

Motion Estimation and Registration of B-Mode Ultrasound Images for Improved Visualisation

By

Yazan Awwad, B.Eng.

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of
Master of Applied Science

The Ottawa-Carleton Institute for
Biomedical Engineering

Department of Systems and Computer Engineering
Carleton University
Ottawa, Ontario, Canada

©2014
Yazan Awwad

Abstract

Medical imaging is an important aspect in today's life. Different imaging techniques are used for diagnostic purposes which help in the early detection of diseases before they develop into more serious issues and cause more damage. Some of these imaging modalities are Magnetic Resonance Imaging (MRI), X-ray, Computed Tomography (CT-scan) and Ultrasound Imaging (US). Ultrasound imaging is known to be the safest among the other imaging modalities since it doesn't send ionised radiation. This thesis presents a method for motion tracking between ultrasound images that originated from moving an ultrasound probe. The motion tracking is done in two steps: In the first step, a modified block-matching search is used to track the motion between frames by tracking the speckle and the tissues in the images. From the modified block-matching approach, the global motion vector can be seen. In the second step, areas with motion vectors in the direction of the probes' motion are put through another motion tracking algorithm called phase correlation. Doing so will improve the overall motion estimation. Phase correlation provides a smoother vector field compared to block-matching. Combining the modified block-matching with phase correlation will increase the accuracy for motion estimation. Finally, after the motion estimation is performed successfully, image registration and averaging is performed to obtain an extended field-of-view-frame. This extended field-of-view shows the information contained in all the separate frames, in one single frame while reducing the overall noise.

Acknowledgments

I would like to thank my two supervisors, Dr. Chris Joslin and Dr. Yuu Ono for their help and support throughout the period the research was conducted. Without their help, this work would have not been possible or finished. Their guidance, comments and suggestions were of extreme help to me throughout the period of the research.

I would like to also thank Dr. Andy Adler for letting us use his ultrasound machine. I would like to also thank all my colleagues at Carleton University for their support and motivation throughout the two years of my program. Special thanks are afforded to Andy Huang whom helped me in using the ultrasound machine for all my experiments at Dr. Ono's lab, and special thanks to my colleague Mahla Abdollahnejad; she helped me in several experiments that involved the design of phantoms. I also would like to thank the department of Systems and Computer Engineering for the nice and comfortable environment they provide to students.

Table of Contents

Abstract	ii
Acknowledgments	iii
Table of Contents	iv
List of Tables	vi
List of Figures	vii
List of Acronyms	x
Chapter 1: Introduction.....	1
1.1. Motivation.....	3
1.2. Objectives.....	8
1.3. Thesis Contributions	8
1.4. Thesis Outline	8
Chapter 2: Background & Literature Review	10
2.1. Ultrasound Imaging.....	10
2.2. Ultrasound Tracking Techniques	16
2.2.1. Tracking by External Devises.....	16
2.2.2. Tracking by Ultrasound images.....	17
2.3. Image Registration in the Medical Field	19
2.3.1. Ultrasound Image Registration	20
2.3.2. Image Registration Techniques.....	22
Chapter 3: Methodology	36
3.1. Overview	36
3.2. Exclusion of No Motion Areas and Areas that cannot be Tracked.....	38
3.3. Estimation of the Motion Using Block Matching	40
3.3.1. Taylor Approximation for Sub-pixel Motion Estimation Accuracy.....	42
3.4. Phase Correlation	45
3.4.1. Integer Based Motion Estimation by Phase Correlation.....	46
3.4.2. Subpixel Based Motion Estimation by Phase Correlation.....	49

3.5.	Ultrasound Image Registration.....	50
3.6.	Averaging the Registered Images	51
3.7.	Summary	53
Chapter 4: Experiments and Results		54
4.1.	Ultrasound Machine	54
4.1.1.	Linear Probe.....	54
4.2.	Phantoms Preparation.....	56
4.2.1.	Phantom Experiment.....	57
4.3.	In Vivo Experiment to Test the Subpixel Phase Correlation	64
4.4.	In Vivo Experiment to Test the Modified Block Matching Approach	73
4.5.	In vivo Experiment to Observe Muscle Contraction.....	88
4.6.	In vivo Experiment to Observe the Effect of Probe compression.....	99
4.7.	Summary and Discussion.....	106
Chapter 5: Conclusions and Future Work.....		111
5.1.	Summary of Contributions	111
5.1.1.	Block-Based Motion Estimation.....	111
5.1.2.	Phase Correlation Motion Estimation.....	112
5.1.3.	Image Registration and Averaging	112
5.2.	Future Work	113

List of Tables

Table 1.1: Advantages and Disadvantages of the Medical Image modalities.....	2
--	---

List of Figures

Figure 1.1: An Ultrasound Phantom of the Hip Joint (left), reproduced from [2]. Normal hip joint (right), reproduced from [3].	3
Figure 1.2: Hip joint with pincer and cam impingement, reproduced from [3].	5
Figure 2.1: Linear Array Transducer.	11
Figure 2.2: Taking an Ultrasound Image on the Hip Region (left). Image of the Hip Region Produced by Linear Probe (right).	12
Figure 2.3: Sector Image Produced on the Biceps Region Using a Sector Transducer.	13
Figure 2.4: Function of the Transmit/Receive Switch.	13
Figure 2.5: Signal Processor Components.	14
Figure 2.6: Image Processor Components.	15
Figure 2.7: An Ultrasound Image Showing Only Speckle Noises.	18
Figure 2.8: Motion Estimation Using Block-Matching.	28
Figure 3.1: Overview of the Work Done in this Thesis.	38
Figure 3.2: An Image Divided into Macro-Blocks to Perform the Block Search on it.	39
Figure 3.3: A Macro-block of Size 16 X 16 Being Searched at Each Possible Location in Predefined Area.	41
Figure 3.4: Motion Vectors Overlaid over a Speckle Noise Image Generated by Ultrasound Imaging.	45
Figure 3.5: (a) Eraser in Designed Ultrasound Phantom. (b) The Eraser in (a) was Shifted by 30 Pixels on the X and Y axis Respectively. (c) The Peak Output from Phase Correlation which Represents the Motion Between (a) and (b).	47
Figure 3.6: (a) Eraser in Ultrasound Phantom. (b) Same Eraser at Different Position due to Probe motion. (c) Same Eraser at another Position due to Probe Motion. (d) Same Eraser at another Position due to Probe Motion. (e) Average of the 180 Registered Images.	52
Figure 4.1: Picus Ultrasound Machine with a Linear Probe.	55
Figure 4.2: First Layer of Gelatine added Some Materials on Top of it.	57
Figure 4.3: Examples of the Frames Obtained on the Eraser in the Phantom.	58
Figure 4.4: Motion Estimation between Eraser Frames on the Y Direction using Integer based Phase Correlation.	59
Figure 4.5: Motion Estimation between Eraser Frames on the Y Direction using Subpixel Based Phase Correlation.	60
Figure 4.6: Motion Estimation between Eraser Frames on the X Direction (depth) using Integer based Phase Correlation.	61
Figure 4.7: Motion Estimation between Eraser Frames on the X Direction using Subpixel based Phase Correlation.	62
Figure 4.8: Sample Registered Images of the Eraser in the Ultrasound Phantom.	63
Figure 4.9: Averaging of the 180 Frames of the Eraser (838 x 440).	64
Figure 4.10: Ultrasound Probe Connected to a Mechanical Arm.	65
Figure 4.11: The Probe moved from Location (a) to Location (b) using a Mechanical Arm.	66
Figure 4.12: Examples of the Frames obtained along the Hip Region while the Probe moved a Distance of 3cm for 4s using a Mechanical Arm.	67

Figure 4.13: Motion Estimation between Hip Region Frames on the Y Direction using Integer Phase Correlation.	68
Figure 4.14: Motion Estimation between Hip Region Frames on the Y Direction using Subpixel Phase Correlation.	69
Figure 4.15: Motion Estimation between Hip Region Frames on the X Direction using Integer Phase Correlation.	70
Figure 4.16: Motion Estimation between Hip Region Frames on the X Direction using Subpixel Phase Correlation.	71
Figure 4.17: Averaging of the 180 Frames of the Hip Region (1025 x 422).	72
Figure 4.18: The Probe Moved from Location (a) to Location (b) while the Motion of the Probe was Provided by the Subject being Tested.	73
Figure 4.19: Motion Estimation between Hip Region Frames on the Y Direction using Integer Phase Correlation.	74
Figure 4.20: Some Outputs obtained using Ultrasound on the Hip Region while the Probe was Moved Manually.	75
Figure 4.21: Macro-block Created on Frame Number 30 of the Hip Experiment (image size 560 x 420).	76
Figure 4.22: Cropped Version of Figure 4.21(image size 560 x 416).	77
Figure 4.23: Motion Vectors Overlaid over Frame 30 (image size 560 x 416).	77
Figure 4.24: (a) Frame 80 of Hip Experiment (Actual size is 560 x 420). (b) Frame 81 of Hip Experiment (Actual size is 560 x 420).	79
Figure 4.25: Motion Vectors Overlaid over Frame 80 of the Hip Experiment.	80
Figure 4.26: Some of the Wrong Motion Vectors Overlaid over Frame 80 of the Hip Experiment.	81
Figure 4.27: Motion Vectors over Frame 80 after Applying the Modified Block Matching Search.	82
Figure 4.28: PSNR for the Original Full Search Block Matching and for the Modified Full Search Block Matching.	84
Figure 4.29: Motion Estimation between Hip Region Frames on the Y Direction using Subpixel Phase Correlation.	85
Figure 4.30: Motion Estimation between Hip Region Frames on the X Direction using Subpixel Phase Correlation.	86
Figure 4.31: The Extended-Field-Of-View Frame of the Hip Experiment.	87
Figure 4.32 : (a) Start Position for the Ultrasound Probe. (b) End Position of the Ultrasound Probe.	88
Figure 4.33: Some Output Frames Obtained During the Experiment of Muscle Contraction. (a) Output while Muscle is Relaxed. (b) Output During Muscle Contraction. (c) Output When Muscles are Fully Contracted. (d)Output when Muscle is back to Relaxation.	90
Figure 4.34: Motion Estimation between Hip Region Frames on the Y Direction using Integer Phase Correlation.	91
Figure 4.35: Frame Number 200 Divided into 16 x 16 Macro-Block from the Experiment Performed on the Upper Part of the Leg (Size 400 x 400).	92
Figure 4.36: (a) Frame 215 of the Upper Leg Experiment. (b) Frame 216 of the Upper Leg Experiment.	93
Figure 4.37: Motion Vectors Overlaid over Frame 215 which Represent the Motion between Frames 215 and 216.	94

Figure 4.38: (a) Frame 238 of the Upper Leg Experiment. (b) Frame 239 of the Upper Leg Experiment.	95
Figure 4.39: Motion Vectors Overlaid over frame 238 which Represent the Motion between Frames 238 and 239.....	96
Figure 4.40: Motion Estimation between Frames on the Y Direction using the Subpixel Phase Correlation Motion Estimation.	97
Figure 4.41: Average of the Registered 900 Frames of the Experiment Performed on the Upper Leg Region (Size of 708 x 428).	98
Figure 4.42 : Some Output Frames obtained During the Experiment of Probe Compression. (a) Output while Muscle is Relaxed with no Compression. (b) Output During the Period Half Way through the Compression Process. (c) Output During the Period of Full Compression Effect of the Probe. (d) Output During back to relaxation mode.....	100
Figure 4.43: Motion Estimation between Hip region Frames on the Y Direction using Integer Phase Correlation.	101
Figure 4.44: (a) Frame 94 of the Upper Leg Experiment to Test Probe Compression Effect. (b) Frame 95 of the Upper Leg Experiment to Test Probe Compression Effect.....	102
Figure 4.45: Motion Vectors Overlaid over Frame 94 which Represent the Motion Vectors between Frames 94 and 95.	103
Figure 4.46: (a) Frame 108 of the Upper Leg Experiment to Test Probe Compression Effect. (b) Frame 109 of the Upper Leg Experiment to Test Probe Compression Effect.....	104
Figure 4.47: Motion Vectors Overlaid over Frame 108 which Represent the Motion Vectors between Frames 108 and 109.	104
Figure 4.48: Motion Estimation between Frames on the Y Direction using the Subpixel Phase Correlation Motion Estimation.	105
Figure 4.49: Average of the Registered 581 Frames of the Experiment Performed on the Upper Leg Region (Size of 559 x 434).	106

List of Acronyms

Abbreviation	Description
MRI	Magnetic Resonance Imaging
X-ray	Fluoroscopy
CT-scan	Computed Tomography
US	Ultrasound
RF	Radio frequency
B-mode	Brightness mode
M-mode	Motion mode
A-mode	Amplitude mode
1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
4D	Four-dimensional
SAD	Sum of Absolute Differences
MAD	Mean of Absolute Differences
MI	Mutual Information
MSE	Mean Squared Error
NCC	Normalized Cross Correlation
LSE	Least Square Estimator
BDM	Block Distortion Measure
ES	Exhaustive Search
TSS	Three Step Search
NTSS	New Three Step Search
SES	Simple and Efficient Search
4SS	Four Step Search
DS	Diamond Search
SVD	Singular Value Decomposition

Chapter 1

Introduction

The use of medical imaging technologies is important in medical procedures. These imaging technologies can be used for diagnostic purposes and are used to guide surgeons through surgeries. It is of high importance that these imaging modalities are accurate to enable surgeons to focus more on the procedure of the surgery rather than on identifying different body tissues which are not usually visible. There are many imaging modalities that could do this job such as Magnetic Resonance Imaging (MRI), fluoroscopy (X-ray), Computed Tomography (CT scan), and Ultrasound (US). For each of these imaging modalities there are advantages and disadvantages. Table 1.1 shows some of the advantages and disadvantages for these imaging modalities.

Table 1.1: Advantages and Disadvantages of the Medical Image Modalities.

Imaging Modality	Advantages	Disadvantages
MRI	<ul style="list-style-type: none"> • Works well in scanning soft-tissues. • Non-invasive. • Clear images. • No ionising radiation. 	<ul style="list-style-type: none"> • High magnetic field. • Long wait times (up to 3-4 months). • Long imaging times (10 minutes to 1 hour). • Expensive.
X-ray	<ul style="list-style-type: none"> • Good for imaging the bones. • Clear outlines. 	<ul style="list-style-type: none"> • Ionised radiation. • Virtually no visibility for soft-tissue.
CT scan	<ul style="list-style-type: none"> • Able to scan all parts of the body. • Faster than MRI (around 5 minutes). • Clear outlines. • Non-invasive. • 3D reconstruction. 	<ul style="list-style-type: none"> • High dose of ionised radiation.
US	<ul style="list-style-type: none"> • No ionising radiations. • Low cost. • Portable. • Short wait time to get it done. • Non-invasive. 	<ul style="list-style-type: none"> • Noisy images.

Among all of these imaging modalities, US is always preferred if it can do the desired job accurately due to the fact that it does not send any ionised radiation into the patient, it is more portable than others, less expensive, doesn't require a long wait time to get it done and it is quick [1,2]. It is important to note that we don't know the long term effects of US imaging. US may

have long term biological effects that we don't know of but it is considered safer than other imaging modalities since it doesn't use ionised radiation. One of the disadvantages for ultrasound imaging is that the images obtained by ultrasound are corrupted by speckle noise which degrades the quality of the images. Speckle noise is inherent in all US images and that makes it possible to use speckle noise as some unique features in the images that can be tracked.

1.1. Motivation

The main motivation for conducting this research is to help in the diagnosis and treatment of a disease that affects the hip joint called Femoro-Acetabular Impingement (FAI). Figure 1.1 shows an ultrasound image of the hip joint and the anatomy of the hip joint. The hip joint consists mainly of the femoral head, the acetabulum, the labrum and the articular cartilage. FAI can injure the labrum and the articular cartilage between the femoral head and the acetabulum which could cause pain and limitation of movement [3].

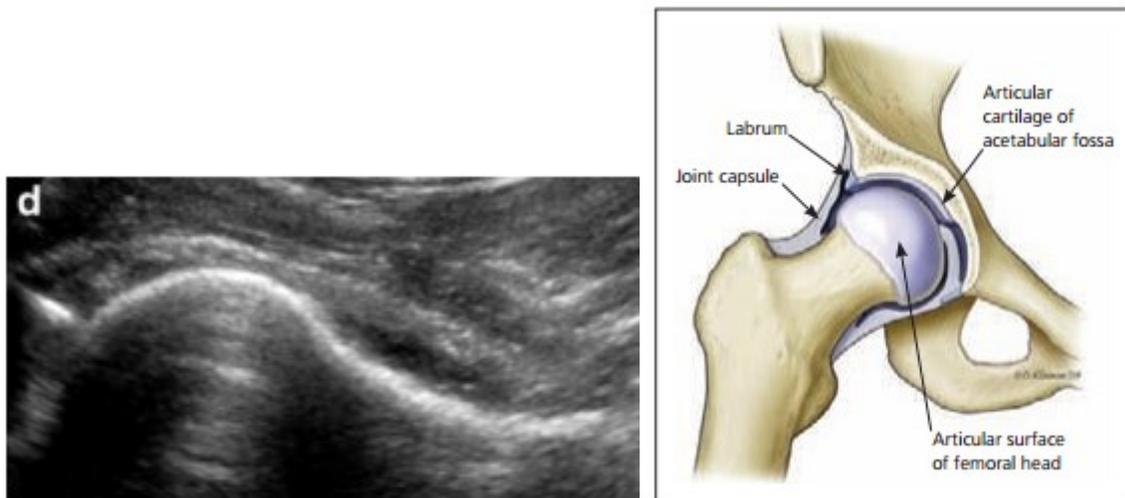


Figure 1.1: An Ultrasound Image of the Hip Joint (left), reproduced from [2]. Normal Hip Joint (right), reproduced from [3].

There are three types of FAI which are: Pincer FAI, Cam FAI and a mix of both. Pincer FAI impingement occurs when an excessive bone grows over the rim of the acetabulum. Pincer impingement usually happens during flexion or internal rotation [3]. However, cam FAI

impingement occurs when the femoral head is not perfectly round [3]. Figure 1.2 shows a hip with a cam and pincer impingement. The major symptoms of FAI are pain in the hip area and reduction in the movement ability at the hip joint. However, if a patient has these symptoms at the hip area, it cannot be concluded that this patient has FAI from just the symptoms. If a patient has the symptoms of FAI, imaging is performed around the patient's hip region to accurately diagnose if the patient has FAI or not, as well as accurately identifying which type of FAI it is so that the proper treatment is given.

X-rays, CT scans and MRI imaging modalities are usually used to diagnose FAI. All of these methods are capable of producing high quality images of the hip joint however; each one of them has its own disadvantages [1]. As was shown in Table 1.1 in the introduction, one of the main disadvantages for X-ray and CT-scans is the ionised radiation that they send to the human body. If used frequently, this type of radiation could be dangerous on the patients. Also, MRI has disadvantages, as shown in Table 1.1, such as being expensive and long waiting time. As a result, one should look for a no risk or lower risk imaging modality that can be used frequently in diagnosing FAI and other diseases without putting the human body in harm's way.

As discussed in the introduction, ultrasound imaging was thought to be the safest between all the other imaging modalities. The radiations, sound waves, which are sent by ultrasound, are non-ionising. These waves are thought to be not harmful to the human body or at least less harmful than the ionised radiations. However, images produced by ultrasound are noisy which makes it harder to diagnose FAI or other diseases but researchers are always looking at ways to improve the quality of these images, so that it becomes widely used in the diagnosing patients. To help researchers in making ultrasound more widely used for diagnosis of diseases, ultrasound imaging was chosen to be used in this thesis.

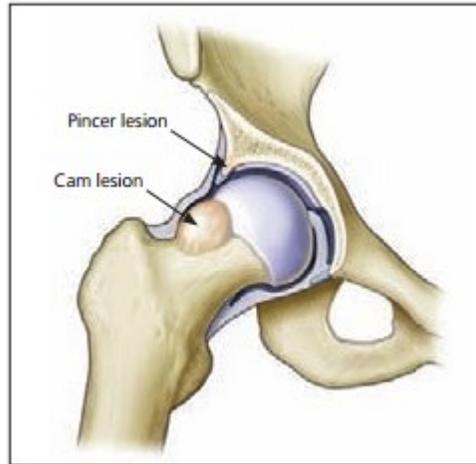


Figure 1.2: Hip joint with pincer and cam impingement, reproduced from [3].

Treatment of FAI varies depending on the degree of damage FAI caused to the hip joint. If damage is small, patients could be asked to change the activities that could lead to increase the damage. But if the damage is big, a surgery would be required to return the hip joint to its normal shape. If a patient has a pincer impingement that caused a large area of excess bones to grow over the acetabulum, the surgery would aim to remove the excess bone that grew over the acetabulum to return the acetabulum to its normal shape. If the patient has a cam impingement, the surgery would aim to return the femoral head into its normal round shape.

Motion tracking of the ultrasound probe through ultrasound images is important and helps physicians in many ways. Motion tracking can be helpful in intra-operative imaging and pre-operative imaging. In intra-operative imaging, if motion tracking is done accurately through US images, it could be used by surgeons as a guide in surgeries which make it easier for surgeons to see tissues that are not usually seen with the human eye. However, motion tracking in intra-operative imaging is beyond the scope of this thesis. This thesis looks at ways of tracking the probe's motion, which can be done through tracking the transformations between the ultrasound images obtained in pre-operative imaging sessions using ultrasound, to obtain a 2D model that visualises a large scale area of the region being imaged. The field-of-view that ultrasound probes can provide is not large when compared to CT scans and MRI. Extending the field-of-view obtained by ultrasound imaging would be a great asset in this field

since it would provide a cost effective and non-invasive alternative to CT scans and MRI imaging. During pre-operative imaging sessions, several ultrasound images can be obtained as the probe is moved from a point to another point. The produced images correspond to the regions scanned by the probe at the location that the probe is pointing at. As the probe is moved, different regions of the scanned area can be seen. To view different parts of the area of interest, which in our case is the hip region, the probe has to be moved across the different parts of the hip region giving an ultrasound image for each part of the hip region. This would require the physician to look at separate images to identify different tissues in the different regions of the hip area. The fact that the physician would need to look at separate images to identify different regions of the area scanned is bothersome, time consuming and increases the chance of human error. As a result, the goal of this thesis is to track the motion between the several images produced in a pre-operative imaging session, so that the images can be aligned with each other and an extended field-of-view frame can be obtained. Also, the extended field-of-view frame would have lower noise than the original frames that were used to obtain the extended field-of-view frame. It is believed that the extended field-of-view would have lower noise since it is obtained from averaging the aligned images. Averaging works like a low pass filter. Motion estimation between the US images has to be done with high accuracy so that the produced extended field-of-view frame accurately displays the different tissues scanned. The procedure described in this thesis for estimating the motion between the ultrasound frames and combining them to give an extended field-of-view would provide several advantages. The first of which, would be having an extended field-of-view frame that visualises a large scale area. Secondly, an extended-field-of-view ultrasound image would accurately provide all the information of several ultrasound images of different regions, in one image with less overall noise than any of the separate images. Finally, an extended-field-of-view ultrasound frame would be of extreme help for physicians to identify different diseases including FAI.

Area based methods were used in this thesis to track the motion between the ultrasound images accurately. The two area-based methods that were thought to be suitable to be used with ultrasound images are a block-matching search and phase correlation. A block-matching motion estimation would be useful in tracking the motion between the ultrasound images because ultrasound images are characterized by noise that is inherent in all the ultrasound images. The intensity values of the noise in the images can be used in area-based methods to

track the motion of the noise between the frames [4]. Therefore, using area-based methods with the ultrasound images would be a good choice since the intensity values of the images are only used to track the motion between the images, without the need of identifying what each tissue represents.

Each US frame that is required to find its motion relative to the subsequent or preceding frame is first to be divided into square blocks of equal size. Each of these square blocks is to be searched within a predefined search area in the reference frame (the frame before or after the current frame) using a suitable search algorithm [6]. Block-matching finds the best matching block between two consecutive frames. The best matching block represents the motion vector of that block [5]. Exhaustive search is to be used to find the best matched block since we are dealing with medical US images and thus we want to obtain the motion with very high accuracy. Exhaustive search would be a good choice since it searches for the best matched block at each possible location in the predefined search area. On the other hand, block-matching could give some false motion vectors if decorrelation between blocks arises and thus the block-matching search is modified to take this fact into account. To improve the accuracy of motion estimation even further, subpixel phase correlation motion estimation is to be used to estimate the motion between the parts of the frames that are observed to have a motion in the direction of the global motion vector.

Phase correlation measures the motion between two fields directly from their phases and is not like block-matching which searches the blocks from luminance matches [7]. Phase correlation gives a smoother vector field and it is more accurate and more computationally efficient [7]. If the modified block-matching approach and phase correlation were used together, the global motion vector would be more accurate than just using one by itself. In this thesis research, both of these motion tracking algorithms are investigated to track the motion through the ultrasound images. From block-matching, false motion vectors will be eliminated to improve the global motion vector field. Then phase correlation will be applied to the compensated images obtained from block-matching to obtain the motion more accurately. After the motion estimation is done accurately, image registration (image alignment) is applied on all the frames using the information obtained from the motion estimation. Finally, all the registered images

for each experiment conducted in this thesis are to be averaged to obtain an extended field-of-view frame.

1.2. Objectives

The main objective of this thesis is to obtain an extended field-of-view ultrasound frame that can be used to visualize a large scale area instead of looking at a multiple of images to see different tissues. To do this, a very accurate and precise motion tracking procedure, using ultrasound images, must be developed. The extended field-of-view frame could be used by physicians in the diagnosis of FAI and other conditions. Experiments are carried out on an ultrasound phantom where known materials were put in the phantom and ultrasound images were obtained on these materials, and afterwards experiments are carried out on the authors' hip region and on the upper leg region to validate the developed procedure.

1.3. Thesis Contributions

A summary of the contributions in this thesis are as follows:

- Modification of the block-matching search using exhaustive search to account for areas of no motion in the US images that could be assigned a false motion vector.
- Modification of the block-matching search to account for macro-blocks that are hard to track due to tissue changing shape from compression or muscle contraction or due to decorrelation of speckle between the frames.
- Evaluation of the modified block-matching search and the subpixel phase correlation search through the use of an ultrasound phantom and through ultrasound images obtained on the hip region and the upper leg region.

1.4. Thesis Outline

This thesis is divided into five chapters organized as follows:

Chapter 2 Background & Literature Review: Provides a review about ultrasound imaging techniques and the different modes of ultrasound images. It discusses issues usually seen in

ultrasound images. It also provides a review about motion estimation algorithms used in the literature and a review of the different image registration techniques that can be used.

Chapter 3 Methodology: It discusses the methodology used to estimate the motion through ultrasound images. It describes the modified block-matching algorithm used in details. It also describes the subpixel phase correlation motion estimation algorithm in details as well.

Chapter 4 Experiments and Results: This chapter presents the experimental set up and results for all the experiments carried out in this research which includes the phantom experiment, the hip region experiments and the upper leg region experiments. It also introduces the characteristics of the ultrasound machine used for this research.

Chapter 5 Conclusions and Future work: This chapter summarizes the work done in this research and it also suggests some aspects for future work.

Chapter 2

Background & Literature Review

2.1. Ultrasound Imaging

In taking ultrasound imaging, there are a number of elements and persons involved, which are the instrument, transducer, patient, sonographer and sinologist. The sonographer is the professional responsible for operating the ultrasound machine to obtain ultrasound images that can be used for diagnosis purposes. The sinologist is the professional responsible for studying the acquired ultrasound images. Ultrasound imaging is based on a pulse-echo technique. The transducer part (Figure 2.1) in ultrasound system produces pulses of ultrasound and sends them to the human body. These pulses then produce echoes at the boundaries of organs inside the body and within the tissues. Then the echoes reflect back to the transducer where it detects them and displays the images on the screen of the ultrasound instrument [8]. The ultrasound instrument is the part that processes echoes that are reflected back and forms an image. This image produced is a gray scale image and the intensity values correspond to the echo strength. The intensity variations produced matches the anatomic structure. As a result, information about the location of different tissues can be found by knowing the direction the pulses entered the human body and by knowing the time it took until the echoes produced came back to the transducer [8].



Figure 2.1: Linear Array Transducer.

Ultrasound transducer sends ultrasonic pulses successively to form an image. Sending only one pulse to the tissues will produce one scan line. This scan line represents the echoes reflected back to the transducer from sending that specific pulse. In a linear array transducer, shown in Figure 2.1, the pulses are sent vertically but each time it is sent from a different start point. Thus, pulses sent from a linear array are all parallel to each other. The reflection of these pulses results in returning many scan lines where each scan line refers to the pulse that was sent and each scan line represents the returned echoes. The sum of all the scan lines that were returned to the transducer, results in forming a cross-sectional image of the anatomy of the region being scanned. This produced image is called a linear image, referring to the linear array used to produce it [8]. An example of a linear image and how it is obtained is shown in Figure 2.2 which is an image produced by imaging the hip region.

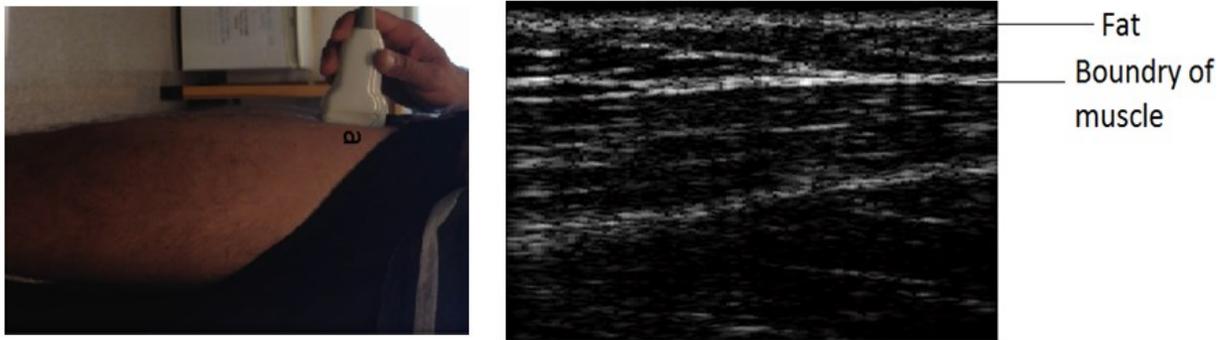


Figure 2.2: Taking an Ultrasound Image on the Hip Region (left). Image of the Hip Region Produced by Linear Probe (right).

Curved array is used to produce ultrasound images which are called sector images. In this method, the transducer generates many pulses from the same starting point but the pulses themselves don't travel in straight as with the linear transducer; they travel in different directions to produce a sector image which has the shape of a slice of a pie as can be seen in Figure 2.3 [8]. This type of transducer was not used in this thesis because the field-of-view of each frame obtained by processing the reflected echoes by the curved array transducer is less than the field-of-view obtained by processing the reflected echoes by the linear array transducer. Also this type of transducers is less comfortable for the person under imaging than the linear transducer as this transducer requires a larger amount of pressure against the skin of the person under testing to obtain an image of sufficient clarity. As the goal in this thesis is to obtain an extended-field-of-view frame that visualizes as large area as possible, a linear transducer was chosen to be used.

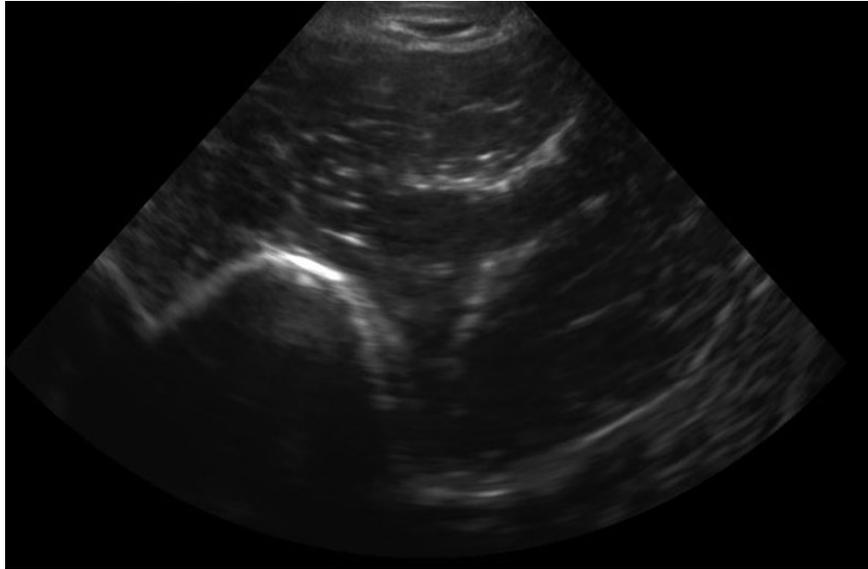


Figure 2.3: Sector Image Produced on the Biceps Region Using a Sector Transducer.

When the echoes are reflected back to the transducer, they will be processed using a beam former, a signal processor, an image processor and a display to display the resulted image. In the beam former, there is an element called Transmit/ Receive (T/R) switch, shown in Figure 2.4. This switch drives the originated electrical pulses to the transducer and then drives the pulses of reflected echoes from the transducer to the amplifiers where the voltage amplitude is increased in order to convert small voltages to larger voltages that can be processed.

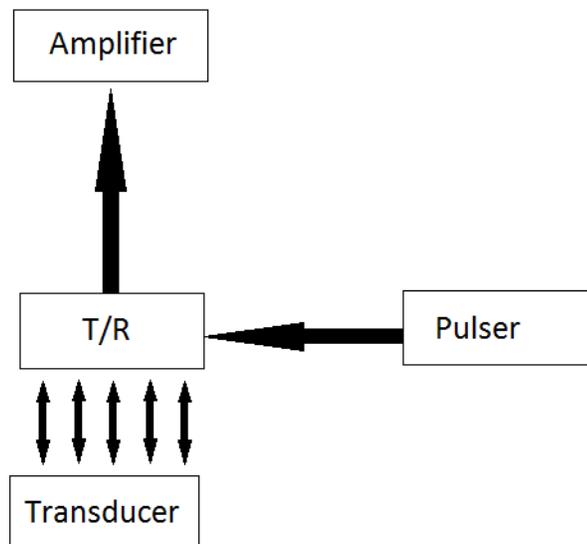


Figure 2.4: Function of the Transmit/Receive Switch.

In the signal processor component, the operations of filtering, detection and compression are carried out as can be seen in Figure 2.5. At this stage, the echo voltages which are in radio frequency (RF) form are converted to a video signals.

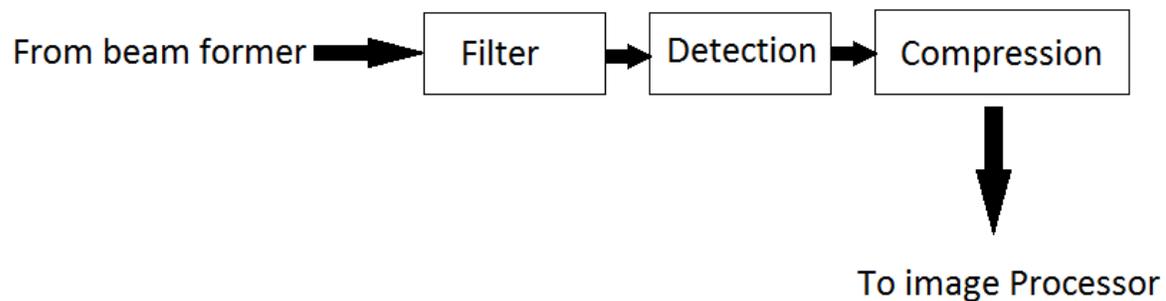


Figure 2.5: Signal Processor Components.

The image processor component, shown in Figure 2.6, is the part responsible for converting the serial scan lines that have been through many processing stages to an actual image that can be displayed on the display screen. When the images are converted into image format, they will be stored in the image memory component of the image processor. After the images have been processed in the image processor, they can be displayed on a display monitor in one of the following modes: B-, M-, or A-modes. The most commonly used one in the clinical environment is the B-mode which stands for Brightness mode and it is also called a gray-scale sonography; this mode is widely used in abdominal and obstetrical ultrasonography [9]. M-mode stands for Motion mode and is used in echocardiography. A-mode stands for Amplitude mode is used in ophthalmology (eyes physiology) [8].

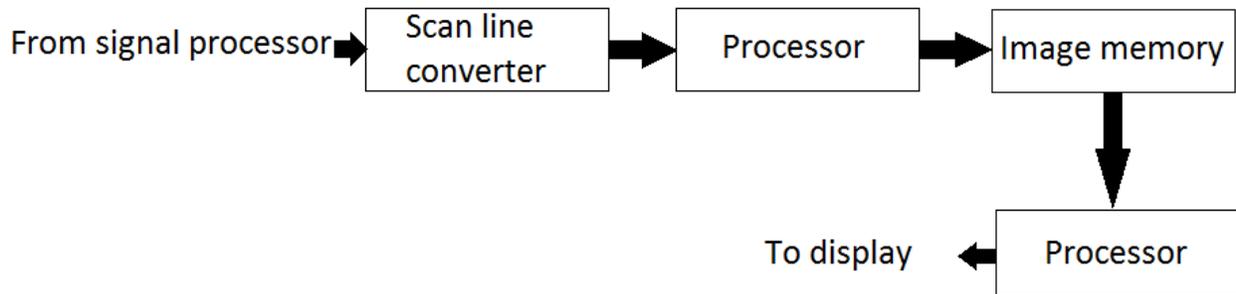


Figure 2.6: Image Processor Components.

In A-mode imaging the ultrasonographer must know where the sound beam is pointed which is generally hard to know. The image of this type looks like a range of mountains and as a result it is hard to see anatomic data from this type of images. M-mode is similar to A-mode, but with a few differences. The M-mode scans measure the position and the motion of the echo while A-mode scans measure the position and amplitude of the echo [9]. In B-mode, the first thing to do is to obtain an A-mode image and replace the spikes with dots. Here the brightness of the dots (intensity of the pixels) corresponds to the amplitude of the echo. A-mode and M-mode were not considered in this thesis because both of them do not provide anatomical structures of the image. B-mode images provide anatomical structures which are represented by different pixel values. By understanding these pixel values and how they change between different frames, it is possible to calculate the motion between frames.

To create a B-mode image, the line of dots will be oriented in the same direction that the ultrasound transducer is pointing. As the transducer is moved, the line of dots representing the reflected echoes will move in the same direction. The display system of the ultrasound will keep a record of all the dots produced as the transducer is moved across the patient and as a result all the echoes that are reflected will be shown on the display at the position they were produced from. This mode produces a two-dimensional (2D) image of the area enclosed by the transducer where the brightness of the dots represents the amplitude of the echoes. In this produced image, the width of the image is equal to the width of the ultrasound transducer. This image provides much more detail of anatomical structure for the ultrasonographer, compared to the A- and M-modes [9].

In this thesis B-mode images will be used to track the motion of the probe since it provides real-time 2D cross-sectional image of an anatomical structures. Using real-time B mode scanners, many frames can be produced each second and as a result a real time image can be seen.

2.2. Ultrasound Probe Tracking Techniques

There are various ways to track the motion of the ultrasound probe. It allows for the construction of an extended field-of-view frame if the motion between the consecutive frames, which resulted from the motion of the probe, is tracked accurately. It also enables the tracking to be used to reconstruct the ultrasound images as 3D shapes which is usually favored because it provides more details about the structure of the tissues being imaged. Tracking can be done by many types of sensors such as acoustic devices, electromagnetic sensors, and optical sensors. Also tracking can be done through image processing by estimating the motion between the consecutive ultrasound images.

2.2.1. Tracking by External Devices

Tracking the motion of the probe using different sensors has been seen in the literature. Acoustic devices have been used to track the motion of the ultrasound probe by Moritz et al. [10] and by King et al. [11]. Electromagnetic sensors have been used to track the motion of the probe by Berg et al. [12] from the raw ultrasound data. Also optical sensors have been used to track the motion of the probe by Tamaki et al. [13] and by Zhang et al. [14]. Using any of these sensors to track the motion of the ultrasound probe requires the use of a calibration technique so that the movement between the acquired ultrasound images and between the tracking sensor attached to the probe is known [15]. For each of these sensors used in the tracking of the ultrasound probe, there are advantages and disadvantages:

The first method of tracking, optical devices, can obtain the motion of the probe with very high accuracy. Also optical devices can capture data at high frequency. On the other hand, optical devices are expensive and using them with ultrasound would make ultrasound an undesired imaging tool since one of the advantages of ultrasound imaging is it being low cost. Also, optical systems require the use of a line-of-sight between the camera and the

dynamic reference body. As a result, this solution for tracking will not be considered for the purpose of this thesis.

The other method of tracking, electromagnetic sensors, does not require the use of a line-of-sight between the camera and the dynamic reference body giving it an advantage over optical devices. However, electromagnetic sensors can obtain the motion of the probe with reduced accuracy compared to optical systems. Also, since this method of tracking involves the use of an electromagnet attached to the probe, the results of tracking would be affected in the presence of any metallic objects in the magnetic field of the used sensor [16]. Therefore, to maintain the accuracy of tracking using electromagnetic sensors, one should make sure that there are no metallic objects within the proximity of the electromagnetic field. Due to this limitation, this method will not be considered in this research as the goal of this thesis is to develop an extended-field-of-view that visualises a large scale area but at the same time try to obtain one that does not put any limitations on the use of such system.

2.2.2. Tracking by Acquired Ultrasound images

Tracking the motion of the probe can be accomplished by tracking the transformations between the acquired ultrasound images. This method of tracking is known in the literature as Speckle Tracking. Speckle noise in ultrasound is inherent in all the images and as a result tracking the speckle could be useful for tracking the tissues that lie underneath it [17]. Figure 2.7 shows an example of a B-mode ultrasound image of speckle. This method of tracking aims on tracking the motion between the 2D ultrasound images accurately without the need of using an external system such as optical and electromagnetic sensors which is preferred. This method is preferred because there is no cost of the use of external sensors and because it does not have any limitations such as the limitation electromagnetic sensors have which is not wanting any metallic objects in the electromagnetic field of the used sensor.



Figure 2.7: An Ultrasound Image Showing Only Speckle Noises.

Speckle noise in ultrasound is caused by the coherent accumulation of random scatters within a resolution cell when scanning a specific region. In other words, it is caused when the transmitted ultrasound signal reflects back from targets that are smaller than the wavelength of the ultrasound which determines the resolution of the system [17]. The ultrasound machine cannot visualize particles that are smaller than the resolution of the scanner and as a result the sum of the reflected waves from these particles would be presented in the ultrasound images as speckle granules (speckle noise) as can be seen in Figure 2.7 [17]. Burckhardt [17], has tested the types of speckle and he found that the speckle caused by the reflections of a large number of particles has a Rayleigh distribution in the intensity of the images. He called this type of speckle a Fully Developed Speckle.

It is true that speckle noise in ultrasound images degrade the quality of the images and makes it harder for the observer to observe different tissues in the images but on the other hand speckle noise in US images can be useful for motion tracking purposes. The information necessary to estimate the position of an US image with respect to the previous and subsequent frames lie within this speckle noise. Therefore, a method of tracking the motion of speckle between the ultrasound images is necessary to be developed.

To determine the motion of speckle between the ultrasound images, image registration procedures can be used. Tracking the speckle between the ultrasound images for the purpose of obtaining an extended field-of-view frame has been seen in literatures [18, 19]. Weng et al. [18], used a block-matching approach to register the ultrasound images, i.e. the transformation parameters between each two consecutive images were estimated. The purpose for their work was to obtain an extended field-of-view. Tirumalai et al. [19], described a method used by Siemens, provided as SieScape, to obtain an extended field-of-view ultrasound frame. SieScap uses a block-matching approach to track the motion between the acquired ultrasound frames. Block-matching makes use of the speckle in the images by tracking its motion between frames. After the motion is estimated, an extended field-of-view of the region scanned by the probe was obtained. Bohs et al. [20], tracked the motion of speckle between the ultrasound images using an interpolation method called grid slope. This method involves the use of block-matching motion estimation. Geimann et al. [21], also used a grid slope method similar to the one used by Bohs et al. [20] to estimate the subsample speckle motion. This method also involves the use of block-matching motion estimation.

Prager et al. [22], estimated the motion between ultrasound frames. They used a modified block-matching search to do the motion estimation. The modified block-matching developed by them took the block size as the full size of the image. The main goal for doing their work is to obtain a 3D freehand ultrasound. Housden et al. [23], built on the work that was developed by Prager et al. [22]. They used the same modified block-matching method but they modified the way of the used interpolation method to obtain the motion with subpixel accuracy on the X and Y directions. The subpixel method developed by Housden et al. [23] involves the use of a Gaussian curve to fit the five correlation values located at the area of the maximum value. Their method was compared with the method developed by Geimann et al. [21] and they found that their method is better in tracking the motion between the images in the presence of elevational motion.

2.3. Image Registration in the Medical Field

Image registration is an important feature that has a large field of research in the recent years. It is the process of aligning two or more images that have the same scene but taken at different

times, different angles, or different sensors [16, 17,18,19]. It is the main tool used to differentiate between two or more images of the same scene [17]. Image registration plays a key role in the medical environment, especially in ultrasound imaging. It can provide the physicians by an extended field-of-view and as a result it enhances the ability of diagnosing different diseases [17]. If the images corresponding to different tissues were looked at separately by the physician, he or she should mentally try to connect these images together to see the overall anatomy of the regions imaged. This process by itself could lead to errors which would result from miss reading the multiple images separately. As a result, if these images can be registered with each other, an extended field-of-view frame can be obtained and can be used by physicians to see the overall anatomy of the scanned area.

2.3.1. Ultrasound Image Registration

There are two types of registration in the medical environment which are inter-modality and intra-modality. Inter-modality registration is a type of registration that is able to register images from different modalities such as registration of an ultrasound image to a CT image. On the other hand, Intra-modality registration is the process of registering images from the same modality such as registering an ultrasound image with another ultrasound image. Registration of the same modality, which in this thesis is registering ultrasound images together, is important in many aspects such as averaging the registered images to obtain an extended field-of-view; it is also useful in the analysis of temporal effects. In the literature, there are few investigations that have been done to test the feasibility of ultrasound image registration. The reasons behind this could be due to the fact that ultrasound images have poor signal to noise ratio compared to other medical imaging modalities, or it could be due to the fact that the organs usually imaged with US are non-rigid. Also the 2D nature of ultrasound could be another reason as now 3D nature image modalities are favored [24].

Rohling et al. [25], registered ultrasound volumes together and then averaged the register volumes. For their registration, they used a block-matching method. This method checks the correlation of the gradient magnitude of the acquired data. They also used a multi-resolution method to improve the 6-dimensional search space so that the best registration is obtained. Schers et al. [26] developed a method for registering ultrasound volumes. In their method, a segmentation of the interface between the bone and soft tissues was obtained. They used

Normalized Cross-Correlation (NCC) to do the registration which is a block-matching method. The method developed by them was found to work in 65% of the tested cases.

Xiao et al. [27] registered 2D ultrasound images of the breast. In their work, they used a block-matching approach to do the registration between the images. They tested their method on designed phantoms and they found that their registration method is essential in averaging registered images due to the good accuracy and fast execution time. Poon et al. [28] extended the field-of-view of ultrasound. They did that through registering 3D ultrasound volumes using a block-matching approach. After testing their algorithm, they found that a block-based registration produced high reconstruction accuracies. Ni et al. [29] registered 3D ultrasound volumes to obtain an extended field-of-view. They used position trackers to perform the registration. They also used two block-matching approaches to correct for misalignments of volumes. Their method was tested and was found to give high accuracy registration. Groves and Rohling [30] registered 2D ultrasound images for the purpose of improving the quality of the obtained image after averaging the registered ultrasound images. They also used a block-matching method to register the images. After testing their algorithm on phantoms, they found that the quality of the obtained image improved compared to the separate images before averaging. Xu et al. [31] registered ultrasound images to obtain the registration parameters with an accurate way. Their registration technique is based on polar transform and block-matching. They used Sum of Absolute Differences (SAD) as the cost function to find the best matched blocks. The use of polar transform was mainly to find any rotation in the images accurately since a rotation in Cartesian domain is represented by a translation on polar domain. Their method showed high accuracy in obtaining the registration parameters. Sandoval and Dillenseger [32] performed registration on 2D ultrasound images and 2D CT images to help in the diagnosis of Atrial Fibrillation treatment. They did the registration using MI method by registering US images of the Cardiac model to CT images of the same cardiac model. Huang et al. [33] developed a new way for rapid image registration of 2D ultrasound images in order to obtain an extended field-of-view. They used normal area based registration technique (block-matching) to register the images. However, to improve their registration technique, they tried to extract the features of the images. After doing that, the blocks were sorted according to the features they have and based on that, the validity of the blocks were tested. Then, only the valid blocks were used for the registration and as a result registration time was reduced significantly.

2.3.2. Image Registration Techniques

Most of image registration methods consist of four stages. The first stage is feature detection. In this stage, different objects in the scene are detected. The second stage is feature matching. In this step, the features detected in one image are matched with its correspondence in the reference image. The third stage is transform model estimation. In this stage, the transformation of the features matched in the current image and the reference image is estimated. The fourth stage is image resampling and transformation. In this stage, the current image is transformed into the reference image by the values estimated in the third stage [34].

Image registration in ultrasound can be divided into two categories which are feature-based registration and intensity-based registration. A feature-based method looks for features in the image such as significant regions (bones and boundaries) and tries to match them with consecutive images. It is known that ultrasound images are corrupted by noise, thus distinctive features cannot be easily identified in these images. As a result, the choice of using a feature-based registration technique is not widely used in ultrasound. On the other hand, intensity-based registration does not look for significant regions in the image but looks at the intensity values of the pixel and try to match their motion between the consecutive frames. This type of registration uses an area-based method for the registration. Area-based methods find the transformations between the consecutive frames. In other words, the main goal of image registration is to be able to identify the relative motion between images so that the motion is estimated accurately and so that the registration of images would be done as accurately possible [31]. These transformations can vary from rigid transformations such as translation, rotation and scaling to linear transformation to affine transformation such as skewed and perspective transforms and to nonlinear wrapping [35]. Area based methods, which are called correlation-like methods or template matching, combine the first two stages of image registration together which are the feature detection stage and the feature matching stage since there is no need to identify the features using area based methods. Using area based methods; the different types of transformation can be obtained by two ways. The first way is to divide an image into equally sized squared blocks which are called macro-blocks and then search each macro-block for its transformation from the current frame to the preceding or subsequent frame. The block that provides the highest correlation is the best matched block and this macro-block will describe the

motion vector which describes the motion between these two macro-blocks [6, 36]. On the other hand, the second way can be done by finding the transformation of the full frame with respect to its preceding or subsequent frame. Using this approach, a full frame will be assigned a motion vector describing its motion with respect to the preceding or subsequent frame. The first choice of first dividing the frame into macro-blocks and then finding the motion for each macro-block has shown success in registering the 2D ultrasound images as was shown in the literature described in the previous section. Some of the area based methods are Mutual Information (MI) methods, optimization methods, correlation-like methods and Fourier methods [34].

MI methods are intensity-based methods. These methods are usually used in multimodal registration. Mutual information methods measure the statistical dependency between two data sets. This type of registration would be useful when registering images from different modalities is required, such as registering an ultrasound image with a CT image. Multimodal registration is an important subject in the medical field. From multimodal registration, diagnosis of images would be easier by registering images from different modalities to see the difference in between them [34]. Therefore, MI methods are not-applicable to be used in this research since we are only dealing with one imaging modality (ultrasound).

Correlation-like methods (Block-matching methods) are widely used in the literature as was discussed in the previous section. There are many correlation-like methods such as Normalized Cross-Correlation (NCC), Sum of Absolute Differences (SAD), Mean Squared Error (MSE), Mean of Absolute Differences (MAD) and many more. NCC is one of the correlation-like methods. NCC measures the similarities between two macro-blocks, one in the current frame and one in the reference frame to find the best matched block which is the one that yields the maximum value for cross-correlation. Cross-correlation works great in scenarios that have a translational motion. In addition, it also works in scenarios that have slight rotation or scaling. SAD is another method of correlation-like methods. This method works by calculating the sum of absolute differences between image intensities between a block in the current frames and a block in the reference frame to look for the best matched block. The macro-block in the reference frame that gives the minimum SAD is the best matched block. MSE is another correlation-like method. It is exactly the same as the SAD but this one finds the square of the intensity difference between the current macro-block and the reference macro-block instead of

finding the absolute difference. MAD is also another correlation-like method. It works by calculating the mean of the absolute difference between a macro-block in the current frame and a macro-block in the reference frame. The macro-block that yields the minimum MAD is the best matched block. Therefore, the position of the macro-block with the lowest MAD value is assigned a motion vector of the magnitude and direction it was found in [34]. MAD is a preferred method in correlation-like methods since the results obtained by MAD are independent from the used block size. In other words, the results will still be valid even if a different block-size other than the one used in a specific research is used. In using MAD to find the correlation between macro-blocks, one should keep in mind that the correlation between ultrasound images is not as good as it is in other imaging modalities or in video images. As a result, a MAD cost function that could be rejected in video systems could be accepted in ultrasound. For example, if a macro-block of size 4 by 4 in a high quality image obtained by a video camera had the following intensities: $\begin{bmatrix} 75 & 90 \\ 120 & 130 \end{bmatrix}$ and then assuming that the person taking the video made a translational motion, it is expected to see the same macro-block with the same intensities in the subsequent frame but of course at a different location. As a result, the MAD cost function between these two macro-blocks will be zero. On the other hand, considering a person is taking ultrasound images along the hip region, by only translating the probe along the hip region, the pixel intensities could change values between the consecutive frames due to the presence of noise in the images and due to the fact that ultrasound images are constructed from the returning echoes and not by high quality optical cameras. For example, if a macro-block in the current frame has the following intensities: $\begin{bmatrix} 75 & 90 \\ 120 & 130 \end{bmatrix}$, some of its intensity values could be changed and its best matched block could be $\begin{bmatrix} 76 & 89 \\ 118 & 129 \end{bmatrix}$. The MAD cost function in this case wouldn't be zero as some of the intensities changed values. As a result, if a threshold is to be used to reject MAD values larger than a specific number, one should keep in mind the fact that the correlation between ultrasound images is not always high.

Fourier methods are intensity based methods as well. Fourier methods are faster than correlation-like methods. Fourier methods require transferring the images into their Fourier representation. One of the most important Fourier methods is phase correlation. Phase correlation works perfectly in scenarios that have a translational motion. Phase correlation finds

the location of the peak that represents the location of the translation on the X and Y directions. It does that through finding the inverse Fourier transform of the cross power spectrum of the current and reference images. Phase correlation shows strength against frequency dependent noise and non-uniform time varying illumination disturbances [34]. As a result, the use of phase correlation would be applicable in this research due to its high performance and speed. The feasibility of using a correlation-like method (MAD block-matching) along with a Fourier based method (phase correlation) is studied and implemented in this research to improve the accuracy of the registration between ultrasound images.

2.3.2.1. Block Based Motion Estimation

As discussed previously, block-based motion estimation algorithms have been used widely in ultrasound. The use of block based motion estimation algorithm is useful in ultrasound since block based methods are area-based methods. Using area-based methods in ultrasound images is useful in estimating the ultrasound probe's motion. This is because speckle intensities in the images along with tissue intensities in the images are both used in tracking the motion between the consecutive frames. In the literature, Strintzis and Kokkinidis [37] performed motion estimation on ultrasound image sequences using a block based approach that is called Maximum Likelihood (ML). Ultrasound images as mentioned previously have a lot of speckle noise, this speckle noise can be described by multiplicative Rayleigh distributed noise or by signal-dependent additive noise [37]. ML estimation can be used for tracking the motion through ultrasound B-mode images through the maximization of the conditional probability density [37]. Strintzis and Kokkinidis [37] tested the ML algorithm on ultrasound B-mode images of the heart and the results obtained by them shows that ML has a high signal-to-noise ratio. Their method was not considered in this thesis because it was only tested on small number of sequences of images on the heart. Golemati et al. [38] used B-mode ultrasound images to track the carotid artery wall motion using region tracking and block matching. It is worth mentioning that in block-matching approaches, all the pixels in a selected block are assumed to have the same velocity vector. Golemati et al. [38] used the Normalized Correlation Coefficient described by Bohs and Trahey 1991 and Huang 1981 as their block-matching algorithm. The technique here is based on speckle tracking in the B-mode ultrasound images and thus speckle tracking between each two consecutive frames was done. Golemati et al. [38] was able to successfully track the

motion of carotid artery but with putting limitations in spatial and temporal resolution [29]. Gastouniotti et al. [39] used a block based matching algorithm for arterial wall motion estimation from B-mode ultrasound images. They used a Kalman-Filter-Based block matching to track the motion between the reference block and the block in motion [39]. Kalman filter is an efficient recursive filter that is able to estimate the current state of a linear dynamic system from sequences of noisy quantities [39]. In their work, conventional block based matching algorithms were joined with Kalman filtering using two cases. In the first case, an updated reference block strategy based on the Kalman filter was presented and in the second case, the estimate of the conventional algorithm was renewed using Kalman filtering. Gastouniotti et al. [39] used a reference block size of 25x17 pixels and a search area of 10x10 pixels. It was observed that tracking using this algorithm reduced the mean tracking error and the Kalman filter applied improves the performance of the motion estimation techniques in low or very low bit rate applications [39]. Their method was mainly developed to estimate the motion of the arterial wall without the need to move the probe. As a result, their method was not considered as it does not have the same goal needed in this thesis (extending the field-of-view resulted from moving the probe). Gao et al. [40] developed a method to track the blood flow from the left ventricle using speckle tracking with Navier-Stokes equations. It is known that blood flows from the left ventricle at high speeds. This fact about the motion of the blood with respect to the typical frame rate of an ultrasound imaging system makes it harder to track its motion using block based matching algorithms [40]. Gao et al. [40] suggested a way to perform the motion tracking of the blood from the left ventricle. Speckle tracking was first performed. A 2D kernel of size 0.34mm x 0.028 rad was used and a search area of 15.3mm X 0.078 rad was used and then the normalized cross-correlation was used to find the best matched block. By doing speckle tracking by itself, estimate of blood motion from the left ventricle would not be correct due to the high speed of blood flow. As a result, Gao et al. [40] suggested a regularization method to regularize the initial motion estimates obtained by speckle tracking with the physical characteristics of the expected velocity field which is described by Navier-Stokes equation. The results obtained by Gao et al. [40] algorithm shows that the accuracy for motion tracking of blood flowing at low velocities from the left ventricle improved when speckle tracking was combined with Navier-Stokes regularization. However, the method failed in tracking the motion of speckle using their algorithm at high velocities since the initial estimates were too noisy [40]. Basarab et al. [41]

proposed a new algorithm that is able to estimate the sub-sample 2D translations of ultrasound signals based on a previous model of images. A Least Square Estimator (LSE) was used to solve for the 2D translation estimation problem. In their method, the proposed estimator is applied directly on signals to find the relative motion which is found by calculating the difference between two signals. According to Basarab et al. [41], the proposed estimator does the job of calculating the motion accurately with no need to interpolate for sub pixel accuracy. The estimator developed is then integrated in a block-matching method to estimate the motion from ultrasound RF images. The results obtained by Basarab et al. [41] were compared to those obtained using SAD cost function on original images and on images integrated by a factor of five. The results obtained using the estimator developed were more accurate than those found using SAD on the original images. On the other hand, the results obtained by the estimator were roughly as accurate as those obtained using SAD on the interpolated images by a factor of five but the computational time for calculating SAD on the interpolated images by a factor of five took ten times longer than just using the estimator developed [41]. However, this method was not considered for the purpose of this thesis as the results obtained were tested on the original RF signals of ultrasound and not on the images obtained after processing the RF signals (B-mode images).

As mentioned earlier, there are a lot of block based algorithms that have been used in the literature. The most important ones are MAD, MSE, SAD and NCC. These all are called cost functions. The value obtained by the cost function called block distortion measure (BDM) which is a value in block-matching that is used to rate the similarities between two blocks [36]. Usually in block matching a search area is defined around the coordinates of the source block to find the best matched block. The source block would be a block at a frame A and the search area is located in another frame, frame B, that comes before or after frame A [6, 36, 42]. The search consists of the area of p pixels on all four sides of the macro-block in the frame being searched. This p value is called the search parameter. When there is large motion between frames, a larger p value should be chosen and as a result the computational time will increase. Usually in block-matching motion estimation, a block size of 16 by 16 pixels is chosen and a search parameter of 7 is chosen [6]. The idea of block-matching can be represented in Figure 2.8.

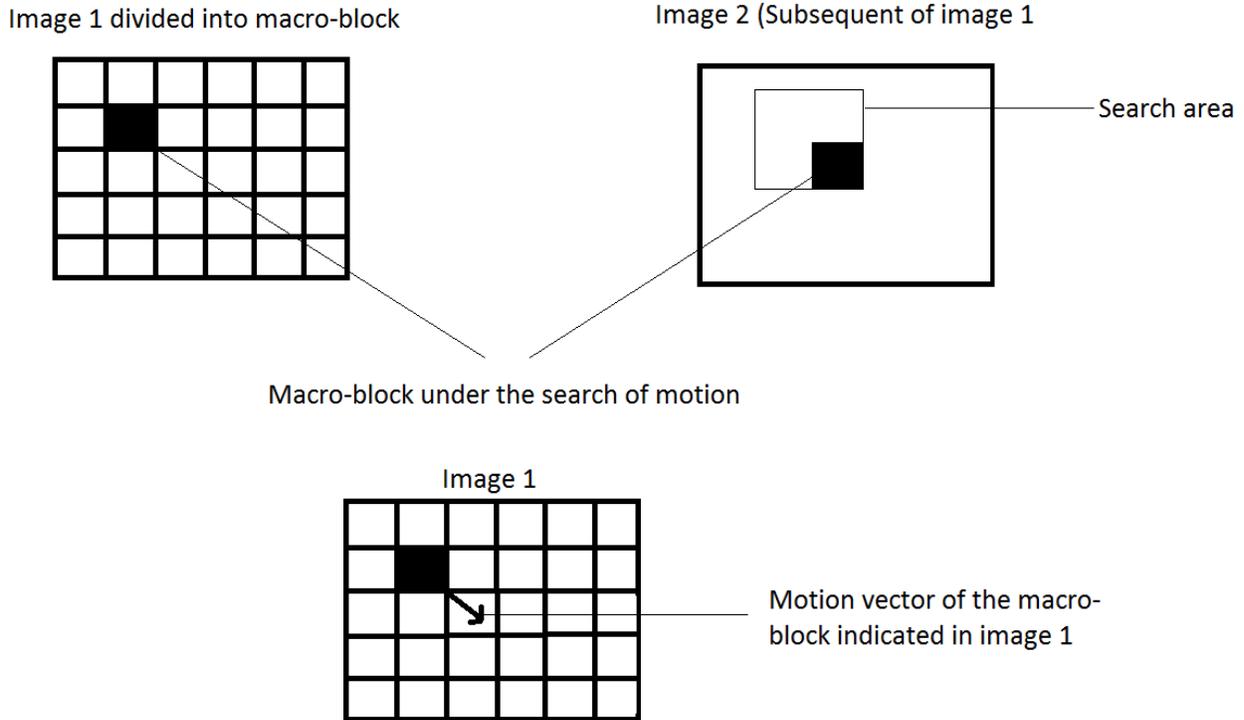


Figure 2.8: Motion Estimation Using Block-Matching.

In the literature, there are many algorithms developed in searching for the best matched block in the search area. Some of these algorithms are Exhaustive Search (ES), Three Step Search (TSS), New Three Step Search (NTSS), Simple and Efficient Search (SES), Four Step Search (4SS), Diamond Search (DS), Two Dimensional Logarithmic Search and many more. These algorithms differ in the way they perform the search for the best matched block in the search area. Each one of them has its own method of performing the search for the best matched block in the search area. Each one of them except for exhaustive search tries to reduce the points to search for the best matched block in the search area to speed up computational time.

Three Step Search (TSS) has many good characteristics such as being simple, with a significantly low computational time and its good performance [6,42]. Lu and Liou [42] proposed a new and efficient algorithm that develops on the TSS algorithm and named it Simple and Efficient Search (SES). SES builds on the TSS algorithm and tried to reduce the computational complexity even more than the TSS does while having a good performance that is comparable to TSS [42]. Zhu and Ma [43] developed another fast algorithm for block-matching

called Diamond Search (DS). The results obtained by Zhu and Ma [43] show that DS performs better than the TSS algorithm with even reducing the computational complexity even further. Nei and Ma [44] developed another fast block matching algorithm that is called Adaptive Rood Pattern Search (ARPS). ARPS uses the motion vector of the macro-block to the left of the current macro-block to predict the motion of the current macro-block [6]. It checks for the best matched block at the place the predicted motion vector is pointed to and it checks at the points that are at a rood pattern distribution where the direction of the arrow is the direction of the predicted motion vector. After reviewing the above algorithms, one stood out. All of the algorithms, except Exhaustive Search, try to reduce the points in the search area used to check for the best matched block in the defined search area. They do so in order to reduce the computational time to look for the best matched block with trying to get results that are as close as possible to the Exhaustive search. As was observed by Barjatya [6], none of the other algorithms (TSS, NTSS, SES, 4SS, DS, ARPS) gave results that are as accurate as the Exhaustive Search. Ideally, for the case of ultrasound images, we would like to reduce the computational time but on the other hand we would not want the accuracy to be affected. As a result the Exhaustive Search method was chosen to be used to look for the best matched block in the defined search area. The search for the best matched block using exhaustive search is going to take a little longer than the other fast algorithms but the results obtained are worth this time since they are the most accurate.

Exhaustive Search is known as the full search algorithm [6]. Exhaustive Search is the simplest method in motion estimation and works by calculating the MAD cost function at each possible location in the search area defined to find the best matched block. This gives a total of $(2p+1) \times (2p+1)$ positions that needs to be checked to find the best matched block, where p as mentioned earlier is the search parameter defined [6,45]. For example, if a search parameter is set to be 7 pixels on all directions of the macro block, 225 positions would need to be checked for each macro-block in the frame to find the best matched block. This makes ES to be the most computationally expensive block matching algorithm. On the other hand, since each possible location is being searched for the best matched block, ES provides the most accurate motion vectors among other block-matching algorithms with the highest PSNR [6, 45]. Due to its high accuracy, it was used in this thesis work to do the job of motion tracking between the ultrasound frames.

After looking at many papers that used different block-matching techniques, it was concluded that a block-matching method that uses a MAD or SAD cost function to find the highest correlation between macro-blocks, would work best with ultrasound images. MAD and SAD are the same but the difference between them is that MAD is normalized by the area of the block size used. In the literature, it was noticed that MAD and SAD were used frequently in ultrasound to find the location of the best matched block. In this research, MAD was chosen to be used since as mentioned earlier, the results obtained using MAD are independent from the block size chosen. After using a block-matching search, an integer value for motion estimation can be obtained. When the motion between ultrasound images is estimated with an integer value, registration of the images would not be accurate. As a result, either interpolation methods, which are not computationally efficient, or some other method has to be used to estimate the motion between macro-blocks with high accuracy. In this thesis it is thought that Taylor series approximation can be used to find the sub-pixel shift after the pixel shift has been determined by block-matching.

2.3.2.2. Taylor Series Approximation in Motion Estimation

Barron et al. [46] discussed several techniques used for optical flow which lie in one of these categories: differential techniques, region-based matching, energy-based methods, and phase-based techniques. Differential techniques which are known as gradient based methods follow the assumption that the brightness of a pattern is invariant over time and as a result they calculate the velocity of images from spatio-temporal intensity derivatives [47]. These derivations come from this assumption and from the first-order Taylor series approximation [48]. Each equation in the image flow constraint equations has two different variables and as a result these equations by themselves would not be enough to find the optical flow [48,49]. Horn and Schunck [50] introduced smoothness of the flow as a second constraint to help in solving the flow constraint equations. Gradient-based approach suffers from a big problem that is the image flow constraint equation is unreliable in areas of the image where the local brightness function is highly nonlinear [48]. Examples of these areas are highly textured regions, areas with motion discontinuities and areas with depth discontinuities [48]. As mentioned earlier, first-order Taylor series approximation is used in the derivation of the image flow constraint equation. The use of the first-order Taylor series approximation leads to inaccuracies when the higher order terms are

highly significant [48]. These higher order terms are usually ignored in finding the image flow constraint equation due to the assumption that the motion time between successive frames is usually small. Most of the time, the motion time between successive frames is consistent and not very small. This fact about the motion time between successive frames makes the image flow constraint equation only reliable when the higher order derivatives of the brightness function are insignificant [48].

Taylor series approximation was chosen to refine the search for the best matched block, along with the Exhaustive Search algorithm to obtain an estimate of the motion between the ultrasound frames with sub-pixel accuracy. Together these algorithms were used to determine the sub-pixel motion estimation of the macro-block under testing. The normal way of obtaining sub-pixel motion estimation in block-matching is interpolation methods. However, interpolation techniques are not computationally efficient. As a result, Taylor series approximation was used to refine the search for the best matched block with sub-pixel accuracy. It was chosen because according to Chan et al. [5], Taylor approximation is able in estimating the supixel shift as long as it is less than one. Chan et al. [5], used a block matching algorithm which used the MAD cost function to estimate the integer shifts between video sequences and then they refined the search for the subpixel shift using a Taylor series approximation. It was found by Chan et al. [5] that Taylor series approximation can be used to estimate the motion between images with sub-pixel accuracy and they found that this takes less time than interpolation methods for sub-pixel motion estimation with giving the same accuracy as interpolation method does. The reason that allows us to use Taylor approximation to estimate the subpixel shift between the ultrasound frames is that the sub-pixel shift is less than one. As mentioned earlier, accuracy is important in this thesis. Thus, the most accurate method of motion estimation is desired. Since, as long as the motion is less than one, Taylor series approximation can estimate the motion with high accuracy, it was chosen to estimate the sub-pixel motion estimation between the ultrasound frames.

2.3.2.3. Phase Correlation Motion Estimation

Phase correlation motion estimation is also an area-based method. Thus, it is an intensity-based registration method. Phase correlation has been used with ultrasound but not as widely as block-matching algorithms. Phase correlation was first reported by Kuglin and Hines [51]. Gilliam et al. [52] proposed a 3D ultrasound volume reconstruction through the use of phase correlation.

They tracked the motion of the probe using B-mode ultrasound image sequences. They tracked the motion between the 2D ultrasound images in order to be able to reconstruct a 3D volume from these 2D images [52]. Gilliam et al. [52] used subpixel phase correlation algorithm which is the one developed by Hoge [53] to estimate the motion between the 2D ultrasound images with subpixel accuracy. The results obtained by Gilliam et al. [52] shows that the developed algorithm which uses the subpixel phase correlation motion estimation is very accurate and provides superior results [52]. Liang [7] compared a block-matching method with phase correlation method in a final project to see which one provides more accurate motion vectors. However, the images were not ultrasound images, they were normal video images obtained by a camera. For both methods, Liang [7] divided each current frame into 16x16 pixel macro-blocks and then for each macro-block, different block-matching algorithms which were listed earlier in this thesis were used and phase correlation was used as well. The results obtained were compared to see which method performs better and computationally faster. The results obtained by Liang [7] showed that phase correlation provided a smoother vector field with lower entropy compared with the four-step log search block matching. Over all, phase correlation was about 20% faster if half-pixel accuracy was not considered. It was observed that block matching performed better in situations that had smaller translational motion where phase correlation performed better in situations where a large scale translational motion arises. Lastly, it was observed that motion compensation was more accurate with higher SNR using phase correlation if the images contained mainly large translational motion with no rotation [7]. Foroosh et al. [54] suggested a way for subpixel accuracy in motion estimation using phase correlation. The main assumption in their work was that images that have subpixel shifts were originally displaced by an integer shift, but due to down sampling, subpixel shifts arises. From down sampling the images, it was observed that the cross power spectrum is made of all polyphase components of a filtered unit impulse and as a result, the output that would be obtained from the inverse Fourier transform of the normalized cross power spectrum would be a down sampled version of a filtered unit impulse [43]. It was observed by Foroosh et al. [54] that the down sampling resulted in more than one coherent peak. Whereas in the case of integer shift, only a single coherent peak is obtained. It was also observed that these coherent peaks in the down sampling case relates to the polyphase transform of a filtered unit impulse which centred at the point of motion. The results obtained by Foroosh et al. [54] showed perfect results in subpixel

motion estimation for the purpose of registration using phase correlation [54]. Hoge [53] made an extension of phase correlation to estimate the motion. Phase correlation is known to work well in any imaging modalities where a translational motion is seen. However, Hoge [53] applied his modified algorithm for phase correlation on MRI images because of the fact that MRI data is typically samples in the spatial domain and that makes it possible to apply phase correlation directly on the raw data of MRI before the reconstruction of MRI images is obtained. Hoge [53] suggested to separate the shifts obtained using the normalized cross-correlation using singular value decomposition (SVD) factorization. After this has been performed, the magnitude of the translation on each direction, X and Y, is obtained using a linear fit to the phase of each component. This developed method has many advantages, such as obtaining subpixel motion without the need to perform interpolation on the data in the spatial domain. As mentioned earlier, this developed algorithm separates the displacement estimation on the vertical and horizontal direction and thus phase unwrapping is reduced to one-dimensional (1D) domain instead of 2D domain. One more advantage for this method is that it is complementary to other subpixel image registration techniques [53]. Hoge [53] tested the developed algorithm on real MRI data and the results obtained were promising. The results obtained showed that the algorithm performs accurately in estimating the motion at subpixel level even with the presence of noise as long as only a translational motion is present [53]. Kumar et al. [55] made some modifications in phase correlation. The steps to compute motion estimation using phase correlation are obtaining the Fourier transform of the images, computing the normalized cross power spectrum, and obtaining the inverse Fourier transform of the normalized cross power spectrum [55]. The most computationally expensive operation between these steps is computing the normalized cross power spectrum. Kumar et al. [55] suggested away to overcome this problem by computing the normalized cross power spectrum from the approximate normalization of complex scalar which is known to be very efficient and hardware friendly. The main goal that Kumar et al. [55] tried to accomplish is to make phase correlation technique hardware friendly and as a result be able to use it in real time. The results obtained by Kumar et al. [55] shows that the proposed algorithm has potential to use phase correlation with low power devices and use it in applications that deals with many images. The use of phase correlation can be also extended to be used in the four-dimensional (4D) domain. Zhang et al. [56] used 3D phase correlation to estimate the respiration-induced tumor motion in 4D. This method is an extension to the 2D motion

estimation using phase correlation to reconstruct a 3D volume. They obtained two CT volumes from different respiration stages and then they modeled the tumor from the shift observed between the two CT volumes. Guizar-Sicairos et al. [57] introduces a way of obtaining subpixel motion estimation. In this method the initial location of the phase correlation peak is first found after upsampling the images by an upsampling factor of two. This was done to try to get an appropriate starting point for the phase correlation that might have more than one high peak with similar values. A matrix multiplication implementation was used to find the location of phase correlation peak with subpixel accuracy. The usefulness of using matrix multiplication is that upsampling would not have to be done on all the image points which were time and power consuming. Upsampling was done around the value of the initial estimate of the phase correlation peak only. This approach is believed to work well in estimating the subpixel shift between ultrasound frames.

Phase correlation is used to track the motion between two frames or two fields directly from their phases. It is unlike block-matching approaches which search for the best matched block from luminance field [7]. This method for motion tracking depends on the property of Fourier transform which states that a shift in the spatial domain relates to a linear phase shift in the Fourier domain [7,52,53,55]. Phase correlation motion estimation is one of the best motion estimation algorithms due to its strong performance, high accuracy and efficient computation [7,53]. Phase correlation finds the motion between frames and represents the motion by a peak located at the positions of shift. For example, if a frame is shifted by 10 pixels on the x axis and 5 pixels on the y axis, the location of the phase correlation peak would be at location (10,5). The algorithm of phase correlation is relatively simple compared to other motion estimation algorithms and gives the motion estimation of high accuracy, especially in cases where only a translational motion is present. The phase correlation algorithm is able to find the integer translation on the X and Y directions. If subpixel motion estimation is required, interpolation methods or some other methods will be needed to be used to estimate the motion with subpixel accuracy. In literature, subpixel accuracy is usually obtained through some interpolation methods such as bilinear, Lagrange and many others [53,54].

After looking at different papers that used phase correlation to estimate the motion between images, especially in cases where only a translational motion is present, phase correlation was

chosen to estimate the motion between the ultrasound frames. However, it was chosen to be used on the compensated frames obtained from the block-matching algorithm that was developed during this thesis. The use of phase correlation is applicable in this thesis since the motion between the ultrasound frames is a translational motion. In this thesis, we were interested in estimating the motion between frames when the probe is moved in the lateral direction. However, phase correlation without interpolation gives an integer estimate of the motion between the ultrasound frames. A method described above by Guizar-Sicairos et al. [57] was used to estimate the subpixel shift between ultrasound images. Applying this approach on the ultrasound images allows the estimation of the motion with subpixel accuracy and at the same time, reduces the computational time that other interpolation methods take.

Chapter 3

Methodology

3.1. Overview

The aim of this thesis work is to track the motion of the ultrasound probe using image processing techniques. This was done in order to extend the field-of-view of the 2D ultrasound probe. 2D B-mode ultrasound images were obtained on the hip region, upper leg region and on an ultrasound phantom that has been designed to validate the results. As mentioned earlier, 2D B-mode ultrasound images are very useful since they show the anatomy of the tissues scanned with the probe. The overview of the work done is shown in Figure 3.1. In all the experiments performed, the ultrasound probe moved in the lateral direction parallel to the linear scan direction. 2D ultrasound images were obtained using a linear probe and motion tracking algorithms have been developed and implemented to do this task. First of all, to test how well phase correlation motion estimation works with ultrasound images, several experiments were conducted. After validating that phase correlation works well for the purpose of this thesis, a

block based motion tracking algorithm was implemented and then tested. The block based motion tracking algorithm was tested in different scenarios. In the hip region, it was tested when a translational motion was obtained. It was also tested for translational motion on the upper leg region in the presence of muscle contraction and probe compression. This block based motion-tracking algorithm works as follows:

- First of all, each 2D ultrasound image frame was divided into a number of macro-blocks of equal size.
- Each macro-block was searched in a predefined area to find the best matched block.
- Applied the integer based phase correlation to choose an appropriate size for the predefined search area.
- The size of the predefined search area is chosen according to the maximum motion obtained by the integer based phase correlation algorithm plus one. If the maximum motion obtained by phase correlation was 10 pixels, the search area for the block-matching search will be set to 11 pixels on all the directions of the macro-block.
- Exhaustive search, which does a search at each possible location that the best matched block could be at, was used.

Before starting the search for the best matched block, blocks with no motion and blocks that are hard to track were excluded from the search to improve the accuracy of motion tracking. This process was done for each frame with respect to the frame that comes before it. First, areas with no motion and areas that are hard to track were excluded from the compensated ultrasound images. After that, motion tracking was done on the compensated ultrasound images using a subpixel phase correlation motion tracking algorithm. Doing so improved the global motion vector. Finally, the ultrasound images obtained were registered with respect to one frame using the motion tracking algorithms that have been implemented. An average of the registered images was taken to obtain an extended field-of-view image which offers lower noise compared to the original frames and shows a bigger area of the anatomy.

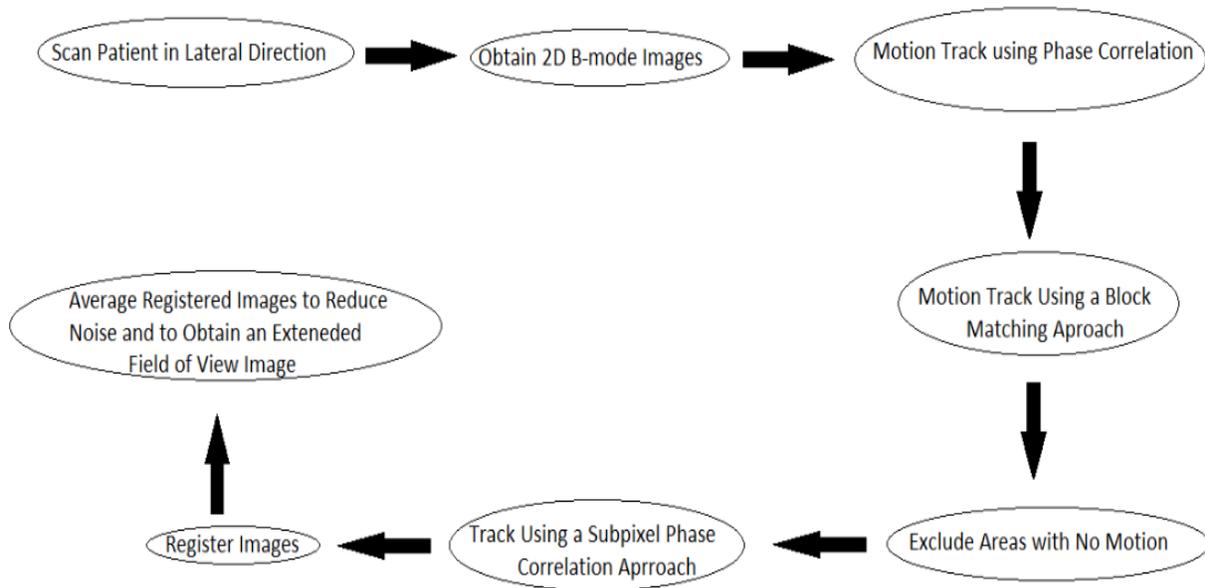


Figure 3.1: Overview of the Work Done in this Thesis.

3.2. Exclusion of No Motion Areas and Areas that are Hard to Track

In ultrasound, there are many areas that change shape due to many reasons such as effects from probe compression and muscle contraction. Also, in ultrasound, there are blank areas that do not represent any tissues or speckle. Therefore a new method has to be developed that can improve the motion tracking method for ultrasound images. Using this method, motion vectors were not assigned to areas in the ultrasound images with no motion nor were they assigned to areas that were hard to track. This was done to improve the accuracy of the tracking algorithm. The ultrasound images obtained went through a block-matching search. Each two consecutive frames were taken and the search was performed on them to determine the relative motion between them. First of all, the current frame was divided into equal sized macro-blocks of size $N \times M$ as can be seen in Figure 3.2 where $N \times M$ were chosen to be 16×16 pixels. It is important that, only blocks that have information that can be tracked, are considered in the search to improve the motion estimation accuracy.

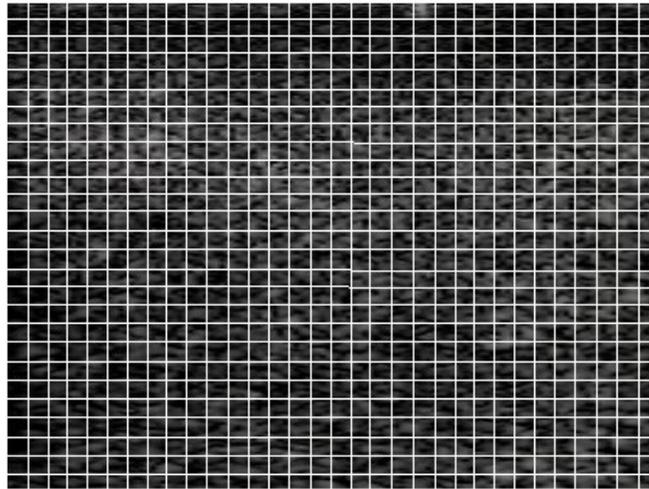


Figure 3.2: An Image Divided into Macro-Blocks to Perform the Block Search on it.

Macro-blocks that contain no motion are found by calculating the standard deviation of the macro-block. If the standard deviation is zero, it was concluded that this macro-block contained no motion and as a result, it was not considered in the tracking since it had a zero motion vector. If the standard deviation of the macro-block turned out to be zero, a zero motion vector was assigned to it without performing the search for its best matched macro-block. A threshold value for the minimum MAD function was set to identify blocks that have uncorrelated motion due to decorrelation of speckle or due to reading errors. As mentioned earlier, a cost function will be used to find the best matched block which in this thesis is the MAD cost function. In ideal cases the MAD function should be equal to zero if the current block and the best matched block had the exact intensities. Ideally this happens if there was only a translational motion between the frames but no change in its intensity values. Since ultrasound images are by their nature susceptible to corruption by noise, some of the values of the intensities could change between blocks in the current frame and blocks in the reference frame. As mentioned earlier in this thesis, speckle noise in ultrasound images correlates between the images and thus can be used to track the tissues underlying it. One of the reasons why some of the blocks could change values between consecutive frames is decorrelation of speckle between frames. In ultrasound images, the tissues that are scanned will not always have fully developed speckle. It is true that speckle noise in ultrasound correlates between frames but this is not always the case. In some cases,

speckle decorrelation appears between the frames and thus tracking these blocks that have speckle decorrelation would be difficult. Another reason why pixel intensities could change between current and reference blocks is the fact that, ultrasound images depend on the reflection of the echoes to form the images as the probe is in motion. The reflection of echoes is not always the same even if it was from the same tissues. It all depends on the position the probe is sending the pulses from and how the ultrasonographer is moving the probe. These reasons make it harder to track these blocks from ultrasound images and thus obtaining a MAD value of zero is not always possible in ultrasound images. But to overcome this problem, a threshold value for the MAD function was used. If the value of the minimum MAD function, between a block in the current frame and a block in the reference frame, is greater than a specific value defined by the user, this block will not be considered in the tracking since tracking it wouldn't yield an accurate motion vector. In this thesis, these blocks were called areas that are hard to track. Tracking using block-matching and using the improvements put into it is very useful and easy. Using a block search, different features of the images need not to be identified; the motion can be tracked from just looking at the intensity values of the blocks. This type of search is an area based search as mentioned earlier.

3.3. Estimation of the Motion Using Block Matching

A block-matching method described by Barjatya [6] and Chan et al. [5] was employed with minor modifications which were described in section 3.2. A block in a current frame was searched for its best matched block in a frame that came before it. The search area to look for the best matched block is set by the user. This search area is different from an experiment to another depending on how fast the probe was moved. In order to know the search area needed to find the best matched block, an integer phase correlation motion estimation is performed (which will be described in section 3.4.1) on the images to know how much the maximum motion between any two consecutive frames was. The maximum motion found is used as the search area after adding one extra pixel to it. For example, if the maximum motion between any two consecutive frames found using the integer phase correlation algorithm was 8 pixels, the search area to search for the best matched block using the block-matching method developed here would be 9 pixels on all directions. An exhaustive search is performed to find the best matched block. Using exhaustive search, also called full search, the search for the best matched block is

performed at each possible location in the defined search area. This search method can be visualized in Figure 3.3 which shows a macro-block of size 16 x 16 being searched on all directions in the search area to find the best matched block. The exhaustive search starts from the top left of the image and then move by a step of the search area defined. Each macro-block will be searched for its best matched block in all the pixels in the search area to the left, right, top and bottom of the macro-block. A cost function will be used to find the cost at every possible location in the search area. The block with the minimum cost function in the search area will be the location that the macro-block moved to and thus the distance between them represents the motion vector for that macro-block. The exhaustive search searches for the best matched block in $(2 \times \text{search area} + 1)$ blocks in the horizontal direction and in $(2 \times \text{search area} + 1)$ in the vertical direction.

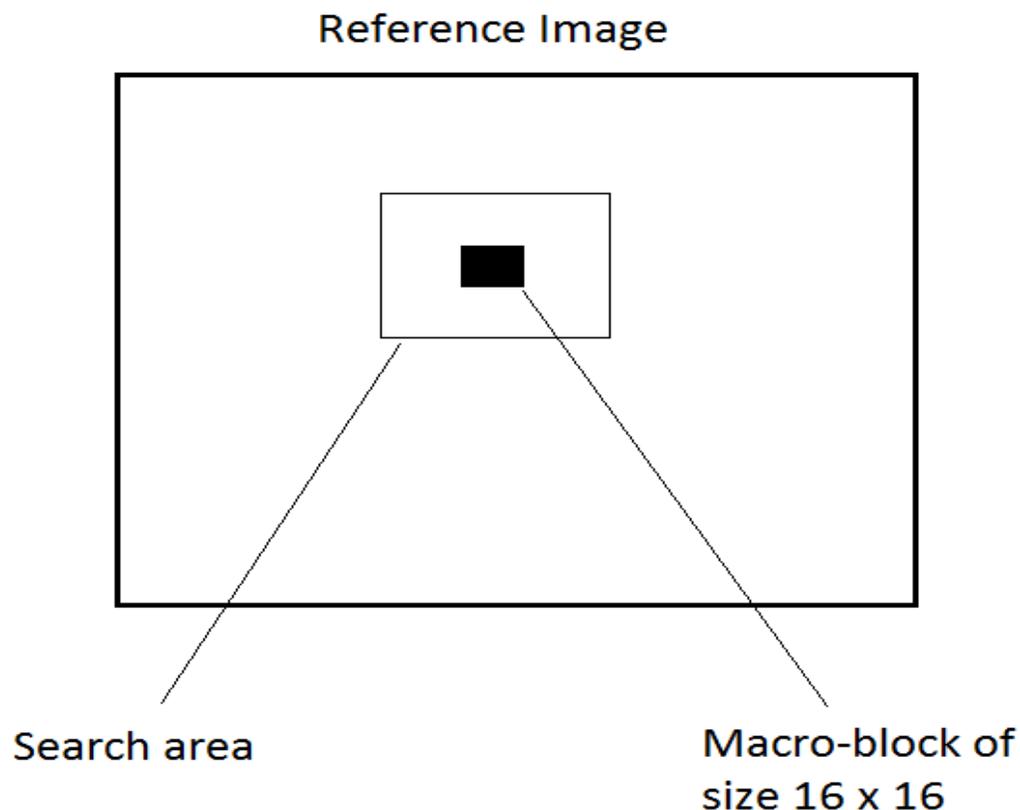


Figure 3.3: A Macro-block of Size 16 X 16 Being Searched at Each Possible Location in Predefined Area.

Mean of absolute difference (MAD) is the cost function used to find the best matched block between two image blocks. This cost function measures the difference between two variables A1 and A2 to see the statistical difference between them. Given two macro-block A1 and A2, the MAD cost function between them can be found as follows:

$$MAD = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |A1_{ij} - A2_{ij}| \quad (3.1)$$

Where N is the length of the macro-block, M is the width of the macro-block. Therefore, N X M is the area of the macro-block. A1_{ij} and A2_{ij} are the pixels being compared between the current macro-block and the reference macro-block. In all the experiments in this thesis, the sizes of N and M are 16 x 16 pixels. The best matched block is the block that yields the minimum cost function MAD since this macro-block is the one with the minimum statistical difference.

According to the position where that best matched block is found, the motion vector for that macro-block is defined. This approach gives the integer shift for that macro-block. Interpolation or some other methods should be used to obtain subpixel accuracy for the motion vector. In this thesis, a method described by Chan [5] was implemented to obtain subpixel accuracy without interpolation and thus following this approach will decrease computational time since interpolation methods are not computationally efficient. The next section describes the method used to obtain subpixel accuracy for the motion vector.

3.3.1. Taylor Approximation for Sub-pixel Motion Estimation Accuracy

So far, accurate motion estimation was obtained by a block search algorithm that has minor changes put into it to improve the accuracy. These minor changes were put to exclude the no motion areas and areas that are hard to track. An integer motion vector was obtained using the modified block-matching algorithm described in section 3.3. Taylor series approximation was used in order to find the subpixel accuracy to register medical ultrasound images accurately and efficiently. It was chosen to be used due its simplicity and efficiency in obtaining the subpixel accuracy compared to interpolation methods. Taylor approximation works in situations where the motion has a value less than one [5]. Thus, using it to estimate the subpixel shift between the ultrasound frames would be applicable. Taylor approximation is an optical flow method used to

find the relative motion [5]. First, an explanation on how motion can be estimated using Taylor series approximation will be given.

Given two consecutive frames $f(x,y)$ and $h(x,y)$ that have motion between them which is defined by equation 3.2:

$$h(x,y) = f(x + \Delta x, y + \Delta y) \quad (3.2)$$

Where Δx is the shift on the x axis and Δy is the shift on the y axis. From equation 3.2, the first order Taylor approximation can be calculated as follows (equation 3.3):

$$h(x,y) = f(x,y) + \Delta x \frac{\partial f(x,y)}{\partial x} + \Delta y \frac{\partial f(x,y)}{\partial y} \quad (3.3)$$

Where in this equation $\frac{\partial f(x,y)}{\partial x}$ is the partial derivative of the 2D image $f(x,y)$ with respect to x axis and $\frac{\partial f(x,y)}{\partial y}$ is the partial derivative of the 2D image $f(x,y)$ with respect to the y axis.

Now to solve for the optimal shift, solving for the minimization of equation 3.4 described below is needed.

$$\Phi(\Delta x, \Delta y) = \sum_{x,y} (h(x,y) - f(x,y) - \Delta x \frac{\partial f(x,y)}{\partial x} - \Delta y \frac{\partial f(x,y)}{\partial y})^2 \quad (3.4)$$

Equation 3.4 is a linear least square problem and thus the shifts on the x and y axis, Δx and Δy respectively are found by setting $\frac{\partial \Phi}{\partial x}$ and $\frac{\partial \Phi}{\partial y}$ to zero and as a result, the solution would be found by the following system of linear equations:

$$\begin{bmatrix} \sum_{x,y} \left(\frac{\partial f(x,y)}{\partial x}\right)^2 & \sum_{x,y} \left(\frac{\partial f(x,y)}{\partial x} \frac{\partial f(x,y)}{\partial y}\right) \\ \sum_{x,y} \left(\frac{\partial f(x,y)}{\partial x} \frac{\partial f(x,y)}{\partial y}\right) & \sum_{x,y} \left(\frac{\partial f(x,y)}{\partial y}\right)^2 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \quad (3.5)$$

$$\begin{bmatrix} \sum_{x,y} (h(x,y) - f(x,y)) \frac{\partial f(x,y)}{\partial x} \\ \sum_{x,y} (h(x,y) - f(x,y)) \frac{\partial f(x,y)}{\partial y} \end{bmatrix}$$

By solving equation 3.5, the solution for motion on the x and y axis will be found. As mentioned earlier in section 3.3, a block-matching approach is used to find the integer shift between macro blocks. Exhaustive search was used as the search method to find the best matched block. Exhaustive search was used because it is the most accurate since it performs the search for the best matched block at every possible location in the search area. Let's assume that the integer shift found by exhaustive search between two blocks, one in the current frame and it is best matched block in the reference frame, is denoted by Δx_1 and Δy_1 where Δx_1 and Δy_1 are the integer motion on the x and y axis respectively. After that has been done, to find the subpixel shift between the blocks, the search is refined using Taylor approximation approach described above. Let's assume that the non-integer shifts are denoted by δx and δy where δx is the subpixel shift on the x axis and δy is the subpixel shift on the y axis. The overall motion on the x and y axis can be shown in equations 3.6 and 3.7 respectively.

$$\Delta x = \Delta x_1 + \delta x \quad (3.6)$$

$$\Delta y = \Delta y_1 + \delta y \quad (3.7)$$

Where Δx and Δy are the total displacement with subpixel accuracy on the x and y axis respectively. Figure 3.4 shows a sample output after performing the entire described algorithm so far.

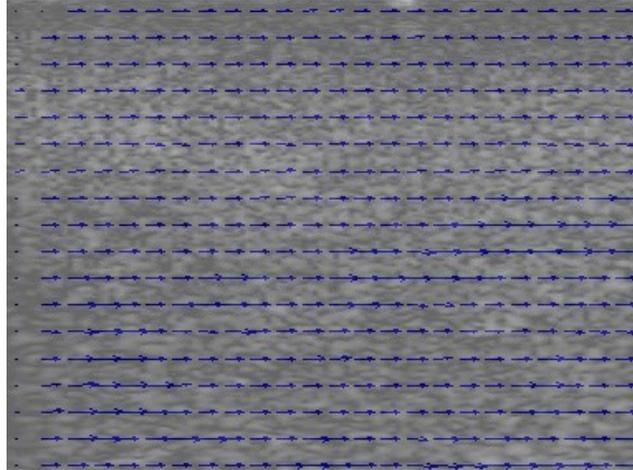


Figure 3.4: Motion Vectors Overlaid over a Speckle Noise Image Generated by Ultrasound Imaging.

Applying the methods described in section 3.3 would be sufficient to perform image registration to obtain an extended field-of-view. However, we thought that if we combine this method with another method which is phase correlation, registration would be improved further. Block-matching methods search the macro-blocks from Luminance matches [6]. On the other hand, phase correlation performs the search directly from the phases of the images in their Fourier domain.

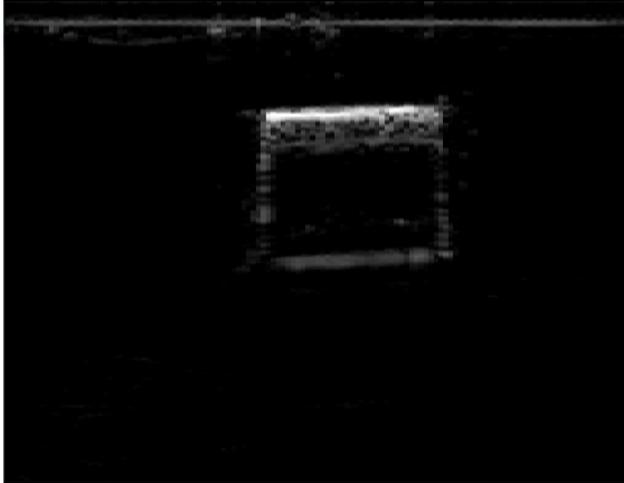
3.4. Phase Correlation

The proposed method in section 3.3 along with phase correlation will improve the accuracy for motion tracking using 2D ultrasound images. Phase correlation as mentioned earlier is a Fourier method and it is known that Fourier methods are robust against noise. As a result, using Phase correlation will benefit us in this research since ultrasound images are noisy. Phase correlation measures the motion between two images directly from their phases and therefore yields more accurate results. Combining the modified block-matching search with phase correlation will improve the motion tracking accuracy even more and as a result the global motion vector will be smoother. This will be done by first finding the motion using the modified block-matching algorithm and then we apply phase correlation on the areas of the images that were found to have

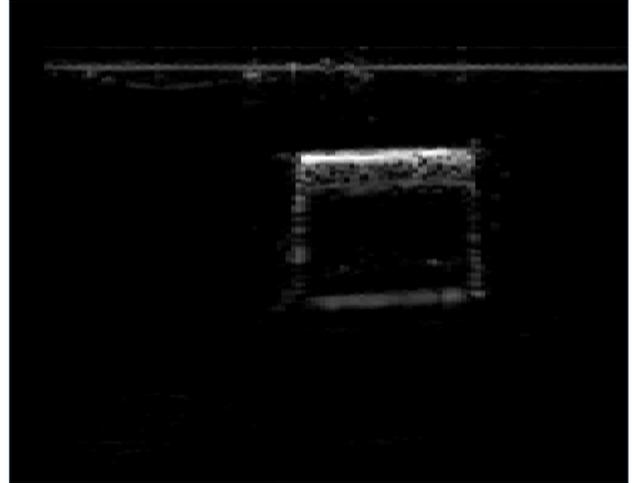
a motion in the direction of the probe's motion, other areas were excluded from passing them to phase correlation.

3.4.1. Integer Based Motion Estimation by Phase Correlation

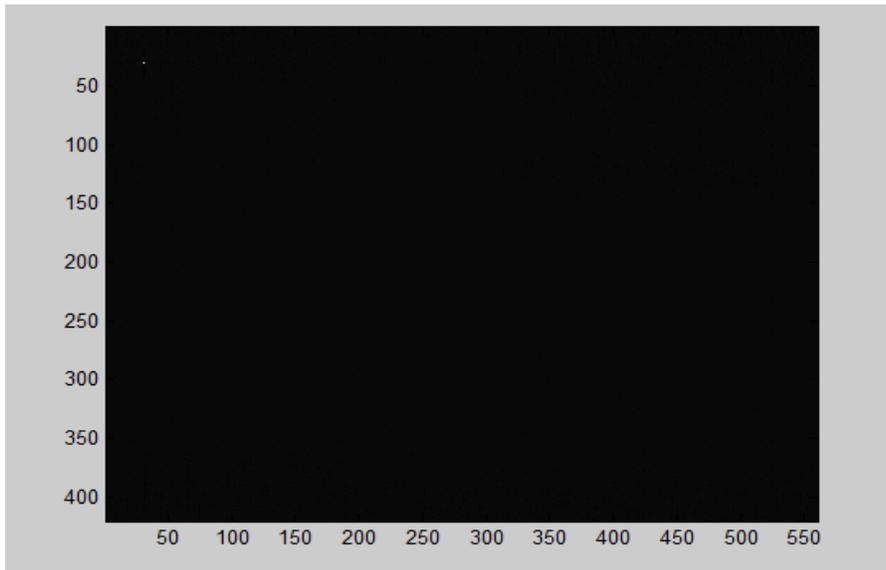
The first step that was done in this thesis was to estimate the motion between the frames using an integer based phase correlation in order to know the maximum motion between any two consecutive frames. This value obtained plus one pixel was used as the search area to look for the best matched block using the modified algorithm for block-matching. Phase correlation is an area based method as well. As a result, identifying the structure of the tissues in the image was not required. Phase correlation is a Fourier method which makes it faster than other motion tracking algorithms [19]. This method works very well and very accurately in cases that have translational motion. The motion between two frames is represented by a peak that shows the exact motion between the two frames (Figure 3.5c). Figure 3.4a shows an eraser in a designed ultrasound phantom. This eraser was translated by 30 pixels on the X and Y axis respectively and the output result is shown in Figure 3.4b. The white dot is the phase correlation peak which in this case is at location (30,30) as can be seen in Figure 3.4c. Since in this thesis we are moving the probe in the lateral direction, the motion between images would be a translational motion as long as the tissues underlying the scan area do not move. So the use of phase correlation method in this thesis would be beneficial.



(a)



(b)



(c)

Figure 3.5: (a) Eraser in Designed Ultrasound Phantom. (b) The Eraser in (a) was Shifted by 30 Pixels on the X and Y axis Respectively. (c) The Peak Output from Phase Correlation which Represents the Motion Between (a) and (b).

The algorithm of phase correlation was relatively simple to implement. If we have two 2D images $f(x,y)$ and $g(x,y)$ where $g(x,y)$ differ from $f(x,y)$ by motion on the x and y axis as follows:

$$g(x,y) = f(x - a, y - b) \quad (3.8)$$

Where a is the displacement on the X axis and b is the displacement on the Y axis. We first obtain the Fourier transform for both images which will be represented by $F(u,v)$ and $G(u,v)$ for the first and second images respectively. Where $G(u,v)$ is

$$G(u,v) = F(u,v) \exp(-j(ua + vb)) \quad (3.9)$$

Where u and v represent the displacements on the X and Y axis respectively in the Fourier domain.

After the Fourier transform for both images have been obtained, the next step in phase correlation would be to calculate the normalized cross-correlation by multiplying the Fourier transform of one image with the conjugate of the second image and then normalize the result by the magnitude of the product as can be seen in equation 3.10.

$$Q(u,v) = \frac{F(u,v)G(u,v)^*}{|F(u,v)G(u,v)^*|} = \exp(-j(ua + vb)) \quad (3.10)$$

Where $F(u,v)$ is the Fourier transform of the first image and $G(u,v)^*$ is the Fourier conjugate of the second image

Finally, to obtain the location of the phase correlation peak (the motion between the two frames), the inverse Fourier transform is applied on $Q(u,v)$ as can be seen in equation 3.11.

$$Q(x,y) = \delta(x - a, y - b) \quad (3.11)$$

Where the output is a Dirac delta function centered at the displacements on the X and Y axis respectively.

Applying the procedure described in section 3.4.1 is able to find the integer shift between the images. However, we are looking to find the motion with subpixel accuracy, thus a method has to be implemented to find the subpixel shift using phase correlation. Usually interpolation methods are used to do this task. Section 3.4.2 describes the method used in this thesis to obtain the subpixel shift using phase correlation.

3.4.2. Subpixel Based Motion Estimation by Phase Correlation

The algorithm described in section 3.4.1 is able to estimate the integer motion between images using phase correlation. This by itself would be sufficient to identify the search area in the modified block-matching search as described before. In order to improve the motion tracking accuracy in general, knowledge of the subpixel displacement is needed. Dealing with ultrasound images, which are usually used for diagnostic purposes of different diseases, requires very accurate motion estimation. A method developed by Guizar-Sicairos et al. [57] was implemented to work for subpixel motion estimation using phase correlation. In this method the initial location of the phase correlation peak is first found using the algorithm described in section 3.4.1 after upsampling the images by an upsampling factor of two. This was done to try to get an appropriate starting point for the phase correlations that might have more than one high peak or peaks with similar values. A matrix multiplication implementation was used to find the location of phase correlation peak with subpixel accuracy. The usefulness of using matrix multiplication is that upsampling would not have to be done on all the image points which were time and power consuming. Only upsampling around the value of the initial estimate of the phase correlation peak was done. Equation 3.12 below is the cross-correlation equation that is used for upsampling the area around its highest peak.

$$Q(u,v) = F(u,v)G(u,v)^* = \exp(-j(\frac{ua}{M} + \frac{vb}{N})) \quad (3.12)$$

Where $F(u,v)$ is the Fourier of the first image, $G(u,v)^*$ is the complex conjugate of the Fourier of the second image, a and b are the displacements on the X and Y axis respectively, M and N are the dimensions of the images. So now to obtain subpixel accuracy for the phase correlation peak, the cross-correlation equation described in equation 3.12 is upsampled in an area of 1.5 x 1.5 pixel around the initial estimate of the phase correlation peak. Three matrices with dimensions of $(1.5k,N)$, (N,M) , $(M,1.5k)$ respectively are designed and then multiplied by each

other where k is the upsampling factor. If we are trying to find the motion between frames to within a 0.01 accuracy of a pixel, the upsampling factor k would be set to 100. After all of that was done, the location of the phase correlation peak with subpixel accuracy is found in the output (1.5k,1.5k). Performing this method for subpixel motion estimation will increase the accuracy for locating the phase correlation peak.

After subpixel accuracy for phase correlation was implemented, this method was applied to the ultrasound images, after compensation has been made for no motion areas and areas that were hard to track using the modified block-matching algorithm. Doing so provided a more accurate global motion vector. The next step that was done is registering the images with respect to a reference frame. The reference frame in all the experiments was chosen to be frame number one.

3.5. Ultrasound Image Registration

Image registration is the process of overlaying many images of the same scene taken at different times or from different angles [19]. Image registration is important in many fields. It is used in remote sensing, in medicine to obtain an extended field-of-view and in many other fields [19]. In this thesis, we are dealing with ultrasound images and the methods described before for motion estimation will be used to register the images with respect to one frame and thus an extended field-of-view will be obtained which will be described in section 3.6. From the modified block-matching algorithm and from the subpixel accuracy of phase correlation, an accurate estimate of the motion between frames is obtained. The next step is to register the images (align them) with respect to one frame. In all the experiments performed to validate the results obtained, at least 180 images were obtained for each experiment depending on which program was used to read the images. If ART.LAB is used, the frame rate will be 30fps and the record time is the last 6 seconds of the probes motion. This yields a total of 180 frames for each experiment. On the other hand, a video capture system was used in some of the experiments to record the images captured by the ultrasound machine. This system records for a longer period with a frame rate of 60fps and as a result a larger number of total frames were obtained which depended on how long the video capture was on. In all of the experiments conducted, the first frame was considered as the ideal frame and all of the subsequent frames were registered with respect to that frame. The translational motion is determined using the modified block-matching

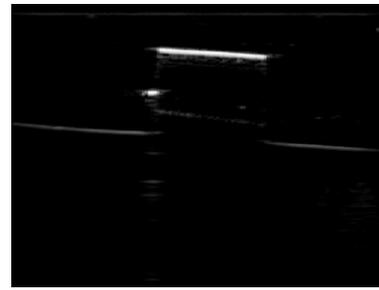
search along with the subpixel motion estimation of phase correlation and according to the X and Y values of the translation motion, the subsequent frames were translated. After doing so, all the images were aligned with respect to one frame and thus all the similar tissues on the hip experiments, the upper leg area experiments or the materials used in the designed ultrasound phantom are aligned above each other.

3.6. Averaging the Registered Images

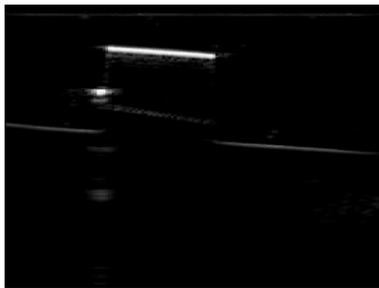
Now that all the images have been registered using the motion estimation methods described in this thesis, averaging them would be beneficial and useful. This was done by pixel wise summation and normalization by the total number of the registered frames. Averaging the registered frames would give an extended field-of-view image which has information from all the frames. This image would be useful in the medical field and if it can be obtained by ultrasound using sensor-less motion tracking algorithms, the benefits of using ultrasound would increase. Ultrasound is favoured to be used over other imaging modalities since it does not send ionised radiation into the body. It is believed that the motion tracking algorithms developed and implemented in this thesis are able to accurately register the images. As a result, averaging these registered images would give us an extended field-of-view image and would help in reducing speckle noise, which ultrasound images are corrupted by, since averaging works like a low pass filter. Figure 3.5 shows some outputs from an experiment on a designed ultrasound phantom and shows the output obtained from averaging the registered images. The original size of each of the ultrasound images obtained was (560 x 420). Some of the outputs obtained are shown in figure 3.5a-d. Figure 3.5e shows the output obtained after registering the total number of frames and then averaging them. It can be seen that the field-of-view was extended and the output frame was of the size 838 x 440. If there was noise in the images, averaging the registered images would have reduced the noise but since the designed ultrasound phantom did not have any noise, speckle noise cannot be seen on those images.



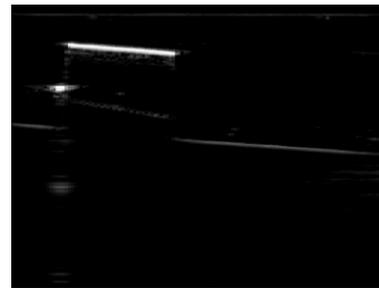
(a)



(c)



(b)



(d)



(e)

Figure 3.6: (a) Eraser in Ultrasound Phantom. (b) Same Eraser at Different Position due to Probe motion. (c) Same Eraser at another Position due to Probe Motion. (d) Same Eraser at another Position due to Probe Motion. (e) Average of the 180 Registered Images.

3.7. Summary

Motion estimation of the ultrasound probe can be obtained by tracking the motion of the 2D B-mode ultrasound images using a sensor-less method. The process of tracking is divided into two parts. The first part is to track the motion using a modified block-matching search algorithm. The second part is to track the compensated frames using phase correlation. After the tracking is done, image registration is performed to register the images with respect to one frame. And finally, averaging the registered frames is done to obtain an extended field-of-view image which has information from many frames while the noise is reduced since averaging works like a low pass filter. The experiments and the results of this tracking method are given in chapter 4 of this thesis.

Chapter 4

Experiments and Results

4.1. Ultrasound Machine

All the experiments in this thesis were carried out using an ultrasound medical imaging machine (Model: Picus, ESAOTE Europe, Maastricht, Netherlands). This machine equips a computer software, ART.LAB, to acquire rf ultrasound signals. Another external video capture system was used to read and store the images in experiments performed on the upper leg area. Figure 4.1 shows an image of the ultrasound machine used. The ultrasound machine has a linear probe which was used in all of the experiments.

4.1.1. Linear Probe

The probe used in this thesis is a linear probe, also shown in Figure 4.1. It contains a 40 mm linear array of transducers. It is made up of four layers. It contains 127 individual transducers that are separated by $315\ \mu\text{m}$ [58]. The first layer has the transducers that are made up of piezoelectric crystals and that occupy $188\ \mu\text{m}$. The second and the third layers are impedance matching layers that are used to enhance the transfer of ultrasound waves into and out of the tissues being imaged. The thicknesses of the layers are $65\ \mu\text{m}$ and $85\ \mu\text{m}$ for the second and third layers respectively. They have a propagation velocity of $2290\ \text{m/s}$ and $2750\ \text{m/s}$ for the second and third layers respectively. The fourth layer is an acoustic lens. This lens is used to focus the ultrasound waves that are sent into the tissues. This layer has a thickness of $1010\ \mu\text{m}$ and a propagation velocity of $975\ \text{m/s}$ [48].

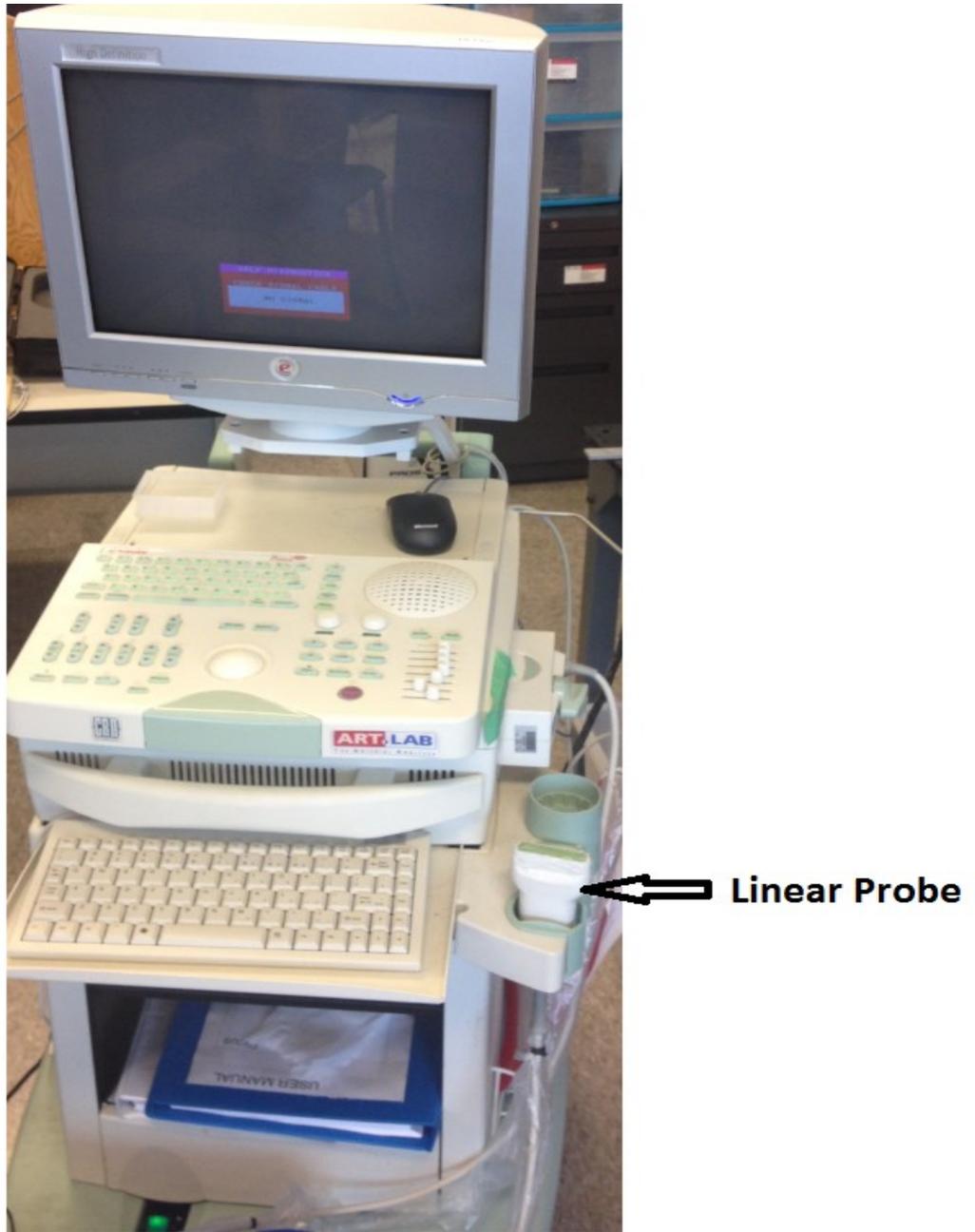


Figure 4.1: Picus Ultrasound Machine with a Linear Probe.

4.2. Phantoms Preparation

An ultrasound phantom was prepared to validate one of the developed algorithms. Since we are not sonographers, designing an ultrasound phantom was necessary in this thesis. We do not have the ability the sonographer has to identify different tissues and the location of them. As a result, in order to overcome this problem, an ultrasound phantom was designed with adding different materials that are known to anyone. The phantom was designed using gelatine products (Knox gelatine) which model the soft tissues of the human body. Knox gelatine is made of collagen in pork's skin and bones. The procedure of designing the phantoms was as follows. First of all, 1 L of water was put to boil. After it was boiled for few minutes, gelatine product was added to it in a container such as the one shown in Figure 4.2. The mixture was then stirred until all the gelatine was completely dissolved. Then, the mixture was put to cool down at room temperature for few minutes and after that it was put in the refrigerator for three hours to completely get solidified. Doing so, a first layer of gelatine was prepared. The next step was to add some known materials such as eraser, glass, tube and so on, on top of the first layer of gelatine after it was completely solidified after being in the refrigerator for three hours. Finally, another layer of gelatine was prepared following the same procedure as for preparing the first layer and was added on top of the added materials. All the mixture was put to cool down at room temperature and then was put in the refrigerator to get solidified. It was kept in the refrigerator until the next day, which is the day of performing the experiment on it. Several ultrasound images were taken on the phantom while the probe was moved in the lateral direction which is the direction of interest in this thesis. After the images were obtained, motion tracking was performed on them to track the motion between the frames. After the motion was tracked, image registration was performed and averaging was performed as well to see how well the motion tracking algorithm works.

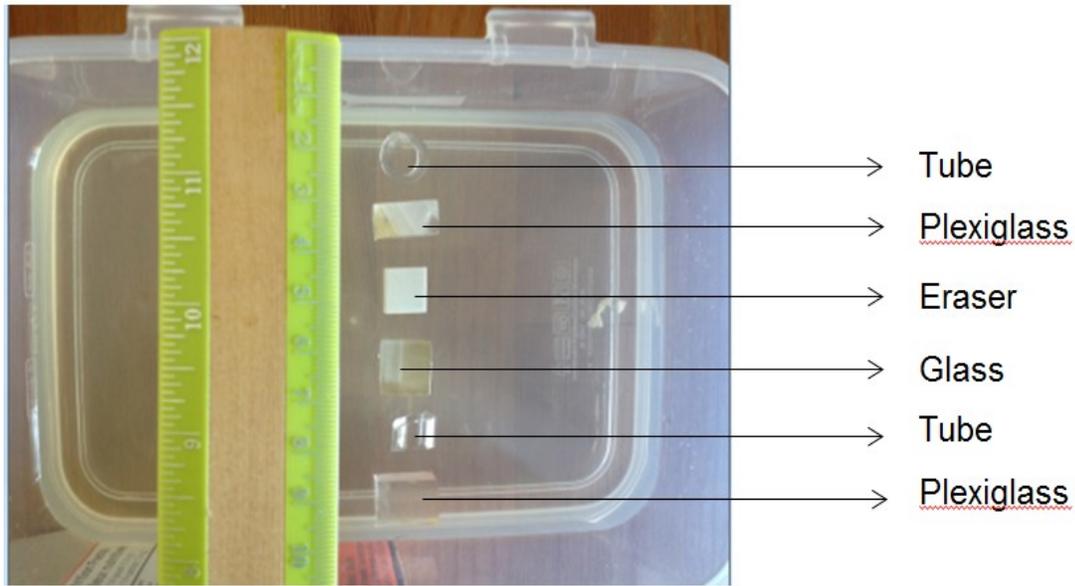
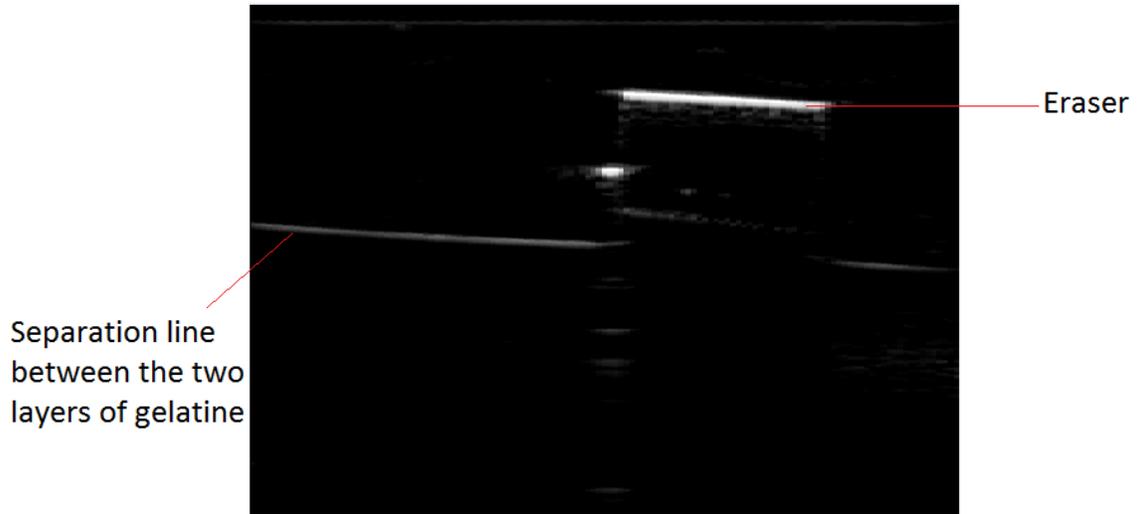


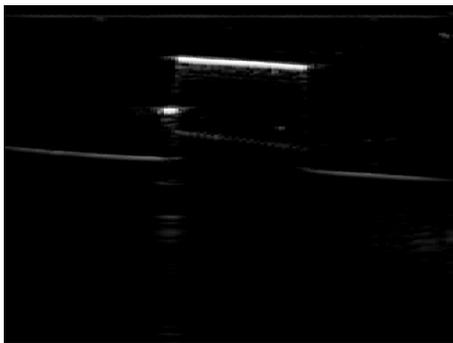
Figure 4.2: First Layer of Gelatine added Some Materials on Top of it.

4.2.1. Phantom Experiment

The designed phantom was used to measure the effectiveness and accuracy of image registration using phase correlation. The probe was moved in the lateral direction above each of the materials in the phantom. The images produced will only have translational motion with no noise as the gelatine is noise free. As mentioned earlier, ART.LAB was used in most of the experiments to record data of the images. ART.LAB was used in the ultrasound phantom experiment. It is able to record the last six seconds of motion of the probe with a frame rate of 30fps. Thus, a total of 180 frames are obtained from each part of the experiment. Each frame is of the size of 560 x 420. After the frames were obtained, tracking the motion between them was performed using phase correlation. The motion between each two consecutive frames was recorded. Some of the output frames on the eraser are shown in Figure 4.3.



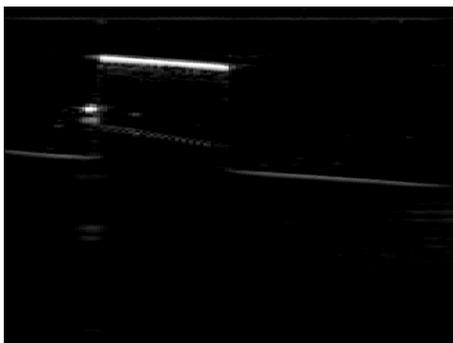
Frame 1



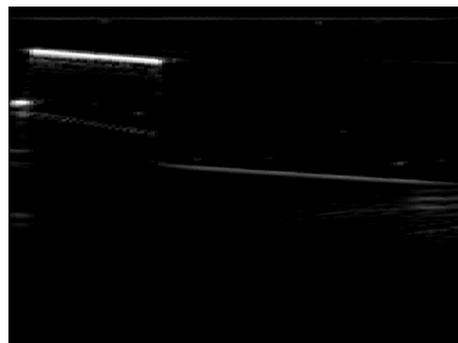
Frame 45



Frame 135



Frame 90



Frame 180

Figure 4.3: Examples of the Frames Obtained on the Eraser in the Phantom.

From these frames, it can be seen that as the probe moved, the location of the eraser changed. Such motion is transitional motion since the probe moved in the lateral direction. Figure 4.4 shows the amount of motion between each two consecutive frames using the integer based phase correlation algorithm described in chapter 3. This motion was obtained from moving the probe along the Y direction which is the direction parallel to the linear scan.

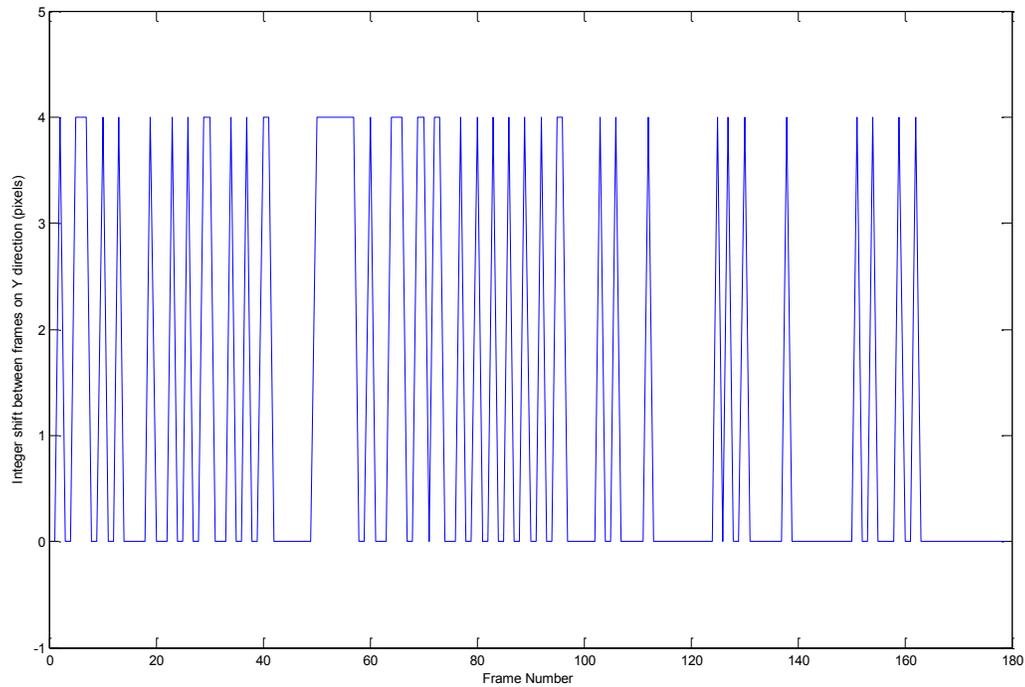


Figure 4.4: Motion Estimation between Eraser Frames on the Y Direction using Integer based Phase Correlation.

In Figure 4.4, it can be seen that the motion was not consistent as the probe did not move at a constant speed. The motion of the probe in the eraser experiment was on and off within the six seconds of motion recording using ART.LAB software. It can be seen in this Figure that the maximum motion was about 4 pixels between successive frames and the minimum motion was about zero pixels between successive frames. Zero pixel difference, meant no motion existed between these frames.

After the integer pixel displacement was found using phase correlation, the subpixel accuracy phase correlation was applied on the same frames to obtain the motion with subpixel accuracy. Figure 4.5 shows the motion between the frames with subpixel accuracy.

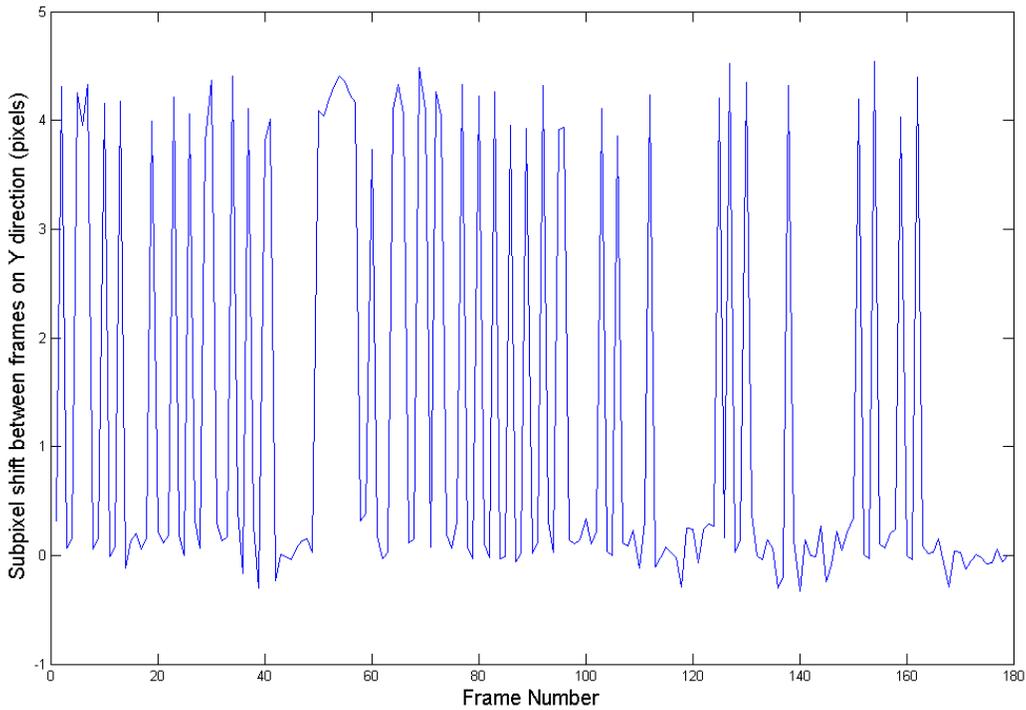


Figure 4.5: Motion Estimation between Eraser Frames on the Y Direction using Subpixel Based Phase Correlation.

It can be seen in Figure 4.5 that the motion between frames was obtained with subpixel accuracy. It also can be seen that the maximum motion between the eraser frames was about 4.54 pixels and the minimum motion was about -0.33 pixels. The negative amount of shift indicates that the motion of the probe moved backward (opposite to the direction of motion). The reason behind this could be due to subject error such as hand shaking while the probe was being moved or the probe stopping within the six seconds of the recorded motion. The accuracy of the motion estimation obtained by the subpixel phase correlation would be judged after image registration, and averaging of the 180 frames is performed.

Using phase correlation, the motion along the X direction, which is the change in depth, can be estimated. It is true that the motion applied on the ultrasound phantom was only along the Y direction but, the pressure the probe was applying on the phantom was not consistent and could have made some changes to the depth in the X direction. Figure 4.6 shows the integer shift on the depth and Figure 4.7 shows the subpixel shift on the depth.

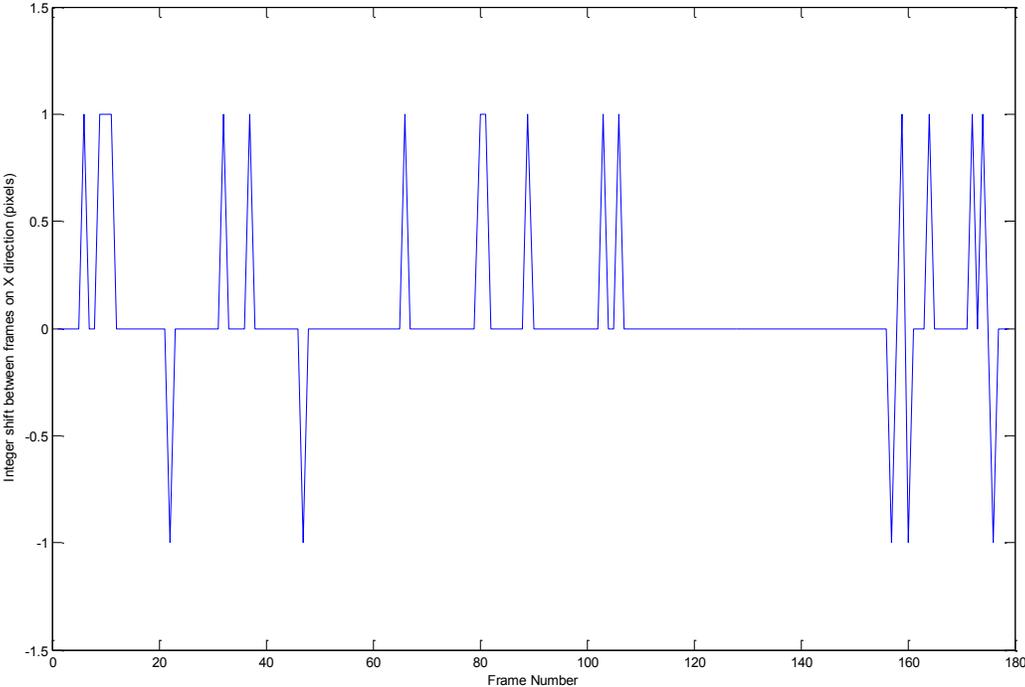


Figure 4.6: Motion Estimation between Eraser Frames on the X Direction (depth) using Integer based Phase Correlation.

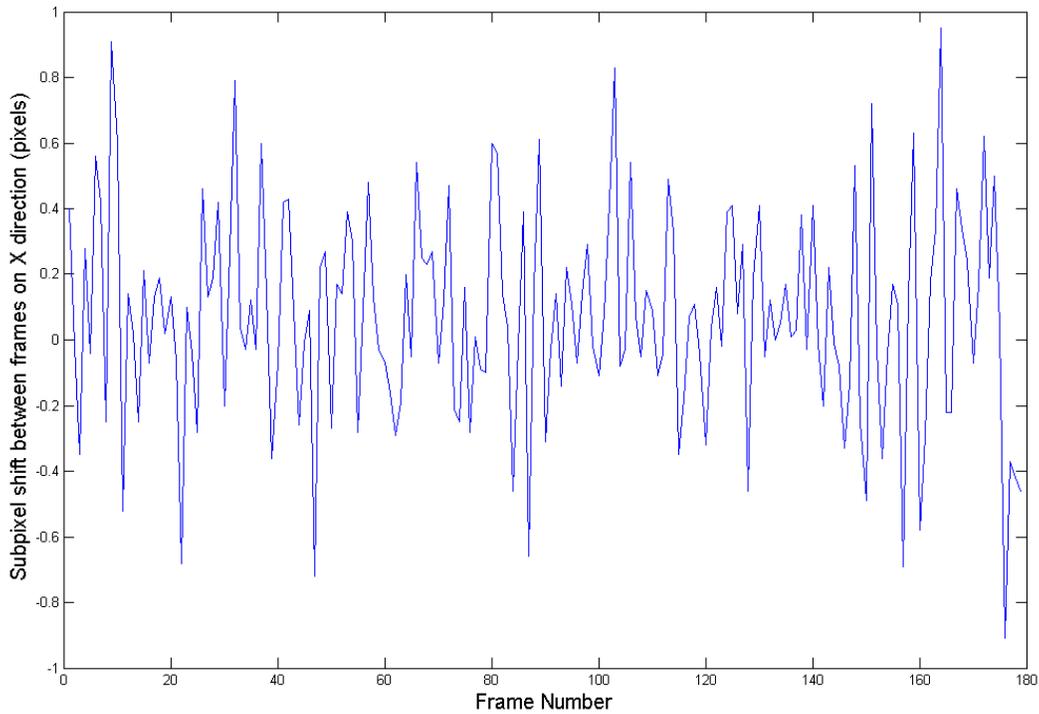


Figure 4.7: Motion Estimation between Eraser Frames on the X Direction using Subpixel based Phase Correlation.

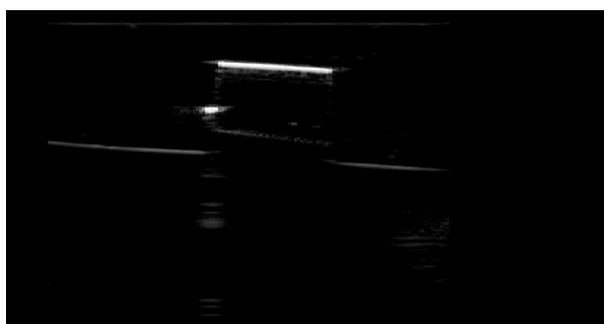
Figure 4.6 shows that the maximum change in depth was about 1 pixel between consecutive frames and the minimum change was about -1 pixel between consecutive frames. Using the subpixel phase correlation algorithm, it can be seen in Figure 4.7 that the maximum motion between consecutive frames was about 0.95 pixels and the minimum motion between consecutive frames was about -0.91 pixels. Again, the accuracy for this motion obtained will be judged after image registration and averaging is performed on the 180 frames.

The next step was to register the images; all the images of the eraser were registered with respect to frame number One by shifting each frame by the amount of motion that was found using phase correlation along the X and Y directions. Due to doing that, the size of the images increased every time a new image is mapped to frame One. After the registration was performed on the 180 frames, they were averaged over the total number of frames. The resulting frame size

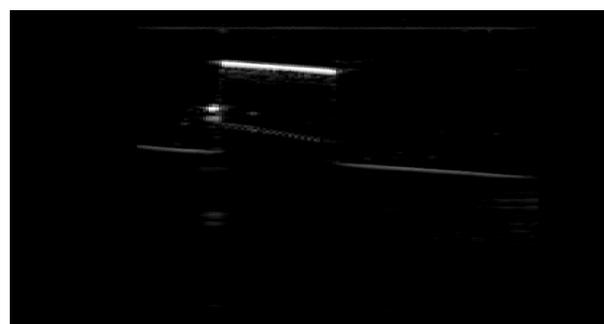
was 838 X 440. Figure 4.8 shows some of the registered frames but their sizes were reduced here for visualization purposes.



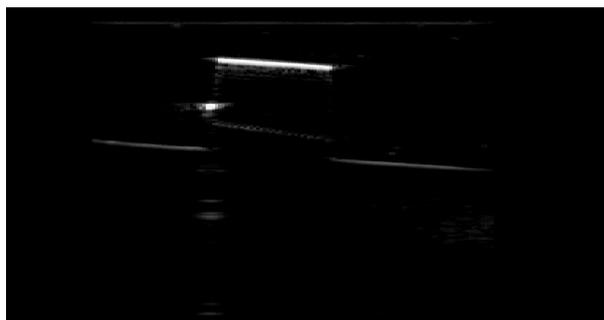
Frame 1 (560 x 420)



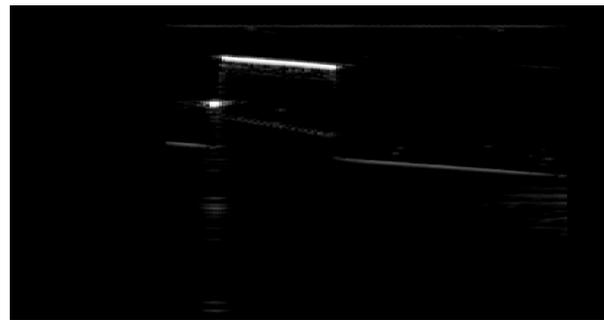
Registered frame 30 with respect to frame 1
(838 x 440)



Registered frame 90 with respect to frame 1
(838 x 440)



Registered frame 60 with respect to frame 1
(838 x 440)



Registered frame 120 with respect to frame
1 (838 x 440)

Figure 4.8: Sample Registered Images of the Eraser in the Ultrasound Phantom.

The result of averaging the registered 180 frames of the eraser is shown in Figure 4.9 below.

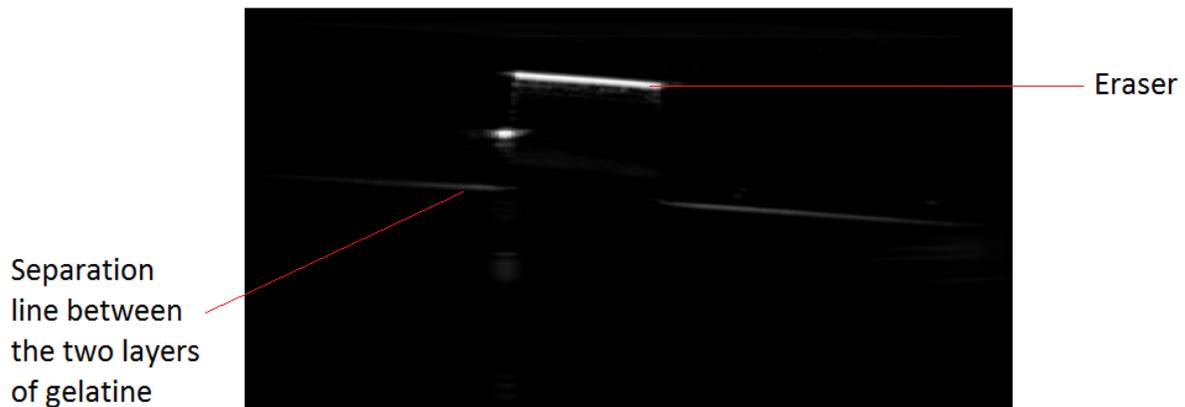


Figure 4.9: Averaging of the 180 Frames of the Eraser (838 x 440).

Figure 4.9 shows the extended field-of-view frame obtained by averaging the 180 registered frames of the eraser. Looking at this frame would make it easier for any observer to view all the information contained in all 180 frames, in one single frame. This was the goal of this thesis. As mentioned earlier, averaging works like a low-pass filter and thus it reduces the noise. It can be seen that the output in Figure 4.9 is clear. The line that appears across the center of the image is the separation between the first and second layers of the gelatine in the prepared phantom. Since the eraser experiment was done on an ultrasound phantom without scatters, speckle noise does not appear here. However, if speckle noise did exist in the images, averaging would reduce the amount of this noise. In obtaining this frame, there was no cost of losing important features that are visible in the separate frames. Once these promising results were obtained from the gelatine phantom, an in vivo experiment on the hip region was performed to test phase correlation, which is described in section 4.3.

4.3. In Vivo Experiment to Test the Subpixel Phase Correlation

An in vivo experiment was performed on the hip region. In this experiment, the motion of the probe was obtained using a mechanical arm which moved at a constant speed. Since the speed was constant, it is assumed that the motion between the frames would be relatively the same. Figure 4.10 shows the linear probe used and how it was connected to the mechanical arm. The

mechanical arm holding the probe can move upwards or downwards and thus the motion of the probe would be in the lateral direction with respect to the hip region.

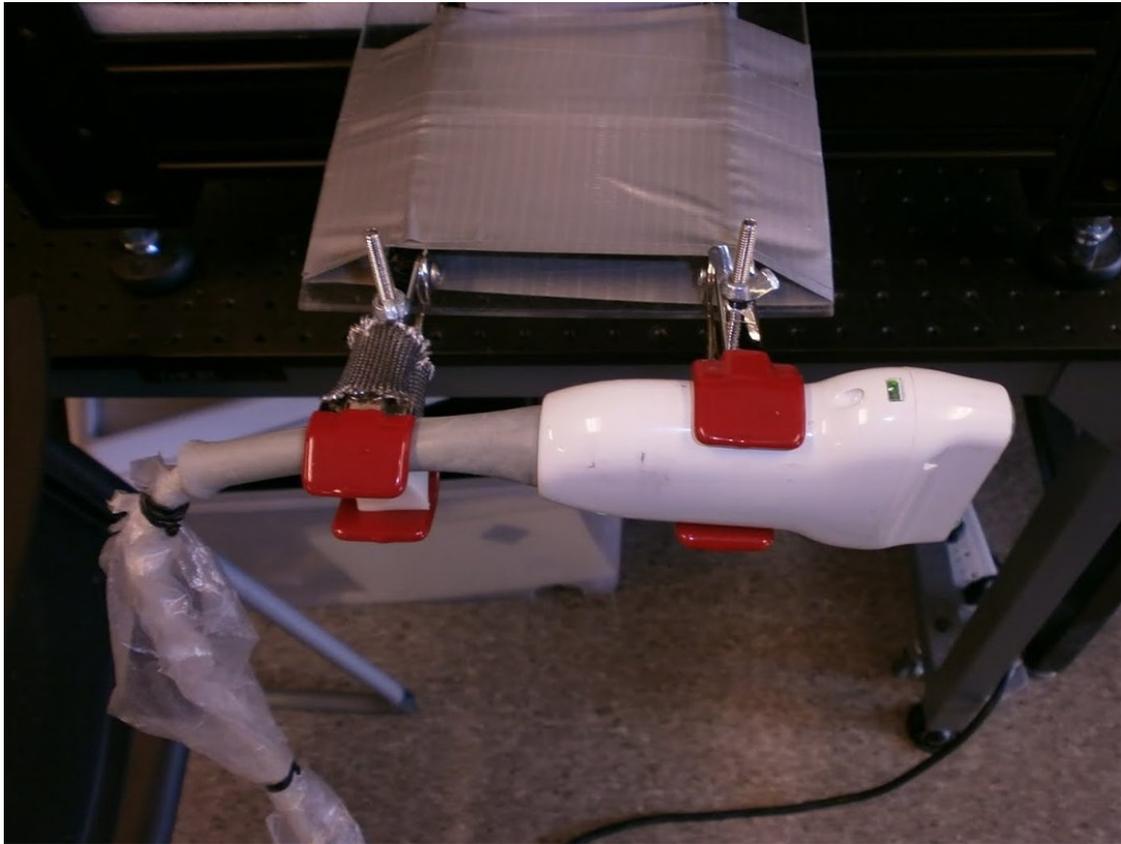


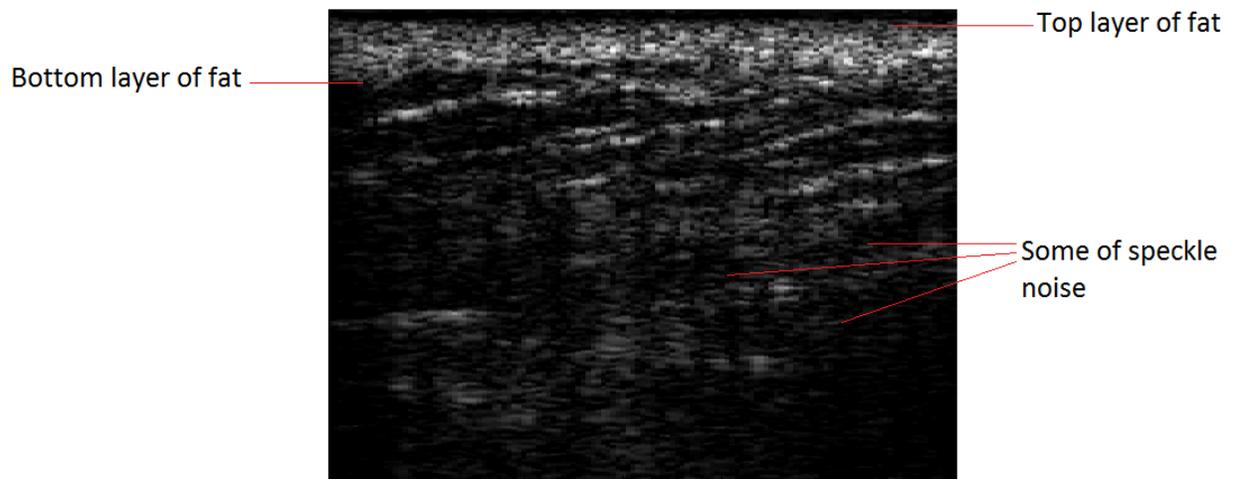
Figure 4.10: Ultrasound Probe Connected to a Mechanical Arm.

In this experiment, the probe was set to move a distance of 3 cm for a period of time of 4s, seen in Figure 4.11. As can be seen in Figure 4.11, the probe moved from location (a) to location (b) establishing a translational motion along the hip region.

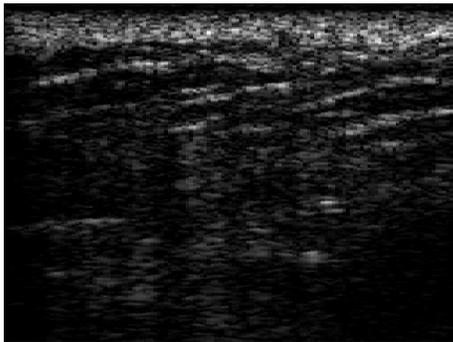


Figure 4.11: The Probe moved from Location (a) to Location (b) using a Mechanical Arm.

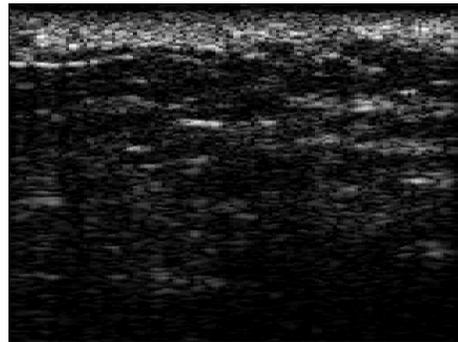
In this experiment, ART.LAB software was used to record the ultrasound images. Since, as mentioned earlier, ART.LAB allows recording of the last 6 seconds of motion with a frame rate of 30fps, 180 frames were obtained. It is expected to see a second of approximately zero motion between the frames, followed by motion between the frames in the 4 seconds of the movement of the mechanical arm, followed by another second of approximately zero motion between the frames. The speed of the mechanical arm was 0.75 cm/s. In this experiment as well, the size of each frame was 560 x 420. Figure 4.12 shows some examples of the outputs obtained in this experiment. In this figure, fat and speckle noise can be observed in the region of scan.



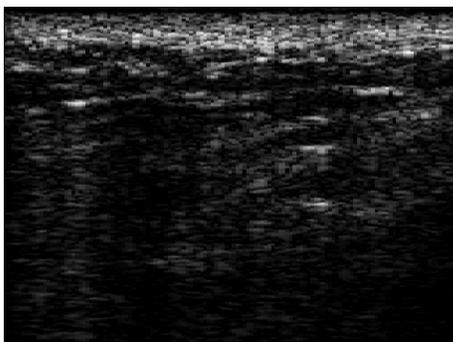
Frame 1



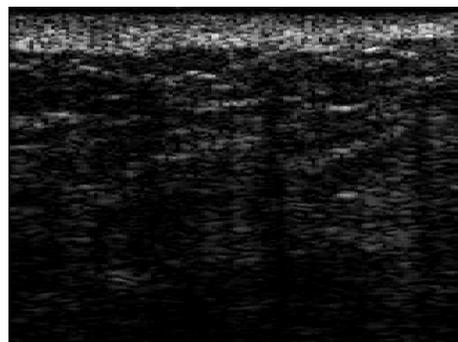
Frame 45



Frame 135



Frame 90



Frame 180

Figure 4.12: Examples of the Frames obtained along the Hip Region while the Probe moved a Distance of 3cm for 4s using a Mechanical Arm.

After the frames were all obtained, integer phase correlation was applied on them. Figure 4.13 shows the amount of motion between each two consecutive frames along the Y direction, which is the direction of motion of the probe.

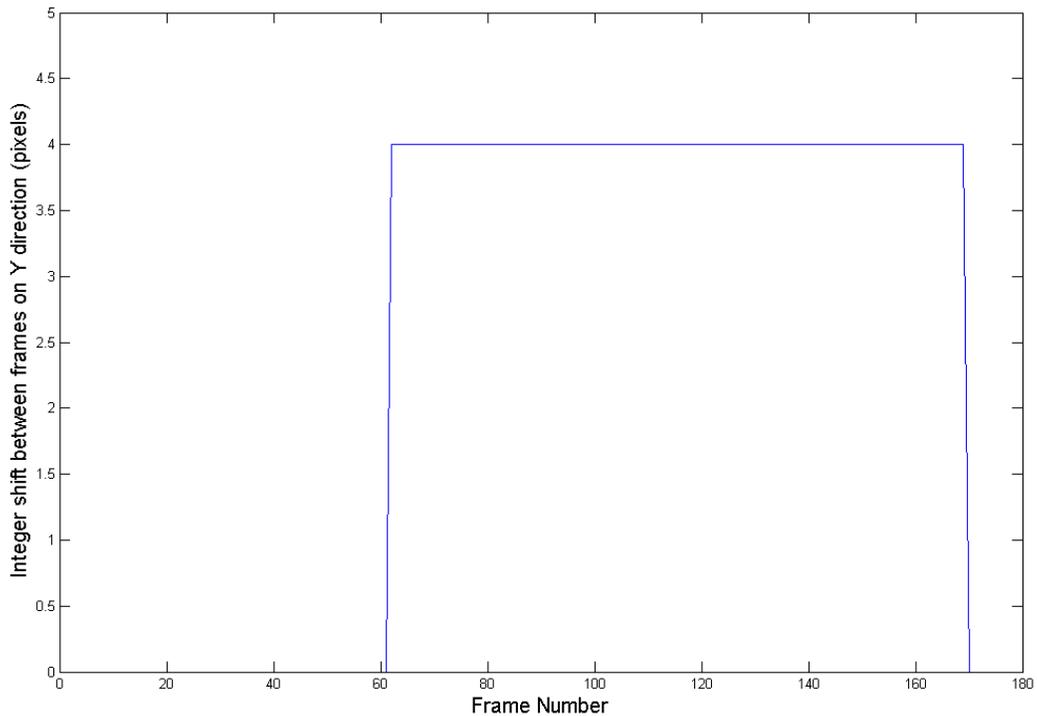


Figure 4.13: Motion Estimation between Hip Region Frames on the Y Direction using Integer Phase Correlation.

From Figure 4.13, it can be seen that from frames zero to 61, there was no motion between the frames which indicates no motion of the probe. Then between frames 61 to 169, the motion was about 4 pixels between each two consecutive frames. Frames 61 to 169 shows the period the probe was in motion using the mechanical arm. It can be seen that there was no motion from frames 170 to 180 which indicates no motion of the probe again.

The probe can cover a width of 4cm in each scan, which is equal to 560 pixels which in turn is the width of each frame obtained from processing the data which was recorded using ART.LAB software. If 560 pixels cover a width of 4cm, then each pixel covers an area of 0.71 mm from

the total width of the frame. As can be seen in Figure 4.13, the maximum motion was about 4 pixels. This 4 pixels stands for 0.029 cm. It was observed that the motion was between frames 61 and 169, which gives a total of 108 frames in motion. If we multiply the 108 frames that were in motion with the distance moved between each of the two consecutive frames in cm, we get: $108 \text{ frames} \times 0.029 \text{ cm} = 3.1 \text{ cm}$ which is approximately equal to 3cm which is the period the probe was in motion using the mechanical arm.

To estimate the motion with subpixel accuracy, the subpixel phase correlation algorithm introduced in this thesis was used and the output obtained on the Y direction is shown in Figure 4.14.

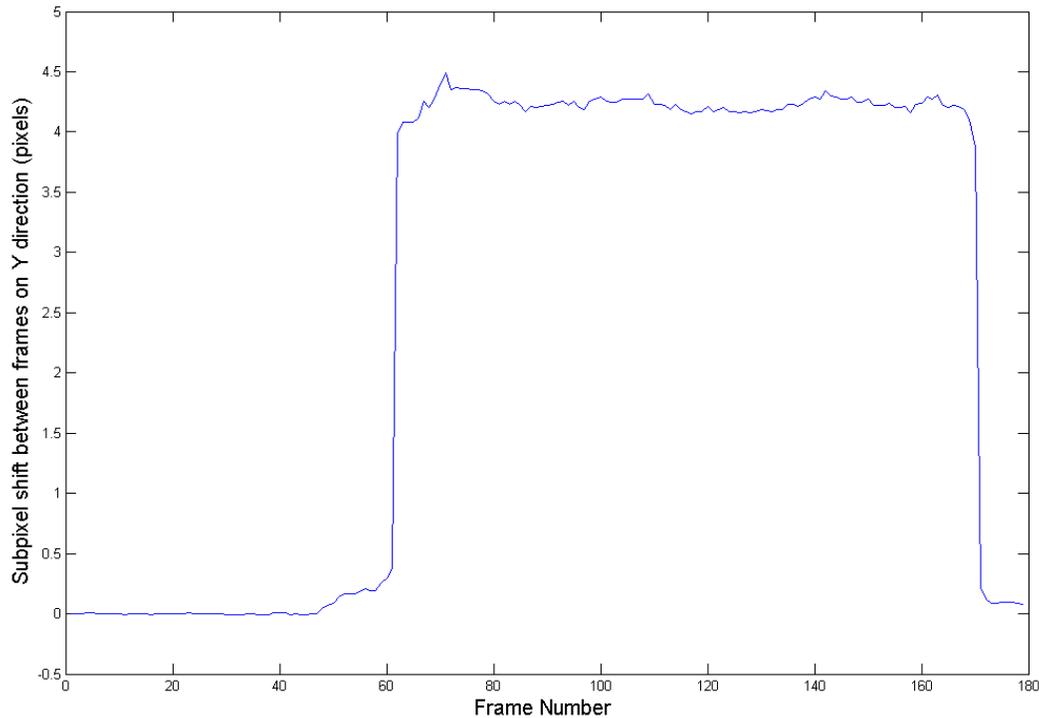


Figure 4.14: Motion Estimation between Hip Region Frames on the Y Direction using Subpixel Phase Correlation.

From Figure 4.14, the same observations seen in Figure 4.13 were obtained here but with subpixel accuracy. It can be seen that the maximum motion obtained was about 4.49 pixels and the minimum motion was zero.

After measuring the displacement along the Y direction, the change in depth between each two consecutive frames was measured using the integer phase correlation algorithm and the subpixel phase correlation algorithm. Figure 4.15 shows the output of the change of depth as the probe was in motion.

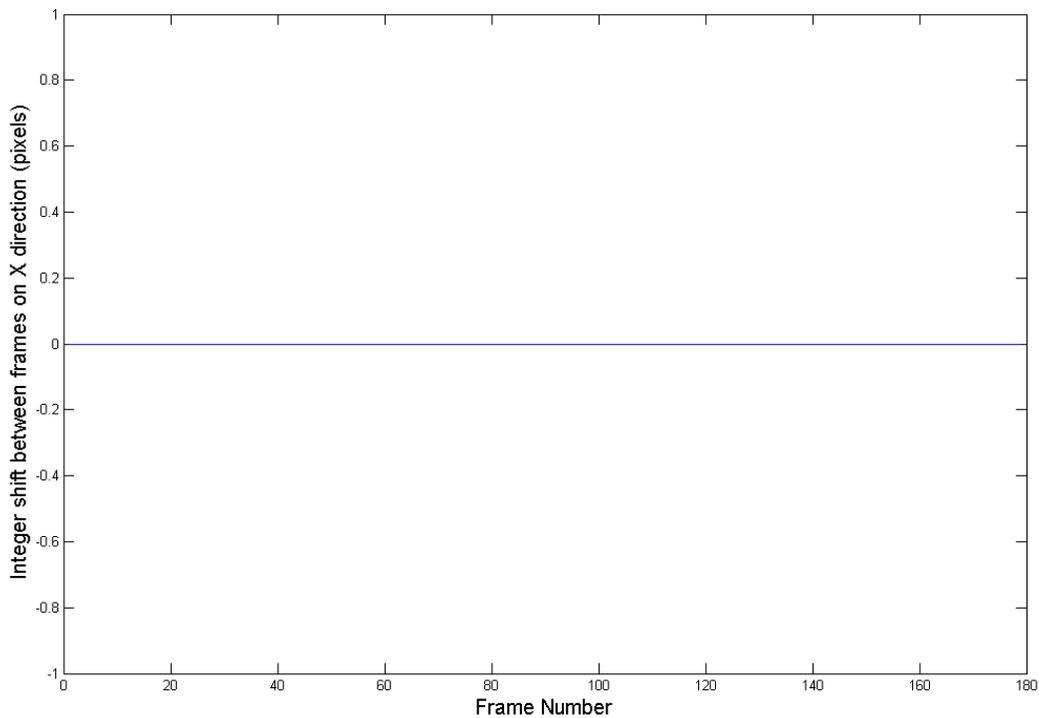


Figure 4.15: Motion Estimation between Hip Region Frames on the X Direction using Integer Phase Correlation.

It can be seen in Figure 4.15 that the change in the X direction between each two consecutive frames was zero using the integer shift phase correlation. In this experiment the change in depth would be around zero, but not exactly zero, which will be shown using the subpixel phase correlation algorithm. The change in depth is approximately zero since the probe was held and

moved with a mechanical arm and thus the pressure applied by the probe on the subject or by the subject on the probe would be the same throughout the period of motion. This of course is, as long as the subject being tested doesn't move their leg or hip during the period the probe is in motion. In this experiment, the subject being tested didn't move and thus it was expected to get a change of depth motion between the frames of approximately zero.

To estimate the change in depth with subpixel accuracy, the subpixel phase correlation algorithm was applied. The output obtained on the X direction is shown in Figure 4.16.

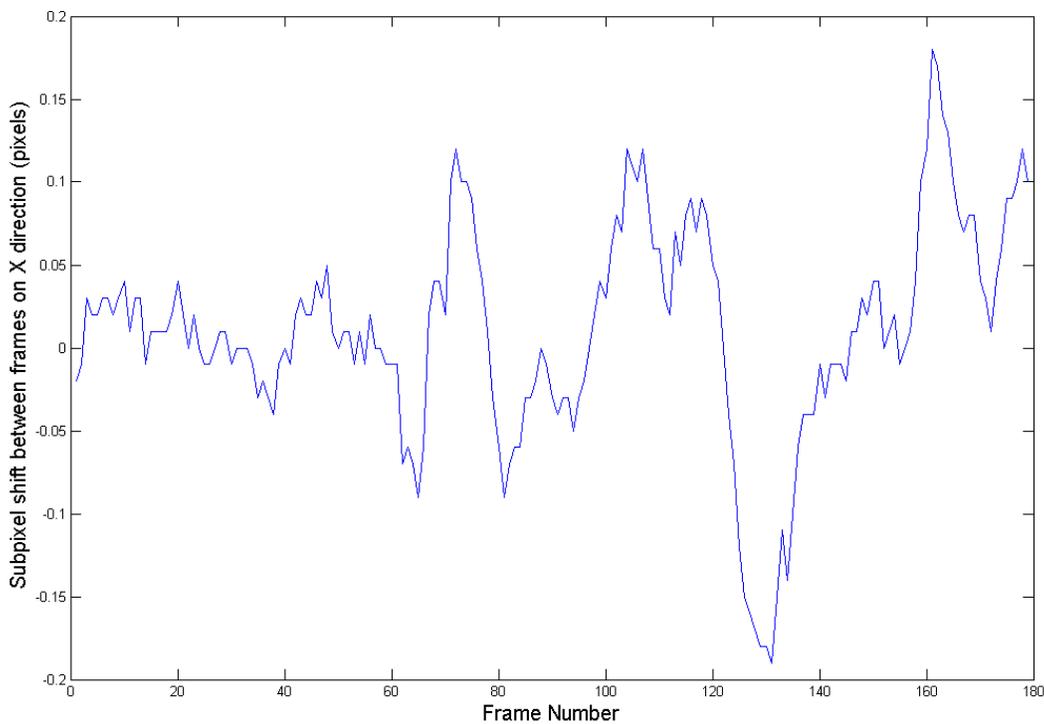


Figure 4.16: Motion Estimation between Hip Region Frames on the X Direction using Subpixel Phase Correlation.

From Figure 4.16, it can be seen that the change in depth between the frames was very small throughout the period of the experiment. The maximum was about 0.18 pixels and minimum was about -0.19 pixels. The fact that this motion was almost zero is because the pressure the probe was pressing on the subject was consistent and the subject tried his best not to move or shake his leg during the motion of the probe.

After motion estimation with subpixel accuracy was obtained using subpixel phase correlation, image registration was applied to the images. All the 180 frames in this experiment were registered with respect to frame 1. After averaging the 180 registered hip region frames, the following output of size 1025 x 422 was obtained which is shown in Figure 4.17. This figure is the extended field-of-view frame obtained in the hip experiment. Looking at the obtained extended field-of-view and the separate 180 frames that were used to obtain it, it can be observed that the tissues visible in the 180 frames are also visible in the extended field-of-view frame which as shown earlier was mostly fat. It was also observed that the overall noise was reduced in the extended field-of-view frame.

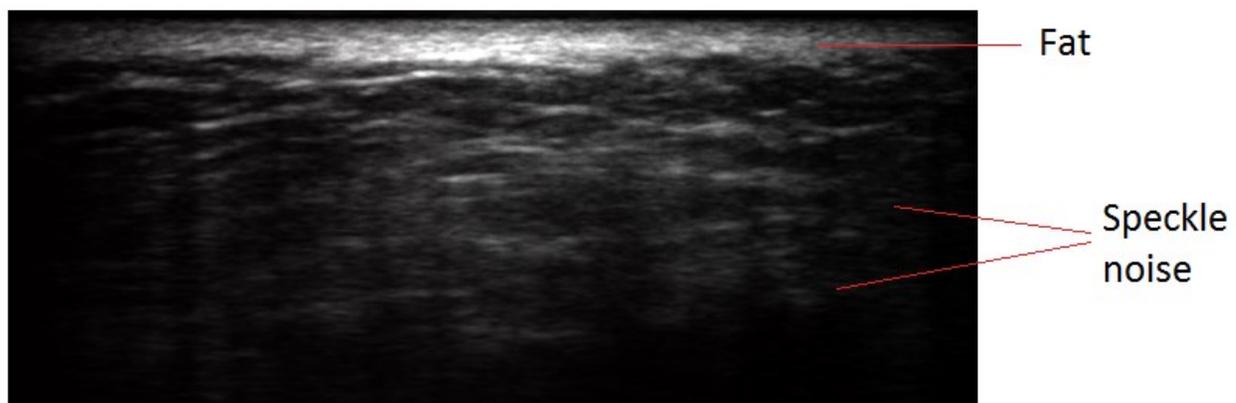


Figure 4.17: Averaging of the 180 Frames of the Hip Region (1025 x 422).

After testing the integer based phase correlation and the subpixel based phase correlation on the prepared ultrasound phantom, it was observed that the extended field-of-view obtained for the eraser was accurate since all the parts that were seen in the 180 eraser frames were also seen in the extended field-of-view frame. According to that, we can say that the extended field-of-view obtained from the hip experiment is also accurate. However, we cannot say that this is 100 % true as the ultrasound phantom did not have speckle noise which was seen in the hip experiment frames. To our knowledge, in obtaining this frame, there was not any cost of losing any important features of the images that were seen in the separate 180 frames. However, this cannot be concluded for sure without validating the accuracy of the motion estimation that was used to obtain the extended field-of-view. The results could be validated by using an electromagnetic

sensor to track the motion of the probe. To account for speckle existence in ultrasound, we went forward with performing the motion estimation using the block-matching algorithm proposed in this thesis since block-matching will account for the existence of speckle noise.

4.4. In Vivo Experiment to Test the Modified Block Matching Approach

The next part in this thesis is motion tracking through a modified block-matching approach which was discussed in chapter 3. Another experiment on the hip region was conducted. In this experiment, the motion of the probe was provided by a human subject as can be seen in Figure 4.18.



Figure 4.18: The Probe Moved from Location (a) to Location (b) while the Motion of the Probe was Provided by the Subject being Tested.

ART.LAB software was used again in this experiment and therefore a total motion of 6 seconds was recorded giving 180 frames at a rate of 30 fps. The probe was moved manually in the lateral direction along the hip region. As a result, a translational motion was obtained but here the motion was not consistent since the probe didn't move at a constant speed as was done using the mechanical arm in the previous experiment. When using a block search algorithm, a block size and a search area should be defined first. In this experiment the block size was chosen to be 16 x 16 pixels. In order to set an appropriate search area, the integer based phase correlation algorithm was applied to the 180 frames to have an idea of the maximum motion between any two consecutive frames. Figure 4.19 shows the motion estimation between the frames along the Y direction.

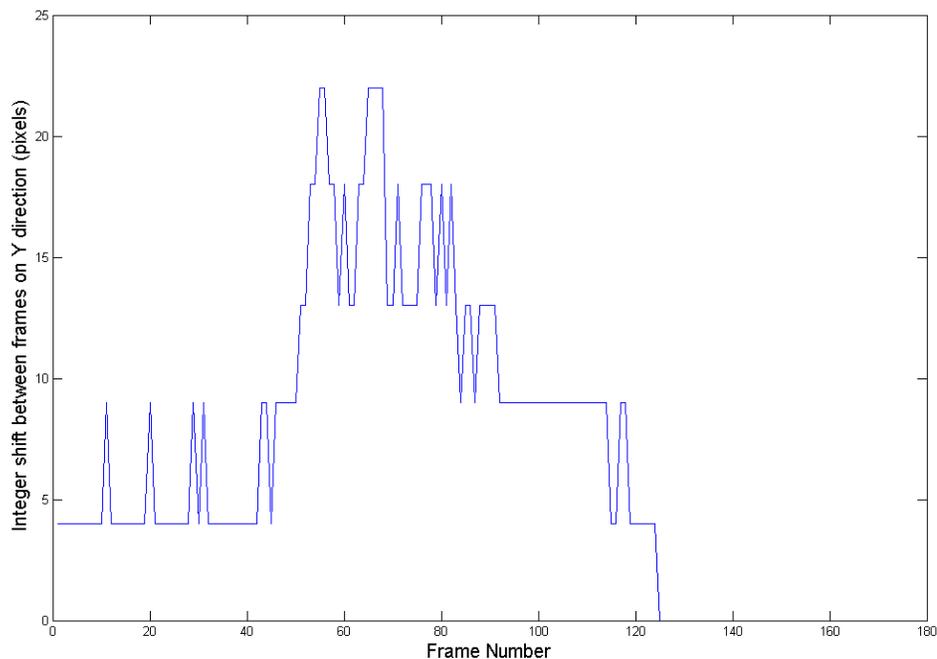
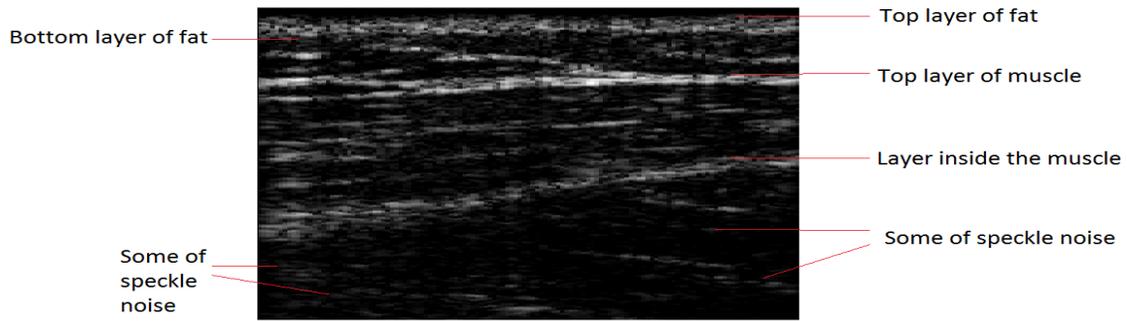


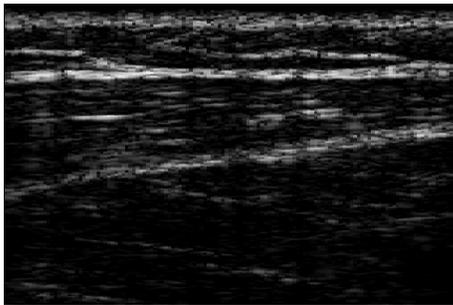
Figure 4.19: Motion Estimation between Hip Region Frames on the Y Direction using Integer Phase Correlation.

From Figure 4.19, it can be seen that the maximum motion was about 23 pixels. In choosing the search area for block matching, we looked for the maximum motion, found using the integer based phase correlation, added one to it and set the number obtained as the search area. As a result, the search area in the experiment was: $23 + 1$, which is equal to 24 pixels. This means

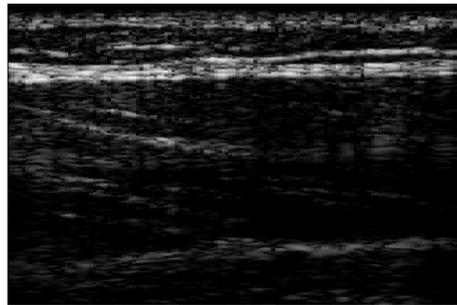
that for each macro-block in the current frame, a search was conducted for its best matched block in an area of 24 pixels surrounding the position of the macro-block, in the preceding frame. Some of the outputs obtained from this experiment are shown in Figure 4.20 which are of size 560 x 420. In this figure, fat, muscle, and speckle noise can be observed. Bone cannot be seen in these figures it is at a location deeper than the depth of the frames.



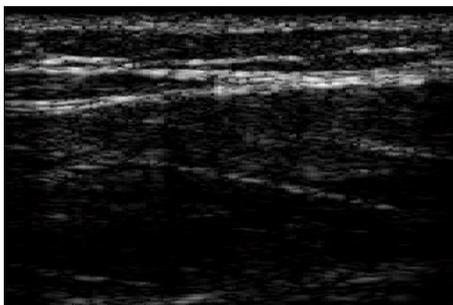
Frame 1



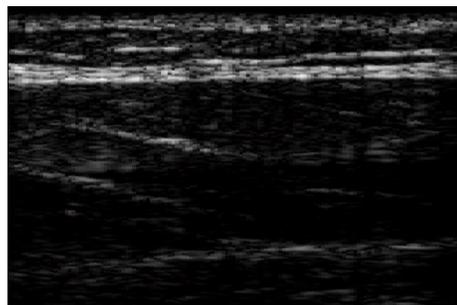
Frame 45



Frame 135



Frame 90



Frame 180

Figure 4.20: Some Outputs obtained using Ultrasound on the Hip Region while the Probe was Moved Manually.

As mentioned earlier, the frames were divided into 16 x 16 macro blocks. An example that shows the macro-blocks created can be seen in Figure 4.21 which shows the macro-blocks of sizes 16 x 16 on frame number 30 of the hip experiment.

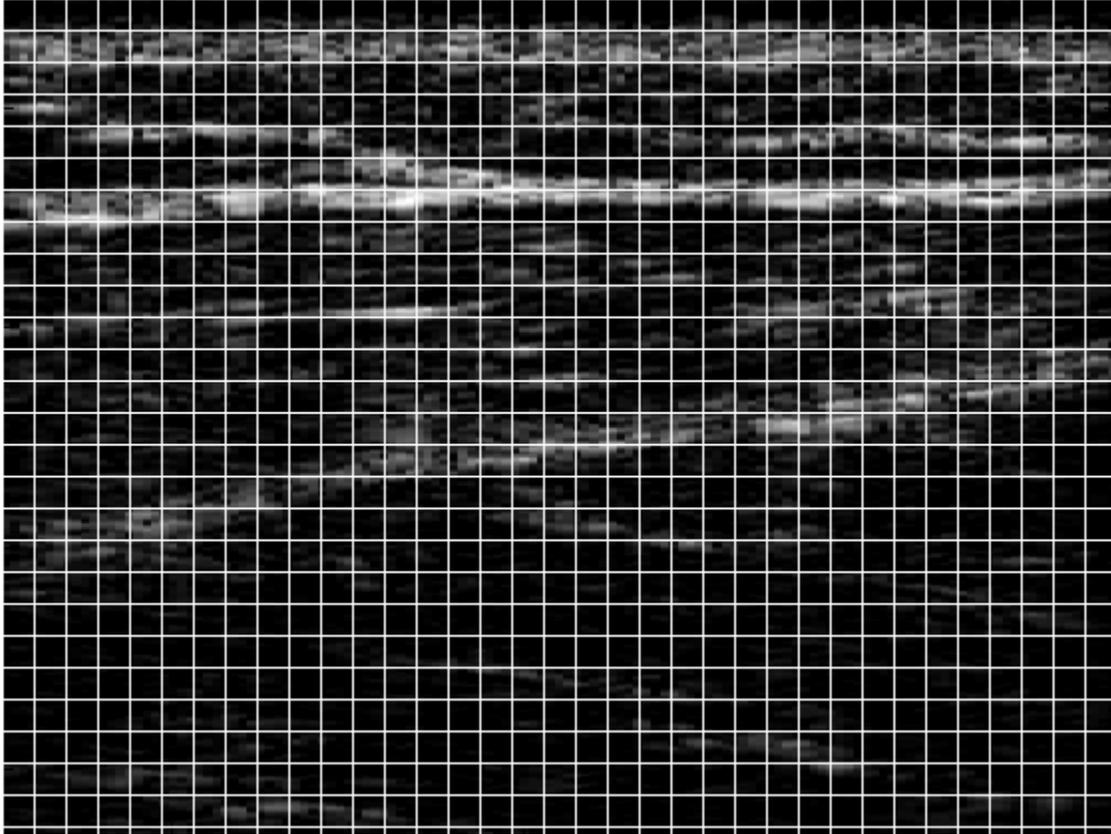


Figure 4.21: Macro-block Created on Frame Number 30 of the Hip Experiment (image size 560 x 420).

In order to get an image that is a multiple of the block size chosen, the images were cropped. As a result, image sizes of 560 x 416 were obtained. An example of that can be seen in Figure 4.22. Each macro-block was searched in an area of 24 pixels on all the directions of the macro-block in the reference frame using exhaustive search block-matching to find the best matched block, which represents the integer motion vector between these two blocks. After the integer motion vector is found, Taylor approximation was applied to each macro-block to find the motion vector with subpixel accuracy. An example of that can be seen in Figure 4.23. Figure 4.23 shows the motion vectors for each macro block between frames 30 and 31 where the directions of the

arrows represent the direction of the probe's motion and the length of the arrows represents the magnitude of the shift.

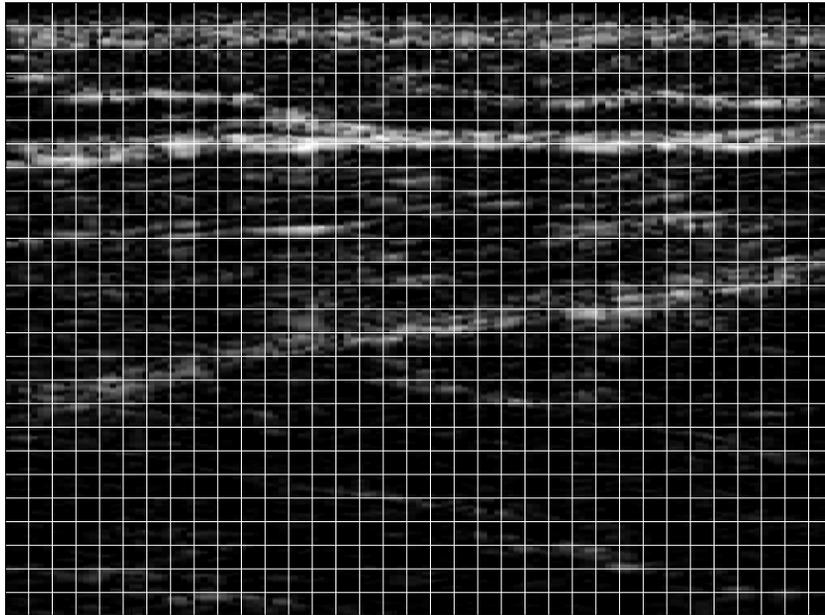


Figure 4.22: Cropped Version of Figure 4.21(image size 560 x 416).

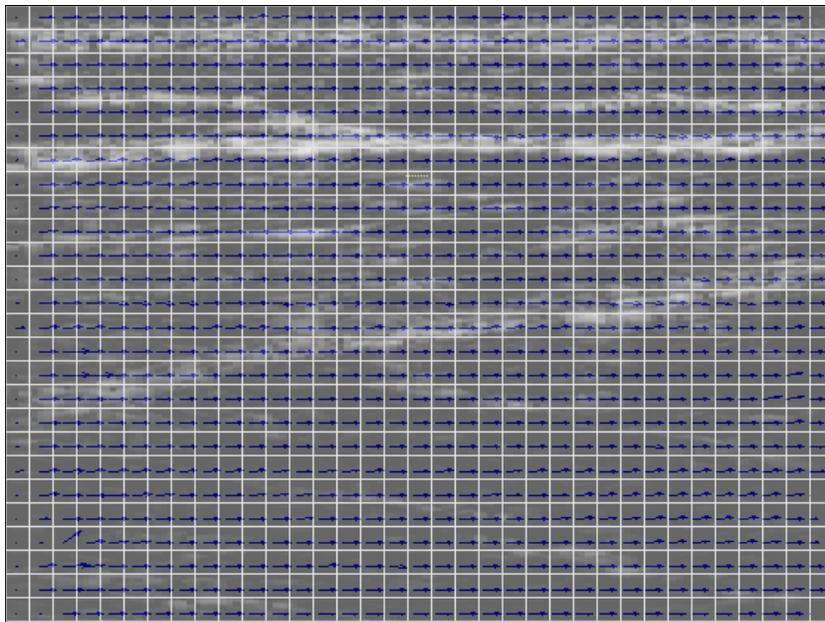
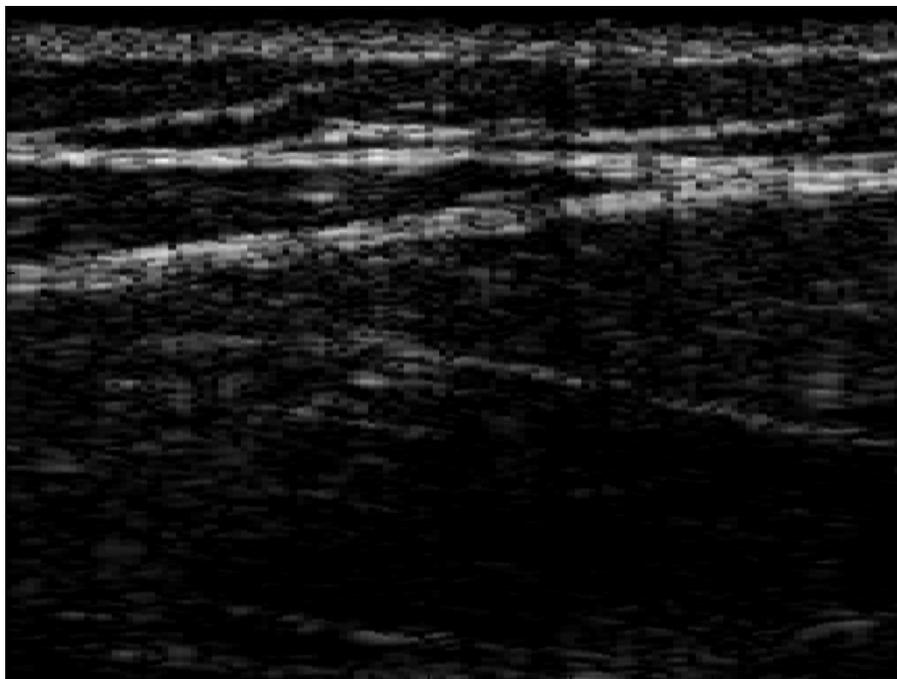
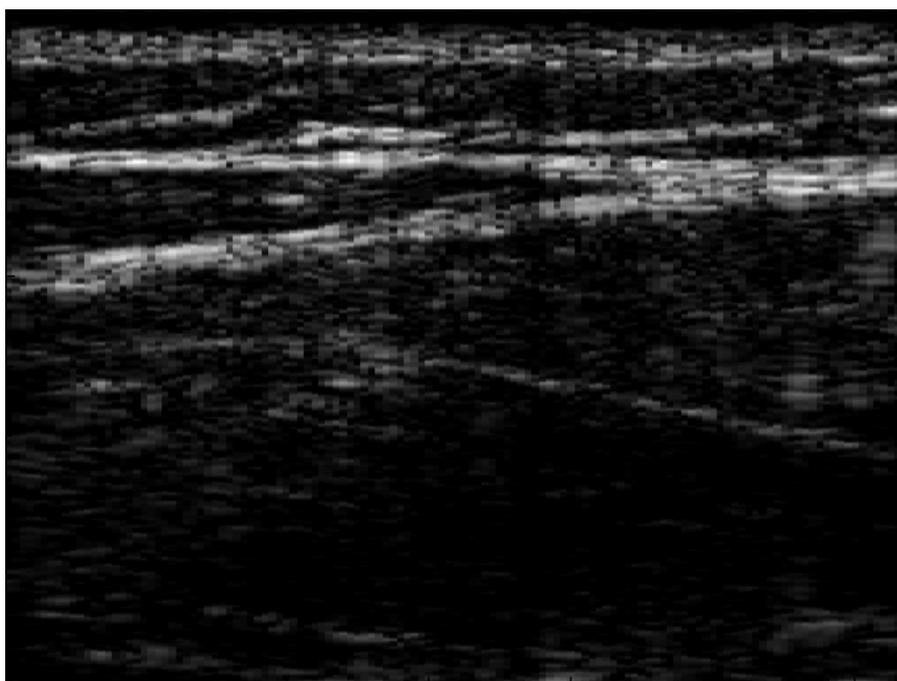


Figure 4.23: Motion Vectors Overlaid over Frame 30 (image size 560 x 416).

As mentioned earlier, speckle noise in ultrasound correlates between frames and thus that helps in using them for tracking motion. This is not always the case since speckle decorrelation also appears in ultrasound. Also, blocks of no motion could affect the accuracy for motion tracking as there could be more than one block near the block of no motion which could give a wrong motion vector. In order to overcome these problems in this thesis, the standard deviation of each macro block was calculated. If the value for the standard deviation was zero, this macro block was not searched for its best matched block and a zero motion vector was assigned to it. As for overcoming the no correlation problem between frames, a threshold was put for the MAD cost function that was used to find the best matched block. The threshold that was used in this experiment was 0.14. If the MAD function for the best matched block was greater than 0.14, this macro-block's motion vector was not found since, the correlation is not high enough between the two macro-blocks and it was concluded that this macro-block is hard to track. This value was chosen after trying several values to eliminate areas that are visible to have false motion vectors. An example showing how the modified block matching search works is shown in Figure 4.24. Figure 4.25 shows the motion vectors for each macro block between frames 80 and 81. If we look at frames 80 and 81 and try to compare the motion between them, we can see that some of the motion vectors shown in Figure 4.25 are false. This is due to the reasons mentioned before which are either decorrelation between the frames and/or assigning a wrong motion vector for blocks which contains no motion. Figure 4.26 shows some of the wrong motion vectors that can be observed by the eye.



(a)



(b)

Figure 4.24: (a) Frame 80 of Hip Experiment (Actual size is 560 x 420). (b) Frame 81 of Hip Experiment (Actual size is 560 x 420).

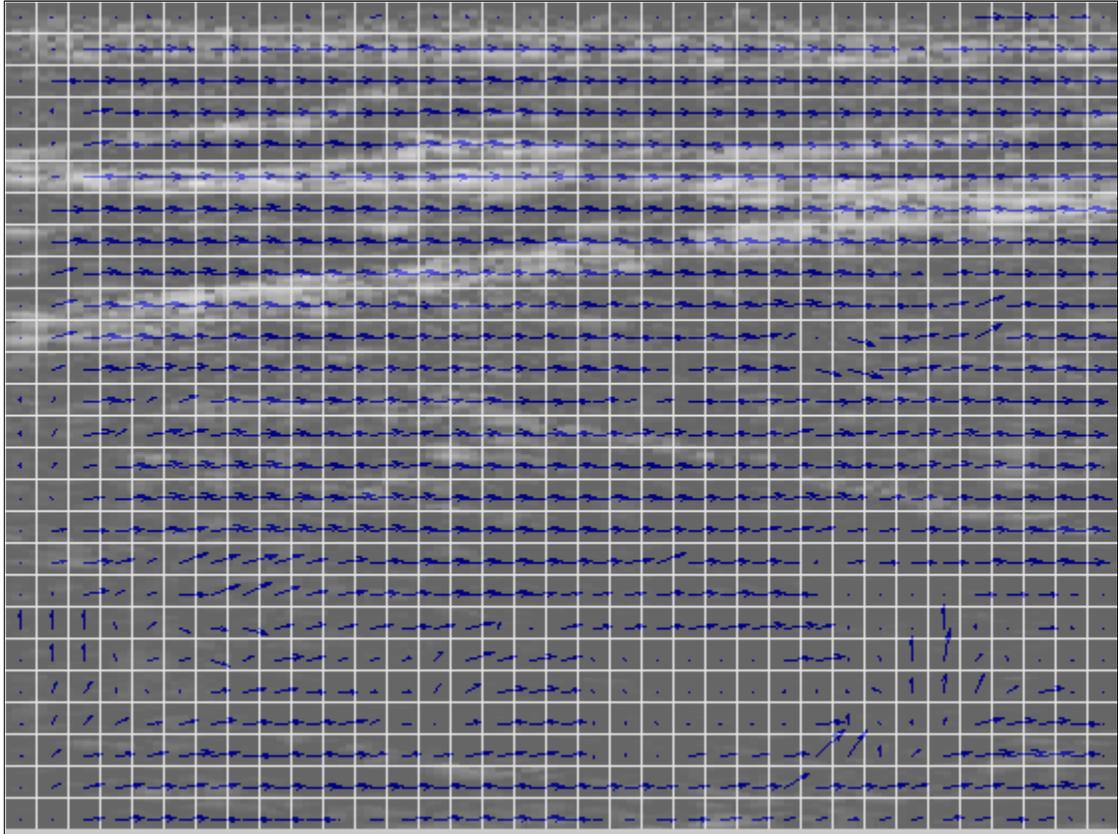


Figure 4.25: Motion Vectors Overlaid over Frame 80 of the Hip Experiment.

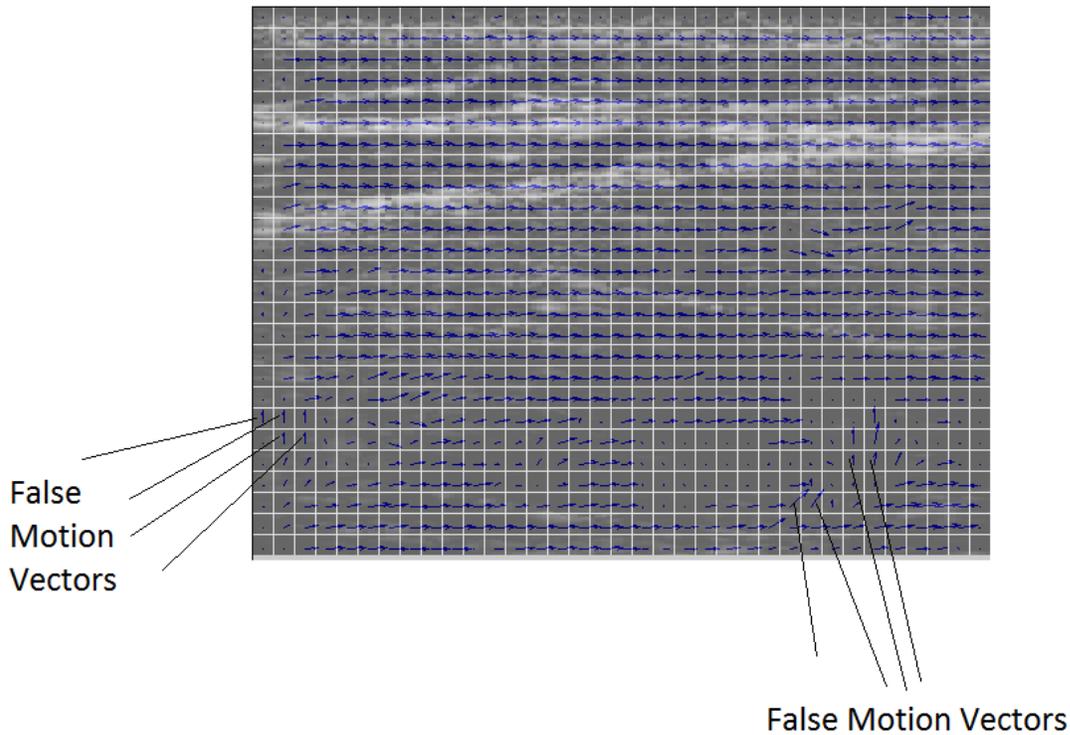


Figure 4.26: Some of the Wrong Motion Vectors Overlaid over Frame 80 of the Hip Experiment.

Figure 4.27 shows the motion vectors between frames 80 and 81 after applying the modified block-matching search. In this Figure, the blocks that were hard to track are indicated by green blocks, the blocks that have no motion are indicated by a dot in the block, and finally the other blocks are indicated by an arrow that shows how much the displacement between the current block and the best matched blocks is. It can be seen in this Figure that the motion vectors that were seen to be false (Figure 4.26) were eliminated and thus the modified algorithm shows success in this area of motion tracking.

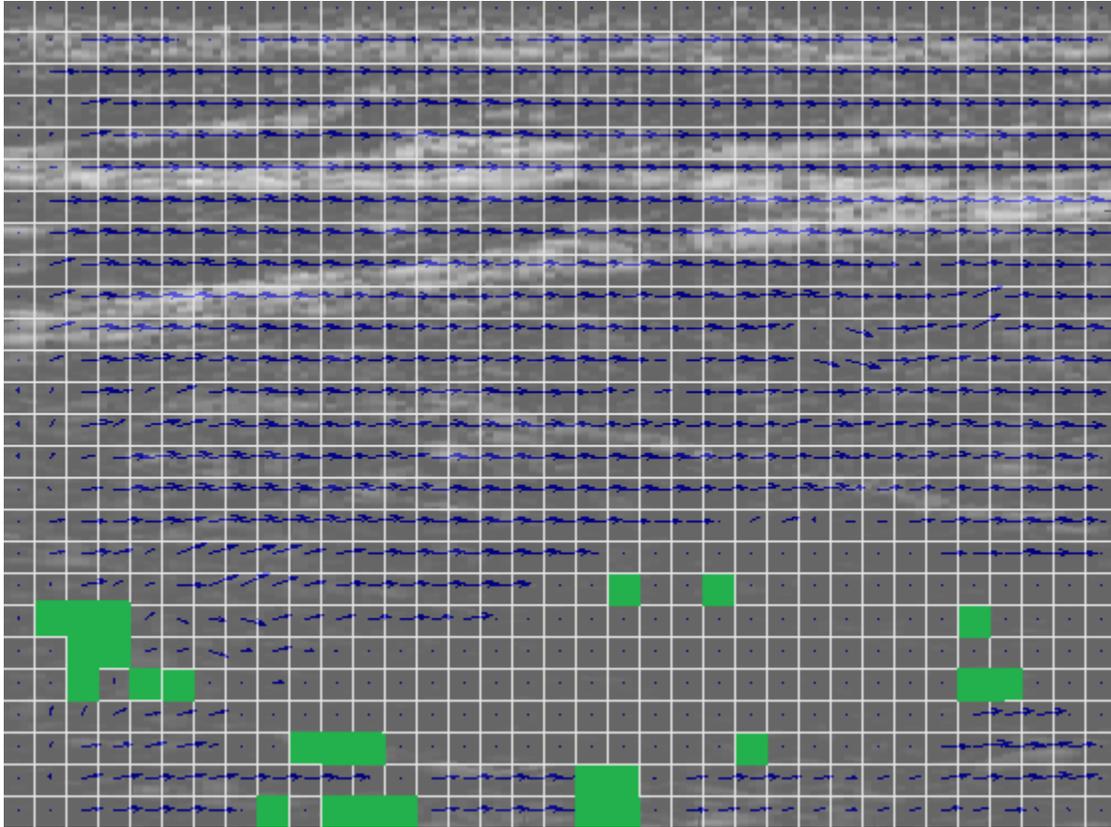


Figure 4.27: Motion Vectors over Frame 80 after Applying the Modified Block Matching Search.

Another reason why some blocks were hard to track could be because of tissue changing its shape. Under compression, tissues change their shape to smaller sizes. In the hip experiment where the probe was moved manually, probe compression could have made some changes to the shapes of tissues, and thus may be the reason why some of the blocks were not tracked.

In order to validate the results obtained by the modified block-matching are better than those obtained using block matching without modification, the signal-to-noise-ratio was calculated. First, signal-to-noise-ratio was obtained from the normal block-matching without modification and the second time it was obtained from the modified block-matching. This was done to see how well the signal improved with the modified block-matching search. In order to obtain the signal-to-noise-ratio, MSE needs to be found. MSE finds the mean squared error between two macro-blocks, one in the current frame and one in the reference frame. The macro-block in the reference frame with the lowest MSE is the best matched block. After finding the MSE for each

macro-block, the signal-to-noise ratio can be found and thus the accuracy of the modified block-matching will be seen. To do so, the MSE was first calculated by the following formula:

$$\text{MSE} = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (A1_{ij} - A2_{ij})^2 \quad (4.1)$$

Where $N \times M$ is the size of the macro-block which was chosen to be 16×16 in this thesis. $A1_{ij}$ and $A2_{ij}$ are the pixels compared between the current macro-block and the reference macro-block respectively. Following that the peak-signal-to-noise ratio (PSNR) was calculated by the following formula:

$$\text{PSNR} = 10 \log_{10} \frac{(\text{Peak to peak value})^2}{\text{MSE}} \quad (4.2)$$

Figure 4.28 shows the output obtained for the PSNR. The red line represents the PSNR for the original full search block-matching and the blue line represent the PSNR for the modified full search block-matching that was developed in this thesis after accounting for areas of no motion and excluding areas that were hard to track. If we take another look at Figure 4.19, we can see that the observed motion was between frames 1 until about frame 125. For this period of motion, it can be seen in Figure 4.28 that the PSNR for the modified block-matching search is relatively better than those for the original block-matching search and as a result we can conclude that the modified block-matching search is expected to give more accurate results in motion estimation using.

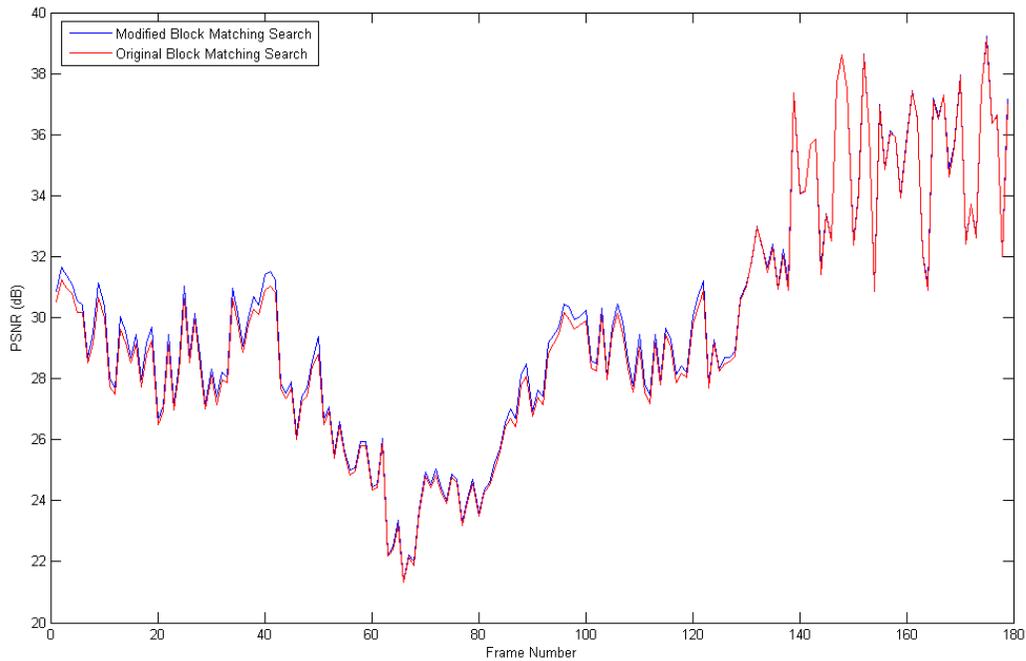


Figure 4.28: PSNR for the Original Full Search Block Matching and for the Modified Full Search Block Matching.

After the modified block-matching search was applied, the images were taken to apply phase correlation on them. The parts that had false motion vectors were eliminated from the images so that phase correlation motion estimation improves. After that was done, image registration was applied and averaging of the total number of frames was done. As a result, an extended field-of-view image was obtained. The subpixel shift using subpixel phase correlation between frames on the Y direction can be seen in Figure 4.29.

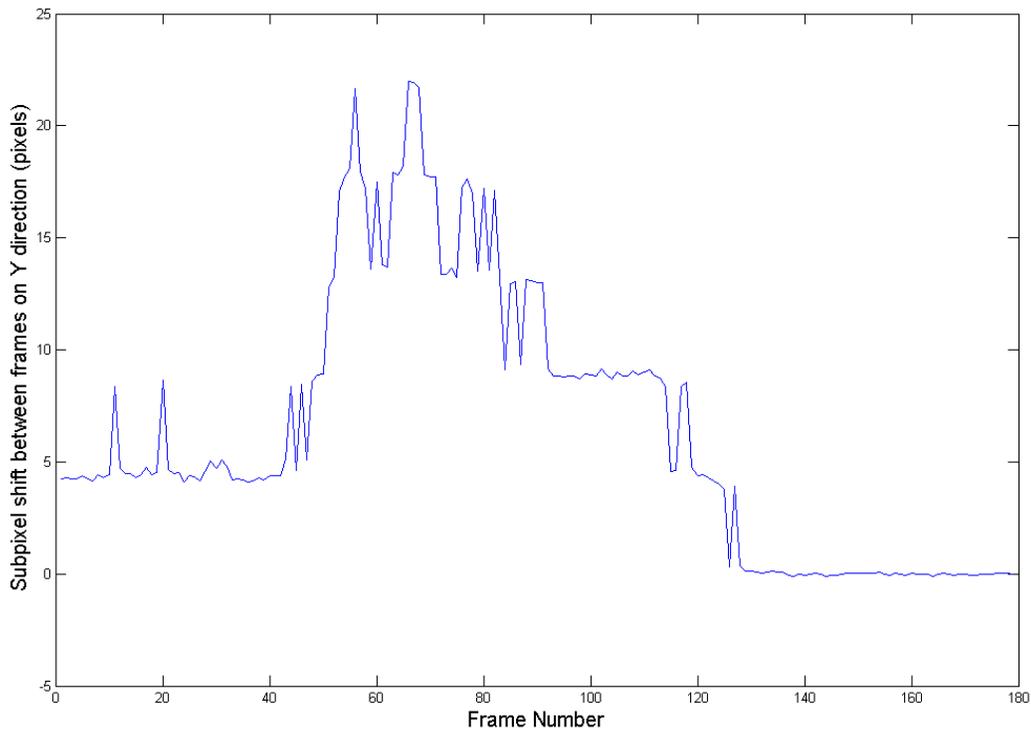


Figure 4.29: Motion Estimation between Hip Region Frames on the Y Direction using Subpixel Phase Correlation.

From Figure 4.29, it can be seen that the motion was not consistent as the probe did not move at a constant speed. The probe's motion started by around 4.2 pixels between frames zero to 45. Then the motion increased gradually to reach a maximum of 22 pixels between frames 66 and 67. The motion then decreased gradually from frame 67 to 180 to reach a minimum of zero. The accuracy of the motion estimation found here would be seen after image registration and averaging is done on all the frames.

After obtaining the motion along the Y direction, the motion along the X direction was obtained using subpixel phase correlation. Figure 4.30 shows the output obtained. It can be seen in this Figure that the minimum change in depth was about -0.1 pixels and the maximum was about 0.53 pixels. It can be seen that the maximum change of depth was between frames 60 until 90 which are the frames that had maximum displacement along the Y direction as was seen in Figure 4.29. The reason behind this is that when the probe was moved manually during the

motion capture, pressure could have been applied against the body being tested changing the depth in that area. Due to the change in depth between frames 60 to 90, it is expected that tissue in that region would change shape. It is true that the change in depth has a maximum of about 0.5 pixels in that region which is not a big number, but this 0.5 pixels of compression could put some effects on the shape of the tissue.

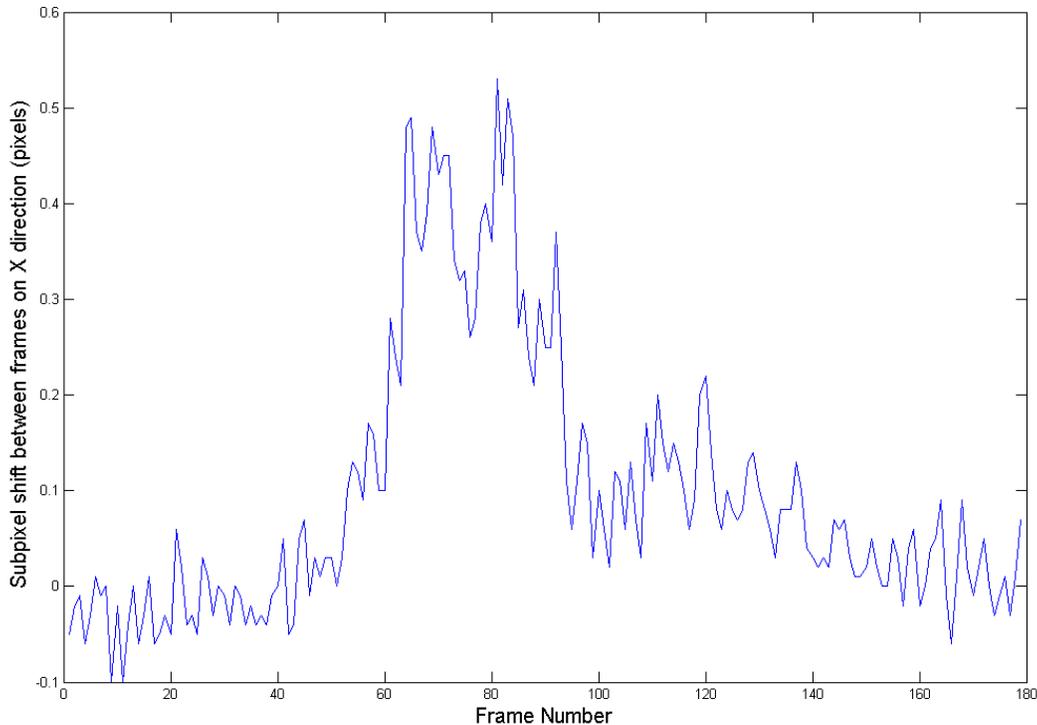


Figure 4.30: Motion Estimation between Hip Region Frames on the X Direction using Subpixel Phase Correlation.

The next step that was done is registering all 180 frames with respect to frame 1 of the experiment and then averaging the registered frames to obtain one frame that has an extended field-of-view. This frame can be seen in Figure 4.31. The actual size of the frame obtained is 1721 x 438 pixels. The size here was reduced for visualisation purposes. Looking at Figure 4.31, one can observe information about the total area of the scan instead of looking at 180 images. Motion tracking to obtain this frame required more work, but the result obtained is worth that work. Accuracy in medical imaging is a key factor for detecting different diseases.

Looking at this extended field-of-view frame and looking at the separate 180 frames, where a sample was shown in figure 4.20, it was observed that the major tissues that were seen in the separate frames are also seen in the extended field-of-view frame, which as described earlier were fat and muscle tissues. It was also observed that the overall noise was reduced in the extended field-of-view frame. Again in this experiment, to our knowledge, in obtaining this frame, there was not any cost of losing any important features of the images that were seen in the separate 180 frames. However, this cannot be concluded for sure without validating the accuracy of the motion estimation that was used to obtain the extended field-of-view.

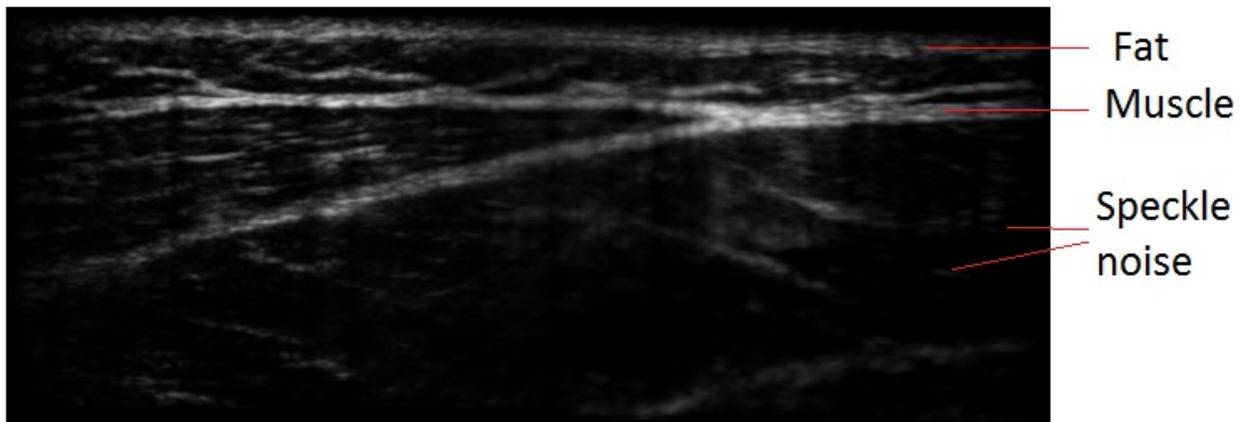


Figure 4.31: The Extended-Field-Of-View Frame of the Hip Experiment.

In order to see how well the modified block-matching search works in the presence of muscle contraction and probe compression, two experiments were conducted on the upper leg region. For the muscle contraction experiment, while the probe was in motion, the subject contracted the muscle in the upper leg region frequently. While for the probe compression experiment, the probe was pushed frequently against the upper leg region to see probe compression effects. These two experiments are explained in details in section 4.5 and 4.6 respectively.

4.5. In vivo Experiment to Observe Muscle Contraction

An experiment was conducted to see how the modified block-matching approach can detect the motion in the presence of muscle contraction. In this experiment the US probe moved manually along the upper part of the leg as can be seen in Figure 4.32. The probe was moved from a point (a) which is 31 cm above the knee to a point (b) which is 16 cm above the knee as can be seen in Figure 4.32 and thus the total motion was 15 cm along the lateral direction. During the period the motion was recorded, the subject contracted his muscles in the upper leg area while the probe was in motion.

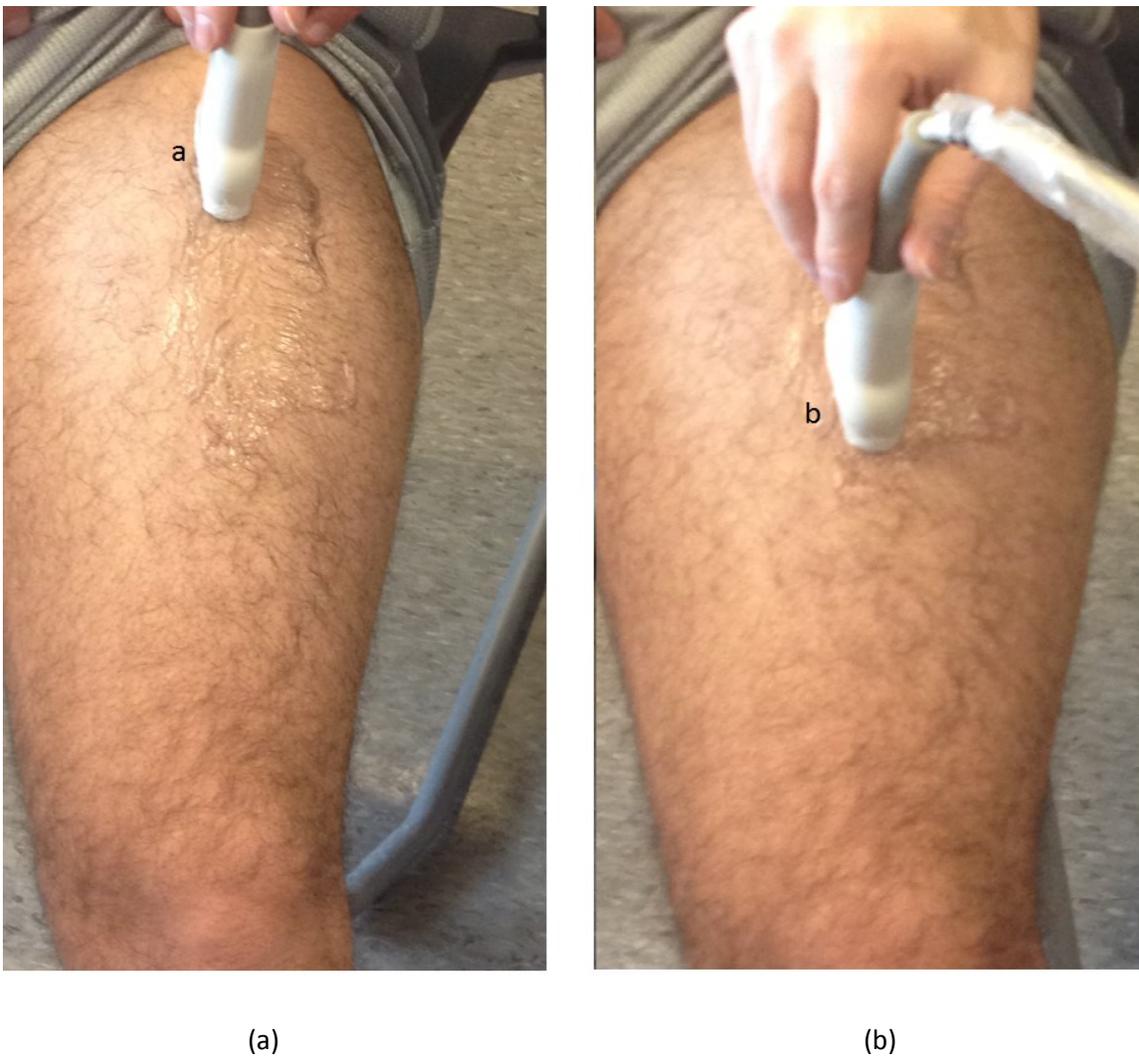
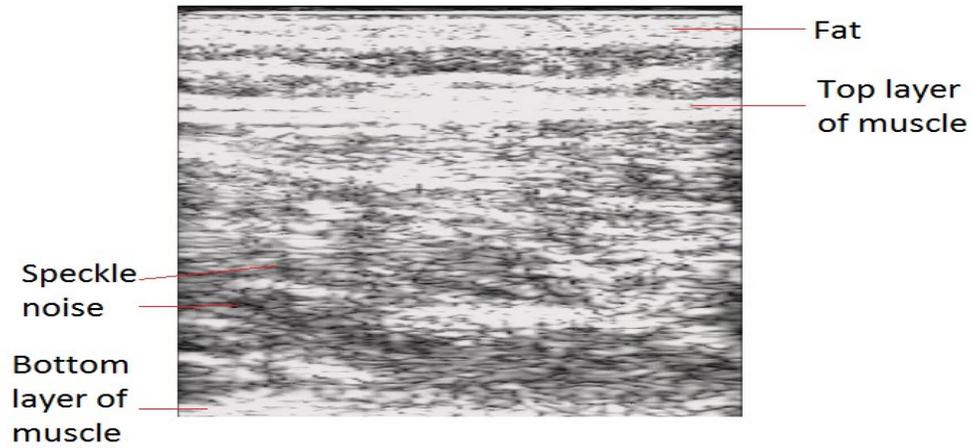
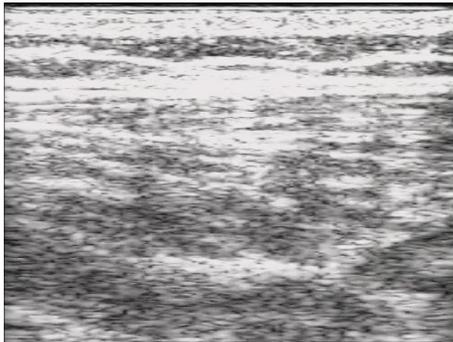


Figure 4.32 : (a) Start Position for the Ultrasound Probe. (b) End Position of the Ultrasound Probe.

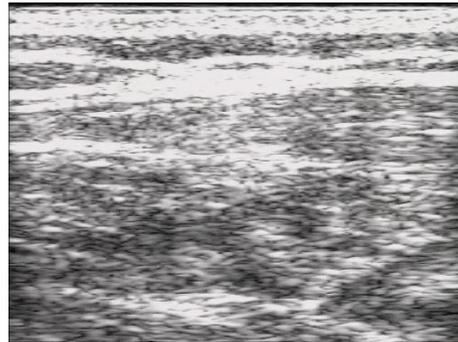
In this experiment, a video capture system was used to capture the motion and thus we had the opportunity to record the motion of the probe for as long as we wanted, unlike the case with ART.LAB which only allowed the recoding of the last six seconds of motion. The period the motion was recorded for in this experiment was 30 seconds. The probe moved at a very low speed of about 0.5cm/sec while the subject contracted the muscles in the upper leg area constantly. The size of each frame obtained using the video capture system in this experiment is 394 x 398. Figure 4.33 shows some of the outputs obtained in this experiment. In this figure, fat, muscle and speckle noise can be observed. It shows the period where we started with no contraction as can be seen in (a) then the period while the muscle was in the contraction process in (b) followed by a fully contracted muscle in (c) and then back to relaxation mode in (d).



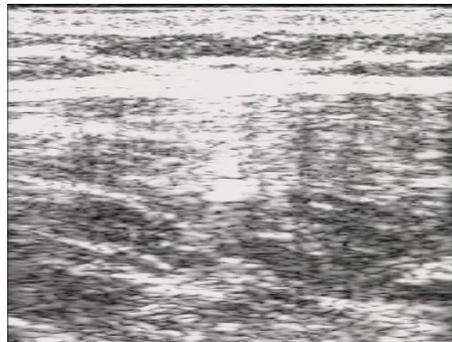
(a) Frame 120



(a) Frame 130



(b) Frame 140



(c) Frame 160

Figure 4.33: Some Output Frames Obtained During the Experiment of Muscle Contraction. (a) Output while Muscle is Relaxed. (b) Output During Muscle Contraction. (c) Output When Muscles are Fully Contracted. (d) Output when Muscle is back to Relaxation.

The total number of frames obtained in this experiment was 900 frames. The first step that was done on the frames was calculating the integer phase correlation between the frames in order to know how much the maximum motion between each two consecutive frames was. This was so an appropriate search area for the modified block-matching search can be set. Figure 4.34 shows the output obtained from the integer based phase correlation. Looking at Figure 4.34 it can be observed that the maximum motion between consecutive frames was about 3 pixels along the direction of motion. Setting the search area used in the modified block-matching search to the maximum number of motion found using the integer based phase correlation plus one, the search area for each macro-block being searched in this experiment was set to 4 pixels.

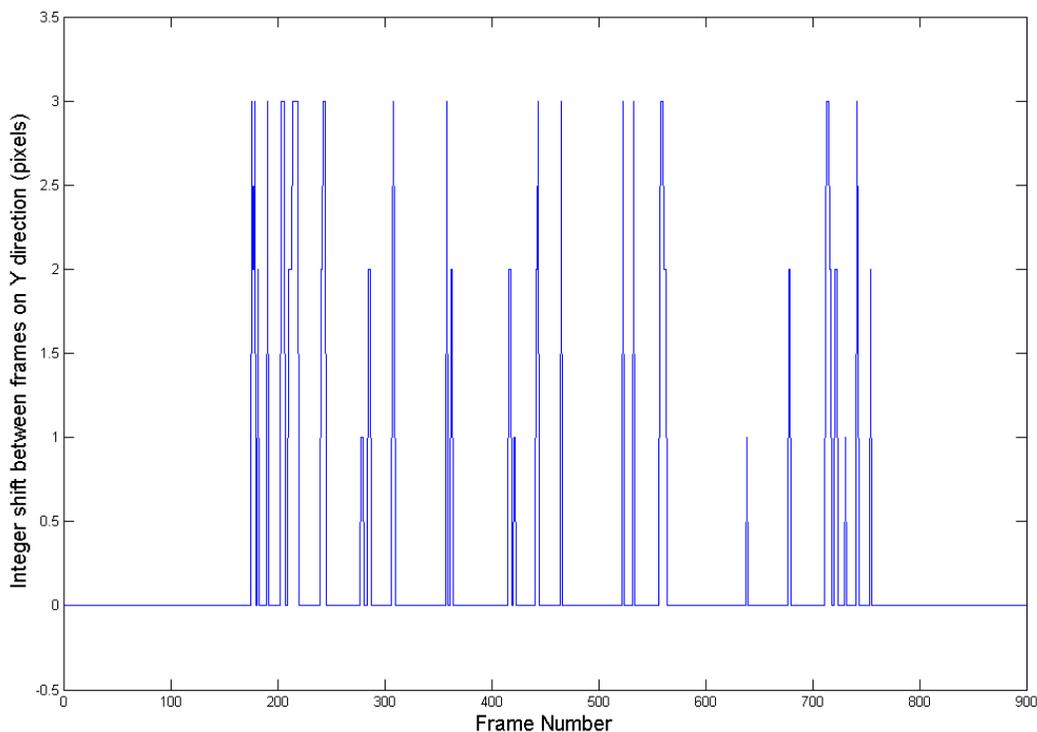


Figure 4.34: Motion Estimation between Hip Region Frames on the Y Direction using Integer Phase Correlation.

All the frames were divided into 16 x 16 macro-blocks and then each macro-block was searched in the search area specified in the frame preceding the current frame being searched. Since the frames obtained in this experiment are of the size 394 x 398, the frames were zero padded to

obtain a size of 400 x 400 so that the frames are made of a multiple of the macro-block size chosen which is 16 x 16. Due to doing so, 25 macro-blocks were obtained on the vertical direction of the image and 25 macro-blocks were obtained on the horizontal direction as well. Figure 4.35 shows an example of an outputted frame that was divided into 16 x 16 macro-blocks after being zero padded to make a size of 400 x 400.

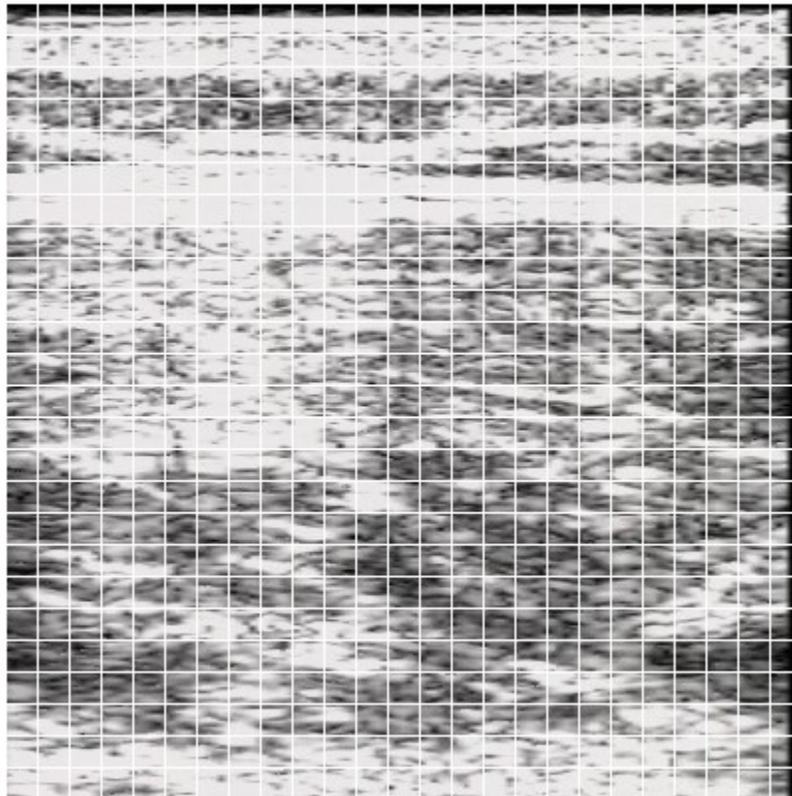
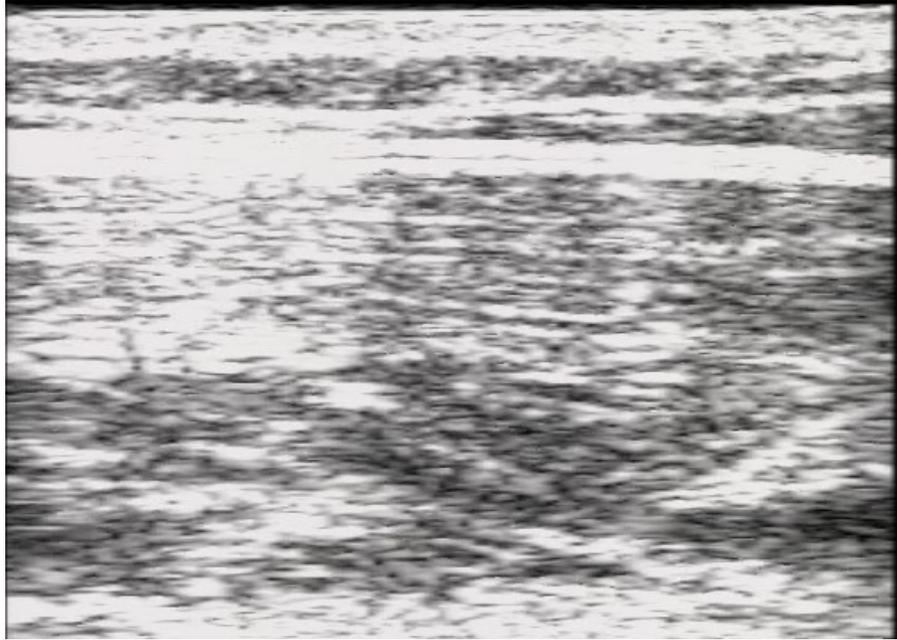
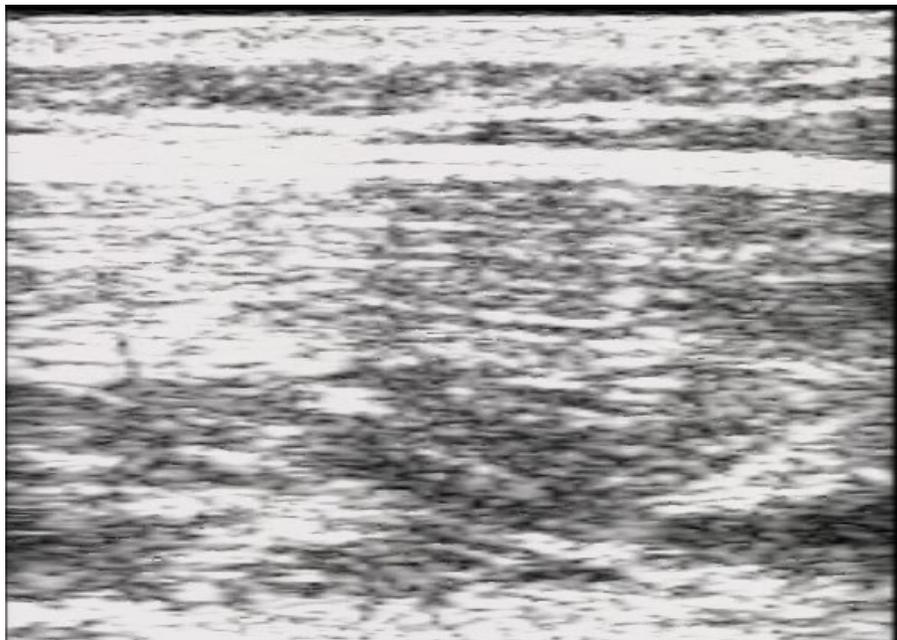


Figure 4.35: Frame Number 200 Divided into 16 x 16 Macro-Block from the Experiment Performed on the Upper Part of the Leg (Size 400 x 400).

After the frames were divided into 16 x 16 macro-blocks, each macro-block was searched in the search area defined in the reference frame. The motion vectors obtained by the modified block-matching search shows that the modified block-matching search is able to track the motion between frames even in the presence of muscle contraction. Figure 4.36 shows case of a translational motion between frames 214 and 215 while the muscle in the upper leg region is relaxed and the probe is in motion.



(a)



(b)

Figure 4.36: (a) Frame 215 of the Upper Leg Experiment. (b) Frame 216 of the Upper Leg Experiment.

Figure 4.37 shows the motion vectors for each of the macro-blocks between frames 215 and 216 originated from the motion between the two frames which were shown in Figure 4.36. Looking at Figure 4.37 it can be seen that the motion vectors obtained are along the direction of motion of the probe and it can be seen from the motion vectors that muscle contraction does not exist between frames 215 and 216.

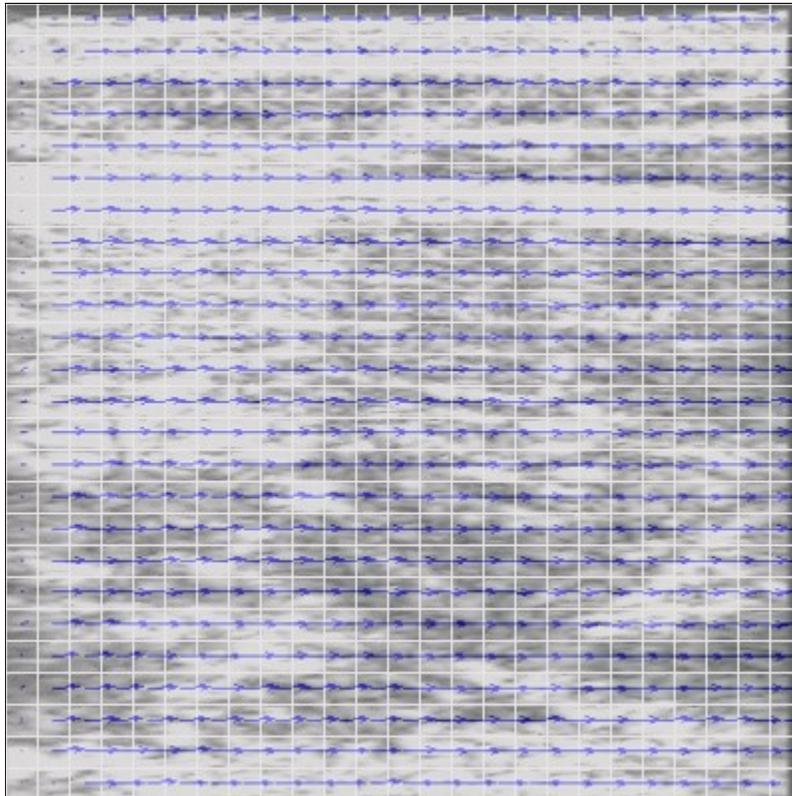
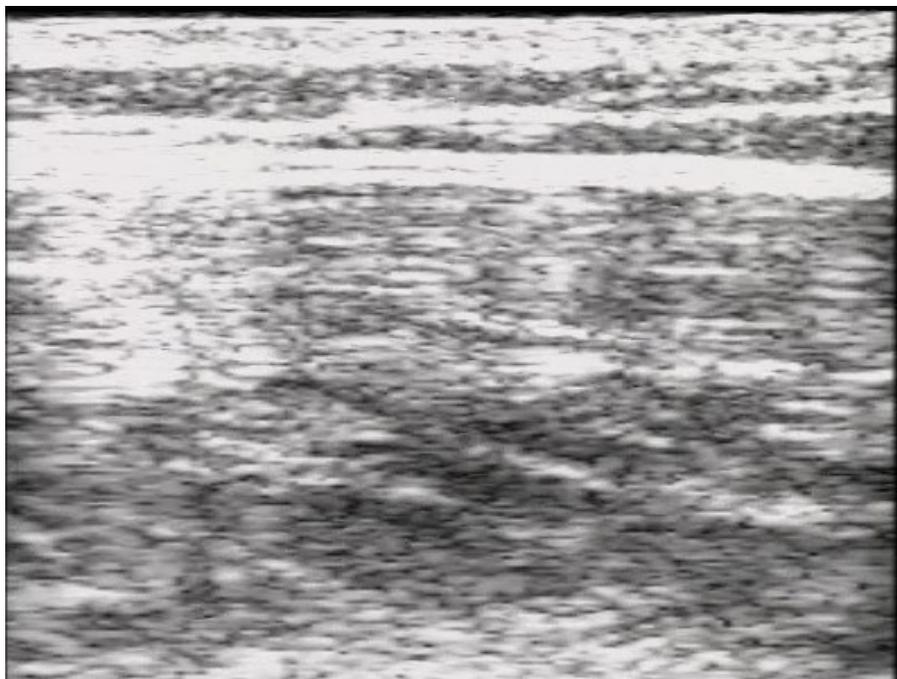
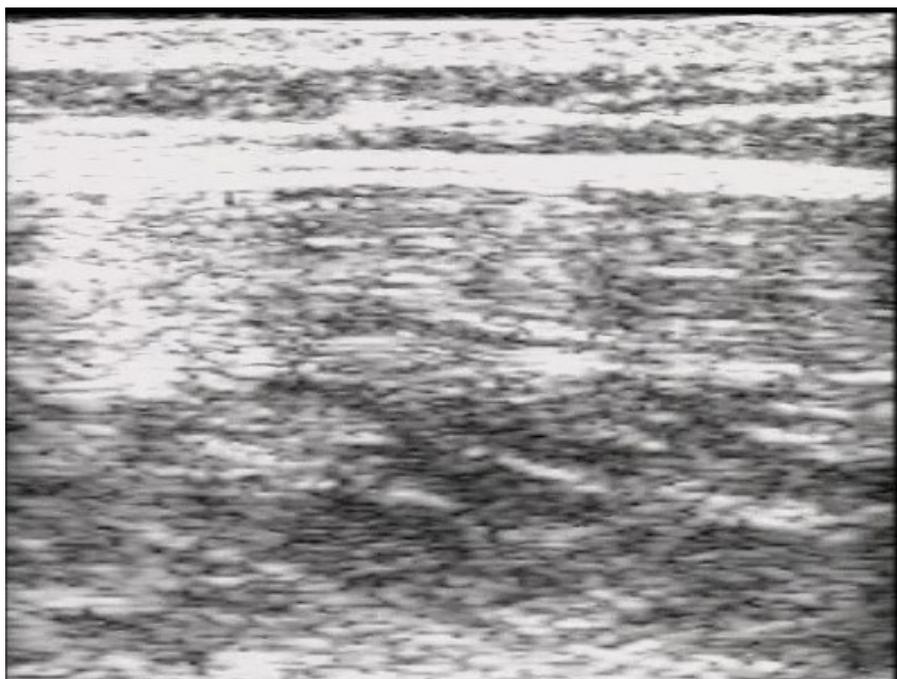


Figure 4.37: Motion Vectors Overlaid over Frame 215 which Represent the Motion between Frames 215 and 216.

Figure 4.38 shows a case of translational motion between frames 238 and 239 while the muscle in the upper leg is in the contraction process. Figure 4.39 shows the motion vectors for each macro-block between frames 238 and 239 which are shown in Figure 4.38.



(a)



(b)

Figure 4.38: (a) Frame 238 of the Upper Leg Experiment. (b) Frame 239 of the Upper Leg Experiment.

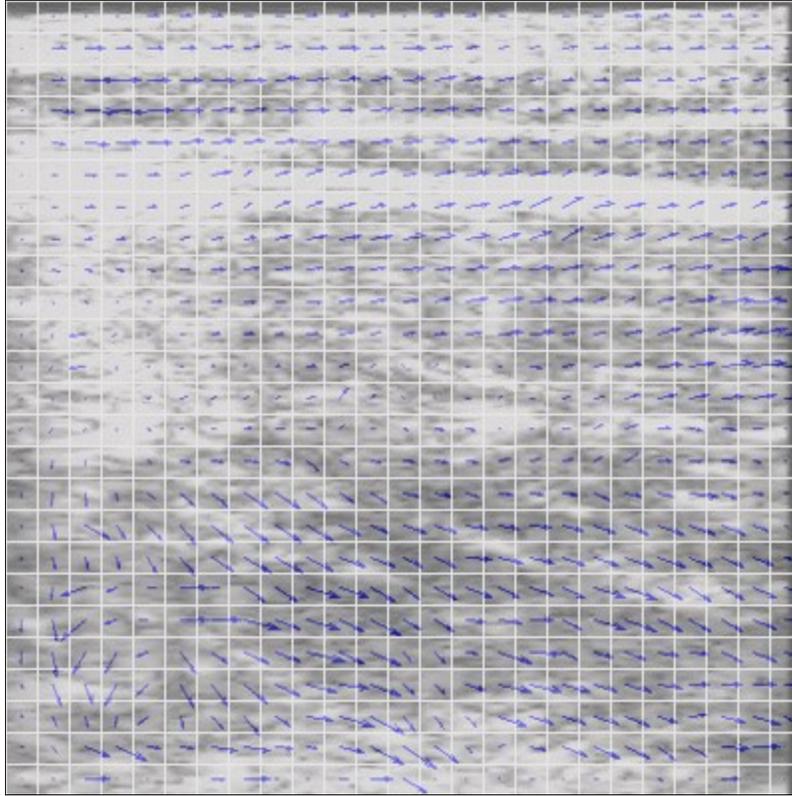


Figure 4.39: Motion Vectors Overlaid over frame 238 which Represent the Motion between Frames 238 and 239.

Looking at Figure 4.39, it can be seen that not all the motion vectors are in the direction of motion of the probe. It can be seen that most of the motion vectors at the bottom of the frame are approaching the bottom right corner of the frame. It also can be seen that some of the motion vectors at the top of the frame are approaching the top right corner of the frame. The fact that the motion vectors are going in these directions is because the muscle is in the contraction process and thus its size is becoming bigger making the top part of the muscle to expand upwards relative to the frame and the bottom part of the muscle to expand downward relative to the frame. Thus it can be concluded that the modified block-matching algorithm is able to track the motion between frames even in the presence of muscle contraction.

After the modified block-matching approach was performed on the frames, the global motion vectors were seen through the frames. As a result, only the parts of the frames that are in the direction of the global motion vectors were taken and were passed through the subpixel phase

correlation algorithm described in this thesis. Figure 4.40 shows the motion estimation between the frames along the Y direction using the subpixel phase correlation algorithm.

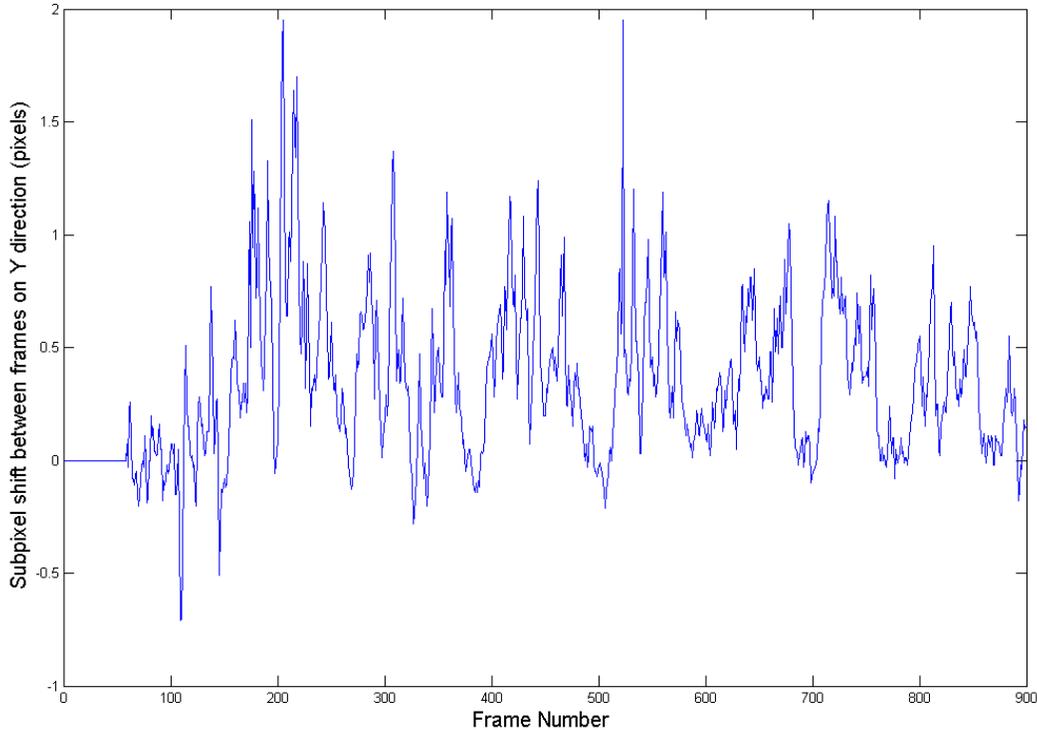


Figure 4.40: Motion Estimation between Frames on the Y Direction using the Subpixel Phase Correlation Motion Estimation.

Looking at Figure 4.40, it can be seen that the maximum shift between frames is about 1.95 pixels and the minimum was about -0.71 pixels. The negative shift could have occurred due to subject under testing pushing the probe back while contracting his muscle in the upper leg area. After the motion estimation was obtained with subpixel accuracy, the original frames were registered using the information obtained from the subpixel motion estimation. All the frames were registered with respect to frame number 1 of the experiment. Finally, the registered frames were all averaged by summing the registered frames and then dividing by the total number of frames. Figure 4.41 shows the extended field-of-view obtained from averaging the registered frames.

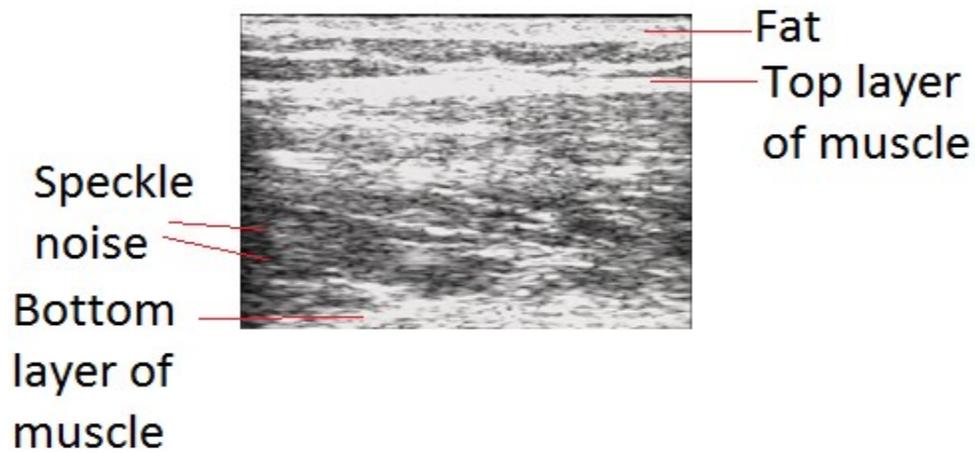


Figure 4.41: Average of the Registered 900 Frames of the Experiment Performed on the Upper Leg Region (Size of 708 x 428).

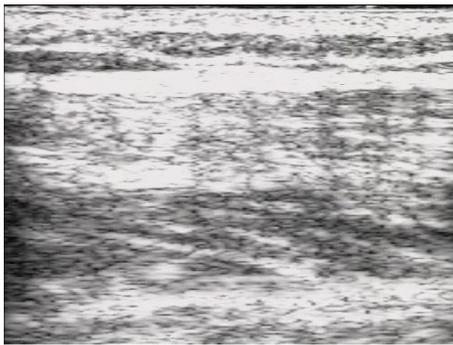
Looking at this extended field-of-view frame and looking at the separate 900 frames that were used to make it, it was observed that the major tissues that were seen in the separate frames are also seen in the extended field-of-view frame, which as described earlier were fat and muscle tissues. It was also observed that the overall noise was reduced in the extended field-of-view frame. Again in this experiment, to our knowledge, in obtaining this frame, there was not any cost of losing any important features of the images that were seen in the separate 900 frames. However, this cannot be concluded for sure without validating the accuracy of the motion estimation that was used to obtain the extended field-of-view.

4.6. In vivo Experiment to Observe the Effect of Probe compression

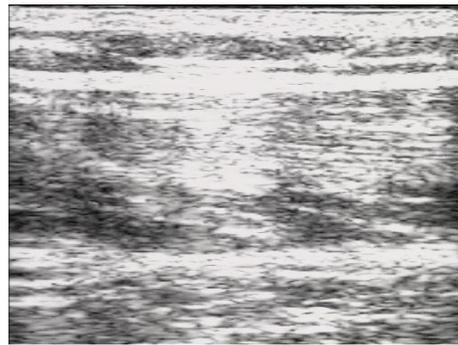
An experiment was conducted to see how well the modified block matching algorithms developed in this thesis work in the presence of probe compression. In this experiment the US probe moved manually along the upper part of the leg exactly as was shown in the previous experiment in Figure 4.32. However, the probe was moved from a point (a) which is 31 cm above the knee to a point (b) which is 21.5 cm above the knee and thus the total motion was 9.5 cm along the lateral direction. However, in this experiment, during the period the motion was recorded, the probe was pressed constantly against the body as the probe was in motion to observe the effects of probe compression. In this experiment, the video capture system used in the previous experiment was used. The total period that was recorded was 19.3 seconds giving an average probe motion speed of 0.49 cm/second. The size of each frame obtained using the video capture system in this experiment was 394 x 398. Figure 4.42 shows some of the outputs obtained in this experiment. Fat, muscle and speckle noise can be observed in this figure. It shows the period where there was no compression effect as can be seen in (a) then the period half way through the compression process in (b) followed by the period of full compression effect of the probe in (c) and then back to relaxation mode in (d).



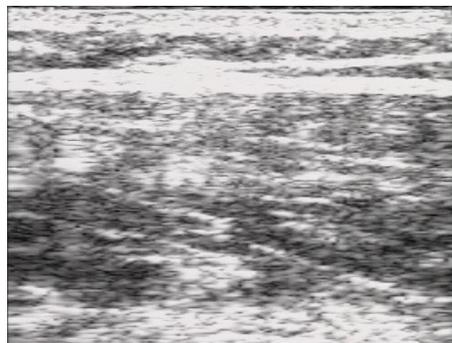
(a) Frame 90



(b) Frame 105



(c) Frame 119



(d) Frame 130

Figure 4.42 : Some Output Frames obtained During the Experiment of Probe Compression. (a) Output while Muscle is Relaxed with no Compression. (b) Output During the Period Half Way through the Compression Process. (c) Output During the Period of Full Compression Effect of the Probe. (d) Output During back to relaxation mode.

The total number of frames obtained in this experiment was 581 frames. Again in this experiment, the first step that was done was calculating the integer shift phase correlation between frames in order to be able to identify a proper search area, to search for the best matched-block using the modified block-matching algorithm search. Figure 4.43 shows the output obtained from the integer based phase correlation. Looking at Figure 4.43 it can be observed that the maximum motion between consecutive frames was about 4 pixels along the direction of motion. As a result, the search area set to look for the best matched block was set to 4 pixels plus 1 pixel which is equal to 5 pixels.

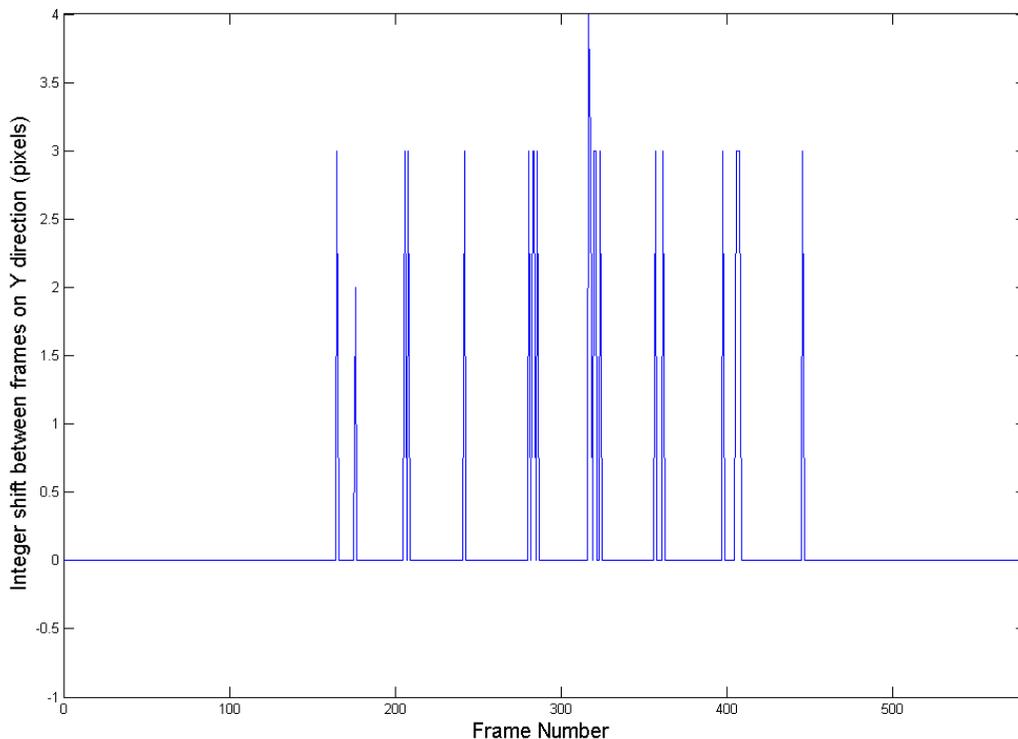
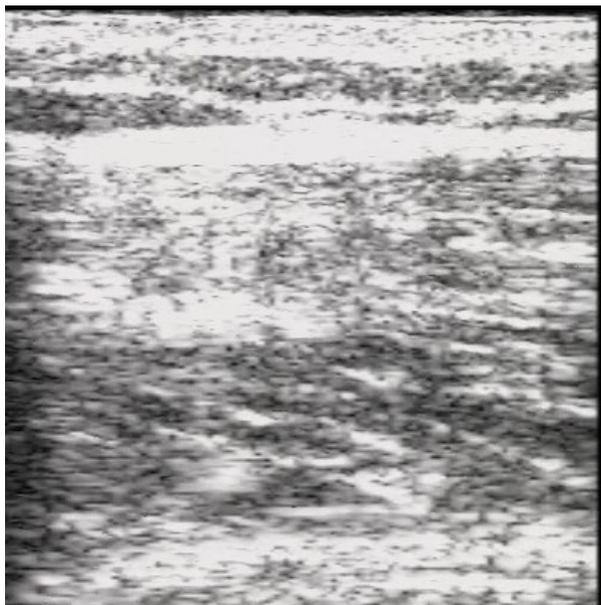


Figure 4.43: Motion Estimation between Hip region Frames on the Y Direction using Integer Phase Correlation.

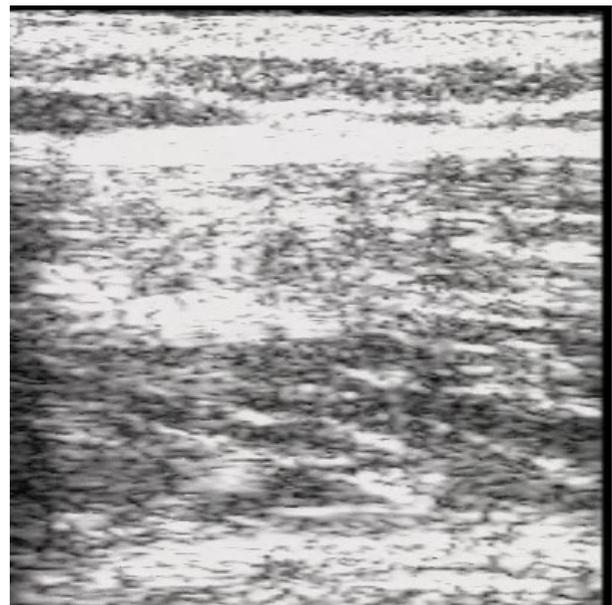
Once again, all the frames were divided into 16 x 16 macro-blocks and then each macro-block was searched in the search area specified in the frame preceding the current frame being searched. Since the frames obtained in this experiment are of the size 394 x 398, just as they

were in the previous experiment, the frames were zero padded to obtain a size of 400 x 400 so that the frames are made of a multiple of the macro-block size chosen, which is 16 x 16.

After the frames were divided into 16 x 16 macro-blocks, each macro-block was searched in the search area defined in the reference frame. The motion vectors obtained by the modified block matching search shows that the modified block matching search was able to track the motion between frames even in the presence of probe compression. The effect that probe compression had on the tissues being imaged was, a change in the size of the muscle underlying the imaged area. Thus, the effects observed are similar to the effects seen in the experiment that was designed to test muscle contraction. Figure 4.44 shows a case of translational motion between frames 94 and 95 without the presence of probe compression.



(a) Frame 94



(b) Frame 95

Figure 4.44: (a) Frame 94 of the Upper Leg Experiment to Test Probe Compression Effect. (b) Frame 95 of the Upper Leg Experiment to Test Probe Compression Effect.

Figure 4.45 shows the motion vectors of each macro-block overlaid over frame 94 which shows the motion between frames 94 and 95. Looking at Figure 4.45 it can be seen that the motion vectors obtained are along the direction of the motion of the probe and thus only a translational motion is observed in this Figure.

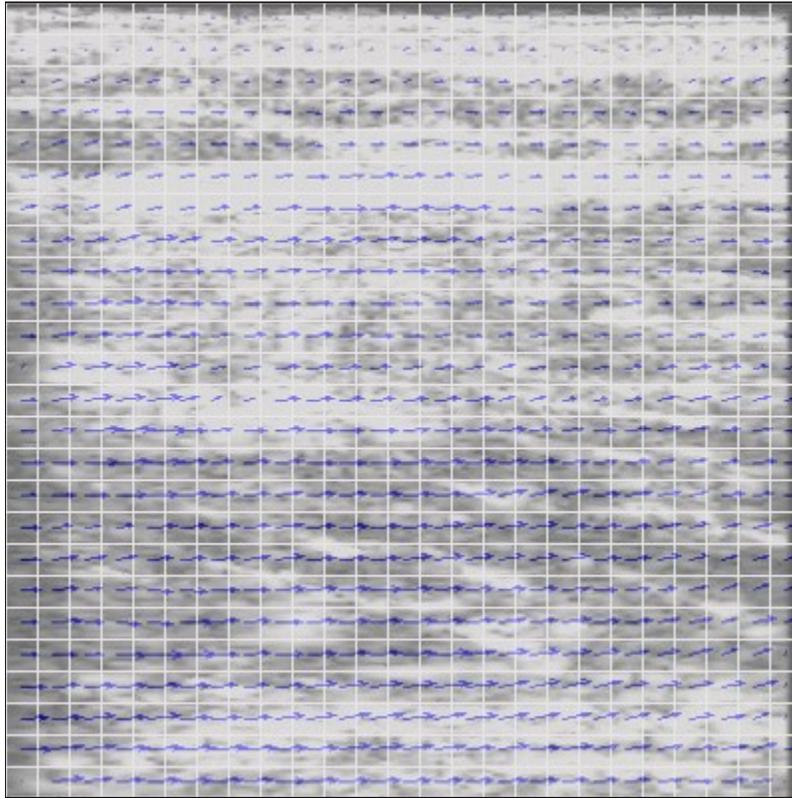
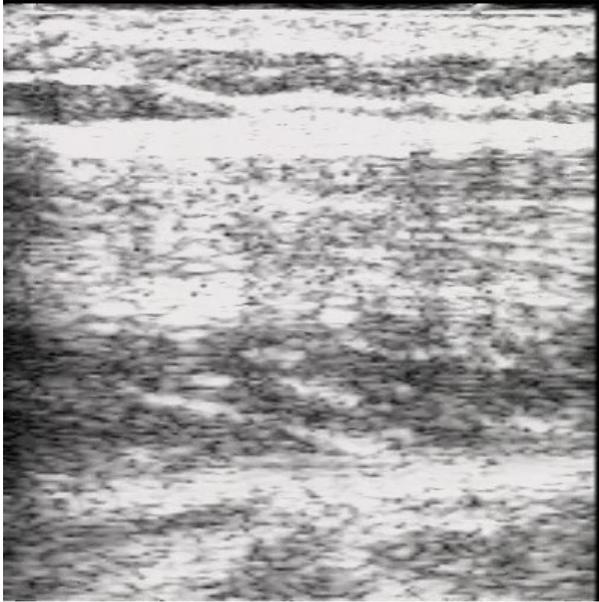
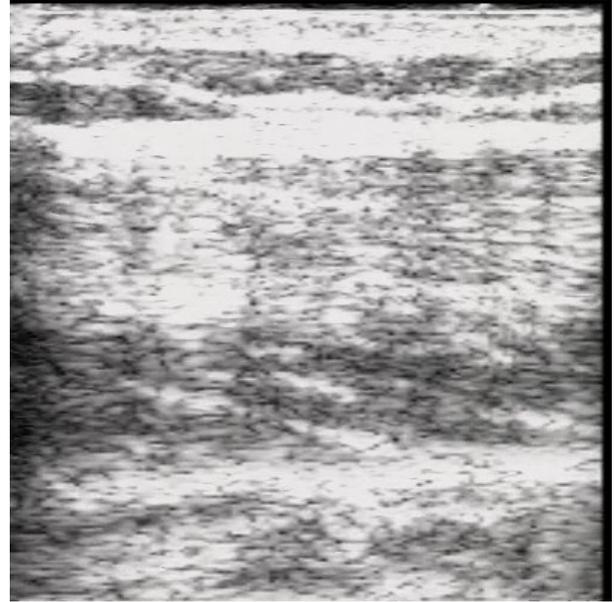


Figure 4.45: Motion Vectors Overlaid over Frame 94 which Represent the Motion Vectors between Frames 94 and 95.

Figure 4.46 shows a case of motion in the presence of probe compression. This Figure shows frames 108 and 109 of the experiment where compression effects can be observed from the muscle becoming smaller as we press more with the probe against the body. Figure 4.47 shows the motion vectors of each macro-block overlaid over frame 108 which shows the motion between frames 108 and 109 while probe compression exists.



(a) Frame 108



(b) Frame 109

Figure 4.46: (a) Frame 108 of the Upper Leg Experiment to Test Probe Compression Effect. (b) Frame 109 of the Upper Leg Experiment to Test Probe Compression Effect.

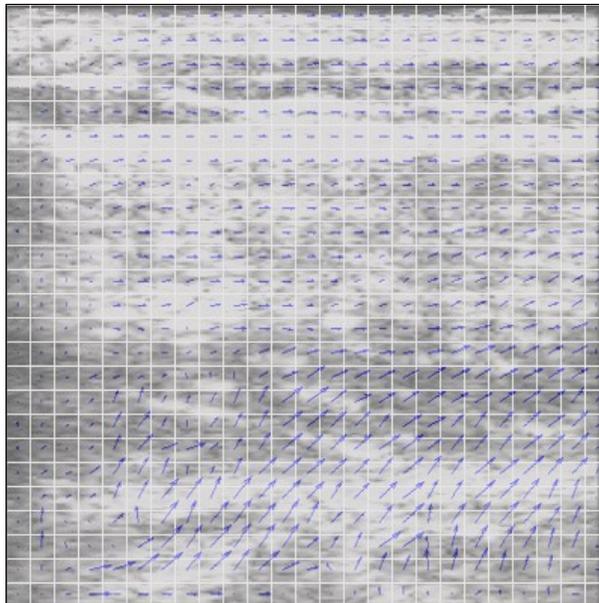


Figure 4.47: Motion Vectors Overlaid over Frame 108 which Represent the Motion Vectors between Frames 108 and 109.

Looking at Figure 4.47, it can be observed that the motion vectors at the bottom of the frame are approaching the top right corner of the frame. This is because the bottom part of the frame represents the bottom part of the muscle in the upper leg area, which during probe compression, moved upward toward the upper part of the muscle and thus the muscle size reduced. As a result, it can be concluded that the modified block matching search works well in cases where probe compression exists.

After the modified block matching approach was performed on the frames. The global motion vectors were seen through the frames. As a result, only the parts of the frames that are in the direction of the global motion vectors were taken and were passed through the subpixel phase correlation algorithm described in this thesis. Figure 4.48 shows the motion estimation between the frames along the Y direction using the subpixel phase correlation algorithm.

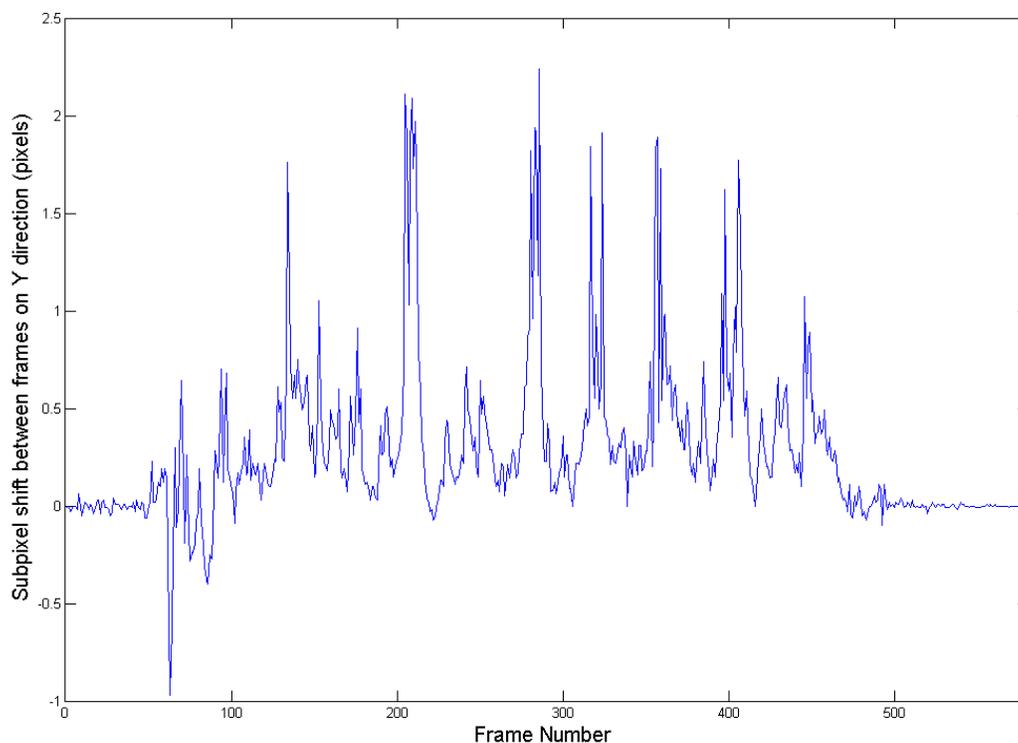


Figure 4.48: Motion Estimation between Frames on the Y Direction using the Subpixel Phase Correlation Motion Estimation.

Finally, the frames were registered and then averaged. Figure 4.49 shows the extended field-of-view obtained from averaging the registered frames. In this figure, fat and muscle that were seen in the separate frames are also seen in this extended field-of-view frame. It can also be seen that the overall noise was reduced. It is believed that there was no cost of losing important features in obtaining the extended field-of-view frame but again this cannot be concluded for sure without validating the motion estimation obtained by the developed algorithms.

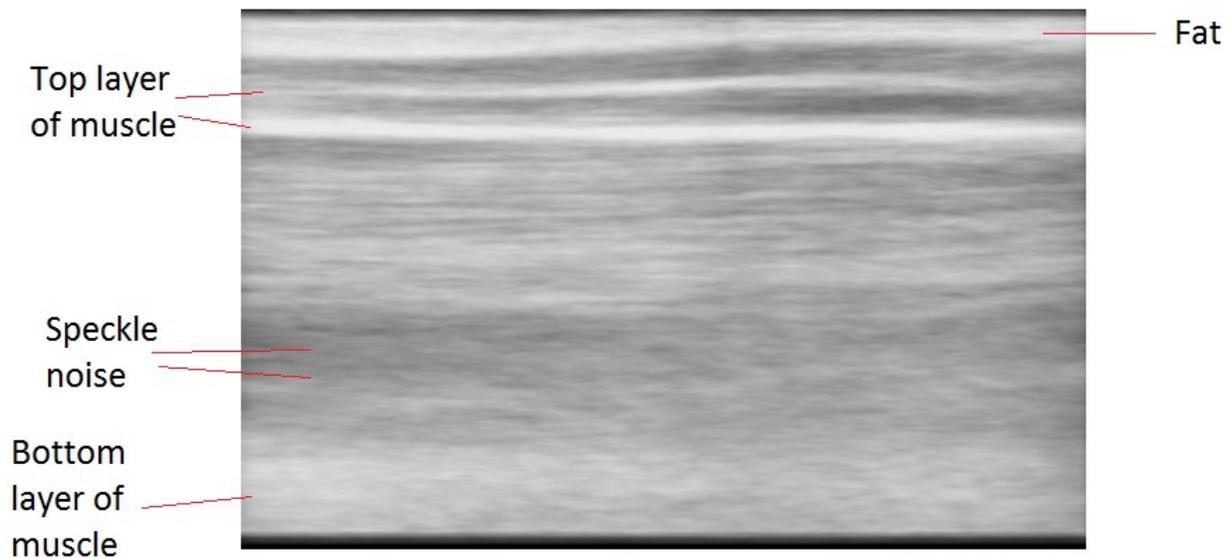


Figure 4.49: Average of the Registered 581 Frames of the Experiment Performed on the Upper Leg Region (Size of 559 x 434).

4.7. Summary and Discussion

Motion estimation between ultrasound images was performed using a modified block search algorithm and phase correlation. Block-matching was very useful in tracking the motion between the ultrasound frames since ultrasound images are characterized by having a lot of speckle noise which usually correlates between the images. Block-matching is an area-based search method. Therefore there was no need to identify different tissues in the images for tracking. As a result, the choice of a block-matching search was very useful in this thesis. An exhaustive search block based method, with some modification, was used because we are dealing with medical images and thus accuracy is required since any small error could cause wrong

information to be observed from the ultrasound images. Phase correlation was also used in this thesis. As mentioned earlier, phase correlation works perfectly in cases that have translational motion between the frames. When the ultrasound probe is moved in the lateral direction, which is the direction of interest in this thesis, a translational motion was obtained between the frames. This was shown in all the experiments performed in this research which included the phantom experiment and the in vivo experiments. US images were obtained using an ultrasound machine of the model Picus, ESAOTE Europe, Maastricht, Netherlands. This machine is accompanied by a linear probe and thus linear images were obtained. Images were first obtained on an US phantom that was designed to test the phase correlation algorithm developed in this thesis. Some of the output images obtained (eraser) were shown in Figure 4.3. Integer phase correlation was first applied to the images and then the subpixel phase correlation was applied on the images. The results that show the motion obtained are shown in Figures 4.4-4.7. From the subpixel phase correlation motion estimation results obtained, all the frames of the eraser were registered to frame 1 and then the average of the total number of registered images was taken. Figure 4.9 showed the output of averaging the total number of the registered frames which we called the extended field-of-view frame. In this output, it was observed that all the information that was seen in the 180 eraser frames was also seen in the extended field-of-view frame. This made us conclude that phase correlation was able to estimate the motion between consecutive frames accurately in the absence of noise since the designed ultrasound phantom was noise free. After the phantom experiment, an in vivo experiment on the experimenter's hip region was conducted to test phase correlation. Motion in this experiment was applied with a mechanical arm and thus motion was consistent. The frames obtained in this experiment had speckle noise. Figures 4.13-4.16 show the outputs obtained from the integer phase correlation and the subpixel phase correlation motion estimation. Figure 4.17 shows the extended field-of-view frame obtained. In this figure, all the important features that were seen in the separate images, which were used to make the extended field-of-view, were also seen in it. It also was observed that the overall noise was reduced. It is believed that there was no cost of losing important features in obtaining the extended field-of-view frame but this cannot be concluded for sure without validating the motion estimation obtained by the developed algorithms. To validate the results, it is suggested to use an ultrasound phantom that contains noise. Also, an electromagnetic sensor or some other type of sensors is suggested to be used for validation purposes. The other part of the research was to

improve the motion estimation between frames even more and thus a full search block-matching algorithm was used with some modifications. The developed algorithm was tested on frames obtained from conducting several experiments around the hip region and the upper leg region. The motion in the experiments was provided manually and thus motion between frames was not consistent. The first experiment was performed on the hip region as was described in section 4.4. Figure 4.26 shows the motion vectors between frames 80 and 81 and Figure 4.27 shows how the motion vectors that were believed to have false motion vectors were corrected for, using the modified block search algorithm by either assigning a zero motion vector to it or by excluding the whole block from the search. It is believed that some blocks of no motion were assigned a false motion vector because ultrasound images have a lot of dark regions and thus some blocks maybe assigned a motion vector for a macro-block with no motion other than its actual best matched block. Also it is believed that some macro-blocks were hard to track due to three main reasons. One of the reasons could be due to the fact that speckle noise in ultrasound images may decorrelate between the frames and thus motion of the macro-blocks that have decorrelation of speckle would be hard to track. Another reason could be due to probe compression. Probe compression against the skin could cause some of the tissues to change their actual shape on the boundaries of the muscles. If some of the tissues change their shape on the boundaries of muscles, correlation between the macro-blocks before and after the tissue changed its shape would be low and thus these macro-blocks will be hard to track. If we take a look at Figure 4.30, we can see the effect of probe compression. The motion on the X direction represents the change in depth between the frames which was obtained from probe compression against the body. From Figure 4.30 it can be seen that the region between frames 62 to 95 was the region with the maximum change of depth which reached about 0.5 pixels at a maximum. Also it was observed that most of the macro-blocks that were hard to track due to having a correlation value of more than 0.14 were in the range between frames 62 to 95. As a result, it can be concluded that, the reason why some of the macro-blocks were hard to track was mainly due to tissue changing its shape due to compression of the probe or the fast motion of the probe. Figure 4.29 shows the PSNR obtained using the information from the original block-matching search and from the modified block-matching search. It was seen in Figure 4.29 that the modified full search block-matching has a relatively higher PSNR in the frames that contained motion compared to the regular full search block-matching. It was also observed that the PNSR

for both the modified block-matching search and the regular block-matching search are the same at areas of no motion. After the modified block-matching search was applied on all the frames, macro-blocks that had motion vectors in the direction of the global motion direction were taken to apply the subpixel phase correlation search on them. By doing so, it is expected to obtain more accurate results for the motion between frames. Finally, all the frames were registered with respect to frame 1 and then averaging of the registered frames was performed to obtain an extended field-of-view frame, shown in Figure 4.31. This frame shows history from the total number of frames of the experiment with noise reduction since averaging the registered frames gave results similar to passing the images through a low pass filter. By looking at the total number of frames and by looking at the extended field-of-view frame obtained, it can be observed that the fat and muscle in the separate frames are all seen in the extended field-of-view frame. It also can be seen that there was no cost of losing important features in the process of obtaining the extended field-of-view frame but again this cannot be concluded for sure without validating the motion estimation obtained by the developed algorithms.

In this thesis, it was also observed that the modified block-matching algorithm developed here works well in the detection of the motion between frames in the presence of muscle contraction, as was shown in the experiment conducted along the upper leg area which was described in section 4.5. Also, it was observed that the modified block-matching search algorithm works well in detecting the motion between frames in the presence of probe compression since in normal cases, probe compression causes the muscles under compression to become smaller in size by bringing the upper and lower parts of the muscle imaged closer to each other without actually changing the shape of the upper and lower parts of the muscles. As a result, it can be concluded that the modified block-matching is able to track the motion between blocks if an individual macro-block is skewed or rotated within the search area used to find its best matched block.

From the all experiment observed, an extended field-of-view was obtained which is the goal of this thesis. However, it can be observed that the extended field-of-view frame obtained from the images that were acquired using ART.LAB was much clearer than the one obtained from the images that were acquired using the video capture. This is because the acquired ultrasound images using ART.LAB are much clearer than the ones acquired using the video capture. Using ART.LAB software, the images go through some filtering stages before displaying them on the

monitor. However, using the video capture software, the images are shown without any stages of filtering. Therefore, the video capture images were noisier.

The developed algorithms were tested on images obtained on the author's hip region and upper leg region. This limitation did not allow us to compare the extended field-of-view on different subjects. To validate the results obtained, it is suggested that a larger number of subjects participate in the experiments. It is also suggested that different subjects with varying ages are chosen to see what could be the differences in the extended field-of-view. It is also suggested to apply the developed algorithms on images of subjects with FAI to see if we can differentiate between them and a healthy subject, using our method for obtaining the extended field-of-view frame. Also in this thesis, we looked for tracking the motion along the Y direction which as mentioned earlier is the direction parallel to the scan. Moving the probe in this direction shows a translational motion between the consecutive frames. However, in reality we would want to track the motion of the probe when it is moved in any direction. Thus, other algorithms should be implemented in order to cover the full areas of probe's motion. When the ultrasonographer images the hip region to diagnose FAI, the ultrasound probe would be moved in the lateral, axial and elevational directions. Therefore, in obtaining an extended field-of-view frame, motion estimation between the consecutive frames should cover all these directions.

In conclusion, a comparison was made between each extended field-of-view and their separate frames by eye. It was observed that all the important features that were seen in the separate frames were also seen in their extended field-of-view frame with an overall reduction in noise. It was believed that there was no cost of losing any important features in the process of obtaining the extended field-of-view. However, this cannot be guaranteed for sure without either validating the results by a speckle phantom or some type of sensor.

Chapter 5

Conclusions and Future Work

5.1. Summary of Contributions

This thesis presented a method for tracking the motion of the ultrasound probe through image processing techniques without the help of any external sensors. Modified block based motion estimation was used along with a subpixel phase correlation motion estimation to estimate the motion between 2D B-mode ultrasound images. The images were then registered using the information obtained from the motion estimation algorithms used. Finally an extended-field-of-view output was obtained.

5.1.1. Block-Based Motion Estimation

A block based motion estimation algorithm was used in this research. In this method, the frames were divided into macro-blocks of the same size then each macro-block was searched for its best matched block in the frame before it which is referred to as the reference frame. In this thesis, a full search block matching search was used where the best matched block was searched for in every location in the search area defined to look for the best matched block. The MAD cost function was used to find the best matched block in the search area defined. Additionally, we developed a method to detect blocks that contains no motion which could be assigned a false motion vector. This was done by calculating the standard deviation of each macro-block; if the

standard deviation for a given macro-block was zero, a zero motion vector is assigned to that macro-block without looking for its best matched block. Also, macro-blocks that were hard to find its motion were excluded from the search. It was concluded that macro-blocks that had a minimum cost function (MAD) of 0.14 or higher to not have high correlation with its best matched block and thus these macro-blocks were eliminated from the search. It is important to exclude these macro-blocks from the search since they could affect the accuracy of the registration. This method was combined with a Fourier method to improve the accuracy of registration even more.

5.1.2. Phase Correlation Motion Estimation

To improve the accuracy of estimating the motion, phase correlation motion estimation was used. Phase correlation gives more accurate results in estimating the motion between frames with less computational time since it measures the motion between frames directly from their phases. We started off by first finding the integer motion between frames using phase correlation which was shown as a peak located at some point at the x and y directions. After that, subpixel phase correlation was estimated following a method developed by Guizar-Sicairos et al. [57] which estimate the motion between frames with low computational time and thus this method can be used in real time. This method was used along with the modified block-matching search by passing the areas that were observed to have motion vectors in the direction of motion of the probe. Doing so helped in improving the accuracy of motion estimation.

5.1.3. Image Registration and Averaging

After the motion estimation was obtained with high accuracy using the algorithms used and developed in this thesis, all the frames were registered with respect to one frame. Results were validated through a phantom experiment which had known materials such as an eraser and through in vivo experiments on the hip area and the upper leg area. The translational parameters on the X and Y directions found by phase correlation were used to register the images with respect to one frame. All the frames were translated by the X and Y values obtained from the motion estimation with respect to frame 1 of each experiment. The last step after translating the frames was averaging the registered images. Averaging was expected to reduce the noise since averaging works like a low pass filter. Averaging the ultrasound images was a good choice since

in ultrasound there are areas with noise that do not correlate between frames and thus averaging the registered images was reduced the noise. Also, from averaging the frames, an extended-field-of-view frame was obtained which contained information from the several frames. The accuracy of the registration through the subpixel phase correlation by itself was seen through the phantom experiment and the in vivo experiment when the probe moved by a mechanical arm. The extended-field-of-view for each of these experiments showed that the registration was done precisely and noise was reduced compared to looking at the total number of the frames. Also, the registration accuracy for the modified block-matching motion estimation and the subpixel phase correlation together was seen through other in vivo experiments on the hip region and the upper leg region but in these experiments the probe motion was provided manually. The extended-field-of-view output obtained from these experiments showed how accurate the performed registration was and it showed how the noise was reduced compared to the total number of frames on hand.

5.2. Future Work

The following points show some of the suggestions for future work:

- Future improvements for speckle tracking can be obtained. It was observed that some macro-blocks do not correlate between consecutive frames. In this thesis we tried to explain the reasons that could be behind this issue. However, we did not find a way to track these macro-blocks. Further improvement can be made by figuring a way to track the motion of these macro-blocks. If these macro-blocks can be tracked somehow, the accuracy of block-matching motion estimation would further improve. Thus, motion estimation in ultrasound with a sensor-less method would be favored more.
- Further improvement in block-matching can be obtained by extracting the features in each block and then identify each element before trying to search for its motion. This way a feature based method along with an intensity based method will be combined to perform the search for the search for the best matched block. Doing so, it is believed that the registration accuracy could even improve more.
- The motion tracking algorithms developed in this thesis will only work in obtaining an extended-field-of-view if the probe moved in the lateral direction. However, in ideal

cases we would like to obtain an extended-field-of-view when the probe is moved in any direction including the elevational direction. As a result, a motion estimation algorithm needs to be developed to track the motion on the elevational direction and then this algorithm will be joined with the work done in this thesis to obtain an extended-field-of-view when the probe is moved in any direction.

- The in-vivo experiments in this thesis were all performed on the author's hip region and upper leg region. It is suggested that different subjects with varying ages are chosen to participate in the experiments to see what could be the differences in the extended field-of-view. It is also suggested to apply the developed algorithms on images of subjects with FAI to see if we can differentiate between them and a healthy subject, using our method for obtaining the extended field-of-view frame.
- For the presented work in this thesis, the field-of-view of 2D linear probes was extended. However, the extended-of-view was obtained in 2D. This work can be extended by obtaining an extended-field-of-view in 3D domain. This can be done through speckle tracking as was seen in the literature.
- The work done was able to obtain an extended field-of-view frame if the probe moved in the lateral direction; the direction parallel to the scan line. However, in order to know how well our method works, the results can be validated using an electromagnetic sensor. Ideally the main goal of our work was to develop a sensor-less method for tracking the motion of the probe. However, the use of an electromagnetic sensor for the purpose of validating the results should not be a problem.

References:

- [1] S. Lerch, A. Kasperczyk, J. Warnecke, and T. berndt, "Evaluation of Cam-type femoroacetabular impingement by ultrasound," *Int Orthop*, vol. 37, no. 5, pp. 783–788, 2013.
- [2] F. M. Buck, J. Hodler, M. Zanetti, C. Dora, and C. W. A. Pfirrmann, "Ultrasound for the evaluation of femoroacetabular impingement of the cam type: Diagnostic performance of qualitative criteria and alpha angle measurements," *Eur Radiol*, vol. 21, pp. 167–175, 2010.
- [3] G. S. Kuhlman and B. G. Domb, "Hip impingement: identifying and treating a common cause of hip pain," *Am Fam Physician*, vol. 80, no. 12, pp. 1429–1434, 2009.
- [4] J. Meunier, "Tissue motion assessment from 3D echographic speckle tracking," *Phys. Med. Biol*, vol. 43, no. 5, pp. 1241–1254, 1998.
- [5] S. H. Chan, D. T. Vo, and T. Q. Nguyen, "Subpixel Motion Estimation Without Interpolation," in *Acoustics Speech and Signal Processing*, Dallas, TX, 2010, pp. 722–725.
- [6] A. Barjatya, "Block Matching Algorithms For Motion Estimation," Utah State University, Technical Report, 2004.
- [7] Y. Liang, "Phase Correlation Motion Estimation," Final Project EE 392J, Stanford University, Winter Quarter 2000, pp. 1-9.
- [8] F. W. Kremkau, *Diagnostic Ultrasound: Principles and Instruments*, Seventh ed., Missouri: Saunders Elsevier, 2006.
- [9] R. J. Bartrum and H. Crow, *Real-time Ultrasound, A Manual for Physicians and Technical Personnel*, Second edition. W.B. Saunders Company, 1983.
- [10] W. Moritz, A. Pearlman, D. Cabe, 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*, Vol. 30, no. 8, pp. 482-492, Aug. 1983.

- [11] D. King, D. Jr, and M. Shao, "Three-dimensional spatial registration and interactive display of position and orientation of real-time ultrasound images," *Journal of Ultrasound in Medicine*, Vol. 9, no. 9, pp. 525-532, 1990.
- [12] 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*, vol. 25, no. 5, pp. 745-753, 1999.
- [13] Y. Tamaki, Y. Sato, M. Nakamoto, T. Sasama, I. Sakita, M. Sekimoto, M. Ohue, N. Tomita, S. Tamura, and M. Monden, "Intraoperative navigation for breast cancer surgery using 3-D ultrasound images," *Computer Aided Surgery*, Vol.4, no. 1, pp. 37-44, 1999.
- [14] W. Zhang, R. Rohling, and D. Pai, "Surface extraction with a three-dimensional freehand ultrasound system," *Ultrasound in Medicine and Biology*, Vol. 30, no. 11, pp. 1461-1473, 2004.
- [15] F. Lindseth, G. Tangen, T. Lang, and J. Bang, "Probe calibration for freehand 3-D ultrasound," *Ultrasound in Medicine and Biology*, vol. 29, no. 11, pp. 1607-1623, 2003.
- [16] M. Nixon, B. McCallum, W. Fright, and N. Price, "The effects of metals and interfering fields on electromagnetic trackers," *Presence: Teleoperators and Virtual Environments*, vol. 7, no. 2, pp.204-218, 1998.
- [17] C. Burckhardt, "Speckle in Ultrasound B-Mode Scans," *IEEE Trans. Sonics Ultrason.*, vol. SU-25, no. 1, pp. 1-6, Jan. 1978.
- [18] L. Weng, A. Tirumalai, C. Lowery, L. Nock, D. Gustafson, P. Behren, and J. Kim, "US extended-field-of-view imaging technology," *Radiology*, vol. 203, no. 3, pp. 877-880, 1997.
- [19] A. Tirumalai, L. Weng, A. Grassmann, M. Li, S. Marquis, P. Sutcliffe, D. Gustafson, J. Kim, C. Basoglu, T. Winter, and Y. Kim, "New ultrasound image display with extended field of view," *In Yongmin Kim, editor, Proceedings of SPIE Medical Imaging 1997*, vol. 3031, pp. 409-419.
- [20] L. N. Bohs, B. J. Geiman, M. E. Anderson, S. C. Gebhart, and G. E. Trahey, "Speckle tracking for multidimensional flow estimation," *Ultrasonics*, vol. 38, no. 1-8, pp. 369-375, Mar. 2000.

- [21] B. Geiman, L. Bohs, M. Anderson, S. Breit, and G. Trahey, "A novel interpolation strategy for estimating subsample speckle motion," *Physics in Medicine and Biology*, vol. 45, no. 6, pp. 1541-1552, 2000.
- [22] R. Prager, A. Gee, G. Treece, C. Cash, and L. Berman, "Sensorless freehand 3-D ultrasound using regression of the echo intensity," *Ultrasound in Medicine and Biology*, vol. 29, no. 3, pp. 437-446, 2003.
- [23] R. Housden, A. Gee, G. Treece, and R. Prager, "Subsample interpolation strategies for sensorless freehand 3-D ultrasound," *Ultrasound in Medicine and Biology*, vol. 32, no. 12, pp. 1897-1904, 2006.
- [24] R. Shekhar and V. Zagrodsky, "Mutual information-based rigid and nonrigid registration of ultrasound volumes," *IEEE Trans. Medical Imaging*, vol. 21, no. 1, pp. 9-22, Jan. 2002.
- [25] R. N. Rohling, A. H. Gee, and L. Berman, "Automatic registration of 3-D ultrasound images," *Ultrasound in Medicine and Biology*, vol. 24, no. 6, pp. 841-854, Jul. 1998.
- [26] J. Schers, J. Troccaz, V. Daanen, C. Fouard, C. Plaskos, and P. Kilian, "3D/4D ultrasound registration of bone," *In Proceedings of IEEE International Ultrasonics Symposium*, pp. 2519-2522, Oct. 2007.
- [27] G. Xiao, M. Brad, and A. Noble, "Nonrigid Registration of 3-D Free-Hand Ultrasound Images of the Breast," *IEEE Trans. on Medical Imaging*, Vol. 21, no. 4, pp. 405-412, April 2002.
- [28] T. Poon, and R. Rohling, "Three-Dimensional Extended Field-Of-View Ultrasound," *Ultrasound in Medicine and Biology*, Vol. 32, no. 3, pp. 357-369, 2006.
- [29] D. Ni, Y. Qu, X. Yang, Y.P. Chui, T. Wong, S. Ho, and P. Heng, "Volumetric ultrasound panorama based on 3D SIFT," *In Medical Image Computing and Computer-Assisted Intervention—MICCAI 2008*, pp. 52-60, 2008.
- [30] A. Groves, and R. Rohling, "Two-Dimensional Spatial Compounding with Warping," *Ultrasound in Medicine and Biology*, Vol. 30, no. 7, pp. 929-942, 2004.

- [31] Z. Xu, C. Tan, L. Wang, and D. C. Liu, "Ultrasound Image Registration Based on Polar Transform and Block Matching," in *5th International Conference, Bioinformatics and Biomedical Engineering*, Wuhan, 2011, pp. 1–4.
- [32] Z. L. Sandoval and J.-L. Dillenseger, "Evaluation of computed tomography to ultrasound 2D image registration for atrial fibrillation treatment," in *Computing in Cardiology Conference (CinC)*, Zaragoza, 2013, pp. 245–248.
- [33] S. Zheng, Q. Huang, M. Lu, L. Jin, and T. Wang, "Rapid Image Registration for Extended-Field-of-View Ultrasound," in *4th International conf., Bioinformatics and Biomedical Engineering*, 2010, pp. 1–4.
- [34] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 997–1000, Oct. 2003.
- [35] H. S. Stone, M. Orchard, E.-C. Chang, and S. Martucci, "A fast direct Fourier-based algorithm for subpixel registration of images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 10, pp. 2235–2243, Oct. 2001.
- [36] T. Ramaprabha and M. M. Sathik, "The role of Block matching algorithm in Stereo image Compression," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 2, no. 6, pp. 167–170, Jun. 2012.
- [37] M. G. Strintzis and I. Kokkinidis, "Maximum likelihood motion estimation in ultrasound image sequences," *IEEE Signal Processing Letters*, vol. 4, no. 6, pp. 156–157, June 1997.
- [38] S. Golemati, A. Sassano, M. J. Lever, A. A. Bharath, S. Dhanjil, and A. N. Nicolaidis, "Carotid artery wall motion estimated from b-mode ultrasound using region tracking and block matching," *Ultrasound in Medicine and Biology*, vol. 29, no. 3, pp. 387–399, Mar. 2003
- [39] A. Gastounioti, S. Golemati, J. Stoitsis, and K. S. Nikita, "Kalman-filter-based block matching for arterial wall motion estimation from B-mode ultrasound," in *IEEE International Conference on Imaging Systems and Techniques*, 2010, pp. 234–239.
- [40] H. Gao, F. Kremer, H. F. Choi, J.-U. Voigt, P. Claus, and J. D'hooge, "Left ventricular 2D flow pattern estimation of the heart by combining speckle tracking with Navier-Stokes based regularization," in *IEEE Ultrasonics Symposium*, 2010, pp. 1068–1071.

- [41] A. Basarab, P. Gueth, H. Liebgott, and P. Delachartre, "Two-dimensional least-squares estimation for motion tracking in ultrasound elastography," in *29th Annual International Conference, Engineering in Medicine and Biology Society*, 2007, pp. 2155–2158.
- [42] J. Lu and M. L. Liou, "A simple and efficient search algorithm for block-matching motion estimation," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 429–433, Apr. 1997.
- [43] S. Zhu and K.-K. Ma, "A new diamond search algorithm for fast block-matching motion estimation," *IEEE Trans. Image Processing*, vol. 9, no. 2, pp. 287–290, Feb. 2000.
- [44] Y. Nei and K.-K. Ma, "Adaptive rood pattern search for fast block-matching motion estimation," *IEEE Trans. Image Processing*, vol. 11, no. 12, pp. 1442–1449, Dec. 2002.
- [45] P. C. Shenolikar and S. P. Narote, "Different Approaches for Motion Estimation," in *Control, Automation, Communication and Energy Conservation*, 2009.
- [46] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, Feb. 1994.
- [47] D. Boukerroui, J. A. Noble, and M. Brady, "Velocity Estimation in Ultrasound Images: A Block Matching Approach," *Information Processing in Medical Imaging*, vol. 2732, pp. 586–598, 2003.
- [48] S. H. Lai and B. C. Vemuri, "Robust and efficient algorithms for optical flow computation," in *International Symposium on Computer Vision*, 1995, pp. 455–460.
- [49] B. K. Horn and B. G. Schunck, "'Determining optical flow': a retrospective," *Artificial Intelligence*, vol. 59, no. 1–2, pp. 81–87, Feb. 1993.
- [50] B. K. Horn and B. G. Schunck, "Determining Optical Flow," in *Techniques and Applications of Image Understanding*, 319, 1981, vol. 281, pp. 319–325.
- [51] C. Kuglin and D. Hines, "The phase correlation image alignment method," *Proc. Int. Conf. Cybernetics Society*, 1978, pp.163 -165.
- [52] A. Gilliam, J. Hossack, and S. T. Acton, "Freehand 3D Ultrasound Volume Reconstruction via Sub-Pixel Phase Correlation," in *IEEE Conference on Image Processing*, 2006, pp. 2537–2540.

- [53] W. Hoge, “A subspace identification extension to the phase correlation method [MRI application],” *IEEE Trans. Medical Imaging*, vol. 22, no. 2, pp. 277–280, Feb. 2003.
- [54] H. Foroosh, J. B. Zerubia, and M. Berthod, “Extension of phase correlation to subpixel registration,” *IEEE Trans. Image Processing*, vol. 11, no. 3, pp. 188–200, Mar. 2002.
- [55] S. Kumar, M. Biswas, and T. . Nguyen, “Efficient phase correlation motion estimation using approximate normalization,” in *Conference Record of the Thirty-Eighth Asilomar Signals, Systems and Computers*, 2004, vol. 2, pp. 1727–1730.
- [56] X. Zhang, N. Homma, M. Abe, N. Sugita, Y. Takai, and M. Yoshizawa, “Volume registration based on 3-D phase correlation for tumor motion estimation in 4-D CT,” in *35th Annual international Conference of the IEEE, Engineering in medicine and Biology*, 2013, pp. 5095–5098.
- [57] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, “Efficient subpixel image registration algorithms,” *Optical Letters*, vol. 33, no. 2, pp. 156–158, Jan. 2008.
- [58] J. I. Silver, “Development of Ultrasound Based Techniques for Measuring Skeletal Muscle Motion,” *Carleton University, Canada (Ottawa)*, 2009.