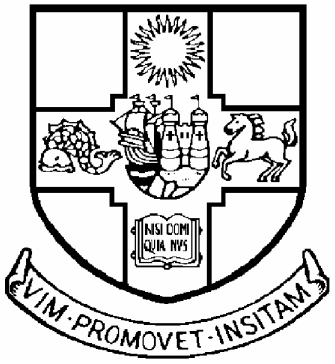
Computer Vision Elastography

James Duncan Revell



A thesis submitted to the University of Bristol in accordance with the requirements for the Degree of Doctor of Philosophy in the Faculty of Engineering, Department of Computer Science.

December, 2004.

36,591 words

Declaration

I declare that the work in this dissertation was carried out inaccordance with the Regulations of the Univeristy of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. The dissertation has not been presented to any other Univeristy for examination either in the United Kingdom or overseas. Any views expressed in the dissertation are those of the author.

James Revell

Abstract

This thesis is concerned with developing a two-dimensional (2D) ultrasound speckle tracking technique to quantify 2D axial and lateral strain fields for studying tissue dynamics. Knowledge of tissue displacement to infer strain characteristics is of major clinical importance. In part, this is due to the lack of simple and accurate non-invasive techniques to measure *in vivo* strain. The established technique for evaluating tissue behaviour using ultrasound is elastography, with conventional methods analysing the raw radio-frequency (RF) data to measure 1D displacements. In contrast, this work develops and applies image processing techniques to analyse the ultrasound images to measure 2D displacements. Measuring meaningful displacement descriptors using ultrasound assumes that speckle motion replicates small tissue deformation.

Current research using ultrasound image processing frequently concentrates on matching selected regions from frame pairs, an extension of RF elastography, to solve this correspondence problem. To assess the feasibility and effectiveness of displacement estimation using ultrasound images, a variable-sized block matching algorithm with a hierarchical full search is proposed. Unlike existing research, development and experimentation focuses on dynamic high frequency ultrasound of the musculoskeletal system. The aim is to provide a non-invasive *in vivo* approach to measuring displacement and strain in tendons and soft tissue for assessing tendinopathy.

With an emphasis on improving block matching approaches, enhancements are explored to extend this framework to sequences, generating trajectories. Further, novel contributions are incorporated to solve known speckle tracking problems. An automatic similarity measure selection is developed, adapting to varying speckle density and strain induced speckle decorrelation. Trajectories are also updated for moving objects in the scanned 3D volume that traverse the 2D plane. Other important challenges dealt with in this work include eliminating user frame pair pre-selection, freehand scanning registration, minimising tracking drift error and measuring displacement accuracy.

Finally, displacement fields are used to derive corresponding axial, lateral and shear strain elastograms. Further, the trajectory fields are used to generate spatiotemporal elastograms, producing object strain histories. The latter enables clinicians to have a unique means of comparing the knowledge of the applied motion to the strain that occurred during the activity, potentially aiding diagnosis. In addition, all results are validated using purpose-made tissue mimicking phantoms and *in vitro* tendon data, prior to analysing *in vivo* clinical sequences.

Key words: Ultrasound, Speckle Tracking, Elastography, Axial Lateral Strain Quantification, Musculoskeletal.

Acknowledgements

This thesis has benefited from the guidance of, and discussions with Dr Majid Mirmehdi in the Computer Vision Group at Bristol University, and from the support and enthusiasm of Dr Donal McNally in the Institute of Biomechanics at Nottingham University. I am indebted to both supervisors, as these discussions and continual support provided much encouragement over the course of this work.

I am appreciative to all the people that volunteered to be scanned, Mr Mark Parry for aiding in the construction of phantom experiments, and to the School of Veterinary Science at Bristol University for providing various *in vitro* specimens. I am also grateful to Professor Mark Batt a consultant in sports medicine at Nottingham, whom at short notice always provided clinical insight on tendon pain and pathology.

Additionally, I would like to acknowledge Professor Barry Thomas whom passed away before I completed. I felt fortunate that we laughed together at MIUA 2004, regarding his student antics at Imperial University.

Finally, I am thankful to family, friends and colleagues that made my time studying enjoyable, and the funding from the BBSRC that made this research possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Many biomechanical investigations require the quantification of loads and displacements or ideally stresses and strains. However, measurements of internal deformation and strain in biological soft tissues present a number of important challenges. Consequently current measurement techniques are far from ideal, particularly for *in vivo* conditions.

Measurements of surface strain such as mercury in silicone rubber [82] or solid state strain gauges [31] require the insertion of a device with barbed prongs into the tissue. They are therefore highly invasive when used for *in vivo* studies, and are subject to serious artefacts if there is a gradient in strain along the length of the prong. Furthermore, they can only measure deformation in one direction and at a single location. An alternative means of measuring surface strain includes video based techniques [60] to obtain two-dimensional (2D) information. However this requires a grid of markers to be placed on the tissue and viewed using one or more cameras. It is therefore unsuitable for *in vivo* measurements.

A method of measuring internal deformations is to use metal markers [86, 66] and a series of radiographs or fluoroscopy. This method is invasive, requiring implantation of metal beads or wires and exposure to ionising radiation, only giving information about deformation at the site of the markers. A similar technique is sonomicrometry [32], where ultrasound crystals are implanted and the distances between them are measured using the time of flight of acoustic

pulses. This still requires the invasive implantation of transducers and only gives information about the movement of these transducers. It is possible to measure deformations by using existing natural landmarks, such as the musculo-tendonous junction [47], using ultrasonography. However, such measurements are limited by the availability of reliably identifiable structures.

The ideal deformation measurement system would be entirely non-invasive and would give at least 2D-deformation information at any point in structures such as muscles, tendons and intervertebral discs in real-time. This specification rules out all surface strain techniques [82, 31, 60], all techniques requiring implanted transducers, beads or wires or relying on ionising radiation [86, 66, 32]. Whilst Magnetic Resonance Imaging (MRI) can be used to give quasi-static measurements of deformation [109], the acquisition time of several minutes (where the target area must be stationary) makes it unsuitable for real-time deformations, for example, from the gait cycle.

Observing the dynamic responses of soft tissues permits tissue physical properties including elasticity and stiffness to be inferred, as well as aiding in diagnosis. This is the principle of sonoelastography, which derives the biomechanical properties of tissue from measurements of motion resulting from an applied perturbation. Our main aim is to yield reliable and yet physically meaningful local descriptions of the motion field in musculoskeletal ultrasound.

1.2 Clinical Application Areas

This work concentrates on developing and applying techniques to musculoskeletal ultrasound video data, quantifying displacement and deformation in tendons and surrounding tissues. Minimal research currently exists on *in vitro* or *in vivo* human tendon mechanical properties [55], especially under loading. Most literature refers to various animal material testing and invasive approaches. Some important clinical applications that would benefit from a non-invasive approach to determining *in vivo* local deformation using musculoskeletal ultrasound are:

1. *Diagnosis.* Tendinopathy has many causes including inflammation, dislocations, tears, ruptures and overuse that result from currently immeasurable amounts of extreme or repetitive strain. An example benefit of this work is the non-invasive examination of

carpel tunnel syndrome, providing a method of measuring repetitive and forceful movements of the hand and wrist during work or leisure activities.

2. *Surgery.* Displacement and deformation information could enable the dynamic behaviour of suturing techniques used to repair tendons to be analysed, determining post-operative strength and performance, as well as the ability to investigate tendon-implant interfaces.

Broader applications include measuring displacements in other anatomical regions, for example, analysing intervertebral disc deformation.

1.3 Aims and Objectives

The principal aim of this work is to analyse ultrasound sequences to evaluate tissue deformation. To achieve this, several objectives are defined and corresponding research conducted:

⋆ To construct *in vitro* groundtruth and capture *in vivo* musculoskeletal sequences.

Throughout, gold standard purpose-made tissue mimicking phantoms and *in vitro* tendon pull data are used to validate estimated displacement, trajectory and strain results, before application to clinical *in vivo* musculoskeletal ultrasound sequences.

⋆ To assess 2D speckle tracking techniques from analysing selected ultrasound frame pairs.

We examine tissue and tendon deformation employing a 2D variable-sized block matching algorithm with a hierarchical full search, refined by extending displacements to sub-pixel accuracy. We show by application that this technique yields quantitatively reliable results.

⋆ To devise a method for measuring local internal motion in ultrasound sequences.

We develop a novel methodology to quantify temporal 2D axial and lateral displacements in ultrasound sequences. The principles of 2D interframe displacements produced by our earlier work (using hierarchical variable block size matching), are extended to quantify trajectories.

Solutions are provided for correcting transducer motion, quantifying objects moving in the 3D volume traversing the 2D plane, and eliminating previous manual frame pair pre-selection. To improve displacement accuracy, a combination of matching measures is also used. The first measure uses the normalised cross correlation for regions of strong signal and the second measure CD_2, specifically for regions of speckle determined by the signal to noise ratio.

- ⋆ To generate axial, lateral and shear strain maps for *in vitro* and *in vivo* sequences.

The biomechanical properties of any target object cannot be measured directly. A stimulus must be provided and the resulting motion assessed using a suitable technique. In our case the stimuli is the kinesis or articulation of the joint. Therefore, we derive 2D strain and strain rate information from the estimated displacements and velocity measurements, produced by our developed speckle tracking method. The computed trajectories are used to measure local temporal strain, deriving a strain history. This enables clinicians to compare the knowledge of the applied stimuli to the strain that occurred during activity. Spatial and temporal results are presented for both complete frames and specific clinical regions of interest.

Figure 1.1 illustrates the distinct areas of research that categorises these objectives, including freehand scanning, interframe and sequences analysis and strain estimation. The various stages shown will be dealt with in detail in the forthcoming chapters.

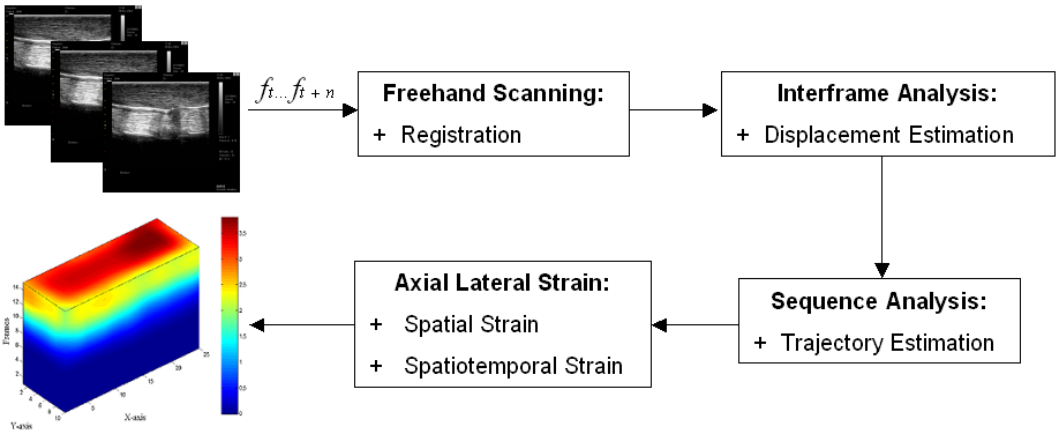


Figure 1.1: Flow diagram illustrating the main stages of research including data capture, interframe, sequence and strain analysis.

1.4 Contributions

The main contributions of this thesis are:

Displacement Estimation. An enhanced multiscale 2D speckle tracking methodology for ultrasound video sequences, using a novel alternating similarity measure approach to block matching motion estimation adapting to local signal.

Trajectory Estimation. An approach to extend interframe displacements to quantify trajectories of speckle motion, incorporating strategies to quantify objects moving in the 3D volume traversing the 2D plane and correcting for trajectory drift error.

Strain Estimation. The development of axial and lateral strain maps and novel spatiotemporal elastograms providing strain history, derived from displacement and trajectory measurements.

1.5 Thesis Overview

The remainder of this thesis is structured as follows:

- ⋆ *Chapter 2 Background,* provides an overview of the current state of research that has been published by others in the fields of elastography and computer vision elastography.
- ⋆ *Chapter 3 Groundtruth* in vitro *and Clinical* in vivo *Datasets,* explains the reasoning for the *in vitro* experiments and describes the various *in vivo* datasets used throughout [1].
- ⋆ *Chapter 4 Interframe Motion Estimation for Displacement Fields,* examines ultrasound tissue motion and deformation using a 2D variable-sized block matching algorithm.
- ⋆ *Chapter 5 Trajectory Fields from Motion Estimation,* presents the development and validation of a B-scan speckle tracking method to determine 2D temporal trajectories.
- ⋆ *Chapter 6 Strain Estimation from Trajectories,* uses the developed speckle tracking framework to quantify strain and elastograms.
- ⋆ *Chapter 7 Conclusions and Future Work,* discusses the proposed method and presents a hypothesis to automatically classify areas of signal and artefact.

[1] Several figures denoted with a ♠ have movies on-line at *http://www.cs.bris.ac.uk/home/revell/cve.html*.

Chapter 2

Background

2.1 Introduction

Chapter 1 considered current physical invasive techniques for measuring displacements in order to study soft tissue biomechanics. These comprised of strain gauges, high-speed video cameras and sonomicrometry. The disadvantages of these techniques included recording strain information at a limited number of specific locations, only evaluating function under arbitrarily chosen isometric conditions, invasive fixation methods and erroneous landmarking approaches. Motivated by these problems, in this thesis internal soft tissue deformation is captured using ultrasound, and thus the focus of our research is conducted into ultrasound motion estimation.

Ultrasound imaging provides a useful and cost effective way of examining *in vitro* and *in vivo* anatomical regions dynamically and in a non-invasive manner. Diagnostic ultrasound is used to view the musculoskeletal system, to detect problems with muscles, tendons, ligaments, joints and soft tissue. Throughout, we analyse a variety of musculoskeletal datasets concentrating on the achilles, patella and digital flexor tendons.

In this chapter we introduce an overview of the necessary ultrasound theory in Section 2.2, describing the B-mode modality, ultrasound artefacts and acoustic speckle. Next, to analyse the dynamic responses of soft tissues, we review elastography research in Section 2.3, followed by a study of the current image analysis algorithms used in similar research in Section 2.4.

2.2 Ultrasonography

The most commonly used modality in medical ultrasound is B-mode imaging, where an ultrasound transducer is placed against the skin directly over the region of interest (ROI). A typical ultrasound transducer employs an array of piezoelectric elements to generate short, $\approx 1\mu$s in duration, broad band pulses (with a centre frequency of about $10-15$ MHz for musculoskeletal work or 3.5–5 MHz for abdominal work). The array size determines the imaging system's aperture. The same transducer also receives the backscattered signals. The transmit signals passing to, and the received signals passing from the array elements, can be individually delayed in time, defining a phased array. Phased arrays are used to electronically steer and focus (beamforming) the sequence of acoustic pulses through the target volume. Processing these echo signals routinely begins at the individual channel (element) level to produce A-lines (A-mode).

The general formation of B-mode sequences commences with (a) RF demodulation or envelope detection storing resulting A-modes in a 2D image matrix, (b) attenuation correction using time gain compensation (TGC) or swept and lateral gains to increase signal amplification from increasing depths, (c) scan conversion (8 bit digitisation) to allow the B-mode to be displayed with a defined resolution (known as a B-scan), and finally (d) logarithmic compression to adjust the large echo dynamic range ($60-100$ dB). This process is illustrated in Figure 2.1. The B-scan sequences captured and analysed throughout, are those processed and displayed by the ultrasound machine, with a uniform dynamic range of intensities between 0 and 255.

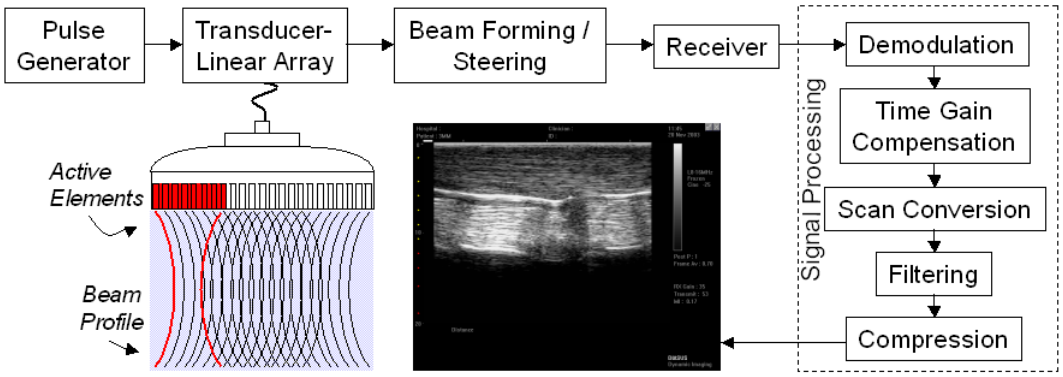


Figure 2.1: The processes used to generate a B-scan. B-scans are composed of a set of axial RF signals representing the response magnitude from a pulse generator using a linear array transducer. Since the response magnitude delays exponentially with depth, it is log-amplified prior to quantisation and display.

Consequently, the B-mode signal is suboptimal compared to the raw RF, with a fine axial resolution (high spatial frequency) but a low lateral resolution. As a result, decorrelation due to relative lateral scatterer (see Section 2.2.2) motion is far more severe than relative axial scatterer motion. Unfortunately, this is the direction of interest when the transducer is held at 90° to acquire longitudinal traversing tendon sections (see Section 2.2.1). The resulting real-time images represent cross sectional or tomographic views of the plane that the beam was swept across, capturing non-rigid motion (without constant shape) with a freehand transducer.

2.2.1 Ultrasound Artefacts

Imaging artefacts refer to components generated in the imaging process rather than true reproductions of the target structures. Ultrasound imaging artefacts include:

1. The *beam width artefact,* caused by a wide beam creating image smearing (intensity blurring) of echogenic objects that are smaller than the beam diameter.

2. The *enhancement* and *shadow artefacts,* caused by attenuation variations creating artificially increased or decreased image intensities respectively.

3. The *mirror image artefact* or *multipath reflection artefact, reverberation artefact* and *split image artefact.* These arise when the ultrasound signal fails to travel in a straight line directly to the reflecting object and back to the transducer.

4. The *speckle motion artefact,* which is the difference between the acoustic speckle motion and the underlying tissue motion.

5. The *speed artefact,* brought about by variances in material speed of sound values. For example, the speed of sound in fat (1450 m/s compared to 1540 m/s in soft tissues) can delay the return signal, resulting in reflections in the B-scan appearing incorrectly further from the transducer.

6. The *ambiguity artefact,* when echoes from all depths do not reach the transducer before the next ultrasound pulse is emitted, usually from a high pulse repetition frequency.

Specifically, for optimal imaging it is necessary to keep the transducer perpendicular to the tendon. An angle $< 85^\circ$ or $> 95^\circ$ between the emitted sound waves and the tendon (or any flat

reflector) cause the majority of reflected waves to not be received by the transducer. Consequently, tendon appearance (see Chapter 3) becomes increasingly hypoechoic as the transducer is tilted. Sequences used in this thesis are captured with the transducer at the optimal position.

2.2.2 Speckle Formation

Ultrasound image analysis in general is complex due to data composition, which is described in terms of speckle formation. Speckle is structured noise from a medium containing many scatterers. Speckle appearance is dependent on bandwidth, frequency and manufacturer of the employed transducer, in addition to the geometry and sub-wavelength structure of the tissue. Echographic speckle texture of imaged tissue is mainly due to intensity scattering, implying structures are smaller than the sampling volume (a product of spatial pulse length and beam cross section). Upon visual inspection speckle consists of a relatively high grey level intensity [28], qualitatively ranging between hyperechoic (bright) to a hypoechoic (dark) domain.

Scatter occurs when small imperfections (scatterers) in the target cause seemingly random reflections and refractions of the sound wave. The textures created do not correspond to underlying structure, but the intensity reflects the local echogenicity of the underlying scatterers. Scatterers account for a decrease in image quality, causing blurring and decreased intensity at impedance boundaries, while within the medium they create speckle. The statistics of the signal depends on the density of scatterers, with a large number of randomly located scatterers following a Rayleigh distribution (known as fully developed speckle). This is discussed further in Section 3.3.3. These conditions are seldom met, resulting in several different statistical models of speckle being used, for example [35][48]. Further, speckle is multiplicative in the sense that as the average grey level of a local area increases so does the noise level.

In this work, we encountered 2 different forms of speckle textures that occur from tendon and surrounding soft tissue. In the tendon region the speckle texture is extremely striated and unidirectional, which is quite contrasting to the typical soft tissue speckle textures. The importance to displacement estimation of this textural difference is studied in Chapter 5.

In general, since the effect of additive sensor noise is much smaller than speckle noise η [1], the formation of each B-scan I can be written as:

$$I(x, y) \approx \eta(x, y)s(x, y) \tag{2.1}$$

where s is the original signal corrupted with multiplicative speckle noise η. Typically, to use the dynamic range of the display, nonlinear signal conversion is commonly employed to show weak echoes on the same scale with strong specular reflections. The logarithmic amplification transforms (2.1) into the traditional signal with additive noise form:

$$\log(I(x,y)) = \log(\eta(x,y)) + \log(s(x,y)) \tag{2.2}$$

2.3 Elastography

Over the last decade, ultrasonic imaging for analysing non-invasive soft tissue mechanical behaviour under applied loading has seen considerable research [6]. Unlike elastic MRI, ultrasound provides low cost, real-time imaging, enriched with coherent speckle textures. The fundamental assumption for ultrasound analysis is that speckle motion adequately represents the underlying tissue motion for small compressions, therefore many techniques use variations of speckle tracking to estimate tissue elasticity.

The established approach to measuring the elasticity of soft tissue is elastography, introduced by Ophir et al. [70], using time-domain speckle tracking to estimate 1D longitudinal axial strain images called elastograms. The technique acquires sets of digitised radio-frequency (RF) echo lines for a ROI for cross correlation analysis, pre and post-deformation. An illustration of the technique is shown in Figure 2.2 and examples of the output stages of the process are shown in Figure 2.3. Conventionally an ultrasound transducer is utilised to compress the tissue by $< 1\%$ (*in vivo*) of the tissue depth [101]. Using a high temporal resolution, elastography assumes the RF signal patterns are constant between acquisitions for small object deformation.

Since, RF elastography has undergone many developments solving limitations of the original approach. To compensate for RF distortions from tissue compression, Varghese et al. [104] temporally stretched (using 1D interpolation) the post-compressed signal. This, however, assumes a priori knowledge of strain magnitude and a constant local deformation. Further, the elastogram signal to noise ratio was increased using elastogram multicompression averaging. Other improvements include multiresolution 1D windows on congruent RF echo lines at all depths, applying overlapping steps along the temporal axis [103].

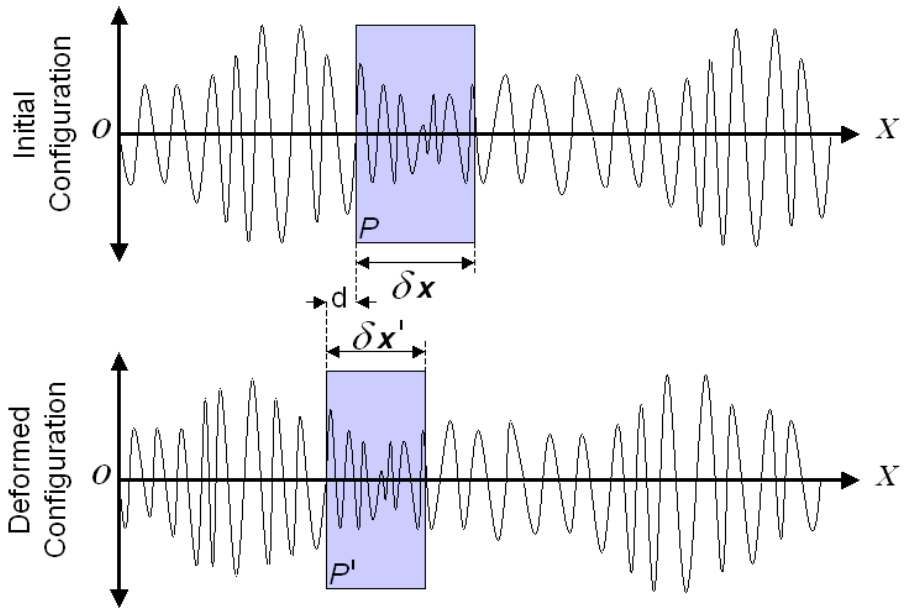
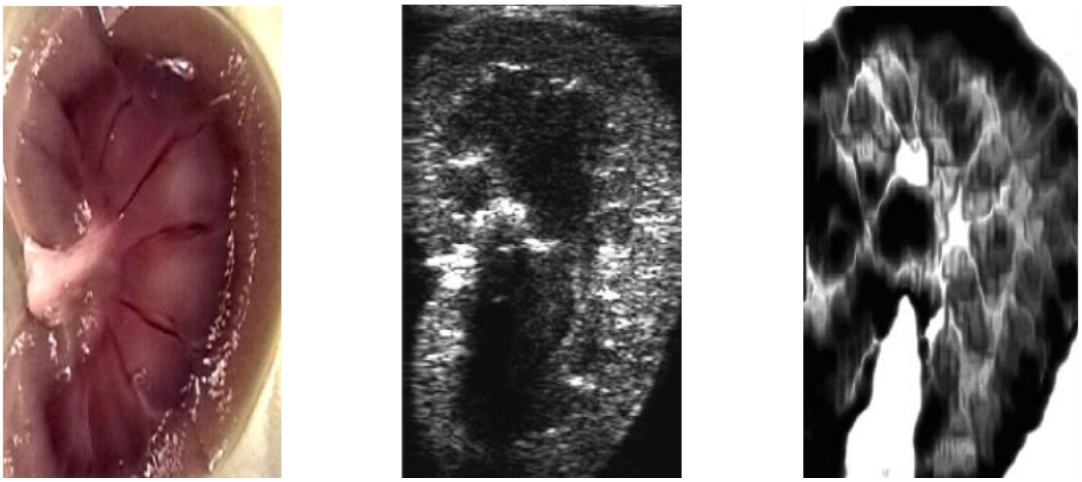


Figure 2.2: Traditional 1D elastography measuring the displacement $\boldsymbol{d}$ between the pre and post-deformed axial RF signals. The shaded regions highlights the signals used for cross correlation.

Until recently, RF elastography only quantified the axial strain component, with the lateral and elevation components both discarded. Lubinski et al. [53] first used axial displacements to compute lateral displacements, under a set of assumptions describing strain boundary conditions. It has since been demonstrated by Konofagou and Ophir [46] that precision lateral displacements can produce elastograms. The lateral elastograms are derived by lateral 1D cross correlating the pre-compressed signal, with the original and a linear interpolation of the neighbouring post-compressed signal. By displaying axial or lateral strain values in a matrix according to their spatial location a grey scale strain image or elastogram can be produced. Recently D'hooge et al. [25] used this RF technique and the sum of absolute differences (SAD) measure for estimating 2D strain (1D axial followed by 1D lateral) for cardiac cycles.

Elastography variations include Miller et al. [63] acquiring RF echo lines from pre and post-heated tissue, assuming the speed of sound in tissue changes as a function of temperature. Alternatively, instead of applying a relatively slowly changing (quasi-static) compression, Lerner et al. [49] and Parker et al. [72] developed sonoelastic imaging, propagating low-frequency vibrating shear waves into tissue creating displacement. A form of Doppler is then used to de-



(a) Photograph. (b) B-mode. (c) Elastogram.

Figure 2.3: An *in vitro* example using RF elastography, from [46], including (a) a photograph of a kidney, (b) the equivalent ultrasound image of the kidney in a gelatin block and (c) corresponding axial elastogram obtained from a 0.5% compression (low strains in black and high strains in white).

tect the tissue vibration response, with an absence of vibration signifying a possible presence of tumours [85]. Pesavento et al. [75] compared both conventional elastography and sonoelastic imaging for real-time application, noting the latter can be prone to aliasing[1] (streaking) and frequency dependent attenuation (shadowing). Later, Nightingale et al. [69] used acoustic radiation force to create displacements, employing a series of ultrasonic pulses targeted at a specific region of tissue. Echoes from a subset of these pulses are then processed yielding time and displacement data for the particular location.

Elastography techniques are predominantly invasive measuring *only* relatively small strains in 1D along the direction of wave propagation. With few exceptions elastography has remained an off-line technique requiring hardware with RF access restricting clinical application. The remainder of this chapter describes a speckle tracking technique, which presents a different approach to elastography by analysing the ultrasound images. Tracking in ultrasound images to estimate displacement information is not aimed at replacing elastography methods, but to offer several important advantages to clinicians, for example, generating strain information without using specific hardware to process RF signals.

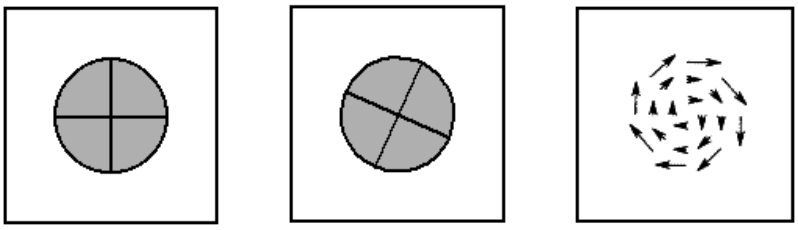
[1] A visible streaking distortion due to the low frequency undersampling of the signal.

2.4 B-scan Speckle Tracking Methodologies

Alternatively, using B-mode images 2D tissue motion can be measured by tracking the movement of the speckle produced by the back-scattering of the ultrasound signal itself, for example [19][33]. Speckle tracking can measure both axial and lateral motion in the image plane from analysing successive ultrasound B-mode frames, without modifying existing hardware. Also, strain is derived as the gradient of these displacements, and can be used as a first approximation for Young's modulus[2]. B-scan speckle tracking methodologies aim to estimate 2D motion (displacement or correspondence) fields from image sequences. B-scans or any still images are defined as a spatial distribution of intensities that are constant with respect to time. Therefore a sequence of ultrasound images or video, is defined as a time varying spatiotemporal intensity pattern, $I_t(x, y)$ where x, y and t are spatiotemporal variables. In this section, we define the brightness change constraint equation, the basis of all methods, followed by providing a literature review of the approaches using 2D speckle tracking methodologies in ultrasound.

2.4.1 Optical Flow for Speckle Tracking

In essence this is an optical flow problem where extracting a visual displacement flow field can explain changes in an image sequence. For example, optical flow is a method that measures the velocity field due to the apparent motion of an objects' brightness pattern in 2D image sequences during a time interval. An artificial example demonstrating the concept is shown in Figure 2.4 from [90], where the observed optical flow is equivalent to the 2D motion field.



(a) f_t (b) f_{t+1} (c) Displacement field.

Figure 2.4: Conceptual example: (a) Reference frame, (b) new frame and (c) optical flow.

[2]Young's modulus is the ratio of stress to strain on the loading plane along the loading direction, see Chapter 6.

The displacement field is defined as a field of vectors that measure the temporal rate of change, $\boldsymbol{d}_t(u,v) = (dx/dt, dy/dt)$ at a particular point (x,y,t), where $u = \frac{dx}{dt}$ is the velocity along the x-axis and $v = \frac{dy}{dt}$ is the velocity along the y-axis.

Optical flow assumes that the intensity at any point remains constant along a motion trajectory:

$$\frac{dI_t(x,y)}{dt} = 0 \tag{2.3}$$

The above is a total derivative expression, denoting the rate of change of intensity along the motion trajectory, which can be written as the well known motion constraint equation:

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0 \tag{2.4}$$

defining the optical flow brightness change constraint (pixel intensity invariance over time), where $u = dx/dt$ and $v = dy/dt$ denote the components of the velocity vector. Further, the spatial coherence constraint assumes the optical flow within a region changes gradually.

A large number of algorithms exist to measure motion and all make the assumption of the brightness change constraint, but in different forms. This has instigated research into motion estimation comparative studies. These include Barron et al. [9] and Galvin et al. [30], allowing displacement accuracy and algorithm applicability issues to be comprehended. However, few of the techniques reviewed in these works are employed for ultrasound speckle tracking. The reason becomes apparent from the following review of those techniques (differential and energy based and block matching), which have been applied to ultrasound data.

2.4.2 Differential Based Techniques

Early approaches include Meunier and Bertrand [62] using differential or gradient based techniques. Unfortunately, success was only reported with synthetic ultrasound images. Poor performance was also observed by Mailloux et al. [57, 58] applying the standard global optical flow of Horn and Schunck [38] to *in vivo* cardiac sequences. The authors observed that input images required smooth spatial derivatives, which they achieved by applying a 15×15 mean filter to each frame. Similarly, Chunke et al. [18] directly applied the Kanade-Lucas-Tomasi [97] local optical flow motion estimation technique, which assumes constant motion within a window. Sühling et al. [92] extended the same local model to account for typical heart motions

such as dilation, contraction and shear using an affine model. Another extension also included estimating displacement vectors at multiple scales and selecting the most promising fit. Baraldi et al. [8] compared the performance of both the Lucas-Kanade [54] and Horn-Schunck [38] differential optical flow algorithms on synthetic 2D echo image sequences. The approaches only differ in the choice of the smoothing term and in the local or global treatment of the domain. Results were comparable to the previous research, with only reliable results from synthetic data and large errors for *in vivo* data. This is partly as a result of acoustic speckle that means there is insufficient gradient information, reducing accurate numerical differentiation of (2.4).

To summarise, it can be observed that these approaches require that the spatiotemporal partial derivatives of the intensity can be estimated with sufficient accuracy. In Section 4.4 a differential based approach is applied and compared to several popular block matching approaches.

2.4.3 Energy Based Techniques

Less popular for ultrasound sequences are the energy based (or frequency based) techniques for motion estimation. These techniques assume that 2D textures have uniform spectral content, and have translational motion with constant velocity in the spatial domain. Under these assumptions, the power spectrum (square of the absolute magnitude of the Fourier transform) of the moving signal is uniformly distributed along a tilted plane, passing through the origin in the spatiotemporal frequency domain [36, 9]. By estimating the orientation of this plane, the displacement components of $\boldsymbol{d}$ can be recovered.

Cooper and Graham [21] addressed the application of estimating motion in transabdominal ultrasound sequences, using energy based techniques. The authors specifically applied Heeger's [37] spectral motion estimation technique for speckle textured sequences. The orientation of the plane, was estimated by locally sampling the power spectrum by convolution using spatial and temporal Gabor filters. Unfortunately, when the signal motion is not in the form of a translational shift, the energy spreads out around this plane, making orientation and location less robust. Consequently, this approach potentially requires the tuning of a large range of optimal filters (a minimum of 12 tuned to different spatial orientations and different temporal frequencies for synthetic data), especially for *in vivo* sequences that contain complex motion. Recently, Cooper et al. [22] implemented an alternative more simplistic scheme to reduce this

problem, whereby an extensive full search is performed that samples the power spectrum using multiple planes, with the plane containing the highest power selected as the best fit.

To summarise, it is apparent that the aperture problem remains a significant issue causing large errors, as in the spatiotemporal frequency domain the power spectrum of the moving signal is restricted to a line (where normal flow, defined in Section 2.5.1, is the gradient of the line). Finally, Adelson and Bergen [2] interestingly showed that certain energy based techniques are equivalent to correlation based methods.

2.4.4 Block Matching Techniques

To date, the most popular approaches to speckle tracking use 2D region-based matching that assume the optical flow (2.4) is constant over a defined region. Frame to frame matching of features is referred to as the correspondence problem. In the literature the term block, region, window, template and kernel are equivalent, all defining a localised, small, constrained, 2D subset of values. The purpose is to locate the best match for a block $I(i, j)$ in frame f_t, somewhere in frame f_{t+1} given numerous blocks of exactly the same dimensions (Figure 2.5).

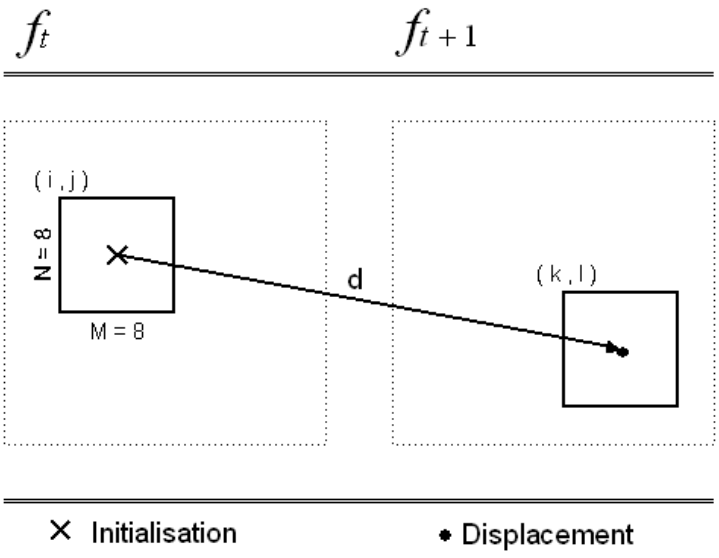


Figure 2.5: FSBM illustrating the reference block in f_t and final displacement in f_{t+1}, using a full search in a defined search region.

Fixed size block matching (FSBM) is a common method in computer vision for matching moving objects from one image to the next. It provides the basis for motion object tracking and segmentation. Many examples of this can be found in [94]. Typical implementations, such as Jain and Jain [40], use fixed size blocks of the order of 16×16 pixels using restricted search regions of 64×64. Interframe motion is quantified by matching sub-regions of a reference frame f_t at time t to a new frame f_{t+1} minimising or maximising various measured criteria. A good match requires the whole block to have undergone the same translational motion. The choice of block size, search range and matching criterion affect the reliability of displacements. Smaller and more numerous blocks can better represent complex motion (but are more likely to converge on inappropriate local minima) than fewer large blocks, that represent only large global motion well. This problem has been investigated by Ribas-Corbera and Neuhoff [81], concluding that the optimum block size can be affected not only by motion vector accuracy, but also by other scene characteristics such as texture and interframe noise.

Early work includes Bohs et al. [12], acquiring displacements in real-time for *in vitro* datasets using the sum of absolute differences (SAD) as a region matching criterion. The authors note that B-mode data, although having a coarser resolution, tended to be more uniform than RF data. Golemati et al. [33] implemented FSBM using correlation coefficients as the matching criterion for speckle tracking, to determine carotid artery wall motion from B-scans. Although multi-resolutions or sub-pixel accuracy were not analysed, the method yielded accurate results.

Pratikakis et al. [79] used a multiresolution block matching scheme for the application of tissue deformation tracking during surgery. Using a multiresolution framework improved the quality of the initial estimates compared to FSBM. Further block matching adaptation is the deformable mesh, applied by Yeung et al. [110] for feature adaptive motion tracking of ultrasound sequences of muscle contractions. The authors manually allocate nodes connecting the mesh elements adaptively to stable speckle patterns, which are less susceptible to temporal decorrelation producing successful non-rigid displacement fields in ultrasound. Later, various researchers (such as Bamber and Bush [7], Shapo et al.[87], Morsy and Von Ramm [65], Wang et al. [107] and Vogt et al. [105]) favoured normalised cross correlation to estimate tissue motion, which is explored later.

2.5 B-scan Speckle Tracking Issues

2D motion estimation solutions suffer from uniqueness, existence and continuity problems:

1. Uniqueness: When displacement components are treated independently, there are two unknowns, leading to a potential aperture problem.
2. Existence: No correspondence solution exists for both covered and uncovered background regions. This is known as the occlusion problem.
3. Continuity: Motion estimation is sensitive to the presence of noise, with small amounts of noise resulting in a large deviation in the continuity of motion estimates.

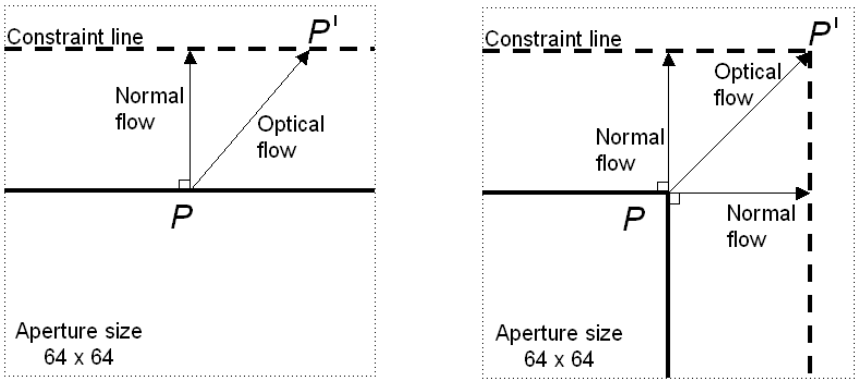
The aperture, occlusion and continuity problems are described in the following subsections.

2.5.1 Aperture Problem

The aperture problem is illustrated in Figure 2.6, where an object moves from P to P' corresponding to the optical flow displacement vector. In Figure 2.6(a) it is not possible to measure both displacement components. Only the vertical motion or normal flow (the motion perpendicular to the object) can be measured. However, in Figure 2.6(b) it is possible to estimate both displacement components, as there are gradients in both perpendicular directions. Therefore, the aperture problem can be overcome for regions of pixel intensities that contain sufficient variation and share the same motion. We show an approach to minimising inaccuracies in displacement estimation caused from the aperture problem, later in Chapter 5.

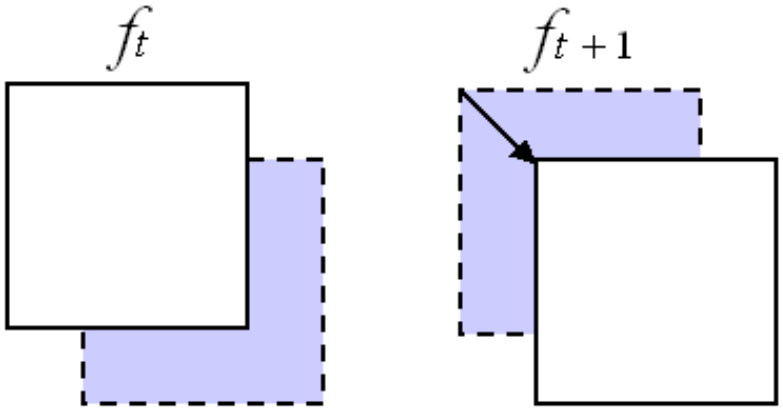
2.5.2 Occlusion Problem

The occlusion problem refers to the covering and uncovering of an object's surface that was only in part of the field of view. This is illustrated in Figure 2.7, where the object indicated by the solid line translates diagonally from frame f_t to f_{t+1}. The shaded region in f_t indicates the future covered region in f_{t+1}, hence no correspondence exists for the pixels in f_{t+1}. The shaded region in f_{t+1} indicates the region uncovered by the motion of the object, and as a result no correspondences exists for the pixels in f_t. Reasons for occlusion in ultrasound sequences are given in Section 5.3.6, including an approach to update for the occlusion of data.



(a) Aperture 1. (b) Aperture 2.

Figure 2.6: The aperture problem: (a) Aperture 1, single displacement component estimated and (b) aperture 2, both displacement components estimated.



(a) Covered. (b) Uncovered.

Figure 2.7: The occlusion problem: (a) Region to be covered (no region in f_{t+1} matches this region in f_t) and (b) region uncovered (no displacement represents the motion in the uncovered region in f_{t+1}).

2.5.3 Continuity Problem

The continuity problem, discontinuous inaccurate displacements caused by noise in the images, can be reduced by sequence pre-processing. Pre-processing using filtering, smooths spatial derivatives leading to improved displacement accuracy, specifically for the differential techniques discussed in Section 2.4.2. Speckle filtering is an area of active research with many

existing techniques. Categories and example filter methods used for ultrasound include:

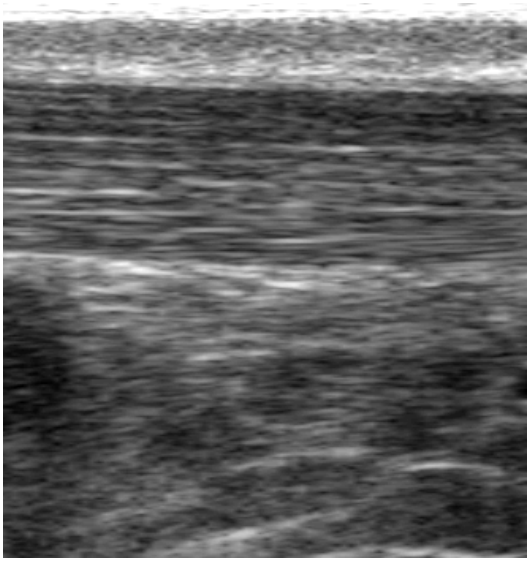
1. Spatial Non-adaptive Filter: Examples include the median and mean low-pass filters (the latter used by Mailloux et al. [56]).

2. Spatial Adaptive Filter: Popular approaches include the classic Frost filter (Frost et al. [29]), adaptive weighted median filter (Loupas et al. [52]), wavelet coefficient thresholding for echocardiographic images (Zong et al. [114]), multiscale nonlinear thresholding (Hao et al. [35]) and lastly real-time anisotropic diffusion (Abd-Elmoniem et al. [1]).

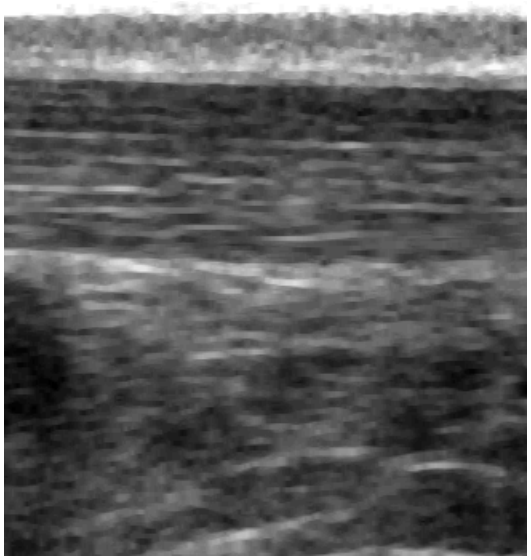
3. Spatiotemporal Filter: Methods range from compounding or multiple frame averaging (Klingler et al.[45], and Friedland and Adam [28] temporally averaging three frames with a weighting on the central frame), to a motion-adaptive biased least mean squares temporal filter (Evans and Nixon [26]).

Figure 2.8 illustrates sample filtering results from several techniques from each of the categories defined above, applied to a typical musculoskeletal ultrasound image. The original and filtered output images from using a low-pass median filter, anisotropic diffusion[3] and multiple frame averaging, are shown in Figures 2.8(a)-2.8(d) respectively. The performance of each filter has received attention elsewhere, however, qualitatively we can observe some general characteristics. First, the median filter has preserved edges but blurred some important fine detail (removing high frequency information). Second, anisotropic diffusion has smoothed the speckle and maintained edge information, generating local homogeneous regions. These regions can cause erroneous displacements, as pixels in f_t can correspond to several pixels within a region in f_{t+1}. Yu and Acton [112] use this technique of extensive speckle over-smoothing for a segmentation scheme. Lastly, frame averaging, which exploits the non-stationarity of speckle by combining a series of input images, produced a blurred image with resolution loss. These example results lead to improving visual understanding, producing homogeneous regions with clearer contrasting regions, and enhanced boundaries that are detrimental for speckle tracking.

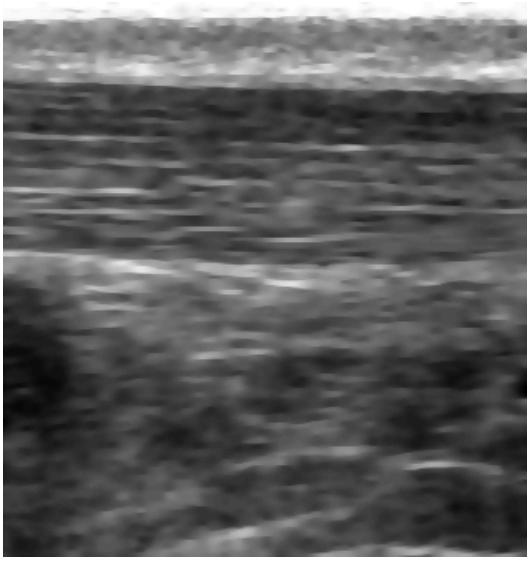
[3]This filtering process updates each pixel depending on its nearest neighbours, selectively smoothing regions of pixels based on the iterative evaluation of a diffusion function, which is typically based on the image gradient magnitude. Further information is available in [74].



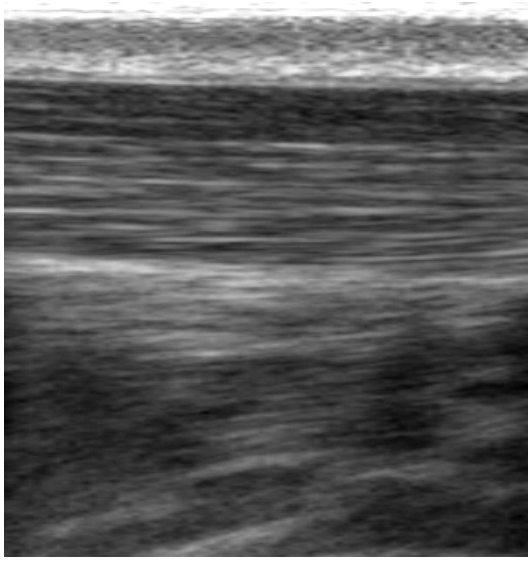
(a) Original test image.



(b) Median filter output.



(c) Anisotropic diffusion output.



(d) Multiple frame averaging output.

Figure 2.8: Speckle filtering: (a) Original image of *in vivo* achilles tendon and tissue, (b) filtered image using 3×3 median filter, (c) filtered image using anisotropic diffusion [74] with 10 iterations, and (d) filtered image using multiple 5 frame averaging.

In this work, we favour displacement estimation with displacement post-processing, rather than speckle filter pre-processing and then displacement estimation. Although much research has been aimed at removing speckle to enhance ultrasound image understanding, many schemes produce increasingly localised homogeneous regions. This is due to features that are the same scale as the speckle being eliminated, observed by [113], impeding local motion estimation.

Filter performance tends to be measured by quantifying edges and boundaries, with speckle preservation and fluctuation reduction measured using the co-occurrence matrix and localised mean and standard deviation, (known as the speckle signal to noise ratio, used later for measuring speckle density). In our situation all echo information is maintained, justifying a region-based motion estimation approach, as discussed in Section 2.4.4, which has some inherent robustness to speckle incoherence and machine noise for speckle tracking. An experiment is conducted in Chapter 4 to measure speckle stability and non-stationary noise. The next section describes the only pre-processing that has occurred as part of capturing the sequences.

2.6 Signal Processing

The objective of any displacement estimation algorithm is to track the same group of scatterers in the image of the pre-deformed tissue, to the image of the post-deformed tissue. Therefore the performance of displacement estimation using B-scans, is dependent on the quality of the images captured. Contrast and resolution are the two most important quality parameters for ultrasound images. For example, poor user gain selection can create regions of saturated intensity or areas of very low inhomogeneous image contrast. This can result in relatively flat correlation surfaces with negligible significant peaks. A range of methods exist for automatically adjusting swept gain functions. The simplest, applies a pre-set TGC function based on the average image properties for the transducer used, to continuously estimating the attenuation coefficient by analysing the backscattered signal properties.

Sequences capturing large motions can also cause the TGC and lateral gain compensation to vary for the same tissue at different frames, although negligible in musculoskeletal sequences due to the presence of strong lateral motion. For sequences of low contrast intensity, normalisation or constant variance enhancement, described in [76], can be applied as a pre-processing

stage. These local methods produce images with the same mean and variance, by subtracting the mean and dividing by the standard deviation of the intensity values in the region.

2.7 Elastography: RF-data vs B-scan

Some comparisons between traditional RF elastography and using B-scans are shown in Table 2.1. Although B-scans are a potentially degraded signal (from machine settings and internal processing, Section 2.2) compared to the RF signal, this is insignificant to the benefits that can be gained. The most favourable attributes include, reduced data storage compared to A-lines, the vast availability of machines with direct digital video capture feeds, and the ability to measure a large range of displacements and strains.

Attributes	**RF-data**	**B-scan**
Modality	A-mode	B-mode
Dimensionality	1D	2D
Signal Quality	Raw	Processed
Hardware Compatibility	Modified for RF access	Full support
Deformation Range	$0<1\%$ *in vivo*	$0<5\%$ *in vivo*
Potential Process Rate	Real-time	Real-time
Clinical Applicability	High	High

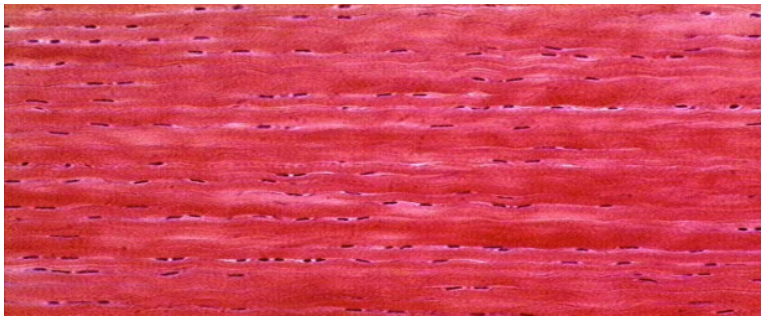
Table 2.1: General comparison of attributes between traditional 1D RF elastography and 2D B-scan computer vision elastography.

2.8 Conclusion

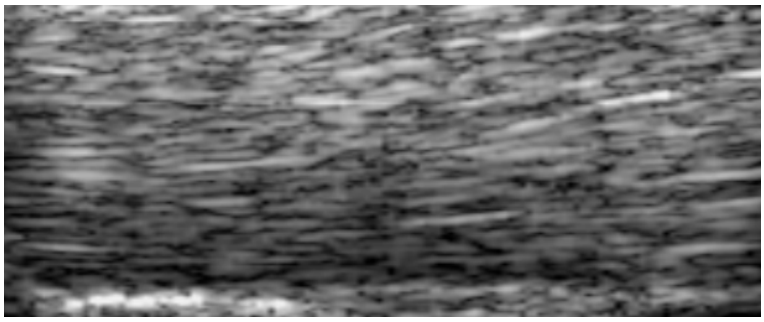
In this chapter, we have attempted to provide an overview of the existing methods to estimate displacements using ultrasound speckle tracking. This has commenced with a summary of the processes required to construct a B-scan using ultrasound, artefacts that can occur and the formation of speckle. A detailed review of the current state of elastography followed, which over

the past decade has developed rapidly, due to the clinical need to further our understanding of the dynamic properties of *in vivo* soft tissue. In comparison to RF elastography, B-scan tracking methodologies have been explored, specifically the differential, energy based and block matching techniques. The advantages of computer vision approaches have been clearly stated, overcoming the lack of available ultrasound machines with RF access and limitation to 1D analysis followed by a further 1D analysis to simulate 2D analysis. Also the disadvantages of B-scan tracking, for example, aperture, occlusion and displacement continuity issues.

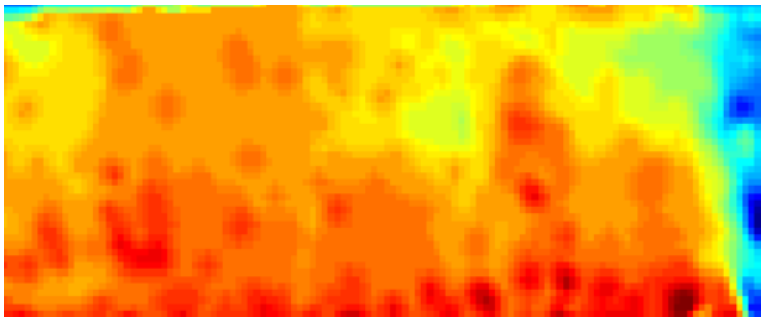
In comparison to the RF elastography example shown in Figure 2.3, we propose an alternative approach to producing strain images, by quantifying dynamic displacement fields from ultrasonographic B-scan sequences. These can then be converted into strain maps, as shown in Figure 2.9, by computing the first derivatives for all displacements. Deriving displacement and strain fields and particularly measuring accuracy are difficult problems, which are addressed in this thesis.



(a) Photograph of tendon.



(b) B-mode of tendon.



(c) Proposed elastogram.

Figure 2.9: An *in vitro* example using computer vision elastography, including (a) a photograph of a tendon section (using a $20\times$ objective lens and stained pink), illustrating the dense regular connective tissue of parallel collagen fibres. Elongated nuclei of fibroblasts can also be observed flattened between the layers of collagen fibres. (b) the ultrasound image and (c) corresponding proposed elastogram.

Chapter 3

Groundtruth *in vitro* and Clinical *in vivo* Datasets

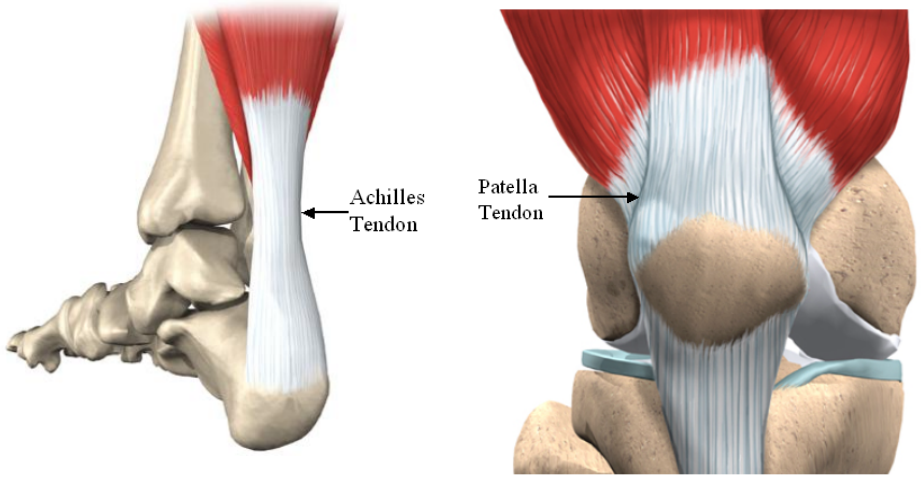
3.1 Introduction

In this thesis we analyse sequences dedicated in purpose and function to musculoskeletal ultrasound. This is an area of high clinical interest due to the demand for improved understanding of tendinopathy resulting from, for example, inflammation, overuse or sport injuries. Hence our motivation for testing datasets of tendon and soft tissue. This chapter is divided into three sections. First, Section 3.2 describes the data of our two main anatomical focus areas. Second, Section 3.3 outlines our tissue mimicking phantom data, tendon pull data and synthetic speckle data. Finally, Section 3.4 concludes by detailing the *in vivo* musculoskeletal datasets.

3.2 Tendon Anatomy: Structure and Function

The first *in vivo* dataset focuses on the calcaneal tendon (or achilles tendon), the thickest and strongest tendon in the human body, shown in Figure 3.1(a). Since Hippocrates first recorded that achilles tendon injury induced acute fevers, choking, derangement of the mind and death [23], achilles tendon disorders have attracted much research. The second dataset focuses on the patella tendon, located at the knee joint, Figure 3.1(b), the largest and most complex joint

in the human body. The patella tendon connects the patella (kneecap) to the tibia (shin bone), part of the extensor mechanism of the knee, which allows the knee to extend (straighten).



(a) Achilles tendon. (b) Patella tendon.

Figure 3.1: Illustrations of the joints: (a) Ankle and (b) knee, both adapted from [34].

Typically tendons are soft (but stiff in tension) collagenous tissues that connect and transmit forces developed by muscles to bones to move and stabilise joints. The junction between the muscle and tendon, myotendinous junction, and the junction between the bone and tendon, osteotendinous junction, are both important clinical areas and provide possible landmarks. Tendons slide over bones and often move through sheaths to direct their path and limit friction. Some long tendons such as the patella or the achilles are not covered by a synovial tendon sheath. Tendon sheaths contain a small amount of fluid, normally seen as a hypoechoic rim on ultrasound images.

Tendons have a hierarchical structure that influences mechanical behaviour. The largest structure in Figure 3.2 is the tendon. The main load carrying element in tendons are collagen molecules that wind themselves together into fibrils; these fibrils are organised into fibres and the fibres into various tissues. The fibril has a crimped (wavy) structure providing a significant mechanical influence. The extensibility of collagen is 1.1%.

A force-extension plot for a typical tendon stretch is shown in Figure 3.3. Most tendons exist in regions *A* and *B* where collagen fibrils start to de-crimp. In region *C* tendon stiffness reduces,

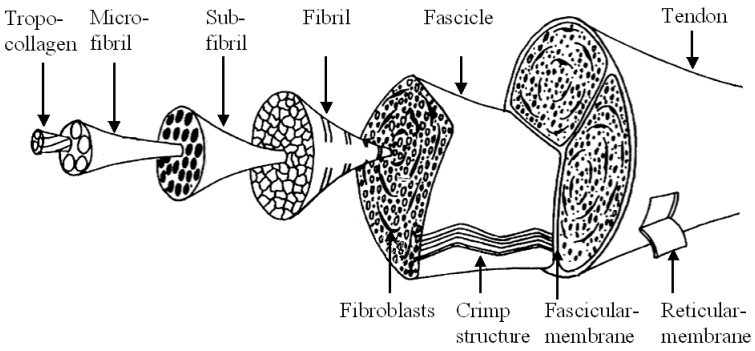


Figure 3.2: Hierarchical structure of tendon, adapted from [44].

fibril damage accumulates and in region *D* tendon failure. Therefore, tendon behaviour depends on the individual crimp structure and failure rate of the collagen fibrils. Tendon collagen fibres break with overuse by repeated $4-8\%$ strain, reported by [43]. It has also been observed that the number of structures visible closely correlates with the ultrasound transducer frequency, and therefore the fascicles are not imagined directly.

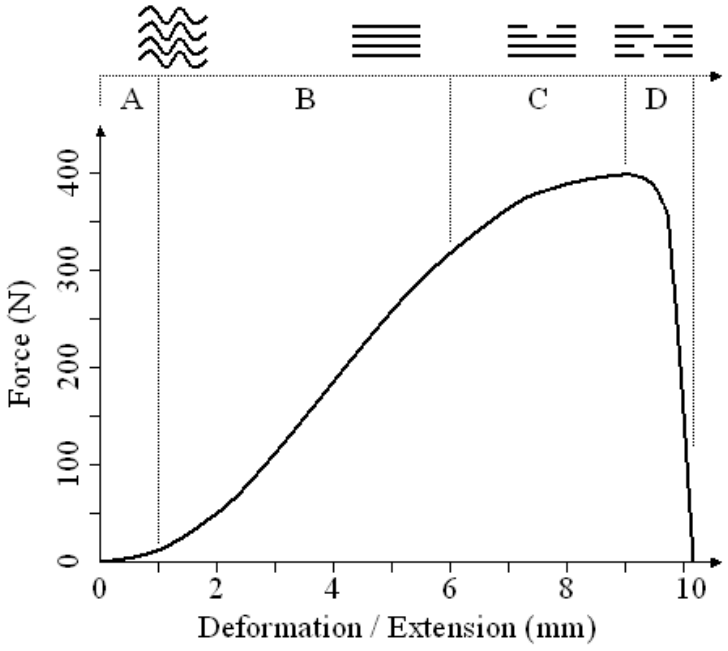


Figure 3.3: Force extension plot for a typical tendon, with sketches of tendon de-crimping, shown at each designated region from A to D.

3.3 Groundtruth *in vitro* Datasets

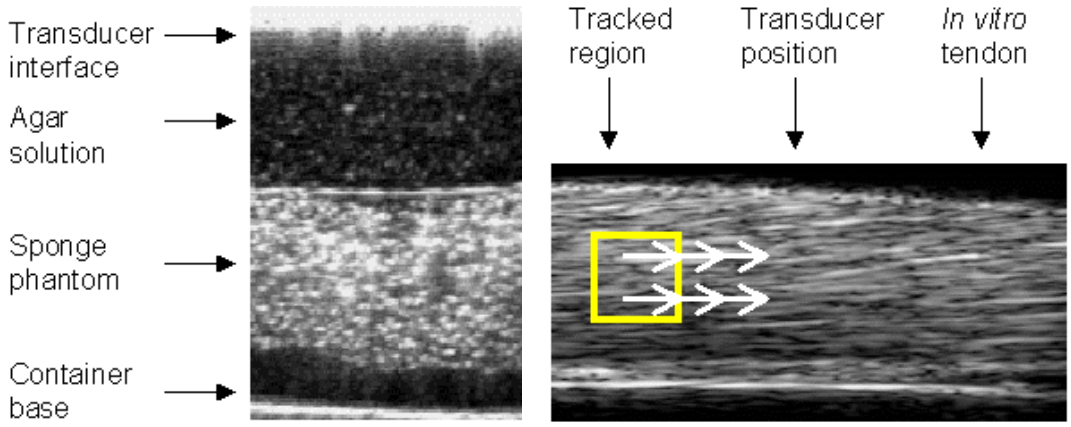
To facilitate performance evaluation of the proposed method for *in vivo* musculoskeletal ultrasound, particularly in regions of tendon and tissue, we setup a tissue mimicking phantom and *in vitro* tendon experiments. These aim to validate the approach in two specific anatomical regions:

1. Tissues, with a high density of scatterers yielding speckle texture.
2. Tendons, differing greatly to tissues with a low density of scatterers and striated structure.

All sequences consist of 30 frames (the default acquisition length using a Diasus Dynamic Imaging ultrasound machine), of 512×512 pixels, captured at 8.6 Hz using a linear array transducer, with a frequency of $8-16$ MHz centred at ≈ 12 MHz (unless otherwise stated) and quantised into 8 bits, requiring 7.75 Mbytes. All sequences were acquired with the ultrasound machine set to operate with a scan depth of 30 mm with a single transmit focus zone. Also, the TGC values were set to obtain equal grey levels at every depth. The resulting captured images have an extremely fine resolution of 50 microns (0.05 mm). Traditional external frame grabbing technology was not used, as most modern machines store displayed frames internally as cine (video). This minimises poor frame rates, aliasing and compression errors, which are associated with previous works, promoting the advantages of computer vision based approaches.

3.3.1 Tissue Mimicking Phantom Data

A tissue mimicking phantom was constructed using a sponge-agarose gel (100%) composite sample (Figure 3.4(a)). The sponge introduced echogenicity to simulate fully developed speckle characteristics (Section 2.2.2), corresponding to the ultrasound of soft tissues. A fixed transducer was positioned on a compression plate (larger than the sample) to ensure uniform global axial deformation. Three experiments using a material testing machine generated sinusoidal ($\sim$), triangular ($\triangle$) and square ($\square$) wave compressive displacements with time, of equal magnitude. Sequences for each deformation from rest, compressing, compressed, decompressing and rest states were continuously captured. We allowed appropriate elapsed time periods between successive tests for the sample to recover to an equilibrium state of being fully re-hydrated from potential fluid loss. The data acquired is summarised in Table 3.1.



(a) Tissue phantom. (b) Tendon *in vitro*.

Figure 3.4: Groundtruths: (a) Tissue mimicking phantom and (b) *in vitro* tendon. ♠

Dataset ID	**Loading Cycle**	**Compression (%)**	**QTY×Frames**
∼	sinusoidal	$0-10$	1×30
△	triangular	$0-10$	1×30
□	square	$0-10$	1×30
TOTAL			**90**

Table 3.1: Summary of the tissue mimicking phantom datasets showing the applied compression.

3.3.2 Tendon Pull Data

In vitro experiments used vertical clamped longitudinal equine tendon specimens of length 100 mm, Figure 3.4(b). Each specimen first underwent several small preconditioning trials and then the applied load, with the loading axis orientated parallel to the fibre direction. A clamped transducer (using an offset gel) scanned continuously in the longitudinal position. Material testing machines (Hounsfield and Zwick Dartec) with increasing load generated several tendon pulls, small and large, to deform the tendon specimen (Figure 3.5). Tendon fibres can be observed with a spatial resolution of lower than 0.1 mm *in vitro*. To enable feature tracking, a landmark was simulated by inserting a 1.6 mm diameter ball-bearing (BB) creating STATIC-LANDMARK1. The data acquired is summarised in Table 3.2.

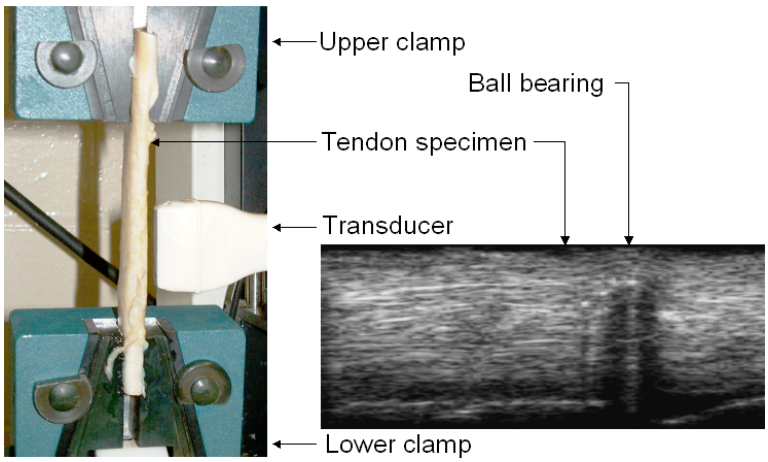


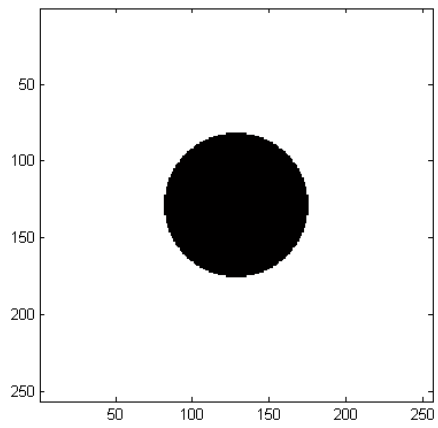
Figure 3.5: Apparatus setup for *in vitro* tendon pulls showing both upper and lower clamps and transducer position alongside the captured B-mode ROI.

Dataset ID	State	Landmark	Clamped	Rate (mm/sec)	QTY×Frames
STATIC-HEALTHY0	healthy	without	single	10	1×30
STATIC-LANDMARK1	healthy	with	single	10	1×30
DYNAMIC-HEALTHY2	healthy	without	double	1	1×30
DYNAMIC-DAMAGED3	damaged	without	double	1	1×30
MUSCLE-HEALTHY4	healthy	without	double	10	2×30
TOTAL					**180**

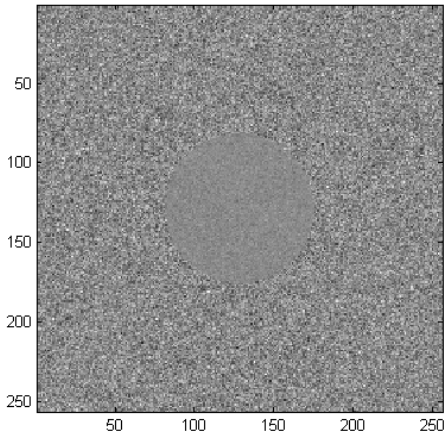
Table 3.2: Summary of the *in vitro* datasets showing clamping, pull rate and quantity.

3.3.3 Synthetic Data

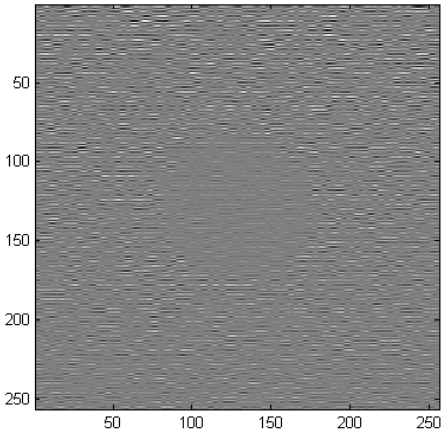
In this work we prefer to use the tissue mimicking phantom and *in vitro* tendon pull data for our gold standard groundtruth, reducing displacement result ambiguity. Much literature analysing ultrasound images use synthetic sequences (generated completely artificially) to test developed algorithms, for example [93]. It is well known that synthetic data rarely synthesises real ultrasound information. However, to first appreciate the formation of speckle, and second to evaluate the advantages of the proposed method in Chapter 5, we also generated spatially uniform and temporally stable speckle textures, as outlined in [62, 42]. The process stages are illustrated in Figures. 3.6(a)-3.6(d), commencing with an initial intensity or echogenicity map



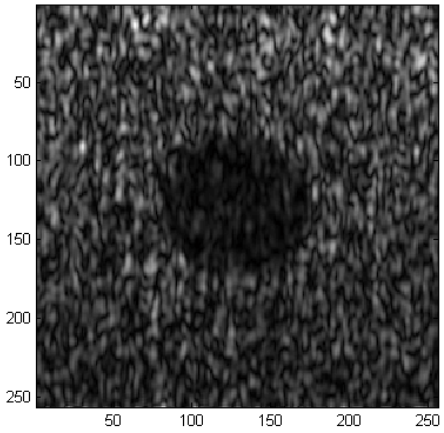
(a) Object intensity or echogenicity map.



(b) Dense scatterers.



(c) RF data.



(d) Simulated speckle image.

Figure 3.6: Datasets: (a) Echogenicity map corresponding to a hypoechoic object. Contrast is -5dB. (b) Scattering function representing scatterers being imaged, weighted by echogenicity map. (c) RF data: PSF convolved with scattering function and (d) Envelope detection using Hilbert transform to produce desired image of echo magnitude (20×20 mm).

in Figure 3.6(a) representing an object of interest, which is a contrasting circular hypoechoic region reproducing [5]. The scattering function $T(x, y)$ is a normally distributed random field, which represents the population of scatterers being imaged, weighted by the echogenicity map, shown in Figure 3.6(b). The point spread function (PSF) $H(x, y)$ is assumed to be a Gaussian-weighted sinusoidal function (a Gabor function, also used by Yu and Acton [113]). Convolving the scattering function with the PSF ($H(x, y) \otimes T(x, y)$) yields the resulting RF echo data

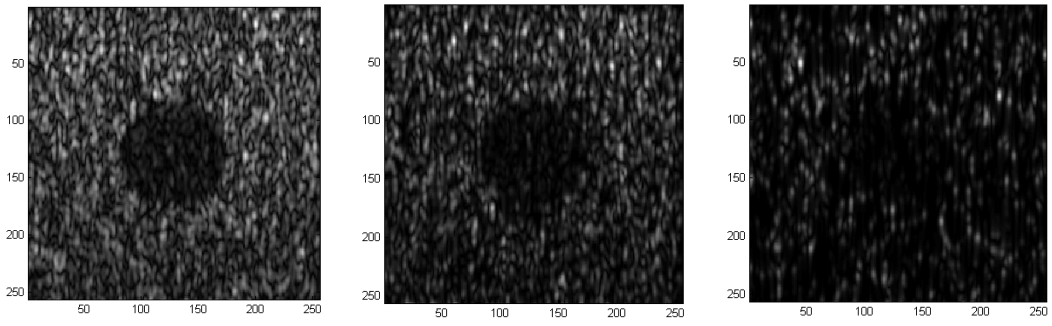
shown in Figure 3.6(c), with envelope detection (magnitude of the Hilbert transform) producing the desired image of echo magnitude in Figure 3.6(d). It can be observed that the speckle has a negative effect on image understanding, obscuring the full details of the circular object.

Two approaches were used to construct various sequences: first, a known motion field was applied to the simulated image I; second, a known motion field was applied to the point scatterer distribution of the tissue T. An affine transformation was used, defined as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} s_x & r_y \\ r_x & s_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{3.1}$$

where intensities at (x, y) are mapped to (x', y') using a translational shift defined by the displacement components $\boldsymbol{d}_t(u, v)$, with s_x and s_y defining the lateral and axial scaling parameters, and r_x and r_y defining the lateral and axial shear parameters. For example, for lateral tension $s_x > 1$ to stretch I, or for lateral compression $0 < s_x < 1$, shrinking I. This transformation typically requires an interpolation scheme; following the interpolation evaluation of Meijering et al. [61] for medical imaging, we use cubic B-spline interpolation.

Different scatterer density causes different observed speckle density, which is shown in Figure 3.7. Therefore speckle density is varied to generate sequences of high and low speckle (for example 100% and 20% respectively), and varying speckle (containing half of each of these). With no underlying structure, there is less reflection from the reduced number of scatterers.



(a) Density $= 100\%$. (b) Density $= 66\%$. (c) Density $= 33\%$.

Figure 3.7: Synthetic datasets: (a) Scatterer density is 100% (full developed speckle), (b) scatterer density is 66% and (c) scatterer density is 33%.

Finally, to measure the robustness of displacement estimation in regions of speckle reduced temporal coherence, we also corrupt each frame with multiplicative Rayleigh distributed noise. Synthetic data is used in Section 5.6 to assess displacement accuracy in varying speckle.

3.4 Clinical *in vivo* Datasets

Our final dataset comes from *in vivo* musculoskeletal ultrasound. We capture and analyse clinically relevant data, demonstrating regions of both tissue and tendon. Tendons are highly anisotropic due to their dense fibrous tissue in which Type 1 collagen fibres are aligned along the tendon's long axis, see Section 2.2.1. While this structure allows tendons to withstand high tensile loads imposed on them during normal activities, it is ill-suited to withstand compressive or shear loading. Although tendons largely function uniaxially, they are multidimensional structures that often undergo multidimensional loading [10]. For example, the tendon-bone insertion sites are structurally complex, and may undergo combinations of tension, torsion, shear, bending, and compression in the course of kinematic joint loading. We have an *in vivo* database including 80 achilles and 80 patella tendon 30-frame sequences, capturing full longitudinal dynamic flexion to extension movements ($< 120°$) from different subjects. A summary is provided in Table 3.3 and labelled example frames appear in Figure 3.8.

Dataset ID	**Freq. (MHz)**	**Subjects**	**QTY$\times$Frames**
Patella	$8-16$	8	80×30
Achilles	$8-16$	9	80×30
Digital Flexor (P2-P3)	$10-22$	5	10×30
Radialis	$10-22$	6	12×30
TOTAL			**5460**

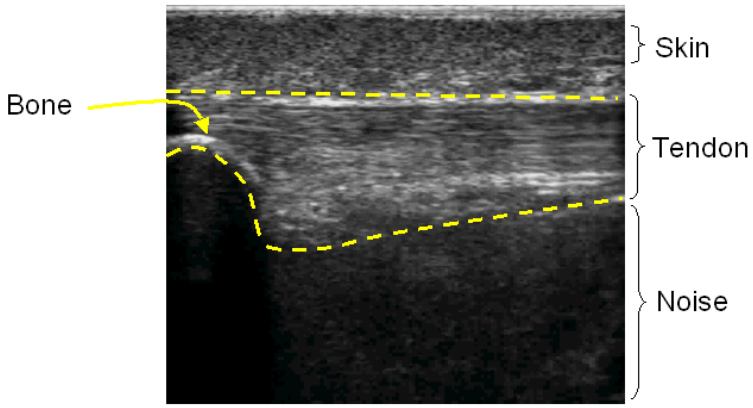
Table 3.3: Summary of the *in vivo* datasets showing tendon type, transducer frequency and quantity.

3.5 Conclusion

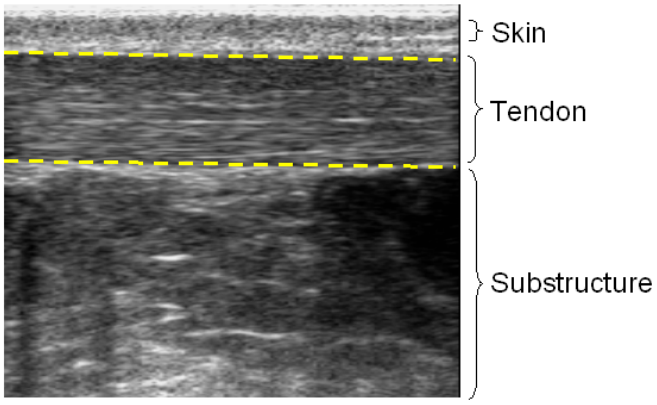
This chapter has described the set of *in vitro* and *in vivo* data we use in this thesis. Due to the challenges of accurate measuring of soft tissue, we generated *in vitro* data focusing on tendon motion, with the tissue mimicking phantom and synthetic data focusing on soft tissue deformation. Furthermore we describe our musculoskeletal *in vivo* data with an emphasis on tendons and surrounding soft tissues.

A major problem for speckle tracking methods is the large variation in captured images. Discrepancies at a high level commonly occur from different machines, user techniques, patients and pathology, with low level discrepancies including machine built-in image processing and filtering software.

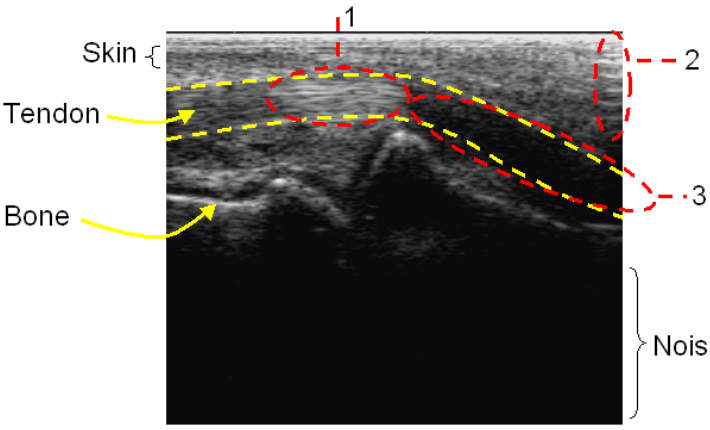
It should be noted that the groundtruth *in vitro* and clinical *in vivo* datasets comply with data protection acts, with subject details removed where applicable. This collection of data will be made publicly available on-line, to provide the first (that we are aware of) collection of ultrasound sequences and displacement results for other researchers to use. All experiments in this thesis use the datasets described here for testing.



(a) Patella tendon.



(b) Achilles tendon.



(c) Digital flexor tendon.

Figure 3.8: Labelled ultrasound frames: (a) Longitudinal section of the patella, (b) achilles and (c) superficial digital flexor tendons (labels 1, 2 and 3 identify enhancement, reverberation and shadow artefacts respectively). ♠

Chapter 4

Interframe Motion Estimation for Displacement Fields

4.1 Introduction

This chapter presents a block matching approach to measuring 2D soft tissue deformation, and assesses the feasibility of tracking regions in musculoskeletal ultrasound sequences. While block matching has been used extensively in the computer vision field, we demonstrate a novel application of the concept to musculoskeletal ultrasound. An iterative variable-size block matching (VSBM) framework is implemented to combine the advantages of using a range of block sizes. This enables us to quantify varying scales of motion, which exist in typical *in vivo* ultrasound image sequences. A hierarchical full search strategy is employed, yielding optimal convergence for all search iterations. The resulting estimated displacements are also refined to sub-pixel accuracy using bi-linear interpolation. To evaluate the performance of VSBM on our ultrasound sequences, we performed a comparative analysis consisting of several relevant approaches, investigating variations of block matching and a differential technique. Our comparisons are primarily focused on internal tendon deformation, hence we concentrate on the accuracy and reliability of displacement estimation using specific *in vitro* groundtruth data. Our study found VSBM produced superior performance due to analysing blocks of varying scales, indicating improvements were made in de-correlated regions, and where motion disparities occurred. In addition, experimental results are shown for *in vivo* clinical sequences.

We begin in Section 4.2 with a fundamental *in vitro* experiment to observe the temporal changes in image content, by measuring interframe stability under static conditions. Next, we describe in Section 4.3 the proposed block matching approach including similarity and dissimilarity matching criteria, block size definition, searching strategy and refinement of displacements to sub-pixel accuracy. Further, displacement estimation algorithms are described in Section 4.4, with accuracy measurements defined in Section 4.5, prior to a comparative analysis detailed in Section 4.6. Displacement field results from *in vivo* sequences are reported in Section 4.7, and conclusions drawn in Section 4.8 on estimating displacements using ultrasound images.

4.2 B-scan Change Detection Considerations

Before commencing with motion estimation, we first conducted a change detection experiment using image differencing [83]. The purpose was to measure the temporal coherence of the observed speckle textures, and existence of any non-stationary noise. By calculating the absolute values of the differences between the corresponding pixels in two images a global difference map was produced. This quantifies the number of corrupted pixels that have a non-zero value, from noise or temporally unstable speckle, assuming a still target object and static transducer to ensure captured frames were automatically registered. Different frame intervals between consecutive and larger periods were also analysed to observe the amount of temporal change.

Different transducer frequencies were compared with bandwidths 5–10, 8–16 and 10–22 MHz. Each was used to capture "perfect" *in vitro* conditions of a stationary tendon in a still water bath with clamped probe to quantify different pixels, and "less than perfect" *in vivo* conditions of a freehand scanning of muscle to locate different pixels. Although substantial research exists using lower frequencies (for example 3–7 MHz for abdominal [20], cardiac and mammography [13]), we focus on the higher frequencies $8-16$ MHz used for musculoskeletal diagnosis. The higher transducer frequencies and wider aperture size[1], produce higher axial and lateral resolutions[2], but a reduced penetration depth. The change in depth is due to attenuation where the signal can reduce by approximately 1 dB/cm/MHz [95].

Results from the perfect conditions experiment are shown in Table 4.1. All tests showed that

[1]The aperture size is the active transmitting and receiving portion of the transducer array.

[2]The resolution in an ultrasound image is a function of the F-number, or size of the aperture.

the transducer frequency had no observable influence on the quantity of pixel variations, and as expected, only a maximum of $\approx 8\%$ of pixels changed. These changes are attributed to the presence of some negligible electronic additive Gaussian noise, rather than unstable speckle. However, sequences captured with less than perfect conditions, highlighted a reduction in correlation (indicated by a large pixel difference) in areas of speckle, tending to occur in the lower regions, shown earlier in Figure 3.8 and here in Figure 4.1. This variation in signal content has an effect on displacement accuracy, which is dealt with later in Chapter 5.

Frame	**Frequency (MHz)**		
Interval	**5-10**	**8-16**	**10-22**
$f_1 \rightarrow f_2$	6.49	4.24	5.43
$f_1 \rightarrow f_{10}$	7.66	4.92	5.98
$f_1 \rightarrow f_{20}$	7.71	5.12	6.02
$f_1 \rightarrow f_{30}$	7.53	6.19	6.15

Table 4.1: Percentage pixel change detection measurements in “perfect” conditions.

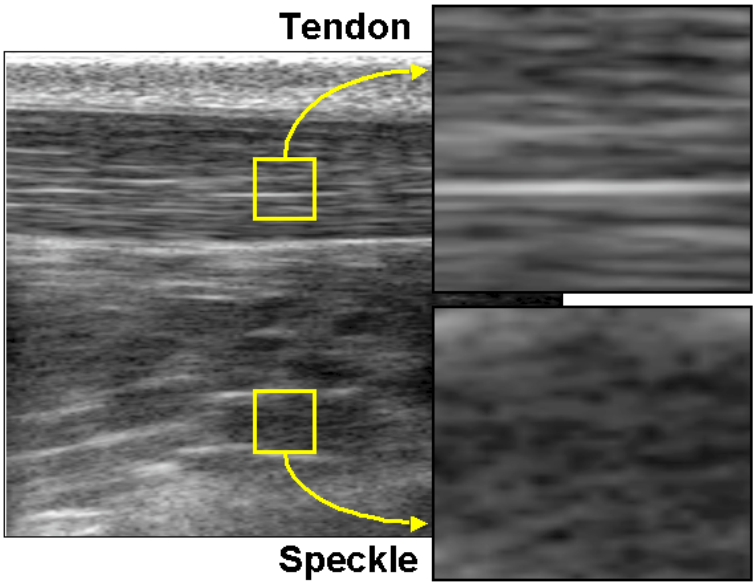


Figure 4.1: Achilles tendon with enlarged images of tendon and speckle.

4.3 Variable Sized Block Matching (VSBM) for Speckle Tracking

The purpose of block matching is to locate the best match for a block $I(i,j)$ in frame f_t, somewhere in frame f_{t+1} given numerous blocks of exactly the same dimensions. We now describe several matching criteria before detailing the criterion used in the VSBM approach, followed by our hierarchical full search strategy and sub-pixel displacement refinement.

4.3.1 Similarity and Dissimilarity Matching Criteria

Various matching criteria are used for locating textures in images, which either measure the similarity or dissimilarity between localised regions. The most popular criterion that has been implemented in hardware, due to having no square operations, is the minimum absolute difference (MAD), which is defined as:

$$\text{MAD}\left(I_{(i,j)}, I'_{(k,l)}\right) = |I[i+x, j+y] - I'[k+x, l+y]| \tag{4.1}$$

Unfortunately, the performance of the MAD criterion deteriorates as the search area becomes larger due to the presence of several possible local minima. It can, however, be incorporated into a matching pixel count (MPC) criterion, which classifies each pixel within each block as either matching or mismatching, where the block yielding the largest MPC produces the optimum displacement $\boldsymbol{d}$, according to:

$$\text{MPC} = \begin{cases} 1 & \text{if MAD} \leq \text{threshold} \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

depending on a predetermined threshold. Another alternative is the sum of square difference (SSD) criterion, defined as:

$$\text{SSD}\left(I_{(i,j)}, I'_{(k,l)}\right) = \sum_{x=1}^{M}\sum_{y=1}^{N}\left(I[i+x, j+y] - I'[k+x, l+y]\right)^2 \tag{4.3}$$

The SSD is not robust to any shift or scaling of intensity. To compensate for shifts in intensity from uneven illumination, the intensity shift-corrected SSD can be used. This subtracts the local mean from each block prior to performing the SSD. Instead of using the mean value, Strintzis and Kokkinidis [91] introduced a weighted term corresponding to the statistical description of ultrasound images. This is discussed further in Chapter 5.

Briefly we mention here the use of the spatial domain normalised cross correlation (NCC), which correlates 2 regions to provide a similarity measure, with local means $\bar{I}$ and $\bar{I}'$ removed. We expand on this later, but the NCC is used in our comparative analysis in Section 4.4.2, and is defined as:

$$\text{NCC}\left(I_{(i,j)}, I'_{(k,l)}\right) = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I[i+x, j+y] - \bar{I})(I'[k+x, l+y] - \bar{I}')}{\sqrt{\sum_{x=1}^{M}\sum_{y=1}^{N}(I[i+x, j+y] - \bar{I})^2 \sum_{x=1}^{M}\sum_{y=1}^{N}(I'[k+x, l+y] - \bar{I}')^2}} \quad (4.4)$$

In this chapter, we quantify the matching of blocks using the mean square error (MSE), following the real-time work of Bohs et al. [12]. We can evaluate the MSE for a pair of sub-images, defined as:

$$\text{MSE}\left(I_{(i,j)}, I'_{(k,l)}\right) = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}\left(I[i+x, j+y] - I'[k+x, l+y]\right)^2 \quad (4.5)$$

where I and I' are regions of size $M \times N$ from f_t and f_{t+1} respectively, and (i, j) are the top left corner co-ordinates of the reference block in f_t as are (k, l) for the candidate blocks in f_{t+1}. Neighbouring pixels give the central pixel and corresponding displacement vector unbiased support. The resulting MSE $\left(I_{(i,j)}, I'_{(k,l)}\right)$ produces a candidate motion vector $\boldsymbol{d}\,(u, v)$ as illustrated earlier in Figure 2.5. The estimate $\hat{\boldsymbol{d}}$ of the motion vector $\boldsymbol{d}$ is taken to be the value of (u, v) that minimises the MSE. The displacement estimate is given by:

$$\hat{\boldsymbol{d}}(u, v) = \arg\min_{(I,\ I')} \text{MSE}\left(I,\ I'\right) \quad (4.6)$$

Minimising the MSE imposes an optical flow constraint on all pixels in a block. The MSE provides a good comparison of computational complexity versus quality.

4.3.2 Hierarchical Search and Variable Block Size Strategy

Improvements in displacement estimation in areas of de-correlation are gained by using a coarse to fine search with varying block sizes. This is due to the fine-tuning of the displacement vector, as the estimate of the vector at the largest scale is passed to the next scale. Each scale encapsulates differing speckle texture, and at the smaller scales, a relatively smaller search region is used as the initial estimate is robust. In some cases where a region is extremely

de-correlated or homogeneous (Section 2.5.3), no displacement can be estimated with any accuracy. De-correlation is typically caused by signal saturation, lack of signal, or large strains.

The hierarchical full search initially determines the optimal matching block in f_{t+1}, by computing (4.5) for all possible candidate displacement vectors at each pixel $I'[k+x, l+y]$ in a search region. To reduce the computational burden and improve displacement accuracy, we define a search region centred about each pixel for which a motion vector $\boldsymbol{d}$ is estimated. The first search iteration initialises with a block size $M \times N$ and a search region of $-q_1 \leq u \leq q_1$ and $-q_2 \leq v \leq q_2$, where q, M and N are determined empirically, but are set at a relatively large size, (Figure 4.2). The block producing the lowest MSE in frame f_{t+1} to the candidate block from the reference frame f_t is then used to initialise the next iteration, where *both the block size and search region* reduce logarithmically. The process of scaling the search region and block dimensions, and using the block producing the lowest MSE as the initialisation for the next iteration, repeats until a block size limit is reached. For each iteration we simultaneously update the reference block from f_t, by halving its dimension, refreshing the MSE matching criterion (shown in Figure 4.2). Every possible block within the search range in f_{t+1} is examined, with block overlapping to improve flow continuity and smoothness. Prior research has involved optimising the search strategy using various search region sampling techniques, for example the adaptive search of [17]. Although optimised searching decreases computation time, they cannot guarantee optimal convergence.

To illustrate our approach we initially present a cropped B-scan from our groundtruth data STATIC-LANDMARK1 under simple translational motion. Block sizes $M \times N$, where $M, N = \{64, 32, 16, 8, 4\}$ were used, with the final iteration matching the candidate superimposed blocks on f_t in Figure 4.3(a). Each iteration is visualised in Figure 4.3(b). Progressively increasing displacements are visible as the block size reduces, with $M \times N = 64$ showing no displacements in the landmark region, to $M \times N = 4$ displaying the final refined displacements. Further, *in vitro* estimated displacement accuracy and displacement field results are shown in Section 4.6, and *in vivo* results in Section 4.7.

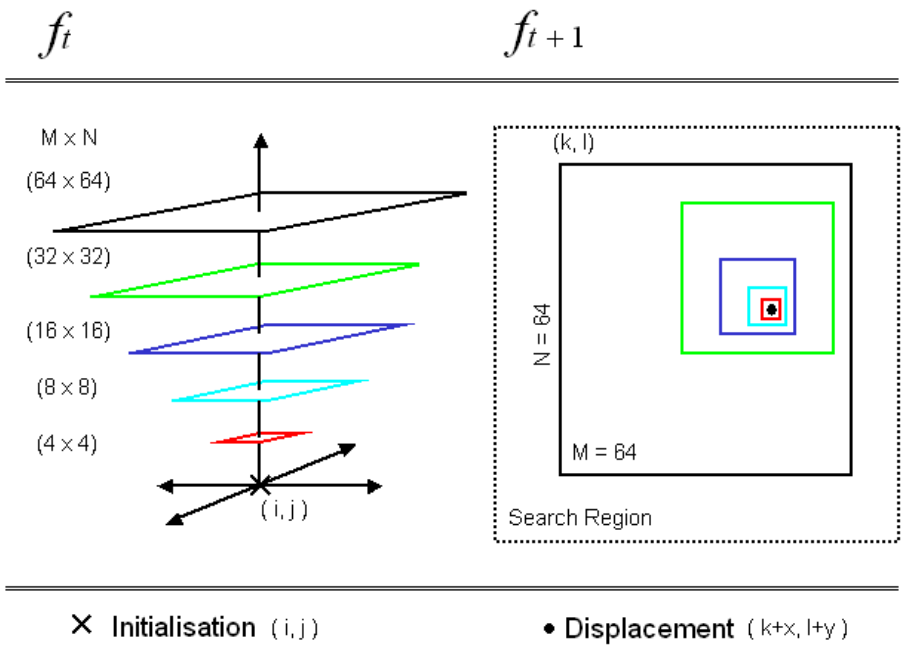


Figure 4.2: Hierarchical full search strategy using VSBM illustrating the reference block in f_t and final displacements in f_{t+1}.

4.3.3 Sub-pixel Displacement Refinement

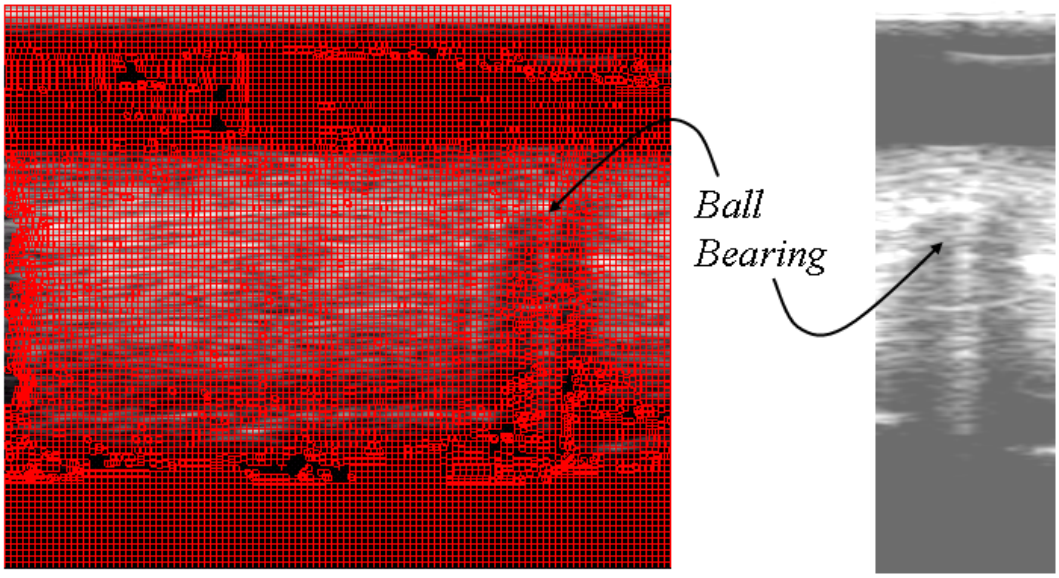
Due to motion not occurring in discrete increments, the VSBM algorithm was extended to perform measurements to sub-pixel accuracy. We computed $\frac{1}{4}$ pixel granularity expanding a $M \times N$ search region to $4\,(M \times N)$ in both f_t and f_{t+1} using bi-linear interpolation. The newly generated pixel P at (x, y) is a weighted sum of the four nearest neighbourhood pixels: $P_1(x_1, y_1)$, $P_2(x_1, y_2)$, $P_3(x_2, y_1)$, $P_4(x_2, y_2)$, so that:

$$P = (1-a)\,(1-b)\,P_1 + a\,(1-b)\,P_2 + abP_3 + (1-a)\,bP_4 \tag{4.7}$$

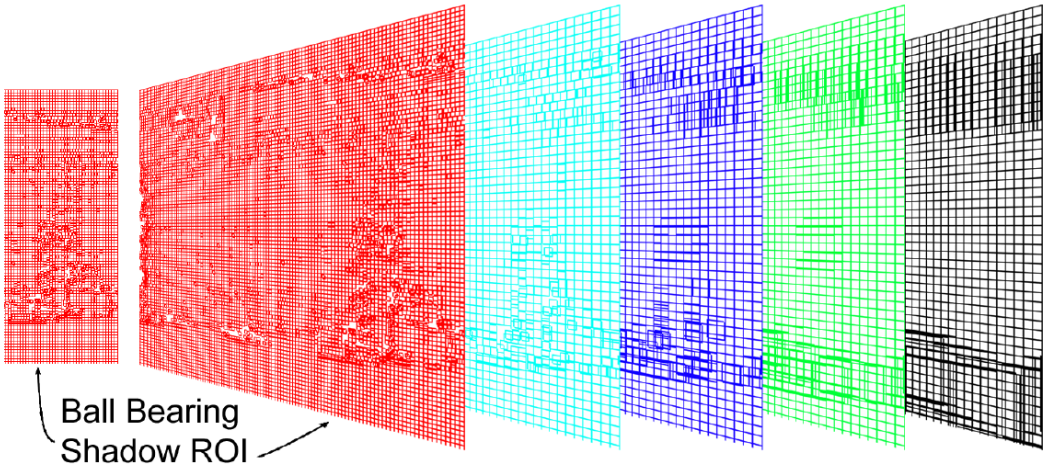
where a and b vary between 0 and 1, defined as:

$$a = \frac{x - x_1}{x_2 - x_1} \qquad b = \frac{y - y_1}{y_2 - y_1} \tag{4.8}$$

Although bi-linear interpolation alters pixel data, reduces contrast and the high frequency components of the image, it is commonly used for image enlargement as image quality is retained.



(a) B-scan f_t from STATIC-LANDMARK1 data set superimposed with candidate blocks.



(b) Hierarchical VSBM in f_t.

Figure 4.3: Multiple block scales: (a) Cropped image of STATIC-LANDMARK1 groundtruth data (ball-bearing landmark located centre right), (b) hierarchical VSBM for 5 levels highlighted using sizes $M \times N$, where $M, N = \{64, 32, 16, 8, 4\}$ (from right to left).

4.4 Comparison of Displacement Estimation Algorithms

This section provides a comparison of the performance of several displacement estimation algorithms to VSBM. Our goal is to determine technique suitability for measuring tendon internal motion. Although many optical flow algorithms perform well for certain sequences few are used for ultrasound. The selection of algorithms used in this comparison was motivated by previous inconsistent and unreliable results from applying the algorithms available[3] and reviewed by Barron et al. [9]. Also, the desire to examine different forms of block matching that have already established success for others (mentioned in Chapter 2). Thus, in the following sections two block based approaches and a robust differential approach are briefly described. These include a fixed size block matching framework, the temporally extended pixel matching framework of Camus [14] and the differential flow algorithm of Proesmans et al. [80][4].

4.4.1 Fixed Size Block Matching (FSBM)

Block matching methods require that the block should be large enough to estimate large displacements, but also small enough so that the displacement remains constant within the region. These two requirements are only solved using multiple block sizes that are defined in VSBM, contrary to the FSBM approach, which is equivalent to only a single scale of VSBM. In this experiment both FSBM and VSBM use the MSE matching criterion, with the FSBM limited to a block size of $M, N = 8$, and VSBM limited to 2 block sizes of $M, N = \{16, 8\}$.

4.4.2 Camus [14] Temporal Pixel Matching

Camus [14] defined a variation on block matching flow algorithms, by utilising the definition of velocity as the change in distance over the change in time. Typical region matching algorithms, for example FSBM and VSBM apply a temporal search that analyses frame pairs (f_t, f_{t+1}) for a constant time interval of $t = 1$, and define the estimated displacement magnitude as the extent of the matching criterion search area. However, Camus proposed that the search be extended in time (f_t, f_{t+n}) for a varying time interval of n frames, but used a constant

[3] Sources available at *ftp://ftp.csd.uwo.ca/pub/vision*

[4] Sources available at *http://www.cs.otago.ac.nz/research/vision/Research/OpticalFlow/opticalflow.html*

reduced search area limited to single pixel shifts. This is in contrast to VSBM, which proposes a constant temporal search but varying increased spatial search. A desirable property of Camus for ultrasound analysis is real-time performance. To operate in real-time, however, typically only a single pixel shift can be measured in a region $M \times N$ where $M, N = 3$. However, for large motions the spatial search region can be extended or images sub-sampled. We apply the former, increasing the search region, to measure the groundtruth 8 pixel shifts at the expense of increasing processing time.

The Camus technique is illustrated in Figure 4.4, demonstrating a 1D temporal search of 2 frames (f_{t+1}, f_{t+2}), and a 2D spatial 3×3 search in f_t. In this example, the displacement for a pixel at $f_t(2,3)$ shifts $(0,1)$ pixels over 2 frames, or a displacement $\boldsymbol{d}_{t=1} = (0, 0.5)$ per frame. From experimentation, we use 2 frames of temporal support and pixel cross correlation defined in (4.4), instead of the MSE that produced unusable results.

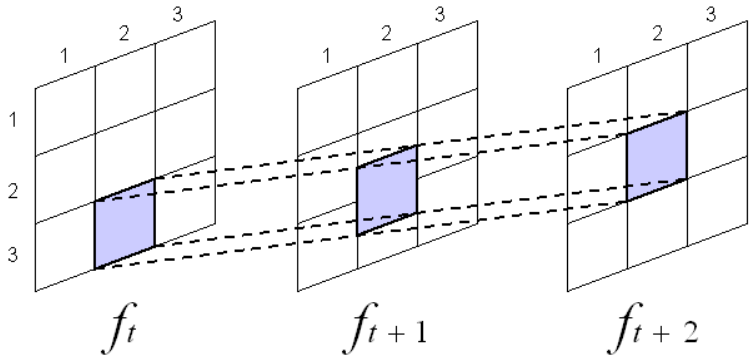


Figure 4.4: A pixel in f_t at $(2,3)$ moves to $(2,2)$ in f_{t+2} producing $\boldsymbol{d}_{t=1}(0, 0.5)$.

4.4.3 Proesmans et al. Differential Approach

The final method, Proesmans et al. [80], is a similar differential technique to Horn and Schunck [38]. However, Proesmans et al. incorporate a matching process into (2.4), and define a method for dealing with discontinuous displacements. As explained earlier, for each point in the initial frame $f_t(x, y)$, the estimated displacement $\boldsymbol{d}_t(u, v)$ is used to locate a corresponding point in the next frame $f_{t+1}(x + u, y + v)$. Proesmans et al. state that if these points have similar brightness, the resulting flow estimate is good, else the flow estimate is moved along the image

gradients, in order to correct for the difference. Further, Proesmans et al. use displacement post-processing to increase neighbouring displacement similarity, and reduce expected discontinuous displacements at object edges. This is achieved by producing consistency maps computed from the magnitude of forward (f_t, f_{t+1}) and backward (f_{t+1}, f_t) displacement fields. The two displacement fields are usually equal but opposite, however, at object boundaries this is not the case, with consistency maps producing high values of inconsistency at object boundaries. The application of anisotropic diffusion is used to smooth displacements, where the diffusion function uses the consistency map to control smoothing instead of image gradient, which differs to the image filtering described in Section 2.5.3. Similarly to Proesmans et al. we set parameters empirically using a smoothness constraint of $\lambda = 30$ using 40 iterations. Galvin et al. [30] observed this technique produced consistently good displacement accuracy and tolerance to noise.

4.5 Interframe Displacement Field Accuracy

In this section we describe the velocity angular error and image reconstruction error, which are used to determine the accuracy of the displacement field $\boldsymbol{d}_t$ produced from each technique. For our *in vitro* sequences an actual displacement field is known, hence both measures are used, except for *in vivo* data using only the image reconstruction error. All error measurements use the full displacement field density, with no outliers removed from thresholding.

4.5.1 Velocity Angular Error

To quantify displacement accuracy the error between the correct groundtruth displacement $\boldsymbol{d}_{t,c} = (u_c, v_c)$, and the estimated displacement $\boldsymbol{d}_{t,e} = (u_e, v_e)$ is measured by the angular error [9], combining errors in magnitude and direction into a single value:

$$\psi = \cos^{-1}(\breve{\boldsymbol{d}}_c \cdot \breve{\boldsymbol{d}}_e) \tag{4.9}$$

where ψ is the angle between the correct spatiotemporal vector $\breve{\boldsymbol{d}}_c = \frac{(u_c, v_c, 1)^T}{\sqrt{u_c^2 + v_c^2 + 1}}$ and, the estimated spatiotemporal vector $\breve{\boldsymbol{d}}_e = \frac{(u_e, v_e, 1)^T}{\sqrt{u_e^2 + v_e^2 + 1}}$.

4.5.2 Image Reconstruction Error

Further, interframe displacement fields are used for displaced frame differencing (DFD), quantifying interframe pixel error. DFD is intensity dependent and provides a measure of displacement quality. If we assume translational motion, then an estimated displacement $\boldsymbol{d}_t(u_e, v_e)$ defines the change in pixel position from $\boldsymbol{x} = (x, y)$ to $\boldsymbol{x'} = (x', y')$ from frame f_t to f_{t+1}: $\boldsymbol{x'} = \boldsymbol{x} + \boldsymbol{d}_t(u_e, v_e)$. Assuming intensity constancy for small (≈ 8 Hz) interframe pairs due to stable scatterers, the displacement field error can be measured using backward warping:

$$e_t(\boldsymbol{x}) = f_t(\boldsymbol{x}) - f_{t+1}(\boldsymbol{x'} - \boldsymbol{d}_t(u_e, v_e)) \tag{4.10}$$

To determine the interframe global error for frame pairs or regions ($M \times N$), we compute:

$$\text{MSE} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} e_t(\boldsymbol{x}(x, y))^2 \tag{4.11}$$

To generate the reconstructed frame at any time in the sequence, blocks are positioned using the corresponding displacements. Pixels with multiple values due to overlapping blocks are averaged, and pixels with no value due to untracked regions use the value up to that frame. Although minimal for frame pairs, blocking artefacts can cause artificially high DFD errors.

4.6 Displacement Comparison Analysis Results

Now we compare the algorithms: FSBM, VSBM, Camus and Proesmans et al.. The purpose is to justify that block based methods are indeed a suitable approach for our data, and to demonstrate the improvements of using variable block sizes.

The comparison uses frames (f_1, f_2) to (f_9, f_{10}) of the STATIC-LANDMARK1 groundtruth sequence. Each frame pair captures the same 8 pixel lateral translational shifts, hence we expect resulting interframe measurements to produce similar results. By analysing these frame pairs we remove displacement accuracy result ambiguity. It is important to note that the STATIC-HEALTHY0 and STATIC-LANDMARK1 groundtruths (different only in BB insertion), are equivalent to a moving transducer capturing a static object. Thus, with no load and only translational motion, the tendon remains static and only shifts in time. Therefore, we would expect good displacement accuracy from each method. Finally, these errors assume that there were no

changes in illumination resulting from transducer gain changes, nor tendon slippage between the clamps, nor small discrepancies from machine accuracy.

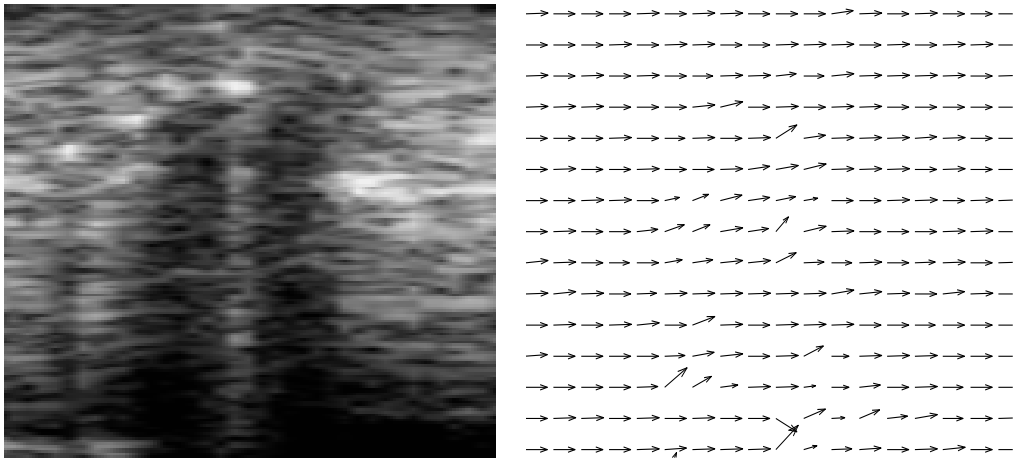
Table 4.2 quantifies the image reconstruction errors, which measure displacement field accuracy for each technique. Before displacement estimation, the pre-warped frame difference (FD) and the NCC for each frame pair was computed. The worst performing, but also the fastest (by far), was the Camus algorithm. Although Liu et al. [51] reported Camus to have one of the best accuracy-efficiency ratios, we observed the efficiency dramatically reduce on our data. We found the main reason for this was that the pixel search uses only a single scale. The next most poor result was Proesmans et al., confirming that differential algorithms suffer on ultrasound data, even with displacement post-processing (this was not performed by the other algorithms). Another source of displacement inaccuracy, prominent with Proesmans et al., were errors induced by the aperture problem, caused by the extremely striated structure of the tendon as mentioned before in Section 3.2. The best results were produced by VSBM for this data (even better results are reported in Chapter 5), and highlighted the benefits of using just 2 block sizes when compared to the errors from FSBM. Erroneous displacements from not optimising the tracked region size were also observed by Chunke et al. [18].

Frame	**PRE-WARP**		**FSBM (MSE)**		**VSBM (MSE)**		**Camus (NCC)**		**Proesmans et al.**	
Pair	**FD**	**CORR%**	**DFD**	**CORR%**	**DFD**	**CORR%**	**DFD**	**CORR%**	**DFD**	**CORR%**
$f_1 \rightarrow f_2$	273.46	97.40	84.78	99.19	**55.33**	**99.47**	214.74	97.87	106.63	98.99
$f_2 \rightarrow f_3$	270.28	97.44	80.12	99.24	**52.98**	**99.50**	215.95	97.96	108.34	98.97
$f_3 \rightarrow f_4$	272.57	97.41	81.79	99.22	**51.12**	**99.51**	214.67	97.91	111.61	98.94
$f_4 \rightarrow f_5$	273.87	97.40	81.98	99.21	**53.70**	**99.49**	216.30	97.96	119.63	98.86
$f_5 \rightarrow f_6$	271.60	97.43	82.07	99.22	**52.13**	**99.51**	219.39	97.93	115.43	98.92
$f_6 \rightarrow f_7$	273.82	97.40	83.32	94.54	**51.37**	**99.51**	217.34	97.94	112.37	98.93
$f_7 \rightarrow f_8$	273.87	97.40	84.54	99.21	**53.63**	**99.49**	223.30	97.89	115.32	98.90
$f_8 \rightarrow f_9$	277.87	97.40	86.68	99.18	**57.60**	**99.45**	219.85	97.92	121.90	98.84
$f_9 \rightarrow f_{10}$	277.84	97.35	84.23	99.24	**59.29**	**99.43**	218.98	97.86	119.05	98.86

Table 4.2: Image reconstruction error measures for FSBM (MSE), VSBM (MSE), Camus (NCC) and Proesmans et al. using STATIC-LANDMARK1.

Figure 4.5(a) is an enlarged region of the tendon with the BB landmark from our groundtruth dataset STATIC-LANDMARK1. The corresponding interframe displacement field is illus-

trated in Figure 4.5(b). This shows the dense continuous linearity of displacements in the lateral axis of motion from left to right, and some expected discontinuities of marginal magnitude located at the BB region from specular reflections. The VSBM and in particular the FSBM, showed several displacement outliers in this region, dissimilarly to the displacement post-processing of Proesmans et al, which smoothed them.



(a) Enlarged ROI of B-scan. (b) Enlarged displacement field.

Figure 4.5: Groundtruth analysis: (a) BB landmark in tendon sample, (b) VSBM displacement field using block sizes $M \times N$, where $M, N = \{32, 16, 8\}$.

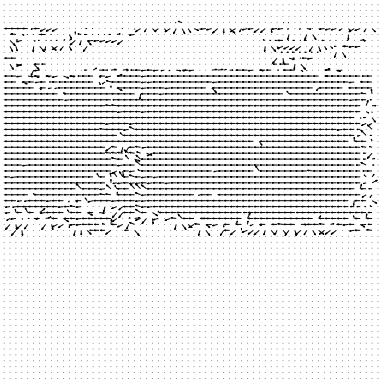
Table 4.3 shows the velocity angular error, which quantifies displacement field accuracy for each technique, using the same frames as above. This measure requires both the actual known and estimated displacement fields, and is dependent on the quality of the actual known data.

In this work, we found the image reconstruction error was overall more reliable, due to not needing the actual known displacements, where discrepancies can occur. The velocity angular error results confirm that Camus and Proesmans et al. produced the worst results (for example mean values of 50.34 and 30.37 respectively for f_5 to f_6), and that the most accurate method was VSBM (for example a mean value of 11.77 for f_5 to f_6).

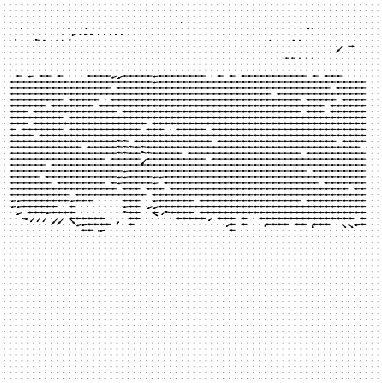
Figure 4.6 shows example displacement fields for the frame pair (f_1, f_2) from each technique using the STATIC-LANDMARK1 groundtruth sequence. It is clear that Proesmans et al. post-displacement processing over-smoothed the displacement fields, and is a contributory factor to its poor results. Using a reduced smoothness constraint and less iterations, however, we observed that displacements became discontinuous and noisy.

Frame Pair	FSBM (MSE)		VSBM (MSE)		Camus (NCC)		Proesmans et al.	
	Mean°	STD Dev°	Mean°	STD Dev°	Mean°	STD Dev°	Mean°	STD Dev°
$f_1 \rightarrow f_2$	14.94	35.45	**9.53**	**29.13**	50.61	26.55	31.64	14.52
$f_2 \rightarrow f_3$	14.65	35.14	**9.61**	**29.29**	50.95	26.17	30.67	14.79
$f_3 \rightarrow f_4$	14.77	35.27	**10.16**	**30.12**	50.57	26.63	30.86	14.57
$f_4 \rightarrow f_5$	16.15	35.53	**9.98**	**29.17**	51.53	26.66	30.16	14.83
$f_5 \rightarrow f_6$	15.22	35.53	**11.77**	**32.08**	50.34	26.69	30.37	14.54
$f_6 \rightarrow f_7$	15.36	35.58	**10.30**	**31.23**	50.57	26.57	30.90	14.48
$f_7 \rightarrow f_8$	15.63	36.04	**10.24**	**31.02**	51.10	26.65	31.74	14.60
$f_8 \rightarrow f_9$	15.69	36.07	**11.73**	**31.39**	51.55	26.57	30.54	15.03
$f_9 \rightarrow f_{10}$	16.16	36.55	**10.09**	**29.13**	51.01	26.59	33.93	16.16

Table 4.3: Velocity angular error measures for FSBM (MSE), VSBM (MSE), Camus (NCC) and Proesmans et al. using STATIC-LANDMARK1.

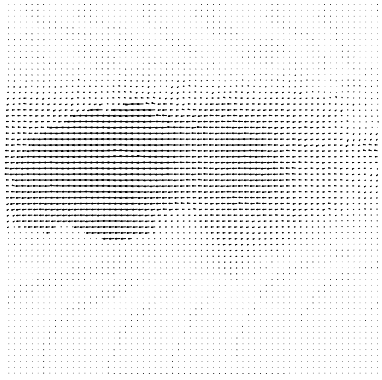


(a) FSBM displacement field.



(b) VSBM displacement field.

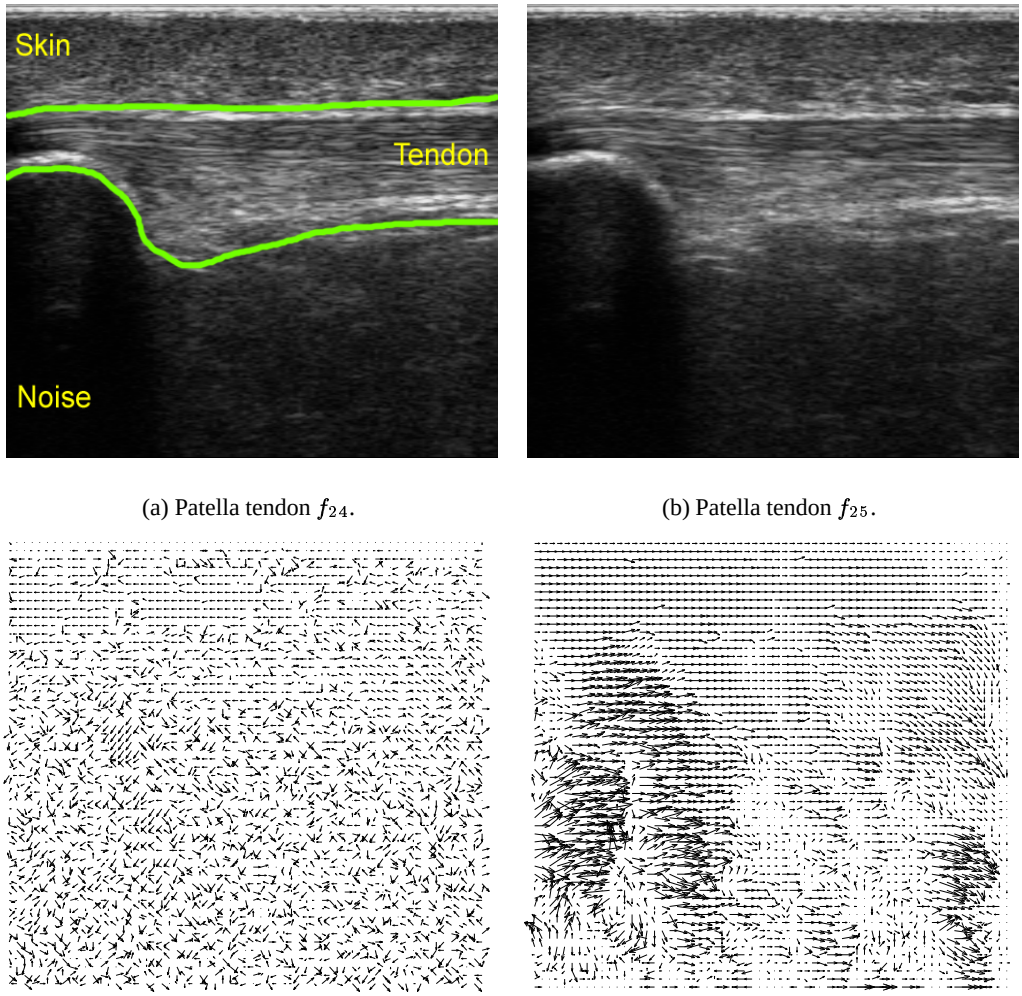
(c) Camus displacement field.



(d) Proesmans et al. displacement field.

Figure 4.6: FSBM, VSBM, Camus and Proesmans et al. for f_1 to f_2 using STATIC-LANDMARK1.

For a purely subjective comparison of VSBM and FSBM, we calculated the interframe displacement field using *in vivo* data. The reason was to appreciate the obvious *visual* improvement of displacements from using multiple block sizes, which becomes an area of further work in Chapter 5. Examples are shown in Figures 4.7(a)-4.7(b) of a patella tendon under extension to flexion motion. The FSBM results in Figure 4.7(c) do not show any convincing output, and are equivalent to the VSBM using only a single block size. The VSBM results in Figure 4.7(d) using multiple block sizes, however, show coherent groups of displacements in the top and centre of the image, corresponding to the motion of the tendon. The noisy regions lower down correspond to the motion of well formed backscatter image noise.



(a) Patella tendon f_{24}.

(b) Patella tendon f_{25}.

(c) FSBM displacement field.

(d) VSBM displacement field.

Figure 4.7: Qualitative *in vivo* comparison: (a-b) B-scans frame pair, with results (c) FSBM ($M \times N = 8$) and (d) VSBM ($M \times N$, where $M, N = \{16, 8\}$).

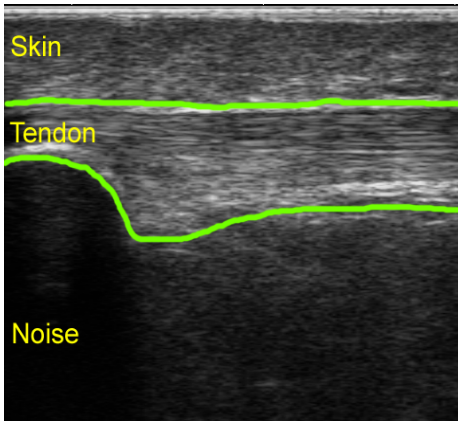
4.7 *In vivo* Musculoskeletal Results

We applied VSBM to *in vivo* frame pairs to establish suitability before further development. Figures 4.8(a), 4.8(c) and 4.8(e) are freehand musculoskeletal B-scans of 2 patella tendons and a digital flexor tendon. These frames are samples from captured sequences of 30 frames over a period of independent flexion to extension motion. Figures 4.8(b), 4.8(d) and 4.8(f) were produced using the VSBM approach using block sizes $M \times N$, where $M, N = \{32, 16, 8\}$. As before, above the tendon is tissue with minimal structure and highly well formed speckle, and below noise. The results illustrate consistent random motion in these lower regions, but in the main tendon area, we obtain subjectively accurate displacement fields in all our sequences. To quantify the accuracy of these results, we again use the image reconstruction error and global correlation, with results shown in Table 4.4. For the examples shown, DFD error was lower than FD error by about 50%, and correlation increased by an approximate 1% with an average of 98.6%. Extensive *in vivo* displacement accuracy results are reported later in Chapter 6.

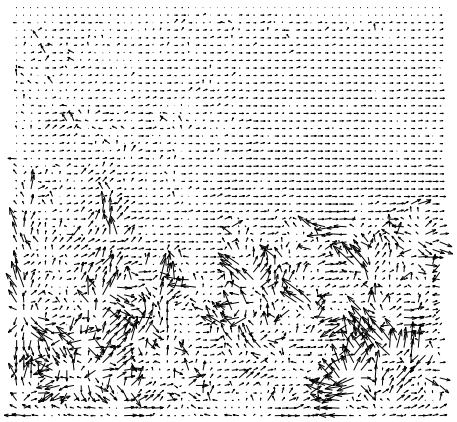
Frame	**PRE-WARP**		**VSBM (MSE)**	
Pair	**FD**	**CORR%**	**DFD**	**CORR%**
Patella Tendon *Case #1*	226.50	97.69	132.20	98.65
Patella Tendon *Case #2*	256.23	97.45	145.91	98.54
Digital Flexor *Case #3*	252.16	97.40	125.21	98.70

Table 4.4: Image reconstruction error measures for VSBM (MSE) using *in vivo* musculoskeletal data.

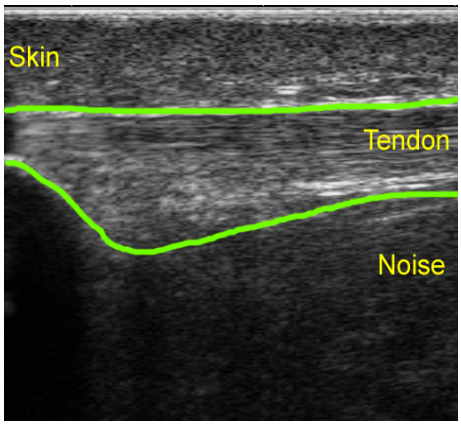
The results also confirmed the different motions occurring from those tendons encapsulated with sheaths and those that are not. From the displacement fields in Figures 4.7(d), 4.8(b) and 4.8(d), the patella tendon is clearly not encapsulated by a sheath, as similar motions from the tendon are replicated in skin and the surrounding tissues. The digital flexor tendon in Figure 4.8(e) is encapsulated by a sheath. This corresponds to the estimated displacement field in Figure 4.8(f), which shows little motion in the upper region (below the skin), followed by a narrow band of uniform displacements in the actual tendon area. This is then followed by regions of varying motions below it caused by noise. Specific regions and enlarged displacements are shown for further clarity in Figure 4.9.



(a) Patella *case #1*.

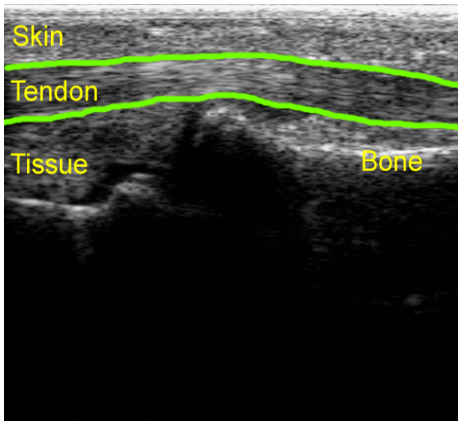


(b) VSBM displacement field.



(c) Patella *case #2*.

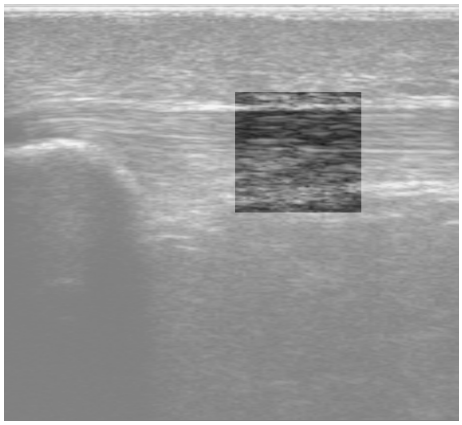
(d) VSBM displacement field.



(e) Digital flexor *case #3*.

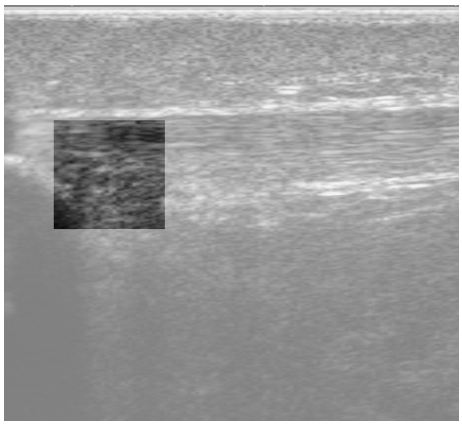
(f) VSBM displacement field.

Figure 4.8: Frames from sequences of freehand *in vivo* musculoskeletal data.

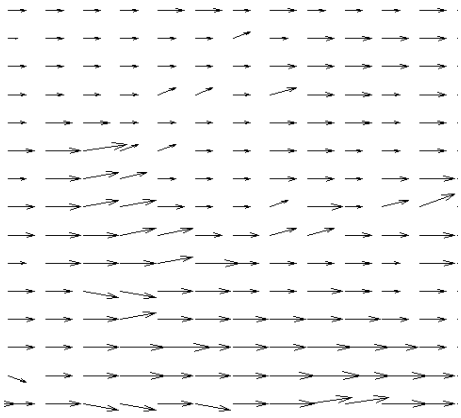


(a) Patella ROI *case #1*.

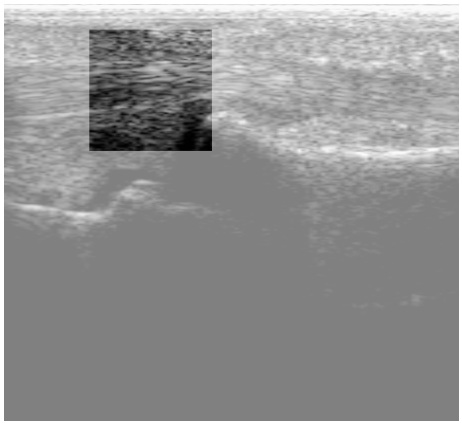
(b) VSBM displacement field.



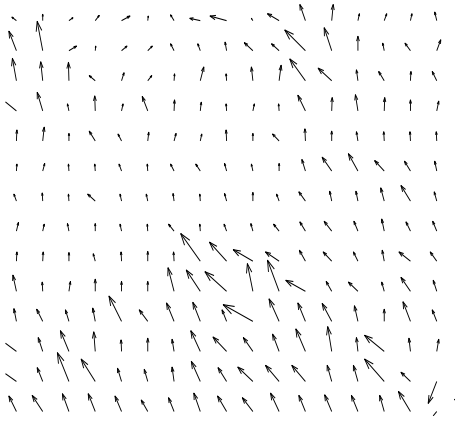
(c) Patella ROI *case #2*.



(d) VSBM displacement field.



(e) Digital flexor ROI *case #3*.



(f) VSBM displacement field.

Figure 4.9: Enlarged regions of displacements for highlighted sections of tendon from *in vivo* sequences.

4.8 Conclusion

This chapter has assessed the feasibility of using 2D speckle tracking techniques for analysing frame pairs of *in vitro* and *in vivo* musculoskeletal ultrasound sequences. The earlier part of this work presented VSBM, and later demonstrated the accuracy of the generated displacement fields for various datasets. It was found that the block size proved critical in locating a peak match between interframe blocks. Additionally, it was observed that if errors occurred at the larger block sizes, then displacement estimates at subsequent block sizes were generally poor. The applied MSE matching criterion proved an adequate quantitative measure of image difference, although obviously unable to reflect the qualitative differences that human viewers perceive. This early work has shown that the assumption of speckle motion replicating the underlying tissue and tendon deformation to be robust for our data.

A comparative analysis was performed to evaluate several approaches to speckle tracking. To measure the accuracy of interframe displacement results, we used the velocity angular error, image reconstruction error and global normalised cross correlation. For validation we used the groundtruth sequence STATIC-LANDMARK1, capturing a tendon with translational motion. As expected Proesmans et al. [80] differential based algorithm was sensitive to noise, with inaccurate numerical differentiation resulting in direct errors in the displacement fields. It was observed, however, that efficient displacement post-processing improved results significantly, but over and under-smoothing produced a detrimental effect. Proesmans et al. and Camus DFD errors ranged between $106.63-119.05$ and $214.32-223.30$ respectively. These were both high and showed inconsistencies for each frame pair, unlike block matching. Interestingly, Camus demonstrated that further temporal support of greater than interframe analysis was even more challenging, implying that interframe analysis is a sufficiently difficult problem. In agreement with Liu [51], we also observed that Camus's method is fast but inefficient, making it inappropriate for our purposes.

A modification for techniques that search for interframe correspondence, is optimising the search region based on the observed motion. For example, a horizontal biased search region would restrict displacement measurement laterally rather than axially, favouring the lateral motion observed in our *in vitro* data. Unfortunately, this would be of limited use for our *in vivo* sequences, which is much more varied, as shown in Figure 4.8 and throughout this thesis.

The latter part of this chapter analysed the more complex motion found in our *in vivo* data, which posed significantly more challenges due to a number of issues. First, ultrasound images are characterised by Rayleigh-governed speckle noise, which yields a low SNR. Second, motion ambiguities due to insufficient representation of spatial information (occlusion and aperture problems), are clearly apparent in regions of image saturation or specular reflection, and in homogeneous regions of weak acoustic scatterers. Third, local regions in each frame that are decorrelated due to strain, or highly correlated from uniform speckle. The final challenge was the fact that motion occurs at various scales. Many of these issues were covered in Chapter 2. We observed that the accuracy produced from using variable block sizes $M \times N$, where $M, N = \{32, 16, 8, 4\}$ for the speckle de-correlation and saturation issues, showed improvements for *in vitro* and *in vivo* data, producing dense displacement fields.

This work confirmed that block matching approaches are suitable for our data, with multiple block sizes improving displacement accuracy. Standard block matching for motion estimation using measure such as MAD (L1-norm), SSD or MSE (L2-norm), are appropriate for image sequences characterised by Laplacian or Gaussian statistics. Unlike the tendon and regions of well-defined structure, regions of dense speckle are best characterised by Rayleigh statistics. Hence, the previous measures are not optimal for motion estimation in ultrasound image sequences. Therefore, in the next chapter we introduce a novel approach of alternating two matching measures, to adapt to image content. Furthermore, we analyse full sequences and not manually selected frame pairs, prominent in existing research.

Chapter 5

Trajectory Fields from Motion Estimation

5.1 Introduction

This chapter provides a comprehensive study of a hierarchical region-based motion estimation technique, using novel refinements to resolve existing speckle tracking issues that were highlighted in Chapter 4. The basis of our approach uses multiple hierarchical region-based matching, encapsulating more unique speckle pattern without restricting spatial resolution. From our earlier experiments, multiscale template analysis produced consistently superior results than single scale or biased shaped and sized templates [12] that correspond to the shape of the object to be tracked. We extend our previous interframe displacement work by using trajectories, which quantify continuous temporal displacement of speckle movement for complete sequences. This improves upon the prominent approach of specifically analysing manually selected frame pairs, as trajectories incorporate interframe displacements for every frame pair for all sequences, yielding measurements of local temporal movement. Trajectory drift correction is also applied, to ensure regions are accurately tracked in space and time. The proposed method corrects for potential freehand transducer motion, and updates for possible changes in the image plane as a result of axial, lateral and elevation movement. To improve displacement accuracy, displacements are refined by automatically selecting between two matching measures to compensate for speckle noise. The first measure uses the normalised cross correlation (NCC)

for regions of strong signal (refined to minimise aperture problems), and the second measure CD_2 is applied in regions of low SNR from speckle, determined by the SNR. Finally we determine displacement field accuracy, and demonstrate trajectory homogeneity by exploiting the trajectory path coherence, using the output as a validation measure.

In this chapter several strategies will be introduced to improve on the estimated displacements produced from block matching methods presented in the previous chapter. These include:

1. To ensure the appropriate use of similarity measures in regions of varying signal content, a novel alternating similarity measure strategy is introduced, which applies either the normalised cross correlation in regions of tendon (low speckle), or the CD_2 measure in regions of soft tissue (dense speckle), indicated by the local image SNR.

2. To improve on specific frame pair analysis, interframe displacements are extended to trajectories that track regions of signal through a sequence, utilising strategies for:

   (a) Trajectory drift correction, to eliminate temporal displacement error from propagating displacement inaccuracy.

   (b) Trajectory updating, to adjust the trajectory fields to include movement in the 2D image caused from movement in the 3D volume.

   (c) Trajectory path coherence, to provide a means of measuring and comparing a trajectory fields temporal displacement activity.

The chapter commences with ultrasound image registration in Section 5.2, before introducing in Section 5.3 the proposed approach to measuring displacements in sequences, describing the aperture problem, combined matching measures and template updating solutions. Next, trajectory path coherence is defined in Section 5.4, and displacement smoothing in Section 5.5. Results from our combined matching measures approach are reported in Section 5.6. Using our proposed method, trajectory fields and accuracy results are presented in Section 5.7 from analysing the tissue mimicking phantom, *in vitro* tendon and *in vivo* data. Finally, we conclude with a discussion in Section 5.8.

5.2 B-scan Image Registration

For each ultrasound image a coordinate system is defined, centred at the position of the transducer with 3 axes as illustrated in Figure 5.1. The axes are labelled as: the axial axis (Y), taking the direction of propagation of the ultrasonic beam, the lateral axis (X), perpendicular to the axial axis, and the elevation axis (Z), perpendicular to both the axial and the lateral axes.

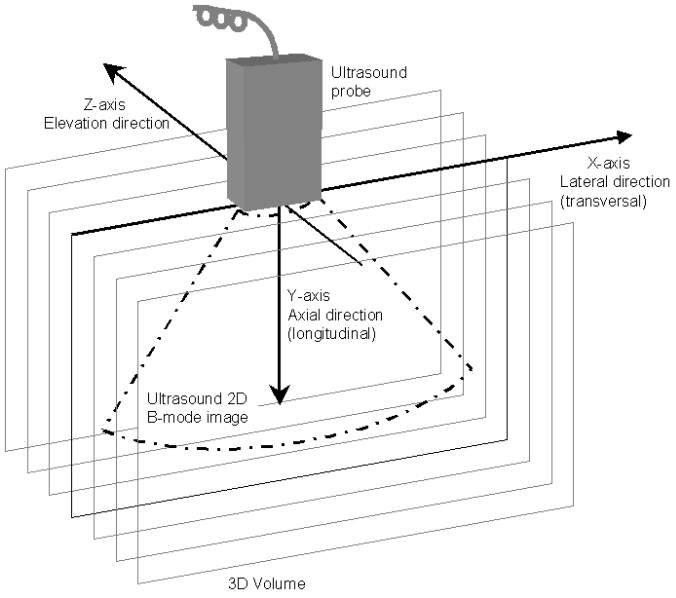


Figure 5.1: Transducer right-handed global image coordinate system, illustrating that each frame captures a 2D slice in a 3D volume.

Issues for all dynamic ultrasound analysis include motion ambiguities that can result from regions of image saturation or specular reflection, homogeneous regions of weak acoustic scatterers, occlusion, aperture problems and speckle decorrelation. Due to B-mode generation, decorrelation due to relative lateral scatterer motion is more severe than axial scatterer motion.

Freehand scanning allows the complete freedom of movement of a transducer, capturing images that are arbitrarily orientated in space. To measure only subject motion rather than transducer motion, either the image must be translated to a fixed coordinate, or the displacements updated with a global value of the transducer motion. In practise, during any clinical freehand ultrasound sequence acquisition, typically both transducer and subject are kept stationary to ensure a reproducible imaging plane. This consistent scanning enables, in the simplest case, for individual images to be manually registered. In this work, we prefer an automated ap-

proach to solve this problem. There has been much active research in automatically registering ultrasound volumes, with methods forming distinct groups:

1. Acoustical Sensors. Signals from an emitter attached to a transducer are received by remote microphones, which are positioned in different orientations, for example [64].

2. Optical Sensors. As above, but the emitters are replaced by infrared light emitting diodes that require a line of sight, for example [99].

3. Magnetic Fields. A receiver is attached to a transducer with a remote transmitter generating a magnetic field, allowing complete freedom of movement, for example [11].

4. Mechanical Systems. These systems have mechanical arms clamping the transducer, which allow various degrees of freedom, for example [88].

5. Image Registration. Global displacements are derived locally [111] and globally [98] using intrinsic measures by analysing the ultrasound images.

Without extrinsic position measurements from a sensor attached to the transducer, we have no prior knowledge of transducer position. Consequently, there is no means of determining the position and orientation of the scan plane relative to the transducer. Using image analysis, we define the registration problem mentioned in point 5. This is a special case of the correspondence problem where frames are rigidly or non-rigidly shifted with respect to each other.

In the previous chapter, interframe motion using single frame pair analysis assumed any transducer motion was negligible or zero. Over time, however, this stationary assumption fails. From observing freehand clinically relevant musculoskeletal B-scans, to correct for the error induced by transducer motion, we calibrate the transducer automatically by using the signal between the transducer and skin inclusively. This region appears in the top most part of the B-scans in Figure 3.8. This has been found to be consistent with all data. Therefore, instead of using non-rigid frame registration, which is known to introduce some interpolation errors and to be more computational, we register the displacement fields. This is achieved by updating the displacement fields with the mean displacement in the identified region. Assuming displacements are accurate with minimal elevation motion and movement is only transducer dependent, this can correct for small transducer lateral motion. During freehand scanning, the transducer is reasonably stationary.

5.3 Proposed 2D Speckle Tracking Methodology

Given an ultrasound sequence, interframe displacements $\boldsymbol{d}_t = (u_e, v_e)$ are estimated generating displacement fields, where $\boldsymbol{d}_t$ defines the displacement of an object at location $\boldsymbol{x} = (x, y)$ to its location $\boldsymbol{x}' = (x+u_e, y+v_e)$ at time $t+1$. The fundamental approach uses a multiscale template matching pyramid initialised by a regular lattice $\mathcal{R}$, sampled by $P \times Q$ (typically 8×8 or 4×4) on a preliminary ultrasound frame f_t. Templates of $M \times N$, where M,$N = \{64, 32, 16, 8, 4\}$ are used to increasingly improve displacement accuracy until $M = P$ and $N = Q$ defining the finest scale. Displacement and trajectory field density is dependent on the initial size of $\mathcal{R}$.

Using NCC as a first measure, the local disparity between reference I and candidate I' templates is determined by identifying the maximum correlation coefficient c_{max}. Instead of using the spatial domain NCC (4.4), for efficiency the NCC is performed in the frequency domain using the FFT defined as:

$$\boldsymbol{c} = \frac{\mathcal{F}^{-1}\{\mathcal{F}^*(I)\mathcal{F}(I')\}}{\sqrt{\int\int |\mathcal{F}(I)|^2 \cdot \int\int |\mathcal{F}(I')|^2}} \qquad \text{leading to} \quad \{0 \leq \boldsymbol{c} \leq 1\} \tag{5.1}$$

where $\mathcal{F}$ is the Fourier transform of templates I and I', $\mathcal{F}^{-1}$ is the inverse Fourier transform, and * is the complex conjugate (accomplishing the reversal of I by negating the imaginary component), with the numerator defining the cross correlation. In order to overcome the problem of possible significant variations in spectral amplitude over the frequency range, due to fluctuations in spatial intensity, it is advantageous to normalise the correlation, dividing by the product of the magnitudes of the spectral components, which increases the sharpness of the correlation peak so that $\boldsymbol{c}$ ranges between 0 and 1. For curvilinear tendons or regions of attenuation artefacts, the mean intensity for each template may not be the same, in these cases intensity invariance can be achieved by subtracting the mean intensity from I and I' prior to (5.1), for significant sized templates. Since the 2D Fourier transform is periodic with the block size $M \times N$, the displacement estimates need to be unwrapped to accommodate negative displacements. Thus the range of displacement estimates is $[-M/2, N/2]$, for example, to estimate displacements within the range $[-32, 32]$, the block size should be at least 64×64 (the largest scale). NCC tracking is sensitive to imaging scale, rotation and perspective distortions. In this context minimal perspective and rotation distortions potentially exist, however, the NCC does enable equal sized patterns to be detected by a rotation distortion of 5° to 10° [68]. By

using multiple template scales, we achieve scale invariance and improved accuracy from local reflectance variations, which can be caused by subtle artefacts present in the images.

The spatial displacement vector $\boldsymbol{d}_t$ is estimated in the x and y directions, after locating c_{max} from the resulting correlation field $\boldsymbol{c}$. For each position i in $\mathcal{R}$ the NCC is performed at multiple template scales, using the previous $\boldsymbol{d}_t$ as candidate template I' offsets. This yields an approximate displacement as we are yet to take into account the aperture problem or speckle. Multiple templates were not constructed using repeated low-pass filtering and sub-sampling as in a Gaussian pyramid [50], as first the intensity distribution of ultrasound images is not Gaussian and second, the templates at lower resolutions contain a distinct lack of any useful signal for correspondence, a problem from filtering mentioned in Chapter 2. Using NCC typically produces strong correlation matching in high frequency data (>10 MHz), whilst assuming an increased speckle SNR from high capture rates and sparse scatterers. Speckle noise, predominantly at lower frequencies (<7 MHz), reduces the NCC matching accuracy, highlighting the necessity of a suitable alternative measure in such regions. Other causes of correlation reduction also include a lack of signal (probe de-coupling or curvilinear tendons), signal saturation (incorrect gain controls or bone), and minimal features, causing problems for any similarity measure. Later in Section 5.3.4 we introduce a new approach to combine similarity measures, aiming to improve displacement accuracy for sequences capturing regions of different signal.

5.3.1 Interframe Displacements to Trajectories

At this stage we have quantified a displacement $\boldsymbol{d}$ for each block in f_t for a lattice $\mathcal{R}$ yielding an approximate interframe displacement field. For sequences, this process is repeated for every frame pair in the sequence to give:

$$\boldsymbol{d}_t \dots \boldsymbol{d}_{t+n} = \text{NCC}\{(f_t, f_{t+1}) \dots (f_{t+n}, f_{t+n+1})\} \tag{5.2}$$

Instead of using the same $\mathcal{R}$ for each frame pair, as above, we redefine $\mathcal{R}$ with the estimated displacements from (f_t, f_{t+1}) for the next frame pair (f_{t+1}, f_{t+2}). Consequently, tracking results in a displacement vector with temporal history. The motion trajectory can be formally defined as a vector function $\boldsymbol{T}_i(t; \boldsymbol{x}, t_0)$, which specifies the horizontal and vertical coordinates at time t of a reference location at $\boldsymbol{x}$ at time t_0. Given the motion trajectory $\boldsymbol{T}_i$, the velocity at

time t' at a location $\boldsymbol{x}'$ along the trajectory can be defined by:

$$V(\boldsymbol{x}', t') = \frac{d\boldsymbol{T}(t; \boldsymbol{x}, t_0)}{dt} \tag{5.3}$$

Trajectories represent the temporal tracking of features through a sequence. A powerful benefit from the trajectory definition is the temporal history, which has a direct relationship with temporal displacement coherence. Example trajectories are shown in Figure 5.4, demonstrating a full 30-frame history of the temporal displacements from $f_1 \rightarrow f_{30}$. However, the amount of displacement history can be changed, for example a moving 5-frame history where displacements measure $f_1 \rightarrow f_5$, $f_2 \rightarrow f_6 \ldots f_n \rightarrow f_{n+5}$, which is shown in Chapter 6.

5.3.2 The Aperture Problem Effect

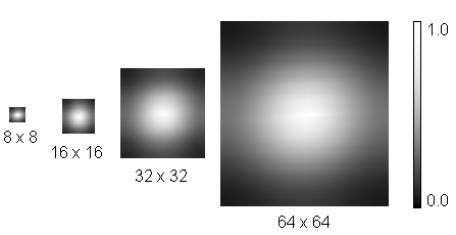
As a result of tendon anisotropic striated structure and tendon linear movement, discussed in Section 3.2, significant aperture problems occur in regions of tendon rather than in regions of speckle texture from soft tissue, hence the above displacements are only approximate. In the perfect case, the correlation field from the NCC would typically contain a single strong peak, however, where the aperture problem is greatest the correlation field typically contains a single strong elliptical shaped peak. Consequently, the highest correlation coefficients, of very similar value, lie within a large displacement range. A further problem, identified by Thomas [96] is the possibility that the correlation field can contain multiple peaks, due to the presence of any velocity gradients within the template, however, in our data this is unusual.

Therefore, we threshold the correlation field $\boldsymbol{c}$ produced from the NCC, to yield a binary field $\boldsymbol{c}_{out}$, where the threshold value is $c_{thresh} = 0.9$, a value that complies with [102], given as:

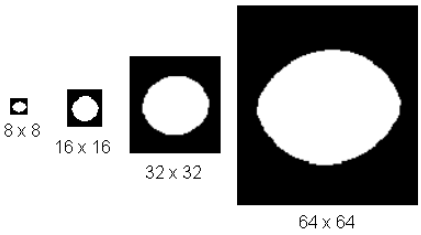
$$\boldsymbol{c}_{out}(i,j) = \begin{cases} 0 & \text{if } \boldsymbol{c}(i,j) < c_{thresh} \\ 1 & \text{otherwise} \end{cases} \tag{5.4}$$

The single most significant flagged region in $\boldsymbol{c}_{out}$ is used to provide a mask, to reduce an exhaustive search in $\boldsymbol{c}$, to ensure the maximum correlation coefficient is located in this region of support. The major and minor axes of the flagged region in $\boldsymbol{c}_{out}$ determines the anisotropy ratio, which quantifies the extent of the aperture problem. This indicates the precision of the location of the peak and displacement accuracy, for example, an equal ratio would indicate no aperture problem. This is repeated for all template scales with minimal computational expense,

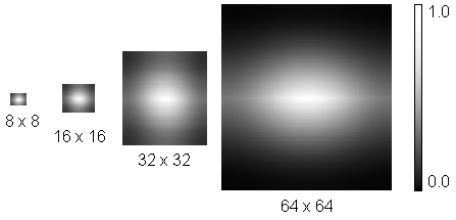
and can suitably indicate anisotropic motion more effectively than 1D RF elastography, due to measuring 2D motion in both axial and lateral directions. To illustrate the anisotropic motion, in Figure 5.2 we show local multiscale correlation fields $\boldsymbol{c}(i,j)$ and $\boldsymbol{c}_{out}(i,j)$ according to (5.4), for a single displacement from *in vivo* tendon data. We have observed that cases of the aperture problem are typically no more extreme than illustrated.



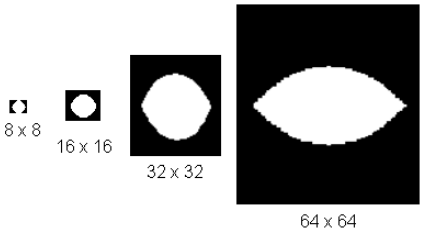
(a) Example $\boldsymbol{c}(i,j)$ with isotropic peak.



(b) Example $\boldsymbol{c}_{out}(i,j)$ with isotropic peak.



(c) Example $\boldsymbol{c}(i,j)$ with anisotropic peak.



(d) Example $\boldsymbol{c}_{out}(i,j)$ with anisotropic peak.

Figure 5.2: Local correlation fields for a single displacement using multiscale template analysis, demonstrating the effects of the aperture problem on the correlation coefficients.

5.3.3 Block Matching Motion Estimation using Maximum Likelihood

As mentioned earlier, to combat reduced NCC accuracy in regions of increased speckle noise, we propose the use of an additional, selectively chosen measure, namely the CD_2 measure introduced by Cohen and Dinstein [20]. Recently, Boukerroui et al. [13] showed that in regions of dense speckle CD_2 is a more precise measure than for example, the NCC or MSE criteria. The reason is due to CD_2 incorporating the characteristics of the ultrasound imaging modality, which considers the multiplicative speckle noise in ultrasound images.

Until now, block matching methods described in Chapter 4 considered the SAD, SSD and MSE matching criteria, which Patras et al. [73] show are equivalent to a maximum likelihood es-

timator, assuming the motion intensity differences follow independent Laplacian or Gaussian distributions respectively. When using block matching techniques with the maximum likelihood estimator, we perform a full search to match a block from the first frame with a block in the second frame, within a predefined search area. The estimated displacement vector is the one maximising a likelihood function, formulated according to the image formation model, instead of minimising an intensity difference criterion.

In this section, we denote the intensities for the reference template I in f_t as $\boldsymbol{a}_i = [a_1, \ldots, a_j]$ and the intensities for the candidate template I' in f_{t+1} as $\boldsymbol{b}_i = [b_1, \ldots, b_j]$ where i is an individual block and j a pixel. Therefore, the estimated displacement vector $\hat{\boldsymbol{d}}$ is found by the maximisation of the following conditional probability density function:

$$\hat{\boldsymbol{d}}_i = \arg\max_{\boldsymbol{d}_i} \; p\left(\boldsymbol{a}_i | \boldsymbol{b}_i, \boldsymbol{d}_i\right) \tag{5.5}$$

which determines the likelihood of locating $\boldsymbol{b}_i$ in f_{t+1} given the original $\boldsymbol{a}_i$ in f_t and a displacement $\boldsymbol{d}_i$. Making the assumption that each block $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ are corrupted with multiplicative speckle noise, then the observed pixel j in block i becomes $a_{ij} = \eta_1 s_{ij}$ and $b_{ij} = \eta_2 s_{ij}$ where s_{ij} is the original signal, and η_1 and η_2 are two independent noise elements with Rayleigh probability density functions. The relationship between the observed noisy pixels in both blocks then becomes:

$$a_{ij} = \eta b_{ij} \qquad \text{where} \qquad \eta = \frac{\eta_1}{\eta_2} \tag{5.6}$$

Taking into account that the noise term η is the division of two Rayleigh distributed random variables, the conditional probability density function is given in [71] and [20] as:

$$p\left(\boldsymbol{a}_i | \boldsymbol{b}_i, \boldsymbol{d}_i\right) = \prod_{j=1}^{MN} \left\{ \frac{2(a_{ij}/b_{ij})^2}{\left((a_{ij}/b_{ij})^2 + 1\right)^2} \right\} \tag{5.7}$$

Using a block matching approach to maximise this equation is equivalent to maximising (5.5), thus, for a specific displacement the corresponding point in the next frame is computed by maximum likelihood using blocks of $M \times N$. Motion estimation based on (5.7) is known as the CD_2 measure, named by Cohen and Dinstein, which we use in our multiscale template matching approach. Finally, the CD_2 measure achieves the best displacement estimates for regions that closely conform to a Rayleigh distribution[1].

[1] Further explanation appears in Appendix A.

5.3.4 Combined Ultrasound Speckle Pattern Similarity Measures

We employ both NCC and CD_2 measures with multiscale template analysis, using the NCC as the primary matching measure due to its high accuracy in low speckle density regions. However, we require an appropriate means of automatically determining the amount of local speckle present to decide which measure to apply. For this we use the SNR, λ, given by the ratio of the mean I_μ to the standard deviation I_σ of the pixels contained in a local region I:

$$\lambda = \frac{I_\mu}{I_\sigma} \tag{5.8}$$

In a typical area of uniform dense ultrasound speckle, Wagner et al. [106] have determined an expected SNR value of $\lambda = 1.91$. We verified this value with *in vivo* data using multiscale template regions of a uniform area, located at the focal zone of the ultrasound beam, and computed the mean SNR for 5 sequences of 30-frames. Results are shown in Figure 5.3, with the mean SNR converged at $\lambda \approx 2$ confirming the use of 1.91. This highlights a requirement to use a tolerance, empirically derived at 25%, to ensure a reasonable speckle sensitivity for *in vivo* images where speckle is seldom uniform. Only regions of $M \times N$, where M,$N \geq 16$ were able to determine reliably that a region contained uniform speckle, hence the previous scale SNR is used for M,$N < 16$. Therefore, we propose to use and apply either the NCC or the CD_2 measures determined by the SNR, which implies the speckle density present in a region:

$$\text{measure} = \begin{cases} \text{NCC} & \text{if } \lambda > (1.25 \times 1.91) \\ \text{CD}_2 & \text{otherwise} \end{cases} \tag{5.9}$$

The SNR increases with a low amount of speckle (reaching ∞ for specular reflection), justifying the NCC measure. However, SNR decreases for high amounts of speckle, indicating that a region should be tracked using the second CD_2 measure. This is evaluated using the associated reference and candidate templates, in a full search with the same extents as the primary NCC measure for the larger scales. The speckle SNR is used as an indicator of speckle content (rather than correlation), as typically featureless regions of uniform speckle produce high correlation coefficients with surrounding speckle. The SNR is also sensitive to other image components, for example, feature boundaries resulting in a low local SNR, therefore we also check that c_{max} is low. This approach allows the proposed method to adapt to image content.

In Figures 5.4(a), 5.4(c) and 5.4(e), we illustrate sample spatial and temporal trajectory fields

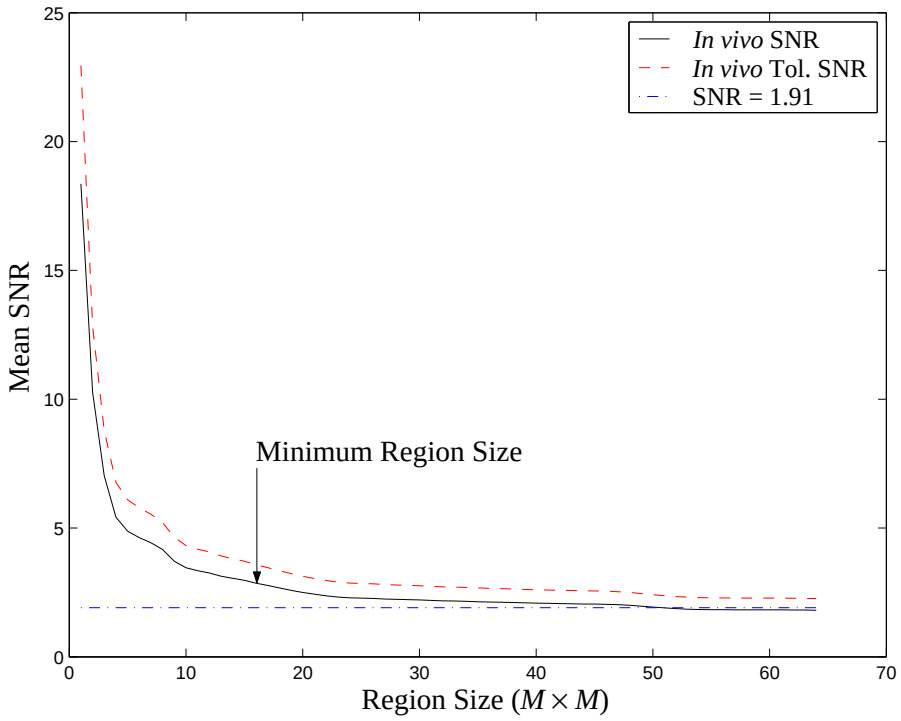
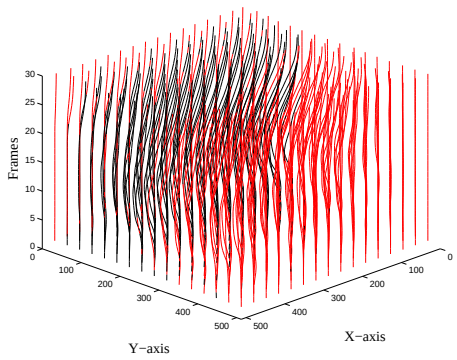


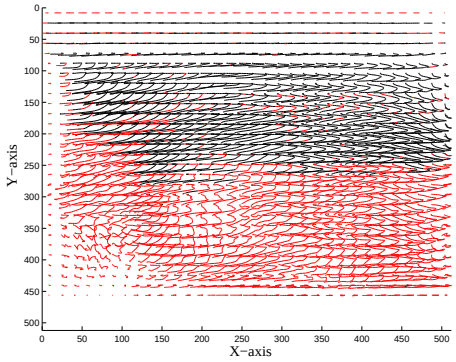
Figure 5.3: A plot of the speckle SNR for multiple block scales in a region of dense speckle from our *in vivo* data. Experimentally using blocks of $M \times N$, where M,$N \geq 16$, the SNR converges closely to the theoretical SNR of 1.91 [106] in regions of Rayleigh statistics.

for a video sequence of the patella, achilles and digital flexor tendons respectively. We also show these as spatial displacement fields in Figures 5.4(b), 5.4(d) and 5.4(f), with a 30-frame displacement history.

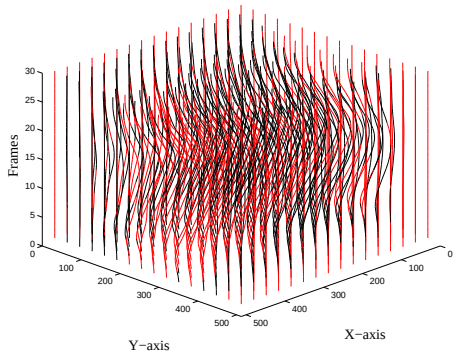
The trajectory field density is reduced for clarity, with each result corresponding to the labelled B-scans in Figure 3.8. To highlight our combined similarity measure approach, NCC usage is indicated in black and CD_2 usage in red. By comparing both B-scans and trajectory fields, NCC usage was mainly in the tendon region and in regions of strong substructure, with CD_2 used elsewhere, as expected. Also, a minority of trajectories use both measures to track the same region, which is indicated by both black and red displacements. This variation can occur from motion blurring and from temporal speckle changes. Additional analysis of our combined similarity measure approach is conducted later in Section 5.6.



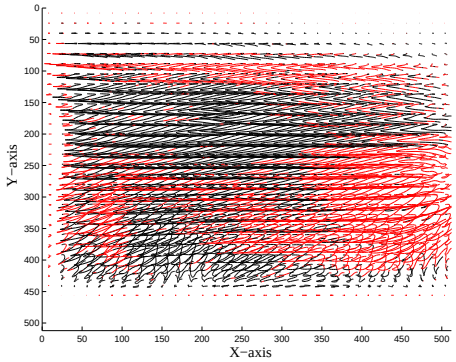
(a) Spatiotemporal trajectory field.



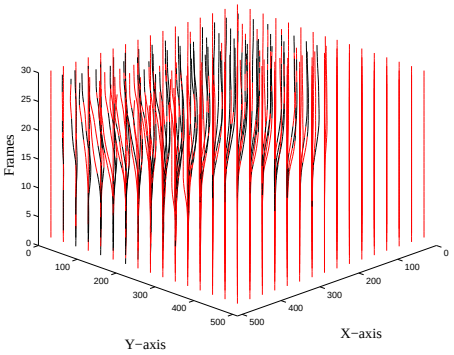
(b) Trajectories with 30-frame history.



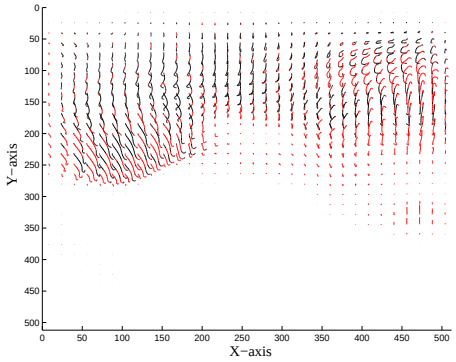
(c) Spatiotemporal trajectory field.



(d) Trajectories with 30-frame history.



(e) Spatiotemporal trajectory field.



(f) Trajectories with 30-frame history.

Figure 5.4: Spatiotemporal and spatial trajectory fields demonstrating a 30-frame displacement history using our combined similarity measure approach. NCC usage is indicated in black and CD_2 usage in red. *In vivo* sequences analysed were: (a-b) Patella, (c-d) achilles and (e-f) digital flexor tendon. ♠

5.3.5 Temporal Trajectory Updating

We now deal with the issue of temporal updating. The underlying assumption behind region-based matching is that the texture of the reference template I remains consistent throughout the captured sequence. For a finite period of time this is acceptable, hence for interframe matching, I is *not updated* prior to matching with the candidate template I':

$$I_{t+1}(\boldsymbol{x}) = I_1(\boldsymbol{x}) \qquad \forall\ t \geq 1 \tag{5.10}$$

For tracking across multiple frames, to generate trajectories, I eventually becomes an inaccurate representation of the tissue being tracked, correlating poorly with candidate templates. As a result, trajectory paths can deviate and this error is known as drift. *Temporal updating* reduces the error by resetting I in every frame with a new template:

$$I_{t+1}(\boldsymbol{x}) = I_t(\boldsymbol{x} + \boldsymbol{d}_t(u_e, v_e)) \qquad \forall\ t \geq 1 \tag{5.11}$$

However, *temporal updating* is limited by an accumulated tracking error, resulting in the updated reference template being inaccurate. Instead, we use *drift updating* (previously only applied to a few points for vehicle tracking in [59]) to eliminate temporal drift error. Trajectory *drift updating* implements *temporal updating*, but realigns I_t to the original reference template I_1, this is illustrated in Figure 5.5.

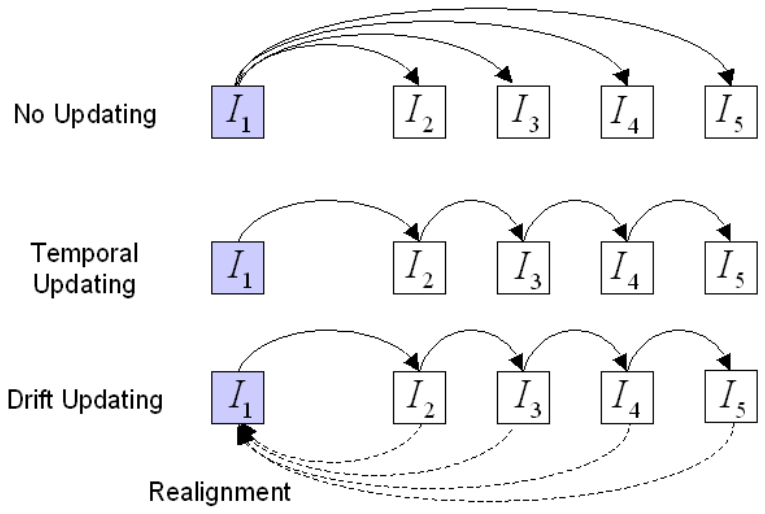


Figure 5.5: Schematic of trajectory matching using no updating, temporal updating and drift updating.

Trajectory *drift updating* applies two matching steps for each displacement:

1. *Temporal updating,* to produce a displacement using the appropriate measure, as described in Section 5.3.4, and

2. Realignment, to correct for drift between reference templates I_t and I_1 using the faster NCC measure, updating the final displacement.

This is used for the largest template scale (automatically correcting for drift in smaller scales), and effectively doubles the number of correspondences, but is important to ensure that the original reference template remains the same throughout. As a check the realignment for the first frame pair should be zero if the displacement from *temporal updating* is correct. The realignment is only used when the correlation is high, as sequences that capture both large motions and deformations cause $I_t \neq I_1$. With hardware limitations of 30 frames and potentially high trajectory density, for the NCC measure the Fourier transforms of each reference template from f_1 are stored, keeping drift computation time minimal. Further extensions are possible at the expense of significantly increasing computation, for example, storing each temporal reference template to update for appearance variation [59], or a Kalman filter as applied in [84], using the blocks at the best match locations as observations to decide the correct reference template.

We compare both *temporal updating* and *drift updating* using the DYNAMIC-HEALTHY2 data. Table 5.1 reports the image reconstruction errors to measure displacement accuracy, using pre-warped FD and DFD measures, where, for example, a DFD of 16 measures a mean intensity error per pixel of 4. In general, both the FD and DFD errors were low (FD $= 14.71$ at $f_1 \rightarrow f_2$), as a result of the reduced pull rate. *Temporal updating* errors (DFD $= 13.75$ at $f_1 \rightarrow f_2$ and DFD $= 15.75$ at $f_9 \rightarrow f_{10}$) increased at a higher rate, than *drift updating* errors (DFD $= 11.13$ at $f_1 \rightarrow f_2$ and DFD $= 12.66$ at $f_9 \rightarrow f_{10}$), which proved more constant. In summary, only slight displacement inaccuracy occurred for sequences of < 10 frames, but next we demonstrate larger errors propagating for sequences of 30 frames.

We used a specific *in vitro* experiment capturing a tendon displaced at a linear rate along the lateral axis. A set of 100 trajectories (10×10) representing a tracked ROI of 40×40 pixels for 30 frames are shown in Figure 5.6. A typical actual trajectory is shown as a dashed red line in Figure 5.6(c), with the estimated trajectories from both drift and temporal updating as

Frame Pair	PRE-WARP FD	Temporal Updating DFD	Drift Updating DFD
$f_1 \rightarrow f_2$	14.71	13.75	11.13
$f_2 \rightarrow f_3$	15.05	13.77	11.37
$f_3 \rightarrow f_4$	15.89	14.74	11.90
$f_4 \rightarrow f_5$	14.56	13.43	10.75
$f_5 \rightarrow f_6$	15.81	13.17	13.42
$f_6 \rightarrow f_7$	14.82	13.41	11.49
$f_7 \rightarrow f_8$	14.96	13.65	11.75
$f_8 \rightarrow f_9$	15.15	14.15	11.95
$f_9 \rightarrow f_{10}$	16.05	15.75	12.66

Table 5.1: Error measures for proposed approach using STATIC-LANDMARK0.

a crossed black and circled blue lines respectively. The estimated trajectories using *temporal updating* rapidly diverged from the actual trajectory, with a final difference of 65 pixels. The estimated trajectories using *drift updating*, however, produced displacements similar to the actual trajectory, with a final difference of 4 pixels.

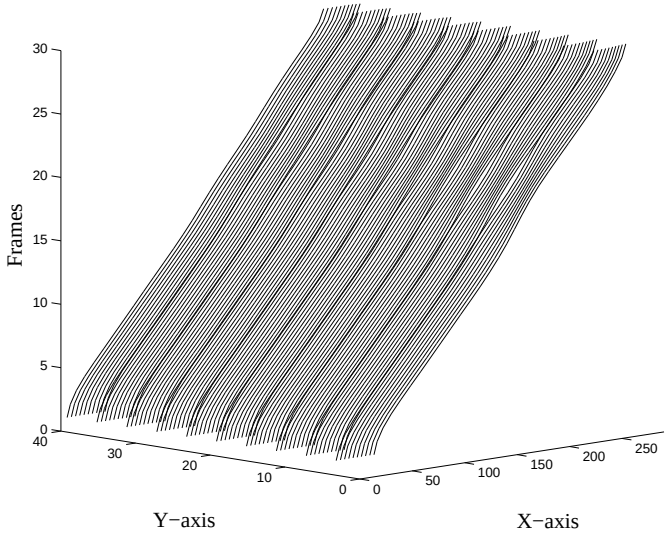
5.3.6 Spatial Trajectory Updating

Finally, tracking, especially in long sequences, requires updating trajectories for feature identification in each consecutive frame f_{t+1}, for new objects *entering* and old objects *exiting.* Potential causes of such trajectory updates are features traversing in the 2D plane, 3D volume, at image boundaries, and occlusion, producing potential trajectory clusters and voids. If the Euclidean distance between the coordinates for a block in $\mathcal{R}$ and all the corresponding neighbouring displacements in the displacement field are $> P, Q$, then a new trajectory is included for that block in $\mathcal{R}$. Also, if a displacement $\boldsymbol{d}_i$ is equal to any of its neighbouring displacements $\boldsymbol{d}_j$, then that displacement is flagged for removal, forming a single trajectory path.

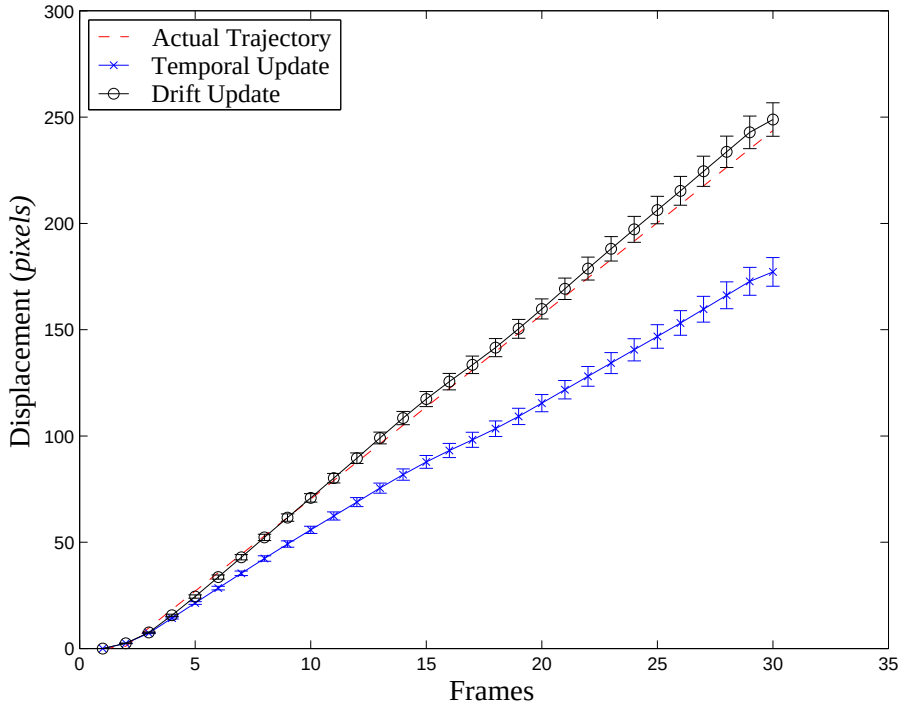
$$\text{Update} = \begin{cases} \boldsymbol{d}_{t,add} & \text{if} \quad \|\boldsymbol{d}_{t,i} - \boldsymbol{d}_{t,j}\|_p > P, Q \\ \boldsymbol{d}_{t,remove} & \text{if} \quad \boldsymbol{d}_{t,i} = \boldsymbol{d}_{t,j} \\ \boldsymbol{d}_{t,i} & \text{otherwise} \end{cases} \tag{5.12}$$

The best example of trajectory updating, occurs at image borders due to the continual motion of the tendon traversing from extension to flexion, this is demonstrated in Figure 5.7.

(a) Actual and enlarged ROI.



(b) Sparse trajectory field.



(c) Drift and temporal updating comparison.

Figure 5.6: *In vitro* tendon for temporal reference template updating. Note: Drift corrected trajectories contain a slight kink from a digitisation induced artefact. ♠

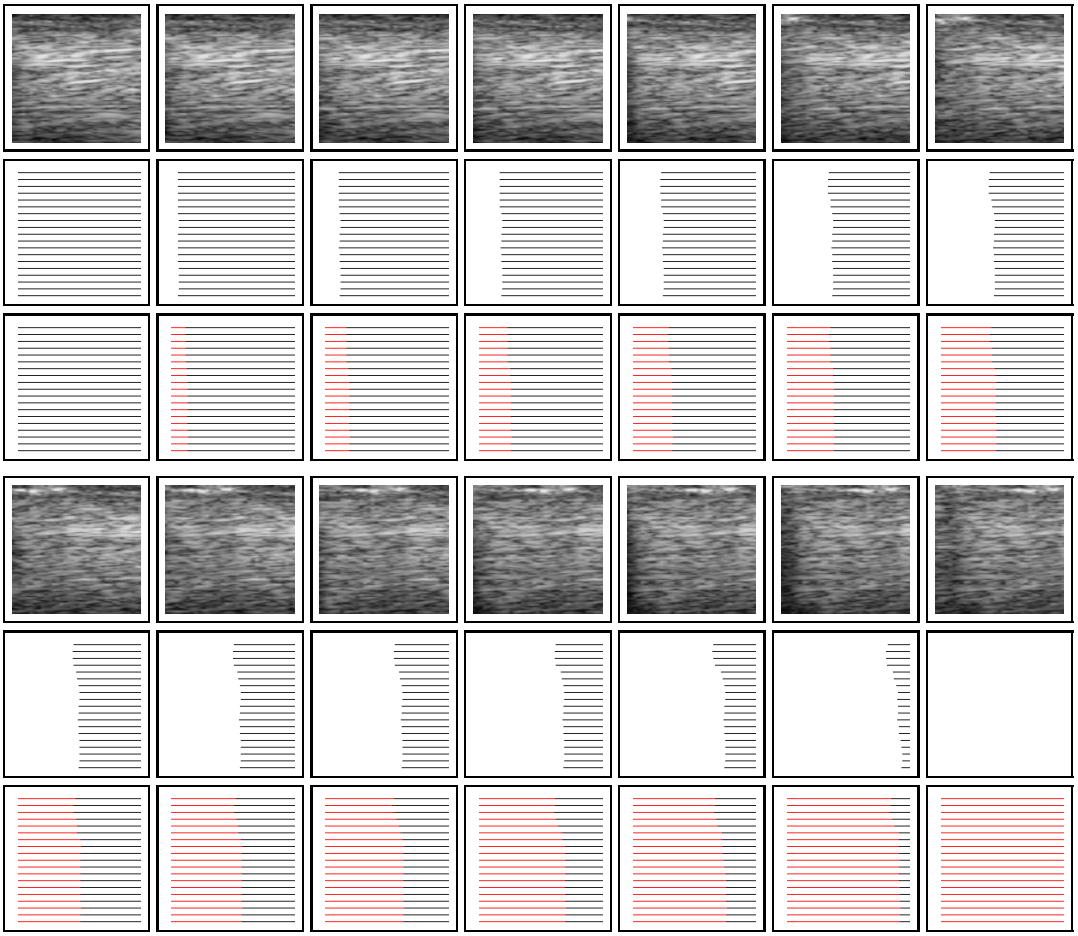
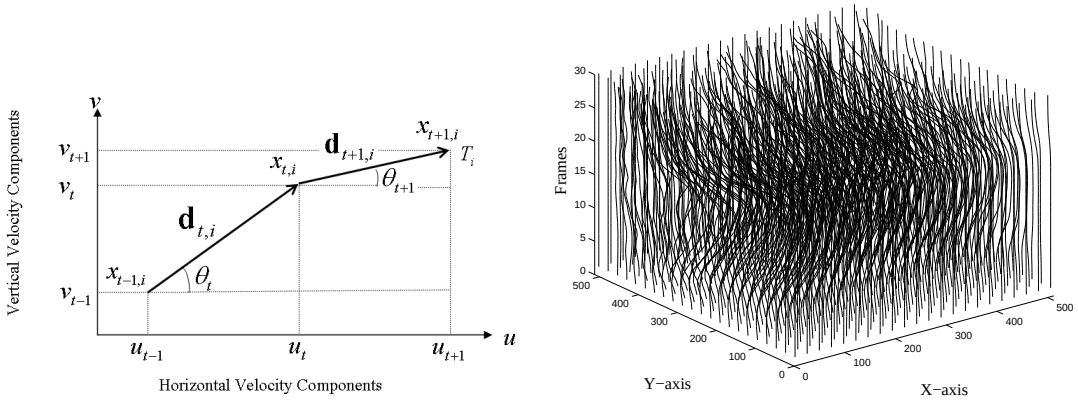


Figure 5.7: Groundtruth trajectory updating. The first and fourth rows show the B-scan ROI of the groundtruth sequence, the second and fifth rows show trajectories without updating (in black), and the third and sixth rows show trajectories with updating (in red). An *in vivo* example is on-line. ♠

We used a restricted aperture size to capture only a local region of tendon from the STATIC-HEALTHY0 sequence. Every other frame of the sequence is shown in the first and fourth row of Figure 5.7. The trajectories for the signal that exist from f_1 are in black. These are shown without updating in the second and fifth rows. The trajectories that are automatically included, as a result of voids in the previous trajectory fields, are shown in the third and sixth rows in red. At the start there are no voids in the trajectory field, as there is no significant motion. At the finish, however, the original trajectories no longer track objects in the scene and have been removed, leaving only new trajectories.

5.4 Trajectory Path Coherence

With modern ultrasound machines providing real-time frame acquisition rates, most object trajectories between consecutive frames have minimal abrupt changes in motion. Consequently, it is possible to formulate the notion of trajectory path coherence (TPC). TPC measures the coherence of motion for an object at any point in an image sequence, as for example used by [41] for tracking feature points in cinematic film sequences. In our case, an object is defined as the speckle captured in each original reference template in the lattice $\mathcal{R}$, for a complete frame or user specified region. The path coherence function Φ represents trajectory deviation, a smoothness measure of the derived object trajectory. Dense trajectory fields quantify the displacement of single or many objects throughout a sequence. We use the TPC to enable individual trajectories within a trajectory field to be compared, producing an indicator of neighbouring trajectory similarity. When a subset of trajectories from this field for a specific ROI is consistent with part of a single object, the expected TPC field would be reasonably homogeneous and within a localised range. Conversely, for a region of artefact or noise, the expected TPC would be varied and fluctuate within a large range. Lastly, under deformation, the expected TPC would gradually change. Consequently, TPC output is important as a trajectory confidence measure, summarising the general path coherency across the trajectory field.



(a) Trajectory schematic.

(b) Dense trajectory field.

Figure 5.8: Trajectory path coherence: (a) Schematic of the displacements that form a trajectory T_i and (b) an *in vivo* dense trajectory field for a 30-frame sequence.

To derive the TPC, each trajectory $\boldsymbol{T}_i$, which was formalised in Section 5.3.1, denotes: $\boldsymbol{T}_i = \{\boldsymbol{x}_{t,i}, \boldsymbol{x}_{t+1,i}, \ldots, \boldsymbol{x}_{t+n,i}\}$, a sequence of positions where $\boldsymbol{x}_{t,i}$ represents a point $\boldsymbol{x}$ at frame f_t along a trajectory i for a sequence of n frames, as shown in Figure 5.8. The deviation in the trajectory path $\Phi_{t,i}$ is defined by the displacement vector $\boldsymbol{d}_{t,i}$ for displacements $\boldsymbol{x}_{t-1,i}$ to $\boldsymbol{x}_{t,i}$ and so on: $\Phi_{t,i} = \Phi(\boldsymbol{x}_{t-1,i}, \boldsymbol{x}_{t,i}, \boldsymbol{x}_{t+1,i}) = \Phi(\boldsymbol{d}_{t,i}, \boldsymbol{d}_{t+1,i})$. Hence, for n frames the trajectory path coherence for a single trajectory $\boldsymbol{T}_i$ with $n-1$ displacements, is defined as:

$$\text{TPC}\ \{\boldsymbol{T}_i\} = \frac{1}{n-1} \sum_{t=1}^{n-1} \Phi_{t,i} \tag{5.13}$$

The expression $\Phi_{t,i}$ in (5.13) uses two normalised terms so that $\Phi \in [0, 1]$, with TPC ranging between 0 for low trajectory coherence and 1 for high trajectory coherence, given as:

$$\Phi_{t,i} = \frac{\alpha + \beta}{2} \qquad\qquad 0 \leq \Phi \leq 1 \tag{5.14}$$

For each trajectory the direction coherence α (change in angle) uses the dot product of the magnitude of displacement vectors $\boldsymbol{d}_{t,i}$ and $\boldsymbol{d}_{t+1,i}$. The sign of the angle of trajectory deviation is not included in the TPC function (absolute value of numerator in (5.15)), as the TPC considers the amount of deviation in the direction of motion, and not the direction in terms of left or right, yielding absolute values of cosines of angles giving numbers between 0 and 1. For example, when $\alpha = 1$, then the angle between the displacement vectors is 0° or 180°, meaning that both displacements $\boldsymbol{d}_{t,i}$ and $\boldsymbol{d}_{t+1,i}$ move in the same direction (high coherence), whilst $\alpha = 0$ would mean the angle between displacements $\boldsymbol{d}_{t,i}$ and $\boldsymbol{d}_{t+1,i}$ is either 90° or 270° thus are unrelated (low coherence).

$$\alpha = \cos(\theta) = \frac{|\boldsymbol{d}_{t,i} \cdot \boldsymbol{d}_{t+1,i}|}{|\boldsymbol{d}_{t,i}|\ |\boldsymbol{d}_{t+1,i}|} \qquad\qquad 0 \leq \alpha \leq 1 \tag{5.15}$$

For each trajectory the speed coherence β (change in magnitude) considers the geometric mean normalised by the arithmetic mean of the magnitude of consecutive displacement vectors $\boldsymbol{d}_{t,i}$ and $\boldsymbol{d}_{t+1,i}$. The deviation ranges between 1 for the same motion (high coherence) and 0 for high activity (low coherence), defined as:

$$\beta = 2\frac{\sqrt{|\boldsymbol{d}_{t,i}|\ |\boldsymbol{d}_{t+1,i}|}}{|\boldsymbol{d}_{t,i}| + |\boldsymbol{d}_{t+1,i}|} \qquad\qquad 0 \leq \beta \leq 1 \tag{5.16}$$

To demonstrate the TPC as a measure of trajectory activity, three different trajectories $\boldsymbol{T}_1$, $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$ are shown in Figure 5.9. The first $\boldsymbol{T}_1$, is a linear trajectory consisting of equal

displacements $\boldsymbol{d}(u, v)$ to provide a benchmark of trajectory coherence. The second and third, $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$, show periodic motion and provide an increasing deviation from $\boldsymbol{T}_1$. TPC results are reported in Table 5.2 with $\boldsymbol{T}_1$ having a total TPC of 1.00, hence perfect trajectory coherence. Trajectories $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$ produce a TPC of 0.73 and 0.50 respectively, with $\boldsymbol{T}_3$ showing the most deviation and activity for these example trajectories.

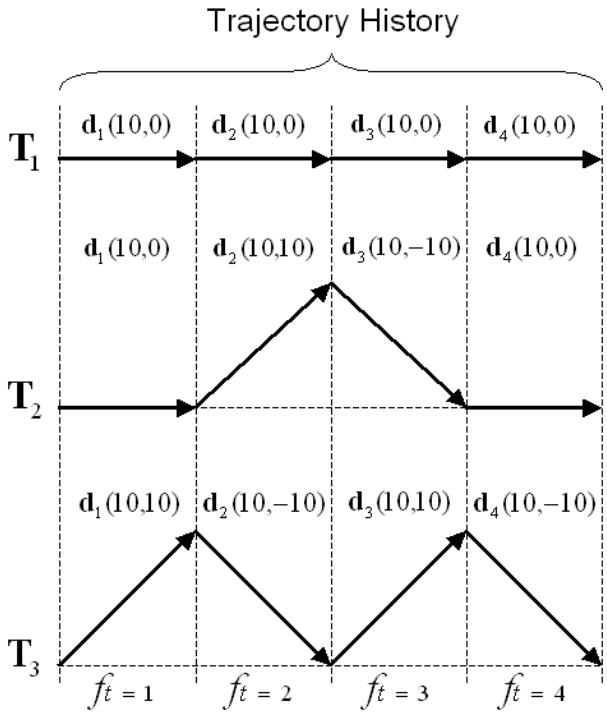


Figure 5.9: Trajectories $\boldsymbol{T}_1$, $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$ with a temporal displacement history of 4 frame pairs.

	$f_1 \rightarrow f_2$				$f_2 \rightarrow f_3$				$f_3 \rightarrow f_4$				
Trajectory	α	β	Φ	**TPC**	α	β	Φ	**TPC**	α	β	Φ	**TPC**	**TPC**
$\boldsymbol{T}_1$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.00**
$\boldsymbol{T}_2$	0.71	0.99	0.85	0.85	0.00	1.00	0.50	0.67	0.71	0.99	0.85	0.73	**0.73**
$\boldsymbol{T}_3$	0.00	1.00	0.50	0.50	0.00	1.00	0.50	0.50	0.00	1.00	0.50	0.50	**0.50**

Table 5.2: Trajectory path coherence for trajectories $\boldsymbol{T}_1$, $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$, with $\boldsymbol{T}_1$ with high temporal coherence and compared to $\boldsymbol{T}_3$ with less displacement coherence.

5.5 Displacement Field Smoothing

Once the combined matching method is applied we perform displacement post-processing. Spurious velocity vectors are inevitable from any tracking process and are not always obvious. Potential causes are from noise or artefacts where multiple block scales have insufficient encapsulated features. Using a coherence based post-processing algorithm, that is the adaptive weighted vector median filter (WVMF) [4], displacements are smoothed if inconsistent with their dominant neighbours whilst preserving motion boundaries.

Given n displacements inside a sliding window (of any size, we use 3×3) where neighbouring displacements are $j = 1, 2, \ldots, n$, the WVMF output is a median displacement vector $\boldsymbol{d}_{t,m}$:

$$\boldsymbol{d}_{t,m} = \text{med}\{\boldsymbol{d}_{t,j\ldots n}\} \tag{5.17}$$

Since the median value must actually be the value of one of the displacement vectors in the neighbourhood, the WVMF does not create new unrealistic values when the filter straddles an edge, preserving edge information.

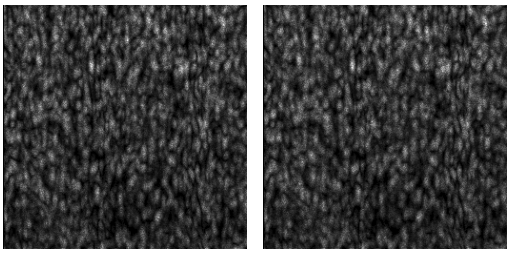
A displacement is substituted with $\boldsymbol{d}_{t,m}$ if the cumulative weighted p-norm distance between $\boldsymbol{d}_{t,m}$ and $\boldsymbol{d}_{t,i}$ is significant (less than the cumulative weighted p-norm distance between an individual $\boldsymbol{d}_{t,i}$ and neighbouring $\boldsymbol{d}_{t,j}$ displacements), expressed as:

$$\text{WVMF} = \begin{cases} \boldsymbol{d}_{t,m} & \text{if} \quad \sum_{i=1}^{n} w_i \|\boldsymbol{d}_{t,m} - \boldsymbol{d}_{t,i}\|_p \;\leq\; \sum_{i=1}^{n} w_i \|\boldsymbol{d}_{t,j} - \boldsymbol{d}_{t,i}\|_p \\ \boldsymbol{d}_{t,i} & \text{otherwise} \end{cases} \tag{5.18}$$

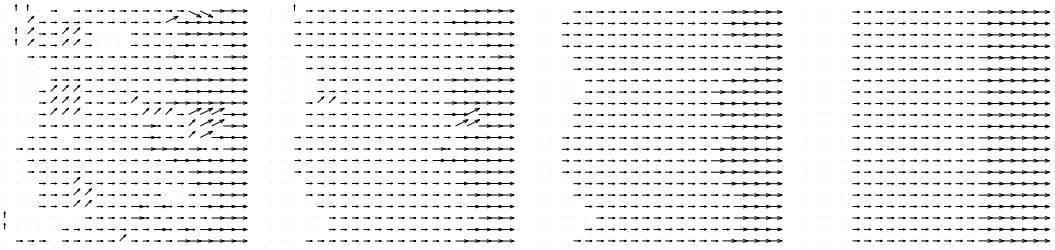
where $p = 2$ denoting the L2-norm. WVMF can be iterative for both interframe displacements and trajectory displacement smoothing with low computation. For our alternating measures approach the weighting uses either c_{max}, or double the CD_2 measure output (where 0.5 is the maximum value as shown in Figure A.1), so that $w \in [0, 1]$ preventing unbiased weighting.

Figures 5.10(a) and 5.10(b) are a corrupted synthetic frame pair (f_1, f_2), as defined in Chapter 3. An initial displacement field is shown in Figure 5.10(c), produced from the NCC matching measure with a block size of 8×8. This appears to have many erroneous displacements and several strong outliers. For this example, WVMF required 2 iterations, where Figures 5.10(d) and 5.10(e) are displacement fields produced from post-processing the original displacement field. Differing to the original displacement field, the final post-processed displacements are

only slightly different to the groundtruth in Figure 5.10(f). The maximum number of iterations was determined by using the DFD image reconstruction error, to prevent over-smoothing, which is detrimental to displacement accuracy. For *in vivo* sequences, we tend to use 2 iterations, unless otherwise stated.



(a) Synthetic f_1. (b) Synthetic f_2.



(c) #0. DFD= 200.3. (d) #1. DFD= 198.0. (e) #2. DFD= 197.1. (f) Groundtruth.

Figure 5.10: Post-processing: (a) Synthetic frames f_1 and (b) f_2, (c) original, (d,e) post-processed displacement fields, and (f) the groundtruth.

5.6 Combined Measure Improvement Analysis

In this section we demonstrate the improvement gained from alternating the matching measures, providing a comparative analysis of our proposed combined NCC/CD_2 approach against the CD_2, NCC and MSE measures. The evaluation of our method is demonstrated using *in vitro* and synthetic speckle data with displacement results using *in vivo* sequences shown later.

Initially, we tracked an *in vitro* region of tendon from the DYNAMIC-HEALTHY2 data. The absolute error between the groundtruth and mean estimated trajectories are summarised in Table 5.3. The NCC consistently produced significantly more accurate displacements compared to the single CD_2 and MSE measures. For example, the NCC produced a maximum error of 5.54 compared to 19.30 and 25.74 from CD_2 and MSE, for a pull of 10 mm. The maximum

error was noticeable near the end of each pull cycle, due to the clamped tendon not returning to its original resting state. The mean error remained low using NCC (2.71, 2.44 and 7.79 for pulls of 3, 6 and 10 mm respectively), and very similar to the combined measures approach (2.69, 2.45 and 7.79 for pulls of 3, 6 and 10 mm respectively). As mentioned later, due to the lack of speckle, using combined measures proved to be only as good as the single NCC measure. Typically, for *in vivo* data the speckle SNR instigates the usage of the NCC in minimal speckled (tendon) regions, and the alternative CD_2 measure in dense speckle areas.

Pull	**MSE Only**			**NCC Only**			**CD_2 Only**			**Combined**		
mm	**Max**	**Mean**	**STD Dev**	**Max**	**Mean**	**STD Dev**	**Max**	**Mean**	**STD Dev**	**Max**	**Mean**	**STD Dev**
3	4.79	3.60	3.53	2.16	2.69	1.21	5.57	3.88	2.23	2.16	2.71	1.22
6	8.97	3.78	4.46	2.62	2.45	1.39	5.32	4.58	2.13	2.63	2.44	1.37
10	25.74	10.35	7.96	5.54	7.79	3.12	19.30	9.00	4.11	5.55	7.79	3.12

Table 5.3: Summary of trajectory absolute error for *in vitro* tendon data for 3 pulls. Settings: WVMF iterations $= 2$ and multiple block scales where $M \times N = \{64, 32, 16, 8\}$.

Sample error results in Table 5.4 show the key improvement from our combined measures approach using synthetic data. Typically, this is observed in sequences with regions of varying speckle density (from multiple objects for example tendon and tissue, as shown in Chapter 3), whereby using the appropriate measure allows for local signal variation; this is important for *in vivo* analysis. Hence, for regions of homogeneous speckle (with a high or low density), the best accuracy is only as good as the appropriate single measure. Thus, Table 5.4 uses synthetic images described in Section 3.3.3, to control speckle density and to generate a groundtruth displacement field. An example of the combined measures approach improvement, is the low mean velocity angular error of 9.62 for varying density speckle, compared to 11.46, 11.66 and 13.48 for CD_2, NCC and MSE respectively.

Consequently, in Table 5.5 we show further results specifically for varying density speckle. These results illustrate a marked improvement compared to using a single measure. For example, for the best case (no applied noise), the combined measures approach produced a mean velocity angular error of 7.06, compared to 7.54, 7.64 and 7.30 for the CD_2, NCC and MSE measures respectively. For the worst case $\sigma = 0.8$, the combined measures approach produced a mean velocity angular error of 11.10, compared to 12.84, 12.10 and 21.28 for the CD_2, NCC

Measures	High Density Speckle (100%)			Low Density Speckle (20%)			Varying Density Speckle		
	Mean°	**STD Dev°**	**DFD**	**Mean°**	**STD Dev°**	**DFD**	**Mean°**	**STD Dev°**	**DFD**
Combined	7.13	6.46	9.22	13.78	15.85	6.56	9.62	10.62	7.48
CD_2 Only	7.13	6.47	9.22	15.95	16.63	7.55	11.46	13.91	9.33
NCC Only	7.39	6.60	9.23	13.72	15.83	6.56	11.66	13.42	8.10
MSE Only	7.58	7.45	13.27	23.74	24.55	14.20	13.48	18.05	10.15

Table 5.4: Interframe velocity angular error and displaced frame difference error. Settings: Affine deformation (4 pixels) sequences without noise using WVMF iterations $= 0$ and a block $M \times N = \{16\}$.

and MSE measures respectively. WVMF noticeably improved all results from just 2 iterations using an 8 neighbourhood region, reducing the velocity angular error, for example 7.06 with and 9.62 without. Similar results were obtained from testing over 100 frame pairs.

Measures	***Uncorrupted***		$\eta_m \sim \sigma = 0.4$		$\eta_m \sim \sigma = 0.8$	
	Mean°	**STD Dev°**	**Mean°**	**STD Dev°**	**Mean°**	**STD Dev°**
Combined	**7.06**	**6.09**	**7.41**	**7.02**	**11.10**	**12.44**
CD_2 Only	7.54	6.92	9.94	10.34	12.84	15.09
NCC Only	7.64	6.09	8.53	9.12	12.10	13.25
MSE Only	7.30	7.49	20.18	18.81	21.28	20.09

Table 5.5: Interframe velocity angular error for 3 cases of noise corrupted frames of varying density speckle. Settings: Affine deformation (4 pixels) sequences with varying noise using WVMF iterations $=$ 2 and a block $M \times N = \{16\}$.

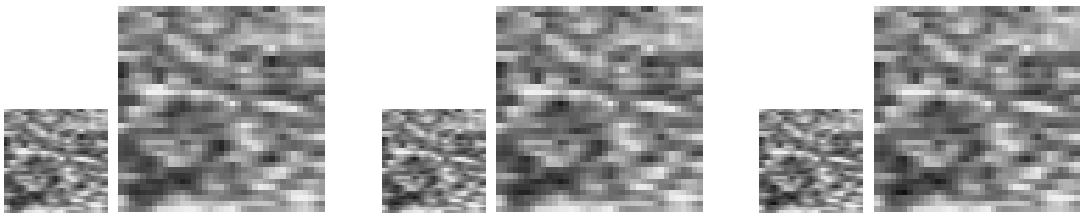
We have demonstrated that using a combination of speckle pattern similarity measures improved displacement performance, validating our approach on synthetic speckle and *in vitro* datasets. Also, we observed displacement accuracy is improved from analysing frames that contain sub-regions ranging from: dense speckle with characteristics that are purely multiplicative Rayleigh, to sparse stable speckle, to minimal speckle mixed with a strong underlying signal, all features typically found in *in vivo* data, as shown later in Table 5.8.

5.7 Proposed Trajectory Results

Subsequently, we evaluate the performance of the proposed method in Section 5.7.1, specifically in regions of fully developed speckle using the tissue mimicking phantom data. In Section 5.7.2 we proceed by establishing the tracking accuracy specifically in regions of *in vitro* tendon. Finally, in Section 5.7.3 we analyse *in vivo* data, with regions of tendon and tissue with potentially varying amounts of speckle. Further details for each datasets appears in Chapter 3.

5.7.1 Tissue Phantom Results

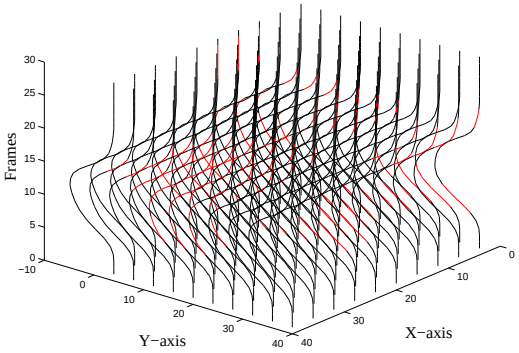
In these experiments we analyse 3 types of tissue mimicking phantom sequences ($\sim$, $\triangle$ and $\square$), each undergoing a compressive strain of 10%, capturing an absolute displacement of 0 to 11 pixels. To measure this movement, 100 trajectories (in a 10×10 region) were initialised at a resolution of $P\times Q$ where $P,Q=4$, tracking a region of 40×40 pixels at the same location for each experiment. These regions, shown in Figure 5.11, displayed very similar stable textures, even after compression and phantom initialisation periods.



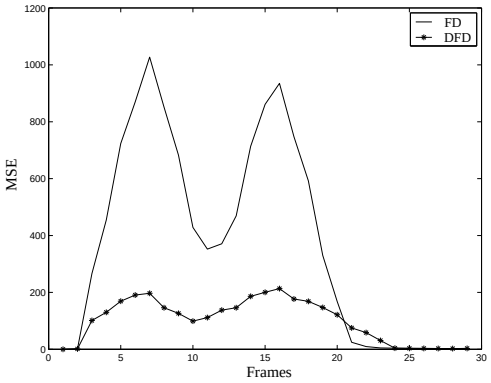
(a) Sinusoidal ($\sim$). (b) Triangular ($\triangle$). (c) Square ($\square$).

Figure 5.11: Actual and enlarged tracked speckle regions of the tissue mimicking phantom. ♠

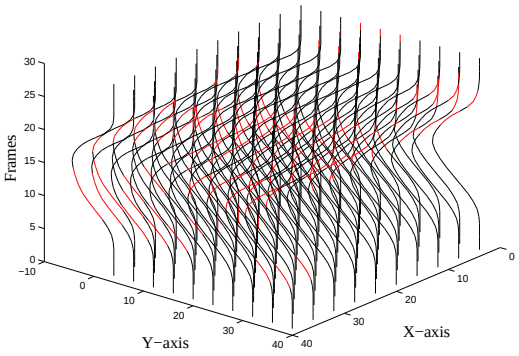
Each experiment used the same parameters: 3 template scales where M,$N=\{32,16,8\}$ and a single iteration of WVMF post-processing (preventing over-smoothing). Figure 5.12 illustrates sparse trajectory field results, with temporal displacements along each trajectory in red, benefiting from the proposed alternative matching measures approach. The sinusoidal, triangular and square displacements produced percentage usage ratios NCC:CD_2 of $89:11$, $87:13$ and $84:16$ respectively. The extremely stable speckle textures resulted in a high NCC usage resulting from an SNR >1.91, with correlation coefficients all >0.98.



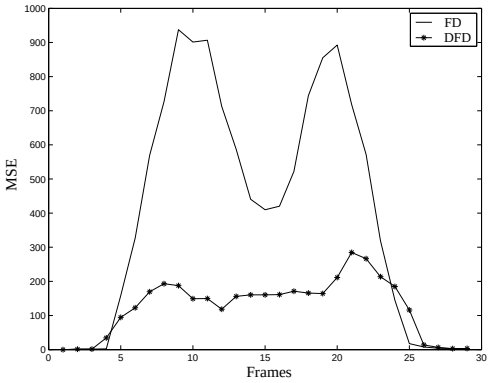
(a) Sinusoidal (∼).



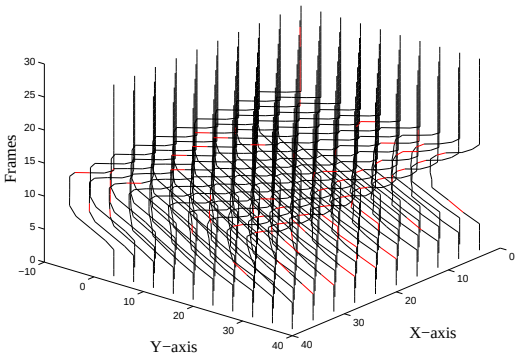
(a) Sinusoidal (∼).



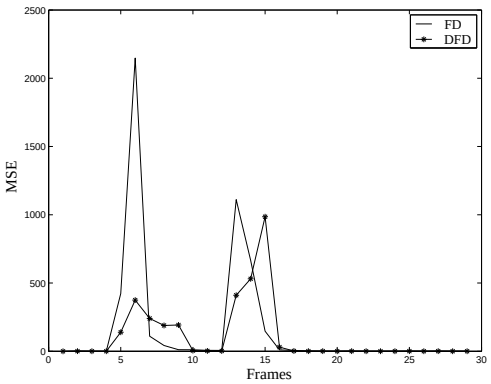
(b) Triangular (△).



(b) Triangular (△).



(c) Square (□).



(c) Square (□).

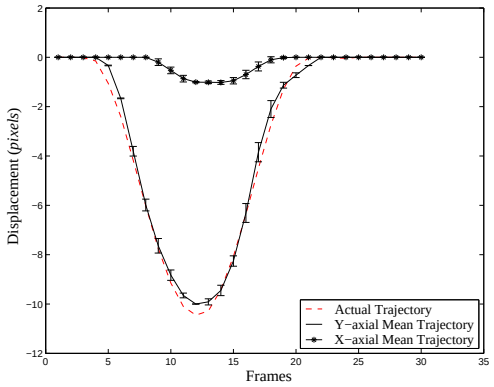
Figure 5.12: Sparse trajectory sets for the tissue mimicking phantom sequences. ♠

Figure 5.13: DFD and FD error for the tissue mimicking phantom sequences.

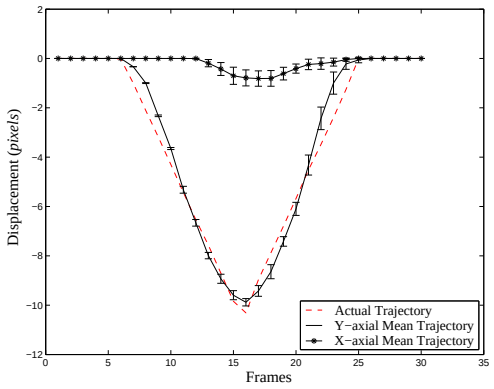
Figure 5.13 illustrates the DFD and FD displacement errors using the MSE criterion, following a set of trajectories. The errors are determined along the trajectory $(f_1, f_2), \ldots, (f_{t+n-1}, f_{t+n})$ instead of comparing to the first frame $(f_1, f_2), \ldots, (f_1, f_n)$, to prevent introducing warping errors. Trajectories before and after loading produced errors of $\text{MSE}_{\sim,\triangle,\square} \approx 0$, indicating perfect displacements. Trajectories during loading produced errors up to $\text{MSE}_{\sim,\triangle} \approx 200$, which is inevitable since speckle changes with tissue phantom motion. Further, two large spikes each of $\text{MSE}_{\square} \approx 400$ and $\text{MSE}_{\square} \approx 900$ present in the DFD results were due to blurring, as the large displacements occurred in small time intervals, shown at f_6 to f_8 and f_{14} to f_{16} in Figure 5.14(c). The FD troughs of $\text{MSE}_{\sim,\triangle}$ occurred due to frames at the maximum displacement being very similar, only different due to some negligible lateral motion. The DFD at these points showed a marked improvement. In comparison, the FD and DFD troughs of $\text{MSE}_{\square}$ were zero.

The mean trajectory and standard deviation for each trajectory field are plotted in Figure 5.14. In general, the estimated trajectories start and finish locations were equal with very similar peak magnitudes (see both x-axial and y-axial mean trajectories in Figure 5.14). This confirmed template drift updating was accurate, eliminating drift and potential strain bias. Each mean trajectory demonstrated a strong similarity to the 1D actual displacements recorded from the groundtruth. The centralised error-bars along each mean trajectory measure the standard deviation in each set of trajectories. Although small, as expected, the largest deviation tended to be at the peak. For instance, the proposed method quantified a mean estimated displacement of -10.5 ± 0.5 pixels. This indicates that the estimated value of the displacement lies between -10 and -11 pixels, compared to the groundtruth displacement of -11 pixels. Under perfect conditions each sequence captured only axial displacements, simulating a 1D movement. In the stationary lateral axis instead of zero (see x-axial mean trajectory in Figure 5.14), however, results showed a small but significant displacement. Although lateral axis displacements were not measured by the test rig, these displacements about the peak suggest that motion did occur.

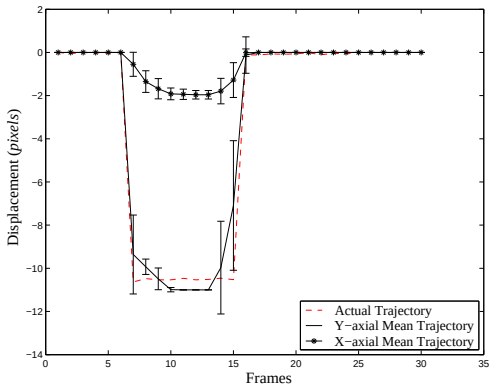
Figure 5.15 shows the ideal linear relationship of the actual trajectory compared to displacements along the estimated mean trajectory. All sequences showed that loaded (black stars) and unloaded (blue triangles) mean displacements correlated extremely well with the actual (red dashed line) displacements.



(a) Sinusoidal (∼).

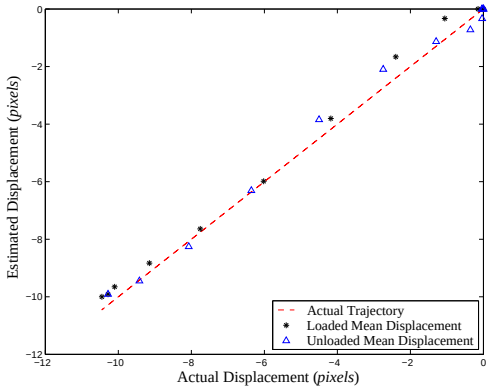


(b) Triangular (△).

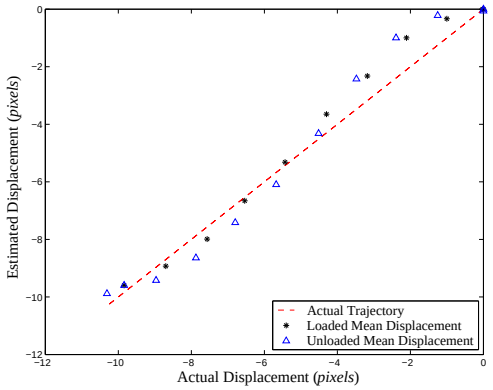


(c) Square (□).

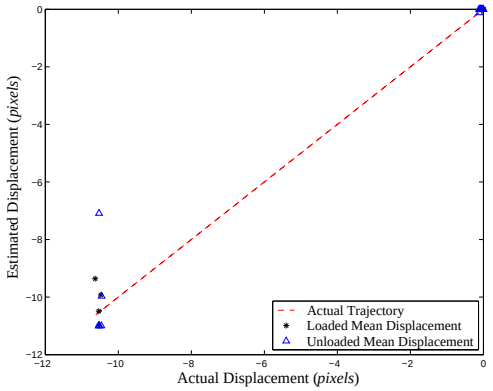
Figure 5.14: Trajectory statistics for the tissue mimicking phantom sequences.



(a) Sinusoidal (∼).



(b) Triangular (△).



(c) Square (□).

Figure 5.15: Estimated mean vs. actual trajectory for the tissue mimicking phantom sequences.

The estimated TPC values ranged between $0.9 \leq \Phi_{\sim} \leq 1.0$, $0.7 \leq \Phi_{\triangle} \leq 0.9$ and $0.9 \leq \Phi_{\square} \leq 1.0$ for the sinusoidal, triangular and square displacements. These were comparable to the actual groundtruth TPC values of $\Phi_{\sim,\triangle,\square} = 0.9$ to 1.0. The TPC values were compact and uniform, consistent with a single object under negligible strain. With higher strain, the TPC and standard deviation increase (shown as error-bars in Figure 5.14), assuming trajectories have zero error.

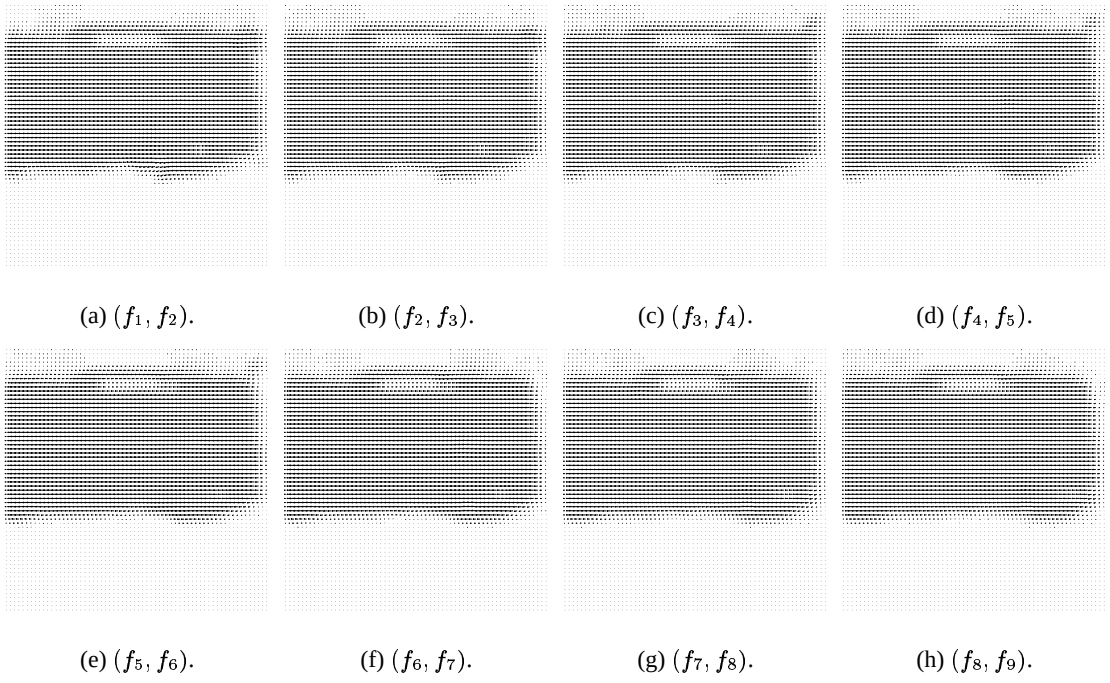
5.7.2 *In vitro* Tendon Results

To provide a comparison to the displacement accuracy results in the previous chapter in Section 4.4, we used our final proposed approach to analyse the same frames (f_1, f_2) to (f_9, f_{10}) of the STATIC-LANDMARK0 sequence, for interframe analysis. Each frame pair captured the same 8 pixel lateral translational shifts, hence we expected resulting measurements to produce similar results, which are shown in Table 5.6. The proposed method produced consistently better performance for all cases, compared to the results from earlier methods in Tables 4.2 and 4.3, which is indicated by all DFD values, global frame correlations (CORR) and velocity angular errors being superior. For example, a DFD of 34.48 at f_1 to f_2 was significantly lower than the FD of 273.46, and DFDs of 84.78, 55.33, 214.74 and 106.63 from FSBM, VSBM, Camus and Proesmans et al. respectively. Example interframe displacement fields for (f_1, f_2) to (f_8, f_9) are shown in Figure 5.16 from the proposed method.

Frame	**PRE-WARP**		**Proposed Approach**			
Pair	**FD**	**CORR%**	**DFD**	**CORR%**	**Mean°**	**STD Dev°**
$f_1 \rightarrow f_2$	273.46	97.40	**34.48**	99.67	6.91	7.69
$f_2 \rightarrow f_3$	270.28	97.44	**32.96**	99.69	6.98	8.05
$f_3 \rightarrow f_4$	272.57	97.41	**33.62**	99.68	7.15	7.94
$f_4 \rightarrow f_5$	273.87	97.40	**36.10**	99.66	7.33	8.40
$f_5 \rightarrow f_6$	271.60	97.43	**33.63**	99.68	6.44	7.93
$f_6 \rightarrow f_7$	274.65	97.40	**32.27**	99.68	7.65	7.70
$f_7 \rightarrow f_8$	273.83	97.40	**36.20**	99.66	7.23	7.70
$f_8 \rightarrow f_9$	273.87	97.40	**36.72**	99.65	7.16	7.69
$f_9 \rightarrow f_{10}$	277.84	97.53	**35.39**	99.66	8.30	7.28

Table 5.6: Error measures for proposed approach using STATIC-LANDMARK0.

Furthermore, the purpose of analysing *in vitro* tendon sequences was to evaluate the feasibility of quantifying lateral tendon displacements specifically for strain estimation using trajectories. We present results for a 30-frame sequence of tendon deformation, undergoing a tensile strain



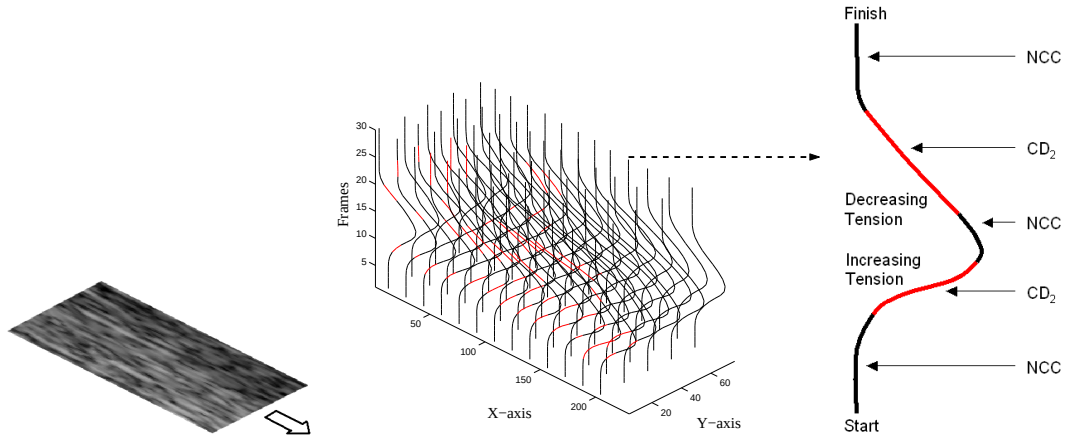
(a) (f_1, f_2). (b) (f_2, f_3). (c) (f_3, f_4). (d) (f_4, f_5).

(e) (f_5, f_6). (f) (f_6, f_7). (g) (f_7, f_8). (h) (f_8, f_9).

Figure 5.16: Displacement fields for (f_1, f_2) to (f_8, f_9) using STATIC-LANDMARK0.

of 3.5%. This encapsulated a recorded groundtruth displacement of 0 to 30 pixels. An example of a tracked region is shown in Figure 5.17(a); in comparison to the tissue mimicking phantom these textures are very different, yet both are visible in musculoskeletal B-scans due to the underlying signal. We initialised 250 trajectories (in a 25×10 region) at a resolution of $P \times Q$ where P,$Q = 8$ tracking a region of 200×80 pixels at the same location for each experiment. Parameters included: 4 template scales where M,$N = \{64, 32, 16, 8\}$ and 3 iterations of WVMF displacement smoothing. Larger scales were used to ensure the displacement range was suitable. Figures 5.17(b) and 5.18(a) illustrate trajectory results where a small minority of temporal displacements along each trajectory are red, to represent CD_2 and the rest black, representing NCC, indicating a usage percentage ratio of NCC:$CD_2 = 93:7$.

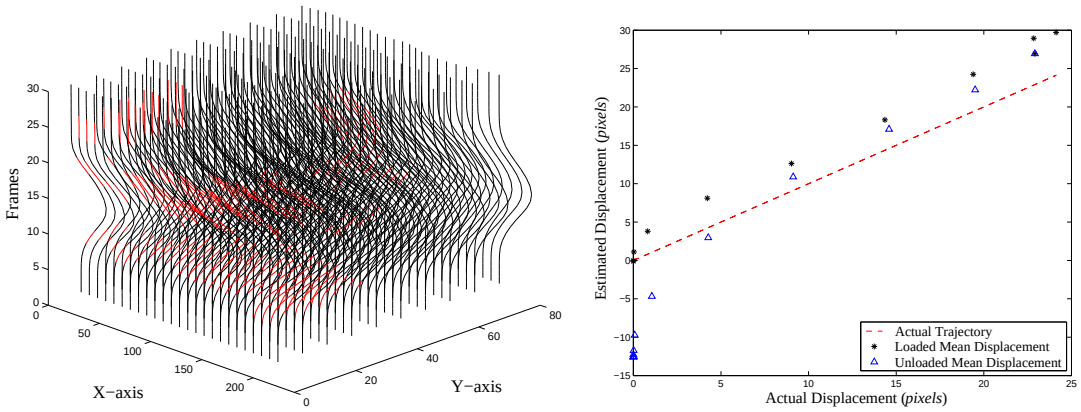
Figure 5.18(b) shows the ideal linear relationship of the actual trajectory compared to displacements along the estimated trajectory. Displacements showed loaded (black stars) and unloaded (blue triangles) mean displacements correlated well with the actual displacements (red dashed line). Two typical inaccuracies were observed during experimentation: first, tendon slippage, tendon straightening (uncrimping) and plastic deformation, all contributing to



(a) Tendon ROI. (b) Sparse trajectory field.

Figure 5.17: Tendon *in vitro* results: (a) Example *in vitro* tendon region and (b) sparse trajectory field.

the tendon failing to return to its original starting position; second, matching accuracy reduced from increasing dissimilarity with larger deformations. We observed that deformations larger than 5% produced inconsistent accuracy. TPC values of $0.6 \leq \Phi \leq 0.77$ resembled the actual groundtruth of 0.6, showing more activity than the phantom results. Further, TPC values gradually increased along the loaded lateral axis and were uniform along the axial axis, corresponding to the applied strain.

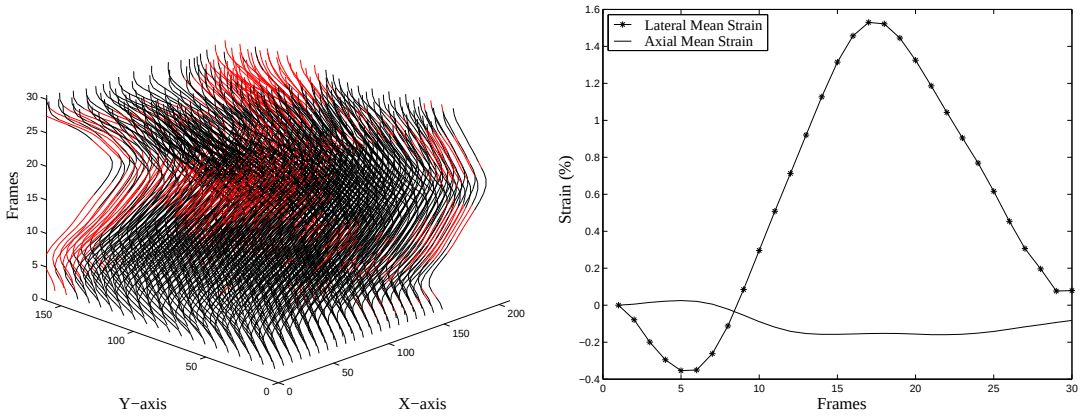


(a) Full trajectory field. (b) Trajectory statistics.

Figure 5.18: Tendon *in vitro* results: (a) Full spatiotemporal trajectory field and (b) estimated mean trajectory vs. actual trajectory relationship.

5.7.3 *In vivo* Musculoskeletal Results

The potential role of this research for *in vivo* elastography is demonstrated by analysing musculoskeletal sequences. A range of clinically relevant data was processed producing trajectory fields and strain, including sequences capturing the achilles, patella, radial flexor tendons and median nerve, only differing to tendons in size. A typical trajectory field and corresponding strain plot are illustrated in Figure 5.19, for a ROI during a complete sequence of resting-flexion-extension-resting motion. Strain increased and decreased with the kinematic motion of the tendon moving from flexion f_1 to f_{15}, to extension f_{16} to f_{30}. Strain magnitude was generally small as no loading was applied for these tests. As expected, results corresponded to our *in vitro* tendon pull groundtruth (as in Figure 5.18(a)). Strain computation is addressed further in Chapter 6, but briefly referred to here for clarity.



(a) Sparse trajectory field.

(b) Strain.

Figure 5.19: Example *in vivo* results: (a) Sparse trajectory field (usage percentage ratio where NCC:CD_2 is $78:22$), and (b) the derived strain.

Further *in vitro* and *in vivo* displacement accuracy results are summarised in Tables 5.7 and 5.8 respectively. Results are compared to the relevant methods mentioned in Chapters 2 and 4, namely single scale template region matching using the NCC correlation [107][105][87] and CD_2 [20] measures. With no *in vivo* groundtruth data, displacement field accuracy is determined solely from the DFD, measured by the MSE. Although not optimal, as it is possible that inaccurate displacements can produce a low DFD error in areas of similar texture (tendon aperture problem), the DFD has proven to be a good indicator, also the CORR value between

the displaced frame and f_{t+1} complements the DFD. The first column contains the FD error, as a measure to compare against each DFD result. The proposed technique outperforms both NCC and CD_2 in most cases, producing consistently lower DFD than FD errors (for example, a DFD of 278.6 compared to a FD of 423.7, and 372.7 and 302.3 for NCC and CD_2 respectively). The percentage usage ratios of NCC:CD_2 for *in vivo* cases ranged between $32:68$ and $86:14$. The NCC appeared less accurate than the proposed method, due to frames capturing a small region of tendon in comparison to speckle and tissue. These results show minor improvements compared to the other methods, but are clinically significant, as errors in displacements directly reduce strain precision.

	PRE-WARP		**NCC Only**		**CD_2 Only**		**Proposed Approach**		
Datasets	**FD**	**CORR %**	**DFD**	**CORR %**	**DFD**	**CORR %**	**DFD**	**CORR %**	**NCC:CD_2**
Phantom $\sim$	147.7	97.55	50.6	99.19	51.5	99.18	**41.2**	**99.32**	92:8
Phantom $\triangle$	104.3	98.29	48.1	99.21	40.6	99.34	**26.4**	**99.58**	91:9
Phantom $\square$	237.4	96.21	401.9	94.60	206.1	96.84	**184.8**	**97.09**	92:8
Tendon *in vitro*	152.3	99.11	241.3	98.62	197.4	98.85	**136.5**	**99.21**	89:11
Tendon *in vitro*	226.6	98.66	215.4	98.74	178.5	98.95	**143.1**	**99.16**	91:9

Table 5.7: Image reconstruction error measures for tissue phantom and *in vitro* tendon. Results are shown for a single scale template region matching using only NCC and only CD_2, and the proposed method using selective use of both measures. The final column shows the usage ratio of these. Frame size was 512×512 pixels.

5.8 Conclusion

We have investigated an automated approach for estimating ultrasound speckle motion, developing a region based method for computer vision elastography. From a systematic validation study analysing tissue mimicking phantom, *in vitro* tendon and *in vivo* clinical musculoskeletal datasets, a range of displacements, trajectory fields and accuracy results has been presented.

Prior to speckle motion estimation, frame registration corrected for any transducer motion, allowing for automatic non-stationary frame calibration, removing the need for traditional transducer clamping apparatus used in RF elastography. The estimated displacements located in the

Datasets	PRE-WARP		NCC Only		CD_2 Only		Proposed Approach		
	FD	CORR %	DFD	CORR %	DFD	CORR %	DFD	CORR %	NCC:CD_2
Achilles tendon	333.2	97.93	341.1	97.69	277.4	98.10	**239.0**	**98.37**	53:47
Achilles tendon	384.2	97.91	437.0	97.60	**324.3**	**98.23**	352.7	98.08	52:48
Achilles tendon	441.4	97.82	422.6	97.93	322.8	98.39	**312.3**	**98.44**	53:47
Achilles tendon	423.7	97.23	372.7	97.57	302.3	98.02	**278.6**	**98.18**	44:56
Patella tendon	208.0	97.98	194.6	98.01	159.7	98.35	**131.1**	**98.65**	37:63
Patella tendon	208.0	97.93	200.5	98.01	167.2	98.33	**133.6**	**98.67**	48:52
Patella tendon	151.8	98.61	218.8	98.02	191.2	98.24	**138.2**	**98.73**	86:14
Patella tendon	266.7	98.34	332.2	97.94	277.3	98.27	**261.6**	**98.37**	61:39
Radial tendon	86.3	99.43	161.6	98.95	123.8	99.18	**84.2**	**99.44**	72:28
Radial tendon	130.5	98.50	192.3	97.83	140.9	98.37	**125.8**	**98.55**	53:47
Radial tendon	212.4	98.10	233.0	97.91	**186.0**	**98.33**	206.9	98.16	58:42
Radial tendon	358.3	95.20	347.0	95.23	323.6	95.55	**278.9**	**96.11**	64:36
Median nerve	349.0	95.20	433.2	94.04	429.2	94.04	**397.6**	**94.49**	52:48
Median nerve	119.3	99.02	162.6	98.67	141.6	98.84	**118.8**	**99.03**	32:68

Table 5.8: Image reconstruction error measures for *in vivo* musculoskeletal data. Results are shown for a single scale template region matching using only NCC and only CD_2, and the proposed method using selective use of both measures. The final column shows the usage ratio of these. Frame size was 512×512 pixels. Best results are shown in bold.

region between the skin and transducer in the B-scans, were used for the global registration of all displacement fields. Using displacements in this landmark region, proved similar to measuring the displacements in a landmark region created from attaching an elastic layer offset to the transducer or target, for example [70].

The accuracy of speckle motion estimation is dependent on the image quality of the registered sequences. We have shown that using multiscale template analysis, with ordered and overlapping sets of detectable velocity ranges, and alternative speckle similarity measures, proved effective for a wide range of datasets. We observed that the speckle SNR made certain that NCC usage was dominant in tendon regions, where speckle density was negligible (non-Rayleigh). Also, the speckle SNR ensured CD_2 usage was dominant in deeper lower frequency regions, where speckle was dense (Rayleigh). With varying ultrasound image content, the benefit from using automatically selected similarity measures, is the improved local displacement accuracy

from tracking a region with the appropriate measure, determined by local image content. In situations where the SNR value may not reliably determine the correct measure to use, a reasonable displacement is still produced that is improved by multiple block scales. Finally, we noted that non-optimal transducer gain selection created areas of low inhomogeneous contrast, which for the NCC produced relatively flat correlation surfaces with negligible peaks.

Physical factors inducing drift include an inhomogeneous strain distribution, target object angular changes and a low frame rate, causing inaccurate displacements. It was also observed, however, that trajectory drift updating successfully eliminated displacement drift error. D'Hooge et al. [24] observed similar drifting inaccuracies when measuring cardiac strain, differing by compensating for drift by assuming strain boundary conditions. This study also highlighted that interframe pixel shifts of < 1 and > 30 proved difficult to track, with large deformations producing weak matching, as the assumption of pixel constant intensity failed.

Performance has been measured using the angular velocity error and image reconstruction DFD error. For *in vivo* data the DFD provides a reconstruction quality measure, which is assured by the fact that quality ranking based on the DFD is quite consistent with the global correlation coefficient. In Section 5.7.2 a comparison is made between our approach and the previous work in Chapter 4. Using our *in vitro* data we have shown improved accuracy for all results with this technique, and also shown low errors using *in vivo* data. This work has illustrated that image based methods provide accurate measurements for a range of displacements in both axial and lateral dimensions. Following on, in the next chapter we use our approach as detailed here, to produce displacements and trajectory fields for determining spatial and temporal strain maps. Using several *in vivo* sequences, we illustrate the process of using 2D speckle tracking for evaluating the biomechanical properties of tendons and soft tissue.

Chapter 6

Strain Estimation from Trajectories

6.1 Introduction

In this chapter a practical alternative to elastography is presented. We generate temporal axial and lateral strain maps derived from our trajectories that measure 2D speckle movement, and provide a novel spatiotemporal strain history for a target subject over time. In chapter 2 we described traditional RF elastography, which generates elastograms quantifying 1D strain from an applied displacement using ultrasound, displaying the longitudinal strain as the axial component of the estimated displacement gradient. Typically, displacements are derived from 1D cross correlation of the RF echo arrival times, pre and post-deformation. Unfortunately, RF elastography drawbacks include the inability to record and store RF data, and limited to measuring small *in vivo* strains of only $< 1\%$, both restricting clinical usability. In this chapter we demonstrate the advantages of an image based approach.

We generate spatiotemporal 2D strain maps, quantifying the strain for a given subject during an input video sequence. This uses the motion trajectories described in the previous chapter, which measure the dynamic responses of musculoskeletal soft tissues and tendons under kinematic loading. We are not aware of any existing research that attempts to quantify a strain history for *in vivo* data, particularly for regions of soft tissue or tendon.

A knowledge of tissue motion is used to derive a second-order strain tensor and strain rate, producing a spatiotemporal visualisation of the applied strain, locally and globally, throughout

time. Providing a 2D temporal visualisation of tissue mechanical properties, has the potential to enable clinicians to uniquely visualise and compare both the dynamic behaviour and axial and lateral strain histories simultaneously in a *non invasive* approach. Furthermore, the potential benefits of using ultrasound image analysis as a tool to measure movement and strain will be discussed. Ultimately, we establish trajectories and elastograms for clinical *in vivo* examples.

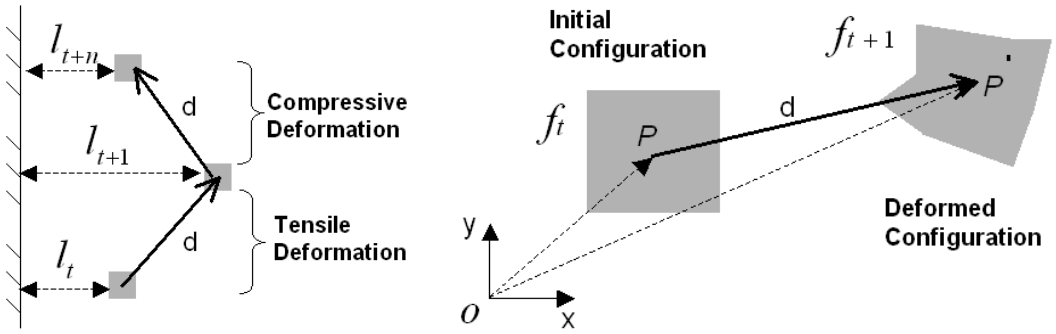
Our results show that the proposed method measures axial and lateral strains of $\varepsilon < 3.5\%$ with good precision for *in vitro* data, with strain accuracy determined by comparisons to the groundtruth 1D strain and displacement error measures. For *in vivo* data, strain accuracy is determined by the image reconstruction error and expert opinion. Additionally, for *in vivo* data, we compare the accuracy of the strain tensor with another strain estimator, decorrelation coefficients, which assume correlation decreases as strain increases. Results conclude that decorrelation coefficients as a strain estimator, are less effective than our strain tensor approach.

The structure of this chapter is as follows. In Section 6.2 we define and formalise stress and strain for 1D analysis, and the 2D strain tensor and temporal strain history for 2D analysis. Next, in Section 6.3, we present strain results for the tissue mimicking phantom and *in vitro* groundtruth data. We then focus on estimating trajectory fields and strain results for *in vivo* musculoskeletal sequences of the patella, achilles and digital flexor tendons, with image reconstruction errors. Then, in Section 6.4, we explore the decorrelation coefficient as a measure of strain. Finally, the chapter is concluded in Section 6.5.

6.2 Strain Estimation from Displacements and Trajectories

Tendons form an integral part of the musculotendinous unit, transmitting the tensions generated in muscle to bone. When the tendon is under load, the deformation is expressed as a strain (the change in linear dimension per unit length), or a stress (the force applied per unit area), with Hooke's Law defining that stress is proportional to strain. Tendon mechanical behaviour is typically influenced by its hierarchical structure, history, strain rates and strain range. In this work, we are concerned with only measuring strain, as this is equivalent to elastography techniques that derive strain from a knowledge of displacements. An important by-product of using the proposed speckle tracking methodology is the possibility of generating 2D strain tensor fields.

In the following sections we define strain from using displacements and trajectories, with examples of both 1D and 2D deformation for an elastic subject shown in Figure 6.1. A reason for quantifying displacement and strain is to categorise tissues based on mechanical properties. For example, soft tissues (such as muscles, fat, cartilage and fibrous tissues) undergo nonuniform displacement or deformation supporting larger strains than hard tissues (such as cancerous tissues, cysts and tumours), which undergo uniform displacement and smaller strains. This assumption can form part of a segmentation scheme that is discussed later in Chapter 7.



(a) 1D Deformation.

(b) 2D Deformation.

Figure 6.1: Schematics showing a local: (a) Point, for 1D elastography and (b) region for 2D B-scans.

6.2.1 1D Axial Strain Analysis

Extension ratios (known as the engineer's strain) are a dimensionless strain value measuring relative movement $\hat{l}/l$, where l and $\hat{l}$ are the original and changed lengths respectively. This leads to 1D Lagrangian strain ε for static object deformation:

$$\varepsilon = \frac{\hat{l}_t - l_{t=0}}{l_{t=0}} = \frac{\hat{l}}{l} - 1 \tag{6.1}$$

where the original length $l_{t=0}$ remains constant for all t. As the tissue phantom and *in vitro* tendon are static objects all motion is deformation as a result of the applied load. For dynamic objects, however, the original length is unknown and varies during the deformation process. By expressing the changed length relative to the length at a previous instant in time, we define the Eulerian strain ε':

$$\varepsilon' = \frac{\hat{l}_{t+\delta t} - l_t}{l_t} \tag{6.2}$$

Finally, temporal integration can produce the total 1D or 2D strain during the time interval.

Currently, RF elastography uses the change in distance between two specific RF signals s_1 and s_2, to directly estimate local strain as the spatial gradient in motion of the RF signal. Using 1D windows positioned near (proximal a) and remote (distal b) to the point of origin, strain is given as:

$$\varepsilon'(s_{1,a}) = \frac{(s_{2,b} - s_{2,a}) - (s_{1,b} - s_{1,a})}{s_{1,b} - s_{1,a}} \tag{6.3}$$

where $s_{1,a}$ and $s_{1,b}$ are the pre-compression echo arrival times from the proximal and distal windows respectively, and $s_{2,a}$ and $s_{2,b}$ are the post-compression echo arrival times from the proximal and distal window respectively.

It must be noted that different strain measures exist (for example Lagrangian and Logarithmic), which produce different results as strains increase. It is most common to use Lagrange strain when solving problems involving large deformations $\varepsilon > 5\%$.

6.2.2 2D Axial and Lateral Strain Analysis

In 2D object deformation there are 4 strain components, axial and lateral components (motion normal to the object border) and shear components (motion parallel to the object border). Second-order strain tensors can describe all axial, lateral and shear strain components, completely defining the state of strain. The Green-Lagrange strain tensor [27], $\boldsymbol{E}$, is a measure of strain that is invariant under rigid body rotations, defined as:

$$\boldsymbol{E} = \frac{1}{2}\left(\boldsymbol{F}^T \cdot \boldsymbol{F} - \boldsymbol{I}\right) \tag{6.4}$$

where $\boldsymbol{F}$ is the deformation gradient and $\boldsymbol{I}$ is the second order identity matrix. We use each displacement vector $\boldsymbol{d}_t(u, v)$ to define the linear Green-Lagrange strain tensor, for small deformations. The tensor defines the two 2D normal directional strains ε_{xx} and ε_{yy}, and the two 2D shear strains ε_{xy} and ε_{yx} for displacement i:

$$\boldsymbol{E}_i = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} \\ \varepsilon_{yx} & \varepsilon_{yy} \end{bmatrix} = \begin{bmatrix} \frac{\partial u_i}{\partial x} & \frac{1}{2}\left(\frac{\partial u_i}{\partial y} + \frac{\partial v_i}{\partial x}\right) \\ \frac{1}{2}\left(\frac{\partial v_i}{\partial x} + \frac{\partial u_i}{\partial y}\right) & \frac{\partial v_i}{\partial y} \end{bmatrix} \tag{6.5}$$

Each component can be derived using (6.1) for static subjects or (6.2) for dynamic subjects. For example, the ε_{xx} and ε_{yy} Lagrangian strain components are:

$$\varepsilon_{xx} = \frac{\partial u_i}{\partial x} = \frac{\hat{u_t} - u_{t=0}}{x_{t=0}} \qquad \text{and} \qquad \varepsilon_{yy} = \frac{\partial v_i}{\partial y} = \frac{\hat{v_t} - v_{t=0}}{y_{t=0}} \tag{6.6}$$

For objects that are homogeneous, isotropic and incompressible, the shear strain components are defined as $\varepsilon_{xy} = \varepsilon_{yx}$. Consequently, by implying that both shear components are complementary, the number of components can be reduced. Finally, in practise the derivatives are estimated using the Sobel operator on the WVMF smoothed displacement fields.

The strain tensor defined in (6.5) is suitable for small strains of $\varepsilon < 1\%$ for *in vivo* analysis, however, for large deformation applications the strain tensor $\boldsymbol{E}$ becomes non-linear, containing quadratic terms. These are derived in [27] using the Taylor series expansion, and given as:

$$\begin{aligned} \varepsilon_{xx} &= \frac{\partial u_i}{\partial x} + \frac{1}{2}\left[\left(\frac{\partial u_i}{\partial x}\right)^2 + \left(\frac{\partial v_i}{\partial x}\right)^2\right] \\ \varepsilon_{yy} &= \frac{\partial v_i}{\partial y} + \frac{1}{2}\left[\left(\frac{\partial u_i}{\partial y}\right)^2 + \left(\frac{\partial v_i}{\partial y}\right)^2\right] \\ \varepsilon_{xy} = \varepsilon_{yx} &= \frac{1}{2}\left(\frac{\partial u_i}{\partial y} + \frac{\partial v_i}{\partial x}\right) + \frac{1}{2}\left[\left(\frac{\partial u_i}{\partial x}\right)\left(\frac{\partial u_i}{\partial y}\right) + \left(\frac{\partial v_i}{\partial x}\right)\left(\frac{\partial v_i}{\partial y}\right)\right] \end{aligned} \tag{6.7}$$

Existing research to illustrate tensor information includes Weinstein et al. [108] and Sigfridsson [89], constructing tensor glyphs, which are geometric representations of the tensor. Typically, ellipses are formed by multiplying points on a circle by the tensor. However, overlapping ellipses become rapidly confusing for dense displacement fields. Alternatively, the components of any second-order tensor can be reduced to an eigenvalue problem, where eigenvalues are a scalar representation that are invariant to direction, quantifying the amount of strain. Splitting the greatest and smallest eigenvalues signifies either tension or compression respectively.

Herein, we prefer to use the Frobenius norm of the strain tensor $\boldsymbol{E}$ to illustrate our strain results, which is invariant to direction and both normal or shear strain, quantifying a single value of strain magnitude, where Tr is the matrix trace, defined as:

$$\|E\|_F = \sqrt{Tr(\boldsymbol{E}^T\boldsymbol{E})} \tag{6.8}$$

Therefore, each value shown in the resulting strain map represents the square root of the sum of all $\boldsymbol{E}$ terms squared, which is determined from a single displacement along a trajectory, mapping a reference template from the first frame through a sequence.

6.2.3 Temporal Strain History

Until now, strain estimation from B-scans has been restricted to only single instances, due to only analysing manually chosen frame pairs, a limitation prevalent in many works [12]. In this work, trajectories are used to quantify a novel spatiotemporal *strain history* that determines a unique axial and lateral strain for every position in f_t through a sequence, measuring strain variations during a target subject's deformation. Consequently, using the strain definition in (6.5), trajectories quantify either accumulated strain for a region tracked from the first frame (relating f_t to f_{t+n}), or incremental strain for a region tracked between frame pairs (relating $f_t, f_{t+1} \dots f_{t+n-1}, f_{t+n}$). These provide unique measurements as a direct result of our motion estimation approach, with *in vivo* results shown in Section 6.3.3.

6.3 Video Sequences, Trajectory Fields and Strain Maps

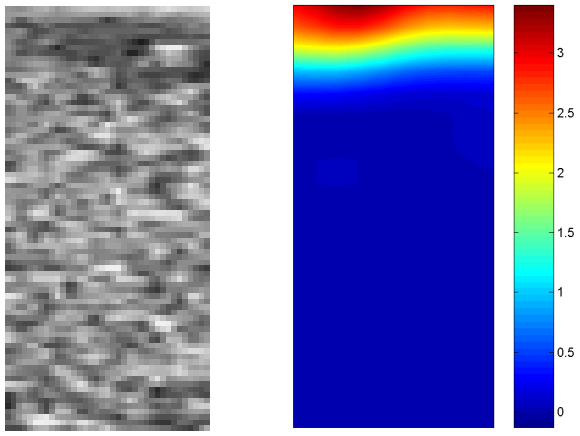
The aim of this section is to demonstrate the complete process of measuring the biomechanical properties of musculoskeletal data using ultrasound video sequences. We first present tissue mimicking phantom and *in vitro* groundtruth strain results, analysing displacements from the previous chapter. We then show sample frames, trajectory fields and strain maps for *in vivo* sequences of the patella, achilles and digital flexor tendons.

6.3.1 Tissue Phantom Results

The tissue mimicking phantom was described in Chapter 3, undergoing a compressive strain of 10%, capturing an absolute displacement of 0 to 11 pixels. The accuracy of the trajectories measured from the tissue mimicking phantom, were shown in Figures 5.13, 5.14 and 5.15. In summary, the proposed method quantified a mean estimated displacement of -10.5 ± 0.5 pixels at the peak displacement, indicating that displacement accuracy at this point was ± 0.5. In this section, we use these displacements to evaluate strain.

Figure 6.2(a) shows a complete axial region of the phantom, alongside its corresponding elastogram in Figure 6.2(b), at the peak displacement. The phantom elastic properties meant deformation was immediately beneath the transducer during contact. The elastogram quantifies a

localised compressive strain concentration in this region of $\varepsilon = 3.2\%$ at the point of load, with the deformation occuring in a minimal area. The remaining phantom displacement was equivalent to the movement of the transducer, and correctly shown as zero strain in the elastogram.



(a) Region. (b) Strain.

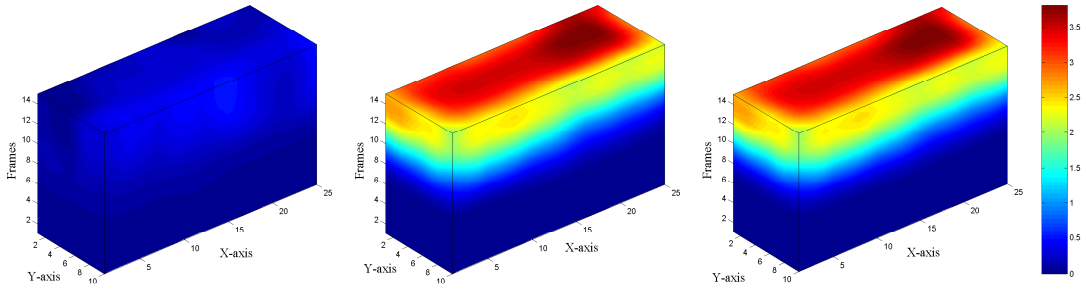
Figure 6.2: Elastogram for an ROI of 40×80 pixels using trajectories in a 10×20 region, for f_1 to f_{13}.

6.3.2 *In vitro* Groundtruth Tendon Results

Additionally, we analysed the *in vitro* tendon sequences to evaluate the feasibility of quantifying lateral tendon strain. We present axial and lateral strain results for a 30-frame sequence of tendon deformation, undergoing a tensile strain of 3.5%, with comparison to a groundruth. The accuracy of the trajectories used were shown previously in Figure 5.18, showing that the proposed method quantified a mean error of $+4$ pixels at the peak of the recorded groundtruth.

Spatiotemporal elastograms are presented in Figure 6.3 showing axial, lateral and strain magnitude results, from using the full trajectory fields generated in Figure 5.17. Each quantifies the applied tensile strain between f_1 to f_{15}, up to the peak displacement. The size and location of the tracked region of tendon result in spatially uniform strains. Axial strains in Figure 6.3(a) show a peak tensile strain of $\varepsilon = 0.3\%$ at f_{15} with strains in general remaining insignificant. Lateral strains in Figure 6.3(b) show a peak tensile strain in red of $\varepsilon \approx 3.5\%$ at f_{15}, showing reasonable spatial uniformity, correlating well with the applied strain shown in Figure 6.4. Both lateral and strain magnitude elastograms correctly show strain incrementing in time. The strain magnitude (normalised strain tensor combining axial, lateral and shear components) in

Figure 6.3(c), shows that the negligible axial strain had little effect on the overall strain.



(a) Axial strain. (b) Lateral strain. (c) Strain magnitude.

Figure 6.3: Spatiotemporal elastograms (axial, lateral and strain magnitude) for f_1 to f_{15}.

Figure 6.4 illustrates the actual strain and estimated mean lateral and axial strain for the tracked region of tendon. This demonstrates a lateral strain error -0.25% at the peak.

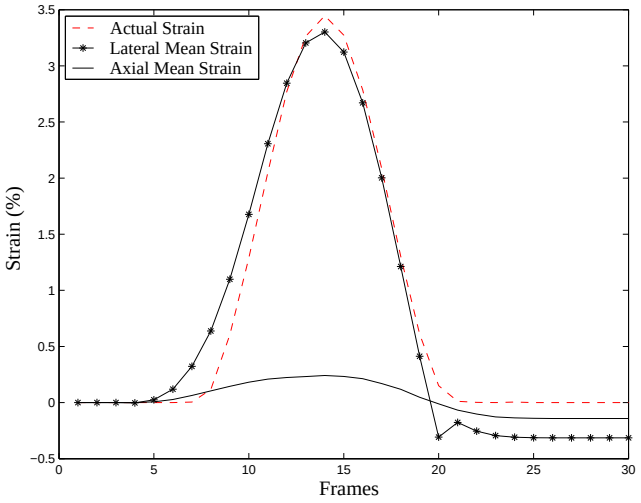


Figure 6.4: Tendon *in vitro* estimated axial and lateral strains using the mean trajectory.

6.3.3 *In vivo* Musculoskeletal Results

In Figures 6.6, 6.10 and 6.14 we show sample frames from our *in vivo* database for a patella, achilles and digital flexor tendons respectively (described in Section 3.4). Each sub-figure corresponds to frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. The maximum displacement and strain potentially exists at the limit of the kinematic movement. All results use the proposed approach with a resolution of $P \times Q$ where $P, Q = 8$ and use multiple block scales $M \times N$, where $M, N = \{64, 32, 16, 8\}$.

6.3.4 *In vivo* Patella Tendon Results

Sample frames showing a patella tendon (Figure 6.5) appear in Figure 6.6. These show approximate tendon flexion motion in f_1 to f_{16} (traversing from right to left), and extension motion onward (traversing from left to right). Each B-scan shows the skin and soft tissue region at the top, the patella tendon in the middle (delineated by the paratenon appearing as a hyperechoic anterior and posterior boundary), and only noise lower down. As a result of scanning the tendon-bone interface, appearing as a prominent signal (centre-left), a hypoechoic region exists from a lack of signal (an attenuation artefact from the highly attenuating bone), and some specular reflection within the tendon. Also, the tendon is striated from the parallel fibre bundles.

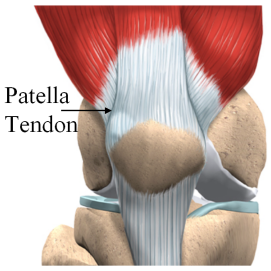


Figure 6.5: Patella.

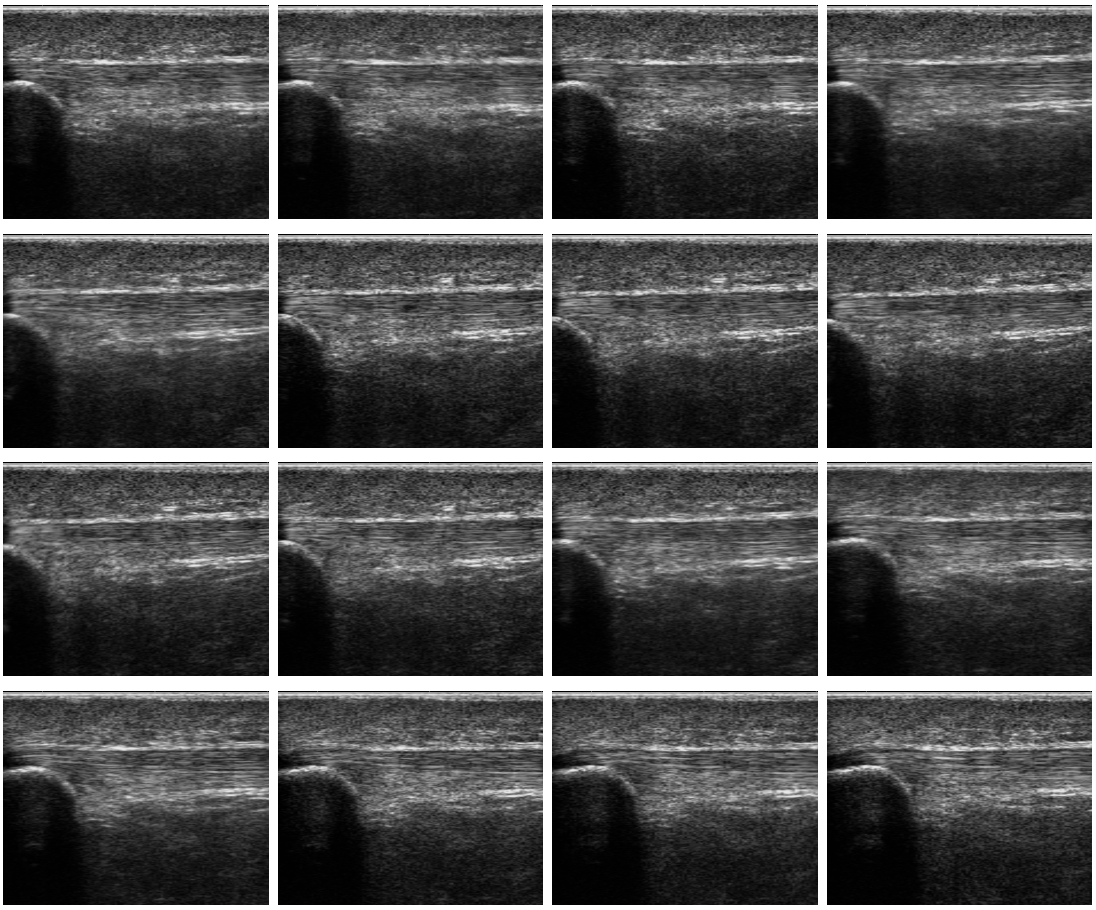


Figure 6.6: Sample *in vivo* B-scans of a longitudinal section of the patella tendon, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

Figure 6.7 shows our trajectory field results. In general, all trajectories demonstrated a common high temporal correlation throughout the sequence. Another observation was that the trajectories within the tendon had relatively linear lateral displacements, and negligible axial displacements. Yet, the trajectories in the noisy region beneath the tendon, instead of being completely discontinuous and random, appeared to mimic some of the lateral tendon motion. This was apparent for all tendons that traverse laterally and not encapsulated by a sheath. Additionally, at the tendon-bone interface that provides a landmark, trajectories formed a dense coherent cluster. Some erroneous trajectories also exist, at the trajectory field edges due to motion at the image borders.

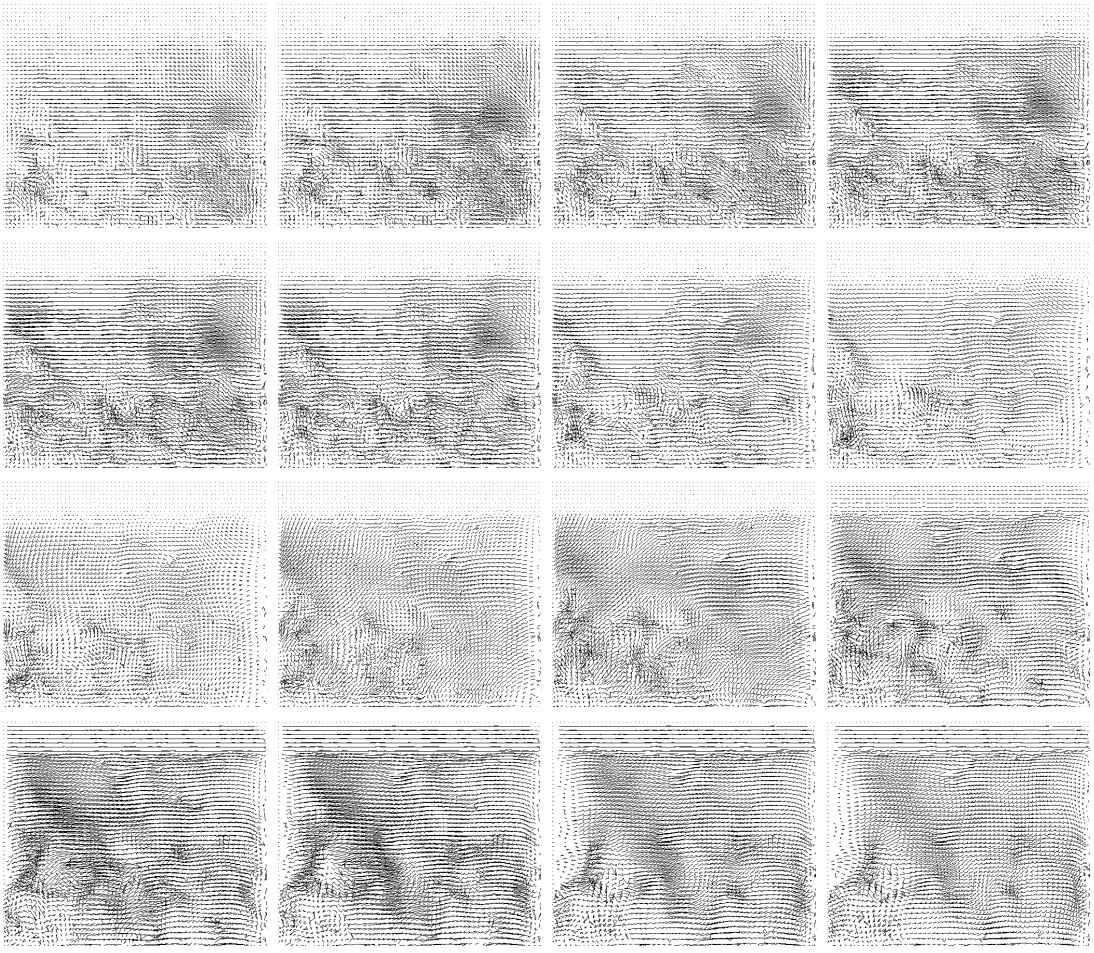


Figure 6.7: Sample trajectory fields from our final proposed approach, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

Trajectory fields were used to produce the corresponding strain maps shown in Figure 6.8. Results provide a temporal history of 2D strain magnitude, generating a history for the complete sequence. As expected, the strain gradually increased to a peak strain occurring at the maximum flexion movement at f_{16} of $\varepsilon = 1.2\%$ (colour coded as dark red). These results do not show a uniform region of strain corresponding to the tendon, however, the internal tendon strain was low, ranging from $\varepsilon = 0.5\%$ (in green) to $\varepsilon = 1\%$ (in orange) at the tendon-bone interface. Interestingly, in $f_8 \rightarrow f_{22}$ real strains are measured in a reverberation artefact (lower-left), shown as repeated diminishing peaks of the strain that occurred within the tendon, which is explored later in Section 7.2. *In vivo* strain maps are difficult to evaluate beyond a subjective expert opinion, however, in Section 6.3.7 we report image reconstruction errors.

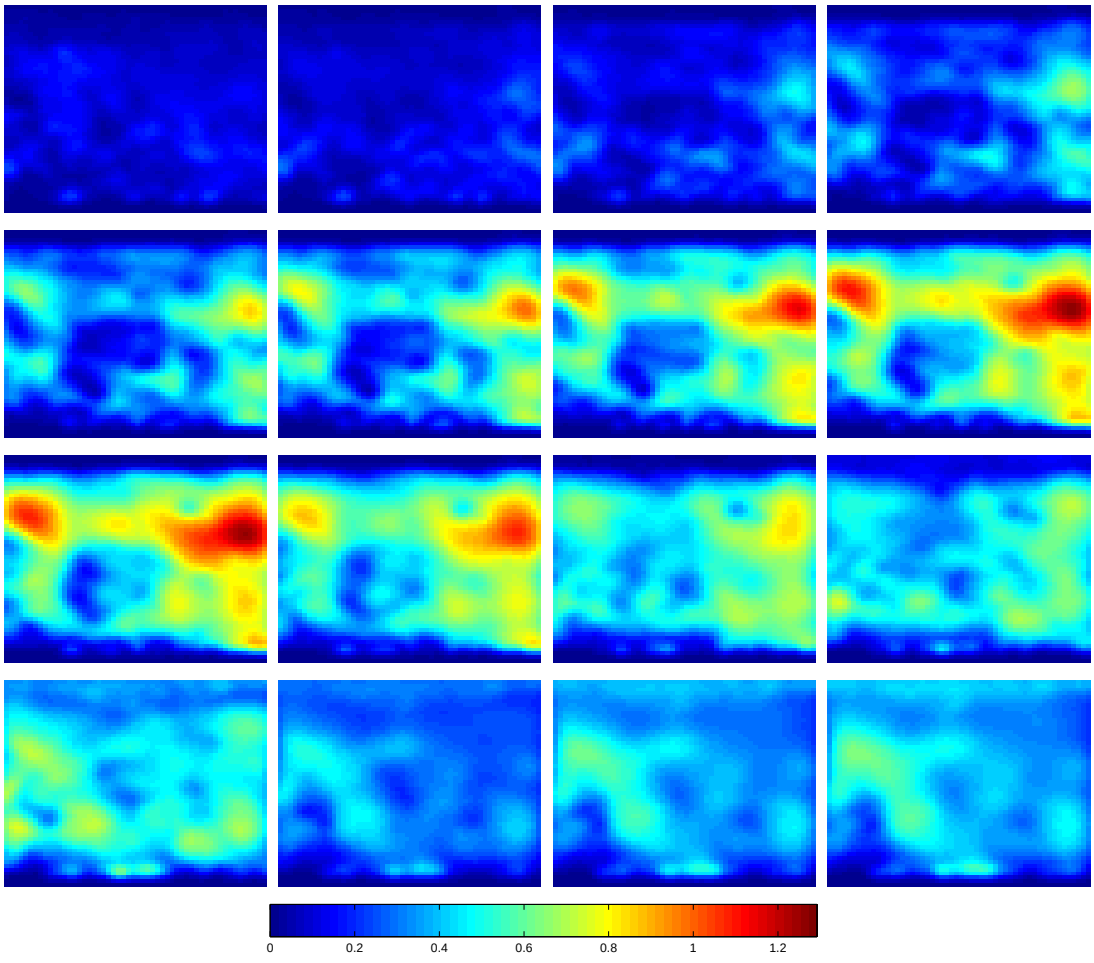


Figure 6.8: Sample 2D elastograms from normalising a second order strain tensor, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

6.3.5 *In vivo* Achilles Tendon Results

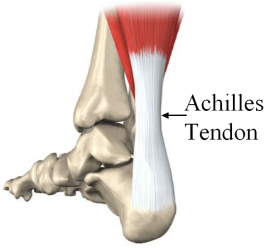


Figure 6.9: Achilles.

Sample frames of an achilles tendon (Figure 6.9) appear in Figure 6.10. These show approximate tendon flexion motion in f_1 to f_{16} (traversing from right to left), and extension motion onward (traversing from left to right). Each B-scan shows the same tiered structure that was observed for the patella tendon sequence. However, differences include an observable hypoechoic region (centre-right), muscle and soft tissue sub-structure beneath the tendon, and the absence of a bone landmark. Also, subtle attenuation artefacts appear distal to the tendon, appearing as slightly darker streaking vertical shadows. It is clear in f_{26} that motion blur (smoothed intensities) has corrupted the B-scan from inconsistent interframe velocity, which we expect to later effect trajectory results, and cause incorrect image reconstruction errors.

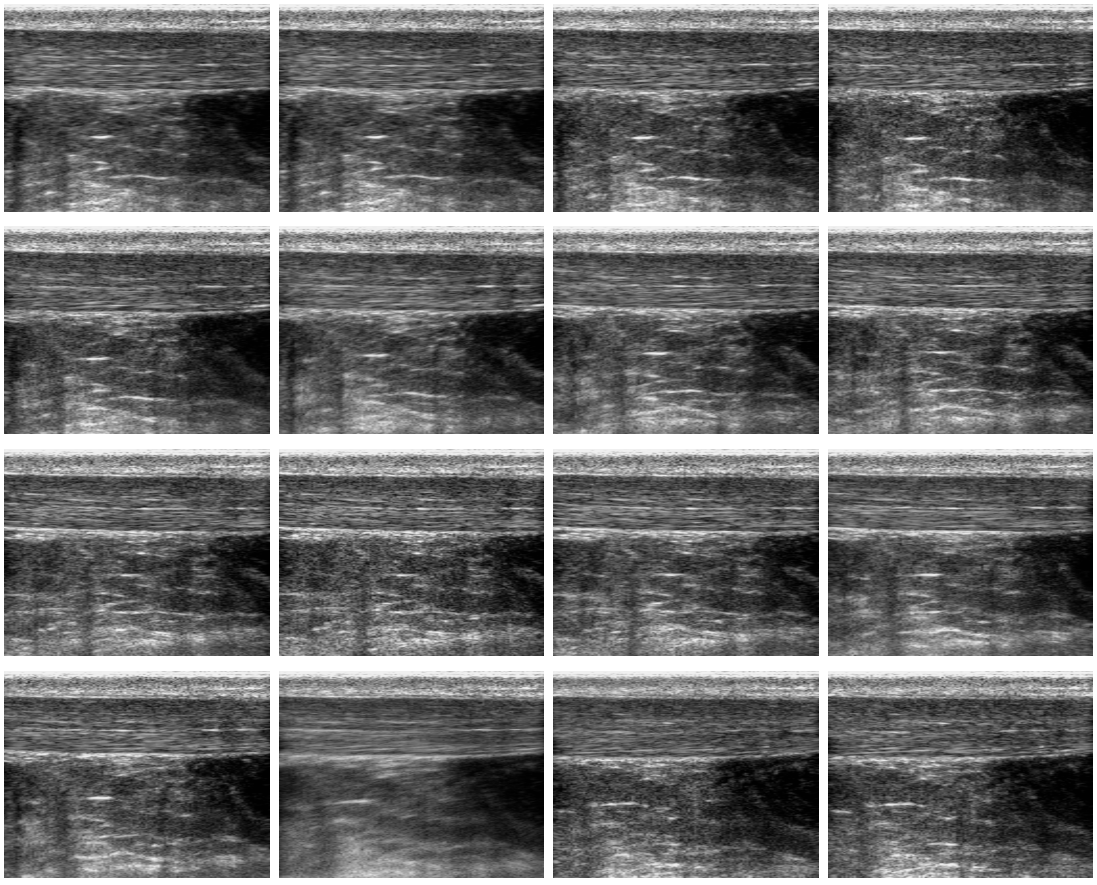


Figure 6.10: Sample *in vivo* B-scans of a longitudinal section of the achilles tendon, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

Figure 6.11 shows our trajectory field results. All trajectory fields shown in the top row have less activity due to the plotted trajectory history being < 5 frames. Therefore, the first trajectory field for each sequence only shows interframe displacements. All results showed spatial and temporal regions of strong correlation, except for a region of discontinuous trajectories at the middle and base of the tendon. This was consistent in several of the trajectory fields. Also, it was observed that WVMF using 3 iterations, produced smooth continuous displacements, yet maintained this region of discontinuity. Therefore, the trajectories in this region must have a high confidence weighting, hence a high correlation or low CD_2 error.

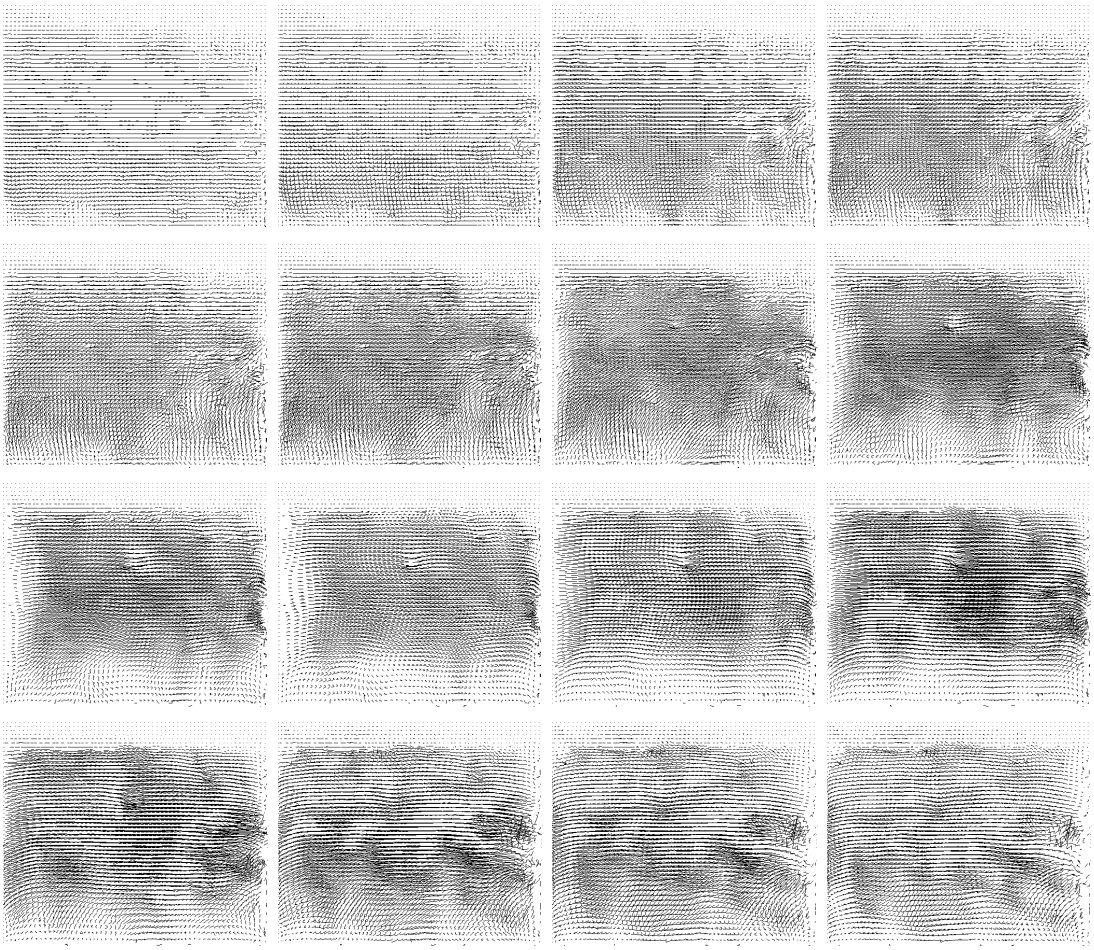


Figure 6.11: Sample trajectory fields from our final proposed approach, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

The strain maps in Figure 6.12 show small global strains of $\varepsilon < 0.5\%$ (in blue) in $f_1 \rightarrow f_8$, as the tendon moved from a fully extended position. As expected, the maximum strain occurred at the limit of flexion motion in f_{16} and f_{18}, where the tendon underwent a strain of $\varepsilon = 1\%$ (in green), and the surrounding sub-structure produced a maximum strain of $\varepsilon = 1.8\%$ (in red). In general, strains were horizontally uniform, with strains in the tendon region lower than the surrounding soft tissue sub-structure at $\varepsilon \approx 0.5\%$ and $\varepsilon \approx 1.1\%$ respectively. The strain maps appear more homogenous than those in Figure 6.8, due to the achilles tendon and sub-structure consisting of a more global motion. Furthermore, the local region of discontinuity observed in the previous trajectory fields did not generate any significant strains, but highlighted the importance of using displacements and strain to detect and measure musculoskeletal tramua.

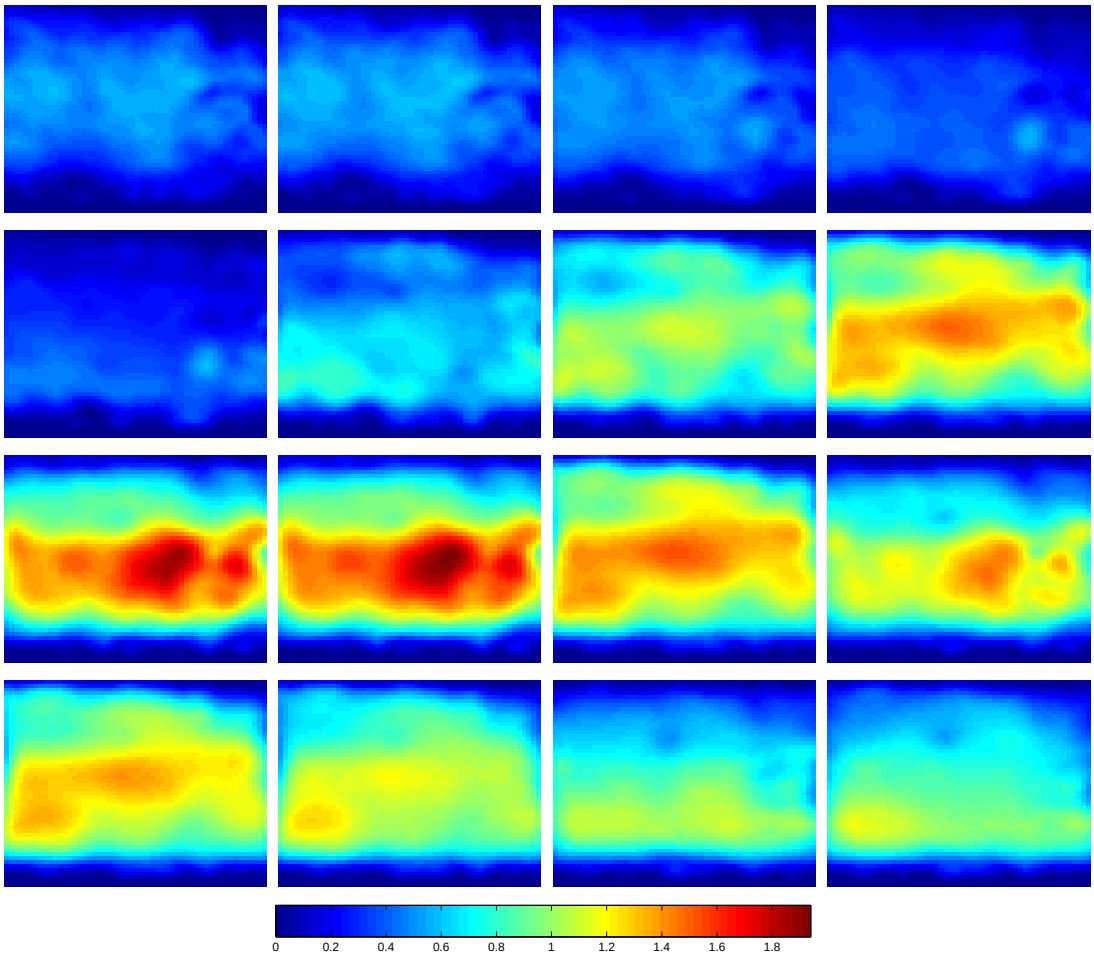


Figure 6.12: Sample 2D elastograms from normalising a second order strain tensor, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

6.3.6 *In vivo* Digital Flexor Tendon Results

Figure 6.13, adapted from [34], illustrates the sheath (in blue), encapsulating the superficial digital flexor tendon, and the bone (in brown). The finger joint captured is the P2 to P3, which is the finger tip. Sample frames from a digital flexor tendon sequence appear in Figure 6.14. These show only approximate tendon flexion motion for f_1 to f_{30}. In the B-scans the tendon sheath and bones appear hyperechoic, delineating the tendon boundary and bone structure respectively, with the region beneath capturing only negligible signal.

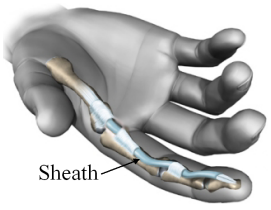


Figure 6.13: Digital flexor tendon sheath and bone.

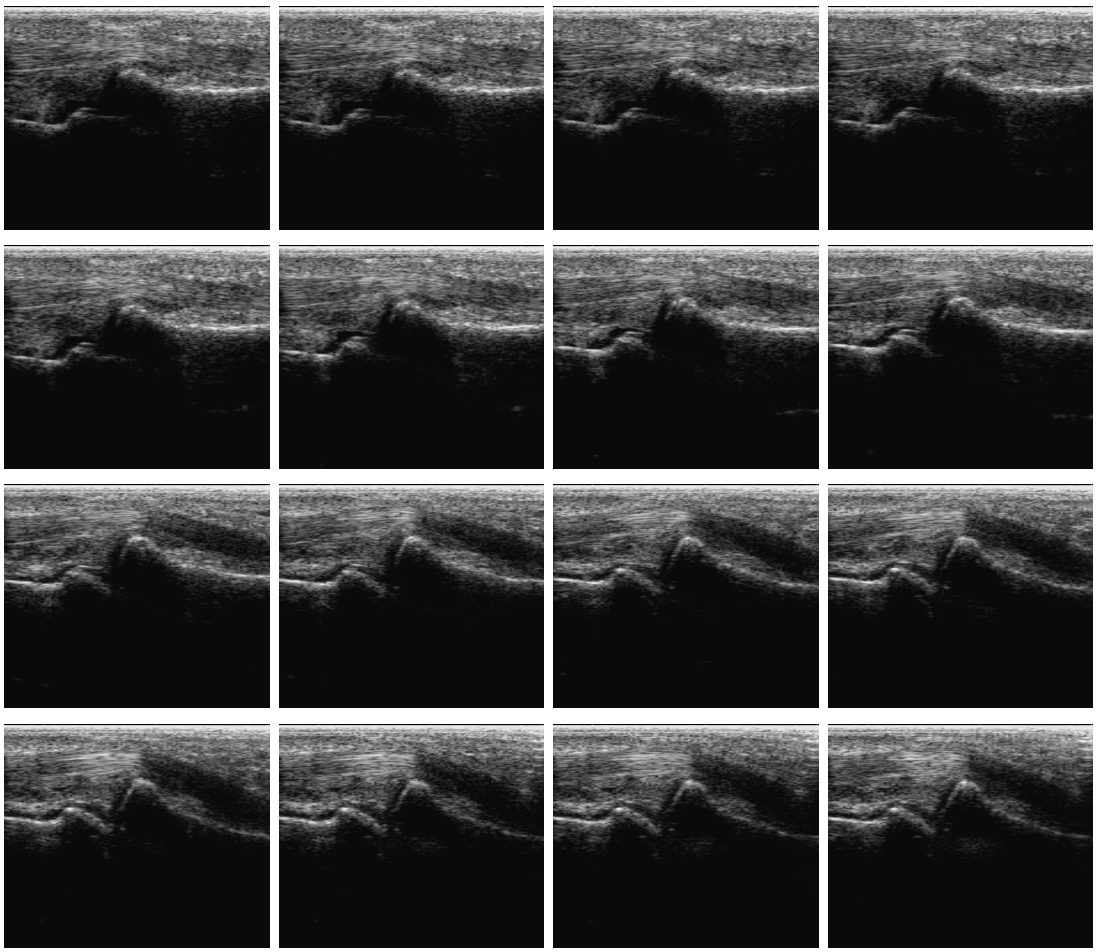


Figure 6.14: Sample *in vivo* B-scans of a longitudinal section of the digital flexor tendon, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

The trajectory fields in Figure 6.15 show the finger tip orientated to the right, which moves downward from extension to flexion. Therefore, maximum motion occurs at the start and end of the curvilinear tendon (located either side of the joint), with least motion distal to the tendon joint (centre). Results show homogeneous regions of trajectories that are constrained to the anatomy. This was also observed in Chapter 4, distinguishing tendons with or without a sheath. An interesting observation was the lack of signal as the tendon moved away from the transducer, becoming increasingly less perpendicular to the beam. Trajectories in this region were smooth, consisting of the same global motion. Some reverberation artefacts under the finger tip (centre-right), produced displacements occurring as a small cluster in f_{12} to f_{24}.

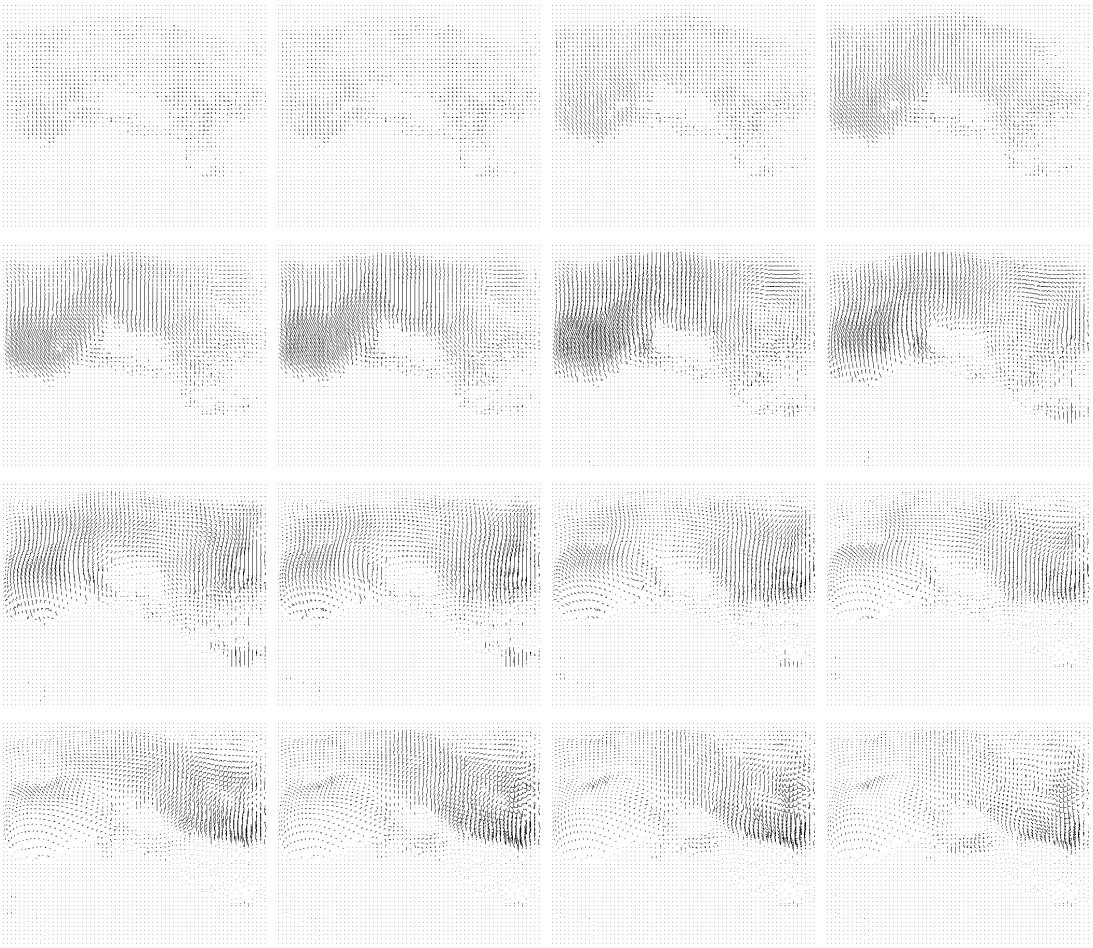


Figure 6.15: Sample trajectory fields from our final proposed approach, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

Figure 6.16 shows the corresponding temporal strain maps providing a strain history for the digital flexor tendon. The hypoechoic region of tendon (centre-right), as expected showed zero strains (in dark blue). Although strains in general were low, there was a substantial localised strain concentration of $\varepsilon = 1.5\%$ on the left of the joint (centre-left). Interestingly, a subtle strain devloped over the finger joint, of approximately $\varepsilon = 0.9\%$ (centre) at f_{12} onward from increasing flexion motion. Due to the occurrence of a reverberation artefact (centre-right), a small region of strain ($\varepsilon = 1\%$) occurred, which did not correspond to the anatomy. This can arise when scanning curvilinear tendons (for example the rotator cuff), due to using tendon dynamic behaviour to determine strain. Finally, a region of stationary trajectories distal to the finger joint (centre), produced a zero strain region in the strain maps.

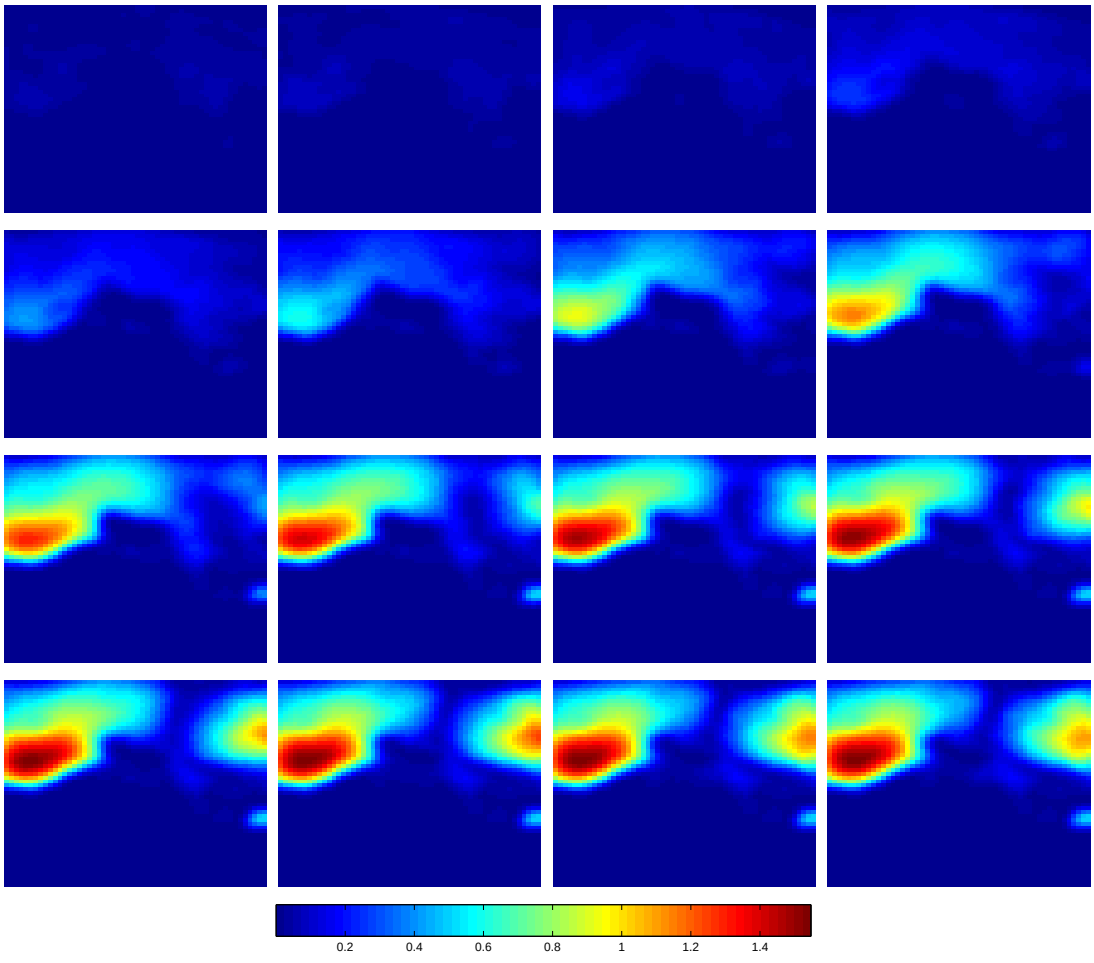


Figure 6.16: Sample 2D elastograms from normalising a second order strain tensor, for frames $f_1 f_2 f_4 f_6$, $f_8 f_{10} f_{12} f_{14}$, $f_{16} f_{18} f_{20} f_{22}$ and $f_{24} f_{26} f_{28} f_{30}$ from the top row to bottom row. ♠

6.3.7 *In vivo* Musculoskeletal Accuracy

The image reconstruction errors for the previous *in vivo* results are summarised in Table 6.1. For the patella, achilles and digital flexor tendon sequences, we show the pre-FD and post-DFD errors, followed by the pre-correlation (pre-CORR) and the post-correlation (post-CORR) measures, as percentages. These show promising results, indicated by 87% of DFD errors being substantially less than the FD errors, and corresponding increased interframe global correlation. The best cases are those that quantify a large difference between DFD and FD error. Some selected cases are highlighted with an [a]. The majority of these results are good, for example, in the patella tendon sequence at f_{14} the FD was 200.92 compared to a DFD of 151.98, and pre and post correlations of 97.96% and 98.45% respectively. Unfortunately, there are some cases that show both a higher DFD than FD error, and lower interframe global correlations. All of these cases are denoted with a [b]. As observed in Section 6.3.5, f_{26} in an achilles tendon sequence was corrupted with motion blurring. Therefore, we anticipated the poor result obtained, where the FD of 250.95 was lower than the DFD of 295.68, with a pre and post correlation of 98.65% and 98.49% respectively. Another example of this is apparent at f_{23} of the same sequence, where the FD of 82.61 was lower than the DFD of 159.01, with a higher pre than post correlation of 99.57% and 99.15% respectively. Upon investigation, the latter cases were due in part to significant global frame pair changes resulting from motion blurring, and not from inaccurate displacements that can be seen in Figure 6.11. Visual inspection revealed regions of blurred speckle caused by the target anatomy having inconsistent interframe velocity.

6.3.8 Elastogram Multicompression Averaging

It has been observed by Varghese et al. [104] that multicompression averaging of the elastograms produced from RF elastography improved the SNR by reducing variations in the strain estimates caused by displacement noise. The SNR of the strain maps, produced from our speckle tracking approach, can also be improved using image based multicompression averaging. Again, the source of noise occurring in the strain maps is from erroneous displacements, but as a result of analysing lateral and complex motions, typically appearing as discontinuities. The averaging technique, reduces the standard deviation of the averaged elastogram by $\sqrt{N}$, where N is the number of strain maps for compression, improving the SNR by a factor of $\sqrt{N}$.

Patella Tendon				Achilles Tendon				Digital Flexor Tendon			
PRE FD	**POST DFD**	**PRE CORR**	**POST CORR**	**PRE FD**	**POST DFD**	**PRE CORR**	**POST CORR**	**PRE FD**	**POST DFD**	**PRE CORR**	**POST CORR**
167.70	164.95	98.28	98.31	188.12	137.97	99.01	99.27	57.52	51.85	99.38	99.44
156.19	148.01	98.28	98.42	417.83	326.97	97.79	98.26	107.26	76.71	98.85	99.18
136.51	120.48	98.62	98.78	**82.61**	**159.01**	**99.57**	**99.15**[b]	42.62	37.42	99.55	99.60
265.51	223.45	97.33	97.76	334.14	275.40	98.22	98.53	**23.32**	**31.11**	**99.75**	**99.67**[b]
232.07	204.35	97.69	97.96	362.25	303.12	98.10	98.41	168.15	102.81	98.22	98.91
222.41	195.46	97.78	98.04	**461.73**	**277.54**	**97.63**	**98.58**[a]	200.72	105.54	97.91	98.90
179.81	180.43	98.20	98.19	417.21	223.22	97.90	98.88	286.19	160.57	97.04	98.33
102.10	**159.92**	**98.99**	**98.42**[b]	464.87	338.07	97.68	98.32	222.29	120.52	97.71	98.76
244.42	221.85	97.85	97.80	373.38	332.57	98.16	98.36	252.16	158.21	97.40	98.35
140.34	**151.74**	**98.59**	**98.47**[b]	413.25	343.60	97.99	98.33	224.09	146.57	97.66	98.45
238.79	216.35	97.59	97.81	388.30	302.82	98.13	98.54	182.94	143.37	98.07	98.48
147.18	146.18	98.49	98.51	356.51	335.54	98.29	98.39	145.05	109.89	98.45	98.82
90.32	91.52	99.08	99.10	338.67	312.01	98.39	98.52	**281.62**	**206.35**	**96.92**	**97.74**[a]
200.92	**151.98**	**97.96**	**98.45**[a]	**226.01**	**245.85**	**98.94**	**98.85**[b]	196.03	134.46	97.80	98.49
228.53	177.52	97.68	98.19	142.06	242.50	99.34	98.87	185.44	136.49	97.89	98.44
190.76	**200.79**	**98.08**	**97.97**[b]	298.41	277.96	98.60	98.69	140.12	110.00	98.38	98.73
222.84	216.56	97.77	97.82	319.11	225.82	98.49	98.93	169.24	124.68	98.03	98.55
186.73	170.45	98.12	98.28	372.20	326.50	98.22	98.43	151.23	114.83	98.23	98.66
291.51	246.65	97.06	97.50	505.41	399.55	97.53	98.05	**65.23**	**74.87**	**99.24**	**99.12**[b]
150.67	139.31	98.48	98.58	**242.10**	**292.52**	**98.83**	**98.58**[b]	113.06	91.70	98.68	98.93
302.78	229.85	96.91	97.64	396.10	342.41	98.08	98.33	134.67	120.01	98.43	98.60
196.32	134.30	97.98	98.61	441.44	378.86	97.82	98.11	171.27	140.23	98.01	98.37[b]
251.47	165.66	97.39	98.26	430.03	338.35	97.81	98.26	159.99	116.36	98.12	98.64
182.67	158.45	98.10	98.37	447.27	352.51	97.61	98.10	**40.46**	**85.85**	**99.53**	**99.00**[b]
174.64	152.58	98.19	98.40	324.40	241.20	98.23	98.67	26.09	38.64	99.70	99.55
191.18	160.40	98.00	98.32	**250.95**	**295.68**	**98.65**	**98.49**[b]	**68.88**	**76.01**	**99.19**	**99.10**[b]
106.05	99.88	99.88	98.97	257.01	236.61	98.63	98.65	90.88	79.74	98.83	99.06
222.29	196.74	97.73	97.98	404.18	380.78	97.87	97.99	**75.90**	**77.70**	**99.11**	**99.11**[b]

Table 6.1: Error measures for *in vivo* data of the patella, achilles and digital flexor tendons, where each row reports interframe analysis.

All the good results, where the DFD is lower than the FD.

[a] Sample best results, where the DFD is much higher than the FD.

[b] All the worst results, where the DFD is lower than the FD.

It is important to only average strain maps from equivalent motions, for example, either a small number of frames or a single kinematic flexion or extension motion.

6.4 Strain Estimation from Signal Decorrelation

So far in this chapter, we have described the process of using displacements produced from our proposed technique to determine strain information. However, an alternative approach to estimating strain, is to exploit the relationship that exists between the correlation coefficient and the applied strain, as applied by Bamber and Bush [7] and Varghese and Ophir [102], using the synthetic data described in Chapter 3. This relationship assumes that as strain increases the correlation coefficient decreases. This is a sensible assumption given that in the most general situation, strain does not have a cyclic nature, as an object can only be deformed once. In ultrasound elasticity imaging, strain induced decorrelation is a major source of error in estimating 2D displacements using correlation techniques. We have shown that this error can be significantly decreased by using multiple block sizes, and by alternating speckle similarity measures, to ensure regions that would produce a poor correlation match are tracked using a measure taking into account the statistics of speckle. Therefore, using the decorrelation-strain relationship has the advantage of not suffering from displacement inaccuracies.

In this section, we measure the signal decorrelation to estimate strain to provide a comparison to our previous *in vivo* strain tensor results and the work in [102]. We determine the peak correlation coefficient c_{max} using the NCC to define the decorrelation $\hat{\varepsilon}$ as an estimate of strain magnitude, given as:

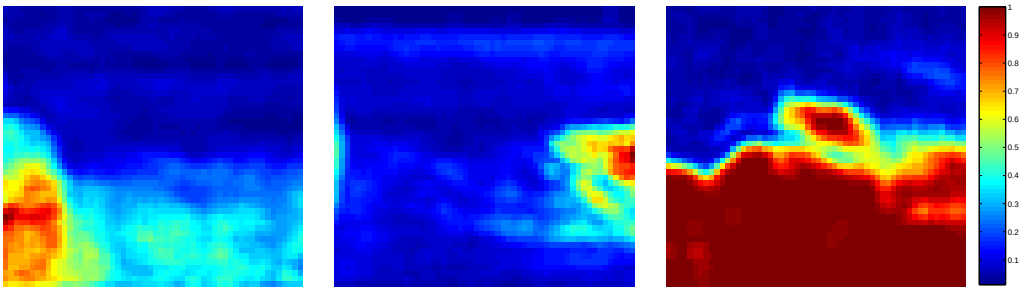
$$\hat{\varepsilon} \approx 1.0 - c_{max} \tag{6.9}$$

Further, we applied multicompression averaging to smooth each decorrelation result, defined in Section 6.3.8, not previously used for the strain tensor results. Also, we specifically analysed frames under the same mechanical motion, for example flexion motion. From the results in the next section, we show that strain estimation from signal decorrelation is not a good method to measure *in vivo* strain, but does provide an indicator of confidence in displacement accuracy.

6.4.1 Signal Decorrelation Results

The signal decorrelation results presented in Figure 6.17, are examples based on the *in vivo* patella, achilles and digital flexor tendon sequences, originally labelled earlier in Chapter 3.

We present 3 sets of decorrelation results. First, Figure 6.17(a) shows the decorrelation result for a period of patella tendon flexion motion. This revealed a singular high decorrelation concentration of $\hat{\varepsilon} = 0.6$ (in orange-yellow), located in the reverberation artefact induced by the patella bone. The remaining tendon section consisted of a uniform low decorrelation of $\hat{\varepsilon} < 0.1$ (in blue), with increasing decorrelation in the region beneath the tendon (in light blue). Second, the decorrelation result in Figure 6.17(b) was produced for a period of achilles tendon flexion motion. Within the tendon and in surrounding substructure, decorrelation was relatively low and uniform at $\hat{\varepsilon} < 0.1$ (in blue). Due to the tendon traversing out of the image plane, some decorrelation was visible at the border ($\hat{\varepsilon} \approx 0.7$), which is not prominent in the strain maps in Figure 6.12. The final experiment used a period of flexion motion of the digital flexor tendon sequence. The decorrelation result produced is shown in Figure 6.17(c). Areas of zero signal are shown with a decorrelation value of $\hat{\varepsilon} = 1$ (in red). A strong decorrelation peak was present just beneath the finger joint, which does correspond to a zero strain region in Figure 6.16. Again, the internal tendon decorrelation was low $\hat{\varepsilon} < 0.1$.



(a) Patella decorrelation. (b) Achilles decorrelation. (c) Digital flexor decorrelation.

Figure 6.17: Decorrelation results for the patella, achilles and digital flexor tendons.

In summary, by comparing the strain maps presented in Figures 6.8, 6.12 and 6.16 to the decorrelation fields in Figures 6.17(a), 6.17(b) and 6.17(c), no two regions correspond. The decorrelation fields proved insensitive to estimating strain, showing no fluctuations or significant values, indicated by large regions of $\hat{\varepsilon} < 0.1$ (in blue). Although, decorrelation provides a measure of scatterer change, from tissue deformation in axial, lateral and elevation axes, and

is invariant to tensile, compressive or shear strains, the decorrelation also indicates trajectory confidence. As a confidence measure, all results showed a low decorrelation (high correlation) from using the NCC, confirming our confidence that our proposed approach in Chapter 5 yielded accurate displacement estimates within the tendon region.

6.5 Conclusion

The objective of this chapter was to generate axial, lateral and shear strain information using *in vitro* and *in vivo* sequences. A problem for all strain measurement techniques is quantifying strain accuracy, specifically for *in vivo* data, as we only have a prior knowledge of what to expect. For example, the accuracy of the *in vitro* RF elastogram in Figure 2.3 produced by Konofagou and Ophir, was determined by simpler *in vitro* phantom experiments, and not from the data it was based on. Although it is popular to measure RF elastogram accuracy by the SNR, this only measures the noise in the strain. However, our 2D speckle tracking technique benefits from quantifying the image reconstruction error and decorrelation measures. These provide a better alternative to the SNR, as both indicate displacement and strain accuracy and confidence respectively, based on the *in vivo* data. For *in vitro* data, strain accuracy is determined by comparisons to the groundtruth 1D strain, and again displacement error measures. Hence, we have demonstrated accurate strain results derived and validated using the improved displacements and trajectories, from our proposed speckle tracking estimation technique reported in Chapter 5. Further, we generated high precision axial and lateral spatial and temporal displacements, and also presented a unique strain history for ultrasound sequences. Results showed measurable axial compressive strains up to $\varepsilon = 3.2\%$ using the tissue mimicking phantom, lateral tensile strains up to $\varepsilon = 3.5\%$ using the *in vitro* tendon data, and a generally low $\varepsilon < 2\%$ strain magnitude for all our *in vivo* sequences, which were captured under zero load.

An advantage of using B-scan image analysis compared to RF elastography, is that the estimated displacements are not dependent on the employed transducer. A further advantage, is the capability of quantifying both axial and lateral components of a second order strain tensor, rather than only the axial components as in RF elastography (exceptions have been described earlier in Chapter 2). Consequently, our technique automatically enables the ability to extract Poisson's ratio (ratio of lateral and axial strain) and Young's modulus (ratio of stress to strain

on the loading plane along the loading direction). Currently, there is little information of this type for soft tissues and tendons. Although the elevation component is also required to fully characterise strain and indeed motion, in order to achieve this we would require 3D image volumes. However, another advantage of using trajectories, is the means to update for 3D volume movement in the 2D plane, which is a current issue with RF elastography.

We observed from our decorrelation study that the decorrelation coefficient was unable to estimate the true value of the strain accurately, confirmed in [102], and that the decorrelation-strain relationship was not valid for our musculoskeletal *in vivo* data. Over a larger number of *in vivo* sequences, the decorrelation coefficients as a measure of strain produced poor repeatability. Lastly, the decorrelation coefficients demonstrated that high correlations > 0.9 were typically observed in regions of tendon. Hence, the correlation provides a good indicator of the confidence in the accuracy of earlier displacements.

Chapter 7

Conclusions and Future Work

The motivation for this thesis was to develop a 2D speckle tracking technique to quantify displacements and deformation using ultrasound video sequences. Our aim was to estimate 2D displacements for determining axial and lateral strain information, focusing in particular on musculoskeletal data captured with a linear array transducer using a frequency of $8-16$ MHz. Each B-scan consisted of two distinct textures from regions of tendon and soft tissue, which later justified the use of *in vitro* tendon and tissue mimicking phantom experiments, to provide a large range of groundtruth data captured under known conditions.

Methods that measure meaningful displacement descriptors using ultrasound, make the assumption that speckle motion replicates small tissue deformation. Typically, for small deformations of $< 5\%$ this assumption is valid. For situations where speckle motion is different from the tissue motion, displacements become less meaningful. This is the principal assumption for RF elastography and B-scan speckle tracking, forming the foundation to obtaining precision displacements and strain estimates. The mainstream approach to 2D speckle tracking consists of analysing selected B-scan frame pairs, in order to produce a single interframe displacement field. For most cases, this requires a manual and repetitive process of selecting and analysing specific frame pairs, to produce displacements that are expertly evaluated, but prone to inter- and intra-observer expert variabilities.

The success of any image based technique is first dependent on the quality and interpretation of the freehand frames captured, and second the suitability of the applied displacement estimation

technique. With modern ultrasound machines producing B-scans of increasing quality, a comparative analysis was undertaken in Chapter 4, to assess the feasibility of speckle tracking techniques in musculoskeletal B-scans. The outcome confirmed that block matching frameworks, which are a logical extension of RF analysis, produced the most comprehensive displacements for our *in vitro* groundtruth and *in vivo* data. In addition, the use of multiple scale blocks proved to have a dramatic increase in displacement accuracy. Notably, differential based algorithms proved sensitive to speckle, with inaccurate numerical differentiation resulting in direct errors in the displacement fields. Also, a significant source of displacement error was apparent from the aperture problem, remaining a substantial challenge for all techniques.

Rather than applying image pre-processing and then performing displacement estimation, we preferred to measure displacements and apply displacement post-processing, a similar approach employed later by Boukerroui et al. [13]. This was due to image pre-processing removing or suppressing useful speckle and important information, altering the displayed signal to be tracked. Also, many ultrasound machines perform internal pre-processing on the RF and B-mode signals, prior to displaying and frame capturing. Due to the employed ultrasound hardware, prior internal pre-processing was consistently controlled and restricted. Instead of B-scan pre-processing, a more important stage is to correct for displayed imaging artefacts present in the B-scans. For example, correcting for non-uniform attenuation artefacts, which has received much attention at the RF signal level [39]. This is mainly an issue for lower frequency ultrasound used for abdominal work, but is also an issue for curvilinear tendons, such as the rotator cuff and digital flexors. The necessity to detect artefacts in order to correct for them, is later discussed as future work.

A notable drawback of block matching approaches was attempting to estimate displacements in featureless regions of uniform speckle. In these cases displacement accuracy was influenced by scale and the selection of a matching criterion. To address the problem, a novel alternating similarity measure approach was developed, ensuring the appropriate measure was used, based on the speckle SNR for the region. The first measure used the normalised cross correlation for regions of strong signal (refined to minimise aperture problems induced from the tendon), and the second measure CD_2 was applied for regions of speckle determined by the speckle SNR. The real improvement in displacement accuracy was obvious from analysing frames that contained sub-regions of differing signal (from tendon and tissue). Further advances included

extending interframe displacements to trajectory fields, whilst integrating drift correction, to represent the temporal tracking of speckle features sequentially through a sequence.

Accuracy was evaluated using synthetic speckle data, *in vitro* tendon pull data and tissue mimicking phantom data. Using synthetic speckle data, the best displacement accuracy was achieved using the alternating matching measures approach, compared to similar block matching methods that used either the NCC, CD_2 or MSE measures only. For example, the combined measures approach produced a DFD of 7.48, compared to CD_2 of 9.33, NCC of 8.10 and MSE of 10.16 using images of varying density speckle. For the *in vitro* tendon pull data, FD errors ranged between 270.28 to 273.87 (lowest and highest), compared to the final proposed technique that achieved DFD errors of 32.27 to 36.74. This was a substantial improvement than previous methods, for example FSBM produced a DFD of 80.12 to 86.68, VSBM a DFD of 51.12 to 59.29, Camus [14] a DFD of 214.76 to 223.30 and Proesmans et al. [80] a DFD of 106.63 to 121.90. For the tissue mimicking phantom data, the maximum $FD_{\sim,\triangle}$ errors were 1000 and 900, with the proposed technique $DFD_{\sim,\triangle}$ errors of 200 for each. Finally, our proposed method for *in vivo* data produced substantially lower DFD than FD errors, and in most cases more accurate displacements than similar existing methods. Determining *in vivo* displacement accuracy is challenging, due to the restrictions of not having a knowledge of a groundtruth. Therefore we used the *in vitro* data extensively for analysis and proof of concept.

Both motion and capture rates had a direct effect on the measured magnitude of deformation and strain. With higher frame rates interframe coherence improved, increasing interframe displacement precision and reducing strain magnitude. The range of accurately measured interframe displacements was approximately 1 to 30 pixels, depending on the size of the largest template. In cases of high deformation, displacement accuracy reduced as the assumption of pixel constancy failed. Temporal reference template updating was used to successfully eliminate trajectory drift error, which was found to be induced by a low frame rate. Trajectory fields proved invariant to a range of capture rates and subject movements, whilst also demonstrating that specific regions could be tracked through complete sequences, eliminating the notion of mere frame matching and improving displacement field temporal coherence.

Improvements in displacement estimates are still possible as indicated by non-zero image reconstruction errors. To further improve displacement accuracy, this work could be extended

to tracking 3D volumes, where a volume of interest (a sphere) is used to measure all axes of motion, instead of a region of interest (a block) that measures just axial and lateral motion. An advantage of trajectories was demonstrated by updating for temporal changes in the image plane from axial, lateral and elevation movement, updating the trajectory fields with new objects entering and exiting from the 3D scanned volume. This provided an enhanced understanding of the displayed B-scans captured during the activity of the target volume.

Currently, there is a lack of real-time displacement and strain estimation systems to provide clinicians with a practical diagnostic tool. To provide an indication of processing time, a displacement field of 64×64 for a frame pair of 512×512, using block scales $M \times N$, where M,$N = \{64, 32, 16, 8\}$ took 65 secs, and a 30-frame sequence 1950 secs. Using the same settings, but for a specific ROI, a displacement field of 10×10 took 2.0 secs, and a 30-frame sequence 60 secs, both generating all inputs and outputs and using a C implementation on an Intel Pentium 4 at 2 GHz. Although not comparable to the frame capture rate of 8.6 Hz, this is certainly faster than RF elastography, and has the potential to be an offline clinical tool operating directly on an ultrasound machine. As expected, the NCC using the FFT was much faster than spatial domain convolution used by the CD_2 measure. The final proposed approach did not use sub-pixel accuracy, although work in Appendix B showed that in some cases sub-pixel accuracy could improve displacement precision. This could be achieved by super-sampling candidate and reference templates for both similarity measures.

Furthermore, we have presented a practical form of elastography, that benefits from estimating displacements directly in 2D using a knowledge of pixel contextual information, rather than the 1D followed by 1D approach used in RF elastography [46]. Also, B-scan analysis does not need specific hardware to process RF signals, which is required by RF elastography. Unfortunately, 2D speckle tracking techniques are all inherently prone to information loss from the demodulation process required to create B-scans, and errors due to both image axes constructed using different methods. Finally, a problem for all elastography techniques is compensating for displacements from cardiovascular or respiratory motion, but this is not an issue for our musculoskeletal data. Our research has shown practical and realistic results are possible, and that the approach is usable in a clinical environment. To date, an executable has been operational in a Biomechanics lab[1] for soft tissue analysis.

[1]Institute of Biomechanics, University of Nottingham, Nottingham, N97 2RD, UK

A novel outcome from the generation of our trajectories, is a unique temporal axial and lateral strain history. This knowledge of temporal strain, provides a measure of the strains occurring for a target subject during physiological activity. Hence, the location of any strain concentration is measured at the precise time it occurs. Many biological tissues undergo a range of deformations in physiological activity. For small deformations $< 3\%$, the small deformation linear strain tensor was defined and applied, which assumes that the change in volume is negligible. However, for larger deformations the non-linear strain tensor was defined. We have measured strains up to 5% with good precision based on the accuracy of the displacements, which is an acceptable range for most soft tissues and tendons.

We also compared the strain results to the decorrelation coefficients produced from the NCC that provide an independent estimate of strain, which is less sensitive to displacement accuracy. The decorrelation coefficients, even with multicompression averaging, demonstrated poor strain estimates compared to the good performance of the strain tensor, which was derived from the estimated displacements. However, the correlations showed consistently high values in regions of tendon, confirming our confidence in displacement accuracy from using the SNR, which instigates the usage of the NCC in this region. Finally, we note that applying an inappropriate single similarity measure, as a result of the local SNR being a poor indicator of speckle content, produces a displacement of less accuracy, but is equivalent to a displacement produced from a standard block matching method.

A lack of quantitative methods to assess *in vivo* strain accuracy has meant that RF elastograms are mainly assessed visually, based on both a prior knowledge of the expected strain distribution and performance on *in vitro* experiments. We measured the displacement accuracy using the velocity angular error and image reconstruction error, for purpose built tissue mimicking phantoms and *in vitro* tendon pull tests, to determine strain precision. Strain results showed that the tissue mimicking phantom produced an interesting local crush zone on impact, rather than a uniform compression, and that *in vitro* tendon pull tests produced valid strains for all sequences. We complemented our *in vitro* results with the well defined and controlled synthetic data, instead of using a graphite phantom. This ensured that errors from the scanning and capturing processes were not an issue. Since we have shown that we can quantify internal motions and deformations within intact structures, it is possible to make direct comparisons to model predictions from finite element models, RF elastography and experimental results.

7.1 Contributions

The contributions of this work were highlighted in Chapter 1 and are now summarised here.

A variable sized block matching method was proposed, which was compared to several relevant existing interfame motion estimation techniques. Superior performance was demonstrated by low image reconstruction and velocity angular errors, using specific *in vitro* groundtruth data.

Further, the method extended traditional interframe displacements to trajectory fields, previously not used in ultrasound analysis, generating accurate temporal displacements of speckle motion by incorporating a strategy to correct for trajectory drift error, whilst also applying trajectory updating to characterise the motion within the target 3D volume. We established that trajectories can represent the motion of a specific region through a complete ultrasound sequence, as well as eliminate the current trend of manual frame pair pre-selection.

Also, a novel alternating similarity measure approach to block matching motion estimation was proposed. Stimulated by the notion that ultrasound images contain regions of varying signal, for example signal from structure with minimal speckle (tendon), and regions of signal comprising of mainly speckle (soft tissue), we demonstrated that alternating two similarity measures is more effective than using a single measure for ultrasound motion estimation.

Utilising a knowledge of trajectories, we presented a unique strain history that measures the temporal axial and lateral strain of a subject undergoing deformation. An important benefit of strain history is the ability to locate strain concentrations at the precise time they occur, a significant advancement compared to previous frame pair analysis and RF elastography.

7.2 Future Work

By quantifying strain, tissue differentiation can be achieved between soft and hard tissues, since they react differently to external pressure. The two basic responses to an applied pressure are displacement and deformation, for which we have proposed a technique to measure in this thesis. Soft tissue tends to deform (undergoing nonuniform displacement), while hard tissue displaces uniformly (translating with rigid body motion), in response to the same external forces. With our strain history, future work will develop a functional segmentation framework

for signal and artefact detection, aiming to segment strain to classify artefacts and tissues, by exploiting the relative amount of displacement deformation for the same applied pressure.

Ultrasound images frequently contain considerable artefacts, yet have minimal noise. The images are composed of strong signals, either from tissues or artefacts. Experienced users can identify these artefacts, but for the inexperienced users artefacts are difficult to evaluate. We propose that strain and a knowledge of the target physical characteristics, can produce a functional segmentation that has significant advantages over existing segmentation techniques.

We have demonstrated that ultrasound sequences can be analysed to give precise full field movement, and mechanical strain information. This information can be used to identify regions that are not moving in a manner expected from soft tissues. For example, a strain map for a soft tissue would either be uniform, slowly changing or discontinuous, where adjacent areas of tissue have different mechanical properties. Regions of the image that are artefact, for example, dense speckle at low frequencies would demonstrate highly disorganised strain maps. Other artefacts like reverberations, signal saturation and attenuation shadowing, would demonstrate differing strain regions. Hence, we propose strain maps can be segmented according to their degree of organisation, into regions that correspond to soft tissue or artefact. Such regions can then be classified based on strain distributions. A benefit is the ability to identify weak boundaries, corresponding to objects with very similar or visually identical speckle textures.

The degree of organisation in strain maps can be formulated as a texture analysis problem, defined by the spatial arrangement of strain tensor components. Two possible approaches are applying the angular second moment (or energy homogeneity measure) of the co-occurrence matrix, or using the local strain tensor components as an external energy force for an active contour model, to force the curve to evolve at the strain object boundary. Previous segmentation approaches using B-scan intensity based texture analysis include the co-occurrence matrix [67][100], speckle statistics [78], Gabor functions [15], active contour models [3], 3D active contour models [16] and region-growing [77]. We are unaware of any published work that exploits the concept of using strain information for a functional segmentation.

Future work is required to assess the concept of using strain information in a segmentation scheme, noting that performance is dependent on displacement accuracy. Finally, this work has shown that computer vision elastography, can suitably measure motion and strain in ultrasound.

Appendix A

Explanation of the CD_2 measure

The CD_2 measure determines the displacement vector $\boldsymbol{d}$ between two blocks in regions of dense speckle. We denote the intensities for the reference template I in f_t as $\boldsymbol{a}_i = [a_1, \ldots, a_j]$ and the intensities for the candidate template I' in f_{t+1} as $\boldsymbol{b}_i = [b_1, \ldots, b_j]$ where i is an individual block and j a pixel. Making the assumption that each block $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ are corrupted with multiplicative speckle noise, then the observed pixel j in block i becomes:

$$a_{ij} = \eta_1 s_{ij} \quad \text{and} \quad b_{ij} = \eta_2 s_{ij} \tag{A.1}$$

where s_{ij} is the original signal, which is corrupted by independent multiplicative noise η_1 and η_2. Noise terms η_1 and η_2 are each characterised by Rayleigh probability density functions:

$$p(\eta_1) = \frac{\pi\eta_1}{2} \exp\left\{-\frac{\pi\eta_1^2}{4}\right\} \qquad \text{where} \qquad \eta_1 > 0 \tag{A.2}$$

$$p(\eta_2) = \frac{\pi\eta_2}{2} \exp\left\{-\frac{\pi\eta_2^2}{4}\right\} \qquad \text{where} \qquad \eta_2 > 0 \tag{A.3}$$

Given A.2 and A.3, by isolating the original noiseless signal s, the relationship between the observed noisy pixels in both blocks then becomes:

$$a_{ij} = \eta b_{ij} \qquad \text{where} \qquad \eta = \frac{\eta_1}{\eta_2} \tag{A.4}$$

where the single noise term η is produced by the division of the two Rayleigh distributed random variables, which for multiplicative noise has the pdf given in [71], defined as:

$$p(\eta) = \frac{2\eta}{\left(\eta^2 + 1\right)^2} \quad \text{where} \quad \eta > 0 \tag{A.5}$$

Finally, the conditional probability density function for this noise distribution is given in [20] as:

$$p\left(\boldsymbol{a}_i|\boldsymbol{b}_i,\boldsymbol{d}_i\right) = \prod_{j=1}^{MN}\left\{\frac{2(a_{ij}/b_{ij})^2}{\left((a_{ij}/b_{ij})^2+1\right)^2}\right\} \tag{A.6}$$

where the nominator term is squared to ensure the peak corresponds to identical pixel values, as shown in Figure A.1. Maximising the above, determines the likelihood of locating $\boldsymbol{b}_i$ in f_{t+1} given the original $\boldsymbol{a}_i$ in f_t and a displacement $\boldsymbol{d}_i$. Motion estimation based on this equation is known as the CD_2 measure, named by Cohen and Dinstein [20].

Figure A.1 shows plots of the Rayleigh distribution in (A.2) for η_1 and η_2 (as a black line), the probability density function in (A.5) (as a blue starred dotted line), and the final probability density function in (A.6) (as a red crossed dashed line).

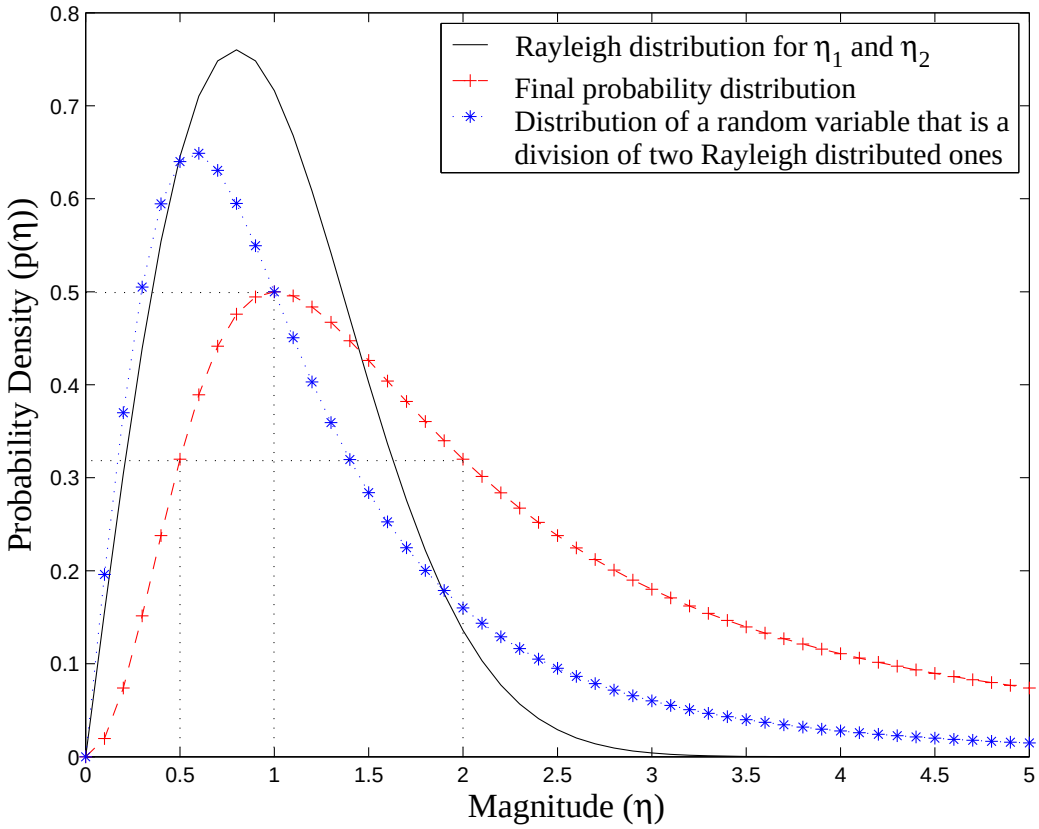


Figure A.1: Plots of a Rayleigh distribution, the final probability distribution, and the distribution of a random variable that is a division of two independent Rayleigh distributed random variables.

Using (A.6), cases of pixel matching in blocks $\boldsymbol{a}$ and $\boldsymbol{b}$ are demonstrated in Table A.1. For example, if pixel $a_{ij} = b_{ij}$ the probability is 0.5, therefore, pixels that are a corresponding match produce a maximum likelihood value (closer to the peak).

a_{ij}	b_{ij}	a_{ij}/b_{ij}	$p(\eta)$
100	100	1.0	0.5
75	80	0.94	0.49
80	75	1.07	0.49
100	200	0.5	0.32
200	100	2.0	0.32
100	20	5.0	0.07
20	100	0.2	0.07

Table A.1: Example pixel values with corresponding probabilities of similarity using the final probability distribution shown as a red crossed dashed line in Figure A.1.

In summary, the CD_2 measure gives the ratio of difference between two blocks, where the only difference is speckle noise, assuming that blocks would otherwise be identical but shifted in time. A block matching framework can compute this for 2D blocks, where the maximum likelihood of all pixels in a block indicates the best displacement $\boldsymbol{d}_i$ for $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$. The CD_2 measure achieves the best displacement estimates for regions that closely conform to a Rayleigh distribution, however, strong signals from structured subjects (such as tendons) and internal hardware processing can alter the signal statistics, which is detrimental to the accuracy of the CD_2 measure. This is problem is tackled with our alternating similarity strategy.

Appendix B

List of Publications

Applied review of ultrasound image feature extraction methods

J. Revell, M. Mirmehdi and D. McNally.

We present an analytical review of specific methods of feature extraction in ultrasound images and implement some of these methods to evaluate their use for our own application, which is the identification of the pulmarus longus tendon. We conclude with the Hough transform, a method that we will use as the basis of future work for our application.

In: 6th Medical Image Understanding and Analysis Conference, A Houston and R Zwiggelaar, editors, pages 173-176. BMVA Press, July 2002.

Variable sized block matching for in vivo musculoskeletal motion analysis

J. Revell, M. Mirmehdi and D. McNally.

We propose a 2D variable sized block matching algorithm with hierarchical exhaustive search to examine tissue deformations in dynamic musculoskeletal ultrasonograhy. Novel aspects include improved matching by updating reference and candidate blocks at each scale and the application area. Performance is quantified on controlled *in vitro* gold standard (groundtruth) sequences and clinical *in vivo* data. We extend the process by refining displacements to sub-pixel accuracy. The proposed technique is validated, by application, to yield quantitatively reliable results.

In: 1st International Conference on Visual Information Engineering, pages 230-233. IEE, July 2003.

Motion trajectories for ultrasound displacement quantification

J. Revell, M. Mirmehdi and D. McNally.

We present a robust methodology to quantify displacements in musculoskeletal ultrasound sequences. This paper extends the principles of 2D interframe displacements produced by our earlier work using hierarchical variable block size matching, to quantify displacement *trajectories.* We provide novel solutions for probe motion, quantification of objects moving in the 3D volume traversing the 2D plane, and improving the temporal coherence of displacements for typical captured sequences, direct from modern ultrasound machines.

In: 7th Medical Image Understanding and Analysis, D. Barber, editor, pages 193-196. BMVA Press, July 2003.

Strain quantification in ultrasound sequences

J. Revell, M. Mirmehdi and D. McNally.

A novel methodology to quantify displacements and strain in musculoskeletal ultrasound sequences is presented. We extend the principles of 2D interframe displacements produced by our earlier work using hierarchical variable block size matching, to quantify displacement *trajectories.* We provide novel solutions for probe motion, quantification of objects moving in the 3D volume traversing the 2D plane, and improving the temporal coherence of displacements for longer image sequences than the frame pairs traditionally applied in ultrasound. We also present trajectory strain that yields a novel strain history for musculoskeletal tendon tissue samples.

In: 14th British Machine Vision Conference, pages 359-368. BMVA Press, September 2003.

Combined Ultrasound Speckle Pattern Similarity Measures

J. Revell, M. Mirmehdi and D. McNally.

We present an enhanced block matching approach to improve displacement accuracy in ultrasound sequences using a combination of matching measures. The first measure uses the normalised cross correlation for regions of strong signal and the second measure CD_2, specifically for regions of speckle determined by the speckle signal to noise ratio. We also show displacement field results for simulated speckle and *in vitro* data.

In: 8th Medical Image Understanding and Analysis, pages 149-153. BMVA Press, September 2004.

Musculoskeletal motion flow fields using hierarchical variable-sized block matching in ultrasonographic video sequences

J. Revell, M. Mirmehdi and D. McNally.

We examine tissue deformations using non-invasive dynamic musculoskeletal ultrasonograhy, and quantify its performance on controlled *in vitro* gold standard (groundtruth) sequences followed by clinical *in vivo* data. The proposed approach employs a two-dimensional variable-sized block matching algorithm with a hierarchical full search. We extend this process by refining displacements to sub-pixel accuracy. We show by application that this technique yields quantitatively reliable results.

In: Journal of Biomechanics, volume 37(4): 511-522, April 2004.

Computer Vision Elastography: Speckle adaptive motion estimation for elastography using ultrasound sequences

J. Revell, M. Mirmehdi and D. McNally.

We present the development and validation of an image based speckle tracking methodology, for determining temporal 2D axial and lateral displacement and strain fields from ultrasound video streams. We refine a multiple scale region matching approach incorporating novel solutions to known speckle tracking problems. Key contributions include automatic similarity measure selection to adapt to varying speckle density, quantifying trajectory fields and spatio-temporal elastograms. Results are validated using tissue mimicking phantoms and *in vitro* data, before applying them to *in vivo* musculoskeletal ultrasound sequences. The method presented has the potential to improve clinical knowledge of tendon pathology from carpel tunnel syndrome, inflammation from implants, sport injuries, and many others.

Accepted in IEEE Transactions in Medical Imaging, 2004.

Bibliography

[1] K.Z. Abd-Elmoniem, A. Youssef, and Y. 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.

[2] E. Adelson and J. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, 2:284–299, 1985.

[3] Y. Akgul, C. Kambhamettu, and M. Stone. Automatic extraction and tracking of the tongue contours. *IEEE Transactions on Medical Imaging*, 18(10):1035–1045, 1999.

[4] L. Alparone, M. Barni, F. Bartolini, and V. Cappellini. Adaptively weighted vector-median filters for motion-field smoothing. In *IEEE International Conference on Acoustic Speech and Signal Processing*, pages 2267–2270, 1996.

[5] M. Anderson and G. Trahey. A seminar on k-space applied to medical ultrasound. Technical report, Department of Biomedical Engineering, Duke University, 1999.

[6] J. Bamber. Ultrasound elasticity imaging: Definition and technology. *Journal of European Radiology*, 9(Suppl. 3):327–330, 1999.

[7] J. Bamber and N. Bush. Freehand elasticity imaging using speckle decorrelation rate. In *22nd Acoustic Imaging*, pages 285–292. Plenum Press, New York, 1996.

[8] P. Baraldi, A. Sarti, C. Lamberti, A. Prandini, and F. Sgallari. Evaluation of differential optical flow techniques on synthesized echo images. *IEEE Transactions on Biomedical Engineering*, 43(3):259–272, 1996.

[9] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. In *12th International Joint Conference of Computer Vision*, pages 43–77, 1994.

[10] M. Benjamin and J. Ralphs. Fibrocartilage in tendons and ligaments - an adaptation to compressive load. *Journal of Anatomy*, 193(4):481–494, 1998.

[11] S. Berg, H. Torp, D. Martens, E. Steen, S. Samstad, I. Hivik, and B. Olstad. Dynamic three-dimensional freehand echocardiography using raw digital ultrasound data. *Ultrasound in Medicine and Biology*, 25(5):745–753, 1999.

[12] L. Bohs, B. Friemel, and G. Trahey. Experimental velocity profiles and volumetric flow via two-dimensional speckle tracking. *Ultrasound in Medicine and Biology*, 21(7):885–898, 1995.

[13] D. Boukerroui, A. Noble, and M. Brady. Velocity estimation in ultrasound images: A block matching approach. In *18th International Conference on Information Processing in Medical Imaging*, volume 2732 of *Lecture Notes in Computer Science*, pages 586–598. Springer-Verlag, 2003.

[14] T. Camus. Real-time quantized optical flow. *Journal of Real-Time Imaging*, 3:71–86, 1997.

[15] C.-M. Chen, H.-S. Lu, and K.-C. Han. A textural approach based on gabor functions for texture edge detection in ultrasound images. *Ultrasound in Medicine and Biology*, 27(4):515–534, 2001.

[16] D.-R. Chen, R.-F. Chang, W.-J. Wu, W.K. Moon, and W.-L. Wu. 3D breast ultrasound segmentation using active contour model. *Ultrasound in Medicine and Biology*, 29(7):1017–1026, 2003.

[17] H. Chung, P Cheung, and N. Yung. Adaptive search center non-linear three step search. In *IEEE International Conference on Image Processing*, volume 2, pages 191–194, 1998.

[18] Y. Chunke, K. Terada, and S. Oe. Motion analysis of echocardiograph using optical flow method. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 1, pages 672–677, 1996.

[19] N. Clark. Design and performance of a joint transform correlator (JTC) system for measuring blood flow and tissue motion. In *Proceedings of SPIE*, volume 2390, pages 44–53, 1995.

[20] B. Cohen and I. Dinstein. New maximum likelihood motion estimation schemes for noisy ultrasound images. In *16th International Conference on Pattern Recognition*, pages 455–463. IEEE CS-Press, 2002.

[21] D. Cooper and J. Graham. Estimating motion in noisy, textured image: Optical flow in medical ultrasound. In *7th British Machine Vision Conference*, pages 585–594. BMVA Press, 1996.

[22] D. Cooper, B. Madsen, and J. Graham. Estimating motion in ultrasound images of the small bowel: Optical flow without image structure. *Lecture Notes in Computer Science*, 2749(7):1017–1026, 2003.

[23] J. Couch. The tendon of achilles. *Canadian Medical Association*, 34:688, 1936.

[24] J. D'hooge, A. Heimdal, F. Jamal, T. Kukulski, B. Bijnens, F. Rademakers, L. Hatle, P. Suetens, and G. R. Sutherland. Regional strain and strain rate measurements by cardiac ultrasound: Principles, implementation and limitations. *European Journal of Echocardiography*, 1:154–170, 2000.

[25] J. D'hooge, E. Konofagou, F. Jamal, A. Heimdal, L. Barrios, B. Bijnens, J. Thoen, F. Van de Werf, G. Southerland, and P. Suetens. Two-Dimensional ultrasonic strain rate measurements of the human heart in vivo. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 49(2):281–286, 2002.

[26] A. Evans and M. Nixon. Biased motion-adaptive temporal filtering for speckle reduction in echocardiography. *IEEE Transactions on Medical Imaging*, 15(1):39–, 1996.

[27] R.T. Fenner. *Engineering elasticity: Application of numerical and analytical techniques.* Ellis Horwood, 1986.

[28] N. Friedland and D. Adam. Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing. *IEEE Transactions on Medical Imaging*, 8(4):344–353, 1989.

[29] V. Frost, J. Stiles, K. Shanmugan, and J. Holtzman. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4:157–166, 1982.

[30] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills. Spatio-temporal approaches to computation of optical flow. In *9th British Machine Vision Conference*, pages 87–95. BMVA Press, 1998.

[31] R. Gill, H. Lim, R. Shatford, E. Toth, M. Voor, and T. Tsai. A comparative analysis of the six-strand double-loop flexor tendon repair and three other techniques: a human cadaveric study. *American Journal of Hand Surgery*, 24(6):1315–1322, 1999.

[32] G. Gillis and A. Biewener. Hindlimb extensor muscle function during jumping and swimming in the toad (bufo marinus). *Journal of Experimental Biology 203 Pt 23*, pages 3547–3563, 2000.

[33] S. Golemati, A. Sassano, M. Lever, A. Bharath, S. Dhanjil, and A. Nicolaides. Carotid artery wall motion in humans, estimated from B-mode ultrasound images. *The Journal of Physiology*, pages 525–536, 2000.

[34] Medical Multimedia Group. http://www.medicalmultimediagroup.com/, 2004.

[35] X. Hao, S. Gao, and X. Gao. A novel multiscale nonlinear thresholding method for ultrasonic speckle. *IEEE Transactions on Medical Imaging*, 18(10):787–794, 1999.

[36] D. Heeger. Model for the extraction of image flow. *Journal of the Optical Society of America*, 4:1455–1471, 1987.

[37] D. Heeger. Optical flow using spatio-temporal filters. *International Journal of Computer Vision*, 1:279–302, 1988.

[38] B. Horn and B. Schunck. Determining optical flow. In *Artificial Intelligence*, volume 17, pages 185–203, 1981.

[39] D. Hughes and F. Duck. Automatic attenuation compensation for ultrasonic imaging. *Ultrasound in Medicine and Biology*, 23(5):651–664, 1997.

[40] J. Jain and A. Jain. Displacement measurement and its application in interframe image coding. *IEEE Transactions on Communications*, 29(12):1799–1808, 1981.

[41] R. Jain and I. Sethi. Finding trajectories of feature points in a monocular image sequence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(1):56–73, 1987.

[42] J. Jenson and P. Munk. Computer phantoms for simulating ultrasound B-mode and CFM images. In S. Lees and L. Ferrari, editors, *23rd Acoustical Imaging Symposium*, pages 75–80, 1997.

[43] L. Jozsa and P. Kannus. *Human Tendons: Anatomy, Physiology, and Pathology*. Human Kinetics. Champaign, Illinois, 1997.

[44] J. Kastelic, A. Galeski, and E. Baer. Hierarchical structure of tendon. *Connective Tissue Research*, 6:11–23, 1978.

[45] J. Klingler, C. Vaughan, T. Fraker, and L. Andrews. Segmentation of echocardiographic images using mathematical morphology. *IEEE Transactions on Biomedical Engineering*, 35(11):925–934, 1988.

[46] E. Konofagou and J. Ophir. A new elastographic method for estimation and imaging of lateral displacements, lateral strains, corrected axial strains and Poisson's ratio in tissues. *Ultrasound in Medicine and Biology*, 24(8):1183–1199, 1998.

[47] K. Kudo, H. Kanehisa, Y. Kawakami, and T. Fukanaga. Growth changes in the elastic properties of human tendon structures. *International Journal of Sports Medicine*, 22(2):138–143, 2001.

[48] M. Ledesma-Carbayo, J. Kybic, M. Sühling, P. Hunziker, M. Desco, A. Santos, and M. Unser. Cardiac ultrasound motion detection by elastic registration exploiting temporal coherence. In *IEEE International Symposium on Biomedical Imaging*, pages 585–588, 2002.

[49] R. Lerner, R. Huang, and K. Parker. Sonoelasticity images derived from ultrasound signals in mechanically vibrated tissues. *Ultrasound in Medicine and Biology*, 16:231–239, 1990.

[50] N. Lin, W. Yu, and J. Duncan. Combinative multi-scale level set framework for echocardiographic image segmentation. In *Journal of Medical Image Analysis*, volume 7, pages 529–537. Elsevier, 2003.

[51] H. Liu, T. Hong, M. Herman, T. Camus, and R. Chellappa. Accuracy vs. efficiency trade-offs in optical flow algorithms. *Journal of Computer Vision and Image Understanding*, 72(3):271–286, 1998.

[52] T. Loupas, W. McDicken, and P. Allan. An adaptive weighted median filter for speckle suppression in medical ultrasound images. *IEEE Transactions on Circuits and Systems*, 36(1):129–135, 1989.

[53] M. Lubinski, S. Emelianov, K. Raghavan, A. Yagle, A. Skovoroda, and M. O'Donnell. Lateral displacement estimation using tissue incompressibility. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 43:247–256, 1996.

[54] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings DARPA Image Understanding Workshop*, pages 121–130. Morgan Kaufmann, 1981.

[55] C. Maganaris and J. Paul. Rapid report: In vivo human tendon mechanical properties. *Journal of Physiology*, 1(521):307–313, 1999.

[56] G. Mailloux, M. Bertrand, and R. Stamfler. Local histogram information content of ultrasound b-mode echographic texture. *Ultrasound in Medicine and Biology*, 11(5):743–750, 1985.

[57] G. Mailloux, A. Bleau, M. Bertrand, and R. Petitclerc. Computer analysis of heart motion from two-dimensional echocardiograms. *IEEE Transactions on Biomedical Engineering*, 34(5):356–364, 1987.

[58] G. Mailloux, F. Langlois, P. Simard, and M. Bertrand. Restoration of the velocity field of the heart from two-dimensional echocardiograms. *IEEE Transactions on Medical Imaging*, 8(2):143–153, 1989.

[59] I. Matthews, T. Ishikawa, and S. Baker. The template update problem. In *14th British Machine Vision Conference*, pages 649–658. BMVA Press, 2003.

[60] L. Meershoek and J. Lanovaz. Sensitivity analysis and application to trotting of a noninvasive method to calculate flexor tendon forces in the equine forelimb. *American Journal of Veterinary Research*, 62(10):1594–1598, 2001.

[61] E. Meijering, W. Niessen, and M. Viergever. Quantitative evaluation of convolution-based methods for medical image interpolation. *Journal of Medical Image Analysis*, 5(2):111–126, 2001.

[62] J. Meunier and M. Bertrand. Ultrasonic texture motion analysis theory and simulation. *IEEE Transactions on Medical Imaging*, 14(2):293–300, 1995.

[63] N. Miller, J. Bamber, and P. Meaney. Fundamental limitations of noninvasive temperature imaging by means of ultrasound echo strain estimation. *Ultrasound in Medicine and Biology*, 28(10):1319–1333, 2002.

[64] W. Moritz, A. Pearlman, D. McCabe, D. Medema, M. Ainsworth, and M. Boles. An ultrasonic technique for imaging the ventricle in three dimensions and calculating its volume. *IEEE Transactions on Biomedical Engineering*, 30(8):482–492, 1983.

[65] A. Morsy and O. Ramm. 3D ultrasound tissue motion tracking using correlation search. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 20:151–159, 1998.

[66] H. Mortensen, O. Skov, and P. Jensen. Early motion of the ankle after operative treatment of a rupture of the achilles tendon. A prospective, randomized clinical and radiographic study. *American Journal of Bone and Joint Surgery*, 81(7):983–990, 1999.

[67] L. Murray, M. Rashidus, and G. Zhenyu. An approach to image classification in ultrasound. In *Proceedings of the 29th IEEE Workshop on Applied Imagery Pattern Recognition*, pages 193–200. IEEE Computer Society, 2000.

[68] D Nair and L Wenzel. Image processing and low-discrepency sequences. *SPIE Advanced Signal Processing Algorithms, Architectures, and Implementations*, 3807(6):102–111, 1999.

[69] K. Nightingale, R. Nightingale, M. Palmeri, and G. Trahey. A finite element model of remote palpation of breast lesions using ultrasonic radiation force: factors affecting tissue displacement. *Ultrasonic Imaging*, 22(1):35–54, 2000.

[70] J. Ophir, I. Céspedes, H. Ponnekanti, Y. Yazdi, and X. Li. Elastography: A quantitative method for imaging the elasticity of biological tissues. *Ultrasonic Imaging*, 13:111–134, 1991.

[71] A. Papoulis. *Probability, random variables and stochastic processes.* McGraw Hill, New York, 1991.

[72] K. Parker, S. Huang, R. Musulin, and R. Lerner. Tissue response to mechanical vibrations for sonoelastic imaging. *Ultrasound in Medicine and Biology*, 16:241–246, 1990.

[73] I. Patras, E. Hendriks, and R. Lagendijk. Confidence measures for block matching motion estimation. In *International Conference on Image Processing*, volume 2, pages 277–280, 2002.

[74] P. Perona and J. Malik. Scale-space and edge detection using anistropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[75] A. Pesavento, A. Lorenz, S. Siebers, and H. Ermert. New real-time strain imaging concepts using diagnostic ultrasound. *Physics in Medicine and Biology*, 45(6):1423–1435, 2000.

[76] M. Petrou and P. Bosdogianni. *Image processing the fundamentals.* John Wiley, 1999.

[77] B. Potocnik and D. Zazula. Automated analysis of a sequence of ovarian ultrasound images. Part 1: Segmentation of single 2D images. *Journal Image and Vision Computing*, 20:217–225, 2002.

[78] R. Prager, A. Gee, G. Treece, and L. Berman. Speckle detection in ultrasound images using first order statistics. Technical Report CUED/F-INFENG/TR 415, Department of Engineering, University of Cambridge, UK, 2001.

[79] I. Pratikakis, C. Barillot, and P. Hellier. Robust multi-scale non-rigid registration of 3D ultrasound images. In *Third International Conference of Scale-Space and Morphology*

in Computer Vision, volume 2106 of *Lecture Notes in Computer Science*, pages 389–397. Springer-Verlag, 2001.

[80] M. Proesmans, L. Van Gool, E. Pauwels, and A. Oosterlinck. Determination of optical flow and its discontinuities using non-linear diffusion. In *3rd European Conference on Computer Vision*, pages 295–304, 1994.

[81] J. Ribas-Corbera and D. Neuhoff. On the optimal block size for block based, motion compensated video coders. In *SPIE Proceedings of Visual Communications and Image Processing*, volume 3024, pages 1132–1143, 1997.

[82] D. Riemersma, A. van den Bogert, M. Jansen, and H. Schamhardt. Tendon strain in the forelimbs as a function of gait and ground characteristics and in vitro limb loading in ponies. *Equine Veterinary Journal*, 28(2):133–138, 1996.

[83] P. Rosin. Thresholding for change detection. In *8th British Machine Vision Conference*, pages 212–221. BMVA Press, 1997.

[84] V. Ruiz, B. Fotopoulos, A. Skodras, and A. Constantinides. An 8x8 block based motion estimation using kalman filter. In *IEEE International Conference on Image Processing*, volume 2, pages 140–143, 1997.

[85] L. Sandrin, M. Tanter, J.-L. Gennisson, S. Catheline, and M. Fink. Shear elasticity probe for soft tissues with 1D transient elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 49(4):436–446, 2002.

[86] R. Seroussi, M. Krag, D. Muller, and M. Pope. Internal deformations of intact and denucleated human lumbar discs subjected to compression, flexion, and extension loads. *Journal of Orthopaedic Research*, 7(1):122–131, 1989.

[87] B. Shapo, J. Crowe, R. Erkamp, S. Emelianav, M. Eberle, and M. O'Donnell. Strain imaging of coronary arteries with intraluminal ultrasound: Experiments on an inhomogeneous phantom. *Ultrasonic Imaging*, 18:173–191, 1996.

[88] J. Shipley, F. Duck, and B. Thomas. Registration of ultrasound breast images acquired from a conical geometry. In *7th Medical Image Understanding and Analysis Conference*, pages 45–48. BMVA Press, 2003.

[89] A. Sigfridsson, T. Ebbers, E. Heiberg, and L. Wigstrm. Tensor field visualisation using adaptive filtering of noise fields combined with glyph rendering. In *IEEE Visualisation*, pages 371–378. IEEE Computer Society, 2002.

[90] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision.* Chapman and Hall, London, 1993.

[91] M. Strintzis and I. Kokkinidis. Maximum likelihood motion estimation in ultrasound image sequences. *IEEE Signal Processing Letters*, 4(6):156–157, 1997.

[92] M. Sühling, M. Arigovindan, P. Hunziker, and M. Unser. Motion analysis of echocardiograms using a local-affine, spatio-temporal model. In *IEEE International Symposium on Biomedical Imaging*, volume 2, pages 573–576, 2002.

[93] M. Sühling, M. Arigovindan, C. Jansen, P. Hunziker, and M. Unser. Myocardial motion analysis and visualization from echocardiograms. In *Progress in Biomedical Optics and Imaging, vol. 4, no. 23*, volume 5032 of *Proceedings of the SPIE International Symposium on Medical Imaging: Image Processing (MI'03)*, pages 306–313, 2003.

[94] A.M. Tekalp. *Digital Video Processing.* Prentice Hall Signal Processing Series, 1995.

[95] J. Thijssen and B. Oosterveld. Texture in tissue echograms: Speckle or information? *American Institute of Ultrasound in Medicine*, 9:215–229, 1990.

[96] G. Thomas. Television motion measurement for datv and other applications. Technical Report 1987/11, BBC Research Department Report RD, 1987.

[97] C. Tomasi and T. Kanade. Detection and tracking of point features. Technical Report CMU-CS-91-132, Carnegie Mellon University, 1991.

[98] G. Treece, R. Prager, A. Gee, and L. Berman. Correction of probe pressure artifacts in freehand 3D ultrasound. In *Medical Image Computing and Computer-Assisted Intervention*, volume 2208 of *Lecture Notes in Computer Science*, pages 283–290. Springer-Verlag, 2001.

[99] J. Trobaugh, D. Trobaugh, and W. Richard. Three-dimensional imaging with stereotactic ultrasonography. *Computerized Medical Imaging and Graphics*, 18(5):315–323, 1994.

[100] F. Valckx and J. Thijssen. Characterization of echographic image texture by co-occurrence matrix parameters. *Ultrasound in Medicine and Biology*, 23(4):559–571, 1997.

[101] T. Varghese, E. Konofagou, J. Ophir, S. Alam, and M. Bilgen. Direct strain estimation in elastography using spectral cross-correlation. *Ultrasound in Medicine and Biology*, 26(9):1525–1537, 2000.

[102] T. Varghese and J. Ophir. Estimating tissue strain from signal decorrelation using the correlation coefficient. *Ultrasound in Medicine and Biology*, 22:1249–1254, 1996.

[103] T. Varghese and J. Ophir. Multiresolution imaging in elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 45:65–75, 1998.

[104] T. Varghese, J. Ophir, and I. Céspedes. Noise reduction in elastograms using temporal stretching with multicompression averaging. *Ultrasound in Medicine and Biology*, 22(8):1043–1052, 1996.

[105] M. Vogt, S. Scharenberg, R. Scharenberg, K. Hoffmann, P. Altmeyer, and H. Ermert. A high frequency ultrasound elastography system for in vivo skin elasticity imaging. In *5th World Congress on Ultrasonics*, pages 393–396, 2003.

[106] R. Wagner, S. Smith, and J. Sandrik. Statistics of speckle in ultrasound B-scans. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):156–163, 1983.

[107] C. Wang, J-M. Deng, G. Ateshian, and C. Hung. An automated approach for direct measurement of two-dimensional strain distributions within articular cartilage under unconfined compression. *Journal of Biomechanical Engineering*, 124:557–567, 2002.

[108] D. Weinstein, G. Kindlmann, and E. Lundberg. Tensorlines: Advection-diffusion based propagation through diffusion tensor fields. In *IEEE Visualisation*, pages 249–253. IEEE Computer Society, 1999.

[109] D. Wisleder, S. Werner, W. Kraemer, S. Fleck, and V. Zatsiorsky. A method to study lumbar spine response to axial compression during magnetic resonance imaging: technical note. *Spine*, 26(18):416–420, 2001.

[110] F. Yeung, S. Levinson, D. Fu, and K. Parker. Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh. *IEEE Transactions on Medical Imaging*, 17(6):945–956, 1998.

[111] F. Yeung, S. Levinson, and K. Parker. Multilevel and motion-based ultrasonic speckle tracking algorithms. *Ultrasound in Medicine and Biology*, 24(3):427–441, 1998.

[112] Y. Yu and S. Acton. Segmentation of ultrasound imagery using anisotropic diffusion. In *35th Asilomar Conference on Signals, Systems and Computers*, pages 1151–1155, 2001.

[113] Y. Yu and S. Acton. Speckle reducing anisotropic diffusion. *IEEE Transactions on Image Processing*, 11(11):1260–1270, 2002.

[114] X. Zong, E. Geiser, A. Laine, and D. Wilson. Homomorphic wavelet shrinkage and feature emphasis for speckle reduction and enhancement of echocardiographic images. *Proceedings of SPIE, Medical Imaging*, 2710:658–667, 1996.