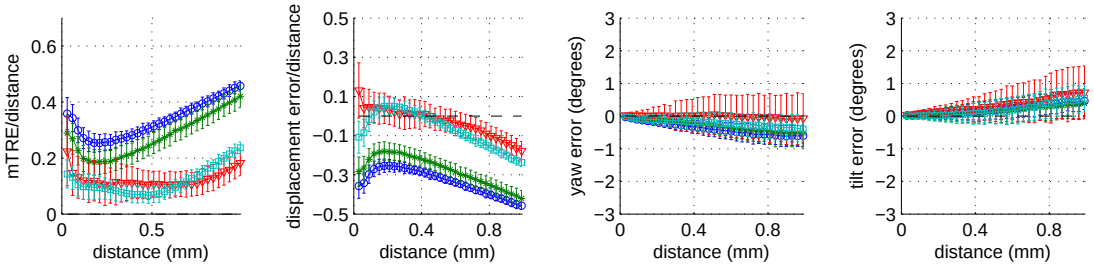
Figure C.11 Distance estimation results for turkey breast, 0.1 mm intervals



(a) Beef brisket 1: 161 frames, 0.03 mm intervals

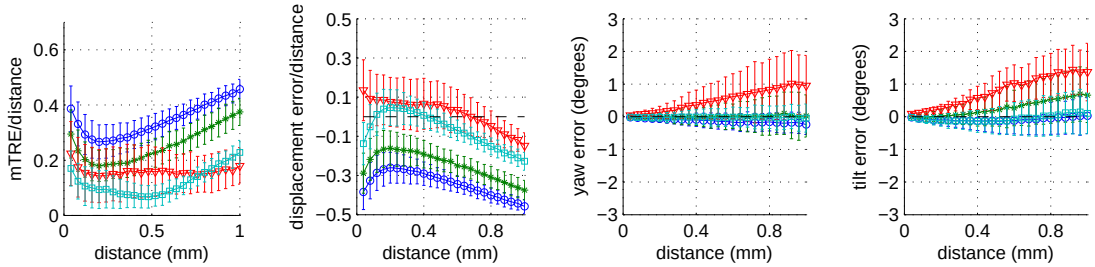


(b) Beef brisket 2: 161 frames, 0.03 mm intervals

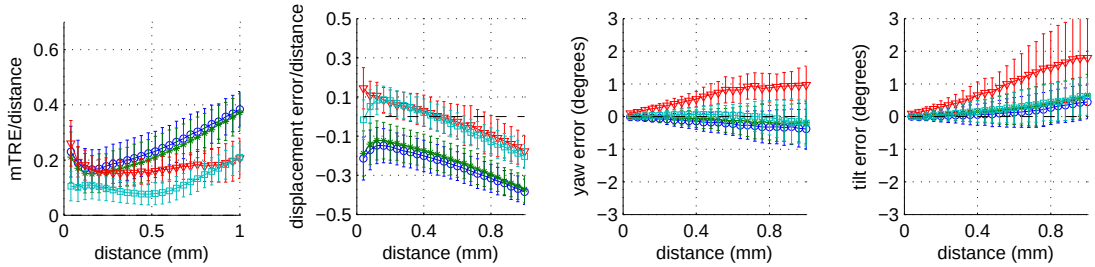


(c) Beef brisket 3: 161 frames, 0.03 mm intervals

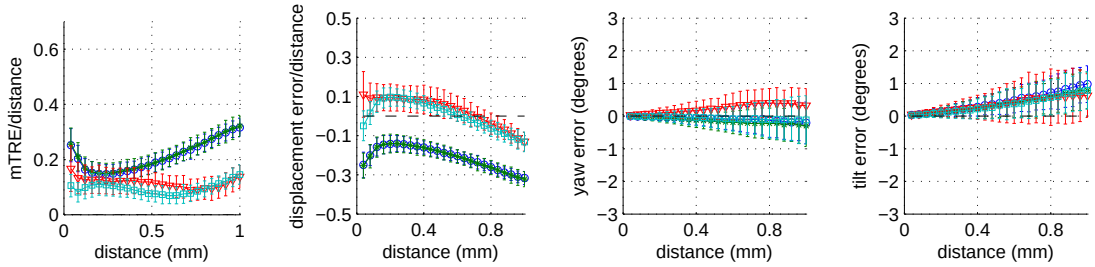
Figure C.12 Distance estimation results for beef brisket, 0.03 mm intervals



(a) Beef brisket 4: 161 frames, 0.04 mm intervals

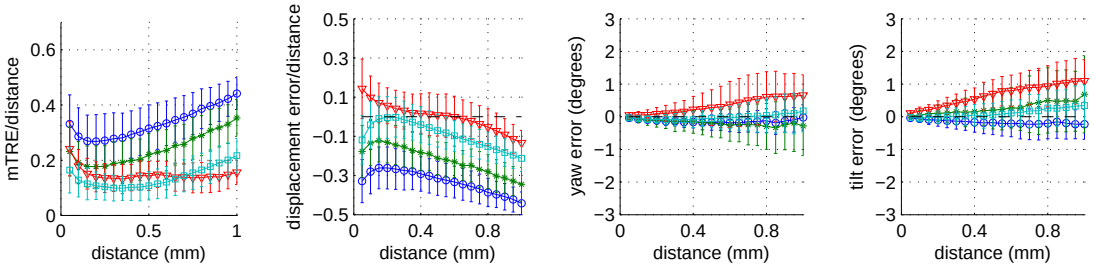


(b) Beef brisket 5: 161 frames, 0.04 mm intervals

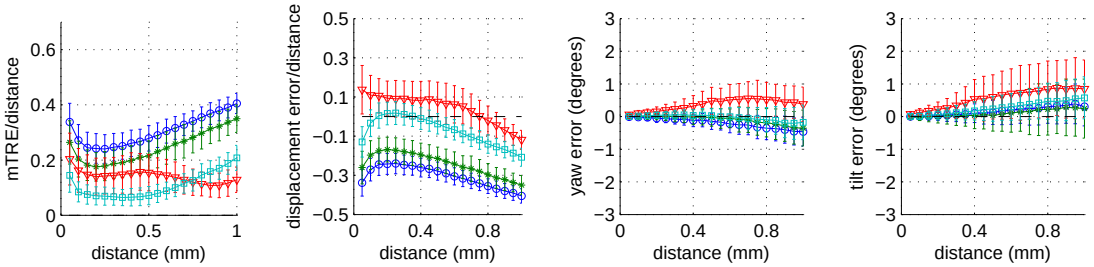


(c) Beef brisket 6: 161 frames, 0.04 mm intervals

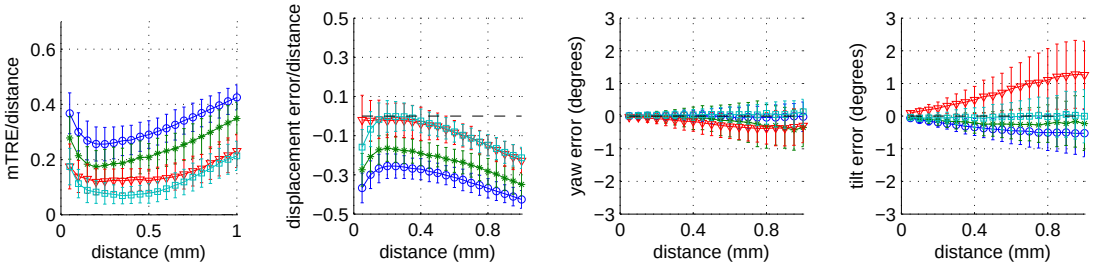
Figure C.13 Distance estimation results for beef brisket, 0.04 mm intervals



(a) Beef brisket 7: 161 frames, 0.05 mm intervals

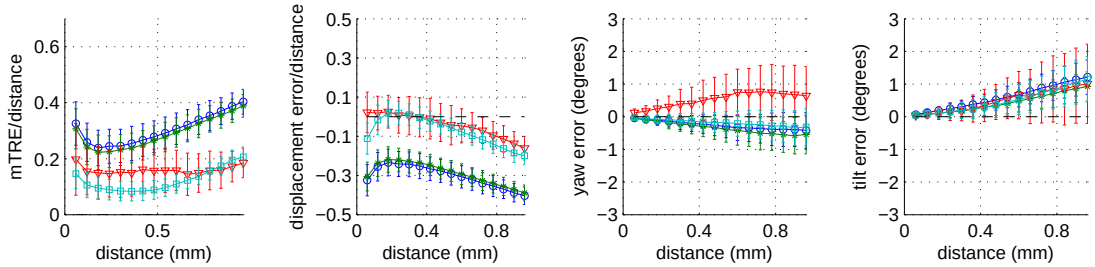


(b) Beef brisket 8: 161 frames, 0.05 mm intervals

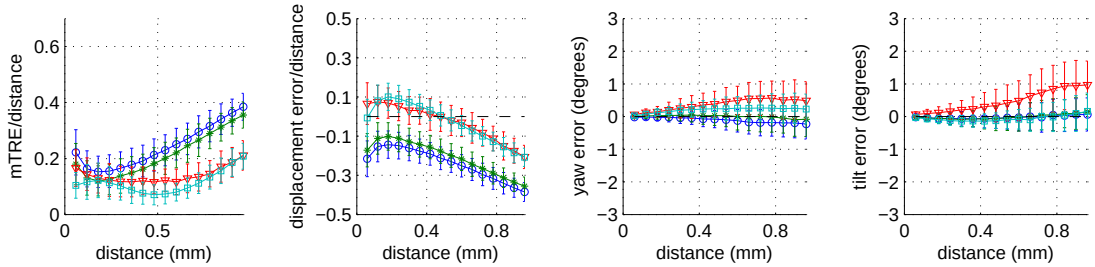


(c) Beef brisket 9: 161 frames, 0.05 mm intervals

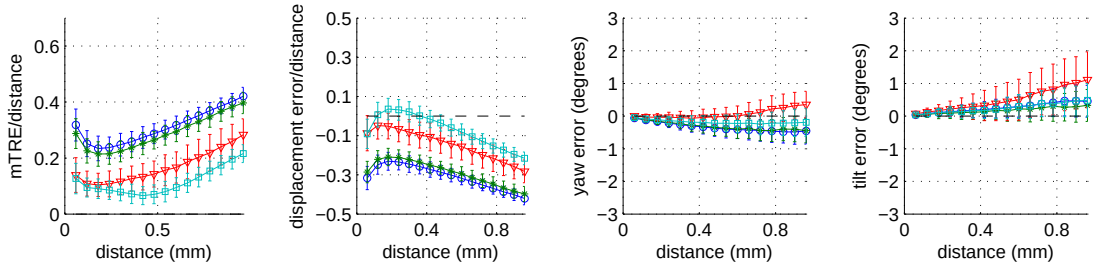
Figure C.14 Distance estimation results for beef brisket, 0.05 mm intervals



(a) Beef brisket 10: 161 frames, 0.06 mm intervals

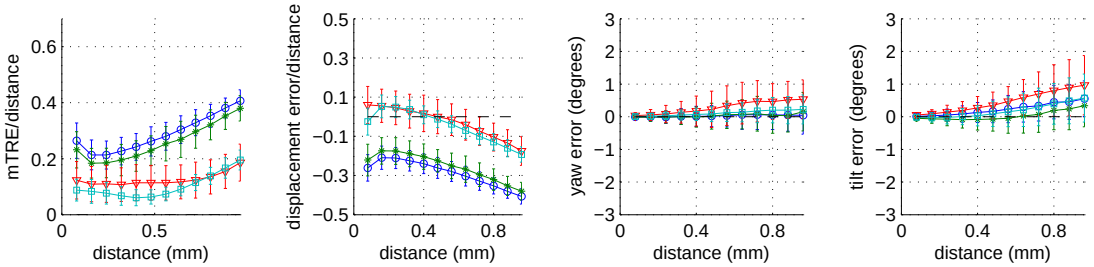


(b) Beef brisket 11: 161 frames, 0.06 mm intervals

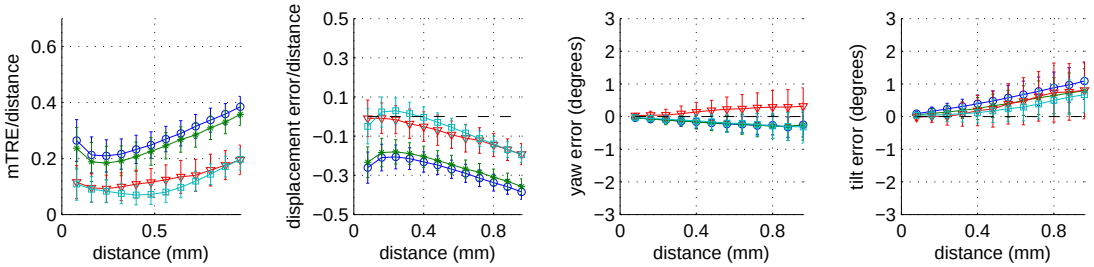


(c) Beef brisket 12: 161 frames, 0.06 mm intervals

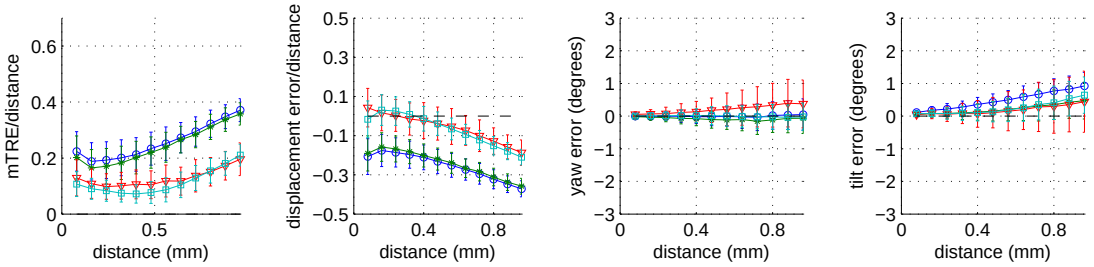
Figure C.15 Distance estimation results for beef brisket, 0.06 mm intervals



(a) Beef brisket 13: 161 frames, 0.08 mm intervals

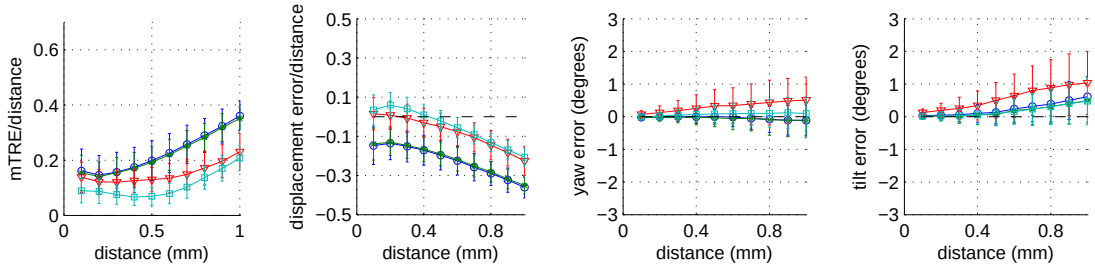


(b) Beef brisket 14: 161 frames, 0.08 mm intervals

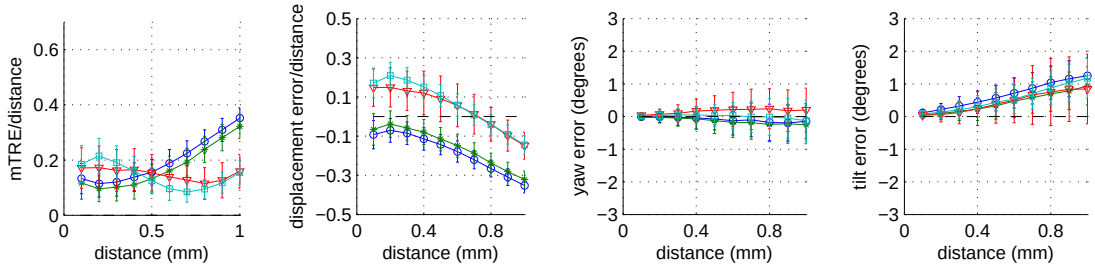


(c) Beef brisket 15: 161 frames, 0.08 mm intervals

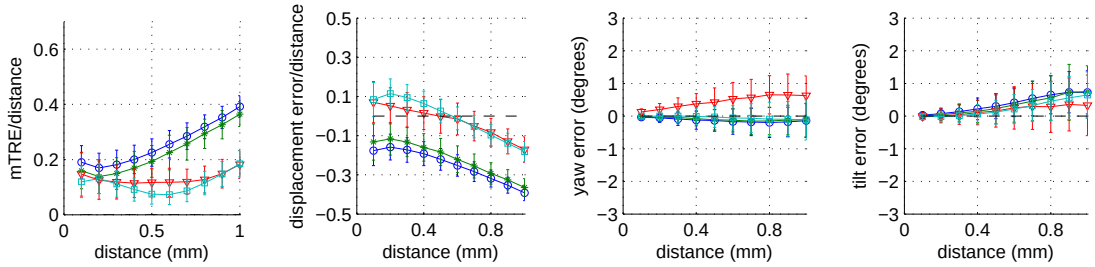
Figure C.16 Distance estimation results for beef brisket, 0.08 mm intervals



(a) Beef brisket 16: 161 frames, 0.1 mm intervals



(b) Beef brisket 17: 161 frames, 0.1 mm intervals



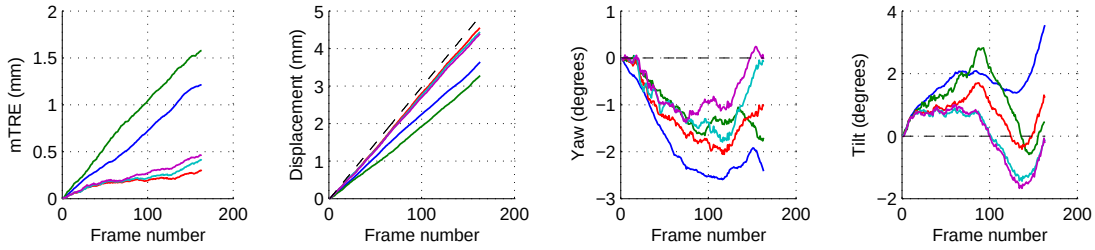
(c) Beef brisket 18: 161 frames, 0.1 mm intervals

Figure C.17 Distance estimation results for beef brisket, 0.1 mm intervals

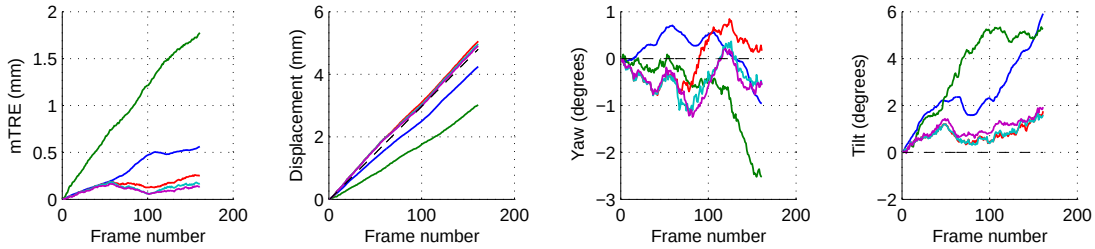
Appendix D

Full set of results for monotonic trajectory estimation in real tissue

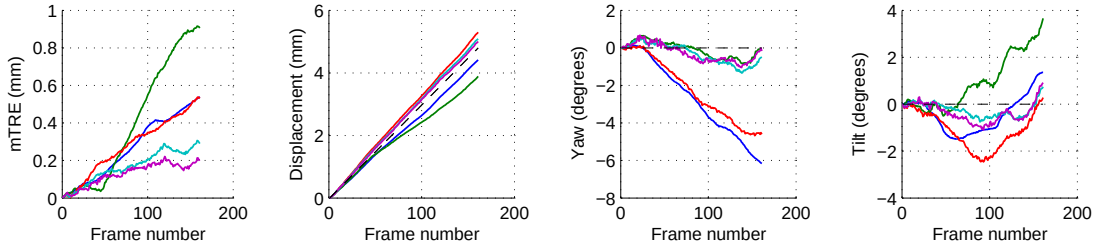
This appendix contains the full set of results obtained for the experiments relating to the recovery of monotonic transducer trajectories in animal tissue scans described in Chapter 7. Every row of each figure represents the results obtained for a single scan. From left to right, the plots in a given row show (1) the progression of the mTRE over frame acquisition steps, (2) the estimated displacement, (3) the estimated yaw (should be zero) and (4) the estimated tilt (should be zero) evaluated using the centroid of the image patches as the centre of rotation. The dashed line corresponds to the ground truth measurements obtained from the sub-millimeter position tracking device. The blue line corresponds to the NN method. The green line corresponds the NMS method (probabilistic fusion approach without measurement selection). The red, cyan and magenta lines represent the HYP-K1, HYP-K2 an HYP-K3 methods, i.e. the iterative measurement selection and fusion approach of Chapter 7 using the κ_1, κ_2 and κ_3 cost functions, respectively, to compute an initial estimate of the trajectory.



(a) Pork tenderloin 1: 163 frames, 0.03 mm intervals

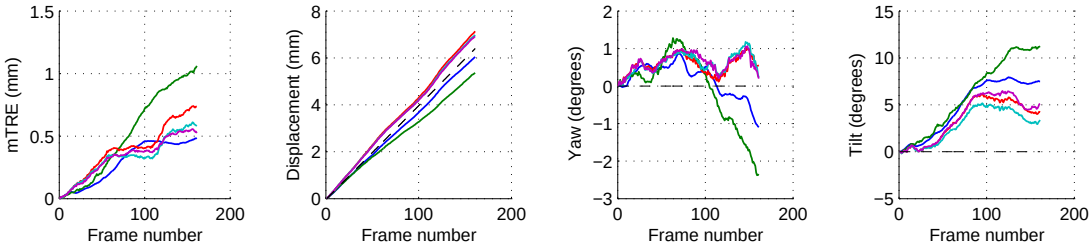


(b) Pork tenderloin 2: 161 frames, 0.03 mm intervals

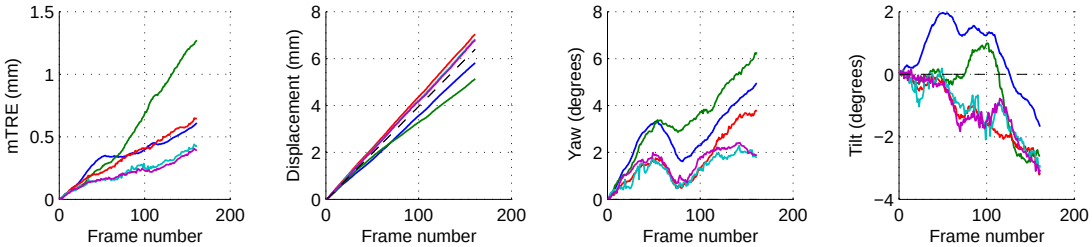


(c) Pork tenderloin 3: 161 frames, 0.03 mm intervals

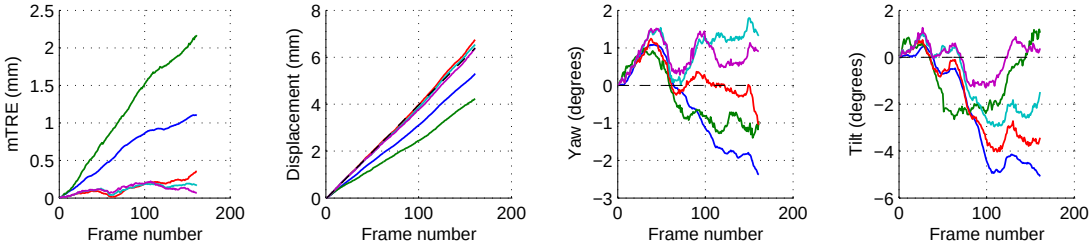
Figure D.1 Pork tenderloin trajectory recovery results, 0.03 mm intervals.



(a) Pork tenderloin 4: 161 frames, 0.04 mm intervals

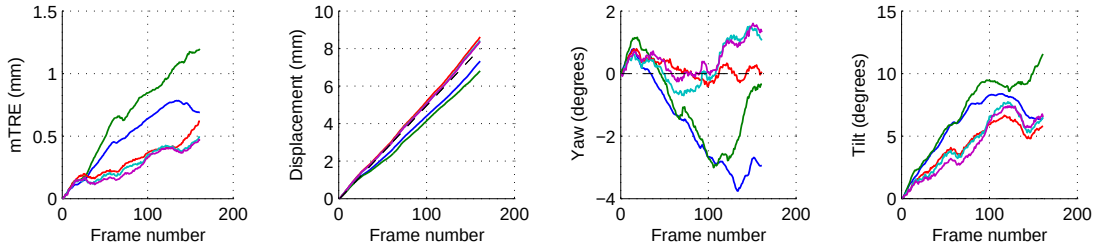


(b) Pork tenderloin 5: 161 frames, 0.04 mm intervals

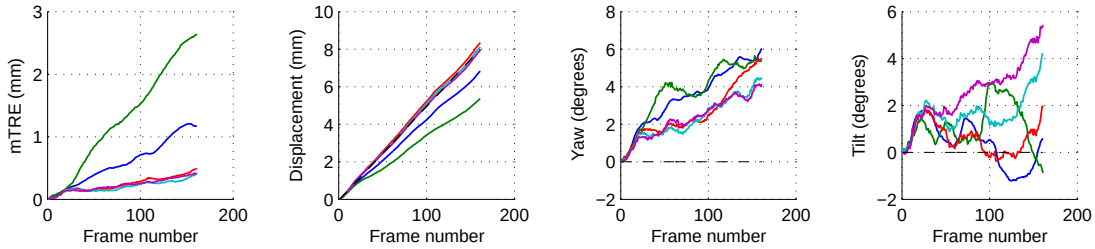


(c) Pork tenderloin 6: 161 frames, 0.04 mm intervals

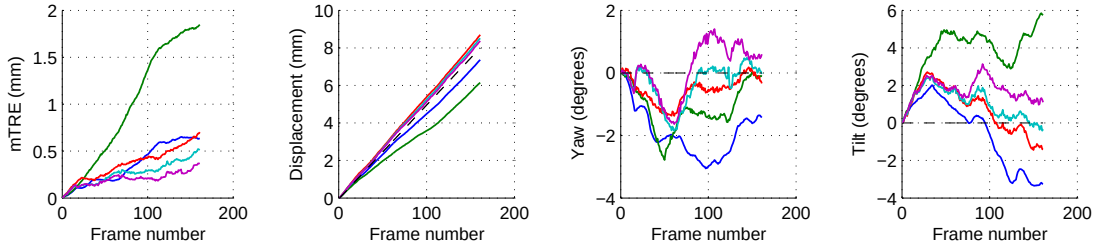
Figure D.2 Pork tenderloin trajectory recovery results, 0.04 mm intervals.



(a) Pork tenderloin 7: 161 frames, 0.05 mm intervals

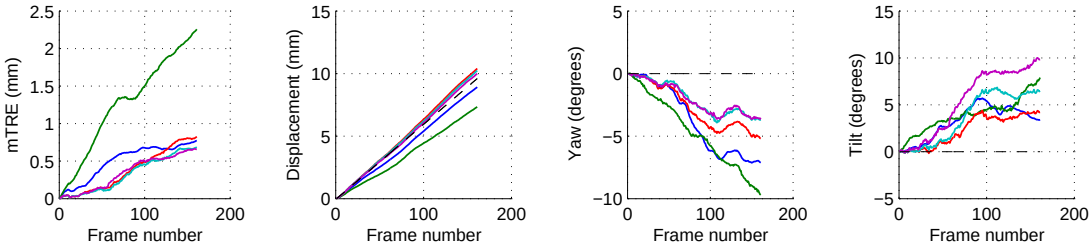


(b) Pork tenderloin 8: 161 frames, 0.05 mm intervals

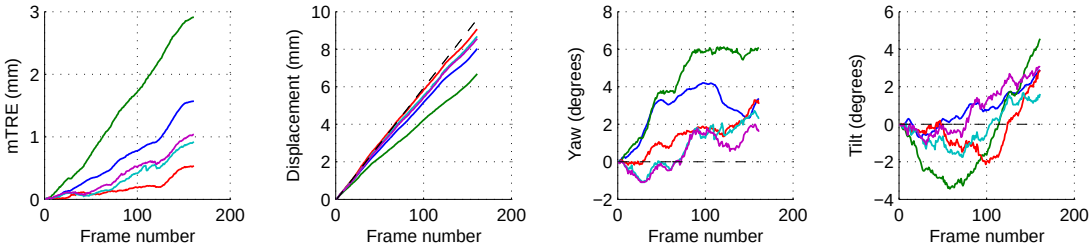


(c) Pork tenderloin 9: 161 frames, 0.05 mm intervals

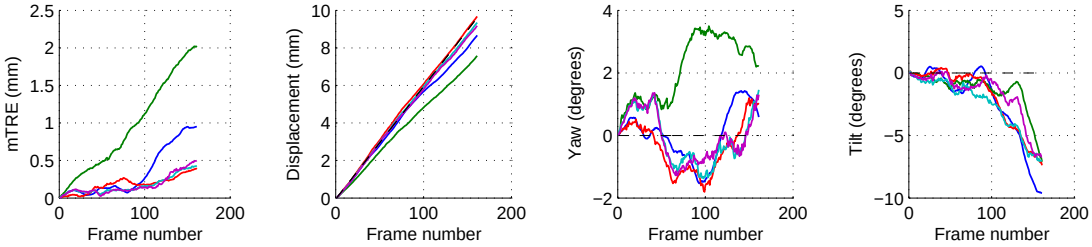
Figure D.3 Pork tenderloin trajectory recovery results, 0.05 mm intervals.



(a) Pork tenderloin 10: 161 frames, 0.06 mm intervals

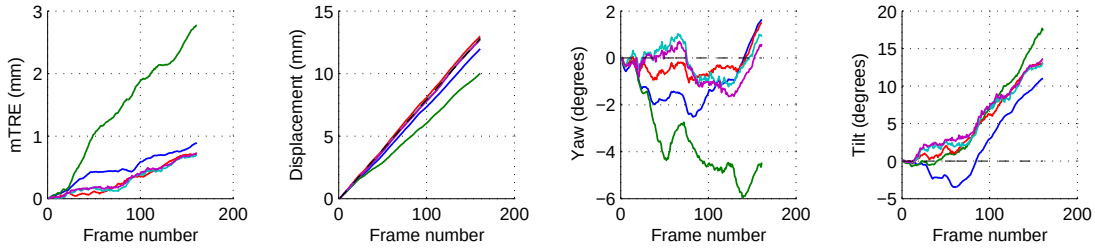


(b) Pork tenderloin 11: 161 frames, 0.06 mm intervals

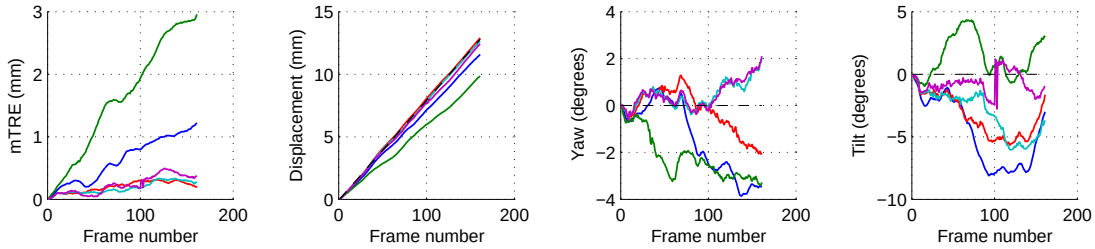


(c) Pork tenderloin 12: 161 frames, 0.06 mm intervals

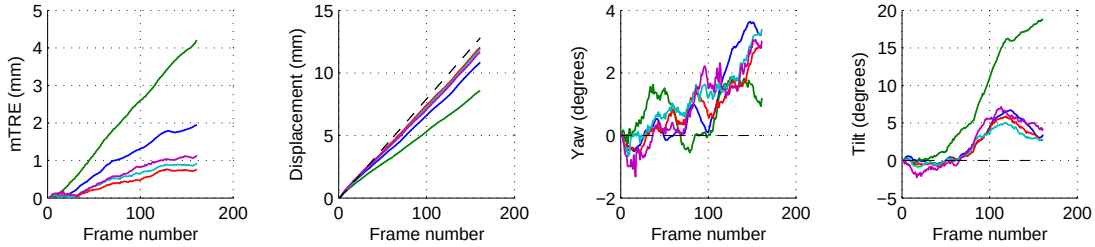
Figure D.4 Pork tenderloin trajectory recovery results, 0.06 mm intervals.



(a) Pork tenderloin 13: 161 frames, 0.08 mm intervals

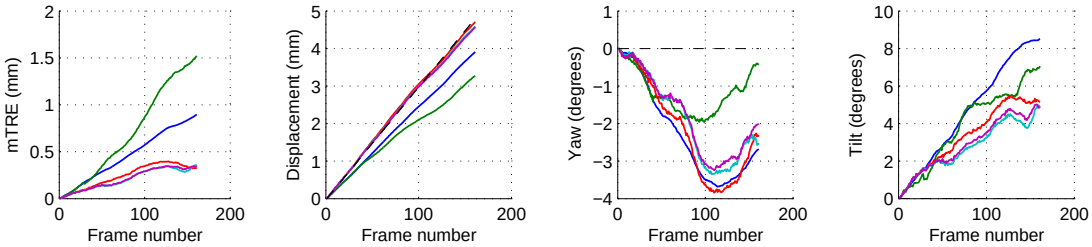


(b) Pork tenderloin 14: 161 frames, 0.08 mm intervals

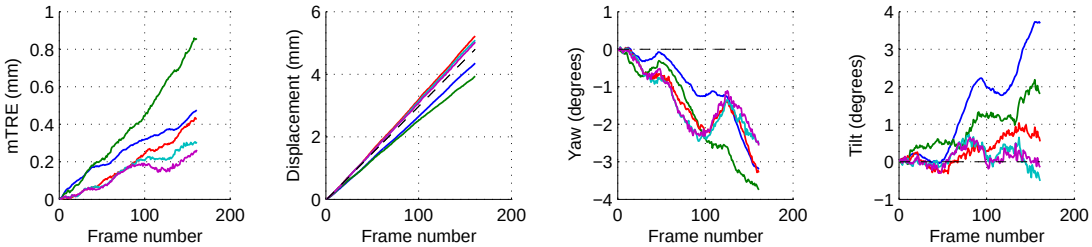


(c) Pork tenderloin 15: 161 frames, 0.08 mm intervals

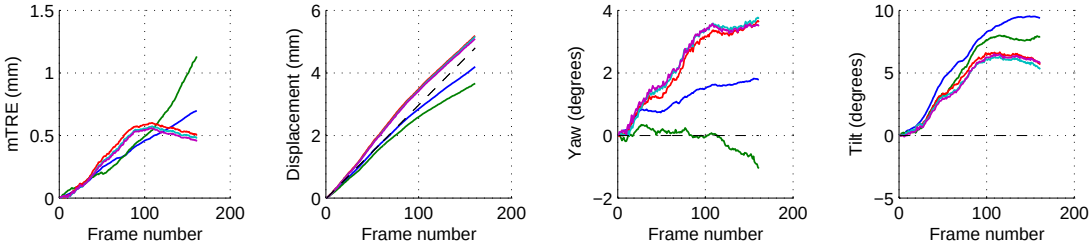
Figure D.5 Pork tenderloin trajectory recovery results, 0.08 mm intervals.



(a) Turkey breast 1: 161 frames, 0.03 mm intervals

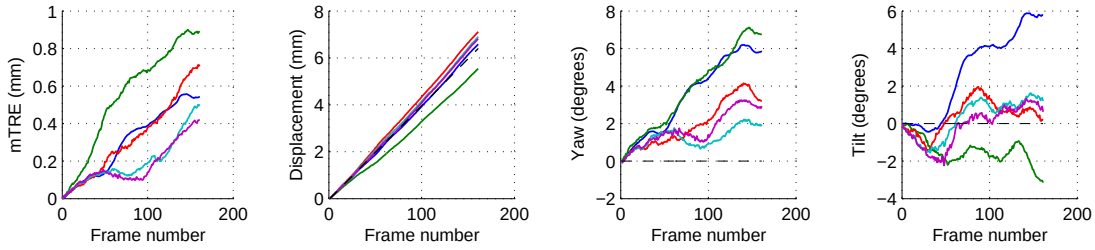


(b) Turkey breast 2: 161 frames, 0.03 mm intervals

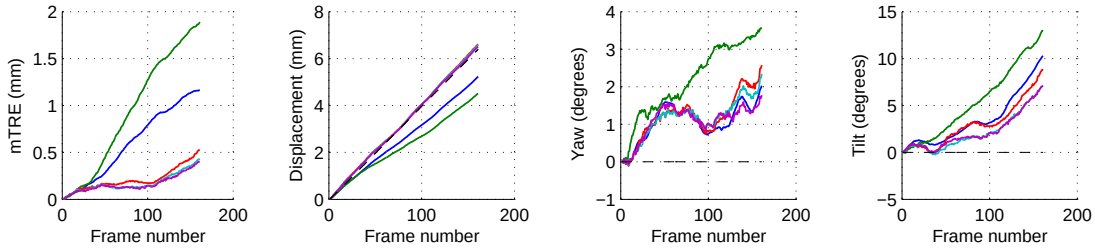


(c) Turkey breast 3: 161 frames, 0.03 mm intervals

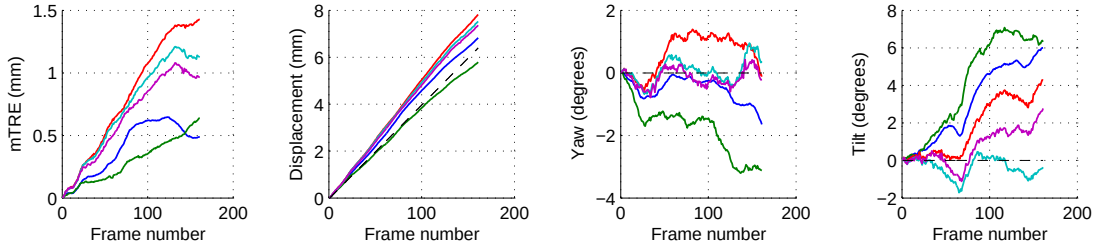
Figure D.6 Turkey breast trajectory recovery results, 0.03 mm intervals.



(a) Turkey breast 4: 161 frames, 0.04 mm intervals

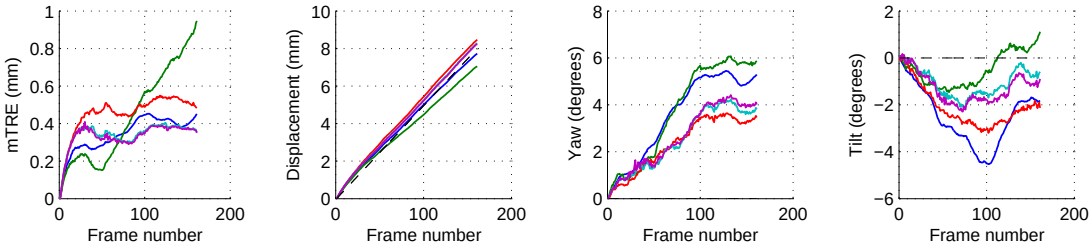


(b) Turkey breast 5: 161 frames, 0.04 mm intervals

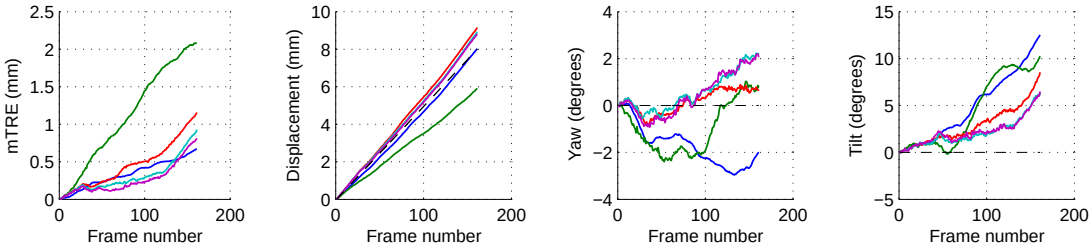


(c) Turkey breast 6: 161 frames, 0.04 mm intervals

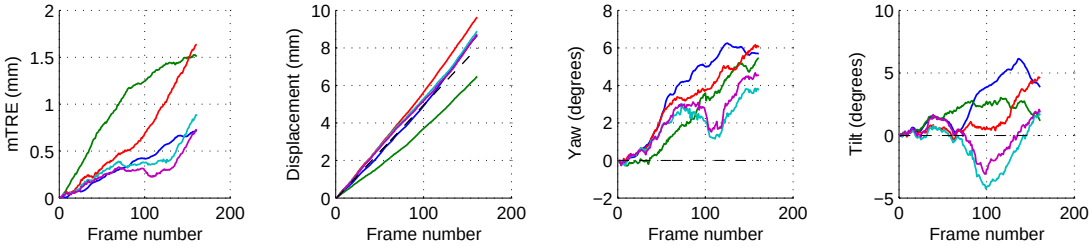
Figure D.7 Turkey breast trajectory recovery results, 0.04 mm intervals.



(a) Turkey breast 7: 161 frames, 0.05 mm intervals

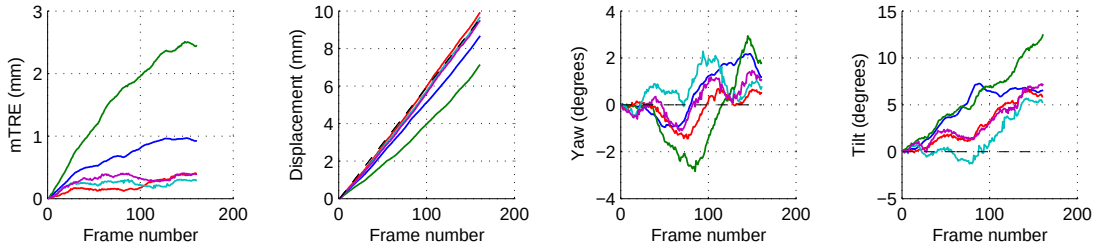


(b) Turkey breast 8: 161 frames, 0.05 mm intervals

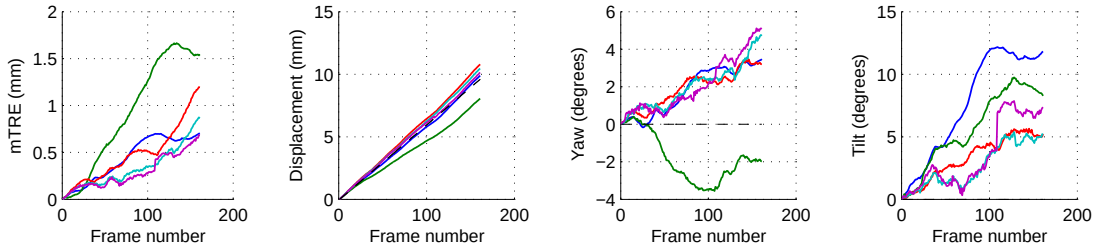


(c) Turkey breast 9: 161 frames, 0.05 mm intervals

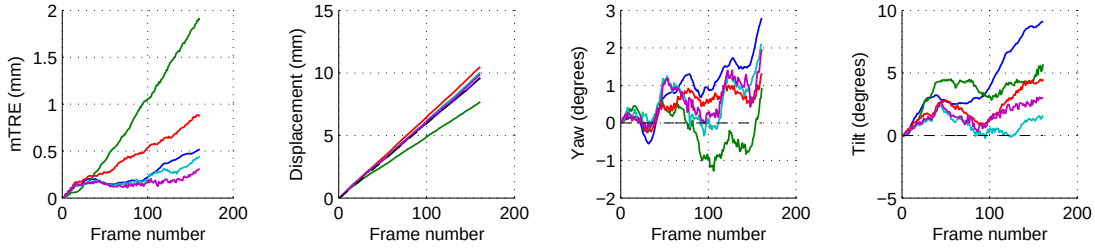
Figure D.8 Turkey breast trajectory recovery results, 0.05 mm intervals.



(a) Turkey breast 10: 161 frames, 0.06 mm intervals

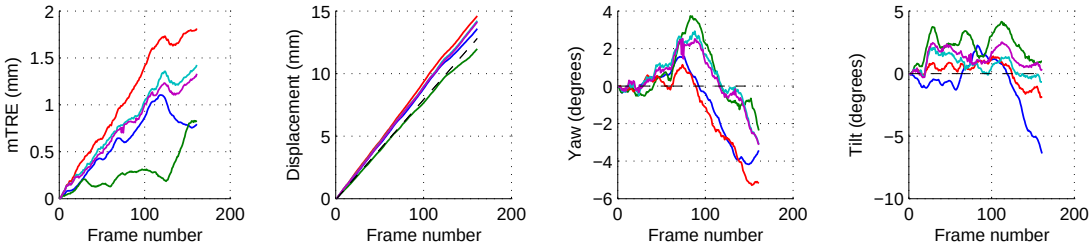


(b) Turkey breast 11: 161 frames, 0.06 mm intervals

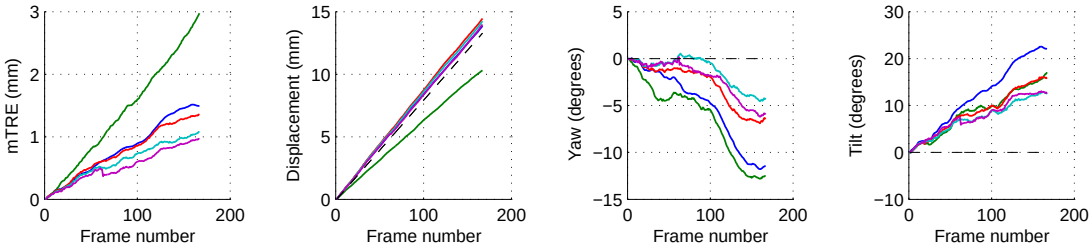


(c) Turkey breast 12: 161 frames, 0.06 mm intervals

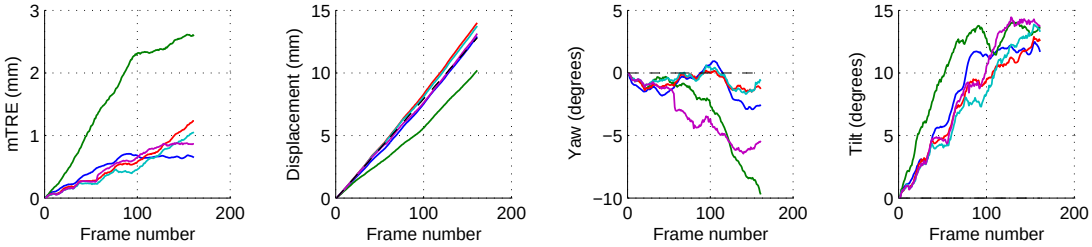
Figure D.9 Turkey breast trajectory recovery results, 0.06 mm intervals.



(a) Turkey breast 13: 161 frames, 0.08 mm intervals

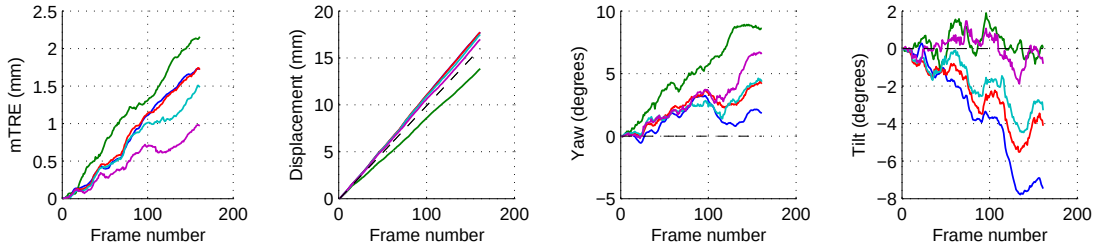


(b) Turkey breast 14: 167 frames, 0.08 mm intervals

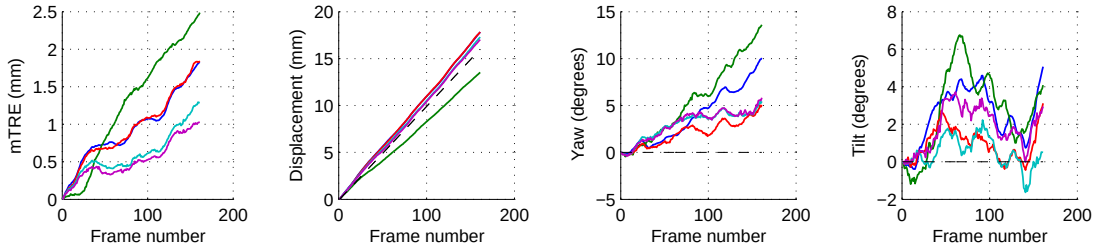


(c) Turkey breast 15: 161 frames, 0.08 mm intervals

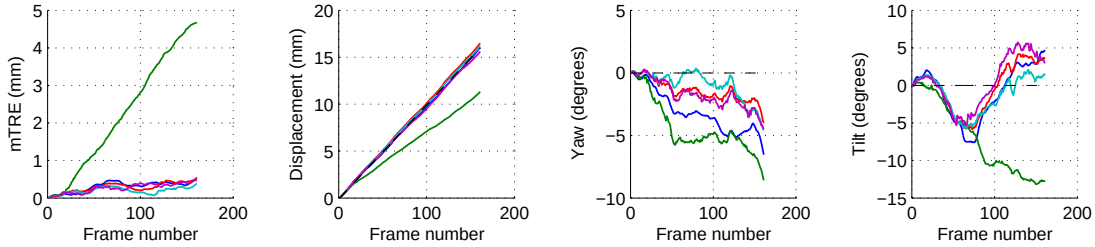
Figure D.10 Turkey breast trajectory recovery results, 0.08 mm intervals.



(a) Turkey breast 16: 161 frames, 0.1 mm intervals

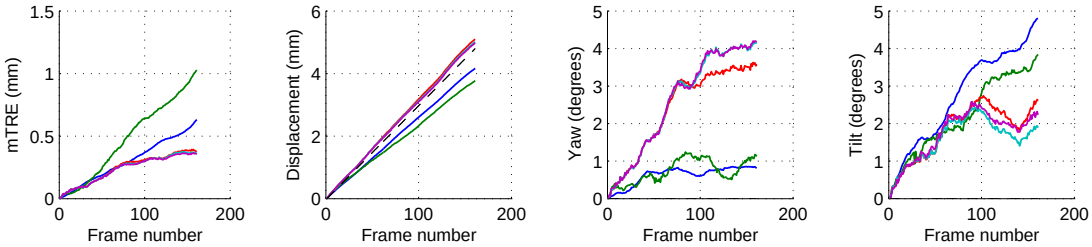


(b) Turkey breast 17: 161 frames, 0.1 mm intervals

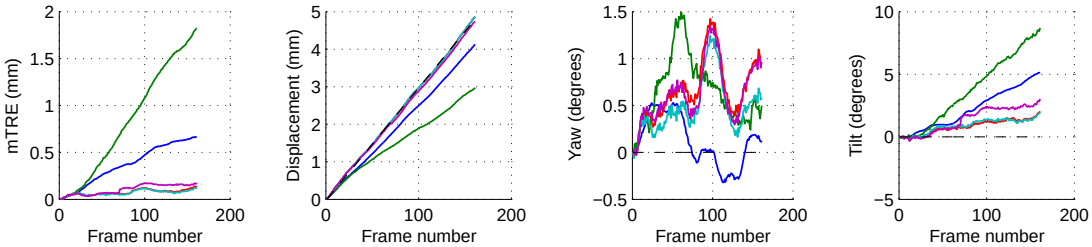


(c) Turkey breast 18: 161 frames, 0.1 mm intervals

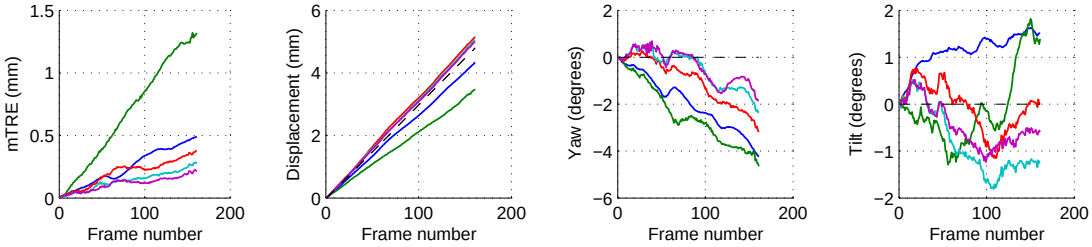
Figure D.11 Turkey breast trajectory recovery results, 0.1 mm intervals.



(a) Beef brisket 1: 161 frames, 0.03 mm intervals

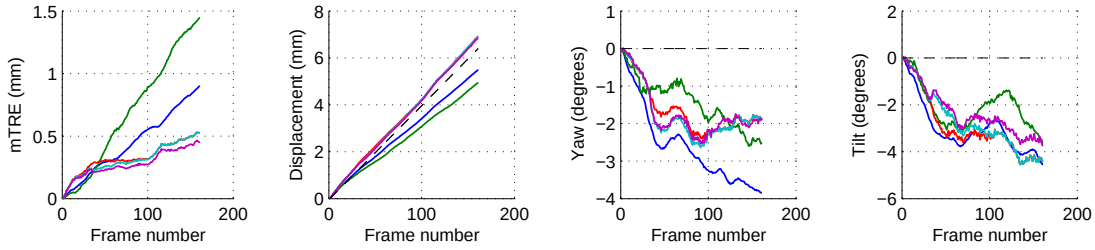


(b) Beef brisket 2: 161 frames, 0.03 mm intervals

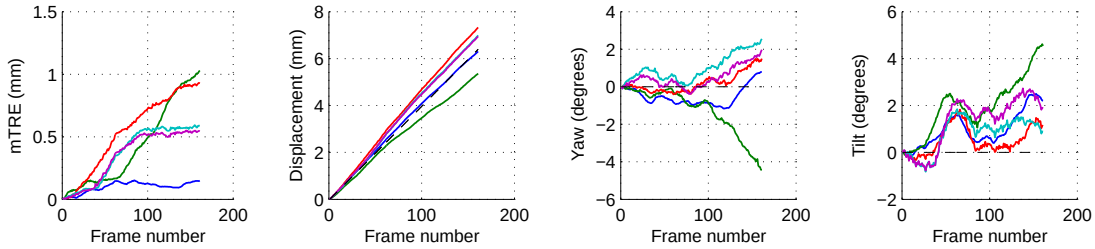


(c) Beef brisket 3: 161 frames, 0.03 mm intervals

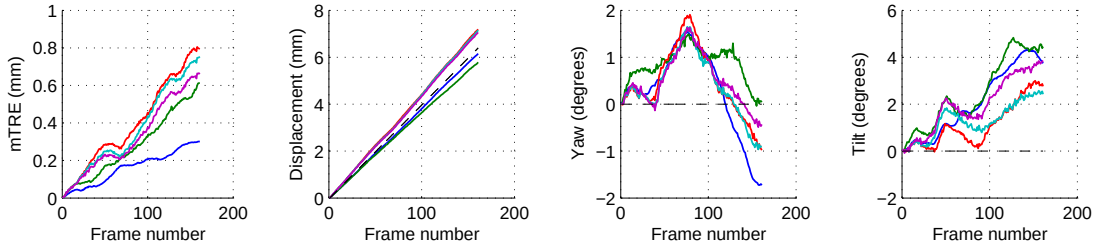
Figure D.12 Beef brisket trajectory recovery results, 0.03 mm intervals.



(a) Beef brisket 4: 161 frames, 0.04 mm intervals

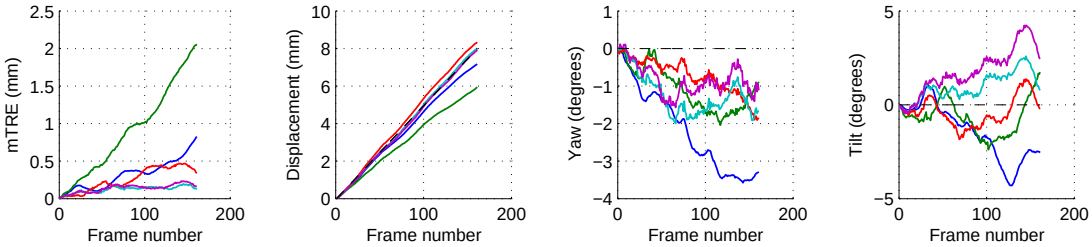


(b) Beef brisket 5: 161 frames, 0.04 mm intervals

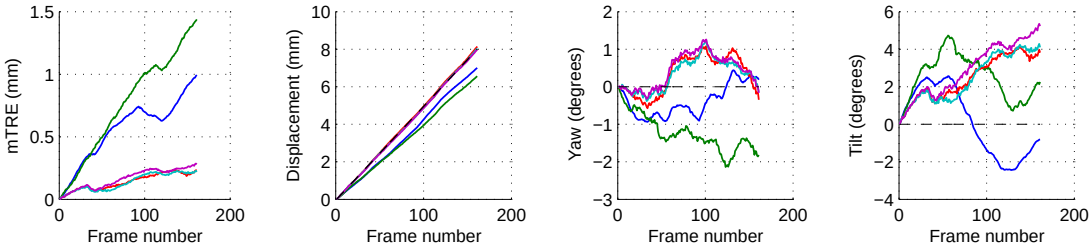


(c) Beef brisket 6: 161 frames, 0.04 mm intervals

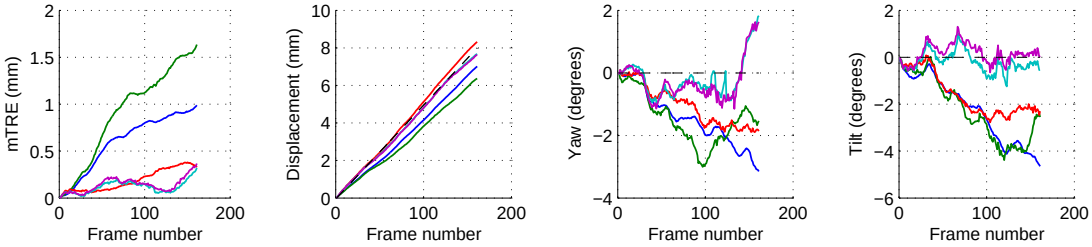
Figure D.13 Beef brisket trajectory recovery results, 0.04 mm intervals.



(a) Beef brisket 7: 161 frames, 0.05 mm intervals

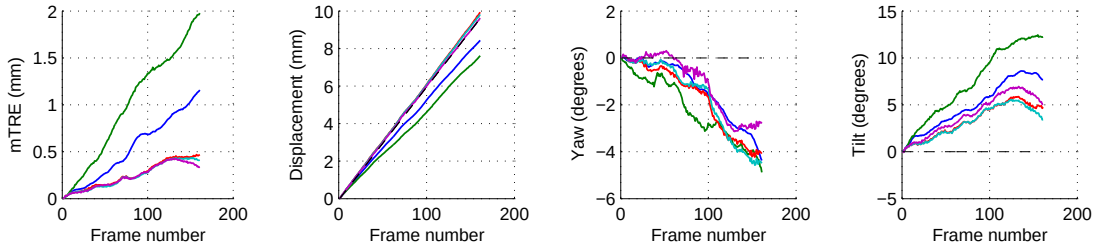


(b) Beef brisket 8: 161 frames, 0.05 mm intervals

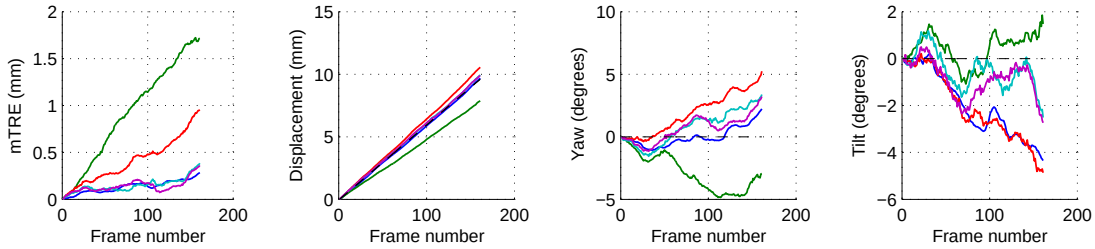


(c) Beef brisket 9: 161 frames, 0.05 mm intervals

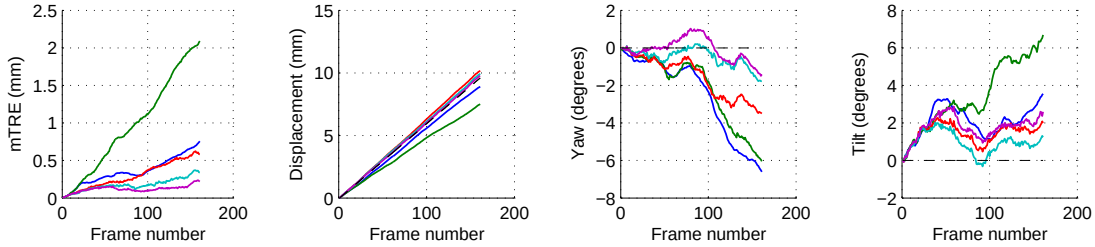
Figure D.14 Beef brisket trajectory recovery results, 0.05 mm intervals.



(a) Beef brisket 10: 161 frames, 0.06 mm intervals

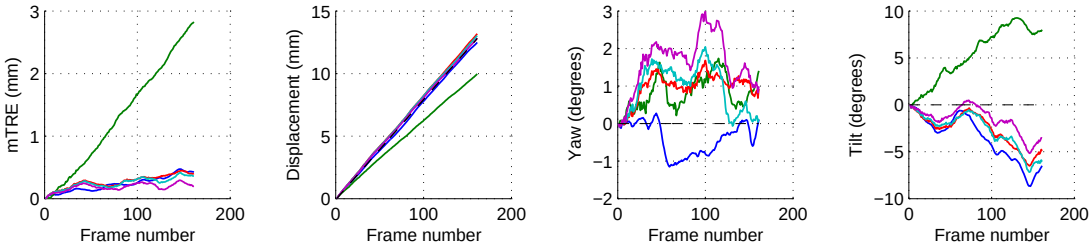


(b) Beef brisket 11: 161 frames, 0.06 mm intervals

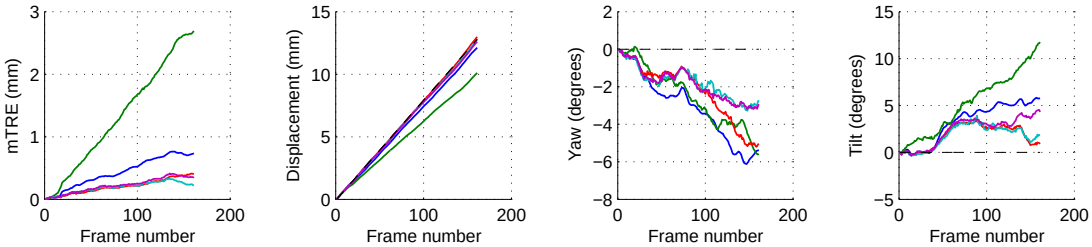


(c) Beef brisket 12: 161 frames, 0.06 mm intervals

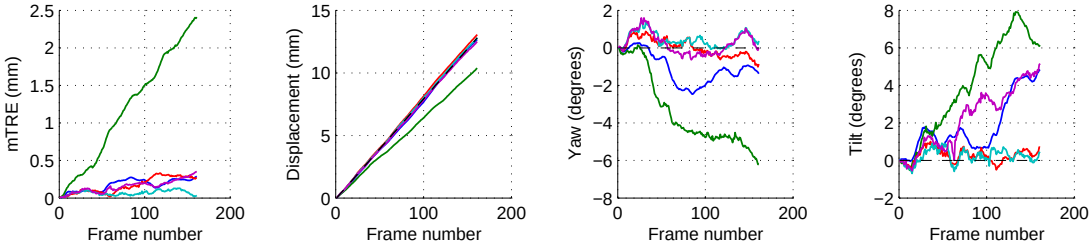
Figure D.15 Beef brisket trajectory recovery results, 0.06 mm intervals.



(a) Beef brisket 13: 161 frames, 0.08 mm intervals

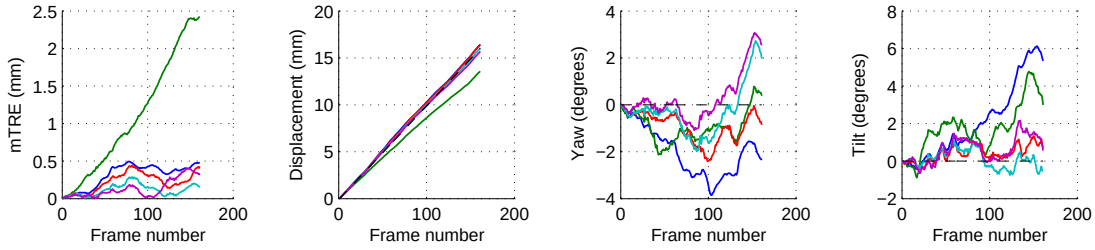


(b) Beef brisket 14: 161 frames, 0.08 mm intervals

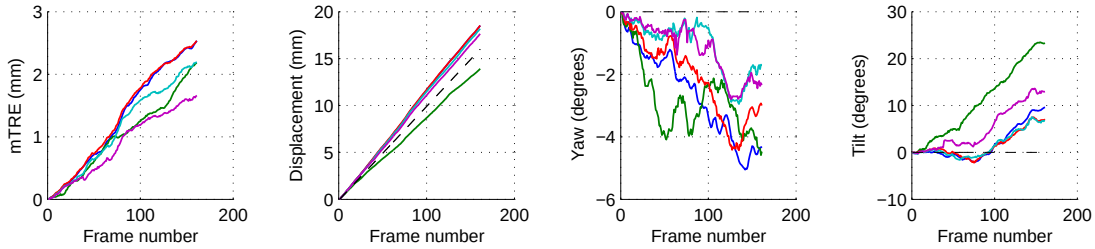


(c) Beef brisket 15: 161 frames, 0.08 mm intervals

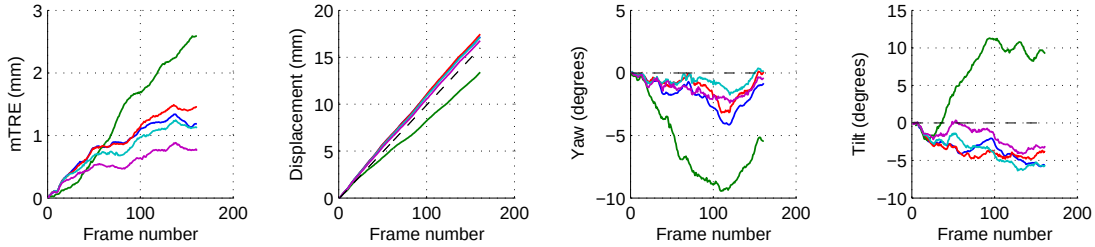
Figure D.16 Beef brisket trajectory recovery results, 0.08 mm intervals.



(a) Beef brisket 16: 161 frames, 0.1 mm intervals



(b) Beef brisket 17: 161 frames, 0.1 mm intervals



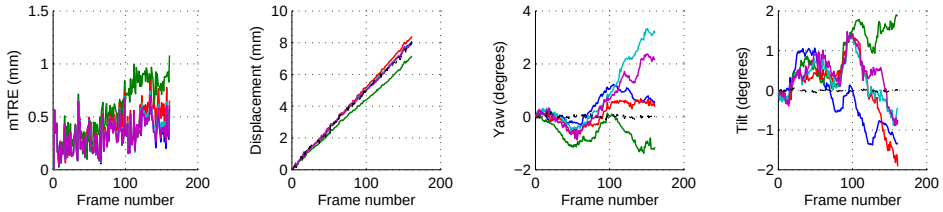
(c) Beef brisket 18: 161 frames, 0.1 mm intervals

Figure D.17 Beef brisket trajectory recovery results, 0.1 mm intervals.

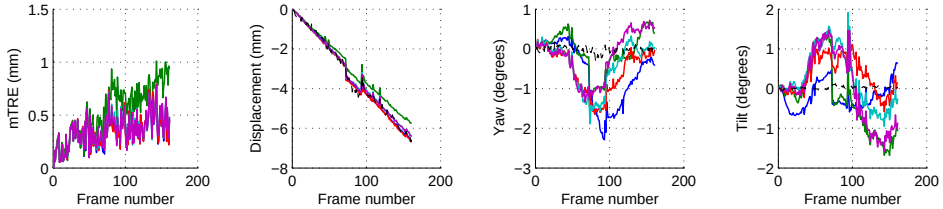
Appendix E

Full set of results for chaotic trajectory estimation in real tissue

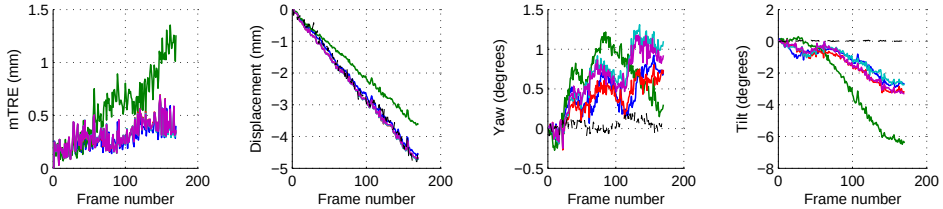
This appendix contains the full set of results obtained for the experiments relating to the recovery of chaotic transducer trajectories in animal tissue scans described in Chapter 7. Every row of each figure represents the results obtained for a single scan. From left to right, the plots in a given row show (1) the progression of the mTRE over frame acquisition steps, (2) the estimated displacement, (3) the estimated yaw (should be zero) and (4) the estimated tilt (should be zero) evaluated using the centroid of the grid of image patches as the centre of rotation. The dashed line represents the measurements recorded by the Polaris optical tracker and used as ground truth for the computation of the mTRE. The blue line corresponds to the NN method. The green line corresponds the NMS method (probabilistic fusion approach without measurement selection). The red, cyan and magenta lines represent the HYP-K1, HYP-K2 an HYP-K3 methods, i.e. the iterative measurement selection and fusion approach of Chapter 7 using the κ_1, κ_2 and κ_3 cost functions, respectively, to compute an initial estimate of the trajectory.



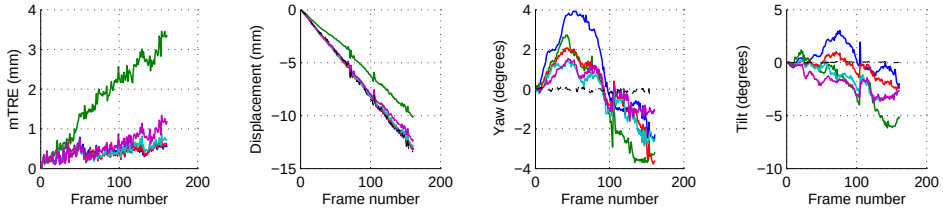
(a) Chaotic pork tenderloin 1



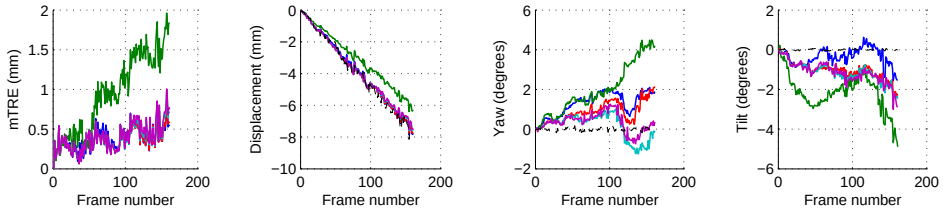
(b) Chaotic pork tenderloin 2



(c) Chaotic pork tenderloin 3

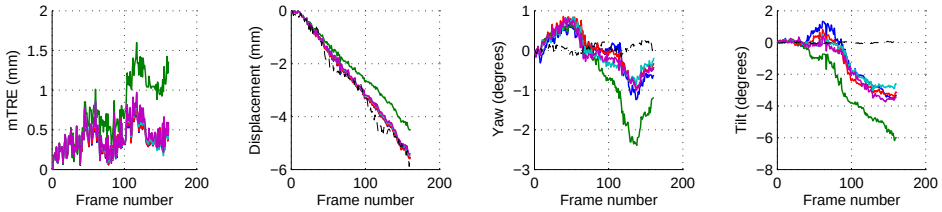


(d) Chaotic pork tenderloin4

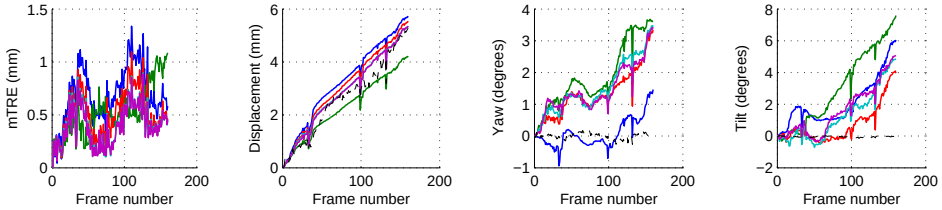


(e) Chaotic pork tenderloin 5

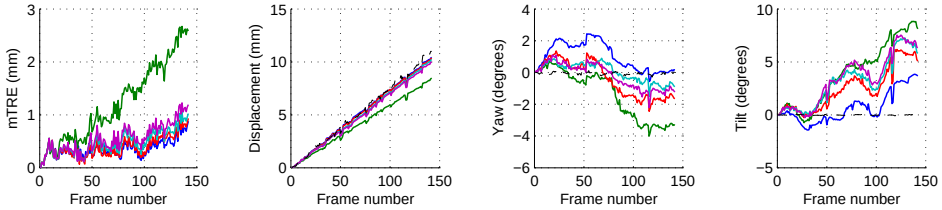
Figure E.1 Pork tenderloin chaotic trajectory recovery results, part 1.



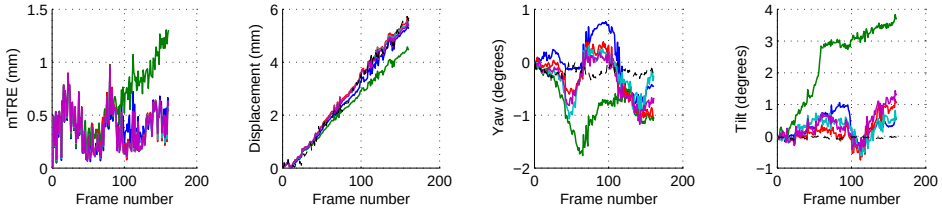
(a) Chaotic pork tenderloin 6



(b) Chaotic pork tenderloin 7

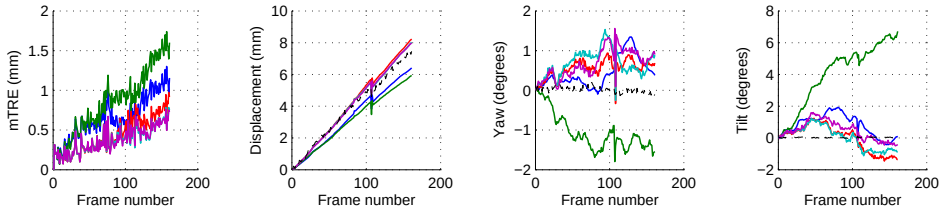


(c) Chaotic pork tenderloin 8

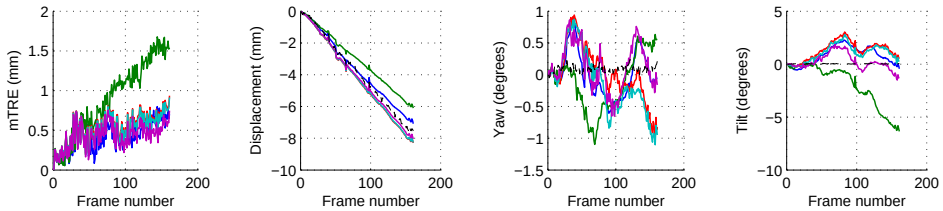


(d) Chaotic pork tenderloin 9

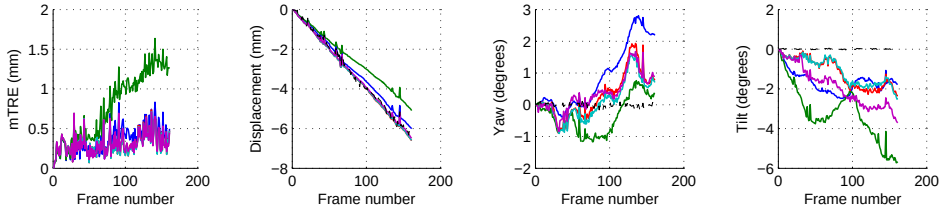
Figure E.2 Pork tenderloin chaotic trajectory recovery results, part 2.



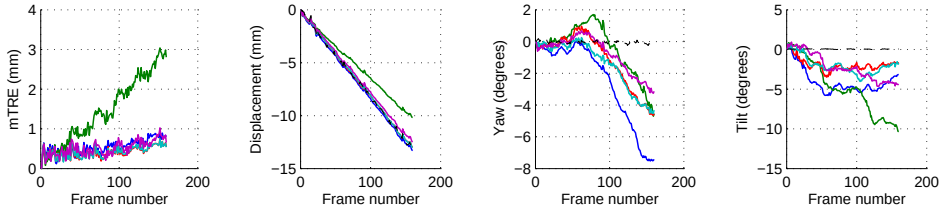
(a) Chaotic beef tenderloin 1



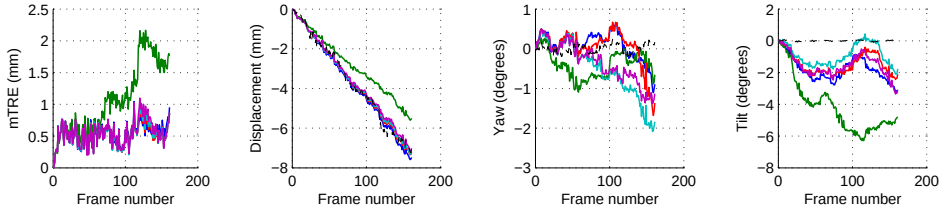
(b) Chaotic beef tenderloin 2



(c) Chaotic beef tenderloin 3

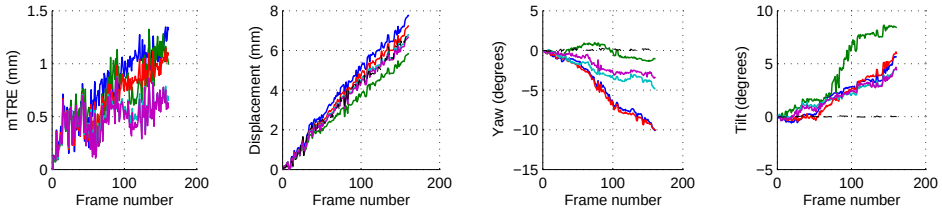


(d) Chaotic beef tenderloin 4

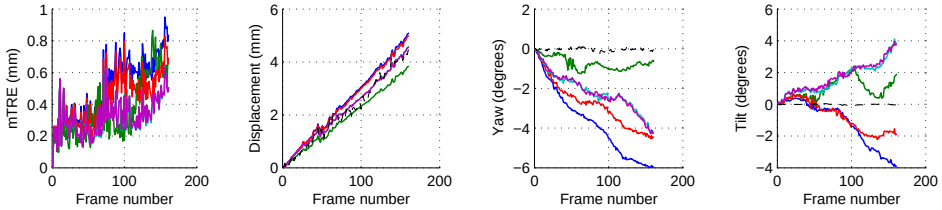


(e) Chaotic beef tenderloin 5

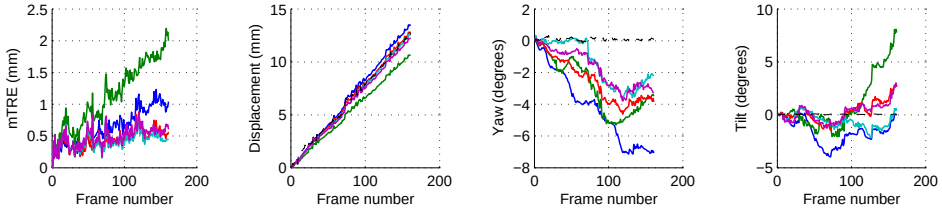
Figure E.3 Beef tenderloin chaotic trajectory recovery results, part 1.



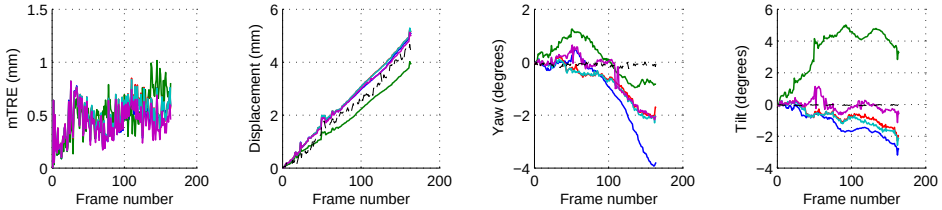
(a) Chaotic beef tenderloin 6



(b) Chaotic beef tenderloin 7



(c) Chaotic beef tenderloin 8



(d) Chaotic beef tenderloin 9

Figure E.4 Beef tenderloin chaotic trajectory recovery results, part 2.

References

[1] Hilbert curve. http://commons.wikimedia.org/wiki/Hilbert_curve, 2008.

[2] K. Z. Abd-Elmoniem, A.-B. M. Youssef, and Y. M. Kadah. Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion. *IEEE Transactions on Biomedical Engineering*, 49(9):997–1014, 2002.

[3] A. Agarwal and B. Triggs. Recovering 3D human pose from monocular images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):44–58, 2006.

[4] M. Agrawal. A Lie algebraic approach for consistent pose registration for general Euclidean motion. In *Proceedings of te IEEE/RSJ Interational Conference on Intelligent Robots and Systems*, pages 1891–1897, 2006.

[5] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[6] A. Ali and R. Logeswaran. A visual probe localization and calibration system for cost-effective computer-aided 3D ultrasound. *Computers in Biology and Medicine*, 37:1141–1147, 2007.

[7] M. E. Anderson and G. E. Trahey. A seminar on k-space applied to medical ultrasound. Technical report, Department of Biomedical Engineering, Duke University, 2006.

[8] D. L. Applegate, R. E. Bixby, V. Chvátal, and W. J. Cook. *The Traveling Salesman Problem: A Computational Study.* Princeton University Press, 2006. See also the Concorde software available from http://www.tsp.gatech.edu/concorde/.

[9] A. J. Baddeley and M. N. M. van Lieshout. Area-interaction point processes. *Annals of the Institue of Statistical Mathematics*, 47(4):601–619, 1995.

[10] R. Baker. Probability estimation and information principles. *Structural Safety*, 9:97–116, 1990.

[11] R. Baker and G. Christakos. Revisiting prior distributions, part II: implications of the physical prior in maximum entropy analysis. *Stochastic Environmental Research and Risk Assessment*, 21(4):427–434, 2007.

[12] J. C. Bamber and C. Daft. Adaptive filtering for reduction of speckle in ultrasonic pulse-echo images. *Ultrasonics*, 24(1):41–44, 1986.

[13] J. C. Bamber and R. J. Dickinson. Ultrasonic B-scanning: a computer simulation. *Physics in Medicine and Biology*, 25(3):463–479, 1980.

[14] C. D. Barry, C. P. Allott, N. W. John, P. M. Mellor, P. A. Arundel, D. S. Thomson, and J. C. Waterton. Three-dimensional freehand ultrasound: image reconstruction and volume analysis. *Ultrasound in Medicine and Biology*, 23(8):1209–1224, 1997.

[15] J. J. Bartholdi III and P. Goldsman. Vertex-labeling algorithms for the Hilbert spacefilling curve. *Software: Practice and Experience*, 31(5):395–408, 2000.

[16] O. Basset, Z. Sun, J. L. Mestas, and G. Gimenez. Texture analysis of ultrasonic images of the prostate by means of co-occurrence matrices. *Ultrasonic Imaging*, 15:218–237, 1993.

[17] M. Baumann, P. Mozer, V. Daanen, and J. Troccaz. Towards 3D ultrasound image based soft tissue tracking: a transrectal ultrasound prostate image alignment system. In *Proceedings of the 10th International Conference on Medical Image Computing and Computer Assisted Intervention, Part II*, pages 26–33, 2007.

[18] J. Bax, D. Cool, L. Gardi, J. Montreuil, E. Gil, J. Bluvol, K. Knight, D. Smith, C. Romagnoli, and A. Fenster. 3D transrectal ultrasound prostate biopsy using a mechanical imaging and needle-guidance system. In *Proceedings of the SPIE Medical Imaging Conference*, 2008.

[19] O. Bernard, B. Touil, J. D'hooge, and D. Friboulet. Statistical modeling of the radio-frequency signal for partially- and fully-developed speckle based on a generalized Gaussian model with application to echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 54(10):2189–2194, 2007.

[20] A. K. Bhandari and N. C. Nanda. Myocardial texture characterization by two-dimensional echocardiography. *American Journal of Cardiology*, 51(5):817–825, 1984.

[21] W. Birkfellner, J. Hummel, E. Wilson, and K. Cleary. Tracking devices. In T. Peters and K. Cleary, editors, *Image-Guided Interventions*, pages 23–44. Springer US, 2008.

[22] W. Birkfellner, F. Watzinger, F. Wanschitz, G. Enislidis, C. Kollmann, D. Rafolt, R. Nowotny, R. Ewers, and H. Bergmann. Systematic distortions in magnetic position digitizers. *Medical Physics*, 25(11):2242–2248, 1998.

[23] J. M. Blackall, G. P. Penney, A. P. King, and D. J. Hawkes. 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*, 24(11):1405–1416, 2005.

[24] L. N. Bohs and G. E. Trahey. A novel method for angle independent ultrasonic imaging of blood flow and tissue motion. *IEEE Transactions on Biomedical Engineering*, 38(3):280–286, 1991.

[25] F. Bonilla-Musoles, F. Raga, and N. G. Osborne. Three-dimensional ultrasound evaluation of ovarian masses. *Gynecologic Oncology*, 59(1):129–35, 1995.

[26] G. Brassard and P. Bratley. *Fundamentals of Algorithmics*. Prentice Hall, 1999.

[27] C. B. Burckhardt. Speckle in ultrasound B-mode scans. *IEEE Transactions on Sonics and Ultrasonics*, 25(1):1–6, 1978.

[28] J. A. Castellanos, J. M. M. Montiel, J. Neira, and J. D. Tardós. The SPmap: a probabilistic framework for simultaneous localization and map building. *IEEE Transactions on Robotics and Automation*, 15(5):948–953, 1999.

[29] R.-F. Chang, W.-J. Wu, D.-R. Chen, W.-M. Chen, W. Shu, J.-H. Lee, and L.-B. Jeng. 3-D US frame positioning using speckle decorrelation and image registration. *Ultrasound in Medicine and Biology*, 29(6):801–812, 2003.

[30] R.-F. Chang, W.-J. Wu, W. K. Moon, and D.-R. Chen. Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis. *Ultrasound in Medicine and Biology*, 9(5):679–686, 2003.

[31] J.-F. Chen, J. B. Fowlkes, P. L. Carson, and J. M. Rubin. Determination of scan-plane motion using speckle decorrelation: theoretical considerations and initial test. *International Journal of Imaging Systems and Technology*, 8(1):38–44, 1997.

[32] J.-F. Chen and L. Weng. System and method for 3-D ultrasound imaging and motion estimation. United States Patent no. 5,876,342, 1999.

[33] J.-F. Chen, J. A. Zagzebski, F. Dong, and E. L. Madsen. Estimating the spatial autocorrelation function for ultrasound scatterers in isotropic media. *Medical Physics*, 25(5):648–655, 1998.

[34] R. C. Chivers. Tissue characterization. *Ultrasound in Medicine and Biology*, 7(1):1–20, 1981.

[35] G. Cloutier, G. Soulez, S. D. Qanadli, P. Teppaz, L. Allard, Z. Qin, F. Cloutier, and L.-G. Durand. A multimodality vascular imaging phantom with fiducial markers visible in DSA, CTA, MRA and ultrasound. *Medical Physics*, 31(6):1424–1433, 2004.

[36] R. S. C. Cobbold. *Foundations of Biomedical Ultrasound*. Oxford University Press, 2007.

[37] X. Compagnion. Embryo at 12 weeks. http://commons.wikimedia.org/wiki/File:Embryo_at_12_weeks.JP 2009.

[38] W. J. Conover. *Practical Nonparametric Statistics*. John Wiley & Sons, 2nd edition, 1980.

[39] D. Cool, D. Downey, J. Izawa, J. Chin, and A. Fenster. 3D prostate model formation from non-parallel 2D ultrasound biopsy images. *Medical Image Analysis*, 10:875–887, 2006.

[40] D. Cool, S. Sherebrin, J. Izawa, J. Chin, and A. Fenster. Design and evaluation of a 3D transrectal ultrasound prostate biopsy system. *Medical Physics*, 35(10):4695–4707, 2008.

[41] D. Cool, S. Sherebrin, J. Izawa, T. Peters, and A. Fenster. Minimization of tool tracking error using fulcrum correction in minimally invasive interventions: application to prostate biopsy procedure. In *Proceedings of the SPIE Medical Imaging Conference*, 2006.

[42] P. Coupé, P. Hellier, X. Morandi, and C. Barillot. Probe trajectory interpolation for 3D reconstruction of freehand ultrasound. *Medical Image Analysis*, 11:604–615, 2007.

[43] R. M. Cramblitt and K. J. Parker. Generation of non-Rayleigh speckle distributions using marked regularity models. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 46(4):867–874, 1999.

[44] D. C. Crawford, D. O. Cosgrove, E. Tohno, W. E. Svensson, B. al Murrani, D. S. Bell, K. Stepniewska, and J. C. Bamber. Visual impact of adaptive speckle reduction on US B-mode images. *Radiology*, 183:555–561, 1992.

[45] G. Dantzig, R. Fulkerson, and S. Johnson. Solution of a large-scale traveling-salesman problem. *Operations Research*, 2:393–410, 1954.

[46] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood via incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[47] O. Deussen, P. Hanrahan, B. Lintermann, R. Měch, M. Pharr, and P. Prusinkiewicz. Realistic modeling and rendering of plant ecosystems. In *Proceedings of ACM SIGGRAPH*, pages 275–286, 1998.

[48] J.-L. Dillenseger, S. Laguitton, and E. Delabrousse. Fast simulation of ultrasound images from a CT volume. *Computers in Biology and Medicine*, 39:180–186, 2009.

[49] M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, 17(3):229–241, 2001.

[50] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, 2001.

[51] R. P. W. Duin. On the choice of smoothing parameters for Parzen estimators of probability density functions. *IEEE Transactions on Computers*, 25(11):1175–1179, 1976.

[52] V. Dutt and J. F. Greenleaf. Ultrasound echo envelope analysis using a homodyned K distribution signal model. *Ultrasonic Imaging*, 16(4):265–287, 1994.

[53] V. Dutt and J. F. Greenleaf. Adaptive speckle reduction filter for log-compressed B-scan images. *IEEE Transactions on Medical Imaging*, 15:802–813, 1996.

[54] V. Dutt and J. F. Greenleaf. Statistics of the log-compressed echo envelope. *Journal of the Acoustical Society of America*, 99(6):3817–3825, 1996.

[55] A. I. Eliazar and R. Parr. Learning probabilistic motion models for mobile robots. In *Proceedings of the International Conference on Machine Learning*, 2004.

[56] T. L. Elliot, D. B. Downey, S. Tong, C. A. McLean, and A. Fenster. Accuracy of prostate volume measurements in vitro using three-dimensional ultrasound. *Academic Radiology*, 3(5):401–406, 1996.

[57] T. Eltoft. Modeling the amplitude statistics of ultrasonic images. *IEEE Transactions on Medical Imaging*, 25(2):229–240, 2006.

[58] E. Esh-Broder, F. Ushakov, T. Imbar, and S. Yagel. Application of free-hand three-dimensional echocardiography in the evaluation of fetal cardiac ejection fraction: a preliminary study. *Ultrasound in Obstetrics and Gynecology*, 23(6):546–551, 2004.

[59] L. Fan and T. Pylvänäinen. Robust scale estimation from ensemble inlier sets for random sample consensus methods. In *Proceedings of the European Conference on Computer Vision, Part III*, pages 182–195, 2008.

[60] F. A. Farrelly, A. Petri, L. Pitolli, and G. Pontuale. Statistical propoerties of acoustic emission signals from metal cutting processes. *Journal of the Acoustical Society of America*, 116(2):981–986, 2004.

[61] H. J. S. Feder, J. J. Leonard, and C. M. Smith. Adaptive mobile robot navigation and mapping. *International Journal of Robotics Research*, 18(7):650–668, 1999.

[62] A. Fenster, D. B. Downey, and H. N. Cardinal. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology*, 46:R67–R99, 2001.

[63] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications of image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[64] R. A. Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–512, 1915.

[65] J. M. Fitzpatrick, D. L. G. Hill, and J. Calvin R. Maurer. Image registration. In M. Sonka and J. M. Fitzpatrick, editors, *Medical Image Processing, Volume II of the Handbook of Medical Imaging*, chapter 8. SPIE Press, 2000.

[66] R. Fletcher. *Practical Methods of Optimization*. Wiley, 2nd edition, 1987.

[67] B. H. Friemel, L. Weng, and T.-J. Teo. 3-dimensional compound ultrasound field of view. United States Patent no. 5,655,535, 1997.

[68] B. H. Friemel, L. Weng, T.-J. Teo, and K. G. Brown. 3-dimensional volume by aggregating ultrasound fields of view. United States Patent no. 5,899,861, 1999.

[69] W. Fung and R. Rohling. Design of an open-architecture ultrasound acquisition system for real-time processing and control. Technical Report UBC-ECE-TR-002, University of British Columbia, 2003.

[70] M. Galassi et al. GNU scientific library reference manual. http://www.gnu.org/software/gsl/, 2007.

[71] C. D. Garson, Y. Li, and J. A. Hossack. Free-hand ultrasound scanning approaches for volume quantification of the mouse heart left ventricle. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 54(5):966–977, 2007.

[72] A. H. Gee, R. J. Housden, P. Hassenpflug, G. M. Treece, and R. W. Prager. Sensorless freehand 3D ultrasound in real tissue: speckle decorrelation without fully developped speckle. *Medical Image Analysis*, 10(2):137–149, 2006.

[73] B. J. Geiman, L. N. Bohs, M. E. Anderson, S. M. Breit, and G. E. Trahey. A novel interpolation strategy for estimating subsample speckle motion. *Physics in Medicine and Biology*, 45:1541–1552, 2000.

[74] B. P. Geiser, W. T. Hatfield, V. V. Kamat, S. C. Miller, L. Y. L. Mo, T. M. Tillman, and B. Yamrom. Method and apparatus for calculating distance between ultrasound images using sum of absolute differences. United States Patent no. 6,106,470, 2000.

[75] G. Georgiou and F. S. Cohen. Statistical characterization of diffuse scattering in ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 45(1):57–64, 1998.

[76] D. G. Gobbi and T. M. Peters. Interactive intra-operative 3D ultrasound reconstruction and visualization. In *Proceedings of the 5th International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 156–163, 2002.

[77] M. J. Gooding, S. Kennedy, and J. A. Noble. Volume segmentation and reconstruction from freehand three-dimensional ultrasound data with application to ovarian follicle measurement. *Ultrasound in Medicine and Biology*, 34(2):183–195, 2008.

[78] M. J. Gooding, S. H. Kennedy, and J. A. Noble. Temporal calibration of freehand three-dimensional ultrasound using image alignment. *Ultrasound in Medicine and Biology*, 31(7):919–927, 2005.

[79] J. W. Goodman. Statistical properties of laser speckle patterns. In J. C. Dainty, editor, *Laser Speckle and Related Phenomena*, pages 9–75. Springer-Verlag, 1975.

[80] A. S. Gopal, A. M. Keller, Z. Shen, P. M. Sapin, K. M. Schroeder, D. L. K. junior., and D. L. King. Three-dimensional echocardiography: in vitro and in vivo validation of left ventricular mass and comparison to conventional echocardiographic methods. *Journal of the American College of Cardiology*, 24:504–513, 1994.

[81] C. Gotsman and M. Lindenbaum. On the metric properties of discrete space-filling curves. *IEEE Transactions on Image Processing*, 5:794–797, 1996.

[82] S. Gunn. Support vector machines for classification and regression. Technical report, University of Southampton, 1998.

[83] X. Hao, S. Gao, and X. Gao. A novel multiscale nonlinear thresholding method for ultrasonic speckle suppressing. *IEEE Transactions on Medical Imaging*, 18(9):787–794, 1999.

[84] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6):610–621, 1973.

[85] P. Hassenpflug, R. Prager, G. Treece, and A. Gee. Distance measurement for sensorless 3D US. In *Proceedings of the 7th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1087–1088, 2004.

[86] P. Hassenpflug, R. W. Prager, G. M. Treece, and A. H. Gee. Speckle classification for sensorless freehand 3-D ultrasound. *Ultrasound in Medicine and Biology*, 31(11):1499–1508, 2005.

[87] K. Helsgaun. An effective implementation of the Lin-Kernighan traveling salesman heuristic. *European Journal of Operational Research*, 126:106–130, 2000.

[88] A. Hernandez, O. Basset, P. Chirossel, and G. Gimenez. Spatial compounding in ultrasonic imaging using an articulated scan arm. *Ultrasound in Medicine and Biology*, 22(2):229–238, 1996.

[89] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46:R1–R45, 2001.

[90] J. A. Hossack. Influence of elevational motion on the degradation of 2D image frame matching. In *Proceedings of the IEEE Ultrasonics Symposium*, pages 1713–1716, 2000.

[91] J. A. Hossack, T. S. Sumanaweera, S. Napel, and J. S. Ha. Quantitative 3-D diagnostic ultrasound imaging using a modified transducer array and an automated image tracking technique. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 49(8):1029–1038, 2002.

[92] R. J. Housden, A. H. Gee, R. W. Prager, and G. M. Treece. Rotational motion in sensorless freehand three-dimensional ultrasound. *Ultrasonics*, 48(5):412–422, 2008.

[93] R. J. Housden, A. H. Gee, G. M. Treece, and R. W. Prager. Sub-sample interpolation strategies for sensorless freehand 3D ultrasound. *Ultrasound in Medicine and Biology*, 32(12):1897–1904, 2006.

[94] R. J. Housden, A. H. Gee, G. M. Treece, and R. W. Prager. Sensorless reconstruction of unconstrained freehand 3D ultrasound data. *Ultrasound in Medicine and Biology*, 33(3):408–419, 2007.

[95] R. J. Housden, G. M. Treece, A. H. Gee, and R. W. Prager. Hybrid systems for reconstruction of freehand 3D ultrasound data. Technical Report CUED/F-INFENG/TR 574, University of Cambridge, 2007.

[96] P.-W. Hsu, R. W. Prager, A. H. Gee, and G. M. Treece. Freehand 3D ultrasound calibration: A review. In C. W. Sensen and B. Hallgrímsson, editors, *Advanced Imaging in Biology and Medicine*, pages 47–84. Springer Berlin Heidelberg, 2009.

[97] P.-W. Hsu, G. M. Treece, R. W. Prager, N. E. Houghton, and A. H. Gee. Comparison of freehand 3-D ultrasound calibration techniques using a stylus. *Ultrasound in Medicine and Biology*, 34(10):1610–1621, 2008.

[98] Q. Huang, M. Lu, Y. Zheng, and Z. Chi. Speckle suppression and contrast enhancement in reconstruction of freehand 3D ultrasound images using an adaptive distance-weighted method. *Applied Acoustics*, 70(1):21–30, 2009.

[99] Q.-H. Huang and Y.-P. Zheng. An adaptive squared-distance-weighted interpolation for volume reconstruction in 3D freehand ultrasound. *Ultrasonics*, 44:73–77, 2006.

[100] A. T. Ihler, J. W. F. III, R. L. Moses, and A. S. Willsky. Nonparametric belief propagation for self-calibration of sensor networks. *IEEE Journal on Selected Areas in Communications*, 23(4):809–819, 2005.

[101] M. F. Insana and D. G. Brown. Acoustic scattering theory applied to soft biological tissues. In K. K. Shung and G. A. Thieme, editors, *Ultrasonic Scattering in Biological Tissues*, pages 76–124. CRC Press, 1993.

[102] M. F. Insana, B. S. Garra, D. G. Brown, and T. H. Shawker. Analysis of ultrasound image texture via generalized Rician statistics. *Optical Engineering*, 25(6):743–748, 1986.

[103] E. Jakeman. On the statistics of K-distributed noise. *Journal of Physics, A: Math. Gen.*, 13:31–48, 1980.

[104] E. Jakeman and R. J. A. Tough. Generalized K distribution: a statistical model for weak scattering. *Journal of the Optical Society of America A*, 4(9):1764–1772, 1987.

[105] M.-A. Janvier, L.-G. Durand, M.-H. Roy Cardinal, I. Renaud, B. Chayer, P. Bigras, J. de Guise, G. Soulez, and G. Cloutier. Performance evaluation of a medical robotic 3D-ultrasound imaging system. *Medical Image Analysis*, 12:275–290, 2008.

[106] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106:620–630, 1957.

[107] J. A. Jensen. A model for the propagation and scattering of ultrasound in tissue. *Journal of the Acoustical Society of America*, 89(1):182–190, 1991.

[108] J. A. Jensen. Field: a program for simulating ultrasound systems. In *Proceedings of the Nordic-Baltic Conference on Biomedical Imaging*, volume 4, pages 351–353, 1996.

[109] J. A. Jensen and P. Munk. Computer phantoms for simulating ultrasound b-mode and cfm images. In *Proceedings of the 23rd Acoustical Imaging Symposium*, pages 75–80, 1997.

[110] J. A. Jensen and N. B. Svendsen. Calculation of pressure fields from arbitrarily shaped, apodized and excited ultrasound transducers. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 39(2):262–267, 1992.

[111] F. Kallel, M. Bertrand, and J. Meunier. Speckle motion artifact under tissue motion. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 41(1):105–122, 1994.

[112] Karmeshu and R. Agrawal. Study of ultrasonic echo envelope based on Nakagami-inverse Gaussian distribution. *Ultrasound in Medicine and Biology*, 32(3):371–376, 2006.

[113] D. A. Kenny, D. A. Kashy, and W. L. Cook. *Dyadic Data Analysis*. Guilford Press, 2006.

[114] R. Khadem, C. C. Yeh, M. Sadeghi-Tehrani, M. R. Bax, J. A. Johnson, J. N. Welch, E. P. Wilkinson, and R. Shahidi. Comparative tracking error analysis of five different optical tracking systems. *Computer Aided Surgery*, 5(2):98–107, 2000.

[115] J. I. Koo and S. B. Park. Speckle reduction with edge preservation in medical ultrasonic images using a homogeneous region growing mean filter (HRGMF). *Ultrasonic Imaging*, 13(3):211–237, 1991.

[116] G. Kossoff, W. J. Garret, D. A. Carpenter, J. Jellins, and M. J. Dadd. Principles and classification of soft tissues by gray scale echography. *Ultrasound in Medicine and Biology*, 2:115–129, 1975.

[117] F. W. Kremkau. *Diagnostic Ultrasound: Principles and Instruments*. Saunders, 6th edition, 2002.

[118] J. F. Krücker, C. R. Meyer, G. L. LeCarpentier, J. B. Fowlkes, and P. L. Carson. 3D spatial compounding of ultrasound images using image-based nonrigid registration. *Ultrasound in Medicine and Biology*, 26(9):1475–1488, 2000.

[119] A. Krupa, G. Fichtinger, and G. D. Hager. Full motion tracking in ultrasound using image speckle information and visual servoing. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2458–2464, 2007.

[120] A. Krupa, G. Fichtinger, and G. D. Hager. Real-time tissue tracking with B-mode ultrasound using speckle and visual servoing. In *Proceedings of the 10th International Conference on Medical Image Computing and Computer Assisted Intervention, Part II*, pages 1–8, 2007.

[121] A. Krupa, G. Fichtinger, and G. D. Hager. Real-time motion stabilization with B-mode ultrasound using image speckle information and visual servoing. *International Journal of Robotics Research*, To appear, 2009.

[122] A. Lagae and P. Dutré. A procedural object distribution function. *ACM Transactions on Graphics*, 24(4):1442–1461, 2005.

[123] L. Landini and L. Verrazzani. Spectral characterization of tissues microstructure by ultrasounds: a stochastic approach. *IEEE Transactions on Ultrasonics, Ferroeletrics and Frequency Control*, 37(5):448–456, 1990.

[124] A. Lang, P. Mousavi, G. Fichtinger, and P. Abolmaesumi. Fusion of electromagnetic tracking with speckle-tracked 3D freehand ultrasound using an unscented Kalman filter. In *Proceedings of the SPIE Medical Imaging Conference*, volume 7265, pages 72651A–1–72651A–12, 2009.

[125] C. Laporte and T. Arbel. Probabilistic speckle decorrelation for 3D ultrasound. In *Proceedings of the 10th International Conference on Medical Image Computing and Computer Assisted Intervention, Part I*, pages 925–932, 2007.

[126] C. Laporte and T. Arbel. Combinatorial and probabilistic fusion of noisy correlation measurements for untracked freehand 3D ultrasound. *IEEE Transactions on Medical Imaging*, 27(7):984–994, 2008.

[127] C. Laporte and T. Arbel. Learning a tissue invariant ultrasound speckle decorrelation model. In *Proceedings of the 6th IEEE International Symposium on Biomedical Imaging*, 2009.

[128] C. Laporte, J. J. Clark, and T. Arbel. A fractal multidimensional ultrasound scatterer distribution model. In *Proceedings of the 4th IEEE International Symposium on Biomedical Imaging*, pages 880–883, 2007.

[129] C. Laporte, J. J. Clark, and T. Arbel. Generalized Poisson 3-D scatterer distributions. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 56(2):410–414, 2009.

[130] N. D. Lawrence. Gaussian Processes — Matlab Software. `http://www.cs.man.ac.uk/~neill/gp`, 2008.

[131] A. Leroy, P. Mozer, Y. Payan, and J. Troccaz. Intensity-based registration of freehand 3D ultrasound and CT-scan images of the kidney. *International Journal of Computer Assisted Radiology and Surgery*, 2(1):31–41, 2007.

[132] M. Li. System and method for 3-D medical imaging using 2-D scan data. United States Patent no. 5,582,173, 1996.

[133] P.-C. Li, C.-Y. Li, and W.-C. Yeh. Tissue motion and elevational speckle decorrelation in freehand 3D ultrasound. *Ultrasonic Imaging*, 24:1–12, 2002.

[134] H. Liebgott, O. Bernard, C. Cachard, and D. Friboulet. Field simulation parameters design for realistic statistical parameters of Radio-Frequency ultrasound images. In *Proceedings of the IEEE Ultrasonics Symposium*, pages 2247–2250, 2007.

[135] E. D. Light, R. E. Davidsen, J. O. Fiering, T. A. Hruschka, and S. W. Smith. Progress in two dimensional arrays for real time volumetric imaging. *Ultrasonic Imaging*, 20:1–15, 1998.

[136] S. Lin and B. Kernighan. An effective heuristic algorithm for the traveling-salesman problem. *Operations Research*, 21:498–516, 1973.

[137] C. P. Loizou, C. S. Pattichis, C. I. Christodoulou, R. S. H. Istepanian, M. Pantziaris, and A. Nicolaides. Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 52(10):1653–1669, 2005.

[138] T. Loupas, W. N. McDicken, and P. L. Allan. An adaptive weighted median filter for speckle suppression in medical images. *IEEE Transactions on Circuits and Systems*, 36(1):129–135, 1989.

[139] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4(4):333–349, 1997.

[140] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[141] D. Marinakis and G. Dudek. Probabilistic self-localization for sensor networks. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, 2006.

[142] M. McCool and E. Fiume. Hierarchical Poisson disk sampling distributions. In *Proceedings of Graphics Interface*, pages 94–105, 1992.

[143] B. McShane, M. Adrian, E. T. Bradlow, and P. S. Fader. Count models based on Weibull interarrival times. *Journal of Business and Economic Statistics*, 26(3):369–378, 2008.

[144] S. Meairs, J. Beyer, and M. Hennerici. Reconstruction and visualization of irregularly sampled three- and four-dimensional ultrasound data for cerebrovascular applications. *Ultrasound in Medecine and Biology*, 26(2):263–272, 2000.

[145] P. Meer, D. Mintz, and A. Rosenfeld. Robust regression methods for computer vision: a review. *International Journal of Computer Vision*, 6(1):59–70, 1991.

[146] L. Mercier, T. Langø, F. Lindseth, and D. L. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in Medicine and Biology*, 31(4):449–471, 2005.

[147] O. V. Michailovich and A. Tannenbaum. Despeckling of medical ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 53(1):64–78, 2006.

[148] G. Y. Mirskii. Statistical errors of measurement of cross-correlation functions. *Measurement Techniques*, 23(9):783–787, 1980.

[149] L. Y. L. Mo, W. T. Hatfield, and S. C. Miller. Method and apparatus for tracking scan plane motion in free-hand three-dimensional ultrasound scanning using adaptive speckle decorrelation. United States Patent 6,012,458, 2000.

[150] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. FastSLAM: a factored solution to the simultaneous localization and mapping problem. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, 2002.

[151] D. C. Morrison, W. N. McDicken, and D. S. A. Smith. A motion artefact in real-time ultrasound scanners. *Ultrasound in Medicine and Biology*, 9(2):201–203, 1983.

[152] V. M. Narayanan, R. C. Molthen, P. Shankar, L. Vergara, and J. Reid. Studies on ultrasonic scattering from quasi-periodic structures. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 44(1):114–124, 1997.

[153] T. R. Nelson and D. H. Pretorius. Interactive acquisition, analysis, and visualization of sonographic volume data. *International Journal of Imaging Systems and Technology*, 8(1), 1997.

[154] J. Neyman and E. L. Scott. Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society B*, 20:1–43, 1958.

[155] J. Ng, R. W. Prager, N. Kingsbury, G. M. Treece, and A. H. Gee. Modelling ultrasound imaging as a linear, shift-variant system. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 53(3):549–563, 2006.

[156] M. A. Nixon, B. C. McCallum, W. R. Fright, and N. B. Price. The effects of metals and interfering fields on electromagnetic trackers. *Presence: Teleoperators and Virtual Environments*, 7(2):204–218, 1998.

[157] P. Novi Inverardi and A. Tagliani. Maximum entropy density estimation from fractional moments. *Communications in Statistics: Theory and Methods*, 32(2):327–345, 2003.

[158] B. J. Oosterveld, J. M. Thijssen, and W. A. Verhoef. Texture of B-mode echograms: 3-D simulations and experiments of the effects of diffraction and scatterer density. *Ultrasonic Imaging*, 7(2):142–160, 1985.

[159] A. Oppelt, editor. *Imaging Systems for Medical Diagnostics*. Publicis Corporate Publishing, 2005.

[160] Ö. Oralkan, A. S. Ergun, C.-H. Cheng, J. A. Johnson, M. Karaman, T. H. Lee, and B. T. Khuri-Yakub. Volumetric ultrasound imaging using 2-D CMUT arrays. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 50(11):1581–1594, 2003.

[161] A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw Hill, 4th edition, 2002.

[162] R. Piessens, E. de Doncker-Kapenga, C. W. Ueberhuber, and D. K. Kahaner. *QUADPACK: a subroutine package for automatic integration*. Springer Verlag, 1983.

[163] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Interpolation artefacts in mutual information-based image registration. *Computer Vision and Image Understanding*, 77:211–232, 2000.

[164] R. W. Prager, A. Gee, G. Treece, and L. Berman. Freehand 3D ultrasound without voxels: volume measurement and visualisation using the Stradx system. *Ultrasonics*, 40:109–115, 2002.

[165] R. W. Prager, A. H. Gee, and L. Berman. Stradx: real-time acquisition and visualization of freehand three-dimensional ultrasound. *Medical Image Analysis*, 3(2):129–140, 1999.

[166] R. W. Prager, A. H. Gee, G. Treece, C. J. C. Cash, and L. H. Berman. Sensorless freehand 3-D ultrasound using regression of the echo intensity. *Ultrasound in Medicine and Biology*, 29(3):437–446, 2003.

[167] R. W. Prager, A. H. Gee, G. M. Treece, and L. H. Berman. Analysis of speckle in ultrasound images using fractional order statistics and the homodyned k distribution. *Ultrasonics*, 40:133–137, 2002.

[168] R. W. Prager, A. H. Gee, G. M. Treece, and L. H. Berman. Decompression and speckle detection for ultrasound images using the homodyned k-distribution. *Pattern Recognition Letters*, 24:705–713, 2003.

[169] R. W. Prager, R. N. Rohling, A. H. Gee, and L. Berman. Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine and Biology*, 24(6):855–869, 1998.

[170] J. Quiñonero-Candela and C. E. Rasmussen. A unifying view of sparse approximate Gaussian process regression. *Journal of Machine Learning Research*, 6:1939–1959, 2005.

[171] J. Quiñonero-Candela, E. Snelson, and O. Williams. Sensible priors for sparse Bayesian learning. Technical Report MSR-TR-2007-121, Microsoft Research, 2007.

[172] U. Raeth, D. Schlaps, B. Limberg, I. Zuna, A. Lorenz, G. V. Kaick, W. J. Lorenz, and B. Kommerell. Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease. *Journal of Clinical Ultrasound*, 13:87–99, 1985.

[173] A. Rahimi, L.-P. Morency, and T. Darrell. Reducing drift in differential tracking. *Computer Vision and Image Understanding*, 109:97–111, 2008.

[174] C. E. Rasmussen and J. Quiñonero-Candela. Healing the relevance vector machine through augmentation. In *Proceedings of the 22nd International Conference on Machine Learning*, 2005.

[175] M. Riccabona, G. A. Fritz, H. Schöllnast, T. Schwarz, M. Deutschmann, and C. J. Mache. Hydronephrotic kidney: pediatric three-dimensional US for relative renal size assessment — initial experience. *Radiology*, 236:276–283, 2005.

[176] H. Rivaz, E. Boctor, and G. Fichtinger. Ultrasound speckle detection using low order moments. In *Proceedings of the IEEE Ultrasonics Symposium*, pages 2092–2095, 2006.

[177] H. Rivaz, E. Boctor, and G. Fichtinger. A robust meshing and calibration approach for sensorless freehand 3D ultrasound. In *Proceedings of the SPIE Medical Imaging Conference*, 2007.

[178] H. Rivaz, R. Zellars, G. Hager, G. Fichtinger, and E. Boctor. Beam steering approach for speckle characterization and out-of-plane motion estimation in real tissue. In *Proceedings of the IEEE Ultrasonics Symposium*, pages 781–784, 2007.

[179] R. Rohling, A. Gee, and L. Berman. A comparison of freehand three-dimensional ultrasound reconstruction techniques. *Medical Image Analysis*, 3(4):339–359, 1999.

[180] R. N. Rohling, A. H. Gee, and L. Berman. Automatic registration of 3-D ultrasound scans. *Ultrasound in Medicine and Biology*, 24(6):841–854, 1998.

[181] F. Rousseau, P. Hellier, and C. Barillot. A novel temporal calibration method for 3-D ultrasound. *IEEE Transactions on Medical Imaging*, 25(8):1108–1112, 2006.

[182] F. Rousseau, P. Hellier, M. M. J. Letteboer, W. J. Niessen, and C. Barillot. Quantitative evaluation of three calibration methods for 3-D freehand ultrasound. *IEEE Transactions on Medical Imaging*, 25(11):1492–1501, 2006.

[183] P. J. Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, 79:871–880, 1984.

[184] P. J. Rousseeuw and A. M. Leroy. *Robust Regression & Outlier Detection*. Wiley, 1987.

[185] M.-H. Roy Cardinal, J. Meunier, G. Soulez, R. L. Maurice, E. Thérasse, and G. Cloutier. Intravascular ultrasound image segmentation: a three-dimensional fast-marching method based on gray level distributions. *IEEE Transactions on Medical Imaging*, 25(5):590–601, 2006.

[186] D. B. Rubin. Iteratively reweighted least squares. In S. Kotz, C. B. Read, N. Balakrishnan, and B. Vidakovic, editors, *Encyclopedia of Statistical Sciences*. Wiley & Sons, 2006.

[187] R. Salim, B. Woelfer, M. Backos, L. Regan, and D. Jurkovic. Reproducibility of three-dimensional ultrasound diagnosis of congenital uterine anomalies. *Ultrasound in Obstetrics and Gynecology*, 21(6):578–582, 2003.

[188] O. Salvado and D. L. Wilson. Removal of local and biased global maxima in intensity-based registration. *Medical Image Analysis*, 11:183–196, 2007.

[189] R. San José-Estépar, M. Martín-Fernández, P. P. Caballero-Martínez, C. Alberola-López, and J. Ruiz-Alzola. A theoretical framework to three-dimensional ultrasound reconstruction from irregularly sampled data. *Ultrasound in Medicine and Biology*, 29(2):255–269, 2003.

[190] J. M. Sanches and J. S. Marques. A Rayleigh reconstruction/interpolation algorithm for 3D ultrasound. *Pattern Recognition Letters*, 21:917–926, 2000.

[191] J. M. Sanches and J. S. Marques. A multiscale algorithm for three-dimensional free-hand ultrasound. *Ultrasound in Medicine and Biology*, 28(8):1029–1040, 2002.

[192] J. M. Sanches and J. S. Marques. Joint image registration and volume reconstruction for 3D ultrasound. *Pattern Recognition Letters*, 24:791–800, 2003.

[193] D. Savéry and G. Cloutier. A point process approach to assess the frequency dependence of ultrasound backscattering by aggregating red blood cells. *Journal of the Acoustical Society of America*, 110(6):3252–3262, 2001.

[194] D. Savéry and G. Cloutier. Effect of red cell clustering and anisotropy on ultrasound blood backscatter: a Monte Carlo study. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 52(1):94–103, 2005.

[195] P. M. Shankar. A model for ultrasonic scattering from tissues based on the K distribution. *Physics in Medicine and Biology*, 40:1633–1649, 1995.

[196] P. M. Shankar. A general statistical model for ultrasonic backscattering from tissues. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 47(3):727–736, 2000.

[197] P. M. Shankar. The use of the compound probability density function in ultrasonic tissue characterization. *Physics in Medicine and Biology*, 49:1007–1015, 2004.

[198] P. M. Shankar, J. M. Reid, H. Ortega, C. W. Piccoli, and B. B. Goldberg. Use of non-Rayleigh statistics for the identification of tumors in ultrasonic B-scans of the breast. *IEEE Transactions on Medical Imaging*, 12(4):687–692, 1993.

[199] D. J. Sheskin. *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman & Hall/CRC, 2nd edition, 2000.

[200] R. Sim and G. Dudek. Learning generative models of scene features. *International Journal of Computer Vision*, 60(1):45–61, 2004.

[201] D. Simon. Intra-operative position sensing and tracking devices. In *Proceedings of the First Joint CVRMed / MRCAS Conference*, pages 62–64, 1997.

[202] R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. In I. Cox and G. Wilfong, editors, *Autonomous Robot Vehicles*, pages 167–193. Springer-Verlag, 1990.

[203] S. W. Smith, H. G. Pavy Jr., and O. T. von Ramm. High-speed ultrasound volumetric imaging system. i. transducer design and beam steering. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 38(2):100–108, 1991.

[204] W. Smith and A. Fenster. Statistical analysis of decorrelation-based transducer tracking for three-dimensional ultrasound. *Medical Physics*, 30(7):1580–1591, 2003.

[205] A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14:199–222, 2004.

[206] E. Snelson. *Flexible and efficient Gaussian process models for machine learning*. PhD thesis, University College London, 2007.

[207] E. Snelson and Z. Ghahramani. Sparse Gaussian processes using pseudo-inputs. In *Advances in Neural Information Processing Systems*, volume 18, 2006.

[208] O. V. Solberg, F. Lindseth, H. Torp, R. E. Blake, and T. A. N. Hernes. Freehand 3D ultrasound reconstruction algorithms — a review. *Ultrasound in Medicine and Biology*, 33(7):991–1009, 2007.

[209] P. C. Tay, B. Li, C. D. Garson, S. T. Acton, and J. A. Hossack. Left ventricle segmentation using model fitting and active surfaces. *Journal of Signal Processing Systems*, 55(1-3):139–156, 2009.

[210] A. Thayananthan, R. Navaratnam, B. Stenger, P. H. S. Torr, and R. Cipolla. Multivariate relevance vector machines for tracking. In *Proceedings of the European Conference on Computer Vision*, 2006.

[211] J. M. Thijssen. Ultrasonic speckle formation, analysis and processing applied to tissue characterization. *Pattern Recognition Letters*, 24:659–675, 2003.

[212] J. M. Thijssen, B. J. Oosterveld, P. C. Hartman, and G. J. E. Rosenbusch. Correlations between acoustic and texture parameters from RF and B-mode liver echograms. *Ultrasound in Medicine and Biology*, 19(1):13–20, 1993.

[213] M. E. Tipping. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.

[214] G. E. Trahey, J. W. Allison, S. W. Smith, and O. T. von Ramm. A quantitative approach to speckle reduction via frequency compounding. *Ultrasonic Imaging*, 8(3):151–164, 1986.

[215] G. E. Trahey, J. W. Allison, and O. T. von Ramm. Angle independent ultrasonic detection of blood flow. *IEEE Transactions on Biomedical Engineering*, 34(12):965–967, 1987.

[216] C. Trantakis, J. Meixensberger, D. Lindner, G. Strauss, G. Grunst, A. Schmidtgen, and S. Arnold. Iterative neuronavigation using 3D ultrasound. A feasibility study. *Neurological Research*, 24(7), 2002.

[217] G. Treece, R. Prager, A. Gee, and L. Berman. 3D ultrasound measurement of large organ volume. *Medical Image Analysis*, 5:41–54, 2001.

[218] G. M. Treece, R. W. Prager, A. H. Gee, and L. Berman. Correction of probe pressure artifacts in freehand 3D ultrasound. *Medical Image Analysis*, 6:199–214, 2002.

[219] T. A. Tuthill, J. F. Krücker, J. B. Fowlkes, and P. L. Carson. Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. *Radiology*, 209(2):575–582, 1998.

[220] T. A. Tuthill, R. H. Sperry, and K. J. Parker. Deviations from Rayleigh statistics in ultrasonic images. *Ultrasonic Imaging*, 10:81–89, 1988.

[221] S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, 1991.

[222] C. M. Urzúa. A class of maximum-entropy multivariate distributions. *Communications in Statistics, Part A, Theory and Methods*, 17(12):4039–4057, 1988.

[223] F. M. J. Valckx and J. M. Thijssen. Characterization of echographic image texture by cooccurrence matrix parameters. *Ultrasound in Medicine and Biology*, 23(4):559–571, 1997.

[224] E. B. van de Kraats, G. P. Penney, D. Tomaževič, T. van Walsum, and W. J. Niessen. Standardized evaluation methodology for 2-D—3-D registration. *IEEE Transactions on Medical Imaging*, 24(9):1177–1189, 2005.

[225] O. T. von Ramm, S. W. Smith, and H. G. Pavy Jr. High-speed ultrasound volumetric imaging system. ii. parallel processing and image display. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 38(2):109–115, 1991.

[226] D. Voorhies. Space-filling curves and a measure of coherence. In J. Arvo, editor, *Graphics Gems II*, pages 26–30. Academic Press, 1991.

[227] R. F. Wagner, M. F. Insana, and D. G. Brown. Statistical properties of radio-frequency and envelope-detected signals with applications to medical ultrasound. *Journal of the Optical Society of America A*, 4(5):910–922, 1987.

[228] R. F. Wagner, M. F. Insana, and S. W. Smith. Fundamental correlation lengths of coherent speckle in medical ultrasonic images. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 35(1):34–44, 1988.

[229] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez. Statistics of speckle in ultrasound B-scans. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):156–163, 1983.

[230] H. Wang and D. Suter. MDPE: a very robust estimator for model fitting and range image segmentation. *International Journal of Computer Vision*, 59(2):139–166, 2004.

[231] G. C. G. Wei and M. A. Tanner. A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. *Journal of the American Statistical Association*, 85(411):699–704, 1990.

[232] E. W. Weisstein. Hilbert curve. From Mathworld – A Wolfram Web Resource http://mathworld.wolfram.com/HilbertCurve.html, 1999.

[233] J. N. Welch, J. A. Johnson, M. R. Bax, R. Badr, and R. Shahidi. A real-time freehand 3D ultrasound system for image-guided surgery. In *Proceedings of the IEEE Ultrasonics Symposium*, pages 1601–1604, 2000.

[234] O. Williams, A. Blake, and R. Cipolla. Sparse Bayesian learning for efficient visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1292–1304, 2005.

[235] B. J. Winer, D. R. Brown, and K. M. Michels. *Statistical Principles in Experimental Design*. McGraw-Hill, 3rd edition, 1991.

[236] R. Winkelmann. Duration dependence and dispersion in count-data models. *Journal of Business and Economic Statistics*, 13(4):467–474, 1995.

[237] H. Wu, K. W. J. Malafant, L. K. Pendridge, P. J. H. Sharpe, and J. Walker. Simulation of two-dimensional point patterns: application of a lattice framework approach. *Ecological Modelling*, 38:299–308, 1987.

[238] W.-J. Wu and W. K. Moon. Ultrasound breast tumor image computer-aided diagnosis with texture and morphological features. *Academic Radiology*, 15:873–880, 2008.

[239] S. Xu, J. Kruecker, P. Guion, N. Glossop, Z. Neeman, P. Choyke, A. K. Singh, and B. J. Wood. Closed-loop control in fused MR-TRUS image-guided prostate biopsy. In *Proceedings of the 10th International Conference on Medical Image Computing and Computer Assisted Intervention, Part I*, pages 128–135, 2007.

[240] F. Yeung, S. F. Levinson, and K. J. Parker. Multilevel and motion model-based ultrasonic speckle tracking algorithms. *Ultrasound in Medicine and Biology*, 24(3):427–441, 1998.

[241] E. Yoxen. Seeing with sound: a study of the development of medical images. In W. E. Bijker, T. P. Hughes, and T. J. Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, pages 281–303. MIT Press, 1987.

[242] W. Yu, P. Yan, A. J. Sinusas, K. Thiele, and J. S. Duncan. Towards pointwise motion tracking in echocardiographic image sequences - comparing the reliability of different features for speckle tracking. *Medical Image Analysis*, 10:495–508, 2006.

[243] X. Yu, T. D. Bui, and A. Krzyzak. Robust estimation for range image segmentation and reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5):530–538, 1994.

[244] Y. Yu and S. T. Acton. Speckle reducing anisotropic diffusion. *IEEE Transactions on Image Processing*, 11(11):1260–1270, 2002.

[245] Y. Zhan and D. Shen. Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method. *IEEE Transactions on Medical Imaging*, 25(3):256–272, 2006.

[246] J. Zhang, Y. Wang, Y. Dong, and Y. Wang. Ultrasonographic feature selection and pattern classification for cervical lymph nodes using support vector machines. *Computer Methods and Programs in Biomedicine*, 88:75–84, 2007.

[247] S. Zhang, W. Yang, R. Yang, B. Ye, L. Chen, W. Ma, and Y. Chen. Noninvasive temperature monitoring in a wide range based on texture of ultrasound images. In *Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention*, 2006.

[248] W. Zhang and J. Košecká. Ensemble method for robust motion estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2006.

[249] S. Zhou, B. Georgescu, X. S. Zhou, and D. Comaniciu. Image based regression using boosting method. In *Proceedings of the 10th International Conference on Computer Vision*, volume 1, pages 541–548, 2005.

[250] X. Zong, A. F. Laine, and E. A. Geiser. Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. *IEEE Transactions on Medical Imaging*, 17(4):532–540, 1998.