A Markov Random Field Based Approach to 3D Mosaicing and Registration Applied to Ultrasound Simulation

by

Jason Francis Kutarnia

Submitted to the Department of Electrical and Computer Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electical and Computer Engineering

at

WORCESTER POLYTECHNIC INSTITUTE

October 2014

Author ....
Department of Electrical and Computer Engineering
August 13, 2014

Certified by ....
Peder C. Pedersen
Professor of Electrical and Computer Engineering
Dissertation Supervisor

This doctoral dissertation has been examined by a committee as follows:

Dr. Sergey N. Makarov ....
Professor, Electrical and Computer Engineering

Dr. Michael A. Gennert ....
Director, Robotics Engineering Program

Dr. Stephen J. Glick ....
Professor of Radiology, Univ. of Mass. Med. School

A Markov Random Field Based Approach to 3D Mosaicing and Registration Applied to Ultrasound Simulation

by

Jason Francis Kutarnia

Submitted to the Department of Electrical and Computer Engineering
on August 13, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electical and Computer Engineering

Abstract

A novel Markov Random Field (MRF) based method for the mosaicing of 3D ultrasound volumes is presented in this dissertation. The motivation for this work is the production of training volumes for an affordable ultrasound simulator, which offers a low-cost/portable training solution for new users of diagnostic ultrasound, by providing the scanning experience essential for developing the necessary psycho-motor skills. It also has the potential for introducing ultrasound instruction into medical education curriculums. The interest in ultrasound training stems in part from the widespread adoption of point-of-care scanners, i.e. low cost portable ultrasound scanning systems in the medical community.

This work develops a novel approach for producing 3D composite image volumes and validates the approach using clinically acquired fetal images from the obstetrics department at the University of Massachusetts Medical School (UMMS). Results using the Visible Human Female dataset as well as an abdominal trauma phantom are also presented. The process is broken down into five distinct steps, which include individual 3D volume acquisition, rigid registration, calculation of a mosaicing function, group-wise non-rigid registration, and finally blending. Each of these steps, common in medical image processing, has been investigated in the context of ultrasound mosaicing and has resulted in improved algorithms. Rigid and non-rigid registration methods are analyzed in a probabilistic framework and their sensitivity to ultrasound shadowing artifacts is studied.

The group-wise non-rigid registration problem is initially formulated as a maximum likelihood estimation, where the joint probability density function is comprised of the partially overlapping ultrasound image volumes. This expression is simplified using a block-matching methodology and the resulting discrete registration energy is shown to be equivalent to a Markov Random Field. Graph based methods common in computer vision are then used for optimization, resulting in a set of transformations that bring the overlapping volumes into alignment. This optimization is parallelized using a fusion approach, where the registration problem is divided into 8 independent sub-problems whose solutions are fused together at the end of each iteration. This

method provided a speedup factor of 3.91 over the single threaded approach with no noticeable reduction in accuracy during our simulations. Furthermore, the registration problem is simplified by introducing a mosaicing function, which partitions the composite volume into regions filled with data from unique partially overlapping source volumes. This mosaicing functions attempts to minimize intensity and gradient differences between adjacent sources in the composite volume.

Experimental results to demonstrate the performance of the group-wise registration algorithm are also presented. This algorithm is initially tested on deformed abdominal image volumes generated using a finite element model of the Visible Human Female to show the accuracy of its calculated displacement fields. In addition, the algorithm is evaluated using real ultrasound data from an abdominal phantom. Finally, composite obstetrics image volumes are constructed using clinical scans of pregnant subjects, where fetal movement makes registration/mosaicing especially difficult.

Our solution to blending, which is the final step of the mosaicing process, is also discussed. The trainee will have a better experience if the volume boundaries are visually seamless, and this usually requires some blending prior to stitching. Also, regions of the volume where no data was collected during scanning should have an ultrasound-like appearance before being displayed in the simulator. This ensures the trainee's visual experience isn't degraded by unrealistic images. A discrete Poisson approach has been adapted to accomplish these tasks. Following this, we will describe how a 4D fetal heart image volume can be constructed from swept 2D ultrasound. A 4D probe, such as the Philips X6-1 xMATRIX Array, would make this task simpler as it can acquire 3D ultrasound volumes of the fetal heart in real-time; However, probes such as these aren't widespread yet.

Once the theory has been introduced, we will describe the clinical component of this dissertation. For the purpose of acquiring actual clinical ultrasound data, from which training datasets were produced, 11 pregnant subjects were scanned by experienced sonographers at the UMMS following an approved IRB protocol. First, we will discuss the software/hardware configuration that was used to conduct these scans, which included some custom mechanical design. With the data collected using this arrangement we generated seamless 3D fetal mosaics, that is, the training datasets, loaded them into our ultrasound training simulator, and then subsequently had them evaluated by the sonographers at the UMMS for accuracy. These mosaics were constructed from the raw scan data using the techniques previously introduced. Specific training objectives were established based on the input from our collaborators in the obstetrics sonography group. Important fetal measurements are reviewed, which form the basis for training in obstetrics ultrasound. Finally clinical images demonstrating the sonographer making fetal measurements in practice, which were acquired directly by the Philips iU22 ultrasound machine from one of our 11 subjects, are compared with screenshots of corresponding images produced by our simulator.

Dissertation Supervisor: Peder C. Pedersen
Title: Professor of Electrical and Computer Engineering

Acknowledgments

First and foremost I would like to express tremendous gratitude towards my advisor, Dr. Peder Pedersen, for giving me the opportunity to contribute to his ultrasound simulation research for all these years. Though our project proved more difficult than we initially imagined, he never lost faith in my ability to finish the work. Without his mentorship and support I would have never been able to complete my PhD. I have certainly learned a great deal while at WPI thanks to him.

I would like to thank our collaborators at the University of Massachusetts Medical School for their tireless efforts during the clinical portion of my dissertation. Dr. Petra Belady, an assistant professor in the Dept. of Obstetrics and Gynecology, assessed our training volumes, giving us invaluable feedback which greatly improved the system. Sonographers Denise Cascione and Kathleen Fitzpatrick handled all the patient scanning, sacrificing many hours of their own time, after work. The quality of the finished product speaks to their impressive clinical skills.

My lab partner Li Liu spent a substantial amount of time and effort using his software engineering expertise to produce an impressive ultrasound simulation system, of which my training volumes are a component. The quality of this work made my job much easier when it became time to test our clinical results.

I would also like to thank the Telemedicine and Advanced Technology Research Center (TATRC) for their funding during the course of my research. It enabled me to really focus on our project. Also, I am grateful to the Dept. of Electrical and Computer Engineering at WPI for giving me a teaching assistantship, which funded a portion of my time here and was a valuable learning experience by itself.

Finally, I would like to thank my parents and brother for their support during the long and difficult road to my PhD. They were ever-present when I needed help and always had encouraging words to say when graduation seemed so distant. I couldn't have done it without them.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Recently increased attention has been paid towards using high fidelity medical simulation as a training tool for clinicians, particularly in medical ultrasound, evidenced by a number of commercial simulators having been developed. Research suggests that traditional methods for medical education, in particular the Halstedian approach of see one, do one, teach one [6], are less effective than Simulation Based Medical Education (SBME) for the acquisition of a wide range of clinical skills. Advanced cardiac life support, laparoscopic surgery, and cardiac auscultation are just a few examples where SBME has been linked to improved patient care [46].

After completing medical school, graduates will participate in residency programs in order to further develop their knowledge base and decision making ability within a specific field of medicine. However, development of their technical skills, which will become the foundation of their profession, is lacking. In order to correct this deficiency researchers have proposed innovative high fidelity surgical simulation devices, ultimately having a positive impact on patient outcomes [69]. The controlled, low intensity practice environment provided by SBME promotes learning and is easy for students to access [29]. After analyzing more than 100 simulation studies comprised of all types of medical simulators, other researchers arrived at the same conclusion, stating that high-fidelity simulators are effective and should complement education in patient care settings [7]. A recent review article on simulation in obstetrics and gynecology [23] also advocates for SBME. Despite the mounting evidence that simulation based education can improve healthcare delivery, ultrasound simulators have yet to be integrated into most training curriculums. Even medical schools which include ultrasound rely on traditional teaching methods [30].

1.1 Motivation for improved ultrasound training

The interest in ultrasound training stems from its widespread adoption in the medical community, which has been expanding due to the availability of low cost portable scanning systems. The cost of the an ultrasound scanning system is much less than the cost of a scanner for other imaging modalities such as CT or MRI, and the industry has reached a point where most medical offices in the U.S. can afford a scanning system. The medical community is currently experiencing a growth in ultrasound based diagnostic procedures, which far outpaces the growth in other imaging modalities [77]. In the past large stand-alone ultrasound scanners were used by sonographers, but approximately 15 years ago the transition to point-of-care (POC) ultrasound began, which resulted in a variety of clinicians beginning to utilize lower cost, portable ultrasound scanners. For example in cardiology there was a 60% increase in the use of POC ultrasound between 2004 and 2009 and also during the same period there was a 28% increase in POC use by non-radiologist physicians [43]. One worrisome development that comes with the increased popularity of POC ultrasound is the lack of formal training for clinical specialists who have begun to use it in their practice [50]. Most education, delivered through short courses or online training, is ad-hoc and lacks procedure or discipline specific standards. Because the use of ultrasound is growing at such a high rate the existing model for education, which requires live subjects for hands-on training, is no longer adequate. Most of the practical hands-on experience comes scanning normal, paid human volunteers, but even such training is costly and of limited availability. The ultrasound manifestations of specific diagnosable conditions are limited to observing pre-recorded video tapes, in which the hands-on learning is eliminated. Simulators could become a key component in meeting standardized training and assessment requirements.

There are many reasons to study the effectiveness of simulation based ultrasound education in fields of medicine where it is applicable, i.e. abdominal trauma, cardiology, obstetrics, etc. For instance fetal abnormalities occur in approximately 2% of pregnancies in the United States, accounting for close to 80,000 events a year, and are often overwhelming for the families. Conventional ultrasound training has not led to better detection [45], which can be attributed to the lack of hands-on experience trainees have scanning these types of patients. Currently only experienced sonographers with a high number of diverse cases can develop the diagnostic ability to handle these rare events. SBME lets all obstetricians experience the detection of unusual fetal abnormalities without requiring live patients to scan.

The effectiveness of simulator training for ultrasound based guided procedures has also been studied. In [60] a simulator was developed for amniocentesis, the practice in which a needle is used to sample a small amount of fluid from the amniotic sac surrounding the developing fetus. In that study, which examined data from 30 trainees, it was shown that the simulator was effective in improving clinician skills. More recent research in [18] demonstrated that the practical image acquisition skills acquired during simulated training were directly applicable to human models in a course designed to teach the Focused Assessment with Sonography for Trauma (FAST) exam. When including simulator based training alongside didactic training in ultrasound-guided central venous catheter insertion, the residents which received the combined training outperformed their peers in aseptic technique and measurements of knowledge [41]. Transoesophageal echocardiography is another area where marked improvements in learning accompanied the introduction of commercially available simulators [61],[12].

Tissue mimicking ultrasound phantoms exist as an alternative to virtual reality based simulation, for instance CAE Healthcare's Blue Phantom line, but are expensive and do not provide a realistic degree of anatomical details. They have been developed for specialized training in areas such as obstetrics, abdominal trauma, and many others. For example, the company CIRS has a line of fetal phantoms and Kyoto Kagaku Co. has developed an abdominal trauma phantom. The problem with this model of instruction is the prohibitive cost of the phantoms, which can reach upwards of $20,000 USD for one Kyoto Kagaku abdominal trauma phantom. In addition they may only cover a few pathological conditions. There is a lack of anatomical variability and once a trainee has mastered a particular phantom there is little training value left in repeat scanning. The images produced from scanning a phantom are often clearer and contain less artifacts than clinical ultrasound, thus they may not properly prepare the trainee for scanning an actual patient. Scanning the phantom also requires the availability of an ultrasound system which runs counter to our goal of delivering a low cost training solution that can be widely deployed.

1.2 Training challenges

The task of training sonographers presents a unique set of challenges. The trainee must develop the psycho-motor skills necessary to position the transducer in the proper plane, ensuring the desired anatomy is imaged by the ultrasound system. Medical ultrasound has relatively poor spatial resolution and suffers from imaging artifacts like speckle and shadowing, while other methods such as CT or MR are acquired at higher resolutions and provide a much clearer image. Also access to willing subjects remains an obstacle to ultrasound training, particularly in the fields of obstetrics and abdominal trauma. Too few opportunities exist to scan patients with varying conditions and the motor skills essential for diagnosis require significant time, instruction, and practice to master. Simulator based training can be used to overcome this bottleneck; however the educational value is dependent upon the simulators realism, which refers to both the image quality and the system's approach to transducer manipulation/tracking.

Ultrasound simulators do not use sophisticated phantoms or require patients to scan, which means that the images must either come from a re-sliced volume stored on the hard disk or be dynamically generated by the CPU/GPU using a model of the anatomy. There are essentially two schools of thought when it comes to producing ultrasound image data for simulation purposes. Proponents of the first school suggest that realistic ultrasound images can be synthesized numerically using elements of ultrasound physics. Although the quality of artificial ultrasound images is steadily improving they are still discernible from clinically acquired images, especially in obstetrics, which possibly limits their training value. One example of this approach can be found in [13], where ray tracing was used in conjunction with a deformable mesh model. The second school of thought suggests that the ideal volume for simulation purposes should be constructed from real ultrasound images of the desired anatomy. This guarantees that US specific features such as speckle or shadowing seem realistic in the simulator and don't appear like a mathematical approximation of the physical process.

This dissertation develops a novel approach for producing these 3D composite image volumes and validates the approach using clinically acquired fetal images from the obstetrics department at the University of Massachusetts Medical School.

1.3 Review of available ultrasound simulator systems

There are a number of companies/institutions developing ultrasound simulators with a wide range of training goals in mind. A recent survey of available systems can be found in [10] where simulators have been classified by type and method of image generation. One such company is Sonosim, which has developed a laptop-based training solution with an extensive library of image volumes. Examples of training modules available include cardiology, abdominal trauma, and OB/GYN. Similar to our system, the ultrasound image displayed in Sonosim's virtual console is also sliced from a 3D volume obtained from a live subject. The crucial difference between our simulator and the Sonosim is in the actual scanning experience, where our simulator system utilizes a sham transducer with 5 degrees of freedom (DoF) whereas Sonosim's only provides 3 DoF. Sonosim only allows the trainee to place the transducer in certain predetermined locations on the body thus the sham transducer is only required to track rotation. Because no translation is allowed on the abdomen the scanning experience isn't as realistic as the one our system provides. This also implies that the trainee doesn't interact with mosaiced image volumes because a separate 3D volume can simply be stored for each possible probe location. Our system uses volumes which have been constructed from many partially overlapping 3D volumes whereas Sonosim does not.

CAE VIMEDIX recently developed an ultrasound simulator with a focus on the thoracic, abdominal, and pelvic cavities. One version of their product concentrates on preparing students to work in the Intensive Coronary Care Unit (ICCU); thus it offers content on the focused cardiac ultrasound exam, the assessment of the lung and pleural space, and the focused assessment with sonography for trauma (FAST) exam among others. Another version of their product offers training in prenatal ultrasound. Users work on developing proficiency in image acquisition, assessment of fetal anatomy, gestational age assessment, and maternal adnexal anatomy. These obstetrics learning modules are closely related to ours, which are briefly described in Chapter 7; however, our simulator adds fetal measurements such as abdominal circumference, biparietal diameter, femur length, and amniotic fluid assessment. While the training goals of the CAE VIMEDIX prenatal ultrasound simulator are comparable to ours, the approaches to simulation are very different. Using the CAE VIMEDIX system the trainee scans a manikin with a tracked probe. This position/orientation information enables the system to display the appropriate ultrasound image on the console. Tracking the transducer as the user scans a manikin adds significant cost to the system because it now requires some type of magnetic tracking solution. Our tracking system is much more cost effective. It derives the probe's position information by capturing images of a minuscule printed pattern, which was attached to the rubber abdomen, using a tiny camera placed in the transducer. More information can be found in [71]. However, the most significant difference between our system and the CAE VIMEDIX systems is the source of the image data. The ultrasound image displayed in the CAE VIMEDIX systems is simulated using a computer model, thus it is not as realistic as those displayed by our training solution.

Simbionix is Israeli company producing medical simulation devices and has been in the market since 1997. Their product line contains several virtual reality surgical trainers which were designed to simulate procedures ranging from anthroscopic knee/shoulder operations to laparoscopic surgery. Their ultrasound training system is known as the U/S Mentor and was built to educate sonographers on basic clinical skills. There are also bedside echocardiography, FAST, abdominal, trans-esophageal echocardiography, and trans-vaginal sonography modules available. This system is similar to the CAE VIMEDIX product in the sense that the trainee is scanning a manikin while observing simulated ultrasound images. One interesting component of this system is the section of the user interface that displays a detailed 3D model of the anatomy being scanned. The simulated ultrasound image is displayed along with the model, helping the sonographer associate the 2D ultrasound image with a 3D representation of the organ. Sonographers need to develop the intuition to orient themselves with the patient's anatomy based solely on 2D ultrasound slices and this supplemental view may expedite the process. The drawback of this system is the simulated ultrasound data, most likely generated from CT volumes, which doesn't match the realism obtained by re-slicing mosaiced clinical ultrasound volumes. Our system is superior to the U/S Mentor in this respect.

Finally Medaphor's ultrasound training systems should be discussed as Medaphor is the only company that incorporates mosaiced 3D volumes obtained from live subjects into their training simulators, much like our simulator does. The company hasn't published the mosaicing algorithm used to produce the composite training volumes in order to maintain their edge in the market; however, based on the advertised screenshots and our experience at various conferences testing their simulators, we infer that they follow the same basic mosaicing steps that are proposed in Figure 1-2. Although non-rigid registration is an important component in both Medaphor's volume generation process and ours, significantly different methods have been proposed in the literature which could have been used to achieve similar results. Our mosaicing method cannot be directly compared to Medaphor's because they have not published their process due to business concerns. Nevertheless, our approach to the group-wise registration of overlapping ultrasound volumes is novel with respect to the published algorithms and will be examined thoroughly in Chapter 3. Medaphor's ScanTrainer product line includes simulators for trans-vaginal and trans-abdominal ultrasound. The main difference between our simulator design and Medaphor's is the way in which the trainee interacts with the sham transducer. The ScanTrainer attaches the transducer to a custom haptic device which is then used to track position/orientation information and give the impression that the trainee is scanning a live subject. Application specific haptic devices were created for both trans-abdominal and trans-vaginal scanning. An issue with this approach is that the haptic device adds thousands of dollars to the cost of an individual system, thus an institution may only be able to afford a few training simulators. Although the elaborate design is impressive this solution doesn't address the need for a low cost training simulator which can provide a hands-on scanning experience to the masses.

Academic research initiatives have also produced a handful of simulators which appeared in the literature over the past 5 years. In [70] a transrectal ultrasound simulator for prostate imaging was developed utilizing a Wiimote attached to a dummy probe. A FAST simulator with patient specific cases was developed by [59]; however this system only allows angling of the transducer (3 DoF) at fixed positions and doesn't provide a true scanning experience. Perk Tutor, An open-source training platform for US-guided need insertions, was introduced in [79]. Finally [26] used simulated ultrasound and fluoroscopy images in a training system designed for prostate brachytherapy.

1.4 Objectives of affordable ultrasound simulator

Ideally simulators can help provide the requisite hands-on scanning experience to every medical professional who needs it. There is a prevalence of scanning systems and not enough qualified sonographers to fully utilize them. We wish to develop a low-cost portable ultrasound training simulator that emulates the actual scanning experience and thus can develop the necessary psycho-motor skills. In addition to increasing ultrasound utilization at home this training system could also impact ultrasound usage overseas. Imaging the World (ITW) is an organization pushing to bring medical expertise to the remote corners of the world, which have been neglected up to this point. The populations in these regions don't have access to advanced imaging technology so many treatable conditions, especially maternal disorders, go undiagnosed which often results in increased morbidity and mortality for mother and child. Due to its portability/low cost ultrasound is an ideal modality to introduce into these areas. For example, ITW has established low cost ultrasound programs in some of the poorer regions of Africa, one of which led to increased attended deliveries at a health clinic in rural Uganda [67]. Because our simulator was designed from the ground up to be low cost it would be a great fit for ultrasound training programs in the developing world. The need for improved ultrasound training is well known in the medical community.

This chapter describes the core motivation behind our research, the main body of which is presented in Chapters 2 through 7. We will introduce a low cost, PC-based simulator which uses training data that has been reconstructed from overlapping 3D ultrasound volumes acquired in a clinical setting. Our philosophy is to provide a library of real ultrasound volumes comprising a specific region of the body and which are acquired from a wide range of subjects where each volume has specific training objectives associated with it. These objectives could as simple as locating anatomical landmarks or more complex such as identifying malignancies or assessing fetal development. A system with a library of training volumes will permit an unlimited number of ultrasound learners to be trained in detection of various medical conditions in a virtual environment. We will also discuss how our approach to ultrasound simulation improves upon the available commercial simulators. Ultimately though, this dissertation is about the novel stitching techniques that were developed to produce the training volumes. It should be noted that throughout this dissertation we will refer to the process of stitching partially overlapping 3D ultrasound volumes as mosaicing. The challenges encountered and techniques developed in order to produce volumes which encompass the entire abdomen are discussed in greater detail within the following chapters.

1.5 Affordable ultrasound simulator overview

A key part of the proposed training model is the simulator itself and a low cost PC-based ultrasound simulator has been designed and implemented. The training is provided by scanning a generic curved compliant scan surface, referred to as the physical scan surface (PSS) with a sham transducer, containing 5 degrees of freedom (DoF) position and orientation sensors, while the PC displays both a virtual subject and a virtual transducer, along with an ultrasound image, obtained from a 3D image volume. Due to the importance of developing the sonographer's psychomotor skills, there is a need for extended (or composite) image volumes and 5 DOF tracking. As a consequence of this, a mosaicing process needed to be developed. The upper half of Figure 1-1 shows a screenshot of the simulator running on a laptop with the major elements labeled. Starting from the right most side we see the image library panel. After mosaicing, the composite 3D ultrasound volumes were assessed for educational value and subsequently loaded into the image library if they were deemed to be instructive. The virtual torso panel, which is displayed above the image library, enables the trainee to visualize the exact location on the human body where he/she has placed transducer. This is necessary since there are no anatomical references on the PSS they are scanning. The instructional field, also shown in Figure 1-1, educates the user on how to perform specific clinical tasks belonging to the training module they are currently working on. Obstetrics training modules have been developed with the help of Dr. Belady and her staff of sonographers at the University of Massachusetts Medical School and are explained in greater detail in Chapter 7. In this screen shot the user is measuring the biparietal diameter of the baby's head. Training modules could be developed in the future with different medical ultrasound applications in mind. An abdominal trauma module is a great example of future work which should be completed. The lower half of Figure 1-1 shows the shame transducer, which has 5 degrees of freedom (DoF), and the physical scan surface used to track the transducers movement on the abdomen. The position and orientation of the sham transducer is tracked along on the scan surface. Position accounts for 2 degrees of freedom, while

3D orientation requires 3 DoF to properly track. The only degree of freedom not currently implemented is the movement of the probe, normal to the scan surface. This is directly related to probe pressure, which would display the internal organs as being deformed if properly simulated. The components described above form the foundation of the simulator and our focus lies in producing realistic mosaics for inclusion within the image library.

The simulator addresses the challenges mentioned in this chapter by offering a realistic scanning experience, using actual fetal ultrasound data in the obstetrics module. We also obtained high quality abdominal image data from a trauma phantom borrowed from Kyoto Kagaku Co., which can be used to train sonographers in the FAST (Focused Assessment for Trauma) exam. The system provides a more realistic scanning experience when compared to competing systems that use simulated ultrasound data or fewer than 5 DoF, as described above. Also, our sham transducer and tracking system can be constructed for a few hundred dollars, encouraging widespread adoption by institutions with tighter budgets. Because our system is so cost effective sonographers may train at home, on their personal laptops, when it is convenient for them. Our vision includes the creation of an online repository of training volumes, so every sonographer has access to constant stream of unique patient cases.

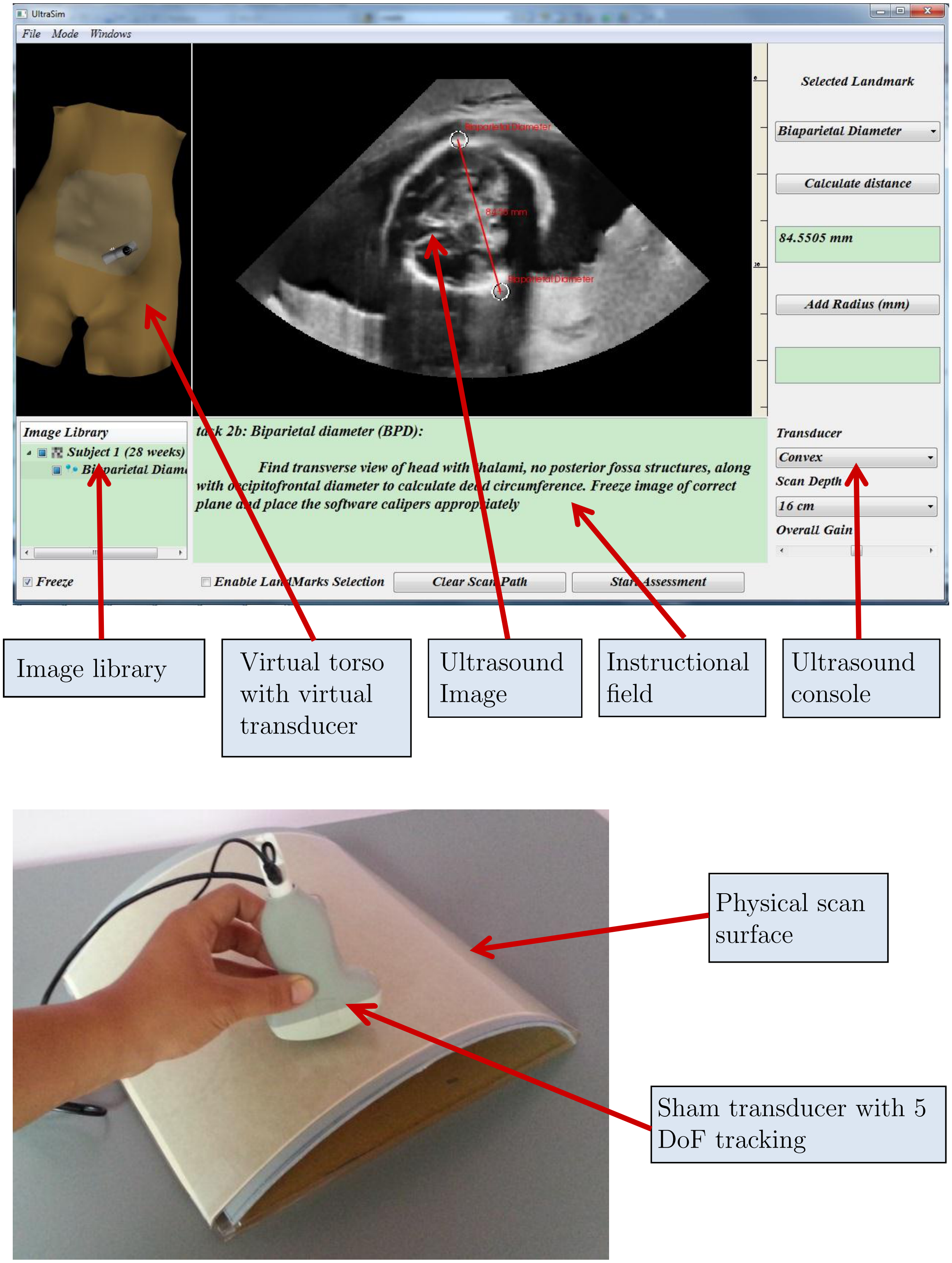


Figure 1-1: Key components of affordable training system

1.6 Overview of mosaicing process

This section discusses the steps which were used to produce training volumes for the affordable ultrasound simulator. Figure 1-2 shows a flowchart illustrating the process to construct a global composite volume from individual partially overlapping 3D volumes. We have broken the process down into 5 image processing steps. This dissertation develops novel techniques to perform steps 2 to 5, focusing on discrete graph based methods which are popular in the computer vision community due to their efficiency/performance. We will briefly discuss each block shown in Figure 1-2 so that the reader has an understanding of all the steps found in our mosaicing pipeline. The first step in ultrasound mosaicing is 2D image/position acquisition, followed by the formulation of 3D volumes with uniformly spaced voxels. The 3D freehand scans were acquired with a Philips iU22 ultrasound system utilizing a convex array coupled with an Ascension Technologies trakSTAR 6 DoF position tracker. Stradwin software [78], developed by the medical imaging group at the University of Cambridge, was used to produce the 3D image volume. This software links position information from the trakSTAR with each captured frame. Essentially the 2D images which were collected during the freehand sweep are stacked to form the 3D volume. Once image acquisition is complete and each frame has been positioned in a coordinate system defined by the position tracker, we are left with a scattered data interpolation problem. This is due to the fact that a 3D uniform grid of samples is needed, and the pixels in the 2D frames acquired in the previous step do not fall on exact voxel locations in the grid. Many techniques can been used to perform this step [73] and we implemented the basic algorithm presented in [40]. Our exact hardware configuration will be discussed in greater detail in Chapter 7, where the clinical results are presented.

Once the individual 3D volumes have been generated it is practical to rigidly register adjacent some widely used similarity metrics in order to make them more resilient in the presence of shadows volumes. This step corrects overall global movement, which could be caused by the patient or fetus shifting between scans. Chapter 2 discusses rigid registration with ultrasound data and introduces improvements to. Based on

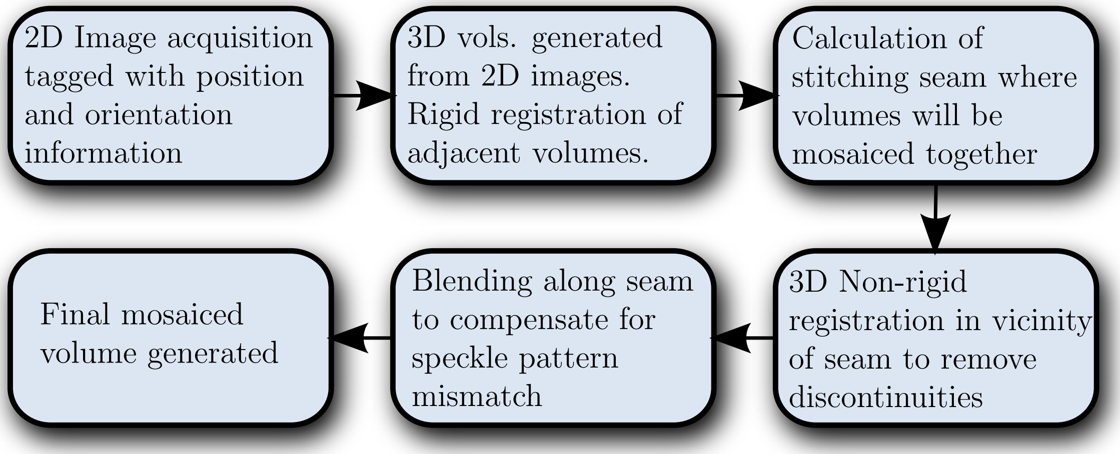


Figure 1-2: Steps to produce composite 3D ultrasound volume

our experience shadowing in clinical ultrasound can heavily affect registration results.

Next the stitching seam is calculated in order to bisect the area of overlap and join two neighboring volumes together. We found that the movement between overlapping abdominal volumes makes alignment difficult thus traditional methods of spatial compounding, the simplest example being a weighted average of overlapping voxels, produces poor results. This is especially true in fetal imaging where the baby may move during a scan, which is why our research included stitching seam calculation.

In our initial research, planes were calculated between each of the overlapping volumes in order to mosaic them together. More optimal stitching surfaces can be formulated, which are discussed in Chapters 3 and 5. The composite volume was constructed by adding each source to it sequentially, thus growing it volume by volume. The stitching plane for each step was required to travel through the overlapping region's center of mass. Also, principal component analysis was used to define a coordinate system which was aligned with the overlapping regions 3 major axes. Aligning the stitching plane's normal with the smallest principle component resulted in a solution that bisected the overlapping region best. Figure 1-3 shows an example where planar seams have been used. The left image shows the area of the abdomen (red volume), which 8 overlapping scans cover when combined. The image also shows the planar seams between each scan where each plane's orientation was calculated by

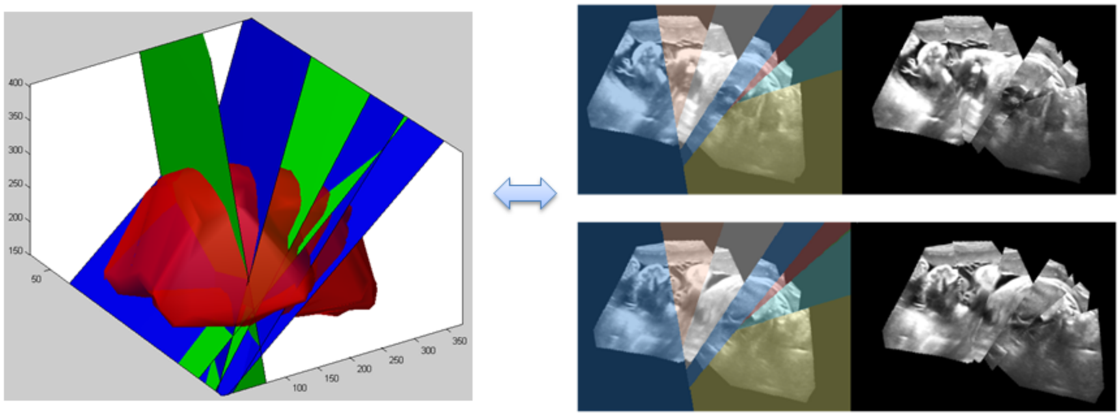


Figure 1-3: Mosaic created using planar seams. Left-most image shows planes calculated using principal component analysis. The right-most column contains two example slices of the composite volume.

applying principal component analysis to the overlapping region. On the right are slices from the combined image volume formed using this method. Because clinical freehand acquired volumes can be oddly shaped/oriented in relation to each other planar seams aren't flexible enough for stitching. This can be seen by the misalignment in the slices. The stitching mask that designates which volume the data originates from is also shown as an overlay on each slice. Interpolation based methods, such as those used for spatial compounding, would require extensive group-wise non-rigid registration to provide reliable results due to movement/deformation in our clinical experiments and so are not presented. Figure 1-4 shows the results of graph-cut based seam selection discussed in Chapter 3. The Left side shows a cross section of the composite volume. Right top image shows a slice with no non-rigid registration performed. Bottom right shows results after non-rigid registration is performed in the vicinity of the seams. This slice is the counterpart to the planar seam slice shown in the top right corner of Figure 1-3. Step 4 is non rigid registration, which is used to remove discontinuities along organ edges spanning more than one volume. A large portion of this dissertation has been dedicated to developing efficient non-rigid group-wise registration techniques. Chapter 3 describes the probabilistic framework for this crucial step and then presents novel techniques to perform it on many partially overlapping ultrasound volumes. The probabilistic framework is linked to a Markov

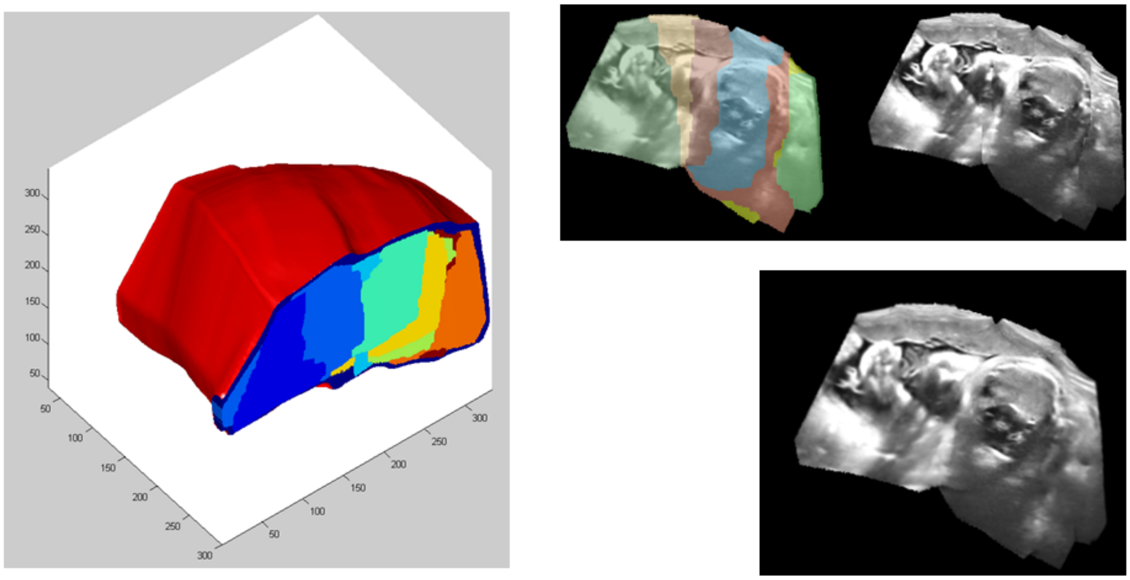


Figure 1-4: Results of graph-cut based approach to mosaicing discussed in Chapter 3

random field, which can be efficiently optimized with graph based techniques common in the computer vision literature. We are essentially performing a group-wise block matching algorithm in the regions of overlap. Also the non-rigid similarity metric presented in Chapter 3 is robust to shadow artifacts and is efficiently evaluated using Fast Fourier Transform (FFT) techniques. The proposed group-wise registration algorithm was then evaluated using synthetic, phantom, and clinical image data.

The final step in the mosaicing process is seam blending which corrects the speckle pattern mismatch between adjacent volumes. This results in a continuous mosaic comprised of several individual 3D volumes. Chapter 6 describes our approach to blending the overlapping sources in the vicinity of the seams. In essence, a 2D Poisson image editing algorithm was adapted to handle multiple 3D ultrasound volumes. This method preserves the distinctive features in each source while forming an unnoticeable transition between adjacent volumes. Figure 1-5 shows a slice from a completed fetal image volume, which was constructed using data attained from a live subject at the University of Massachusetts Medical School. The left-most image displays the slice along with an overlay designating the source volume for each area. There are 7 identifiable source volumes in this mosaic. The middle image is before non-rigid

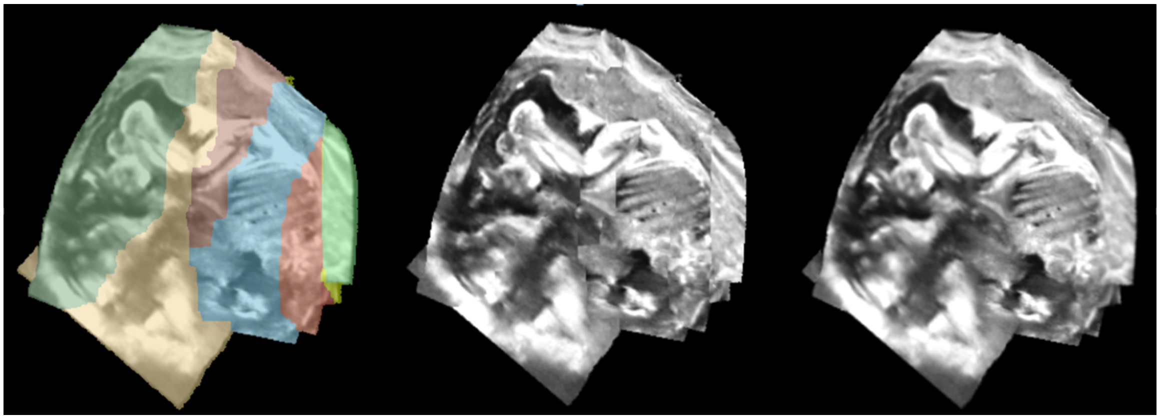


Figure 1-5: A completed fetal image volume. Left image shows a colored overlay designating the source volume for each region of the mosaic. Middle image shows slice without non-rigid registration. Right image shows result after non-rigid registration.

registration where the errors between source volumes are apparent. The final result, after the discontinuities have been removed using non-rigid registration and the source volumes blended together, can be seen in the right-most image of Figure 1-5.

1.7 Organization of dissertation

The material presented in the following chapters is roughly organized according to the flowchart in Figure 1-2. Chapter 2 describes ultrasound specific rigid registration techniques. Following this, Chapter 3 begins with a discussion on the non-rigid alignment of multiple (≥ 3) overlapping volumes. Seam selection is then introduced in the context of group-wise non-rigid registration because it is considered to be an additional pre-processing step specific to ultrasound imaging. Chapter 4 presents experimental results obtained using the theory developed in the previous chapter. Chapter 5 presents another approach to seam selection where we attempt to globally optimize a spline surface using particle swarm methods. This research direction was dropped due to the effectiveness of the Markov random field approach. Chapter 6 reviews the theory behind the Poisson based blending technique. Chapter 7 presents the clinical component of this dissertation in which 12 patients were scanned at the University of Massachusetts Medical School. This process generated a library of fetal training volumes to be used in the simulator. Chapter 7 also discusses obstetrics training goals and how our system can be used to educate sonographers on how to perform important tasks such as gestational age assessment and fetal position. Chapter 8 describes our approach to imaging the fetal heart using freehand ultrasound. Finally, conclusions and future work are presented.

Chapter 2

Rigid registration with shadows

Rigid registration of 3D ultrasound (US) volumes acquired from human subjects in a clinical environment introduces an array of difficulties, which aren't normally encountered in a controlled laboratory setting. Ultrasound shadowing, which can either result from the acoustic properties of the subject's internal anatomy or because of strong reflections from bone structures or gas, becomes a major issue when aligning one ultrasound volume with another. This process is referred to as US-US registration. Shadowing artifacts coupled with patient movement can prevent common ultrasound similarity metrics from converging to the correct solution, or even converging to a reasonable solution at all. This makes rigid registration in the presence of heavy shadowing an interesting problem to study. In this chapter we will introduce a similarity metric which addresses these problems while also allowing for efficient optimization of the desired transformation.

2.1 Introduction

As discussed in Chapter 1, our primary goal of producing 3D composite fetal image volumes for an ultrasound training simulator requires stitching together several partially overlapping volumes obtained by an obstetrics sonographer. Fetal ultrasound presents additional difficulties to registration algorithms due to fetal movement during scanning as well as occasional heavy shadowing which can be seen in the B-mode images. Also the patient is more likely to move during the prolonged time it takes to conduct a complete scan of the abdomen, which is referred to as global movement. The first step in the mosaicing process, which all subsequent steps rely upon, is to rigidly align the overlapping volumes to account for this motion. Based on the experience gained from scanning 12 pregnant women at UMass Medical School's obstetrics department, we concluded that the existing registration similarity metrics commonly used, such as mutual information, were unable to properly align the heavily shadowed volumes obtained. A robust similarity metric should be employed to ensure proper alignment prior to non-rigid registration and stitching. Due to our particular application this metric must be resilient in the presence of shadows, thus image registration models which ignore their effects perform poorly on our fetal data.

The key to our similarity metric's computation will be the identification of regions with poor ultrasound transmission. Intuitively these regions should contribute less to the computation of the metric. Recently, a few novel methods for identifying shadows in ultrasound images have been presented in the literature [28],[34]. In [28] the authors determine shadows using a combined geometrical/statistical approach. They apply their method to reconstruction and US-US registration showing an improvement when incorporating shadow detection. The author's approach to registration is to mask out the shadowed regions from the computation of the SSD (sum of squared differences), which means this metric doesn't take into account the unique properties of ultrasound noise.

2.2 Pair-wise registration in a probabilistic framework: a review

Before derivation of the more robust ultrasound similarity measures it is appropriate to explain the Maximum Likelihood Estimation (MLE) framework for pair-wise image registration [65], where two overlapping volumes are considered. In this discussion the transformation type hasn't been defined yet so it may be rigid or non-rigid; however, the results presented in this chapter use a rigid transformation model. In the pair-wise case one volume is considered the scene, the other volume is considered the model, and the goal of the optimization process to is determine a transformation from the scene coordinate system to the model coordinate system. We will give a brief overview; a more detailed elaboration on this interesting perspective of image registration can be found in [65]. Consider an image to be a grid of voxels where the intensity of each voxel is a random variable (r.v.), which may take on an intensity value from the set $\mathcal{I} = \{0, 1, ..., 255\}$. Thus an image is a collection of random variables where the probability of a voxel assuming a specific value from $\mathcal{I}$ is given by $\mathrm{P}(I(x) = i)$ where $i \in \mathcal{I}$. It is usually assumed in the registration process that voxels are i.i.d. (independent and identically distributed) as we do but others have developed similarity metrics that do not require this condition. In terms of probability maximization we would like to calculate

$$\hat{T} = \underset{T}{\operatorname{argmax}} \; \mathrm{P}(I, J, T) \tag{2.1}$$

where the source volumes I and J are fixed at their observed values. The solution for (2.1) is an optimized transformation specified by $\hat{T}$. This equation can be further manipulated using Bayes' theorem to give us the likelihood function in terms of probabilities we are able to compute using the acquired image data as follows,

$$\begin{aligned} \mathrm{P}(I, J, T) &= \mathrm{P}(T|I, J)\,\mathrm{P}(I, J) \\ &= \frac{\mathrm{P}(I, J|T)\,\mathrm{P}(T)}{\mathrm{P}(I, J)}\,\mathrm{P}(I, J) \\ &= \mathrm{P}(J|I, T)\,\mathrm{P}(I|T)\,\mathrm{P}(T) \\ &\approx \mathrm{P}(J|I, T)\,\mathrm{P}(T). \end{aligned} \tag{2.2}$$

The terms $\mathrm{P}(J|I, T)$ and $\mathrm{P}(I, J|T)$ are usually not distinguished by the literature in the likelihood function. From this we see that the log likelihood function for the

MLE framework is

$$\mathcal{L}(T; I, J) \;=\; \ln \mathrm{P}(J|I,T) + \ln \mathrm{P}(T), \tag{2.3}$$

If certain assumptions are made we can derive a joint probability distribution for the two image volumes in the pairwise registration problem, and then use this distribution to derive a number of similarity metrics in a maximum likelihood framework. The derivation of the sum of squared differences (SSD) is shown in this chapter, but other metrics like the correlation ratio and mutual information can also be derived using this framework [65]. In the likelihood function the image volumes are fixed to their observed intensity values, and the transformation parameters are varied to maximize the likelihood of observing both images. In the case of pairwise registration we see that the matching or similarity term is given by $\mathrm{P}(J|I,T)$ and the regularization term is given by $\mathrm{P}(T)$. This is an intuitive result because we would like to penalize transformations which are not physically realistic and the regularization term provides a means to do that. The calculation of $\mathrm{P}(J|I,T)$ can be simplified using the assumption that the voxels of J are independent of each other given the volume I. Thus we can calculate this term using the following equation,

$$\mathrm{P}(J|I,T) = \prod_{x_k \in \Omega_J} \mathrm{P}\left(j_k | i_k^{\downarrow}\right) \text{ where } \begin{cases} i_k^{\downarrow} = I(T(x_k)) \\ j_k = J(x_k) \end{cases} \tag{2.4}$$

where Ω_J denotes the voxel grid of image J and $x \in \Omega_J$. The arrow notation simply means that the transformation has been applied. This equation essentially compares the transformed version of I to J voxel by voxel and expresses the fact that the volume J is a random variable, which is dependent on I. This makes sense because pair-wise image registration usually assumes that volume J is produced by imaging a deformed version of I. The imaging process which produced J is not necessarily assumed to be identical to the process that produced volume I, which enables the formulation of multi-modal methods that are outside the scope of this chapter. The transformation is from J's coordinate system to I's coordinate system; thus it should be noted that the voxel intensity $i_k^{\downarrow}$ will have to be interpolated. If the image creation process is simplified and certain assumptions are made the likelihood function can derived explicitly. A simple model which expresses the relationship between the fixed image J and the moving image I is

$$J(x) = f(I(T(x))) + \varepsilon(x). \tag{2.5}$$

The function $f : \mathbb{R} \to \mathbb{R}$ in (2.5) is the intensity mapping, which models differences in voxel values of the same structure attained during different acquisitions, and where the function $\varepsilon(x)$ represents noise. The volume J can also be thought of as a function of the random variables I and ε. Volume J is referred to as the fixed image in this equation because volume I is transformed into J's coordinate system for comparison. As stated before, the acquisitions to be registered do not necessarily use the same imaging modality so this function can be extremely important. The Correlation Ratio [66] is an example of a similarity measure that uses a functional relationship between intensity values in multi-modality image registration. Measures based on Mutual Information [62] are even more general as there is no functional relationship assumed, only a statistical one; however, since we are concerned with mono-modality, specifically ultrasound-ultrasound (US-US) registration of many overlapping volumes, we are more interested in the computational advantages of simple similarity measures rather than measures suited for multi-modality registration.

In the simplest case we presume $\varepsilon(x)$ to be white Gaussian noise and the intensity mapping in (2.5) to be identity, meaning that matching tissue classes, such as fetal bone, in volumes I and J are given the same intensity value. Based on these assumptions it can easily be shown that the similarity measure which presents itself in the MLE framework is just the sum of squared differences or SSD. Here the conditional densities are Gaussian and can be written as

$$\mathrm{P}\left(j_k = j | i_k^{\downarrow} = i\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(j-i)^2}{2\sigma^2}\right)$$

which makes the similarity term in the log likelihood function,

$$\begin{aligned}
\ln \mathrm{P}(J|I,T) &= \ln \prod_{x_k \in \Omega_J} \mathrm{P}\left(j_k | i_k^{\downarrow}\right) \\
&= \sum_{x_k \in \Omega_J} \ln \mathrm{P}\left(j_k | i_k^{\downarrow}\right) \\
&= |\Omega_J| \ln \sqrt{2\pi}\sigma + \frac{1}{2\sigma^2} \sum_{x \in \Omega_I} \left(j_k - i_k^{\downarrow}\right) \\
&\approx \sum_{x \in \Omega_I} \left(j_k - i_k^{\downarrow}\right)^2 .
\end{aligned} \tag{2.6}$$

The final term in (2.6) is seen to be the sum of squared differences, which can be useful in mono-modality registration as long as the intensity value for matching tissue types in volumes I and J are identical. This is not the case when registering different types of MRI scans together or MRI-CT scans, but for US-US registration it can be used if the shadowing/reflectance artifacts associated with varying probe position are ignored and the speed of SSD is important to the application. Other US-US similarity metrics exist [3], which assume noise models more consistent with observed clinical ultrasound images; however, they require more computational effort to calculate, and we have seen good results using SSD.

2.3 Simple models for ultrasound shadowing

As discussed above, registration similarity metrics are typically constructed using probability theory by assuming an image formulation model and then deriving an expression for either $\mathrm{P}(I|J,T)$ or $\mathrm{P}(J|I,T)$, where I, J are the images and T is the transformation between the two. The simplest model to account for shadowing during ultrasound image formation is shown below,

$$\begin{aligned}
I(\mathbf{x}) &= M_I(\mathbf{x})\, S(T_I(\mathbf{x})) + \varepsilon_I \\
J(\mathbf{x}) &= M_J(\mathbf{x})\, S(T_J(\mathbf{x})) + \varepsilon_J
\end{aligned} \tag{2.7}$$

where $I(\mathbf{x})$ and $J(\mathbf{x})$ are the acquired ultrasound volumes and $S(\mathbf{x})$ represents the ideal image of the anatomy. In (2.7) the ideal image has been deformed, had its intensity scaled and has also been corrupted with additive Gaussian noise, ε_I and ε_J respectively. We will show in section 2.5 that if you assume this image formulation model then we can use the probabilistic registration framework presented in section 2.4 [65] to derive the simple algorithm of [28]. We will also consider the ultrasound image formulation model developed in [17], which accounts for the unique statistical properties of ultrasound noise and has become popular in the literature due to its improved registration accuracy. This intensity based similarity metric is known as CD2 and it simply considers ultrasound images to be realizations of the ideal anatomical signal, which has been corrupted by multiplicative noise. By considering two images of the same source, which have both been corrupted by Rayleigh noise, it is possible to derive an expression measuring the similarity between them. Also, it is symmetric in a sense because the two ultrasound images, I and J, have been corrupted by independent and identically distributed noise processes. CD2 is derived using a model similar to the following,

$$\begin{aligned} I(\mathbf{x}) &= \eta_I M_I(\mathbf{x}) S(T_I(\mathbf{x})) \\ J(\mathbf{x}) &= \eta_J M_J(\mathbf{x}) S(T_J(\mathbf{x})) \end{aligned} \tag{2.8}$$

The key difference between (2.7) and (2.8), the first being linked to the sum of squared differences similarity metric, is that multiplicative Rayleigh noise is used in (2.8) instead of additive Gaussian noise. In both equations the source $S(\mathbf{x})$ is scaled by masks $M_I(\mathbf{x})$ and $M_J(\mathbf{x})$, which is unique to our model and added to make the registration similarity metric robust to shadowing. The original CD2 metric didn't incorporate these masks, which will account for the amount of ultrasound transmission at each voxel. These masks may be Boolean valued such that $M : \mathbb{R}^3 \rightarrow \{0, 1\}$, or they could be intensity maps quantifying the degree of ultrasound transmission at a particular voxel.

Since running the algorithm in [28] results in a binary image mask, it does not provide much information about the quality of ultrasound transmission at each location. In [34] the authors develop a novel ultrasound confidence measure which assigns a value $v \in [0, 1]$ for each voxel in the B-mode image. This value can also be thought of as a measure of ultrasound transmission, thus for areas labeled 1 there is complete transmission and for areas labeled 0 there is none. Their approach builds upon the much cited work of [27] who introduced random walk segmentation to the medical imaging field. [34] added ultrasound specific constraints to the model resulting in their confidence measure. They applied their work to ultrasound reconstruction as well as the multi-modality task of CT-US registration and demonstrated improved accuracy. This algorithm was used during the generation of the training volumes presented by this dissertation to calculate all of the masks required for the rigid and non-rigid registration steps.

Our contribution is to incorporate this transmission measure into existing ultrasound similarity metrics, which already take into account the unique properties of speckle noise, in order to improve registration of partially overlapping B-mode image volumes containing shadows. It has been shown in the literature that ultrasound specific metrics outperform more general ones [17],[3] such as the SSD used in [28], and also mutual information [83]. Utilizing a probabilistic framework for symmetric registration we will extend the popular CD2 metric [17] by integrating an ultrasound transmission measurement using the masks in (2.8). We will show that in certain situations, an example being abdominal ultrasound where the patient may have shifted during scanning, the original CD2 metric could fail to converge. The simple extension derived in this chapter helps alleviate this problem without adding a significant computational burden.

2.4 Probabilistic tools for analysis of symmetric models

In this section we will review the basic probabilistic framework used to analyze the symmetric registration problem we have presented in section 2.3, where two overlapping ultrasound volumes are acquired from an anatomical source. Chapter 3 will extend the pair-wise registration concepts presented here to a group-wise setting, which better reflects the mosaicing procedure used to generate training volumes for the simulation system since the actual scanning produces more than two overlapping volumes needing alignment. The following framework was derived in [65] and is presented here as review. In this work an image volume is considered a realization of a random process which has corrupted the ideal source "signal". A few assumptions will need to be made before the bulk of the derivation is presented. These assumptions are presented in the context of image I but apply to J as well.

The first assumption is that the voxels in I are conditionally independent given the source S which is stated as

$$\mathrm{P}\left(I|S\right) = \prod_{\mathbf{x}_k \in \Omega_I} \mathrm{P}\left(i_k|S\right) \tag{2.9}$$

where $i_k = I\left(\mathbf{x}_k\right)$ and Ω_I is the uniform grid of voxel coordinates associated with image I. The second assumption, which allows simplification of the expression in equation (2.9) even farther, is that the voxel intensity at i_k depends only its corresponding voxel intensity in the source, denoted as $s_k^{\downarrow} = S\left(T_I\left(\mathbf{x}_k\right)\right)$. Combining these ideas one can write

$$\mathrm{P}\left(I|S, T_I\right) = \prod_{\mathbf{x}_k \in \Omega_I} \mathrm{P}\left(i_k|S, T_I\right) = \prod_{\mathbf{x}_k \in \Omega_I} \mathrm{P}\left(i_k|s_k^{\downarrow}\right) \tag{2.10}$$

The second assumption effectively means that the noise corrupting voxel i_k can only depend on $s_k^{\downarrow}$, thus it must be context free. Basically, this means that the neighboring intensity values of a source voxel do not influence the noise at that particular voxel when it is imaged.

As stated before, the goal is to align volume I with volume J by deforming each one to match in the region of overlap. For the remainder of this chapter we will assume that J is already aligned with S for the sake of brevity, which implies that $T_J = Id$. Also since there is only one transformation of interest left, let $T_I = T$. A simple extension allows the deformation of both volumes, which will become essential in the subsequent group-wise registration chapter.

As discussed, the most likely transformation between I and J is the one that maximizes (2.3), which requires us to calculate the joint probability of I and J during optimization. The expression in (2.3) doesn't contain the source volume S; thus the main challenge of formulating any symmetric registration metric, which is robust to ultrasound shadowing, becomes integrating out the source probability since we have no explicit knowledge of it. Canceling the volume S from the models presented in (2.7) and (2.8) can be done using a bit of algebraic manipulation; however it is instructive to use this framework since more sophisticated models might not present an algebraic solution. In models where the source can't be eliminated from the expression for $\mathrm{P}(I|J,T)$, techniques such as expectation maximization can be used. The following derivation will result in an expression that can be used for any symmetric image formation model.

Starting with the joint probability of the image pair and using the fact that the images in the pair are independent of each other given the source volume, the following expression can be written,

$$\begin{aligned} \mathrm{P}(I,J|T) &= \int \mathrm{P}(I,J|S,T)\,\mathrm{P}(S)\,dS \\ &= \int \mathrm{P}(I|S,T)\,\mathrm{P}(J|S)\,\mathrm{P}(S)\,dS \end{aligned} \tag{2.11}$$

In order to simplify (2.11) we will assume that the voxels of S are independently distributed and as a consequence we find that $\mathrm{P}(S) = \prod\limits_{\mathbf{x}_k \in \mathcal{S}} \mathrm{P}(s_k)$, where $\mathcal{S}$ is the set of voxel coordinates aligned with the source volume. Finally let T be an injection mapping, which means that it maps distinct voxels from I to distinct voxels in S, i.e.

$T : \Omega_I \rightarrow \mathcal{S}$. Since T is an injection mapping it is possible to consider the inverse transformation from $\mathcal{S}$ to Ω_I and write $i_k^{\uparrow} = I\left(T^{-1}\left(\mathbf{x}_k\right)\right)$, where $\mathbf{x}_k \in \mathcal{S}$. Care must be taken since not all voxels in S necessarily have corresponding voxels in I. Let the voxels in S which are matched to voxels in I be denoted as $\mathcal{S}_I \equiv T^{-1}\left(\Omega_I\right)$. This also applies to J as well because despite already being aligned with S it may only contain a limited view of the scene, implying $\Omega_J \subseteq \mathcal{S}$. Using the result from (2.10) the preceding equation can be rewritten as

$$
\begin{aligned}
&\mathrm{P}\left(I, J|T\right) = \int \prod_{\mathbf{x}_k \in \mathcal{S}_I} \mathrm{P}\left(i_k^{\uparrow}|s_k\right) \prod_{\mathbf{x}_k \in \mathcal{S}_J} \mathrm{P}\left(j_k|s_k\right) \mathrm{P}\left(s_k\right) \prod_{\mathbf{x}_k \in \mathcal{S}_I \cup \mathcal{S}_J} ds_k \\
&= \left(\prod_{\mathbf{x}_k \in \mathcal{S}_I \setminus \mathcal{S}_J} \int \mathrm{P}\left(i_k^{\uparrow}|s_k\right) \mathrm{P}\left(s_k\right) ds_k\right) \left(\prod_{\mathbf{x}_k \in \mathcal{S}_I \cap \mathcal{S}_J} \int \mathrm{P}\left(i_k^{\uparrow}|s_k\right) \mathrm{P}\left(j_k|s_k\right) \mathrm{P}\left(s_k\right) ds_k\right) \\
&\quad \left(\prod_{\mathbf{x}_k \in \mathcal{S}_J \setminus \mathcal{S}_I} \int \mathrm{P}\left(j_k|s_k\right) \mathrm{P}\left(s_k\right) ds_k\right) \\
&= \left(\prod_{\mathbf{x}_k \in \mathcal{S}_I \setminus \mathcal{S}_J} \mathrm{P}\left(i_k\right)\right) \left(\prod_{\mathbf{x}_k \in \mathcal{S}_I \cap \mathcal{S}_J} \int \mathrm{P}\left(i_k^{\uparrow}|s_k\right) \mathrm{P}\left(s_k|j_k\right) ds_k\right) \left(\prod_{\mathbf{x}_k \in \mathcal{S}_I \cap \mathcal{S}_J} \mathrm{P}\left(j_k\right)\right) \\
&\quad \left(\prod_{\mathbf{x}_k \in \mathcal{S}_J \setminus \mathcal{S}_I} \mathrm{P}\left(j_k\right)\right) \\
&= \left(\prod_{\mathbf{x}_k \in \mathcal{S}_I \setminus \mathcal{S}_J} \mathrm{P}\left(i_k\right)\right) \left(\prod_{\mathbf{x}_k \in \mathcal{S}_I \cap \mathcal{S}_J} \int \mathrm{P}\left(i_k^{\uparrow}|s_k\right) \mathrm{P}\left(s_k|j_k\right) ds_k\right) \left(\prod_{\mathbf{x}_k \in \mathcal{S}_J} \mathrm{P}\left(j_k\right)\right) \\
&= \mathrm{P}\left(I|J, T\right) \mathrm{P}\left(J\right)
\end{aligned}
\tag{2.12}
$$

We are concerned with calculating the similarity measure in the region of overlap and ignore the prior probabilities in (2.12). To summarize, the goal of the derivation in (2.12) was to end up with an expression for $\mathrm{P}\left(I|J, T\right)$, in which we could plug in our symmetric image formulation models from (2.7),(2.8) and then subsequently integrate out the source S. This is because in practice we only have data from the two acquired volumes I and J. The log likelihood of (2.12) is typically used, which increases the numerical stability of the expression since we are no longer multiplying large quantities of extremely small probabilities. The final form of the registration similarity measure is shown below,

$$\mathcal{L}(T; I, J) = \sum_{\mathbf{x}_k \in \mathcal{S}_I \cap \mathcal{S}_J} \ln \left[\int \mathrm{P}\left(i_k^{\uparrow} | s_k\right) \mathrm{P}\left(s_k | j_k\right) ds_k \right] \tag{2.13}$$

This equation allows us to calculate the joint density between two volumes which may not have been corrupted by the same imaging process. In the next section (2.13) will be applied to (2.7) and (2.8) to demonstrate how it can be used. An example of this technique applied to a more complicated imaging model, which ends up requiring the use of an expectation maximization algorithm, can be found in appendix B of [65]. Application of (2.13) in that case results in a mixture of Gaussians, which is a well studied problem with no known explicit solution. Due to the poor quality of the registration images our ultrasound application doesn't require anything more sophisticated than the model given by (2.8).

2.5 Improved similarity metrics

In this section we will analyze the models from (2.7) and (2.8) using (2.13) and then explain the intuitive equations that result. Starting with (2.7) the first step is to find the probabilities $\mathrm{P}\left(i_k^{\uparrow} | s_k\right)$ and $\mathrm{P}\left(s_k | j_k\right)$, which is simple because the additive noise is assumed to be Gaussian. The first expression is the probability of observing the voxel intensity $i_k^{\uparrow}$ given it's corresponding source voxel intensity s_k. The second expression is basically looking the other direction, i.e. what is the probability of a source voxel having a specific intensity value, given that its corresponding voxel in the acquired image has a known value. These probabilities are listed below,

$$\begin{aligned} \mathrm{P}\left(i_k^{\uparrow} | s_k\right) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-\left(i_k^{\uparrow} - m_{I,k}^{\uparrow} s_k\right)^2}{2\sigma^2}\right) \\ \mathrm{P}\left(s_k | j_k\right) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-\left(j_k - m_{J,k} s_k\right)^2}{2\sigma^2}\right) \end{aligned} \tag{2.14}$$

Substituting (2.14) into (2.13) and then carrying out the integration results in

$$\mathcal{L}\left(T;I,J\right) = -\sum_{\mathbf{x}_k \in \mathcal{S}_I \cap \mathcal{S}_J} \frac{\left(m_{J,k} i_k^{\uparrow} - m_{I,k}^{\uparrow} j_k\right)^2}{2\sigma^2 \left(\left(m_{I,k}^{\uparrow}\right)^2 + \left(m_{J,k}\right)^2\right)} + \frac{1}{2} \ln\left[\left(m_{I,k}^{\uparrow}\right)^2 + \left(m_{J,k}\right)^2\right] \tag{2.15}$$

where the transform independent terms have been dropped because they do not affect the registration result. Also we have assumed that $m_{I,k}^{\uparrow}$ and $m_{J,k}$ are both greater than zero, which is valid because locations where either one is zero provides no information to the registration algorithm and so these voxels are eliminated from the calculation. Looking at (2.15) we see that if $m_{I,k}^{\uparrow} = m_{J,k} = 1$, which implies confidence in the image data from the region of overlap, then the optimization of the expression simplifies to minimizing the sum of squared differences. This is essentially the same similarity metric used in [28], where the sum of squared difference calculation is performed using the set of voxels from the overlapping region which lie outside of areas classified as shadows. Equation (2.15) is interesting because it allows the mask to lie in the range $0 < m \leq 1$ and not just take on Boolean values indicating whether or not a shadow is present. For example voxels with a low ultrasound transmission value may still contain useful information for the registration algorithm and should be weighted appropriately. As $m_{J,k} \to 0$ it is increasingly unlikely that j_k captured the original value of s_k and the sensitivity of the similarity measure with respect to $i_k^{\uparrow}$ decreases.

Next we will examine the model in (2.8) and show how the minor addition of $M_I(\mathbf{x})$ and $M_I(\mathbf{x})$ increase its robustness in the presence of shadows. The original model, to which we added these terms, was shown to outperform the sum of squared differences due to the ultrasound specific nature of multiplicative Rayleigh noise [17],[83]. Also the authors of the metric take into account the log compression of ultrasound images. Starting with the model in (2.8) the natural logarithm is taken on each side results in

$$\begin{aligned} \tilde{I}(\mathbf{x}) &= \tilde{M}_I(\mathbf{x}) + \tilde{S}(T_I(\mathbf{x})) + \ln \eta_I \\ \tilde{J}(\mathbf{x}) &= \tilde{M}_J(\mathbf{x}) + \tilde{S}(T_J(\mathbf{x})) + \ln \eta_J \end{aligned} \tag{2.16}$$

where the superscript tilde signifies the natural logarithm operation, for example $\tilde{I}(\mathbf{x}) = \ln I(\mathbf{x})$. As before, the goal will be to derive an equation for the likelihood in (2.13), i.e. $\mathcal{L}\left(T; \tilde{I}, \tilde{J}\right) \approx \ln \mathrm{P}\left(\tilde{I}|\tilde{J}, T\right)$, which quantifies the similarity between $\tilde{I}\left(T^{-1}(\mathbf{x})\right)$ and $\tilde{J}(\mathbf{x})$ in the region of overlap. First we seek to calculate $\mathrm{P}\left(\tilde{i}_k^{\uparrow}|\tilde{s}_k\right)$ and $\mathrm{P}\left(\tilde{s}_k|\tilde{j}_k\right)$ just as we did in the previous example, since their expressions are required by equation (2.13) and will be used to derive $\mathrm{P}\left(\tilde{I}|\tilde{J}, T\right)$. The probability density function for the multiplicative Rayleigh noise [56] that is used in this ultrasound image formulation model is

$$\mathrm{P}(\eta) = \frac{\pi}{2}\eta \exp\left(-\frac{\pi\eta^2}{4}\right) \tag{2.17}$$

Using the fact that $\tilde{I}$, given $\tilde{S}$ and the transformation T, is simply a function of the random variable η, we can derive the following expressions for the desired probabilities,

$$\begin{aligned} \mathrm{P}\left(\tilde{i}_k^{\uparrow}|\tilde{s}_k\right) &= \frac{\pi}{2}\left(\frac{i_k^{\uparrow}}{m_{I,k}^{\uparrow} s_k}\right)^2 \exp\left(-\frac{\pi}{4}\left(\frac{i_k^{\uparrow}}{m_{I,k}^{\uparrow} s_k}\right)^2\right) \\ \mathrm{P}\left(\tilde{s}_k|\tilde{j}_k\right) &= \frac{\pi}{2}\left(\frac{j_k}{m_{J,k} s_k}\right)^2 \exp\left(-\frac{\pi}{4}\left(\frac{j_k}{m_{J,k} s_k}\right)^2\right) \end{aligned} \tag{2.18}$$

Where $\tilde{i}_k^{\uparrow} = \ln I\left(T^{-1}(\mathbf{x}_k)\right)$. Also, note that the probabilities in (2.18) are calculated using the original voxel intensity values, without the logarithmic compression applied in (2.16). Inserting the probabilities from (2.18) into (2.13) and the carrying out the integration will give the log likelihood expression representing the similarity metric.

This integration is shown below,

$$
\begin{aligned}
\mathcal{L}\left(T;\tilde{I},\tilde{J}\right) &= \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J} \ln\left[\int \mathrm{P}\left(\tilde{i}_k^{\uparrow}|\tilde{s}_k\right)\mathrm{P}\left(\tilde{s}_k|\tilde{j}_k\right)d\tilde{s}_k\right] \\
&= \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J} \ln\left[\int \frac{\pi^2}{4(s_k)^4}\left(\frac{i_k^{\uparrow}j_k}{m_{I,k}^{\uparrow}m_{J,k}}\right)^2 \exp\left(-\frac{\pi}{4}\frac{\left(fori_k^{\uparrow}\right)^2+\left(m_{I,k}^{\uparrow}j_k\right)^2}{m_{I,k}^{\uparrow}m_{J,k}(s_k)^2}\right)d\tilde{s}_k\right] \\
&= \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J} \ln\left[\int \frac{\pi^2}{4(s_k)^5}\left(\frac{i_k^{\uparrow}j_k}{m_{I,k}^{\uparrow}m_{J,k}}\right)^2 \exp\left(-\frac{\pi}{4}\frac{\left(m_{J,k}i_k^{\uparrow}\right)^2+\left(m_{I,k}^{\uparrow}j_k\right)^2}{m_{I,k}^{\uparrow}m_{J,k}(s_k)^2}\right)ds_k\right] \\
&= \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J} \ln\left[2\left(\frac{m_{J,k}i_k^{\uparrow}}{m_{I,k}^{\uparrow}j_k}\right)^2\left(\left(\frac{m_{J,k}i_k^{\uparrow}}{m_{I,k}^{\uparrow}j_k}\right)^2+1\right)^{-2}\right]
\end{aligned}
\tag{2.19}
$$

Determining the optimal rigid transformation for the model described in (2.16) is done by maximizing the likelihood equation in (2.19) with respect to the transformation. Since we would rather calculate the transformation from the source's coordinate system to volume I's coordinate system, and thus deform I into S, the optimization will be done w.r.t. the inverse of T designated as T^{-1}. This becomes important when more than 2 volumes need to be registered to the common source coordinate system. For numerical stability it should be noted that $m_{I,k}^{\uparrow}, m_{J,k}$ must be > 0 in the area of overlap evaluated by (2.19). Once again if we assume that $m_{I,k}^{\uparrow}, m_{J,k}$ are Boolean valued masks, where $m_{I,k}^{\uparrow}, m_{J,k} = 1$ indicates a valid uncorrupted voxel intensity, then (2.19) reduces to simply calculating the CD2 metric in the area of overlap determined to have no shadow artifacts. Equation (2.19) enables us to use of a range of mask values, which weight a voxel's contribution to the similarity metric by how confident we are in its intensity value. Using the hyperbolic cosine identity $\cosh(x) = \frac{e^x+e^{-x}}{2}$, the maximization of (2.19) can be rewritten as the minimization of the following expression,

$$\begin{aligned}\mathcal{L}\left(T^{-1};\tilde{I},\tilde{J}\right) &= \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J} \ln\left[\cosh\left(\ln\left(m_{J,k}i_k^{\uparrow}\right)-\ln\left(m_{I,k}^{\uparrow}j_k\right)\right)\right] \\ &= \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J} \ln\left[\cosh\left(\ln\left(i_k^{\uparrow}\right)-\ln\left(j_k\right)+\ln\left(m_{J,k}\right)-\ln\left(m_{I,k}^{\uparrow}\right)\right)\right] \\ &= \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J} \ln\left[\cosh\left(\tilde{i}_k^{\uparrow}-\tilde{j}_k+\tilde{m}_{J,k}-\tilde{m}_{I,k}^{\uparrow}\right)\right]\end{aligned} \tag{2.20}$$

This formulation makes explicit differentiation of (2.20) w.r.t the parameters of T^{-1} simple, enabling the use of efficient gradient based optimization algorithms. We are essentially registering the volumes produced by taking the natural logarithm of the original data thus in (2.20) we only need to take the logarithm of the image data once, when the registration initially starts. The first order derivative of (2.20) can be calculated as follows,

$$\frac{\partial}{\partial\mathbf{\Theta}}\mathcal{S}\left(T^{-1};\tilde{I},\tilde{J}\right) = \sum_{\mathbf{x}_k\in\mathcal{S}_I\cap\mathcal{S}_J}\left[\nabla\tilde{I}\left(\mathbf{y}\right)-\nabla\tilde{M}_I\left(\mathbf{y}\right)\right]\Big|_{\mathbf{y}=T^{-1}(\mathbf{x}_k)} \tanh\left[\tilde{i}_k^{\uparrow}-\tilde{j}_k+\tilde{m}_{J,k}-\tilde{m}_{I,k}^{\uparrow}\right]\frac{\partial T^{-1}}{\partial\mathbf{\Theta}} \tag{2.21}$$

where we use the convention that the gradient is laid out as a row vector and $\mathbf{\Theta}$ represents the parameters of the transformation. The Quasi-Newton optimization algorithm that we employed only required the first order derivative for acceptable results; however second order methods can be readily be derived using (2.21). Also, direct-search methods such as the Nelder-Mead algorithm work well due to the low dimensionality of rigid registration and the efficient similarity metric in (2.20). For the remainder of the chapter the metric in (2.20) will be referred to as CD2 with shadow masks (CD2+S) stemming from the addition of shadow information to improve its robustness.

We primarily use this registration method for the rough alignment of clinically acquired, partially overlapping ultrasound volumes and so a rigid alignment is employed, which means that T^{-1} is just a 4x4 transformation matrix. The six degrees of

Figure 2-1: Source image used for 2D synthetic experiments

freedom can remove most of the misalignment due to patient shifting during scanning. The group-wise registration algorithm proposed to correct non-rigid deformation will be discussed in the following chapter.

2.6 Registration results

2.6.1 Experiments using 2D synthetic data

Initial registration experiments were conducted using synthetic images, which were produced from a user supplied source image that was corrupted with noise/shadow artifacts based on the image formulation model described by (2.8). The source image we used is shown in Figure 2-1. The images corresponding to I and J, which will be registered, are shown in Figure 2-2. Significant noise and shadow artifacts have been added to the source image from Figure 2-1 in order to test how sensitive the SSD and CD2 similarity metrics are to the variable conditions found in a clinically acquired data. Figure 2-3 shows the hand drawn masks that correspond to the image pair shown above. It is somewhat unrealistic to assume that this information is known; however, we demonstrate in the next section that an estimate is often good enough to provide significant improvements to the capture range of the rigid registration algorithm. In order to add an element of uncertainty to the confidence masks we redrew them after the images in Figure 2-2 were generated, thus they are somewhat different than the masks which produced the pair and can be thought of as an estimate. These

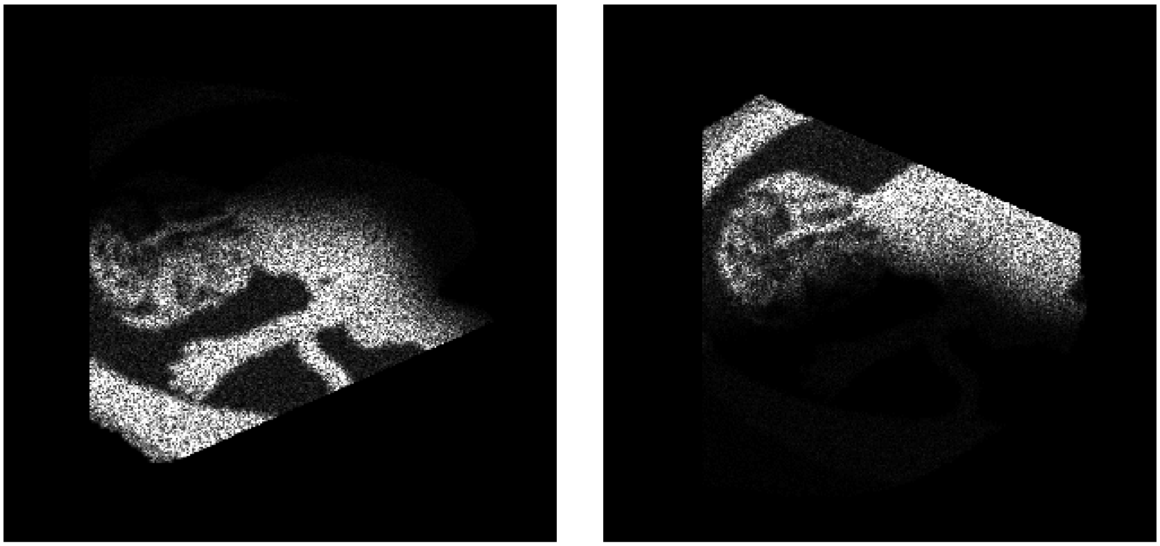


Figure 2-2: Shadow and noise artifacts added to source to produce registration images

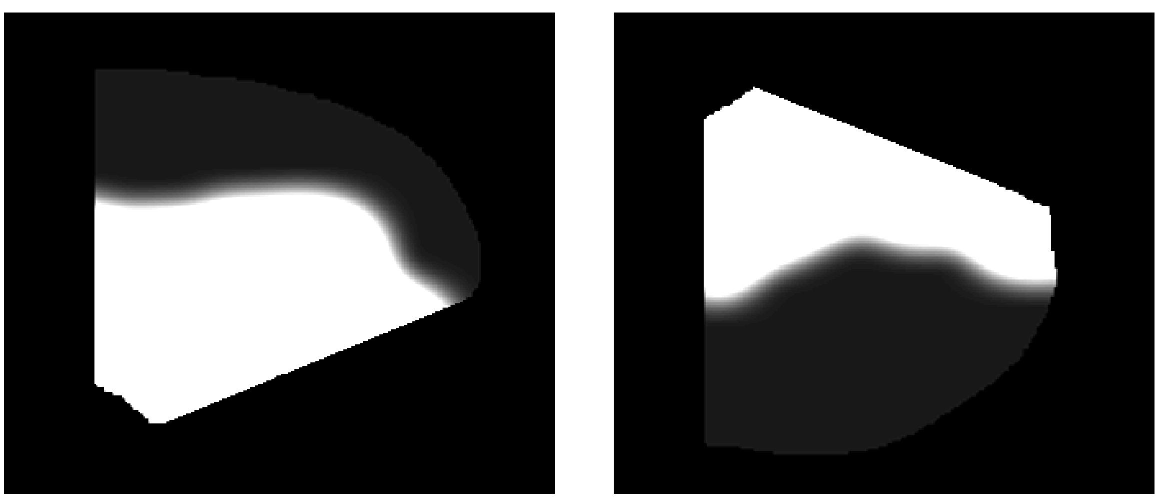


Figure 2-3: Estimated ultrasound transmission masks

images have been purposely synthesized with the intent to simulate the issues we encountered while registering 3D swept ultrasound volumes which have been acquired on opposite sides of the abdomen using transverse slices of a pregnant subject. In our clinical experiments the opposing volumes exhibit strong shadowing on the far side of the fetus due to its acoustic properties. Most of the acoustic power is lost as the ultrasound signal travels through the fetus, thus one volume will contain plenty of detail regarding the positioning of the arms and legs, while the volume acquired from the opposite side won't show these extremities at all. Instead, this volume will have a clear view of the vertebrae which implies that a combination of these two views is ideal. The expectation is that this combination will generate a high quality training volume, which contains realistic images of the fetus from both directions. When this composite volume is incorporated into the training simulator the user can identify structures in each side of the abdomen and is free to place the transducer in any position/orientation they wish. We believe that the synthetic images in Figure 2-2 provide a close enough approximation to the issue we just described to be used in our initial experiments, which will give us some insight into the effectiveness of (2.20).

In order to test the sensitivity of the similarity metrics we apply simple transformations, such as translation and rotation, to the pair in Figure 2-2 and subsequently plot the metric's value versus its transformation parameter. Ideally there should be a clear maximum with few local minimums for the optimizer to get stuck in. Also it should be noted that the similarity metrics are normalized by the amount of overlap, since this changes significantly as they are transformed.

The plot in Figure 2-4 shows how the similarity metric varies as we move the pair out of alignment using a translation transformation. Since the image pair is initially aligned perfectly the maximum in each similarity plot should occur at $(0, 0)$, meaning no translation has been applied in any direction. It is obvious from Figure 2-4 that registration using either SSD or CD2 as a similarity metric in this situation is impossible. No maximum can be identified in either of these plots, implying that any optimization method used would just eventually push the image pair completely apart. The region around $(0, 0)$, where a maximum in the similarity should be obvious, is barely discernible from the surrounding values. SSD and CD2 are equally bad in this case. The most interesting plot in Figure 2-4 corresponds to the recently introduced measure from (2.20). The correct global maximum at $(0, 0)$ is easily distinguished in this plot and with no apparent local minima any optimization method should be successful here. This initial experiment suggests that accounting for shadows in the similarity metric can have a large impact on the registration accuracy of images with substantial artifacts.

Our next experiment measured the sensitivity of the similarity measures to purely rotational transformations. Image I was rotated from -6° to $+6^\circ$ around a point positioned adjacent to the fetus's chest. A plot showing rotational misalignment versus similarity was produced for SSD, CD2, and CD2+S. Interpolation algorithms are especially sensitive to purely rotational transformations and methods such as bilinear interpolation gave poor results when evaluating the similarity metric at varying degrees of rotation. The resulting similarity plots using bilinear and cubic interpolation were very noisy due to the smoothing effect that small rotations have on the resampled image. This occurs because the noise properties of the images, which the similarity metrics are based on, are distorted due to the smoothing effect. Interpolation methods such as Lanczos resampling do a better job of maintaining the original image's noise properties, increasing the smoothness of the registration energy function. Nearest neighbor interpolation is the simplest method which maintains a majority of the noise after interpolation and for this reason we used it in our rotational experiments. The results of the rotational experiment for all 3 similarity metrics, using the image pair from Figure 2-2, are shown in Figure 2-5. The maximum of the similarity measure should occur at 0° since the images were already in alignment. It is apparent from Figure 2-5 that the only metric which would allow efficient/accurate optimization in this instance is CD2+S. The SSD metric is very noisy compared to CD2+S, which means that it requires more sophisticated optimization methods for convergence, but most importantly the SSD metric gives an incorrect global optimum around .02 radians. If the rotational misalignment of the images happen to be negative then there's a chance for convergence to the correct solution of 0° if a gradient based algorithm was

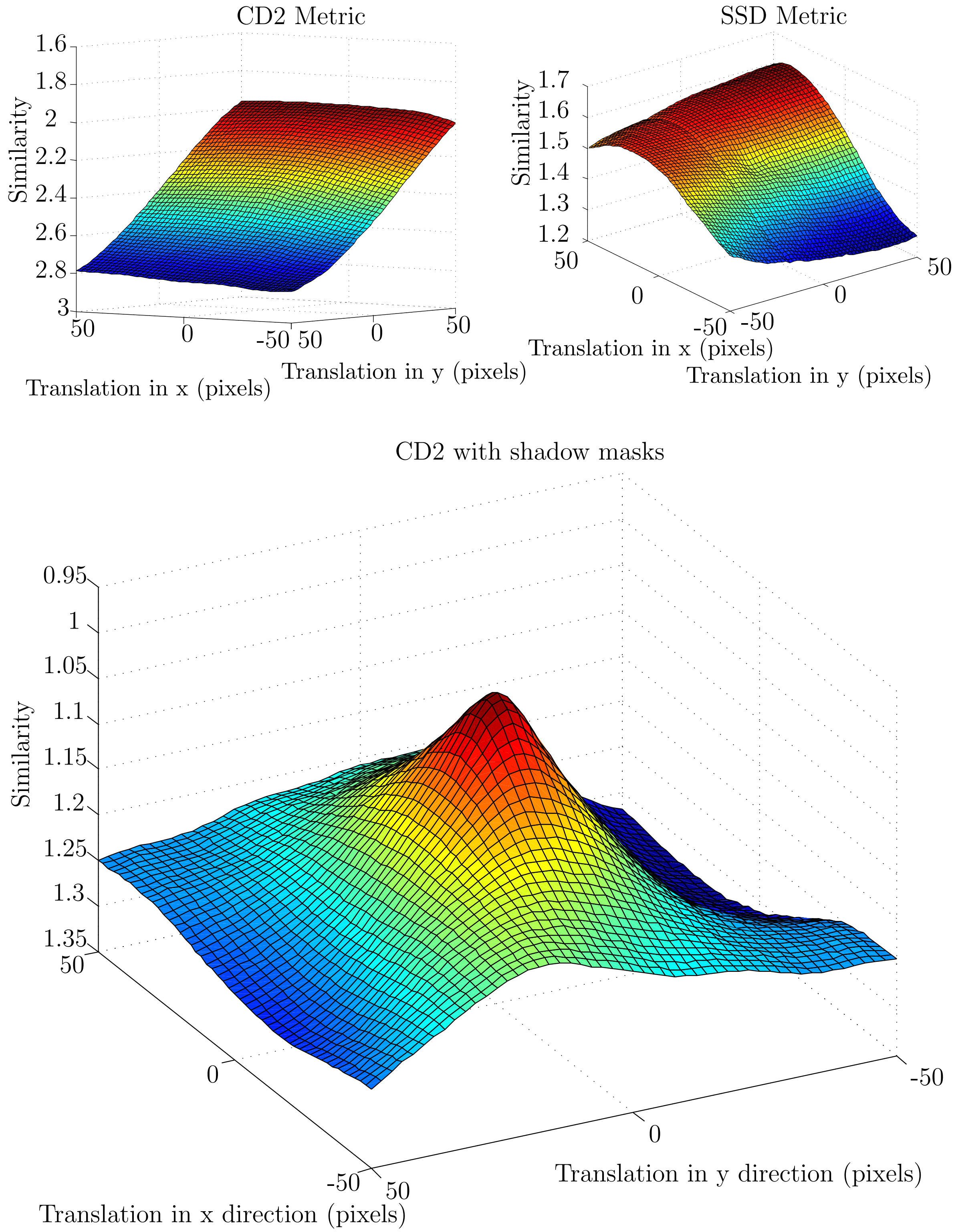


Figure 2-4: Plots showing the inability of standard similarity measures (SSD,CD2) to undue translational movement between synthetic images. The bottom plot shows the similarity energy versus translation misalignment for the new CD2 metric, which incorporates shadow detection. The correct global maximum is clearly discernible in this case, especially when compared to the plots of SSD/CD2 found in the first row.

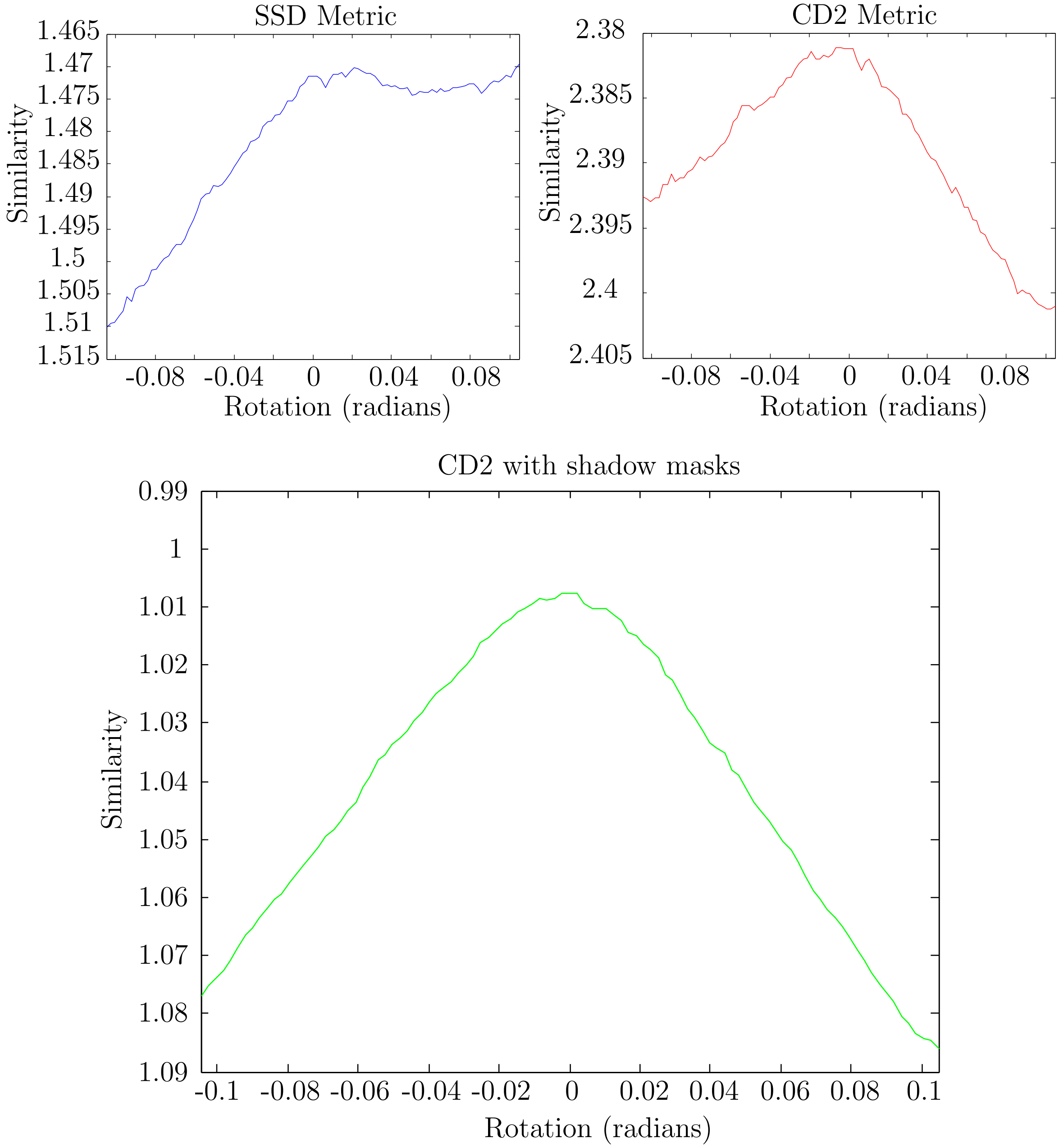


Figure 2-5: Similarity measure experiments for rotational misalignment. CD2+S solidly beats SSD and appears smoother than it's counterpart CD2.

chosen; however, based on Figure 2-5, any hope of convergence to the correct solution is gone if the initial misalignment is greater than $0°$. Taking a closer look at the plot for CD2 we notice the positive effect of accounting for multiplicative Rayleigh noise. The incorrect global optimum, which the SSD metric had around .02 radians, is gone but the plot of the registration similarity still displays a significant amount of noise. These small local minima seen in this plot present a problem for gradient based optimization methods. Another advantage of using CD2 over SSD for this problem is that for positive angles of misalignment the CD2 plot looks more conducive to optimization when compared to the analogous range of the SSD plot. Finally it is clear that for this experiment CD2+S is the winner. The plot of registration similarity is much smoother than the competing methods, with few if any local minima, which makes optimization much easier. One flaw, which no metric overcame, was a plateau around the correct global maximum that can be attributed to the intentional inaccuracy of the shadow masks. Results from the experiment using rotational misalignment agree with the results from our experiment using translational misalignment so we conclude that CD2+S passes its initial assessment.

The investigation of (2.20) using simple synthetic data suggests that it is worth the extra CPU time to estimate per-voxel ultrasound transmission levels, since their incorporation into the registration algorithm offers improved accuracy and increases its capture range. The next section will demonstrate this concept using reconstructed swept 3D ultrasound volumes of an abdominal trauma phantom.

2.6.2 Experiments using abdominal trauma phantom

In this section the improved registration result we achieved on synthetic data from an ultrasound phantom, after incorporating the confidence measure into our model, is validated using real ultrasound data from a trauma phantom. The FAST/Acute abdomen phantom, which was borrowed from Kyoto Kagaku company, simulates the presence of free intra-peritoneal fluid in trauma patients. The phantom was taken to UMass Medical School where it was scanned using a Phillips iU-22 ultrasound machine with a C5-1 transducer, which provides excellent abdominal image quality.

Swept 3D volumes were created by scanning the phantom along multiple linear paths from superior to inferior, acquiring transverse images along each scan path. In our experiments we choose two partially overlapping volumes which contained several distinctive features in the region of overlap but were also degraded by shadowing artifacts common in abdominal imaging.

In the first sequence of tests the two volumes were initially registered using the similarity metric in (2.20), whose result was verified upon careful examination. The newly transformed volume would serve as the ground truth and assumed to be perfectly aligned with its neighbor. Next the similarity values for CD2 and CD2+S were plotted as the volumes were forced out of alignment using a translation. Since the volumes are initially aligned, the similarity value for (0,0)cm, corresponding to no misalignment, should be the global maximum. These results can be seen in the top row of Figure 2-6 and demonstrate the advantage CD2+S has over CD2. The plot corresponding to CD2 has a local minimum at (0,0)cm, which allows correction in the range of translational misalignments that would converge to this point. The problem with this plot is that the global maximum occurs at the incorrect transformation and is located around (-4,-4)cm. If the initial misalignment is somewhere near (-2,-2)cm then the optimization algorithm would actually push the volumes further apart. This issue can be corrected by using CD2+S, whose improved similarity function is plotted in the top right of Figure 2-6. We can see a clear global maximum at the correct location and a much more realistic surface plot. As the volumes are moved out of alignment the similarity drops steeply in all directions. In the next experiment we used the same methodology but replace the translation with a rotation. The results are shown in the bottom half of Figure 2-6. Once again the similarity plot of CD2+S is substantially better compared to the original CD2 metric. The plot for CD2 appears noisy with no local minimum at 0°, thus there is limited hope for convergence unless the initial misalignment is minor. CD2+S once again addresses this issue, providing the optimization algorithm an ideal surface containing a global maximum at 0°.

The goal of the next experiment was to test the reliability of the similarity metrics on misalignments including both rotational and translation components. This was

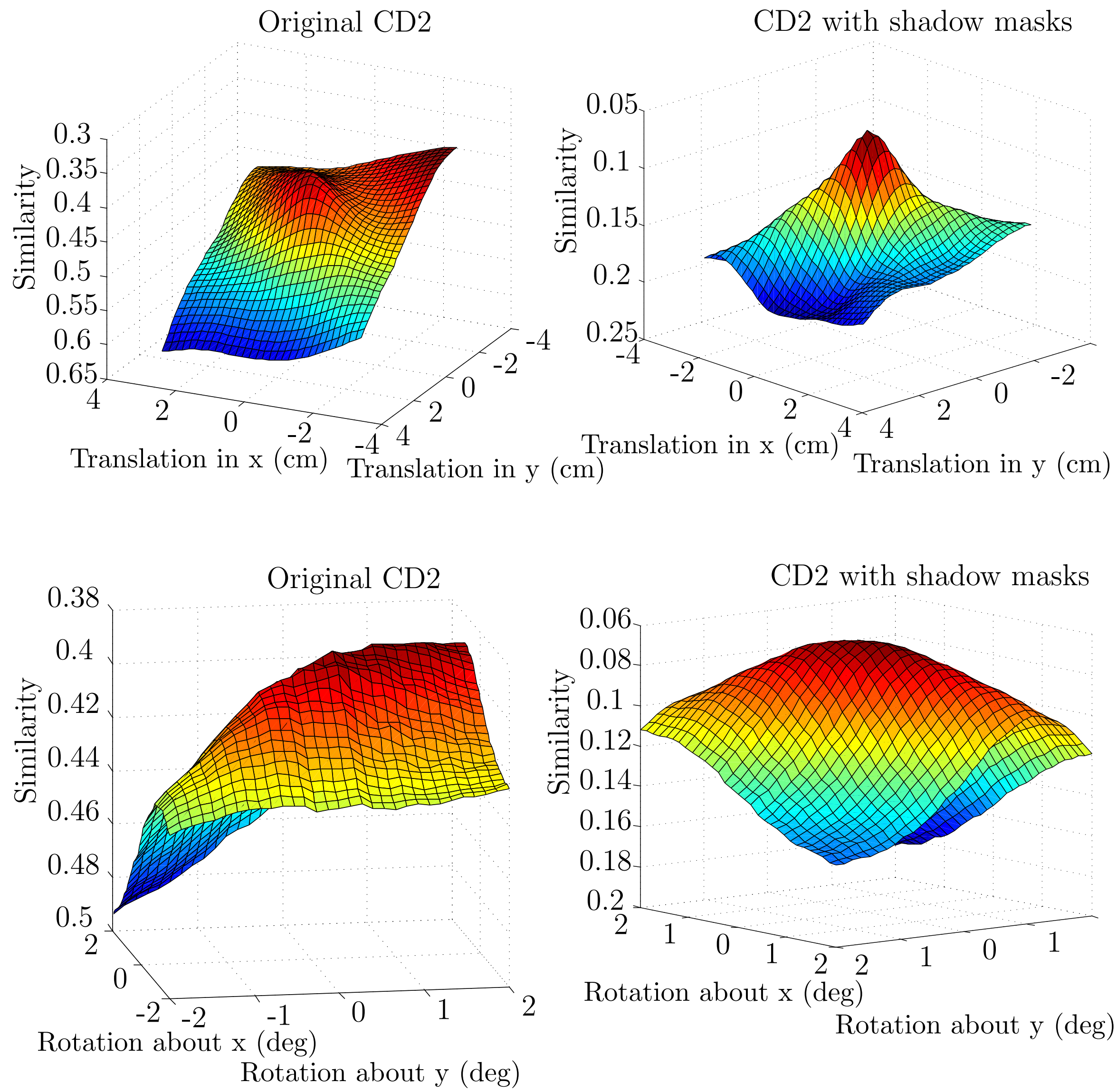


Figure 2-6: Results using abdominal trauma phantom dataset comparing CD2 to CD2+S for 3D registration of overlapping volumes. Translation and rotation are considered with CD2+S outperforming standard CD2 in the experiments.

done by applying an arbitrary rigid transformation to one of the volumes in the pair and subsequently attempting to undue the misalignment. The CD2 and CD2+S metrics were optimized using the Gauss-Newton algorithm, utilizing the gradient given by (2.21). Results from one iteration of the test are given in Figure 2-7. The initial misalignment contained a $2°$ rotation around the x-axis and a .58cm translation along the z-axis. In order to demonstrate the improvement in registration accuracy of CD2+S over CD2 we compared the visual quality of the composite volumes produced after rigid registration using each algorithm. The composite volumes were constructed by taking the mean at each voxel location in the area of overlap. The two columns of Figure 2-7 are comprised of three orthogonal planes taken from the composite volumes produced using registration results from CD2 and CD2+S respectively. The images in the first column of Figure 2-7 clearly show that CD2 failed to converge to the correct solution. The blurry grey outlines of misaligned structures are evident in the first column of images when compared with the optimal reconstruction shown in the second column, which was completed using CD2+S.

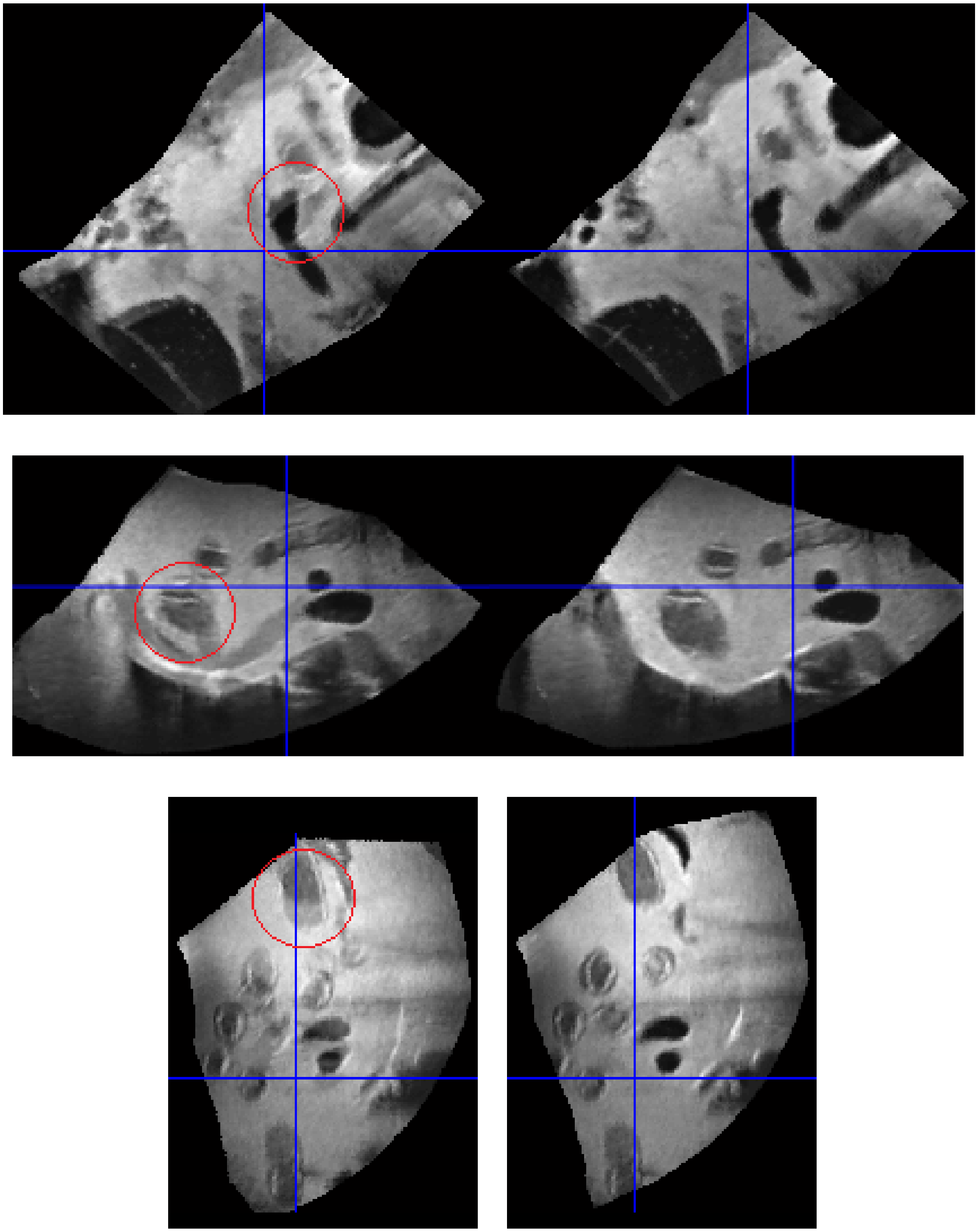


Figure 2-7: Reconstruction results after registration using CD2 and CD2+S. First column corresponds to CD2 where the circles highlight areas that appear incorrectly because CD2 failed to align the volumes. Second column shows results for CD2+S, which has correctly registered the volumes.

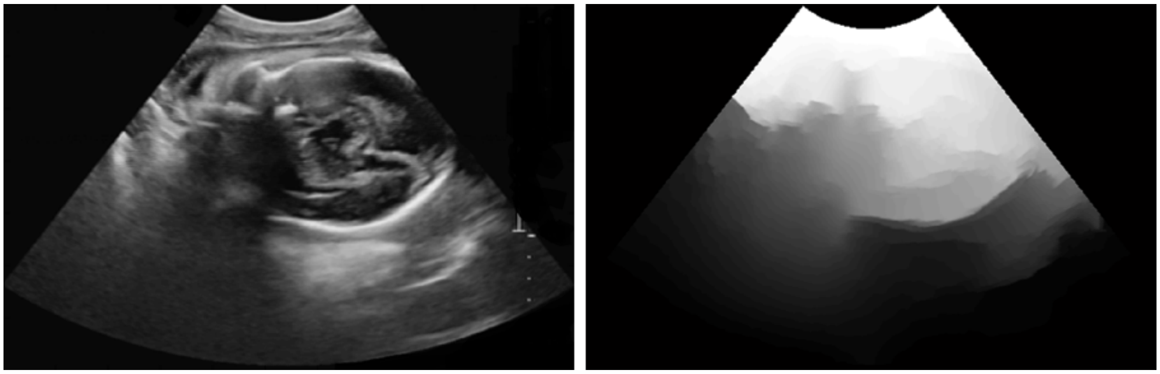


Figure 2-8: Example of shadow mask used in the registration of fetal ultrasound volumes

2.7 Conclusions

In this chapter we have shown that accounting for shadows during US-US registration can result in gains in accuracy and capture range. Convincing results suggesting the use of CD2+S over CD2 were presented for both synthetic and real ultrasound data. We employed the similarity metric of (2.20) in the rigid registration step of the mosaicing procedure, which was then applied to produce the fetal reconstructions that are presented in this dissertation. Since the shadow masks could most likely be calculated in real-time with improved implementations of the algorithms available, they should be included when doing intensity based US-US registration. In 2-8, the shadow mask calculated using the approach developed in [34] is shown for an ultrasound image of the fetal skull, which was acquired in a clinical setting. This example image was taken during the construction of one of the training volumes discussed in Chapter 7. The high intensity regions of the shadow mask in 2-8 correspond to the fetal skull, which is intuitive because the ultrasound echoes are strong here. Likewise, in the heavily shadowed region of the ultrasound image, the mask approaches 0 due to the lack of anatomical features. The similarity measures developed here could also be easily adapted to the group-wise non-rigid registration algorithm presented in Chapter 3 due to its block-based approach.

Chapter 3

Group-wise registration and seam selection: theory

In this chapter we present a group-wise non-rigid registration/mosaicing algorithm based on block-matching, which is developed within a probabilistic framework. The discrete form of its energy functional is linked to a Markov Random Field (MRF) containing double and triple cliques which can effectively be optimized with current MRF optimization algorithms popular in computer vision. Also, the registration problem is simplified by introducing a mosaicing function which partitions the composite volume into regions filled with data from unique partially overlapping source volumes.

3.1 Introduction and prior work

We will develop a novel group-wise image registration/stitching algorithm capable of producing composite volumes with structural continuity from many partially overlapping 3D sources. The definition of a discrete energy functional, used to partition the composite volume into regions corresponding to each unique source, is one contribution of this work and will be referred to as a mosaicing function. The second contribution is the development of a group-wise block matching algorithm that is derived using a probabilistic framework and enhanced with the precomputed mosaicing function defined later. The motivation for this research was to produce a composite obstetrics volume for use in an ultrasound training simulator, where the image volume encompasses the mother's entire abdomen. Unique challenges include fetal movements and extensive shadowing, which caused volumes acquired from one side of the abdomen to have few features in common with volumes acquired from the opposite side.

This work builds on the research of [82], where rigid and non-rigid group-wise registration techniques were developed for ultrasound using a probabilistic framework. In [83] rigid registration strategies were developed for 3D ultrasound mosaicing, and in [82] temporal group-wise non-rigid registration was applied to model liver motion in 4D. Our work links the accumulated pairwise estimates framework, developed for group-wise registration, to a discrete Markov Random Field and formulates an effective method to deal with many partially overlapping volumes.

Previous work on extended field of view ultrasound includes the registration techniques developed in [63] where a set of volumes acquired from a 3D ultrasound machine was stitched together by linking position information from a tracking system to each 3D volume. Non-rigid registration was performed and the volumes were compounded by averaging pixel values in overlapping regions. *In vitro* experiments were performed by stitching together volumes obtained from a fetal phantom. The composite volume is grown by using pair-wise registration techniques and sequentially adding adjacent volumes until all volumes have been incorporated.

Similarity metrics are used to guide registration algorithms, and those which have been designed with specific modalities in mind can boost the accuracy of the resulting deformation fields [3]. For example, methods designed for ultrasound have taken into account the unique statistical properties of signal noise and also the speckle correlation between volumes acquired close to one another in time. An intensity based metric, which minimizes the residual complexity between two images, was presented in [51] and was shown to be superior to mutual information. Other metrics [47] incorporate local phase information because it is argued that it can provide more local structural information than intensity. In [52] an improved rigid registration technique using a 3D scale invariant feature transform was developed. These methods have been shown to outperform traditional metrics such as mutual information, or sum of squared differences in experimental settings but they are costly to compute and don't account for the shadow artifacts which are prevalent in clinical ultrasound. An efficient similarity measure, which can handle large deformations between many partially overlapping volumes (≥ 3), is desired. To this end we choose instead to implement an improved sum of squared differences metric capable of dealing with shadowing artifacts, which dominated the registration error when using clinical data acquired from live subjects. Our metric can be efficiently calculated in the frequency domain and effectively handles obscured regions.

Image registration using models based on MRF theory has been proposed by researchers in the past and shown to produce satisfying results. In [25], [75] a pair-wise hybrid geometric/iconic algorithm based on MRF theory was proposed and validated using lung CT data. A group-wise registration method based on MRFs was presented in [74] and tested on 2D MR human skeletal muscle images. The previously cited discrete methods use first order derivatives to regularize the transformation fields, which we found to be problematic when dealing with partially overlapping volumes. In [38] a higher order regularization term based on second derivatives was proposed and used to perform pair-wise 2D registration. In [39] this higher order smoothness term was extended to 3D and used in the pair-wise registration of MR volumes. We extend this discrete second order smoothness term for use in a group-wise setting.

Our work is unique in that, starting from a probabilistic framework, we develop a pair of MRF energies for the joint group-wise registration/mosaicing problem and then show how they may be efficiently optimized using current computer vision approaches. In contrast to previous efforts we consider the mosaicing function as input to our registration method in order to focus the similarity metric on the most influential regions. This novel approach has been specifically designed for the group-wise registration of many partially overlapping volumes. Choosing an optimal mosaicing function reduces unwanted image artifacts in the final composite volume as well as simplifies the group-wise registration problem, leading to increased computational performance. Existing compounding methods require precise alignment between volumes, which is not achievable with freehand 3D fetal ultrasound, making our method a feasible alternative. For the sake of speed we have chosen to implement an improved version of the sum of squared differences (SSD) similarity metric, where our modification increases the metric's robustness in the presence of shadows. Accounting for these ultrasound specific artifacts is novel in the context of group-wise non-rigid registration. This gives us the ability to quickly compute all the similarity information needed for the entire registration process efficiently by using the FFT method derived in this chapter, which occurs before any optimization is performed. Using the FFT for block matching was previously proposed in [55] for pair-wise multi-modal rigid registration and in this chapter we extend the method for non-rigid registration in a group-wise setting, taking into account ultrasound specific shadowing artifacts. The data required for the optimization of the MRF is precomputed and stored for easy access by whichever optimization algorithm is chosen. Due to the pairwise nature of the similarity metric coupled with its fast pre-computation we can include additional volumes with minimal effort. Existing similarity metric data is reused for the group-wise registration problem involving additional volumes. In addition we present some results about the parallelization of image registration through deformation field fusion of independent solutions.

3.1.1 Initial trials and algorithm overview

The movements and deformations associated with capturing multiple obstetrics ultrasound image volumes from subjects in a clinical setting means that non-rigid registration plays an important role in producing a seamless composite image volume. In our preliminary research we performed pairwise non-rigid registration/stitching between neighboring volumes and produced a final composite volume by either following this procedure in a pyramid fashion or sequentially adding a source to the composite volume one by one. A more intuitive and satisfying result can be obtained by considering the deformation of all volumes simultaneously using group-wise registration techniques. The advantage of using group-wise registration is that displacement fields produced after optimization, which define the non-rigid transformations necessary to bring the source volumes into alignment within a common coordinate system, are linked together so that the optimization will produce a result that doesn't favor deformation of certain volumes over others.

All source volume transformations are linked together within a discrete summation of terms resulting from a Markov Random Field (MRF) model of the registration problem, which is optimized using graph based techniques. This MRF function contains two types of terms. The first term measures the alignment between the overlapping volumes in each region. We define these regions as 3D blocks so that the registration can be thought of as a weighted block matching algorithm. The next term measures the elastic properties of each displacement field using a discrete 2nd derivative. The elastic regularization term prevents the block matching term from producing physically unrealistic deformations. We found that a 2nd order term is required due to the fact that a 1st order term caused undesirable shearing effects when using registration masks. In addition the 2nd derivative based regularization is invariant to linear transformations such as rotation. The complete MRF function representing this registration problem contains components dependent on double and triple cliques, which simply means that terms are dependent on 2 or 3 discrete variables. These discrete variables are assigned labels using the terminology from the MRF optimization literature.

In our problem the label set contains all the possible displacements in a limited 3D window. The displacement window can be calculated to ensure that the final transformation remains diffeomorphic. This means the transformation is invertible or free of folds which aren't physically possible. Hierarchal group-wise registration is performed in an iterative fashion, which relaxes the elastic regularization between iterations so that it becomes fluid-like. The terms in the group-wise registration MRF are not submodular so we must use advanced optimization techniques to find a solution. Non-submodular energies with discrete arguments, such as the one formulated in this chapter, can be minimized using techniques for Quadratic Pseudo Boolean Optimization (QPBO). The Boros, Hammer and Sun (BHS) algorithm [36] was developed for this purpose and it is a key component in our registration algorithm. We use the alpha expansion technique to optimize each displacement field while taking advantage of the BHS algorithm instead of a basic graph cut for each expansion step. The algorithm is implemented in a computationally efficient way by splitting the computation up to different CPU cores using a technique called QPBO fusion which was first introduced in the computer vision literature for the purpose of stereo vision. QPBO fusion breaks the registration problem into 8 separate registration problems where each considers a limited set of the solution space. At the conclusion these 8 solutions are fused together to produce one final solution that contains the best parts of each.

3.2 Group-wise registration in a probabilistic framework

The group-wise registration problem will first be described in a maximum likelihood framework. The reason for this is that a MRF is probabilistic by definition although this point-of-view is sometimes glossed over in the computer vision literature. If given N overlapping ultrasound image volumes, which are simply functions from three-dimensional space mapped to one-dimensional intensity $\mathcal{V} = \{I_1, ..., I_N\}$ where the mapping is $I_n : \mathbb{R}^3 \rightarrow \mathbb{R}$, our goal is to calculate N three-dimensional deformation fields that bring the volumes into alignment in the overlapping regions. Furthermore these registration regions can be intelligently defined using a mosaicing function which will be described in a later section. We define a scene coordinate space in $\mathbb{R}^3$ and seek to transform all N volumes into this space where the final mosaiced volume will reside. Let the set of transformations be defined as $\mathcal{T} = \{T_1, ..., T_N\}$ where $T_n : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Figure 3-1 demonstrates the group-wise registration concept in the context of aligning multiple overlapping 3D ultrasound volumes. In order to produce the composite volume in this example, transformation functions for each source must be optimized according to some matching criteria. The common coordinate system, which represents the space where the final mosaiced volume will be created, is shown

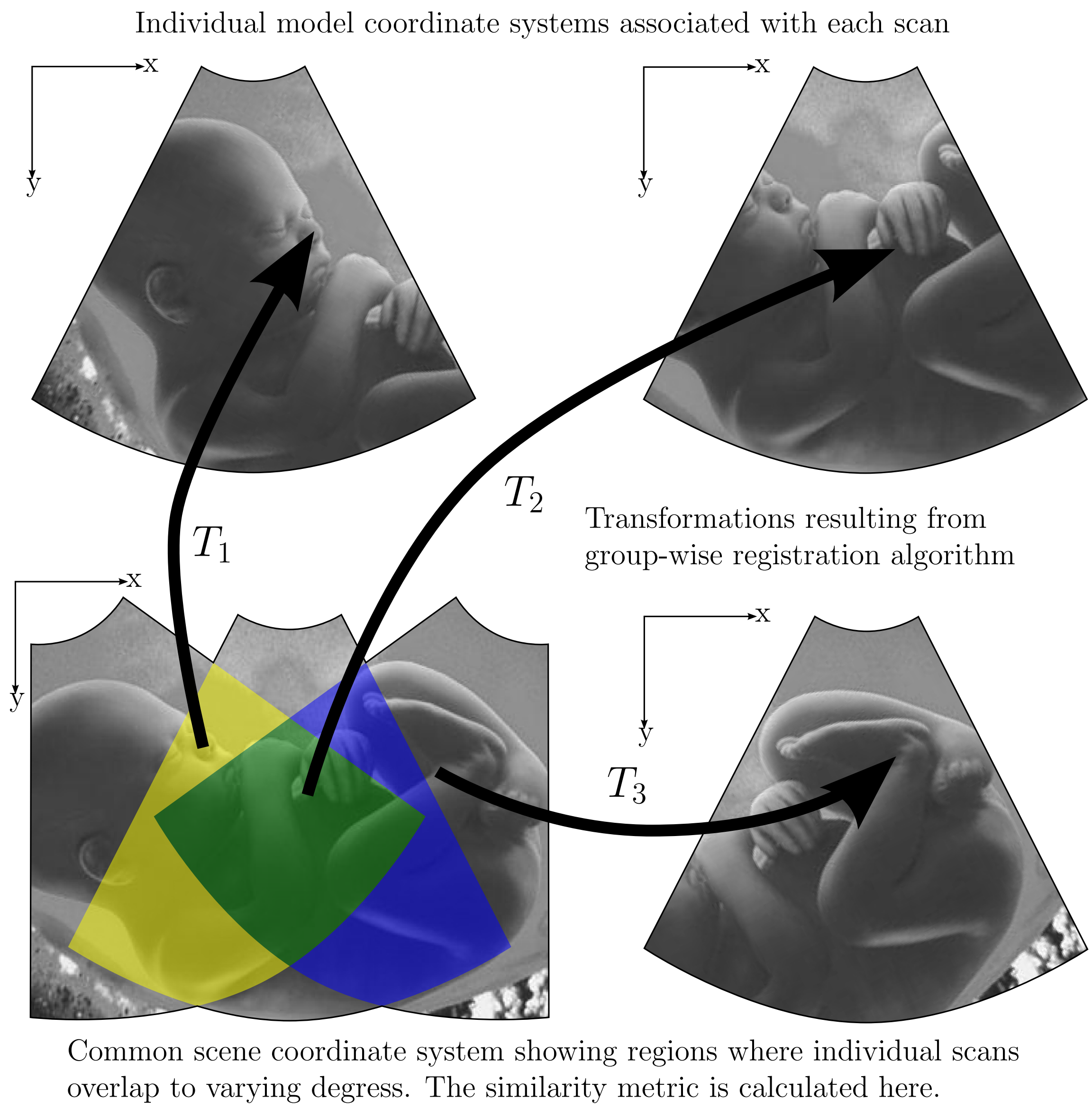


Figure 3-1: Group-wise non-rigid registration being used to mosaic multiple overlapping images

in the lower left corner. Optimized transformations are calculated from the common coordinate system to each of the individual volume coordinate systems where the optimization process evaluates the alignment in the regions of overlap. In this case the yellow/blue regions represent pairwise overlap, and the middle green region designates where all 3 source volumes overlap. The resulting problem is a fairly complicated group-wise registration which can be formulated as an energy minimization stated as

$$\hat{\mathcal{T}} = \underset{\mathcal{T}}{\operatorname{argmin}}\ \mathcal{E}\left(\mathcal{T}; I_1, ..., I_N\right). \tag{3.1}$$

In the equation above $\mathcal{E}$ represents the registration energy that we wish to minimize. A lower value of $\mathcal{E}$ indicates better alignment among the source volumes contained in $\mathcal{V}$, measured in the regions of overlap. The function $\mathcal{E}$ is usually split and written as a sum of two terms, $\mathcal{E} = \mathcal{M} + \mathcal{R}$, which represents the total energy to be minimized. This can be looked at as a combination of the image matching criteria and a deformation field regularization term, $\mathcal{M}$ and $\mathcal{R}$, respectively. The matching term measures alignment of image features and is referred to as a similarity metric while the regularization term penalizes unlikely transformations such as those containing very large first or second derivatives. We are minimizing (3.1) over the entire set of transformations.

Our strategy is to perform a series of group-wise translational alignments at a fixed number of control points, considering only the region of the source volume that is influenced by each control point. Control points are locations on a grid with uniform spacing which define the transformation of each source. Thus each source in $\mathcal{V}$ is given its own set of control points. This alignment process is iteratively repeated to give the final result. The registration procedure can be considered a group-wise block matching algorithm which can be analyzed in a maximum likelihood framework. In the following discussion the transformation model has not been defined, an assumption which has no effect on the validity of the equations.

Since we have reviewed the maximum-likelihood estimation (MLE) formulation for the pairwise registration problem in Chapter 2, where one volume remains fixed and the second volume is transformed into the coordinate system of the first, we will now describe how this can be extended to a group of volumes where all are deformed simultaneously in a group-wise registration framework. The following group-wise similarity metric is known as accumulated pair-wise estimates; it was derived in [83] and used for the rigid registration of multiple ultrasound volumes acquired from a fetal phantom. This approach was later extended to temporal deformable group-wise registration where a 3D ultrasound system was used to model liver motion [82]. We will see that this formulation can be directly linked to the discrete graph based MRF framework that was used for the efficient registration of multiple partially overlapping fetal image volumes in our experiments. Previously the term $\mathrm{P}\left(J|I,T\right)$ was used to describe the similarity between two image volumes given the transformation and can be found in the MLE formulation of equation (2.4). When deriving a similarity measure for the group-wise registration of multiple volumes the following likelihood equation can be used

$$\hat{\mathcal{T}} = \underset{\mathcal{T}}{\operatorname{argmax}} \ln \mathrm{P}\left(I_1, ..., I_n|\mathcal{T}\right). \tag{3.2}$$

In this equation the joint probability is over all possible image volumes and is conditioned on the set of transformations. $\mathcal{T}$ contains a transformation for each volume and the maximum likelihood framework seeks to find the set of transformations that best explains the observed image volumes by maximizing the probability in (3.2). In MLE the volumes are fixed to their observed values and $\mathcal{T}$ is varied until some maximum is reached. In this formulation one particular volume is not favored as being fixed while all others are transformed to its coordinate system. This avoids the issue of the algorithm choosing a fixed volume, which may be significantly out of alignment with the rest, while the remaining volumes are fairly close to being registered to each other. This issue would also cause the outlier volume to be forced to converge with the group instead of drawing the others toward it as in the case of choosing a fixed volume. In the derivation that follows all image volumes are assumed to be conditionally independent of each other thus given a realization of one image volume all other volumes are independent. This is reasonable if we view the other image volumes as being realizations of a random process which corrupts the given image volume. The first step in the derivation [82] is to rewrite $\mathrm{P}(I_1, ..., I_n|\mathcal{T})$ as the product of n conditional densities with respect to some image volume I_n. This can be done using the product rule and the property of conditional independence $\mathrm{P}(A, B|C) = \mathrm{P}(A|C)\,\mathrm{P}(B|C)$, and so we can write

$$\begin{aligned} \mathrm{P}(I_1, ..., I_n|\mathcal{T}) &= \mathrm{P}(I_1, ..., I_{n-1}|I_n, \mathcal{T})\,\mathrm{P}(I_n|\mathcal{T}) \\ &= \mathrm{P}(I_n|\mathcal{T}) \prod_{i=1}^{n-1} \mathrm{P}(I_i|I_n, \mathcal{T}) \,. \end{aligned} \tag{3.3}$$

In order to make the likelihood function symmetric with respect to all the volumes the n^{th} power of the joint density is taken and (3.3) is employed n times, iterating through the image volumes so that each takes a turn being the given or fixed volume, which results in

$$\mathrm{P}(I_1, ..., I_n|\mathcal{T})^n = \left(\prod_{i=1}^{n} \mathrm{P}(I_i|\mathcal{T})\right) \left(\prod_{i=1}^{n} \prod_{j\neq i}^{n} \mathrm{P}(I_j|I_i, \mathcal{T})\right). \tag{3.4}$$

Finally the logarithm is applied to the joint probability function in (3.4) resulting in an expression for the likelihood function $\ln \mathrm{P}(I_1, ..., I_n|\mathcal{T})$ which can be used in practice to evaluate the similarity of multiple overlapping image volumes. This is shown in the following equation:

$$\begin{aligned} \ln \mathrm{P}(I_1, ..., I_n|\mathcal{T}) &= \frac{1}{n} \sum_{i=1}^{n} \ln \mathrm{P}(I_i|\mathcal{T}) + \frac{1}{n} \sum_{i=1}^{n} \sum_{j\neq i}^{n} \ln \mathrm{P}(I_j|I_i, \mathcal{T}) \\ &\approx \sum_{i=1}^{n} \sum_{j\neq i}^{n} \ln \mathrm{P}(I_j|I_i, \mathcal{T}). \end{aligned} \tag{3.5}$$

From (3.5) we see that the higher order joint density representing the group-wise registration problem can be estimated using the sum of pairwise densities. Here we are ignoring the prior probabilities of observing each image and only concentrating on the conditional probability density functions. This is a nice result because it will form the theoretical foundation for our discrete MRF based formulation of the registration problem. Equation (3.5) intuitively states that if the pair-wise similarity in each possible combination of image volumes is calculated in the region of overlap and then summed together, it results in a group-wise measure of alignment. If we assume i.i.d. coordinate samples and functional/probabilistic intensity mappings between image volumes it is possible to derive group-wise similarity metrics based on popular pairwise metrics such as SSD, Correlation Ratio, and Mutual Information. For our work with mono-modal ultrasound registration, the speed advantage of group-wise SSD was chosen over the versatility of other measures, but there is no reason these could not be applied using the registration framework described in this chapter. Notice that the terms inside the summation of (3.5) take the same form as the similarity term from (2.3), which is just the maximum likelihood formulation of the pairwise registration problem. Each term only considers two volumes at once and thus if we make the same assumptions as we did to arrive at the sum of squared differences similarity metric in (2.6) we arrive at the group-wise SSD metric shown below

$$\ln \mathrm{P}\left(I_1, ..., I_n | \mathcal{T}\right) = \sum_{i=1}^{n} \sum_{j \neq i}^{n} \overbrace{\left[\sum_{x \in \Omega} \left(I_{j,k}^{\downarrow} - I_{i,k}^{\downarrow} \right)^2 \right]}^{\text{SSD Comp. Between Vols.}} \text{ where } \begin{cases} I_{i,k}^{\downarrow} = I_i\left(T_i\left(x_k\right)\right) \\ I_{j,k}^{\downarrow} = I_j\left(T_j\left(x_k\right)\right) \end{cases} . \quad (3.6)$$

It is important to note that (3.6) considers the SSD computation over the entire image and will need to be modified in order to be used in the group-wise block matching algorithm.

3.2.1 The transformation function

We choose to use a parametric transformation model where the parameters are the displacements of the control points in 3D space. The set of control points forms a sparse representation of the dense deformation field, and the number of points is usually far fewer than the number of image voxels. As briefly stated before and shown in Figure 3-1, the transformations must be from the common scene coordinate system to the individual model coordinate systems containing the sources. A transformation in the other direction would require scattered data interpolation to produce the deformed volumes. Following this we can construct the deformed volume voxel-wise by using the transformation function to calculate each voxel's original location in its respective source volume. The displacement of each voxel is computed as a weighted linear combination of control point displacements $\mathcal{D}_n = \{\mathbf{d}_{n,1}, ..., \mathbf{d}_{n,m}, ..., \mathbf{d}_{n,M}\}$ where $\mathbf{d}_{n,m} \in \mathbb{R}^3$:

$$D_n(x) = \sum_{m=1}^{M} w_m(x)\mathbf{d}_{n,m}. \quad (3.7)$$

In this expression n specifies the source volume, m identifies the control point which $\mathbf{d}_{n,m}$ corresponds to, and M is the total number of points. The value weighting function $w_m(x)$ represents the amount of influence that control point m has at location x. Using displacement function in (3.7) the transformation at each voxel can be calculated as $T_n(x) = x + D_n(x)$. Each volume to be registered has a unique grid of control points in the scene coordinate system, and the grids are initially co-located before the registration algorithm optimizes the position of their control points. Thus we are trying to determine the displacements in the set $\{\mathcal{D}_1, ..., \mathcal{D}_n, ..., \mathcal{D}_N\}$ where each $\mathcal{D}_n$ contains the control points that define the transformation of volume n. In medical image registration the weighting functions are often chosen to be B-splines and this type of free form deformation has been thoroughly studied. In our experiments the weighting functions were chosen from the tri-cubic convolution interpolation algorithm. Because these weighting functions force the deformation field at each control point location to take the exact value of the control point displacement, a larger step is taken during iterations of the algorithm. This is beneficial because the computational requirement to optimize N 3-dimensional displacement fields simultaneously is large.

The regularization of the transformation field is necessary in the case of general deformable registration or the problem is ill-posed; however it is not theoretically required by the block matching algorithm presented in this chapter due to control point displacements being limited to an arbitrary 3D window in the discrete formulation. The control point displacements should adhere to the constraints in [68], which result in a diffeomorphic transformation field after registration. It is still used to enforce smoothness on the deformation field in the case of partially overlapping ultrasound volumes. If two volumes are registered in the region of overlap it is important for the deformation field to smoothly extend into the non-overlapping region of the image volumes because there is no similarity measure to guide the registration process here. We impose regularization on the deformation field using second order as opposed to first order derivatives because this type of regularization is invariant to linear transformations; thus rigid alignments are not penalized during deformable registration.

3.3 Group-wise block matching as probability maximization

In this section a group-wise block matching algorithm will be described based on the probabilistic concepts discussed previously. Also the mathematical assumptions made during the development of the group-wise similarity measure presented in this chapter will be elaborated on. As discussed above in the context of group-wise image registration, each source volume is given a set of control points placed in the common scene coordinate system. All transformations are initially identity, i.e. $\mathcal{D}_n = \left\{ (0,0,0)^\top_{n,1}, ..., (0,0,0)^\top_{n,m}, ..., (0,0,0)^\top_{n,M} \right\}$ for each source volume n, and every location in the common scene's uniformly spaced grid has N coinciding control points, one for each source image. In addition to concepts already presented we propose that each source image volume I_n should be broken down into overlapping blocks, or sub images, which are centered on each grid point. The set of control point locations will be denoted as $\mathcal{C} = \{\mathbf{c}_b | \mathbf{c}_b \in \mathbb{R}^3 \wedge 1 \leqslant b \leqslant B\}$ where B denotes the total number of blocks or control points and will form a rectangular grid of uniform spacing h in the scene coordinate system. Let $I_{n,\{1,...,B\}} = \{I_{n,1}, ..., I_{n,b}, ..., I_{n,B}\}$ represent the set of B overlapping blocks which comprise image volume n and are centered on the control points. The location of each block in the scene coordinate system can be determined using the block index b. The number of blocks in the grid along the x, y, z dimensions is denoted by B^x, B^y, B^z, respectively. Also we will define a block as a cubic region centered on a control point whose sides are equal to $2h$. Individual voxels within a block will be indexed as offsets from the grid coordinates that the block is centered on using the set $\mathcal{B} = \{\mathbf{x} | \mathbf{x} \in \mathbb{Z}^3 \wedge -h \leqslant x_n \leqslant h\}$. The index of the control points increases fastest along the x dimension, followed by y and is slowest along the z dimension. Let the each $T_n \in \mathcal{T}$ be the transformation associated with image volume I_n which is defined by a set of displacement vectors corresponding to the control points associated with I_n and where the direction of the transformation is from the common scene coordinate system to the individual model system. Defining the transformations to be in this direction makes interpolating the model images in the common scene coordinate system simple. Using the notation introduced above the displacement vectors for source n will be indexed using the notation $\mathcal{D}_n = \{\mathbf{d}_{n,1}, ..., \mathbf{d}_{n,b}, ..., \mathbf{d}_{n,B}\}$ where $\mathbf{d}_{n,b} \in \mathbb{R}^3$ and $\mathbf{d}_{n,b} = \begin{bmatrix} d^x_{n,b} & d^y_{n,b} & d^z_{n,b} \end{bmatrix}$. The complete solution, consisting of all the displacement vectors for the N source volumes, is denoted as $\mathcal{D} = \{\mathcal{D}_1, ..., \mathcal{D}_n, ..., \mathcal{D}_N\}$. Simplified notation for the translation and formation of blocks will be helpful in providing a cleaner presentation. Thus transformed regions near control point $\mathbf{c}_b$ from image volume I_n will be represented using

$$\begin{aligned} I^{\downarrow}_{n,b} &= I_n\left(\mathbf{x} + T_n\left(\mathbf{x}\right)\right) \\ &= I_n\left(\mathbf{x} + \mathbf{c}_b + \mathbf{d}_{n,b} + NR_{n,b}\left(\mathbf{x}\right)\right) \\ &\approx I_n\left(\mathbf{x} + \mathbf{c}_b + \mathbf{d}_{n,b}\right). \end{aligned} \tag{3.8}$$

In (3.8) the deformation is modeled using a translational component and a non-rigid component. The $NR_{n,b}\left(\mathbf{x}\right) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ term models the non-rigid correction of the transformation and is included for completeness but is removed because the deformation field is estimated using block-matching. When it is included, $\mathbf{d}_{n,b} + NR_{n,b}\left(\mathbf{x}\right)$ represents the exact transformation of volume I_n. The symbol $I^{\downarrow}_{n,b}$ represents a transformed version of I_n where the downward arrow is used to signify translation. This transformation centers the region of volume I_n, surrounding control point $\mathbf{c}_b$, at the origin of a provisional coordinate system associated with block b, which is only used to calculate the similarity between image blocks of the set $I_{\{1,...,N\},b}$ and their search windows within the overlapping I_n. The translational component $\mathbf{d}_{n,b}$ is considered to be the exact deformation at the control point $\mathbf{c}_b$ and the term $NR_{n,b}$ corrects for non-translational movement away from $\mathbf{c}_b$. Thus as the control point grid becomes finer the non-rigid term has less of an effect.

The idea of the algorithm is that at each grid location a group-wise block matching is performed. If a grid location is common to many ($\geqslant 3$) overlapping volumes it will be shown that the sum of pairwise terms can serve as the group-wise measure of alignment for this set of blocks. In a registration algorithm based on block matching it is assumed that the deformation of the image volume can be locally described by the translation of a block in a larger search window. The smaller the blocks are the more valid this assumption is. The dense deformation field can then be calculated by fitting some function to the block displacement data; we choose to use a bi-cubic interpolation function based on equation (3.7); however free-form deformation splines would also work.

Using the recently introduced notation, the registration problem is stated as

$$\begin{aligned}\hat{\mathcal{T}} &= \underset{\mathcal{T}}{\operatorname{argmax}}\ \mathrm{P}\left(T_1,...,T_N|I_1,...,I_N\right)\\ &= \underset{\mathcal{T}}{\operatorname{argmax}}\ \frac{\mathrm{P}\left(I_1,...,I_N|T_1,...,T_N\right)\mathrm{P}\left(T_1,...,T_N\right)}{\mathrm{P}\left(I_1,...,I_N\right)}\\ &\approx \underset{\mathcal{T}}{\operatorname{argmax}}\ \mathrm{P}\left(I_1,...,I_N|T_1,...,T_N\right)\mathrm{P}\left(T_1,...,T_N\right)\end{aligned} \tag{3.9}$$

where the prior probability on the source volumes is ignored as before but could be utilized as part of an extension in the future. The desired set of transformations which maximizes this probability is denoted as $\hat{\mathcal{T}}$. Equation (3.9) is a combination of the similarity metric and the prior probability on the set of transformations. Given a fixed image volume all blocks belonging to other volumes are conditionally independent. Taking the natural log of the argument in the last line of (3.9) yields

$$\hat{\mathcal{T}} \approx \underset{\mathcal{T}}{\operatorname{argmax}}\left(\ln \mathrm{P}\left(I_1,...,I_N|T_1,...,T_N\right) + \ln \mathrm{P}\left(T_1,...,T_N\right)\right). \tag{3.10}$$

First we will concentrate on simplifying the term $\ln \mathrm{P}\left(I_1,...,I_N|T_1,...,T_N\right)$, which represents the block based similarity measure. The derivation of an expression representing this probability, which can then be evaluated in practice, parallels the work of [83] introduced previously in this chapter. It will be shown that the modeling assumptions made by our approach will yield an energy function that can be efficiently optimized using graph based methods once it has been discretized. The block formulation process for multiple image volumes, given a scene volume $I^{\downarrow}_{i,b}$, can be approximated as

$$\begin{aligned} I_{j,b}(\mathbf{x}) &= f\left(I^{\downarrow}_{i,b}\left(\mathbf{x} - \mathbf{d}_{j,b} - NR_{j,b}(\mathbf{x})\right)\right) + \varepsilon \text{ where } \mathbf{x} \in \mathcal{B} \\ &\approx f\left(I^{\downarrow}_{i,b}\left(\mathbf{x} - \mathbf{d}_{j,b}\right)\right) + \varepsilon \end{aligned} \tag{3.11}$$

where local non-rigid correction term $NR_{j,b}$ is once again ignored in the final result of (3.11), and where $NR_{j,b}$ represents the error in assuming that the motion between volumes can be modeled as displaced blocks. Noise is represented by the term ε. This equation characterizes the block $I_{j,b}$ as being a transformed and corrupted sub-image of the transformed source or scene image $I^{\downarrow}_{i,b}$. The scene image $I^{\downarrow}_{i,b}$ is simply one of the source images which underwent a translation specified by $\mathbf{d}_{i,b}$. If the movement between volumes is only translational then the non-rigid correction term will be zero. Also as the block size decreases so does the importance of the correction factor because a smaller number of features will be captured in each block. In order to derive the similarity measure we assume that each block or sub-image is the result of a separate imaging process. This assumption will enable the development of an efficient groups-wise metric. Including the non-rigid correction terms means that the transformation is exact for all values of $\mathbf{x}$ and ensures that the neighboring blocks of an image volume are coherent, meaning that in a stationary scene, image blocks should have identical voxel intensities in the regions of overlap assuming the noise from the block formulation process can be known and subtracted. An analogous and intuitive way to think of the process of capturing blocks or sub-images would be the way overlapping ultrasound volumes are captured from the same subject. Each volume is the result of a separate imaging process which is what (3.11) implies about block formation. Although we form $I_{j,\{1,\ldots,B\}}$ from an intact image volume each $I_{j,b}$

is assumed independent of the others for the purpose of deriving the metric.

Starting with equation (3.5) and noting that an image estimated by overlapping blocks produced using (3.11) will contain superfluous random variables (voxel intensities) in regions of overlap, a group-wise probabilistic block matching term can be constructed as follows,

$$\begin{aligned}
\ln \mathrm{P}\left(I_1, ..., I_N | T_1, ..., T_N\right) &\approx \sum_{i=1}^{N} \sum_{j \neq i}^{N} \ln \mathrm{P}\left(I_j | I_i, T_i, T_j\right) \\
&\approx \sum_{i=1}^{N} \sum_{j \neq i}^{N} \ln \mathrm{P}\left(I_{j,\{1,...,B\}} | I_i, T_i, T_j\right) \\
&= \sum_{i=1}^{N} \sum_{j \neq i}^{N} \sum_{b=1}^{B} \ln \mathrm{P}\left(I_{j,b} | I_i, \mathbf{d}_{i,b}, NR_{i,b}, \mathbf{d}_{j,b}, NR_{j,b}\right) \\
&\approx \sum_{i=1}^{N} \sum_{j \neq i}^{N} \sum_{b=1}^{B} \ln \mathrm{P}\left(I_{j,b} | I^{\downarrow}_{i,b}, \mathbf{d}_{j,b}\right).
\end{aligned} \tag{3.12}$$

The terms in the first line signify the probability of observing I_j given I_i and both transformations. The second line of (3.12) follows the line of reasoning that a complete image or scene can be estimated by overlapping blocks which contain highly dependent random variables in common regions. The third line of (3.12) is the result of applying the conditional independence property to the blocks which form an image volume. This property originates from the block formulation process described in (3.11). Due to our model of the imaging process which produced $I_{j,\{1,...,B\}}$ from $I^{\downarrow}_i$, we can say that the overlapping blocks of I_j are independent given the set of control point displacement vectors T_j and their corresponding non-rigid correction terms. The final result of (3.12) ignores the non-rigid component as before which if included would ensure the spatial coherence of the overlapping blocks; however, this omission facilitates the derivation of the registration solution as a group-wise probabilistic block-matching scheme. The approximation of (3.12) becomes increasingly valid as the block size decreases and also allows us to develop a computationally efficient implementation. The last term in (3.12) can be given an intuitive description. The group-wise similarity metric is calculated by iterating through each image in the outer summation, with each taking a turn as the scene from which all other images are assumed to be generated from. Next the scene image's pairwise block based similarity between it and the remaining volumes are calculated. Finally all the similarity results are summed. Note that in the final result of (3.12) the transformed version of I_i is used as the scene image and not the original acquired volume. This is important because it corresponds to fixing the transformation of volume I_i during the calculation of $\mathrm{P}\left(I_{j,b}|I^{\downarrow}_{i,b},\mathbf{d}_{j,b}\right)$.

If we assume zero mean Gaussian noise then the term $\mathrm{P}\left(I_{j,b}|I^{\downarrow}_{i,b},\mathbf{d}_{j,b}\right)$ can be simplified to the SSD metric calculated over the block designated by the index b,

$$\begin{aligned}\mathrm{P}\left(I_{j,b}|I^{\downarrow}_{i,b},\mathbf{d}_{j,b}\right) &\approx \overbrace{\sum_{x\in\mathcal{B}}\left[I_{j,b}\left(\mathbf{x}\right)-I^{\downarrow}_{i,b}\left(\mathbf{x}-\mathbf{d}_{j,b}\right)\right]^2}^{\text{SSD between vols. i,j over block b}} \\ &= \sum_{x\in\mathcal{B}}\left[I_{j,b}\left(\mathbf{x}\right)-I_i\left(\mathbf{x}+\mathbf{c}_b+\left(\mathbf{d}_{i,b}-\mathbf{d}_{j,b}\right)\right)\right]^2.\end{aligned} \tag{3.13}$$

The block-wise SSD metric can be evaluated very fast using FFT techniques [21] and will be elaborated on further. For this reason we choose it for our registration experiments although there is no reason that more complex similarity measures could not be used in this block-matching framework. The final likelihood term for the SSD based group-wise registration algorithm is

$$\ln\mathrm{P}\left(I_1,...,I_N|T_1,...,T_N\right) \approx \sum_{i=1}^{N}\sum_{j\neq i}^{N}\sum_{b=1}^{B}\sum_{x\in\mathcal{B}}\left[I_{j,b}\left(\mathbf{x}\right)-I_i\left(\mathbf{x}+\mathbf{c}_b+\left(\mathbf{d}_{i,b}-\mathbf{d}_{j,b}\right)\right)\right]^2 \tag{3.14}$$

The term inside the summations of (3.14) is simply a block-wise similarity measure which was the goal from the outset. It takes into account translation by both image volumes I_i and I_j in its calculation but chooses fixed blocks of I_j to match within larger windows of I_i enabling efficient calculation. The origins of the windows and blocks are determined by the location of each $\mathbf{c}_b$. During the calculation of (3.14) the only volume which undergoes movement is I_i. The measure is still symmetric because each image volume takes a turn as the scene volume I_i.

Let us assume that not all image volumes overlap which is realistic considering our primary application is ultrasound mosaicing. Let $\mathcal{P} = \{i, j\}$ be the set of volume pairs which have some amount of overlap. The right side of (3.14) can now be rewritten to show the symmetry of the metric,

$$\ln \mathrm{P}\left(I_1, ..., I_N | T_1, ..., T_N\right) \approx \sum_{i,j \in \mathcal{P}} \sum_{b=1}^{B} \sum_{x \in \mathcal{B}} \left[I_{j,b}\left(\mathbf{x}\right) - I_i\left(\mathbf{x} + \mathbf{c}_b + \left(\mathbf{d}_{i,b} - \mathbf{d}_{j,b}\right)\right)\right]^2 + \left[I_{i,b}\left(\mathbf{x}\right) - I_j\left(\mathbf{x} + \mathbf{c}_b + \left(\mathbf{d}_{j,b} - \mathbf{d}_{i,b}\right)\right)\right]^2. \quad (3.15)$$

Thus for each pair of images that overlap we use symmetric block-matching to calculate the similarity in the region of overlap and then sum these values together for all valid pairs which results in the group-wise registration metric. This intuitive result stems from the probabilistic formulation of the group-wise registration problem. In (3.15) we see that the term that drives the registration process is $\mathbf{d}_{i,b} - \mathbf{d}_{j,b} = \tau$, which is the difference in translation between overlapping pairs at location $\mathbf{c}_b$. This means that in a group-wise setting only the relative movement between overlapping image volumes is a factor in the similarity metric, which is an obvious but satisfying result. There were two main assumptions that were used to produce (3.15). The first was that all regional deformation around a control point was translational, i.e. we ignored $NR_{n,b}$, and the second was that images could be estimated by overlapping blocks which were produced by separate imaging processes of a fixed scene image. We also assumed Gaussian noise and an identity mapping between voxel intensities in the block imaging process. The inclusion of the non-rigid components $NR_{n,b}$ as a low order parametric term could be considered in the future.

Before the transformation regularization term $\ln \mathrm{P}\left(T_1, ..., T_N\right)$ from (3.10) is explained, the displacement window for the control points must be discretized. Equation (3.15) already uses a discrete value for the image index x; however nothing has been formally stated regarding transformation variables $T_{i,b}$ and $T_{j,b}$. Each control point is allowed to move within a fixed window during the optimization process to be discussed shortly. Assuming a uniform grid of control points, let the displacement window be

$$\mathcal{W} = \left\{ \mathbf{d} | \mathbf{d} \in \mathbb{Z}^3 \wedge -\tfrac{1}{2}h \leqslant d_n \leqslant \tfrac{1}{2}h \right\}. \tag{3.16}$$

By limiting the displacement window, our search space is smaller.

The transformation regularization for a registration problem which contains many partially overlapping image volumes should be a 2nd order regularizer. A regularization based on the second derivative of the transformation function is invariant with respect to linear transformations, which would be desired when for instance the deformation contains mostly rotational components. Minimizing the second derivative as opposed to the first also has the effect of producing a smoother displacement field. In an area of overlap the field's value is dominated by the similarity metric; however, in a region with no other overlapping volumes the field is totally dependent on the regularization energy to determine its value. Shearing artifacts may occur in the transition between the overlapping and non-overlapping regions if only a first order regularizer is used. For each image the set of discrete control point displacements defining the transformation T_n are used to estimate its 2nd derivative using a central difference scheme. Each T_n is regularized independently from the rest of the transformations $T_{\{1,\ldots n-1,n+1,\ldots N\}}$, and the total regularization energy for an individual volume is calculated by discretely summing up each control points contribution. The 2nd order derivatives along the x, y, z dimensions of the deformation field are estimated at the control points by the following equations:

$$\begin{aligned}
\left.\frac{\partial^2 T_n}{\partial x^2}\right|_{\mathbf{x}=\mathbf{c}_b} &\approx \frac{\mathbf{d}_{n,b-1} - 2\mathbf{d}_b + \mathbf{d}_{n,b+1}}{h^2} \\
\left.\frac{\partial^2 T_n}{\partial y^2}\right|_{\mathbf{x}=\mathbf{c}_b} &\approx \frac{\mathbf{d}_{n,b-B^x} - 2\mathbf{d}_{n,b} + \mathbf{d}_{n,b+B^x}}{h^2} \\
\left.\frac{\partial^2 T_n}{\partial z^2}\right|_{\mathbf{x}=\mathbf{c}_b} &\approx \frac{\mathbf{d}_{n,b-B^xB^y} - 2\mathbf{d}_{n,b} + \mathbf{d}_{n,b+B^xB^y}}{h^2}.
\end{aligned} \tag{3.17}$$

Due to the discretization of the registration problem, the regularization of the deformation field only considers the translation at each control point in its calculation.

This is a necessary approximation required to use the graph based optimization methods to be discussed. For volume I_n, the regularization contribution associated with control point $\mathbf{c}_b$ is found by summing the L1 vector norms of the terms from (3.17). This contribution is expressed as

$$\mathcal{R}_{n,b}\left(\mathbf{d}_{n,b}\right) = \left|\frac{\mathbf{d}_{n,b-1} - 2\mathbf{d}_{n,b} + \mathbf{d}_{n,b+1}}{h^2}\right| + \left|\frac{\mathbf{d}_{n,b-B^x} - 2\mathbf{d}_{n,b} + \mathbf{d}_{n,b+B^x}}{h^2}\right| + \left|\frac{\mathbf{d}_{n,b-B^xB^y} - 2\mathbf{d}_{n,b} + \mathbf{d}_{n,b+B^xB^y}}{h^2}\right| \tag{3.18}$$

$$\approx \left|\left.\frac{\partial^2 T_n}{\partial x^2}\right|_{\mathbf{x}=\mathbf{c}_b}\right| + \left|\left.\frac{\partial^2 T_n}{\partial y^2}\right|_{\mathbf{x}=\mathbf{c}_b}\right| + \left|\left.\frac{\partial^2 T_n}{\partial z^2}\right|_{\mathbf{x}=\mathbf{c}_b}\right|.$$

Finally the total regularization of the continuous transformation T_n is approximated as the sum of discrete terms calculated by applying (3.18) to every control point associated with T_n. The regularization energy for image volume I_n is $\mathcal{R}_n\left(\mathcal{D}_n\right) = \frac{1}{B}\sum_{b=1}^{B}\mathcal{R}_{n,b}\left(\mathbf{d}_{n,b}\right)$. The total energy for the group-wise registration problem is found by summing the regularization energy for each individual image volume as $\mathcal{R} = \sum_{n=1}^{N}\mathcal{R}_n\left(\mathcal{D}_n\right)$. This term is substituted into in the registration formulation where λ controls how much influence the regularization has on the registration process. A large value of lambda would overpower the similarity metric and force a rigid transformation of the image volume. Although direct integration of (3.18) with the similarity metric presented in (3.15) would result in the formation of an optimizable discrete registration energy functional we would like to link it to the probabilistic term $\ln \mathrm{P}\left(T_1, ..., T_N\right)$ and give some meaning to what λ represents. First, since it is assumed that the prior probability of all transformations should be independent of each other, the term is rewritten as $\ln \prod_{n=1}^{N} \mathrm{P}\left(T_n\right)$. Now assume that the sum of 2nd derivative approximations, which represent the average regularization energy of a transformation T_n, is a discrete random variable that resembles an exponential function and models the prior probability $\mathrm{P}\left(T_n\right)$. The r.v. is discrete because there are a finite number of values that $\mathcal{R}_n$ can assume due to the discrete nature of the transformation window.

The expression for the prior probability is $\mathrm{P}(T_n) \approx \mathrm{P}(\mathcal{D}_n) = \frac{1}{\kappa}\exp(\lambda \mathcal{R}_n(\mathcal{D}_n))$. This random variable is parameterized by λ, which controls its variance. The constant κ is needed to ensure that it is a valid probability density. The joint prior of all transformations $\ln \mathrm{P}(T_1, ..., T_N)$ becomes

$$\begin{aligned}
\ln \prod_{n=1}^{N} \mathrm{P}(T_n) &\approx \sum_{n=1}^{N} \ln \frac{1}{\kappa} \exp(\lambda \mathcal{R}_n(\mathcal{D}_n)) \\
&= n \ln \frac{1}{\kappa} + \lambda \sum_{n=1}^{N} \mathcal{R}_n(\mathcal{D}_n) \\
&= const + \lambda \sum_{n=1}^{N} \mathcal{R}_n(\mathcal{D}_n).
\end{aligned} \tag{3.19}$$

The final result of (3.19) is the same term used in the registration process minus the constant which has no effect on the optimization. This result makes sense because as we increase λ more influence is given to the regularization term which stiffens the deformation field. This causes the variance of the probability density function, which models the average 2nd derivative values of each transformation, to shrink thus making the prior probability of a non-rigid transformation less likely. By linking (3.19) with the probabilistic framework, discussed previously, we see that an exponential like distribution of the regularization terms is implied and that the registration parameter has a simple relationship to the variance of this distribution. It is interesting to note that the expected value of the transformation 2nd derivative random variable is not zero; however, this doesn't interfere with the optimization.

3.3.1 Efficient evaluation of group-wise SSD metric

Our goal in this section is to introduce a computationally efficient way to evaluate (3.15), utilizing some properties of the Discrete Fast Fourier Transformation (FFT). The number of block matching operations required for a single iteration of the group-wise registration algorithm for multiple ($|\mathcal{V}| > 2$) full resolution volumes can be calculated as $2\,|\mathcal{C}| \begin{pmatrix} |\mathcal{V}| \\ 2 \end{pmatrix}$, where the last term is a binomial coefficient if we assume that all volumes overlap completely. Because the number of times this calculation is performed grows non-linearly with respect the number of volumes, we found that (3.15) may contain more than 105 3-dimensional block matching operations when $|\mathcal{V}| > 3$. For example, in a registration problem with five completely overlapping volumes and an 18x18x18 control point grid the number of block matching operations is found to be $2|\mathcal{C}|\begin{pmatrix}|\mathcal{V}|\\2\end{pmatrix} = 2\left(18^3\right)10 = 116640$. This makes it an serious computational bottleneck. The method of calculation will be shown in the context of a single block-matching operation, and simple bookkeeping will be required to fully evaluate all the block-matching terms of (3.15). The full calculation of the group-wise metric includes the two outer summations where all overlapping volume pairs are considered. Also care must be taken in the instance where the two volumes fail to overlap at $\mathbf{c}_b$. First we will rewrite the innermost summation of (3.15) in terms of $\tau = \mathbf{d}_{i,b} - \mathbf{d}_{j,b}$ which is a vector variable representing the relative displacement between image volumes I_i and I_j at location $\mathbf{c}_b$. The inner summation can be written as

$$\mathrm{SSD}_b^{i,j}(\tau) = \sum_{x \in \mathcal{B}} \left[I_{j,b}(\mathbf{x}) - I_i(\mathbf{x} + \mathbf{c}_b + \tau)\right]^2 + \left[I_{i,b}(\mathbf{x}) - I_j(\mathbf{x} + \mathbf{c}_b - \tau)\right]^2. \quad (3.20)$$

The inclusion of the term $\mathbf{c}_b$ in the transformation above is necessary in order to shift the region surrounding that control point to the origin of the block coordinate system we defined earlier. It simply allows the summation expressing the similarity to be calculated using the set $\mathcal{B}$. In [21] the authors show the existence of an optimal correlation, which can be used to efficiently solve the SSD block-matching problem while maintaining the accuracy of the naive full search method. They show that it is possible to calculate the movement vector that minimizes the value of the SSD metric using two forward FFTs and one inverse FFT which executes 17.8 times faster than the full search method on the video sequences tested. It is noted in this dissertation that the reduction in execution time is not only due to the arithmetic gains by doing the computation in the frequency domain, but also due to the highly optimized FFT libraries available. Special low-level instructions and architectural specific speed-ups can greatly improve performance. Furthermore GPUs are highly suited for FFT computation. One can naively break down (3.20) into two SSD calculations and assume that it will take 6 FFT operations to calculate; however this can be reduced to four FFT operations, three forward and one inverse, by expanding (3.20) and doing some manipulation in the spatial domain. Expanding equation (3.20) results in the following summation

$$\begin{aligned}\mathrm{SSD}_b^{i,j}(\tau) = \sum_{x\in\mathcal{B}} I_{j,b}(\mathbf{x})^2 - 2I_{j,b}(\mathbf{x})\, I_i(\mathbf{x}+\mathbf{c}_b+\tau)^2 I_i(\mathbf{x}+\mathbf{c}_b+\tau)^2 + \\ I_{i,b}(\mathbf{x})^2 - 2I_{i,b}(\mathbf{x})\, I_j(\mathbf{x}+\mathbf{c}_b-\tau)^2 + I_j(\mathbf{x}+\mathbf{c}_b-\tau)^2.\end{aligned} \quad (3.21)$$

The goal will be to write this expression as the result of two correlation operations between complex functions to be defined shortly. The correlations can be performed in the frequency domain using the circular cross-correlation theorem $f[n]\otimes g[n] \underset{DFT}{\longleftrightarrow} F[k]\,\bar{G}[k]$. First it will be necessary to rewrite (3.21) so that the arguments of I_i, I_j match, which can be done by defining some intermediary functions. Let $I_j^{'}(x) = I_j(\mathbf{c}_b - x)$ be a time reversed and translated version of I_j which centers the region surrounding control point $\mathbf{c}_b$ at the origin. Also let $I_i^{'}(x) = I_i(\mathbf{x}+\mathbf{c}_b)$ and $I_{i,b}^{'}(x) = I_{i,b}(-x)$. Note that the terms $I_{j,b}(\mathbf{x})^2, I_{i,b}(\mathbf{x})^2$ are independent of τ, which indicates that they influence the similarity metric by a constant value and thus can be dropped from (3.21) with no effect on the registration result. This equation can be rewritten as

$$\begin{aligned}\mathrm{SSD}_b^{i,j}(\tau) &= \sum_{x\in\mathcal{B}} -2\left(I_{j,b}(\mathbf{x})\, I_i^{'}(\mathbf{x}+\tau)^2 + I_{i,b}(\mathbf{x})\, I_j^{'}(\tau-\mathbf{x})^2\right) + I_i^{'}(\mathbf{x}+\tau)^2 + I_j^{'}(\tau-\mathbf{x})^2 \\ &= \sum_{x\in\mathcal{B}} -2\left(I_{j,b}(\mathbf{x})\, I_i^{'}(\mathbf{x}+\tau)^2 + I_{i,b}(-\mathbf{x})\, I_j^{'}(\mathbf{x}+\tau)^2\right) + I_i^{'}(\mathbf{x}+\tau)^2 + I_j^{'}(\mathbf{x}+\tau)^2 \\ &= \sum_{x\in\mathcal{B}} -2\left(I_{j,b}(\mathbf{x})\, I_i^{'}(\mathbf{x}+\tau)^2 + I_{i,b}^{'}(\mathbf{x})\, I_j^{'}(\mathbf{x}+\tau)^2\right) + I_i^{'}(\mathbf{x}+\tau)^2 + I_j^{'}(\mathbf{x}+\tau)^2.\end{aligned} \quad (3.22)$$

The change of variables $\mathbf{x} \rightarrow -\mathbf{x}$ in the second line of (3.22) is possible because the region $\mathcal{B}$ is centered on the origin thus the summation doesn't change if we simply iterate through this set in reverse. Equation (3.22) contains four correlation operations where the first two measure the similarity between an image block and a region from the designated search window of its paired image. At this point it is possible to define complex functions which will be used to express $\mathrm{SSD}_b^{i,j}(\tau)$ as an efficient calculation performed in the frequency domain. Let us define complex functions f, g, h, m as follows

$$
\begin{aligned}
f &= I_i'(\mathbf{x}) + jI_j'(\mathbf{x}) \\
g &= \begin{cases} -2\left[I_{j,b}(\mathbf{x}) + jI_{i,b}'(\mathbf{x})\right] & \text{if } \mathbf{x} \in \mathcal{B} \\ 0 & \text{otherwise} \end{cases} \\
h &= I_j'(\mathbf{x})^2 + I_i'(\mathbf{x})^2 \\
m &= \begin{cases} 1 & \text{if } \mathbf{x} \in \mathcal{B} \\ 0 & \text{otherwise} \end{cases} .
\end{aligned}
\tag{3.23}
$$

Optimization of (3.24) is equivalent to optimization of (3.22) when applying the functions defined in (3.23).

$$
\mathrm{SSD}_b^{i,j}(\tau) = \mathrm{Real}\left[\sum_{\mathbf{x}\in\mathcal{B}} g(\mathbf{x})\, f(\mathbf{x}+\tau)\right] + \sum_{\mathbf{x}\in\mathcal{B}} m(\mathbf{x})\, h(\mathbf{x}+\tau). \tag{3.24}
$$

Due to the fixed displacement window, τ is known to be limited to $\{\tau | \tau \in \mathbb{Z}^3 \wedge -h \leqslant \tau_n \leqslant h\}$ thus the search window for each block matching operation can be defined as $\mathcal{S} = \{\mathbf{x} | \mathbf{x} \in \mathbb{Z}^3 \wedge -2h \leqslant x_n \leqslant 2h\}$. Next we calculate DFTs which match the size of $\mathcal{S}$. Because we are performing a circular correlation the function g is padded with zeros to match the size of f which is associated with the search window. For example if the control point spacing h is set to eight voxels, then the DFTs will be 31x31x31 due to the definition of the search window $\mathcal{S}$. In this case the final dimensions of $\mathrm{SSD}_b^{i,j}(\tau)$ should be 17x17x17 in order to match the limits of τ. Also note that the DFT of m only has to be computed once at the beginning of the registration procedure because it is a constant mask, thus its computational burden is negligible. Equation (3.24) can be computed for each block using three forward FFTs and one inverse FFT as

$$\mathrm{Real}\left[\mathrm{IFFT}\left(\mathrm{F}\left(\mathbf{u}\right)\bar{G}\left(\mathbf{u}\right)+\mathrm{H}\left(\mathbf{u}\right)\bar{M}\left(\mathbf{u}\right)\right)\right] \tag{3.25}$$

Because the image block initially is centered in the search window before the FFT is taken and due to the effects of the circular correlation, the results from (3.25) must be rearranged before being assigned to $\mathrm{SSD}_b^{i,j}\left(\tau\right)$. Using the example where $h=8$ we find that the size of $\mathrm{SSD}_b^{i,j}\left(\tau\right)$ is 31x31x31 after computation of (3.25). Care must be taken to extract the correct 17x17x17 region from this result because it will be stored in memory for the duration of the registration and referenced by the graph based optimization algorithm multiple times during its progression, looking up the value $\mathrm{SSD}_b^{i,j}\left(\tau\right)$ for particular i,j,b,τ values. Precomputing $\mathrm{SSD}_b^{i,j}\left(\tau\right)$ for each valid i,j,b,τ combination using the efficient FFT technique described above generates all the data terms which may be required by the graph-based optimization algorithm during registration. It should be mentioned that padding the argument of the FFT so that its dimensions are powers of two will result in a speedup of the FFT algorithm. For example functions of size 32x32x32 were faster to transform than functions of size 31x31x31.

3.3.2 Increasing the robustness of our group-wise block matching term in the presence of shadows

As demonstrated in Chapter 2, shadowing artifacts can substantially hinder efforts to register partially overlapping ultrasound volumes. These artifacts are very common in fetal ultrasound, which is the area our stitching efforts are concentrated in, thus our algorithm should attempt account for them. Our approach to compensate for shadows during non-rigid registration involves a pre-processing detection step followed by group-wise registration with a modified similarity metric. Shadow detection can be accomplished using a few different techniques [28] and we have chosen to implement [34]. The basic idea is to ignore a voxel location's contribution to the similarity metric if it lies inside a region deemed to be occluded. Finally the total similarity value is normalized by the number of voxels that were used in its calculation. The normalization is required in order to prevent the registration algorithm from forcing structures apart in the areas dominated by shadows. Essentially the metric we are minimizing becomes the average error per voxel in non-shadow regions. In the following discussion let us assume that M_i and M_j are Boolean masks identifying valid image regions that are free of artifacts. It is easy to incorporate these masks into the block matching framework presented in this dissertation by modifying (3.20) and this modification is shown below,

$$\begin{aligned}
\mathrm{OVR}_b^{i,j}(\boldsymbol{\tau}) &= \sum_{\mathbf{x}\in\mathcal{B}} M_{j,b}(\mathbf{x})\, M_i(\mathbf{x}+\mathbf{c}_b+\boldsymbol{\tau}) + M_{i,b}(\mathbf{x})\, M_j(\mathbf{x}+\mathbf{c}_b-\boldsymbol{\tau}) \\
\mathrm{SHD}_b^{i,j}(\boldsymbol{\tau}) &= \frac{1}{\mathrm{OVR}_b^{i,j}(\boldsymbol{\tau})} \sum_{\mathbf{x}\in\mathcal{B}} \left[M_i(\mathbf{x}+\mathbf{c}_b+\boldsymbol{\tau})\, I_{j,b}(\mathbf{x}) - M_{j,b}(\mathbf{x})\, I_i(\mathbf{x}+\mathbf{c}_b+\boldsymbol{\tau})\right]^2 \\
&\qquad + \left[M_j(\mathbf{x}+\mathbf{c}_b-\boldsymbol{\tau})\, I_{i,b}(\mathbf{x}) - M_{i,b}(\mathbf{x})\, I_j(\mathbf{x}+\mathbf{c}_b-\boldsymbol{\tau})\right]^2
\end{aligned} \tag{3.26}$$

The function $\mathrm{OVR}_b^{i,j}$ is the normalization factor which gives the number of unobscured voxels that volume i and volume j have in common around control point $\mathbf{c}_b$ when their relative offset is $\boldsymbol{\tau}$. The function $\mathrm{SHD}_b^{i,j}$ calculates the sum of squared differences while ignoring shadow regions. This can be seen by noticing that the term inside the summation of $\mathrm{SHD}_b^{i,j}$ evaluates to 0 if the voxel location $\mathbf{x}$ is obstructed by shadows in both image volumes, and this result is to be expected. Let us also assume that the masks calculated during preprocessing were used to eliminate those obscured regions in the ultrasound volumes via setting their voxels equal to 0. Since we have determined that those areas were occluded during preprocessing we can assume their voxel intensities just represent noise anyway. Performing this step prior to registration ensures that the term inside the summation of $\mathrm{SHD}_b^{i,j}$ is only non-zero as long as $M_i = 1$ and $M_j = 1$, which is an intuitive result. Equation (3.26) simply calculates the sum of squared differences using voxels from the overlapping region which haven't been obscured by shadowing artifacts. In order to simplify the presentation and provide equations which are easily translated into code the notation $I_{i,\mathcal{S}_b}(\mathbf{x})$ will be used to designate the search window around a control point $\mathbf{c}_b$, thus $I_{i,\mathcal{S}_b}(\mathbf{x}) = I_i(\mathbf{x} + \mathbf{c}_b)$ where $\mathbf{x} \in \mathcal{S}$. Using the methodology presented in the previous section we can efficiently calculate all possible values of (3.26) for varying $\boldsymbol{\tau}$ by employing multiple FFT operations. Let us define the following complex functions f, g, h, k, m, n which will be used in the calculation of $\text{SHD}_b^{i,j}$,

$$f = \begin{cases} I_{j,b}(\mathbf{x})^2 - 2jI_{j,b}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{B} \\ 0 \text{ otherwise} \end{cases} \qquad g = \begin{cases} M_{i,w_b}(\mathbf{x}) + jI_{i,w_b}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{S} \\ 0 \text{ otherwise} \end{cases}$$

$$h = \begin{cases} I'_{i,b}(\mathbf{x})^2 - 2jI'_{i,b}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{B} \\ 0 \text{ otherwise} \end{cases} \qquad k = \begin{cases} M'_{j,w_b}(\mathbf{x}) + jI'_{j,w_b}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{S} \\ 0 \text{ otherwise} \end{cases}$$

$$m = \begin{cases} M_{j,b}(\mathbf{x}) + jM'_{i,b}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{B} \\ 0 \text{ otherwise} \end{cases} \qquad n = \begin{cases} I_{i,w_b}(\mathbf{x})^2 + jI'_{j,w_b}(\mathbf{x})^2 & \text{if } \mathbf{x} \in \mathcal{S} \\ 0 \text{ otherwise} \end{cases}$$

and also define the functions s, t to be used in the calculation of $\text{OVR}_b^{i,j}$,

$$s = \begin{cases} M_{j,b}(\mathbf{x}) + jM'_{i,b}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{B} \\ 0 \text{ otherwise} \end{cases} \qquad t = \begin{cases} M_{i,w_b}(\mathbf{x}) + jM'_{j,w_b}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{S} \\ 0 \text{ otherwise} \end{cases}$$

where the prime notation indicates a time reversal here, i.e. $I'_{i,b}(\mathbf{x}) = I_{i,b}(-\mathbf{x})$. Using these functions we see that the following equations are equivalent to those presented in (3.26),

$$\begin{aligned} \text{OVR}_b^{i,j}(\boldsymbol{\tau}) &= \text{Real}\left[\sum_{\mathbf{x} \in \mathcal{B}} \bar{s}(\mathbf{x})\, t(\mathbf{x} + \boldsymbol{\tau})\right] \\ \text{SHD}_b^{i,j}(\boldsymbol{\tau}) &= \frac{1}{\text{OVR}_b^{i,j}(\boldsymbol{\tau})} \text{Real}\left[\sum_{\mathbf{x} \in \mathcal{B}} \bar{f}(\mathbf{x})\, g(\mathbf{x} + \boldsymbol{\tau}) + \bar{h}(\mathbf{x})\, k(\mathbf{x} + \boldsymbol{\tau}) + \bar{m}(\mathbf{x})\, n(\mathbf{x} + \boldsymbol{\tau})\right] \end{aligned} \tag{3.27}$$

These correlation operations can be written in the frequency domain using well known identity

$$(x \otimes y)_n = \sum_{m=0}^{N-1} \bar{x}[m]\, \mathrm{y}[m+n] \overset{\text{DFT}}{\leftrightarrow} \bar{X}[k]\, Y[k] \tag{3.28}$$

Thus the final similarity measure is calculated using

$$\begin{aligned}\mathrm{OVR}_b^{i,j}(\boldsymbol{\tau}) &= \mathrm{Real}\left(\mathrm{IDFT}\left[\bar{S}(\mathbf{u})\,T(\mathbf{u})\right]\right)\\ \mathrm{SHD}_b^{i,j}(\boldsymbol{\tau}) &= \frac{1}{\mathrm{OVR}_b^{i,j}(\boldsymbol{\tau})}\,\mathrm{Real}\left(\mathrm{IDFT}\left[\bar{F}(\mathbf{u})\,G(\mathbf{u})+\bar{H}(\mathbf{u})\,K(\mathbf{u})+\bar{M}(\mathbf{u})\,N(\mathbf{u})\right]\right)\end{aligned} \tag{3.29}$$

With this approach we achieved a speedup factor of 10.31 over the naive implementation which is a substantial improvement. Our results suggest that an efficient group-wise similarity measure, which is robust to ultrasound shadowing, can be easily implemented using FFT operations. To our knowledge this is the first attempt at a group-wise metric for ultrasound mosaicing which accounts for shadowing artifacts such as commonly experienced in a clinical setting. The two block-wise similarity measures discussed in this chapter are readily interchangeable in the following discussion on MRF optimization even though we frequently refer to (3.15) instead of (3.26) due to simplicity. It should be noted that all ultrasound registration experiments described in this chapter were conducted using the similarity measure defined in (3.26), due to its ability to handle heavily shadowed volumes.

3.3.3 Formulation of registration energy

In this section the complete registration energy is formulated, which is a combination of the similarity measure and the transformation regularization developed previously. Using the well-known Hammersley Clifford theorem [8] it will also be shown that this registration energy forms a Markov Random Field (MRF) in which each node corresponds to the displacement of its associated control point. An undirected graphical model $\mathcal{G}$ is called a MRF if two nodes are conditionally independent given the values of the nodes separating them. This can be stated as $P(X_i|X_{\mathcal{G}}/i) = P(X_i|X_{N_i})$ where $X_{\mathcal{G}\setminus i}$ refers to all the nodes except X_i and X_{N_i} refers to the neighborhood of X_i, i.e. all nodes which are connected to X_i. The Hammersley Clifford theorem states that if the joint probability function takes on a specific form called a Gibbs distribution then it also forms a MRF. A proof can be found in [16]. It will be shown in this section that our probabilistic formulation of the group-wise registration problem forms a Gibbs distribution with double and triple clique terms. Efficient inference methods commonly used on MRFs in computer vision will be used to find the optimal transformations to bring the overlapping volumes into alignment. Formulation of the registration energy is given as

$$\mathcal{E}(\mathcal{D}) = \sum_{i=1}^{N} \left[\sum_{j>i}^{N} \sum_{b=1}^{B} \mathrm{SSD}_b^{i,j}\left(\mathbf{d}_{i,b} - \mathbf{d}_{j,b}\right) + \frac{\lambda}{B} \sum_{b=1}^{B} \left(\left| \frac{\mathbf{d}_{i,b-1} - 2\mathbf{d}_{i,b} + \mathbf{d}_{i,b+1}}{h^2} \right| + \left| \frac{\mathbf{d}_{i,b-B^x} - 2\mathbf{d}_{i,b} + \mathbf{d}_{i,b+B^x}}{h^2} \right| + \left| \frac{\mathbf{d}_{i,b-B^xB^y} - 2\mathbf{d}_{i,b} + \mathbf{d}_{i,b+B^xB^y}}{h^2} \right| \right) \right]. \quad (3.30)$$

The first thing to note about this equation is the grouping of random variables from the set $\mathcal{D}$ into distinct terms. Like before, each discrete variable $\mathbf{d}_{n,b}$ corresponds to the displacement of a control point located at $\mathbf{c}_b$ and associated with volume n. The groupings are referred to as cliques and upon inspection it can be seen that there are terms in (3.30) which are functions of double and triple cliques. The similarity terms $\mathrm{SSD}_b^{i,j}\left(\mathbf{d}_{i,b} - \mathbf{d}_{j,b}\right)$ are examples of double clique terms due to them being a function of two displacement variables. Each double clique term links the displacements of co-located control points belonging to a pair of different volumes. Notice in the $\mathrm{SSD}_b^{i,j}\left(\mathbf{d}_{i,b} - \mathbf{d}_{j,b}\right)$ terms that the volume identifiers i, j are different but the location identifier b is identical. As discussed before, this measures the block-wise similarity between volumes at a region surrounding $\mathbf{c}_b$ and can be thought of as inter-volume terms. The triple clique terms found in the second row of (3.30) measures the regularization energy of the transformations. Each control point in volume n belongs to three triple cliques, one corresponding to each dimension of the volume. Looking at the indexes in these terms it is apparent that all triple clique terms are intra-volume terms since the volume identifier i is constant. This results from our assumption that the prior probabilities of transformations in the set $\mathcal{T}$ are independent. We can also write (3.30) in a more compact way where double (inter-volume) and triple (intra-volume) cliques of nodes are collected into the sets $\{\mathbf{d}_{i,b}, \mathbf{d}_{j,b}\} \in \mathcal{N}_{Inter}$ and

$\{\mathbf{d}_{i,a}, \mathbf{d}_{i,b}, \mathbf{d}_{i,c}\} \in \mathcal{N}_{Intra}$ respectively. Equation (3.30) is now written as

$$\mathcal{E}(\mathcal{D}) = \sum_{\substack{\{\mathbf{d}_{i,b},\mathbf{d}_{j,b}\}\\ \in \mathcal{N}_{Inter}}} \mathrm{SSD}_b^{i,j}(\mathbf{d}_{i,b} - \mathbf{d}_{j,b}) + \frac{\lambda}{B} \sum_{\substack{\{\mathbf{d}_{i,a},\mathbf{d}_{i,b},\mathbf{d}_{i,c}\}\\ \in \mathcal{N}_{Intra}}} \left| \frac{\mathbf{d}_{i,a} - 2\mathbf{d}_{i,b} + \mathbf{d}_{i,c}}{h^2} \right|, \quad (3.31)$$

which can be described by a graphical model where the cliques are indicated using edges. This is the most common form found in the computer vision literature. Due to the Hammersley-Clifford theorem and the fact each of these terms was derived from a probabilistic expression using the maximum likelihood framework we can arrive at the conclusion that (3.31) can be represented as a Markov Random Field where each node in the model corresponds to a single discrete displacement variable $\mathbf{d}_{n,b}$.

Although it is not necessary for the understanding of the registration method, the Hammersley-Clifford theorem will be reviewed since it forms the foundation for significant developments in computer vision. A probability distribution, such as $\mathrm{P}(\mathcal{D})$ in our case, defined on an undirected graphical model $\mathcal{G}$ is called a Gibbs distribution if it can be factored into positive functions whose arguments are the cliques which characterize the nodes and edges of $\mathcal{G}$. That is,

$$\mathrm{P}(\mathcal{D}) = \frac{1}{Z} \prod_{c \in C_g} \phi_c(\mathcal{D}^c), \quad (3.32)$$

where C_g is the set of all maximal cliques in $\mathcal{G}$ and Z is the normalization constant. A maximal cliques refers to the largest set of fully connected nodes. If a set of nodes forms a maximal clique then adding any other node from $\mathcal{G}$ breaks the fully connected property of the set. We identify the displacement values $\mathbf{d}_{n,b}$ of the nodes for an individual clique using the notation $\mathcal{D}^c$. In our registration problem the set of C_g contains the double and triple cliques found in (3.31) which are associated with the functions ϕ_c that compute the similarity and regularization terms. (3.32) turns into the summation of (3.31) after taking the natural logarithm. A partial graph corresponding to a two dimensional version of our registration problem is shown in Figure 3-2 for clarity. The 3D version is a straightforward extension of this 2D version. Figure 3-2 shows control point nodes corresponding to three overlapping 2D images and

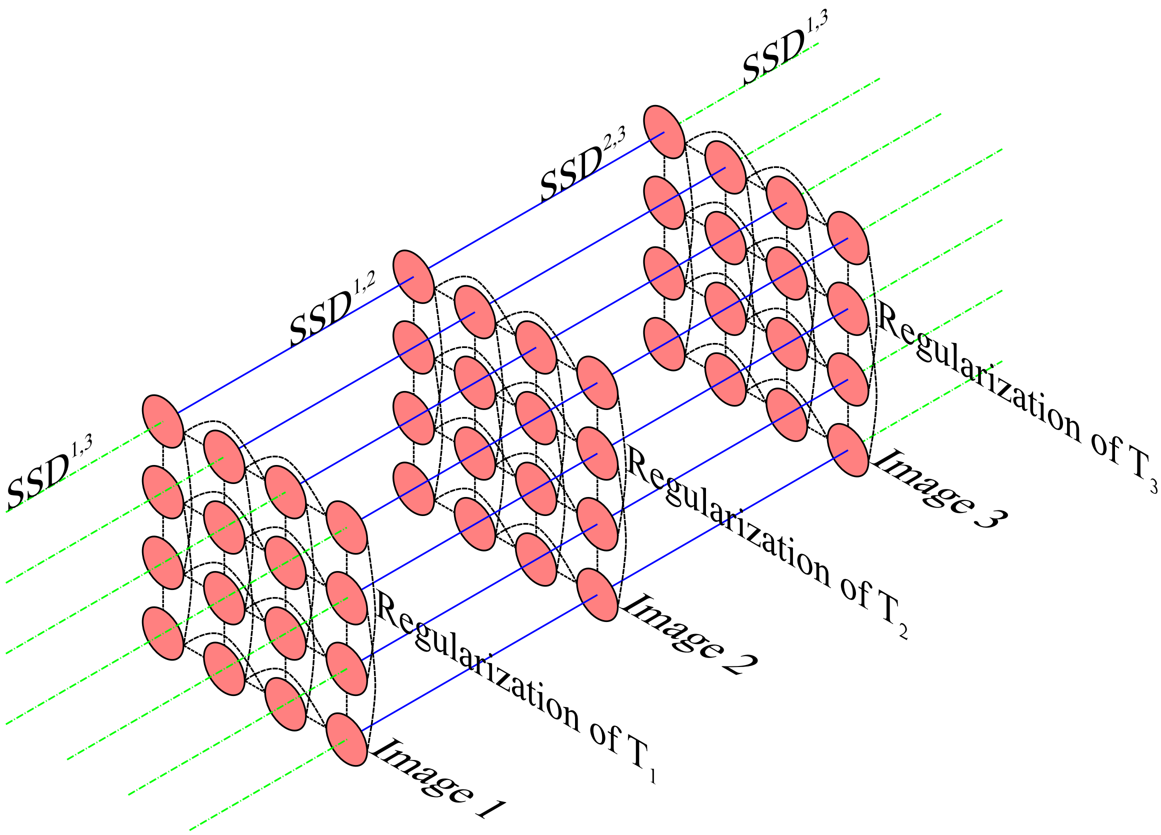


Figure 3-2: Graph for Markov Random Field representing registration problem with three overlapping 2D images. Each control point node is represented by a circle, with its latent value being defined as a displacement vector. Probabilistic dependencies between nodes are modeled by the edges connecting them. For example, similarity dependencies are edges which extend between images while smoothness dependencies are represented by edges between nodes of the same image.

demonstrates the intra and inter-volume dependencies. The maximal cliques which are grouped into the terms of (3.31) are also easily seen in Figure 3-2. Using this graph we desire to make an inference on the displacement values for each node. In the optimization section the method chosen to make this inference will be discussed in detail. The double clique terms corresponding to the similarity measure are modeled by edges between the images while the triple clique terms corresponding to the regularization are modeled by the edges confined to each image's grid of nodes. Not all edges are shown in this figure, but can be deduced by (3.31). We can make some statements about the probabilistic dependencies of these displacement nodes using the Hammersley-Clifford theorem which has found much use since its formulation in 1974. The theorem tells us that the definition of a Gibbs distribution is equivalent to the definition of a Markov Random Field, which means that in the graphical model $\mathcal{G}$ described above two nodes are conditionally independent whenever they are separated by evidence nodes. Evidence nodes have displacements which have been given specific values. In our problem this implies that for every node $\mathbf{d}_{n,b}$ in our graph the following conditional property holds,

$$\begin{aligned} &\mathrm{P}\left(\mathbf{d}_{n,b} \,|\mathcal{D}\backslash\mathbf{d}_{n,b}\right) = \\ &\quad \mathrm{P}\left(\mathbf{d}_{n,b} \,\middle|\mathbf{d}_{\{1\ldots n-1,n+1\ldots N\},b}, \mathbf{d}_{n,b+1}, \mathbf{d}_{n,b-1}, \mathbf{d}_{n,b+B^x}, \mathbf{d}_{n,b-B^x}, \mathbf{d}_{n,b+B^xB^y}, \mathbf{d}_{n,b-B^xB^y}\right) \end{aligned} \tag{3.33}$$

This is a nice representation of the displacement dependencies in our registration method. It defines the local Markov property which characterizes this group-wise block matching algorithm. This works well for group-wise image registration because these conditional properties of Markov Random Fields lead to implicit global dependencies between the volumes.

3.4 Efficient optimization of registration energy using fusion techniques

Since our displacement label space $\mathcal{W}$ is large, i.e. if control point spacing h=8 then $|\mathcal{W}| = 9^3$ using the definition of $\mathcal{W}$ from (3.16), and the MRF representing the registration energy given by (3.30) contains a considerable amount of double and triple clique terms which may not be submodular, we are left with a substantial NP-hard optimization problem. Our strategy to optimize this energy will be to employ a parallelized alpha-expansion technique which was recently developed in [42]. It should be noted that each triple clique term in our energy function is expanded into six double clique terms before optimization, since the methods we employ operate on pair-wise potentials. A key component of this algorithm will be the optimization of non-submodular binary-labeled MRFs which are sub-problems of the registration procedure. The BHS algorithm presented in [36] will be used for this purpose. The general idea of the fusion technique is to combine different suboptimal solutions in an intelligent way in order to produce a better solution with lower registration energy. In our application the unique suboptimal solutions will result from exploring different regions of the transformation space independently and in parallel using multiple CPU threads. These solutions are then fused in a principled way that will be described in this section. The authors of [42] applied this algorithm to the common computer vision problem of calculating optical flow. They coined their approach Fusion-Flow, which is a discrete/continuous optimization scheme based on fusion moves that combines the advantages of both. Different optical flow approaches have diverse strengths and weaknesses, and a superior solution may result from a combination of unique approaches fused together to get the final result. For example in [42] the authors fused flow solutions resulting from two classic optical flow algorithms, the Lucas-Kanade solution which behaves well in textured regions but is useless in smooth areas, and the Horn-Schunck method which gives a good approximation but over smooths motion discontinuities. The authors produced around 200 proposal solutions by varying the parameters of each algorithm and fused them together using the approach will be discussed in detail below. The final fused result was very close to the ground truth demonstrating the effectiveness of the concept.

Before explaining the registration optimization algorithm, the minimization of binary MRFs is briefly reviewed. The registration energy will need to be expressed in this form. These problems take the form

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \Phi_i(x_i) + \sum_{(i,j) \in \mathcal{N}} \Psi_{ij}(x_i, x_j) \quad \text{where} \quad \mathbf{x} \in \mathcal{L}^{|\mathcal{V}|} \tag{3.34}$$

where $\mathcal{V}$ is a set of nodes, $\mathcal{N}$ is a set of undirected edges connecting pairs of nodes, and $\mathcal{L}$ is the set of labels. In the registration energy these nodes correspond to control points, but in other computer vision applications such as stereo vision they usually correspond to pixels or voxels. The labeling $\mathbf{x}$ assigns each node a label from the space $\mathcal{L}$ and the goal of the optimization is to find $\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}}\, E(\mathbf{x})$. The terms Φ and Ψ are real valued functions referred to as unary and pairwise potentials, which are used to calculate the energy of a labeling. These symbols are commonly found in the literature and so are used here in order to introduce the graph cut theory; however they are not equivalent to the symbols which were used earlier to define the registration problem. It is well known that if $\mathcal{L} = \{0, 1\}$ and the Φ and Ψ terms satisfy a certain condition called sub-modularity then a globally optimal labeling can be found by solving a minimum cut problem on a specially constructed graph. Minimization of (3.34) with this binary label space is known as quadratic pseudo-Boolean optimization (QPBO) [36] where quadratic refers to the highest ordered term being pairwise. The submodularity condition for a discrete binary labeling problem, which is analogous to the convex property of continuous functions, can be stated as

$$\Psi_{ij}(0,0) + \Psi_{ij}(1,1) \leqslant \Psi_{ij}(1,0) + \Psi_{ij}(0,1) \tag{3.35}$$

The minimization of equations which take the form of (3.34) can be accomplished using a popular algorithm known as alpha expansion, which is based on graph cuts. The idea behind alpha expansion is to choose a label, which in this application would be one of the possible displacements from the discretized window $\mathcal{W}$, and then perform a binary graph cut optimization which results in that label expanding its footprint in the final solution vector. Because graph cuts are inherently binary, each label is expanded individually during optimization until no label expansion produces a lower energy value. During each cycle of alpha expansion we label the nodes corresponding to the control points defining each transformation as 0 if they should retain their current displacement value or 1 if they should take on the value of the displacement label currently being expanded, also known as alpha. If the pairwise terms in (3.34) satisfy the submodularity constraint given by (3.35) then each expansion produces a globally optimal result. Moves that are globally optimal during each cycle of alpha expansion make this optimization method resilient to getting trapped in local minima. This condition implies that neighboring nodes in the graph construction must be encouraged to take the same label, which is usually the case in medical imaging applications such as segmentation. Alpha expansion often outperforms optimization methods based on local gradients because during each expansion an exponential amount of moves is considered, and the globally optimal move can be chosen efficiently due to advances in max flow/min cut algorithms. Because each cycle results in a globally optimal result for that expansion sub problem the final result after all labels are expanded is very good. When the submodularity condition is satisfied, the energy after optimization using alpha expansion is within a known bound from the globally optimal solution, and can be calculated using the Ψ_{ij} terms.

This theory will now be discussed in the context of group-wise image registration, and an efficient method to calculate all the non-rigid transformations will be developed. Alpha expansion has been effectively used to perform traditional pair-wise registration in [72] where their contribution outperformed other registration algorithms including both continuous approaches such as the diffeomorphic demons found in the Insight Toolkit and also discrete approaches found in [24]. Their idea was to perform alpha expansion using the set of labels $\mathcal{W}$. It should be noted that using a maximum-cut computation to perform each expansion step is only valid for sub-modular functions, which is what [72] dealt with. The restriction to pairwise registration coupled with a 1st order transformation regularizer guarantees sub-modularity of the registration energy. However, in the group-wise registration formulation of (3.30) no guarantees can be made on the sub-modularity of the terms so a more sophisticated optimization method must be applied for each expansion step. This is due to the group-wise nature of the algorithm coupled with the 2nd order regularizer.

This is where the BHS algorithm is utilized. It effectively optimizes non-submodular binary labeled MRFs. As discussed above, the global optimum for a binary labeled submodular function can be computed exactly; however, this isn't possible during the alpha-expansion steps in the registration procedure because submodularity isn't guaranteed. What is possible however is a partial result where each node that has been labeled by the BHS algorithm is part of the globally optimal solution. Thus the BHS algorithm is executed on a non-submodular binary labeling problem during each expansion and produces an output vector $\mathbf{x}$ which contains 0,1 or ?. The nodes given values 0 or 1 are considered labeled and part of the global optimal solution whereas the nodes designated as ? are considered unlabeled. In our experience the registration algorithm produces very few unlabeled nodes, less than .01%, which are simply left with the default displacement of zero. This was also the result of the stereo vision experiments in [36]. A very important property of the BHS algorithm, which is required for its use within alpha-expansion, is called the persistence property. This states that if we replace the unlabeled nodes of the output vector $\mathbf{x}$ with labels from an arbitrary vector $\mathbf{y}$ the energy of the composite vector $\mathbf{x}'$ is guaranteed to be less than or equal to $\mathcal{E}(\mathbf{y})$. This is an interesting but not obvious property which results from the fact that any labeled node must be part of a globally optimal solution. Thus in our application only labeled nodes or control points are assigned displacement values from the set $\mathcal{W}$ while the rest remain in their initial location. The registration energy in (3.30) reduces to the form shown in (3.34) as discussed but it should be noted that the pairwise terms increase exponentially with respect to the number of control points. This is due to the 2nd order regularizer which necessitates the use of triple cliques. In addition to the computational burden from the additional pairwise functions used to express the triple clique terms there is also overhead associated with using the more complex BHS algorithm which is necessary due to the non-submodularity of the problem. In light of this complexity we propose an approach where mutually exclusive regions in the solution space are explored in parallel and determine that this method enables efficient and accurate calculation of the transformation functions associated with each volume. In the 3-dimension application of ultrasound mosaicing the cubic search window is separated into 8 distinct regions corresponding to the 8 corners of the cube. Figure 3-3 demonstrates how the label space $\mathcal{W}$ is broken down into separate simplified registration problems which are then computed in parallel. This figure also explains how the 8 suboptimal solutions which were computed initially by parallel CPU threads are fused into a final enhanced solution. There is very little overhead associated with the fusion process since the label space is so large. For example each fusion step requires the BHS algorithm to run once whereas computing an initial solution using alpha-expansion with a limited displacement label set requires 92 BHS executions if $|\mathcal{W}| = 9^3$.

Consider two registration solutions $\mathbf{d}^0 \in \mathcal{W}_0^{|\mathcal{C}|}$ and $\mathbf{d}^1 \in \mathcal{W}_1^{|\mathcal{C}|}$ where $|\mathcal{C}|$ is simply the number of control points formally stated as the cardinality of the set $\mathcal{C}$. The solutions are referred to as labellings where the subscript indicates the region of the displacement window $\mathcal{W}$ that was considered during their calculation. Also note that they are vectors whose lengths are equal to the number of control points. The goal of the fusion step is to determine a composite labeling $\mathbf{d}^c$ where the displacement vector of each node must come from either $\mathbf{d}^0$ or $\mathbf{d}^1$. Thus $\mathbf{d}^c \in (\mathcal{W}_0 \cup \mathcal{W}_1)^{|\mathcal{C}|}$ and can be expressed as a linear combination of $\mathbf{d}^0$ and $\mathbf{d}^1$ using a binary vector $\mathbf{y} \in \{0,1\}^{|\mathcal{C}|}$ as $\mathbf{d}^c(\mathbf{y}) = \mathbf{d}^0 \bullet (1 - \mathbf{y}) + \mathbf{d}^1 \bullet \mathbf{y}$ where the multiplication is done element-wise. The binary vector $\mathbf{y}$ has an element associated with every displacement node and can be indexed using the same notation that was used to index the displacement vector $\mathbf{d}$. Also, all inter and intra volume nodal relationships or cliques are identical in the new MRF problem. Using this construction for $\mathbf{d}^c$ the fusion of two registration solutions can be written as a non-submodular binary-labeled MRF which can be solved using the BHS algorithm discussed earlier. The new binary optimization problem is shown

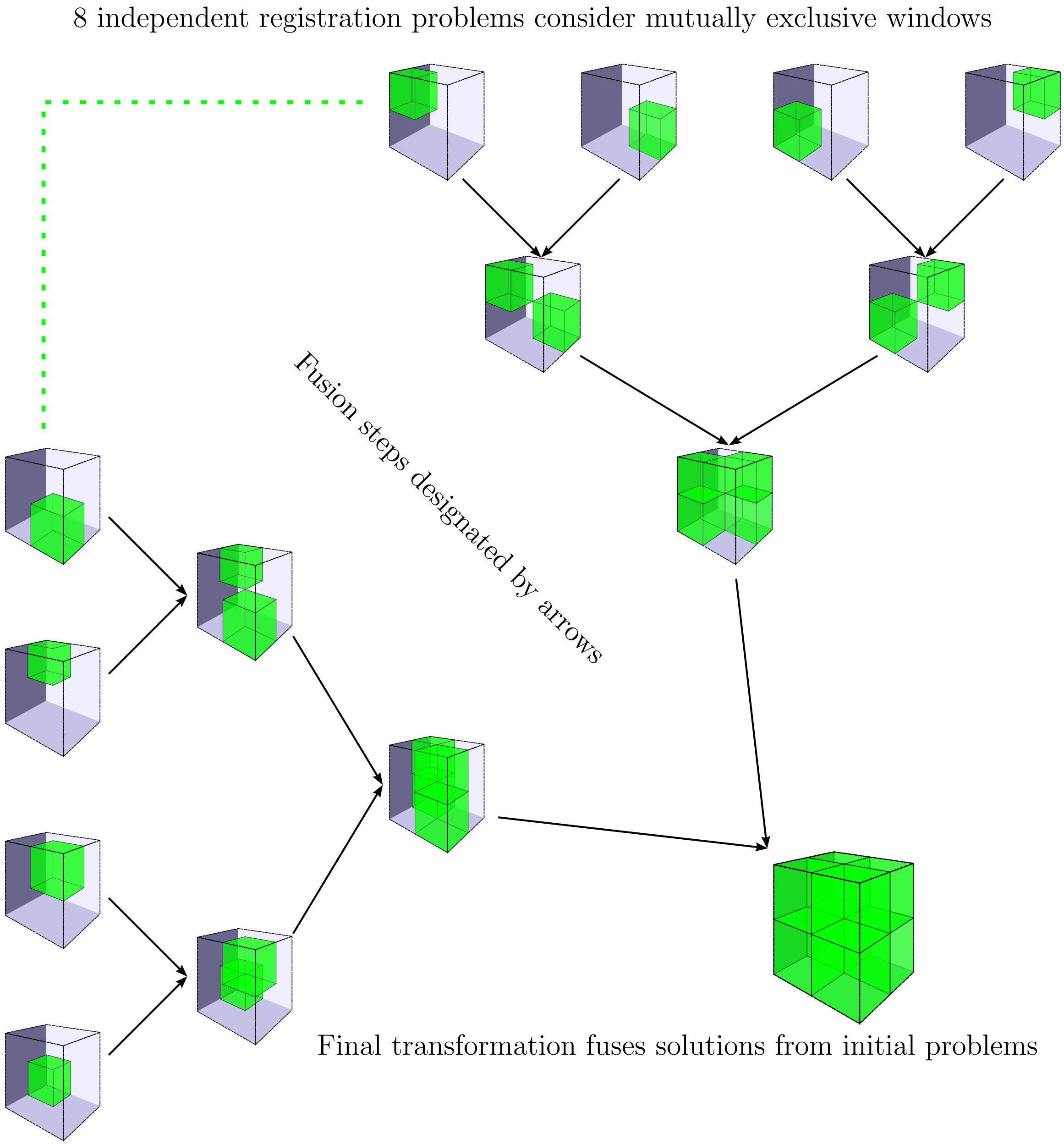


Figure 3-3: Displacement label space is segmented into 8 regions which are evaluated in parallel. Fusion steps detail how final registration solution is calculated.

below

$$\mathcal{E}\left(\mathbf{d}^{c}\left(\mathbf{y}\right)\right) = \sum_{\substack{\left\{\mathbf{y}_{i,b},\mathbf{y}_{j,b}\right\} \\ \in \mathcal{N}_{Inter}}} \Psi_{b}^{i,j}(\mathbf{y}_{i,b}, \mathbf{y}_{j,b}) + \sum_{\substack{\left\{\mathbf{y}_{i,a},\mathbf{y}_{i,b},\mathbf{y}_{i,c}\right\} \\ \in \mathcal{N}_{Intra}}} \Theta_{a,b,c}^{i}(\mathbf{y}_{i,a}, \mathbf{y}_{i,b}, \mathbf{y}_{i,c}) \tag{3.36}$$

where the new potentials Ψ_{ij} and Θ_{ijk} are defined as

$$\Psi_{b}^{i,j} = \begin{bmatrix} \mathrm{SSD}_{b}^{i,j}\left(\mathbf{d}_{i,b}^{0} - \mathbf{d}_{j,b}^{0}\right) & \mathrm{SSD}_{b}^{i,j}\left(\mathbf{d}_{i,b}^{0} - \mathbf{d}_{j,b}^{1}\right) \\ \mathrm{SSD}_{b}^{i,j}\left(\mathbf{d}_{i,b}^{1} - \mathbf{d}_{j,b}^{0}\right) & \mathrm{SSD}_{b}^{i,j}\left(\mathbf{d}_{i,b}^{1} - \mathbf{d}_{j,b}^{1}\right) \end{bmatrix} \tag{3.37}$$

and

$$\begin{aligned} \Theta_{a,b,c}^{i}\left(\mathbf{y}_{i,a}, \mathbf{y}_{i,b}, \mathbf{y}_{i,c} = 0\right) &= \begin{bmatrix} \left|\frac{\mathbf{d}_{i,a}^{0}-2\mathbf{d}_{i,b}^{0}+\mathbf{d}_{i,c}^{0}}{h^{2}}\right| & \left|\frac{\mathbf{d}_{i,a}^{0}-2\mathbf{d}_{i,b}^{1}+\mathbf{d}_{i,c}^{0}}{h^{2}}\right| \\ \left|\frac{\mathbf{d}_{i,a}^{1}-2\mathbf{d}_{i,b}^{0}+\mathbf{d}_{i,c}^{0}}{h^{2}}\right| & \left|\frac{\mathbf{d}_{i,a}^{1}-2\mathbf{d}_{i,b}^{1}+\mathbf{d}_{i,c}^{0}}{h^{2}}\right| \end{bmatrix} \\ \Theta_{a,b,c}^{i}\left(\mathbf{y}_{i,a}, \mathbf{y}_{i,b}, \mathbf{y}_{i,c} = 1\right) &= \begin{bmatrix} \left|\frac{\mathbf{d}_{i,a}^{0}-2\mathbf{d}_{i,b}^{0}+\mathbf{d}_{i,c}^{1}}{h^{2}}\right| & \left|\frac{\mathbf{d}_{i,a}^{0}-2\mathbf{d}_{i,b}^{1}+\mathbf{d}_{i,c}^{1}}{h^{2}}\right| \\ \left|\frac{\mathbf{d}_{i,a}^{1}-2\mathbf{d}_{i,b}^{0}+\mathbf{d}_{i,c}^{1}}{h^{2}}\right| & \left|\frac{\mathbf{d}_{i,a}^{1}-2\mathbf{d}_{i,b}^{1}+\mathbf{d}_{i,c}^{1}}{h^{2}}\right| \end{bmatrix} \end{aligned} \tag{3.38}$$

The optimization of the fusion energy in (3.36) produces a binary solution vector $\mathbf{y}$ which equals 0 at nodes that are designated to use the displacement value from $\mathbf{d}^{0}$ and 1 at nodes that are designated to use the displacement value from $\mathbf{d}^{0}$. The binary vector $\mathbf{y}$ is used to generate the improved composite solution $\mathbf{d}^{c}$. The procedure described above is known as fusion and will be designated by the operator $\odot$. The pseudo-code for registration energy optimization is shown in algorithm 1. An intuitive example that shows the usefulness of fusing two suboptimal solutions in the context of group-wise image registration is given in Figure 3-4. In Figure 3-4 the goal is to align two simulated images where translational movement has occurred near the baby's face. The dots indicate control point locations which need to be labeled with displacement values from $\mathcal{W}$. In this simple 2-dimensional example we are only concerned with control point $\mathbf{c}_{b}$ which is at the center of the cropped images. As shown in Figure 3-1 the goal is to align these two images in a common coordinate system by calculating a transformation function for each. The control point spacing in this example is h, and we will allow the points to be displaced a maximum of $\frac{h}{2}$ in any direction. The displacement window is broken into four regions, and this example concentrates on the

Algorithm 1 Parallelized alpha-expansion for MRF group-wise registration

```
Split displacement window W into W^1...W^8
Calculate all block-based similarity terms
d^1 ← ((0,0,0)^T, ..., (0,0,0)^T)
for several sweeps do
    d^2...d^8 ← d^1
    parfor i ∈ (1...8) do
        for α ∈ W^i do
            d^i ← d^i ⊙ α
        end for
    end parfor
    Wait for all threads
    d^1 ← d^1 ⊙ d^2 // 1st level of fusion
    d^3 ← d^3 ⊙ d^4
    d^5 ← d^5 ⊙ d^6
    d^7 ← d^7 ⊙ d^8

    d^1 ← d^1 ⊙ d^3 // 2nd level of fusion
    d^5 ← d^5 ⊙ d^7

    d^1 ← d^1 ⊙ d^5 // Final result
end for
return d^1
```

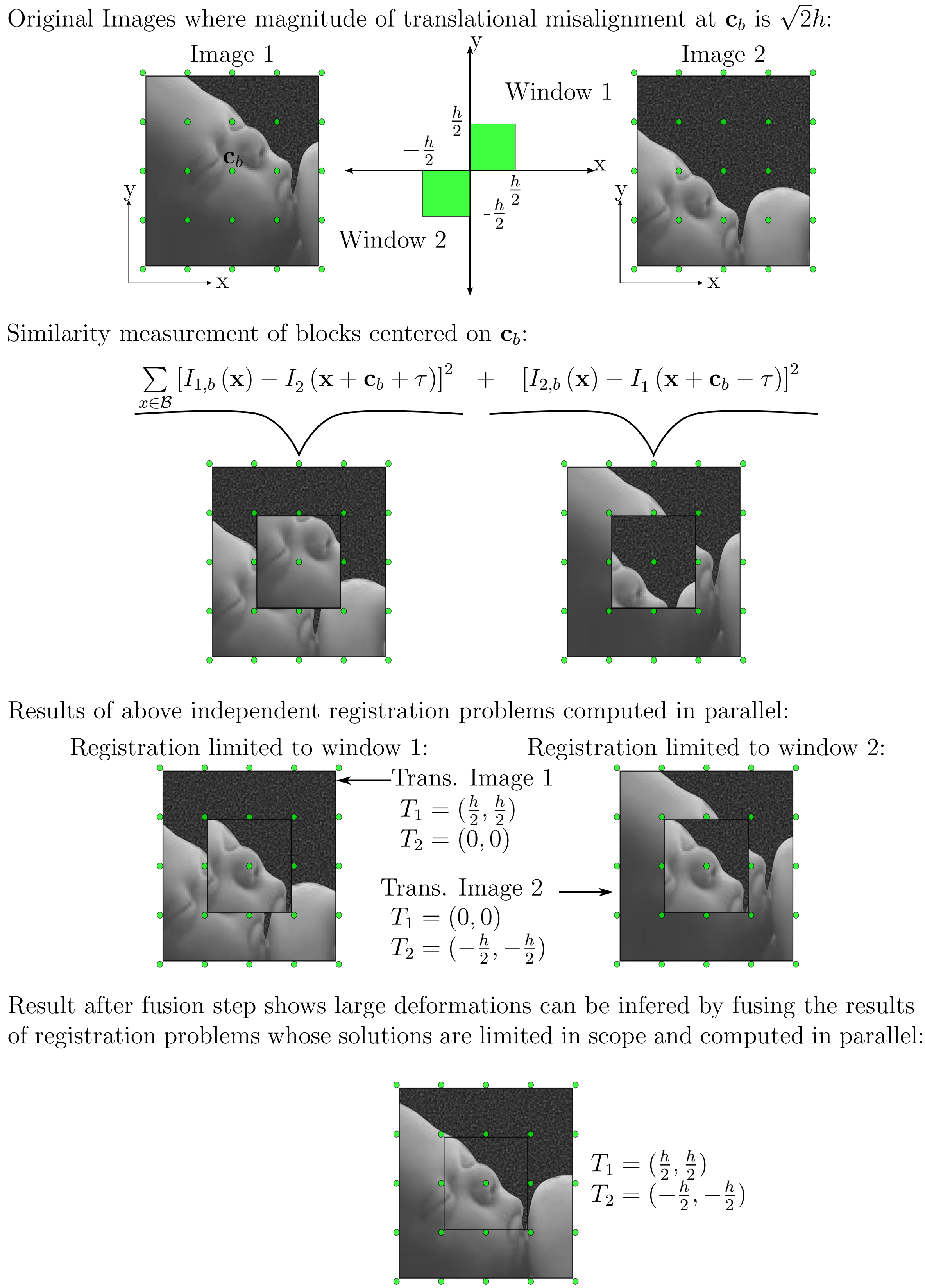


Figure 3-4: Simple example showing fusion techniques ability to compute large deformations

upper right and lower left quadrants of this space. The similarity measure in (3.30) contains pairwise block-matching terms, and in this example the term corresponding to control point $\mathbf{c}_b$ is isolated. Figure 3-4 displays the block/window associated with each half of the expression $\sum_{x \in \mathcal{B}} \left[I_{j,b}(\mathbf{x}) - I_i(\mathbf{x} + \mathbf{c}_b + \tau)\right]^2 + \left[I_{i,b}(\mathbf{x}) - I_j(\mathbf{x} + \mathbf{c}_b - \tau)\right]^2$. Two registration problems are formulated and solved using discrete displacement values from the set $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively. The resulting transformations are also shown in this figure along with the resulting images overlaid on top of each other. In the left-half of the figure, where we have forced the control point displacements of the solution to take values from $\mathcal{W}_1$, it is apparent that no values from this set will improve the similarity measure at $\mathbf{c}_b$ by transforming Image 2. However the alignment at $\mathbf{c}_b$ can be improved by translating Image 1 by $\left(\frac{h}{2}, \frac{h}{2}\right)$, which belongs in the set $\mathcal{W}_1$. Analogously for dual registration problem, where the window was been limited to $\mathcal{W}_2$, we observe that no displacement values from this set can improve alignment by transforming Image 1. In this case Image 2 is translated by $\left(\frac{h}{2}, \frac{h}{2}\right)$ to maximize the similarity metric. As seen in the individual solutions of these dual registration problems are far from perfect though they have slightly improved the translational alignment at $\mathbf{c}_b$. Once these two registration problems have been solved we wish to fuse them together to produce a single solution which increases the similarity metric by taking the best components of each. In the final row the fused solution is given which combines the transformation of Image 1 computed from the first registration problem with the transformation of Image 2 computed from the second registration problem. The resulting images are perfectly aligned at $\mathbf{c}_b$. The extension of this concept to three dimensions and increasing the number of volumes beyond two is straightforward and results in the same advantages, namely a large speedup when using multiple processing cores.

Although we used the fusion technique in order to improve the efficiency of our group-wise registration algorithm by computing sub-optimal solutions in parallel one can think of other registration applications of this algorithm. For example one may find that improved registration may result from the fusion of solutions calculated using different similarity measures or regularization techniques.

3.5 A modified MRF for focused registration

The group-wise registration algorithm was developed with the intent to stitch together overlapping 3D image volumes, with ultrasound being the primary modality for volume generation. Because we are dealing with living structures, including a fetus that typically exhibits some motion during the ultrasound scanning, registration is both difficult and essential when attempting to construct an accurate training volume from clinical data. Thus a seam selection algorithm was developed which would enable the group-wise registration algorithm to focus on regions of overlap where anatomical features are similarly represented in all overlapping volumes and the movement between volumes in this region therefore is hopefully minimized. To this end we borrowed concepts from the common image processing task of seamless image stitching, albeit slightly modified for multiple overlapping 3D ultrasound volumes. We consider this step a preprocessing task before group-wise non-rigid registration. Ultimately this step labels the voxels in the composite volume's common scene coordinate system with a source ID. Thus the label space for this problem is defined as $\mathcal{L} = \{1..N\}$. This leads to mutually exclusive regions of the abdomen being derived from a particular source volume. Seams are implicitly defined as the transitions between regions with different source IDs or labels. The regions are chosen in a way that attempts to minimize intensity and gradient differences between volumes at the seams. We would expect these differences to be minimal in areas with limited movement and well aligned features. Next we will express seam selection as a discrete optimization problem in the form of (3.34), which will also be minimized using the graph cut based alpha expansion technique. In this instance each element of the vector variable $\mathbf{x}$ corresponds to a single voxel in the common scene coordinate system.

In seam selection preprocessing step, where we wish to label each voxel in the global image volume with its corresponding source volume, the random variable $\mathbf{x}$ described above may take on values from 1 to N, where N is the number of sources. The sources are partially overlapping and may comprise the entire abdominal region of the subject. The unary terms prevent the optimization process from labeling the voxels in the global image volume with source IDs, which contain no meaningful data in that location. The source ID is an element of $\mathcal{L}$ and indicates which volume the composite should acquire its data from at this voxel location. We have redefined T in this section so it no longer represents a registration transformation. As before the 3D ultrasound source volumes are designated as $I_1.....I_N$ where $I_n : \mathbb{R}^3 \rightarrow \mathbb{R}$ and now let $T : \mathbb{R} \rightarrow \mathbb{R}^3$ map the voxel index to its 3D spatial coordinates. Because of the common scene space is comprised of a number of irregularly shaped overlapping 3D ultrasound volumes, not every voxel location in this space lies in the domain of each source volume. The unary term found in (3.34) is used by our mosaicing application to guarantee that the entire domain of the composite volume, $\bigcup_{n=1}^{N} \text{Domain}(I_n)$, is comprised of valid image data and is calculated as

$$\Phi_i(x_i) = \begin{cases} 0 \text{ if } T(i) \in \text{Domain}(I_{x_i}) \\ \infty \text{ Otherwise} \end{cases} \tag{3.39}$$

This term ensures that the voxel in the composite volume associated with x_i is assigned a source that coincides with this location.

The pairwise terms control where the surface boundaries are placed between overlapping volumes in the composite image. We wish to limit the intensity and gradient differences across the boundary, which will result in the most continuous composite volume. This approach has been used in an interactive digital photo montage framework [2] and extended for use in 3D ultrasound mosaicing here. The following expression gives the pairwise interaction terms

$$\begin{aligned} \Psi_{ij}(x_i, x_j) = & \left\| I_{x_i}(T(i)) - I_{x_j}(T(i)) \right\| + \left\| I_{x_i}(T(j)) - I_{x_j}(T(j)) \right\| + \\ & \left\| \nabla I_{x_i}(T(i)) - \nabla I_{x_j}(T(i)) \right\| + \left\| \nabla I_{x_i}(T(j)) - \nabla I_{x_j}(T(j)) \right\| \end{aligned} \tag{3.40}$$

where i,j are neighboring voxels. The first term in (3.40) measures the intensity difference if neighboring voxel locations in the composite volume are labeled with different sources. If neighboring locations in the composite volume are labeled with the same source then this term is 0 and no seam runs through this region. The end result is the creation of a visually appealing and probably more anatomically correct composite image volume because overlapping sources are stitched together in regions that closely match in intensity value. The second term in (3.40) penalizes gradient differences along the seams in the composite volume which limits discontinuities along surfaces between neighboring sources. These terms are necessary for the creation of large composite volumes because subject movement and organ deformation during acquisition coupled with ultrasound imaging artifacts result in overlapping sources with large inconsistencies. These inconsistencies presented themselves as large intensity differences between multiple overlapping volumes in certain regions which couldn't be eliminated with non-rigid image registration. This was especially apparent when stitching fetal ultrasound volumes together.

It is easy to show that the terms in (3.40) satisfy the submodularity constraint given in (3.35). We see that $\Psi_{ij}(1,1)$ evaluates to 0 because it signifies the neighboring voxels both taking the new label specified by alpha. Using the triangle inequality, $\|X+Y\| \leqslant \|X\| + \|Y\|$, we get the following result

$$\begin{aligned}
\left\|I_{x_i}(T(i)) - I_{x_j}(T(i))\right\| &\leqslant \|I_{x_i}(T(i)) - I_{\alpha}(T(i))\| + \left\|I_{\alpha}(T(i)) - I_{x_j}(T(i))\right\| \\
\left\|I_{x_i}(T(j)) - I_{x_j}(T(j))\right\| &\leqslant \|I_{x_i}(T(j)) - I_{\alpha}(T(j))\| + \left\|I_{\alpha}(T(j)) - I_{x_j}(T(j))\right\|
\end{aligned} \tag{3.41}$$

This proves the submodularity of the energy function. Ideally neighboring locations in the composite ultrasound image volume should come from the same source volume if possible.

Substituting the newly defined terms $\Phi_i(x_i)$ and $\Psi_{ij}(x_i, x_j)$ from (3.39) and (3.40) into (3.34) and then optimizing using alpha-expansion results in a labeling that can be used to produce a mosaiced volume as well as guide group-wise registration. This labeling implicitly defines the 3-dimensional seams between overlapping volumes in the common scene coordinate system. Let the final labeling be represented as $\Gamma(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}$ where $\mathbf{x} \in \mathbb{R}^3$ is a location in the common scene. In the regions of the composite volume where no sources intersect, that is $\Gamma(\mathbf{x}) = 0$, because there is no information available to determine what the mosaiced volume should look like here.

This function defines how the individual sources could be distributed in the composite volume in an intelligent way to minimize artifacts caused by gradient and intensity discontinuities between volumes. It is considered a mosaicing function and we wish to use it as a guide for the registration algorithm.

The basic idea is to consider a fixed region in the neighborhood of the implied seams and only compute the similarity metric from (3.30) inside this region. The 2nd order regularization will guide the transformation function outside of areas which aren't included in the similarity metric calculation. There are many advantages to performing registration of multiple overlapping ultrasound volumes in this way. Firstly the performance of group-wise registration in terms of computation time is greatly improved due to the decreased complexity of the graph representing the registration problem. Secondly the influence on the registration process from shadowing artifacts and fetal movement can be minimized by registration around the seams. For instance it might be beneficial to ignore movement of the baby's extremities during registration and only concentrate the baby's abdominal movement. Heavily shadowed regions could easily be penalized by the seam selection algorithm above resulting in a mosaicing function which produces few shadows in the composite volume and thus will mostly be ignored by the registration algorithm. Now we will explain how the MRF function in equation (3.30) is altered to take into account the mosaicing function that has been calculated as a preprocessing step.

Since the registration energy expressed by (3.30) is composed of many thousands of discrete terms the simplest way to introduce the mosaicing function is to use it to crop the similarity terms that may negatively affect the registration outcome. Thus certain $\mathrm{SSD}_b^{i,j}\left(\mathbf{d}_{i,b} - \mathbf{d}_{j,b}\right)$ terms are removed from the registration energy if they lie too far away from the junction of multiple image volumes. The process for cropping the pairwise similarity terms can be formalized after introducing a few utility functions. Let $\mathcal{D}_i\left(\mathbf{x}\right)$ be a distance function which measures the minimum distance from point $\mathbf{x}$ to the region in the composite volume labeled with source i. Thus $\mathcal{D}_i\left(\mathbf{x}\right)$ can be defined using set notation as

$$\text{let } \Omega_i = \left\{ \mathbf{x} | \mathbf{x} \in \mathbb{Z}^3 \wedge \Gamma(\mathbf{x}) = i \right\}$$
$$\mathcal{D}_i(\mathbf{x}) = \begin{cases} 0 \text{ if } \mathbf{x} \in \Omega_i \\ \inf\limits_{\mathbf{y} \in \Omega_i} \|\mathbf{x} - \mathbf{y}\| \text{ if } \mathbf{x} \in \Omega_i^c \text{ where } \Omega_i^c \text{ is the complement of } \Omega_i \end{cases} \tag{3.42}$$

Thus Ω_i is the set of voxels in the composite image volume which has been designated to source i and the infimum measures the closest distance from any voxel to this region. Next a parameter will be defined that controls the number of terms which are cropped from the registration energy. This parameter has the effect of limiting the registration to regions surrounding the source boundaries. Let ς be the maximum distance from a seam that we wish to consider when calculating the similarity metric. This means that the width of the total region influencing the pair-wise alignment terms will be 2ς. The last step before expressing the new focused registration energy will to define an indicator function that is used to removes blocks from consideration when they are greater than 2ς from a seam. The indicator function $\delta_b^{i,j}$ is defined as

$$\delta_b^{i,j}(\mathbf{c}_b) = \begin{cases} 1 \text{ if } \mathcal{D}_i(\mathbf{c}_b) \leqslant \varsigma \text{ and } \mathcal{D}_j(\mathbf{c}_b) \leqslant \varsigma \\ 0 \text{ otherwise} \end{cases} \tag{3.43}$$

Using this indicator function the focused registration energy is can now be defined as

$$\mathcal{E}(\mathbf{d}) = \sum_{\substack{\{\mathbf{d}_{i,b}, \mathbf{d}_{j,b}\} \\ \in \mathcal{N}_{Inter}}} \delta_b^{i,j}(\mathbf{c}_b)\, \mathrm{SSD}_b^{i,j}(\mathbf{d}_{i,b} - \mathbf{d}_{j,b}) + \frac{\lambda}{B} \sum_{\substack{\{\mathbf{d}_{i,a}, \mathbf{d}_{i,b}, \mathbf{d}_{i,c}\} \\ \in \mathcal{N}_{Intra}}} \left| \frac{\mathbf{d}_{i,a} - 2\mathbf{d}_{i,b} + \mathbf{d}_{i,c}}{h^2} \right| \tag{3.44}$$

The only difference between (3.44) and (3.31) is the addition of the indicator function used to cancel the unwanted block-based similarity terms. The indicator function can also be thought of as eliminating specific inter-volume edges from the MRF graph defining the problem in Figure 3-2. As ς increases, more blocks are considered by the registration algorithm until all overlapping regions influence the registration solution. This would be the complete group-wise registration algorithm that is usually considered. The advantage of this formulation for focused registration is that the regions in which the similarity metric is calculated may have greater than two source volumes. There is no constraint on the number of volumes considered at each control point. Without further modification of the registration energy it is possible to have regions where only pair-wise alignment is performed, coupled with regions that consider the complete group-wise registration problem. Thus different degrees of volume interaction may be optimized during the same execution.

This concept is demonstrated in Figure 3-5 which shows the results of image volume stitching using our algorithm. In the upper left is a color coded labeling which assigns each voxel in the common scene coordinate system to a particular source volume. This image is just a 2D slice of $\Gamma(\mathbf{x})$ which was discussed above. The upper right corner shows the distance maps $\mathcal{D}_i(\mathbf{x})$ for each source volume which can easily be calculated using built-in functions common in most image processing toolboxes. The most interesting part of Figure 3-5 is the center image (a) which shows varying degrees of volume interaction and also demonstrates how these regions are dependent on the parameter ς. The numbers in the center image indicate which volumes are considered during the similarity measure calculation. For the particular value of ς chosen in this example there are four regions where pair-wise alignment is measured and two regions where group-wise alignment is measured using three out of the four volumes. We see that in the upper right corner (a) volumes 1,3,4 interact with each other and in the lower right volumes 1,2,3 interact. The dark regions in the image (a), which contain no numbers, are solely influenced by the 2nd order regularization terms in (3.44). By constructing the registration problem as a discrete Markov random field we find that the focused registration concept, which has proven useful during 3D volume mosaicing, can easily be integrated into our algorithm.

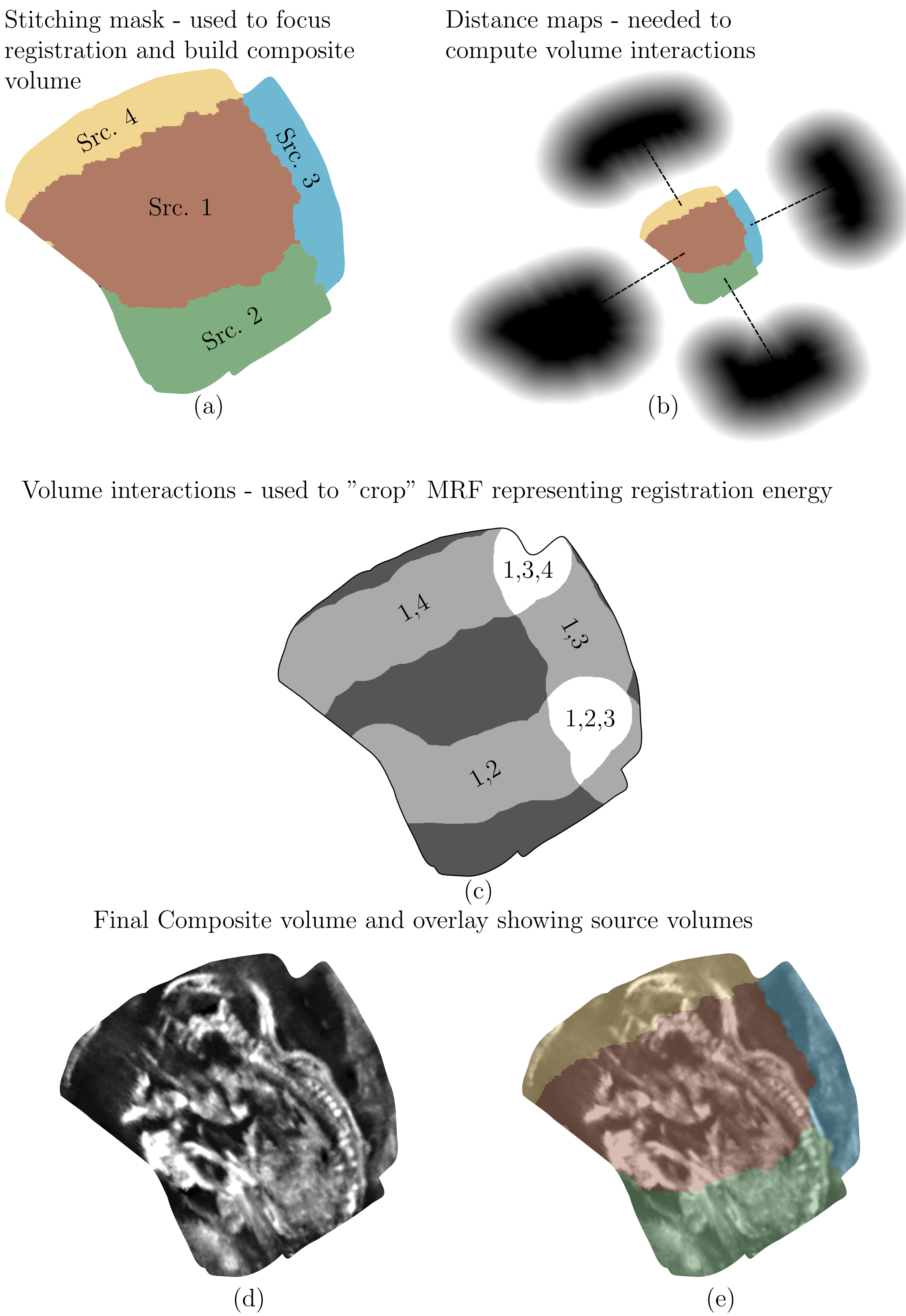


Figure 3-5: Mosaicing function is used to focus registration.

3.5.1 Summary of registration concepts and application to 3D mosaicing

This section will briefly summarize the registration and seam concepts discussed previously with the overall goal of generating composite 3D volumes from overlapping scans. Simulated and clinical results will be presented in the next section. Given a set of partially overlapping 3D volumes we desire to form a mosaiced volume which fills as much of the composite coordinate system as possible. To this end the first step is to perform seam selection using the algorithm discussed above. This results in labeled regions of the composite coordinate system where each label corresponds to a particular source volume. At this point it is possible to create a mosaiced volume by simply assigning each voxel in the composite coordinate system with the intensity value from its designated source. Let the final image be designated as $\mathrm{F}(\mathbf{x})$ which is constructed as follows,

$$\mathrm{F}(\mathbf{x}) = \begin{cases} I_{\Gamma(\mathbf{x})}(\mathbf{x}) & \text{if } \Gamma(\mathbf{x}) > 0 \\ 0 \text{ otherwise} & \end{cases} . \tag{3.45}$$

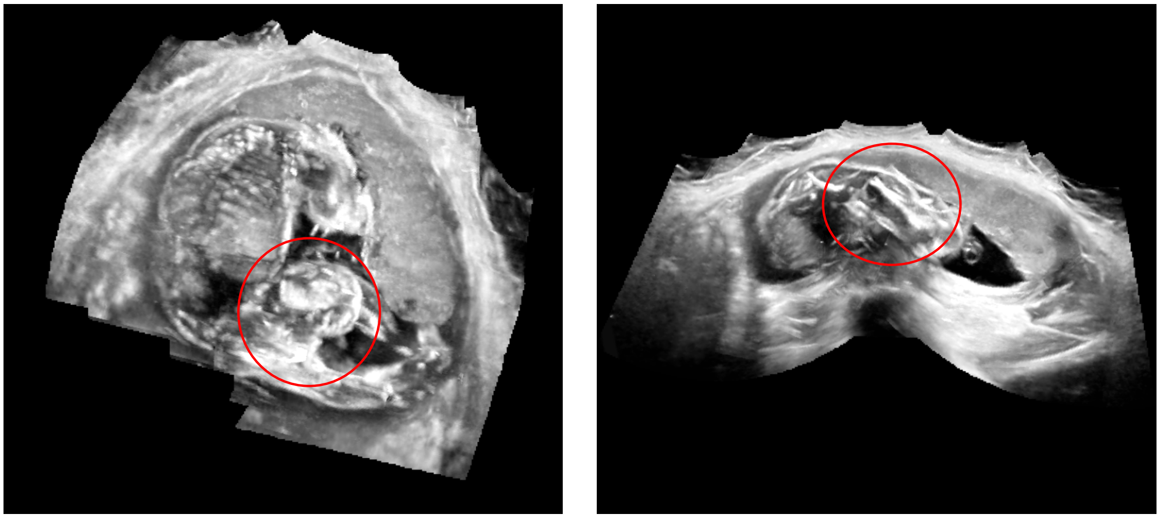
where $\Gamma(\mathbf{x})$ is the final labeling. In order to improve the visual quality of the composite volume we perform group-wise registration before mosaicing by minimizing the energy of equation (3.44) using the parallel alpha-expansion technique. After group-wise registration is completed and before the final mosaic is constructed the seams are recalculated to account for large deformations of the source volumes which may be part of the registration solution. Finally equation (3.45) is used to construct the composite volume. We have also experimented with a Poisson blending technique, which improves the visual quality around the seams without eliminating edges or causing substantial blurring.

It is also possible to construct ultrasound mosaics using spatial compounding techniques. For example, one of the simplest methods is to use the maximum intensity value of all overlapping sources at each voxel in the composite. This type of approach works well when creating mosaics of stationary organs such as the liver, or those with predictable motion patterns such as the heart; however, they give poor results when faced with large non-deterministic movement, which is the type that occurs during fetal ultrasound scanning. Since our clinical application is the construction of fetal mosaics for use in a training simulator our proposed method outperforms spatial compounding techniques, which require almost perfect alignment between volumes in the composite space. Figure 3-6 compares the mosaicing results of our proposed method with a simple spatial compounding approach. The first row shows two slices from a composite volume created using the maximum intensity technique, where the red circles indicate poor quality in the constructed mosaic. In this case fetal leg movement between scans was too great to correct with non-rigid registration, resulting in a composite volume that is unsuitable for training purposes. The second row of Figure 3-6 shows the same source volumes stitched using our proposed mosaicing function. The errors seen in the top two slices have been correctly dealt with resulting in a composite volume that can be included in the simulator. In the last row the mosaicing function is overlaid onto the composite slices showing how the five overlapping volumes are stitched together.

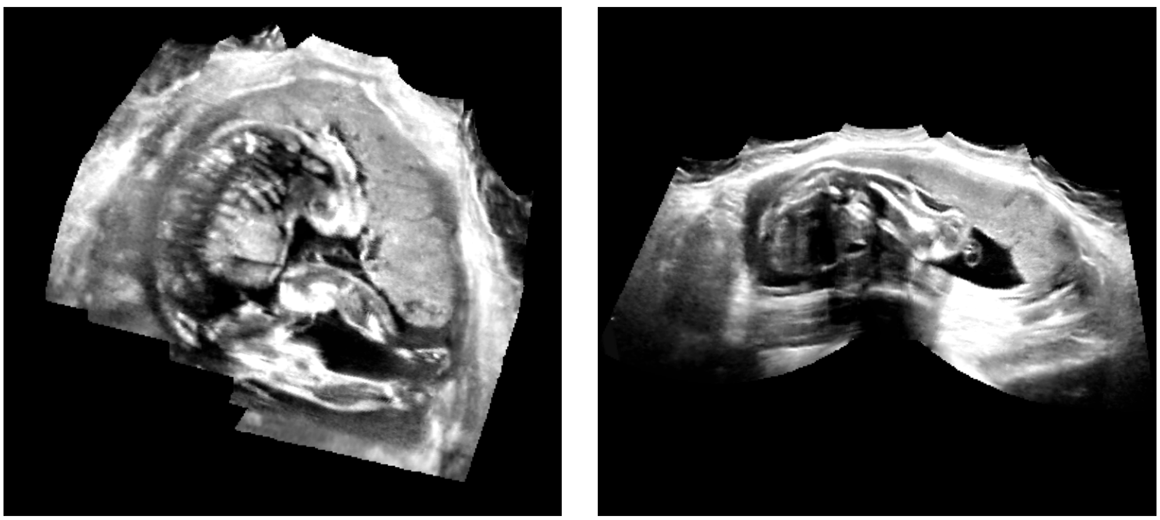
3.5.2 Discussion on consistent image registration in MRF framework

In group-wise registration it is desirable to transform the source volumes into an average reference shape in the scene coordinate system. This concept has been discussed in the literature concerning both the continuous [9] [5] [53] and discrete domains [76]. Our ultrasound mosaicing problem doesn't require consistent image registration techniques to achieve good results but since we are using a group-wise method they should be mentioned. In the pairwise case ($N = 2$) the registration problem becomes the optimization of two control point lattices which are initially overlaid on top of each other in the scene coordinate system. In this example we see that the resulting transformed sources will represent the true average shape of both volumes if $D_1(\mathbf{x}) + D_2(\mathbf{x}) = 0$ for $\forall \mathbf{x} \in \Omega$ where Ω represents the domain of the composite

Coronal/transverse slices formed using maximum intensities in overlapping areas:



Coronal/transverse slices formed using mosaicing function:



Mosaicing function overlaid onto slices from composite volume:

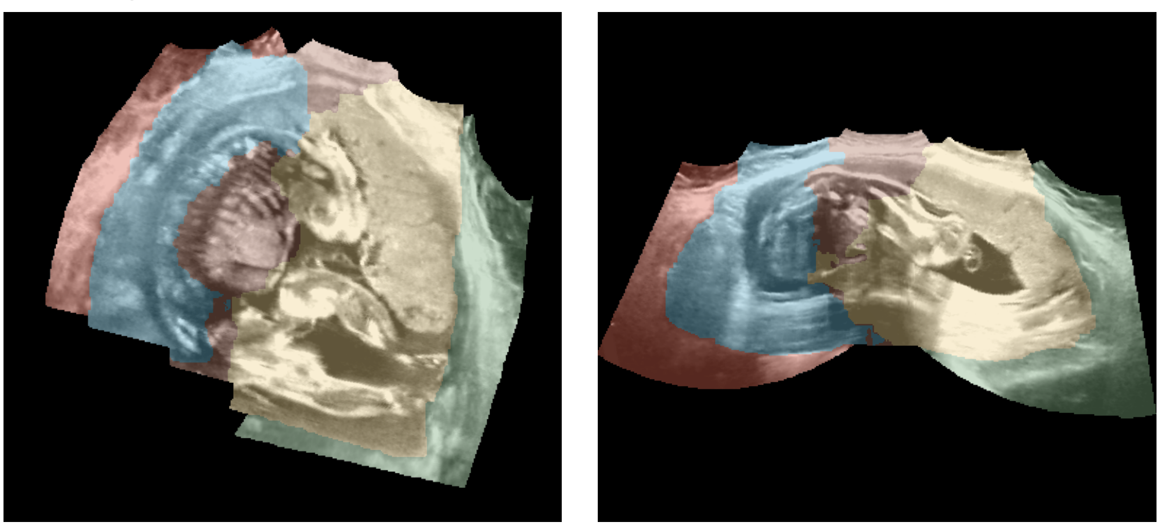


Figure 3-6: Comparison between spatial compounding technique and our proposed mosaicing function using fetal ultrasound data

volume in the scene coordinate system. For group-wise registration problems where $N > 2$ the condition for consistent registration is formulated as

$$\sum_{n=1}^{N} D_n(\mathbf{x}) = 0 \qquad \forall \mathbf{x} \in \Omega. \tag{3.46}$$

Because the transformation at each point is linearly dependent on the control point displacements the condition in (3.46) can be simplified to an expression only involving these values. If the sum of control point displacements at each grid location is zero then (3.46) holds as long as the transformation is modeled in the form given by (3.7) [9]. Thus the final registration problem including the new consistency constraint is

$$\begin{aligned} &\min \mathcal{E}(\mathcal{D}) \\ &\text{subject to } \sum_{n=1}^{N} \mathbf{d}_{n,b} = 0 \qquad \forall b : \mathbf{c}_b \in \mathcal{C}. \end{aligned} \tag{3.47}$$

There are two possible ways to incorporate the new constraint into the MRF framework that was presented in this chapter. The first is to add a consistency term to the registration energy in (3.30) in order to penalize transformations that violate the constraint of (3.47). This term might look like $\kappa \sum_{n=1}^{N} \sum_{b}^{B} \mathbf{d}_{n,b}$. The problem with this new term is that for registration regions encompassing many overlapping volumes it becomes dependent on higher order cliques which are known to put a tremendous computational burden on current MRF optimization algorithms. The consistency term of the new MRF would add numerous edges to the already complicated graph from Figure 3-2. Much research is currently underway to develop optimization methods for high order Markov Random Fields due to their rich descriptive features. However, for our task we did not find the incorporation of a consistency term necessary to produce good results. A more efficient method to obtain results that satisfy (3.47) is to project the solution computed from the optimization of (3.30) onto the hyper-plane where the linear system of consistency constraints is satisfied. This can be written as a set of three quadratic optimization problems, one for each dimension, for which many efficient solvers exist. We found that the projection can be computed in one or two seconds when dealing with multiple overlapping clinical ultrasound volumes. If we form a binary matrix $\mathrm{A} \in \{0,1\}^{|\mathcal{C}|\times|\mathcal{D}|}$ to represent the linear consistency constraint of (3.47) and write the solution set $\mathcal{D}$ as a collection of vectors $\mathbf{d}$ the quadratic optimization problem becomes

$$\begin{aligned}
&\min_{\mathbf{x}^x} \|\mathbf{x}^x - \mathbf{d}^x\| && \text{subject to}\,\mathrm{A}\,\mathbf{x}^x = 0 \\
&\min_{\mathbf{x}^y} \|\mathbf{x}^y - \mathbf{d}^y\| && \text{subject to}\,\mathrm{A}\,\mathbf{x}^y = 0 \\
&\min_{\mathbf{x}^z} \|\mathbf{x}^z - \mathbf{d}^z\| && \text{subject to}\;\mathrm{A}\,\mathbf{x}^z = 0 \\
&\text{projection satisfying constraints:} && \hat{\mathcal{D}} = \left[\mathbf{x}^x\ \mathbf{x}^y\ \mathbf{x}^z\right].
\end{aligned} \tag{3.48}$$

In (3.48) the superscript is used to specify the dimension. After satisfying the consistency requirement for control point displacements in each dimension the final solution vector $\hat{\mathbf{d}}$ is obtained by concatenating these results. The matrix A is sparse and can be constructed by

$$\mathrm{A}_{b,|\mathcal{C}|(n-1)+b} = \begin{cases} 1 & \text{if volume } n \text{ has a control point at } \mathbf{c}_b \\ 0 & \text{otherwise} \end{cases}. \tag{3.49}$$

Each row of this matrix corresponds to a single control point designated by b and sums the displacement values from all the sources which overlap at this control point. The minimization problems in (3.48) can also be rewritten as $\min \frac{1}{2}\left(\mathbf{x}^d\right)^\top \mathbf{x}^d - \left(\mathbf{d}^d\right)^\top \mathbf{x}^d$ where $d \in \{x,y,z\}$ specifies dimension. Coupled with the equality constraint this is a quadratic optimization problem that can be easily solved. This formulation will produce a result satisfying the constraints of (3.47); however, the registration energy may increase. This is important in cases where the solution vector must be altered substantially because the modification of the transformation function could cause unchecked spikes in the regularization energy. A remedy for this problem could be to guide the solution vector toward the consistency constraint's hyper-surface at different points during the optimization procedure instead of only at the end. The system in (3.48) may be solved during optimization whenever the functions $\mathrm{A}\,\mathbf{x}^x, \mathrm{A}\,\mathbf{x}^y, \mathrm{A}\,\mathbf{x}^z$ cross some user specified bound. Further optimization after reinitializing the solution vector using (3.48) will correct the spike in regularization energy and also lead to a set of transformations which are closer to representing the mean shape. We do not need the solution of the group-wise registration to satisfy the consistency requirement in order to produce good results for our application so extensive experiments using (3.48) have not been conducted; however in simulated trials using a finite element model of the Visible Human Dataset the algorithm produced transformations that aligned the overlapping volumes as well as satisfied (3.47). More experimentation will be required to suggest the frequency and placement of the re-initialization procedure in the group-wise registration algorithm presented in this chapter.

Chapter 4

Group-wise registration and seam selection: experimental results

In this chapter we present experimental results to demonstrate the performance of the group-wise registration algorithm. It is initially tested on deformed abdominal image volumes generated using a finite element model of the Visible Human Female to show the accuracy of its calculated displacement fields. Also results using real ultrasound data from an abdominal phantom are presented. Finally composite obstetrics image volumes are constructed using clinical scans of pregnant subjects, where fetal movement makes registration/mosaicing especially difficult. In addition, results are presented suggesting that a fusion approach to MRF registration can produce accurate displacement fields much faster than standard approaches.

4.1 Quantitative results using Visible Human Dataset

In order to validate this procedure we constructed a finite element model (FEM) using the Visible Human Female dataset [1] then subsequently deformed it with varying pressure applied to the surface in order to produce three overlapping and uniquely deformed image volumes. The FEM simulations, which produced each of the deformed volumes, were done using the software package Comsol. Nodes on the surface of the model, which were identified to be adjacent to the current transducer scan path, were displaced into the body approx. 25 mm. in order to simulate transducer pressure. Displacement was also applied internally in the direction perpendicular to the transverse plane to simulate breathing. The kidneys, liver, stomach, and intestines were segmented during construction of the FEM model using ITK-Snap [86], and subsequently given appropriate mechanical properties to produce a more realistic deformation field. This procedure simulates the process of mosaicing partially overlapping 3D image volumes with large deformations. Although these images are visually different from ultrasound and do not suffer from the same artifacts, this framework provides a way to quantify the performance of the mosaicing/registration algorithm on multiple overlapping 3D volumes by using the deformation fields provided by the finite element simulations as the gold standard. The voxels in the Visible Human Female data that we used in this experiment are cubic with 1 mm sides. The procedure is outlined in Figure 4-1. The first step is to uniquely deform the finite element model three times and extract distinct but overlapping regions from each deformation. This results in a transformation function from the deformed coordinate system to the original coordinate system of the Visible Human Female for each region. The colored rectangles shown identify the individual regions extracted during the deformation process. We applied a maximum deformation of 3 centimeters to the center region and 1.5 centimeters to the rightmost region. The leftmost region was undeformed. Breathing artifacts were also included by applying a force to the model in the inferior direction. The right and left volumes each partially overlap with the middle volume but they have no voxels in common with each other; thus, the registration algorithm should be able to use the information contained in these 2 overlapping regions to bring all 3 volumes into alignment for the purpose of mosaicing. The two top images in Figure 4-2 show stitching results before the registration procedure has been applied. It is apparent that the misalignment is significant at the organ boundaries as well as the surface of the skin and will require large displacement fields to correct.

Using the simultaneous registration procedure described earlier we create an MRF inference problem which can be solved using quadratic pseudo Boolean optimization that will bring each volume into alignment with its neighbor in the region of overlap.

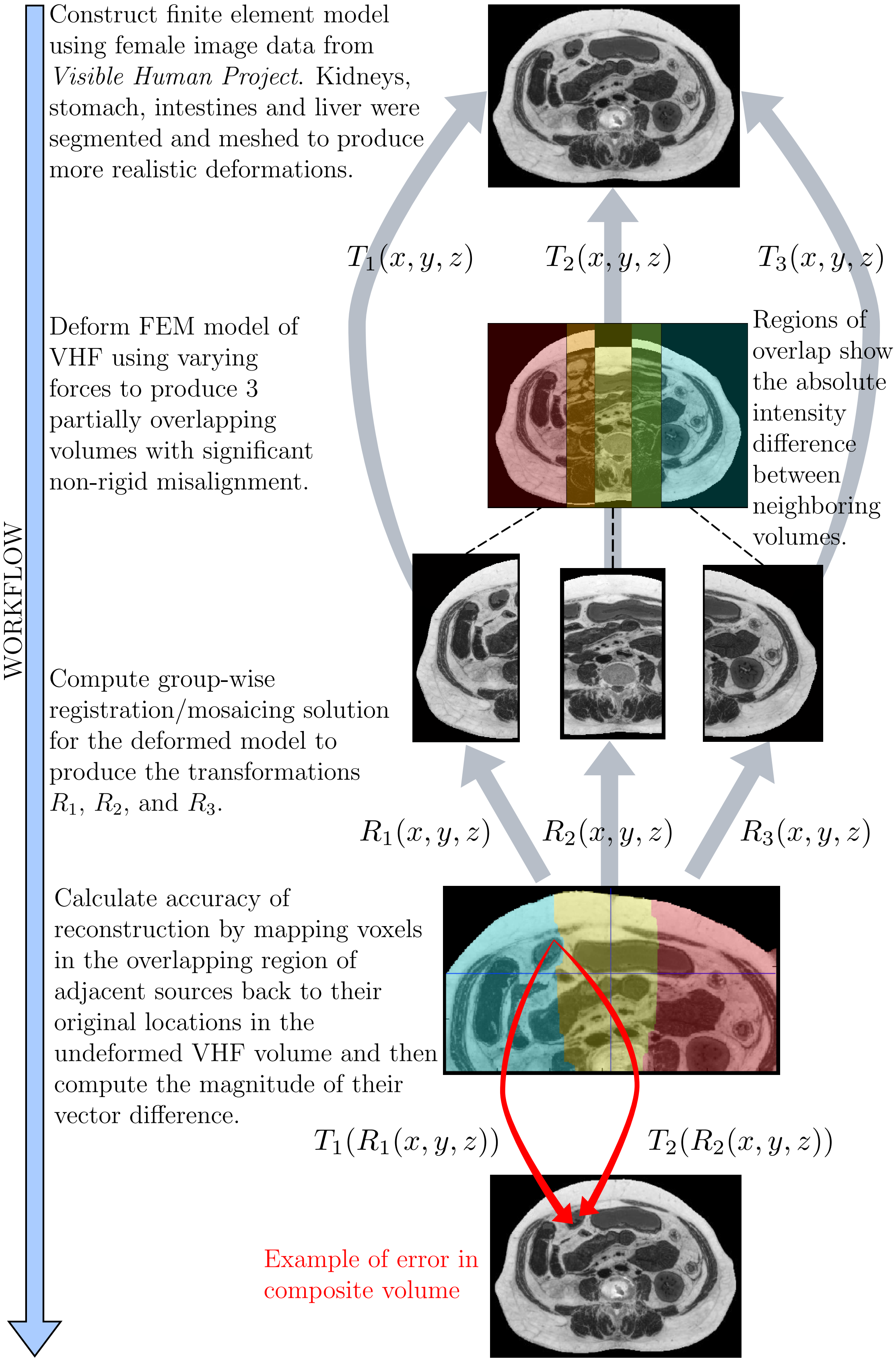


Figure 4-1: VHF validation procedure used to quantify mosaicing algorithms performance

Every volume is elastically linked together by the MRF construction so displacement in the left volume can be influenced by displacement in the right volume even though they have no overlapping region. An implicit linkage between two non overlapping image volumes is intuitive in the case of mosaicing because structures may span more than two volumes. The output of the registration algorithm in this case is three displacement fields, with one corresponding to each region of the visible human female that was extracted during the finite element modeling. A slice from the fully mosaiced volume is shown in the bottom of Figure 4-2. The result is a continuous 3D volume where the extreme discontinuities in organs spanning more than one volume have been corrected.

This situation is similar to what is encountered during the several minute long scanning in a clinical setting where the goal is to acquire 8-10 overlapping ultrasound volumes encompassing the entire abdomen. Patient breathing and fetal movement result in large deformations between neighboring ultrasound volumes, and this experiment attempts to simulate the process of mosaicing highly deformed structures.

Qualitative results for the registration quality of our algorithm are shown in Figure 4-2. However, a finite element model allows us to quantify the performance of our algorithm by calculating the reconstruction error at each voxel location in a region surrounding the seam, which is shown in the bottom half of Figure 4-1. By composing the transformation function from the registration algorithm with the transformation from the finite element model, we can form a function which maps coordinates in the final mosaiced volume to coordinates in the original undeformed Visible Human Female dataset. Anatomical consistency can be defined to mean the error (in mm) associated with voxels near the seam, where the voxels are common to more than one registered volume. It is calculated by mapping voxel locations in the overlapping regions back to the coordinate system of the undeformed finite element model by utilizing the composed transformations shown in Figure 4-1. Locations in overlapping regions are associated with ≥ 2 composed transformations; ideally, locations near the seam between two or more regions in the mosaiced volume should map to the same location of the original undeformed coordinate system, regardless of which composed transformation is used.

When a single voxel maps to two different locations in the original coordinate system, depending on the transformation, the distance between these two points is calculated and becomes the metric for consistency in the region surrounding the seam. Good anatomical consistency implies that anatomy has been reconstructed correctly from multiple overlapping image volumes. Figure 4-3 shows slices where consistency has been measure before and after the registration process has been completed. When no non-rigid registration has been completed it is expected that the anatomical consistency between volumes would be very low and this is shown in the top row of Figure 4-3. The overlapping regions between the 3 volumes are shown with the intensity corresponding to the degree of anatomical inconsistency at each voxel. After registration, when the structure has been mosaiced correctly, the consistency has significantly improved, as shown in the bottom row of Figure 4-3. There is some error along the skin surface and this is most likely due to the fact that there are no structures to align in the fatty layer.

The accuracy/performance results of our Visible Human Female experiments are compiled in Table 4.1. The mean error in consistency, calculated over the regions of overlap shown in Figure 4-3, is given in column 2. The standard deviation of this error is given in column 3. For these calculations we don't consider the error in the external fatty layer because the absence of structures made it difficult to register. The results show the algorithm's ability to reconstruct anatomy from multiple partially overlapping volumes with sub-pixel accuracy using this dataset. Experiments with simulated data were used to verify the correctness of the algorithm and to measure the effects of parallelization on speed/accuracy. It was not designed to mimic the mosaicing process when using clinical ultrasound data which is discussed in the next section. Interestingly we found that the parallelization of the registration algorithm using the fusion technique did not affect the accuracy of the reconstruction as much as we thought it might. Splitting the registration between one, four, and eight processes still produced highly accurate mosaics of the finite element model and the speedup was impressive. The registration time in Table 4.1 includes the processing time to build

Table 4.1: Visible Human Female mosaicing accuracy/speed

# Processes	Mean Error (mm)	Std. Dev. (mm)	Exec. Time (secs)	Speedup
1	0.8067	0.8566	3176.0	1.00
4	0.8278	0.6035	1470.0	2.16
8	0.7271	0.4585	811.6	3.91

the similarity metric data-structure, which is done before any MRF optimization, and thus could be shortened by increased parallelization of that calculation. This affects the speedup factor in Table 4.1.

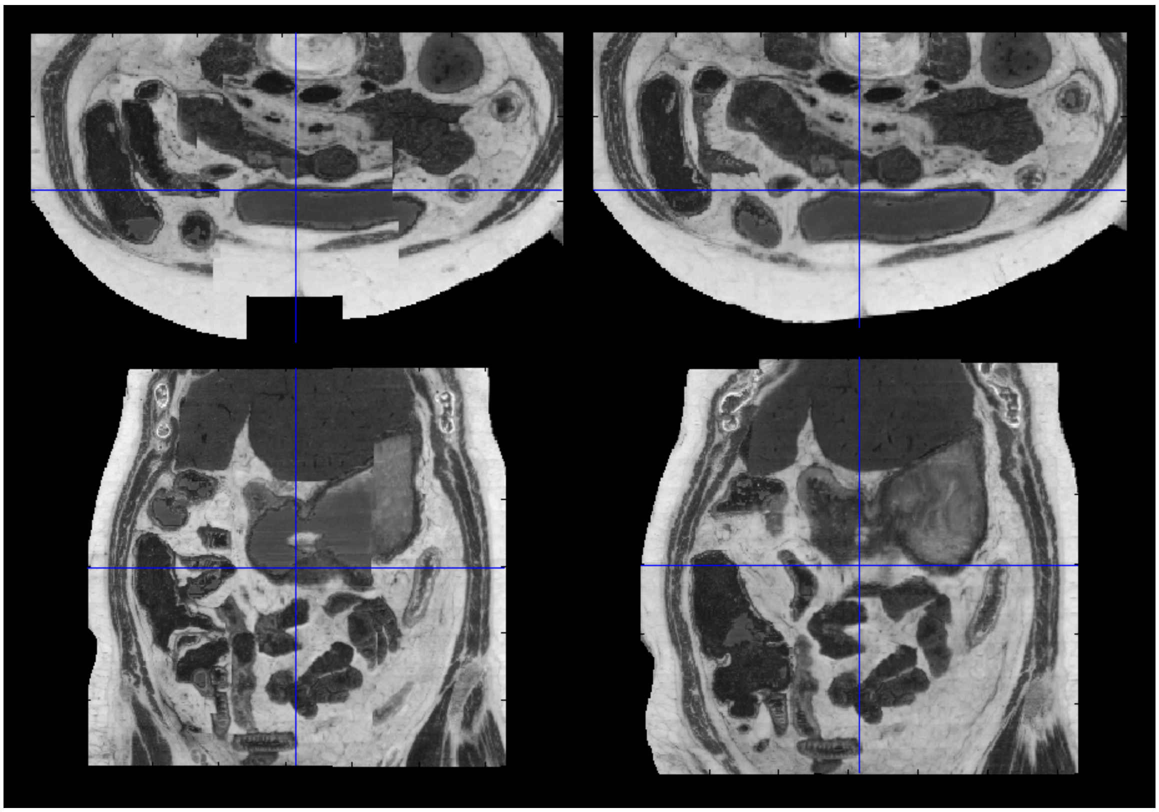
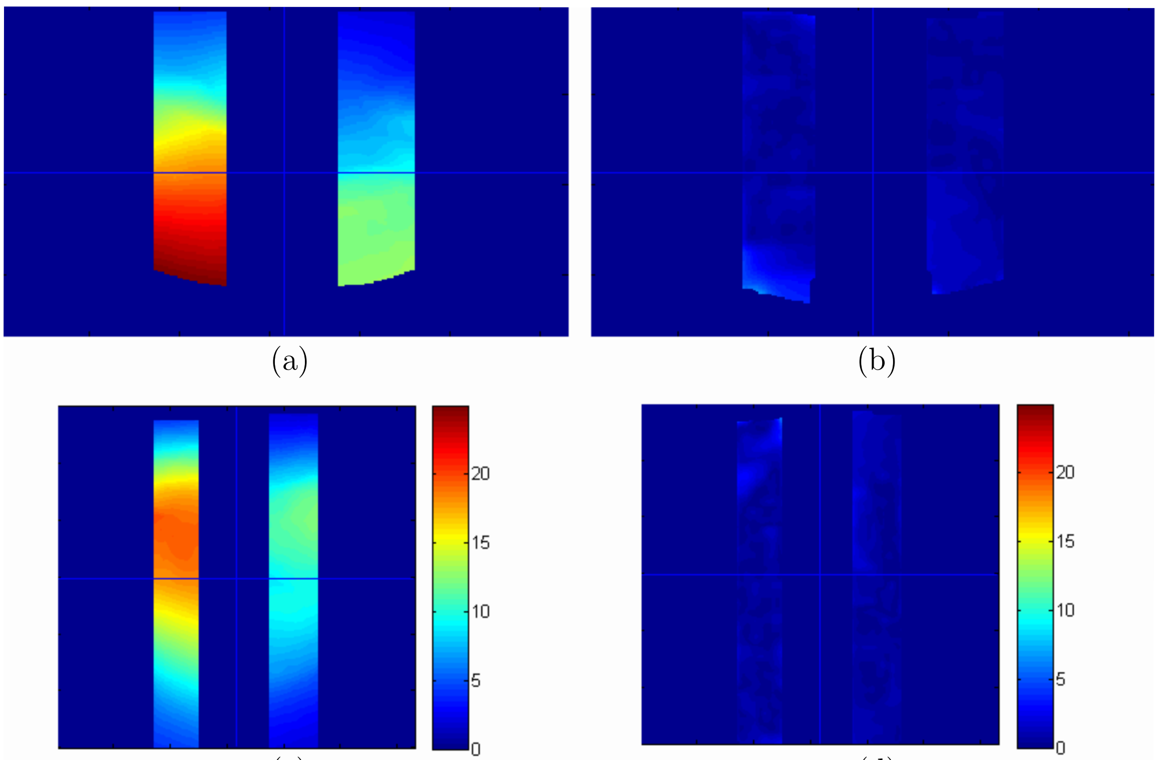


Figure 4-2: VHF stitching results showing the qualitative performance of our algorithm. The left half shows slices from the mosaiced volume when no registration is performed. The right half demonstrates how our algorithm can reconstruct the anatomy seamlessly, using multiple partially overlapping volumes. All images were formed using $\Gamma(\mathbf{x})$
.



(a) (b)

(c) (d)

Figure 4-3: Vector magnitude of misalignment in overlapping regions of Visible Human Female FEM before and after group-wise registration. (a) Axial slice of model before registration. (a) Axial slice of model after registration. (c) Coronal slice of model before registration. (c) Coronal slice of model after registration. Error is shown in mm.

4.2 Validation of group-wise registration algorithm using FAST/ER abdominal phantom

This section describes the procedure that we have used to measure the effectiveness of the group-wise registration algorithm on ultrasound data acquired from an abdominal phantom. Previously we have showed that the algorithm can effectively register 3 overlapping volumes generated from the Visible Human Female dataset; however, these volumes don't present the difficulty that US image volumes do. Our goal in the previous section was to quantify and compare the registration accuracy for varying degrees of parallelization as well as to demonstrate the ability of the algorithm to handle large deformations at the surface. This section compliments those results by demonstrating that our algorithm can effectively register several partially overlapping ultrasound volumes attained from an US phantom. An abdominal US phantoms is ideal for this verification because the internal anatomy remains static between overlapping sweeps, thus the non-rigid misalignment between volumes can be more easily controlled. Also visualizing registration improvement is more difficult with live subjects due to artifacts caused by abdominal gases or fetal movement between sweeps, which are uncontrollable. After the acquisition procedure we applied an additional large non-rigid deformation to each volume in order to make the experiment more realistic. This is a necessary because scanning a phantom doesn't the simulate motion artifacts caused by a patient's movement/breathing and the deformation resulting from the non-constant pressure of the transducer against the skin.

The abdominal phantom used in this section was borrowed from the Kyoto Kagaku Co. and was designed to provide training in the FAST (Focused Assessment with Sonography for Trauma) procedure. To this end it contains various internal injuries that are detectable via US such as the presence of free intra-peritoneal or pericardial fluid in traumatic patients. The Philips iU22 Ultrasound system was used in combination with a C5-1 convex array transducer, on which was mounted an Ascension trakSTAR 6 DoF position sensor. The video output of the scanner was connected to a laptop running Stradwin software [78]. Sonographers at UMass Medical School

Number of overlapping volumes per voxel in transverse slice of FASTFAN:

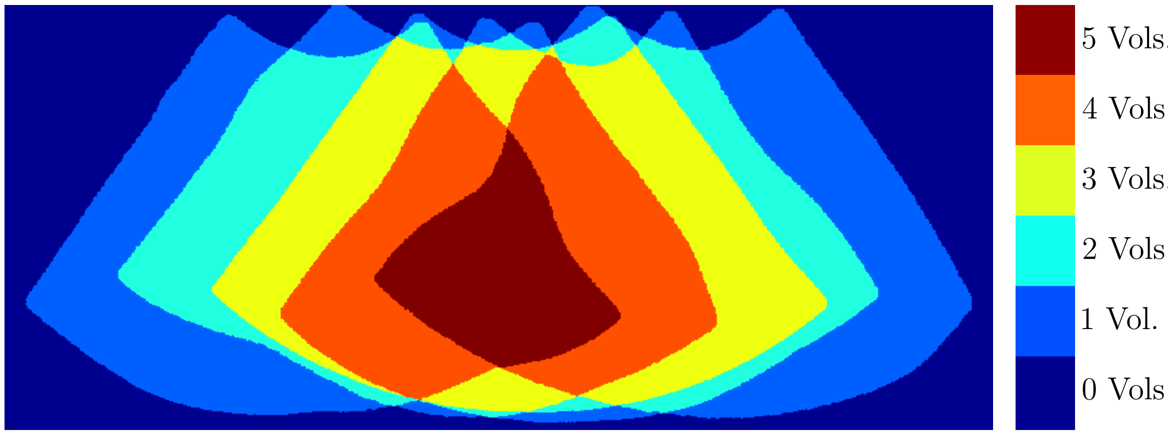


Figure 4-4: Degree of overlap increases as we move towards center mass. An intelligent group-wise registration algorithm designed for ultrasound mosaicing should be able to deal with varying degrees of overlap.

Table 4.2: Thin plate spline deformation field statistics

	Volume 1	Volume 2	Volume 4	Volume 5
Max disp. (cm)	1.7446	1.9888	1.9430	2.2609
Mean disp. (cm)	0.6828	0.6955	0.6792	0.6737

performed the scanning for us and produced five partially overlapping volumes using swept 3D ultrasound, which encompassed the majority of the phantom. Figure 4-4 shows the degree of overlap at each voxel location in the composite volume. In the outer regions of the composite, where only 2 volumes overlap, a pair-wise similarity metric is used; however we see that as many as 5 volumes overlap at each voxel close to the center; thus our group-wise similarity metric is useful here. Non-rigid motion between overlapping volumes was simulated by applying a random deformation fields to 4 of the 5 volumes using thin plate splines [11]. Each volume was deformed independently of the others by randomly positioning approximately 75 control points inside each volume and subsequently assigning a random displacement value based on a uniform distribution. Table 4.2 provides some statistics on the each volume's unique displacement fields. It should also be noted that the depth of the C5-1 transducer was set to 16 cm. Most common ultrasound mosaicing tools, such as Stradwin [78], generate the composite volume by defining planar boundaries between overlapping volumes. We chose this seam selection method to validate the group-wise registration algorithm due to its popularity and also due to the fact that planar boundaries are unaware of the image misalignment through which they pass, thus the improvement in the composite volume before and after registration can be substantial.

Figure 4-5 shows the procedure used to generate the initial composite volume before our algorithm was applied. The top row of images in Figure 4-5 are slices from the five original volumes obtained using swept 3D ultrasound. Dashed arrows denote the application of the random non-rigid transformations that were based on thin plate splines, while the second row of images show how the original slices are deformed. Thin plate splines are popular because they have a closed-form solution, the interpolation is smooth, and there are no free parameters to adjust. The bending of a thin metal plate is the physical analogy usually used to explain the model. Since 4 out of 5 volumes are randomly deformed, the net effects of the transformations in the overlapping areas are considerable because each adds to the misalignment. Two slices from the composite volume, which was produced using planar seams, are shown in the 3rd row of Figure 4-5. Significant misalignment between volumes is obvious and will need to be corrected before using the data in a US simulator. Instead of sequentially performing pair-wise non-rigid registration and thus inefficiently growing the composite result one additional volume at a time, our technique simultaneously aligns all 5 overlapping volumes from the abdominal FAST/ER FAN phantom at once, requiring only a single energy function, as described in (3.31). The last row in Figure 4-5 qualitatively shows the effect of the random transformations on each original volume. The vector plot in the lower left part of Figure 4-5 is one example of a random thin plate spline deformation, specifically the deformation applied to volume 2. The lower right shows the difference images between each original volume and its deformed version. Major abdominal structures are significantly displaced between original and deformed volumes.

Next we will describe how the various parameters of our algorithm were set for the validation experiment. We used a multi-resolution approach based on a Gaussian pyramid with 3 levels. Each subsequent level was produced by convolving the volume

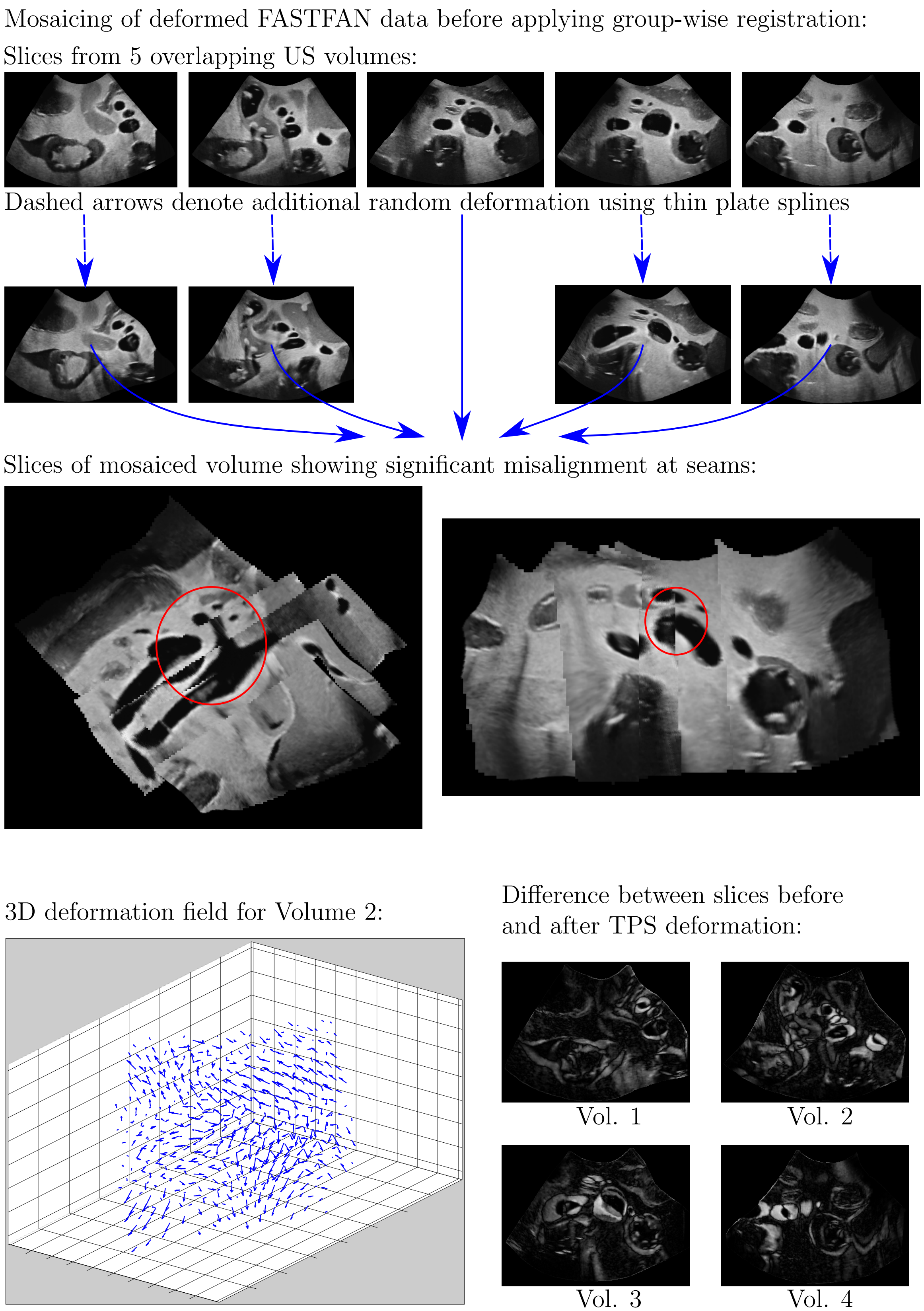


Figure 4-5: Diagram showing the procedure used to deform 4 of 5 overlapping volumes acquired from the FAST/ER FAN

with a 3D version of the typical Gaussian kernel $w = \frac{1}{16}\begin{bmatrix}1 & 4 & 6 & 4 & 1\end{bmatrix}$ and then down-sampling by a factor of 2 [14]. During registration all 5 volumes were permitted to deform in order to achieve alignment in the regions of overlap. It is also possible to fix individual volumes in place if desired, thus transforming the associated pair-wise inter-volume energy terms from (3.31) into unary terms because now one volume isn't permitted to move. Because the C5-1 transducer can produce sector US images with a wide angle, we limited the potential interactions by removing pair-wise terms from (3.31) if they corresponded to a situation where i,j were farther than 2 volumes apart from each other in the composite. This had no effect on the quality of the reconstruction and since MRF optimization speed is directly dependent on the number of nodes in the problem it makes sense to crop these unwanted terms. We performed 6 iterations at the lowest resolution, where the original volumes were down-sampled by a factor of 4, then 8 iterations at the middle resolution and finally 1 iteration at the original resolution, where each voxel was approximately 0.48 mm^3. The regularization parameter in (3.31), denoted by lambda, was allowed to relax during optimization. This ensures that the initial iterations of the registration algorithm results in more rigid transformations, which attempt to globally align the overlapping volumes, before allowing more fluid-like transformations.

We produced qualitative and quantitative results using the experimental procedure described above. Figures 4-6 through 4-8 show improvement after registration, and we see that the misalignment between structures spanning more than one image volume is almost entirely eliminated. The planar seam mask for each slice is also shown in the bottom of the figures and identifies which source each voxel in the composite volume should get its intensity value from. The next result we present quantifies the intensity differences between overlapping volumes in each region of the composite, which have been designated by the planar seam mask. Referencing the seam mask in Figure 4-6, for each region labeled as volume n we use its adjacent volumes $n-1$ and $n+1$ to calculate an error measure as follows

$$\begin{aligned} \text{DIFF}(\mathbf{x}) = \frac{1}{3}\Big(\left|I_{\Gamma(\mathbf{x})+1}(\mathbf{x}) - I_{\Gamma(\mathbf{x})}(\mathbf{x})\right| + \\ \left|I_{\Gamma(\mathbf{x})}(\mathbf{x}) - I_{\Gamma(\mathbf{x})-1}(\mathbf{x})\right| + \\ \left|I_{\Gamma(\mathbf{x})+1}(\mathbf{x}) - I_{\Gamma(\mathbf{x})-1}(\mathbf{x})\right|\Big) \end{aligned} \quad (4.1)$$

where is the mosaicing function described previously. Care must be taken when evaluating (4.1) for voxels that belong to the outer regions of the composite volume. In that case only a single term from (4.1) contributes to the difference measure since that region only has one adjacent volume. The error measure described by (4.1) simply adds the difference errors between all possible combinations of adjacent overlapping volumes at a specific voxel location in the composite volume. The result of (4.1) before and after group-wise registration was performed is presented in Figures 4-6 through 4-8. The improvement is substantial with much of the difference error being removed. This image gives a better visualization of our algorithms ability to correct misalignment between multiple (in our case 5) volumes when compared to the mosaiced results since we aren't restricted to evaluating the error at just the boundary between adjacent volumes.

In order to quantify our registration results for the FAST/ER FAN data, where the correct transformations are unknown and therefore cannot be used for comparison, we compute an adaptation of sum of squared differences error measure by squaring each term in (4.1) and summing all valid voxels. This result is then normalized by the number of contributing terms and can be thought of as a group-wise mean squared error measure. Before registration this measure was 1581 and after group-wise registration it shrank to 707. We achieved an improvement factor of 2.24 over the entire volume. We anticipate that this factor isn't greater due to the noisy nature of ultrasound images; However, the anatomical structures were well aligned after registration, which is the most important aspect in ultrasound mosaicing and which can be seen in Figures 4-6 through 4-8. We should note that incomplete structures residing on a volumes edge posed some difficulty. This can be attributed to the large deformation we applied after acquisition which substantially changed the degree of overlap in some cases. Remaining errors are also due to the variation in the position/orientation of the transducer between volumes which results in differing image intensities for the same structure. Despite these difficulties, high quality mosaicing results can be easily produced using clinical ultrasound data from live patients, and our experiments using fetal data show this.

In order to understand the size of the problem we tracked the number of nodes pair-wise energy terms needed for each level of the Gaussian pyramid. This would be advantageous in the future if we wish to implement more advanced optimization methods that are designed for large-scale non-submodular MRFs containing triple cliques. At every pyramid level in the FAST/ER FAN experiment the labeled nodes defined 5 non-rigid transformations, 1 corresponding to each source. The lowest resolution required approximately 9,000 pair-wise energy terms in order to evaluate the group-wise similarity metric, the second lowest required 69,000, and the highest resolution need 281,000. We should note that pre-computation using the FFT method at the highest resolution required $\left(17^3 \text{ search window size}\right)(281000 \text{ nodes})(2 \text{ bytes per uint16}) = 2.76\,\text{GB}$ to hold the SSD data from (3.31). We also found that optimization at the highest resolution didn't offer substantial improvement over registration at $\frac{1}{4}$ resolution, as the relatively sizable structures were pretty well aligned by this point. However, ultrasound images containing finer features would certainly benefit from full resolution registration.

The FAST/ER FAN data was also used to quantify the performance increase that we achieved by pre-computing the pair-wise similarity terms in (3.26) using the FFT approach. In order to strictly measure the improvement in optimization speed the interpolation time wasn't included in the timing results. Both the FFT and the naive approach were implemented as Matlab functions and care was taken to vectorize the naive approach, utilizing Matlab's strengths where possible. This ensured that our comparison wasn't unfairly biased. We measured the amount of time it took to complete an iteration of alpha expansion, consisting of only one sweep over the label set, at the middle resolution of our pyramid scheme. As stated before, registration at this resolution required 69000 pair-wise energy terms and produced high-quality results with the FAST/ER FAN data thus we believe this to be an ideal setting to evaluate the performance of both approaches. Alpha expansion required 1556 seconds using the FFT based approach and 16040 seconds using the naive approach giving a speedup factor of 10.31. Most alpha expansion schemes use multiple passes over the label set and since the 1556 seconds already includes the preprocessing time the speedup factor between the FFT and naive approaches would only increase as the number of passes increased. We conclude that the FFT approach is the most efficient method to compute the group-wise similarity measure in (3.31) using an MRF optimization scheme.

These results suggest that intra-volume misalignment of many (≥ 2) partially overlapping ultrasound volumes can be effectively corrected using an intuitive group-wise MRF approach. The most important factor to consider when constructing ultrasound mosaics is how well major anatomical structures align when spanning more than one volume. Based on the results presented our algorithm does a good job when dealing with many partially overlapping volumes.

Future work could compare the accuracy of our improved group-wise SSD metric shown in (3.26), which is robustness in the presence of ultrasound shadowing artifacts, to the other modern similarity metrics mentioned in the introduction using an abdominal phantom such as the FAST/ER FAN. One possible way to conduct this investigation would be to initially create a gold-standard from the original US volumes through careful acquisition and registration. This wouldn't be difficult due to the rigid nature of the FAST/ER FAN. Major structures would be segmented in the original image volumes, prior to deformation with thin plate splines. Due to the high quality US images we saw using the FAST/ER FAN this step should also be fairly straightforward and can be accomplished using ITK Snap or some other segmentation tool. The DICE coefficient, which measures the agreement between sets, can be used to quantify the alignment between the segmented structures in each undeformed volume. Anatomy obscured by shadows shouldn't be counted in the calculation of the DICE coefficient. A coefficient of 1 indicates perfect set agreement thus before deformation this coefficient should be close to 1 when comparing all segmented structures in the overlapping regions. After artificial deformation, various configurations of the stitching algorithm would be applied to bring these volumes into alignment. Because all transformations are known we have the ability to map the initial segmentation results to the new coordinate system produced by the group-wise registration algorithm. Finally, to quantify the algorithms accuracy the DICE coefficient needs to be calculated again between the segmented organs in the regions of overlap. This result would tell us how well the anatomical contours in each individual volume are aligned and also give us some idea of our proximity to an ideal solution. Since ultrasound is inherently noisy and lacks fine detail we believe that tracking and comparing organ boundaries using segmentation and the DICE coefficient is a good way to judge the registration accuracy of a group-wise method. The DICE coefficient has also been used in segmentation validation [88] as well as in papers analyzing registration performance [35]. The next section presents clinical results where multiple overlapping fetal image volumes were mosaiced.

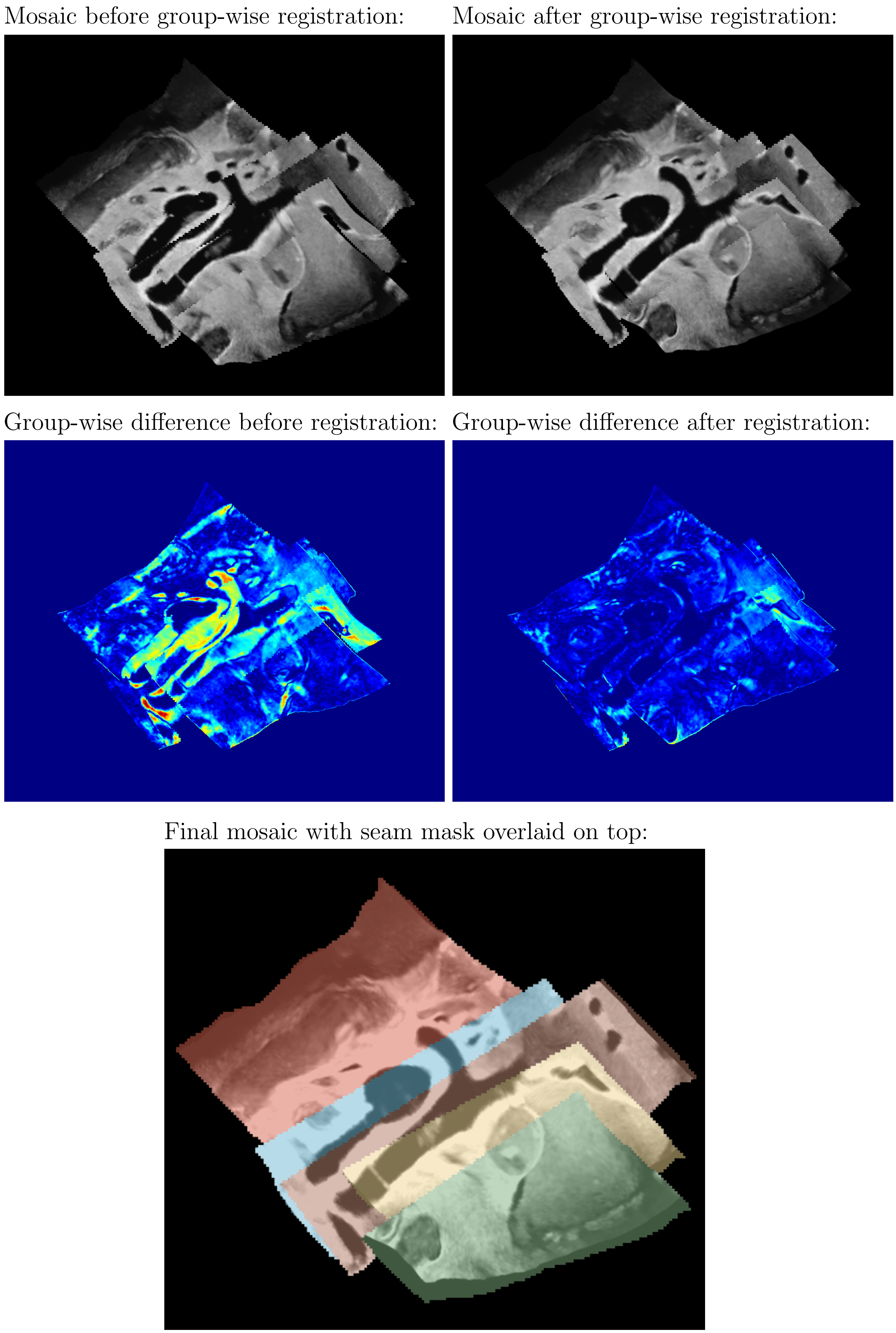


Figure 4-6: Result of registration showing improvement in coronal slice of mosaiced volume

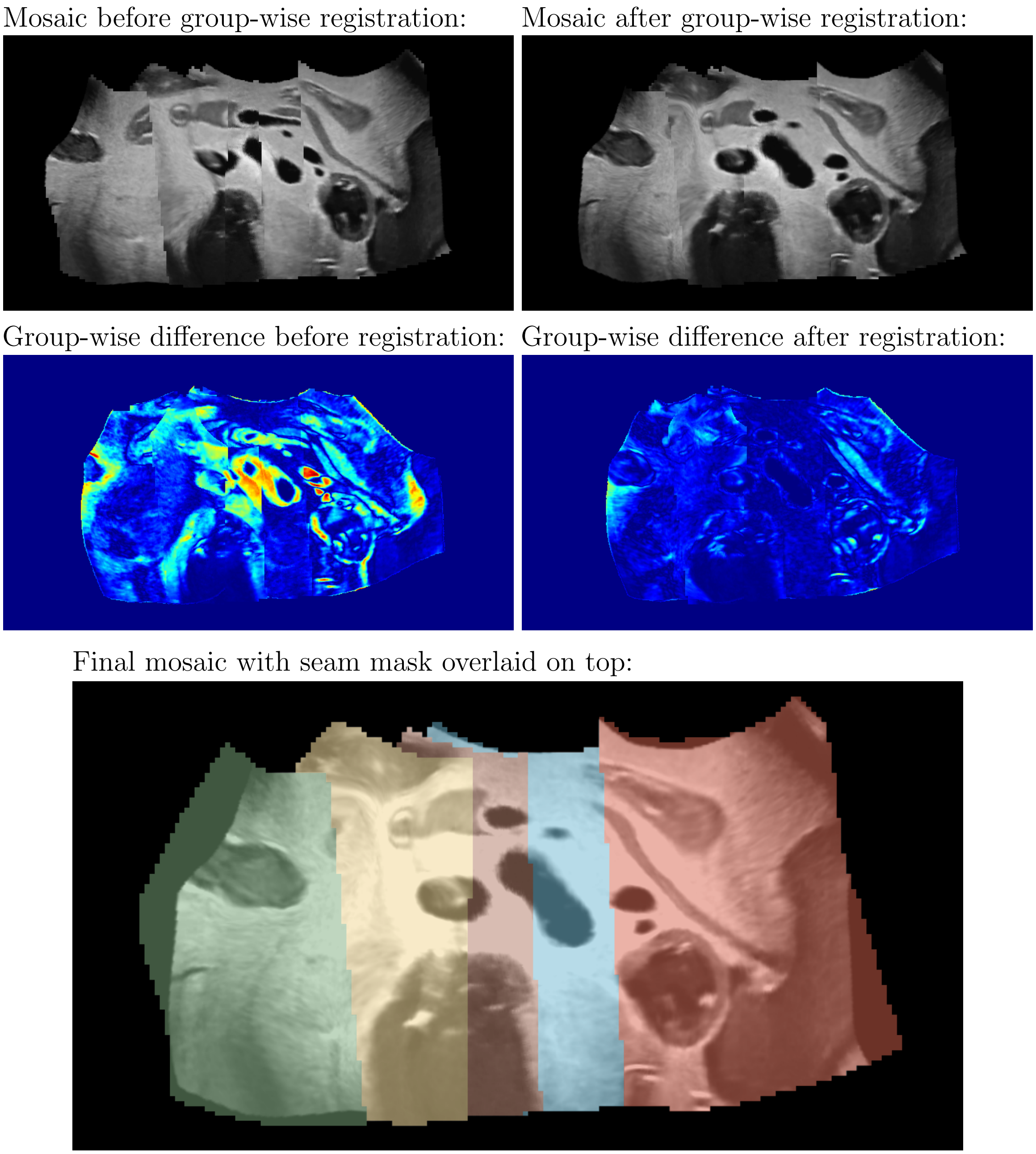


Figure 4-7: Result of registration showing improvement in slice orthogonal to planes displayed in Figures 4-6 and 4-8

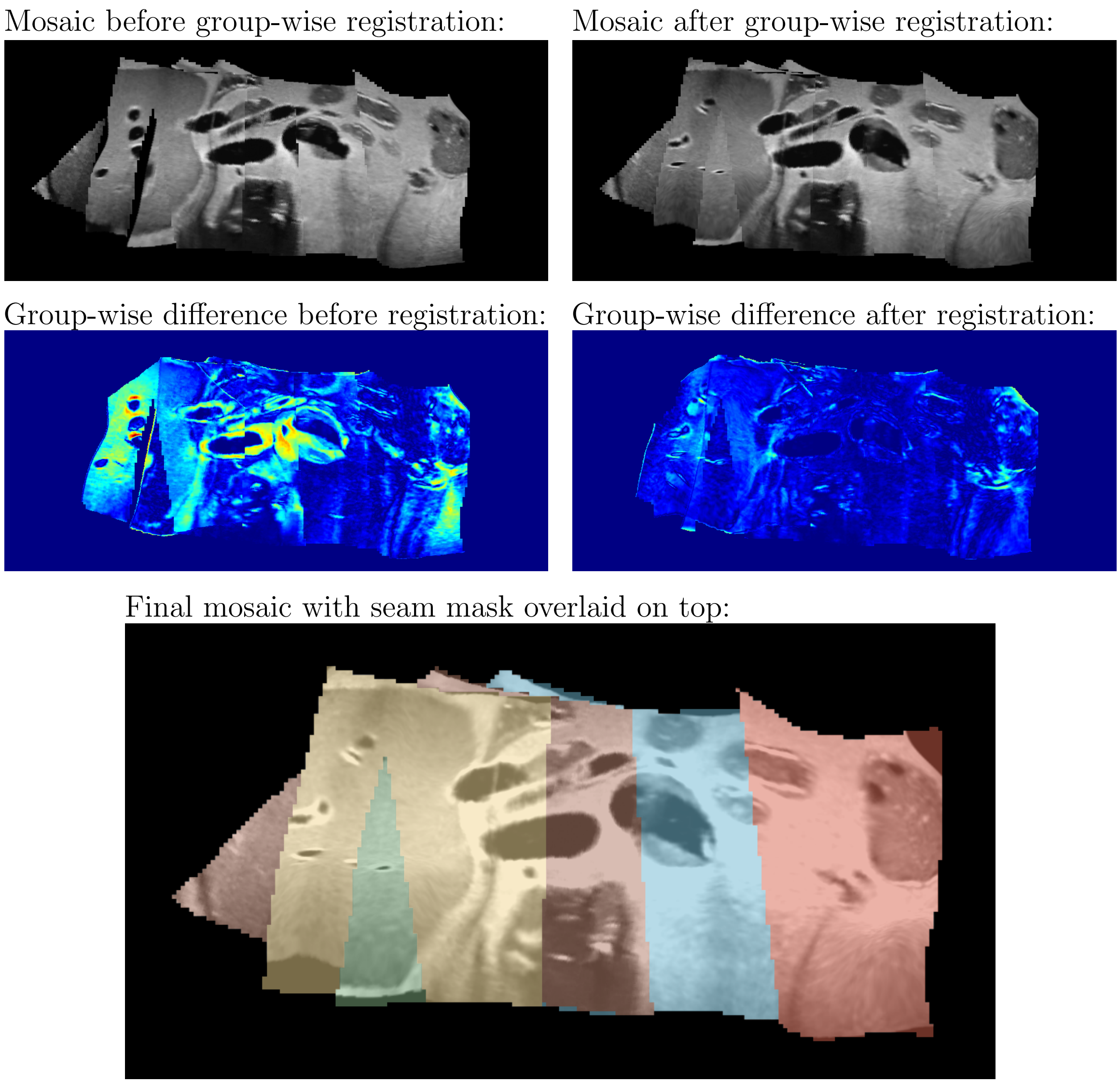
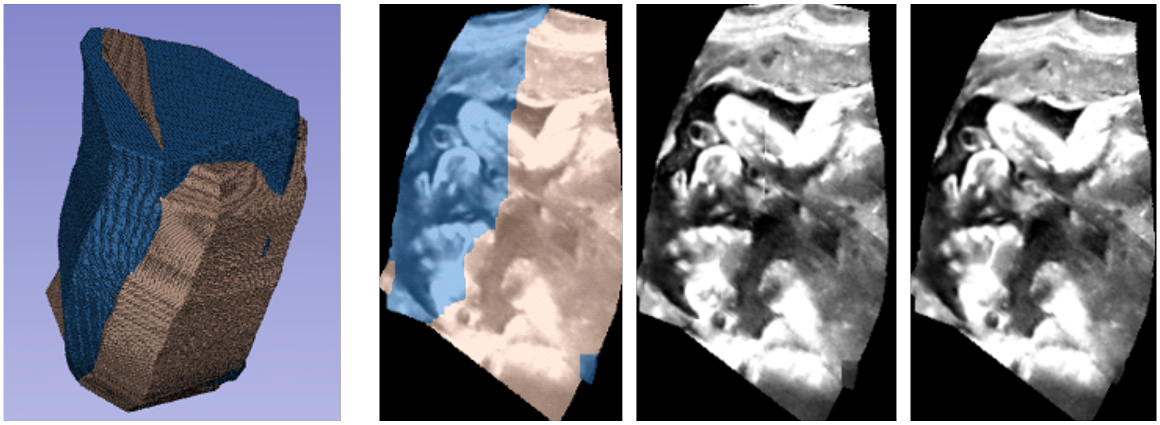


Figure 4-8: Result of registration showing improvement in slice orthogonal to planes displayed in Figures 4-6 and 4-7

4.3 Results from clinical ultrasound

In order to validate our approaches we used clinical ultrasound data obtained from obstetrics patients at the University of Massachusetts Medical School. With the same set-up that was used to scan the FAST/ER phantom we collected individual volumes from the patients along multiple overlapping linear scan paths. The best results were achieved when stitching together overlapping volumes produced from 2D ultrasound images acquired in the mother's sagittal plane. The first step after 3D image acquisition is to rigidly register the multiple volumes together in order to remove the global misalignments caused by shifts of anatomical structures during the scanning; such shifts are due to breathing, variation in muscle tone, fetal movements and non-uniform probe pressure. The rigid registration can be accomplished using the highly optimized modular algorithms found in the Insight Toolkit [32]. We found that the Insight Toolkit worked well enough for our initial rough alignment but it should be mentioned that in [83] group-wise rigid registration of overlapping ultrasound volumes is discussed in detail and a more robust algorithm is developed. Since we require only a rough rigid alignment for our group-wise non-rigid registration algorithm to work well the improvements in [83] were not necessary. Once this step is complete the volumes are stitched together using the approach described above.

Initial clinical experiments verified that the registration/mosaicing algorithm worked correctly with two overlapping ultrasound volumes. In Figure 4-9 a volumetric rendering of the mosaicing function, which was calculated for two overlapping volumes using our graph based algorithm, is shown in image (a). Every voxel colored blue receives its value from volume 1 and every voxel colored gray receives its value from volume 2. Also displayed in Figure 4-9 are two slices from the composite volume, one before non-rigid registration denoted as (c) and one after non-rigid registration denoted as (d). The misalignment along the limbs of the fetus is obvious. We also observe that the mosaicing function chooses to pass through the arm of the fetus suggesting that the misalignment in the rest of the overlapping region is worse than this area. In fact the reason that the seam selection algorithm chooses this area is



(a) (b) (c) (d)

Figure 4-9: Pairwise mosaicing function with non-rigid registration. (a) A volumetric rendering of the mosaicing function. (b) A slice of the composite volume with the mosaicing function overlaid. (c) Same slice before non-rigid registration. There is a definite discontinuity in the fetal arm. (d) Same slice from final composite volume.

due to extensive shadowing artifacts which are often encountered during fetal ultrasound. In this experiment volume 1 shows the orientation and position of the limbs of the fetus in detail while volume 2 only contains a limited view of this region due to shadowing. Thus the mosaicing function chooses the region where the overlapping volumes contain similar structures and by focusing the registration algorithm in this area better results are obtained. Image (d) demonstrates the improvement after non-rigid registration and the effectiveness of focusing on common structures during alignment. Using the entire overlapping region for alignment of fetal volumes gives unusable results in this experiment due to the disparity between the structures viewable in volume 1 versus the structures viewable in volume 2. Also the fact that we choose an optimal mosaicing function reduces the amount of registration necessary to bring the adjacent volumes into alignment in the vicinity of the seam. The inability to easily and effectively register multiple overlapping volumes of clinical obstetrics data makes ultrasound compounding methods based on various weighting schemes a bad choice for 3D ultrasound mosaicing in our experiments [34]. The deformation field in the rest of the volume relies entirely on the 2nd order regularization of the transformation field.

Finally we present results from the complete abdominal reconstruction of three obstetrics patients that were 26-30 weeks along in their pregnancy. During clinical scanning at UMass Medical School 10+ individual image volumes were acquired from each subject encompassing the entire fetus and placenta. We collected such a large number of scans because swept 2D ultrasound was used to generate the volumes and is very sensitive to fetal movement. During construction of the composite volume some of the individual volumes were found to be unusable making redundant scanning of the subject necessary. Any sudden movements of the baby would ruin an active scan. We found that it only requires around 5-6 carefully placed volumes to reconstruct the anatomy of interest in our case. Figures 4-10, 4-11, and 4-12 show slices of the composite volume created from each of these subjects. The overall image quality is very good and the slices show anatomical details which wouldn't be visible in a single scan. For example in the top of Figure 4-10 amniotic fluid is visible on both sides of the fetus which typically doesn't occur during clinical scanning due to the shadowing effects of obstetric ultrasound. Also visible are details of the baby's limbs and spine which are hard to capture simultaneously. The mosaicing function was chosen to construct the composite volume from the registered source volumes as opposed to a voxel-wise weighting scheme because the overlapping source volumes were much too different in regions far away from the optimized seams. Again, this is mostly due to fetal movement and shadowing. Correcting the misalignment between volumes caused by the fetus shifting its limbs is more difficult than correcting for predictable movement such as elastic deformation or a heartbeat which is why the mosaicing function is used. We do not require the limbs in all source volumes to be aligned, only the limbs in the sources specifically used to construct this region of the composite volume. The right side of each figure demonstrates how the mosaicing function partitions the composite volume into regions corresponding to each of the overlapping sources. By looking at the image continuity in Figures 4-10, 4-11, and 4-12 we see that stitching algorithm is effective at producing visually satisfying results from multiple partially overlapping ultrasound volumes.

Our primary purpose for developing the algorithms explained in this chapter is to produce anatomically correct abdominal ultrasound volumes for use in obstetrics ultrasound simulators. Thus in order to evaluate the training value and correctness of each composite volume we loaded them into our laptop based US simulator [57] and encouraged the sonographers at University of Massachusetts Medical School to compare the simulator experience to that of scanning a live patient. Our simulator allows the user to re-slice the composite volume in any orientation/position by utilizing a sham transducer with a 5 degree of freedom tracking system. From an initial evaluation the sonographers were impressed with the quality and realism that our composite volumes added to the training experience. The main issue was a slight blurring in some re-slices that are used to calculate certain fetal measurements such as abdominal circumference. Scanning procedures which focus on areas of the fetus important to clinicians and careful selection ensuring we use the best image data to construct the composite volume should improve this aspect.

4.4 Conclusions

The novel algorithm we presented in this chapter essentially requires two MRF optimizations to mosaic N partially overlapping ultrasound volumes. The first optimization determines how the volumes should be stitched together by attempting to minimize the intensity/gradient differences between adjacent volumes. This could be used as preprocessing step to more sophisticated compounding techniques if desired. The second optimization determines N deformation fields which bring the overlapping volumes into alignment in regions of overlap. Precomputation of the similarity metric using FFTs and a focused registration energy resulted in a very efficient mosaicing algorithm. Performance was further enhanced by splitting the solution space into distinct regions, exploring them in parallel, then fusing the results.

The framework developed here is modular, which allows improvements in optimization or similarity measure to be easily integrated into the algorithm. As more efficient higher order MRF optimization methods are developed their speed/accuracy should be tested on the registration energy in (3.44). Improved MRF optimization could enable the use of more sophisticated registration models, such as those including

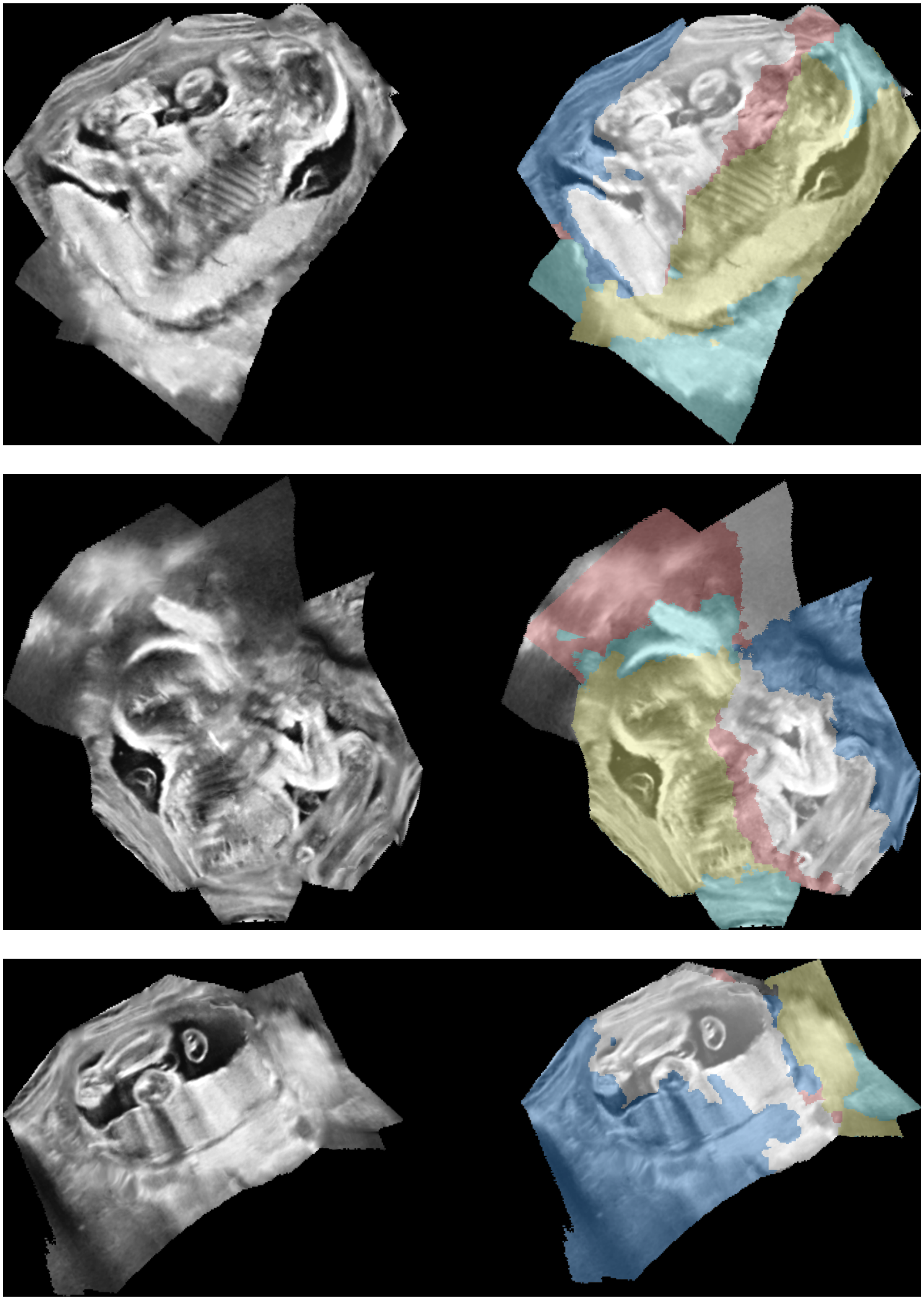


Figure 4-10: Re-slices of composite volume created from subject 1. Rightmost images show how final volume is partitioned into regions corresponding to overlapping source volumes.

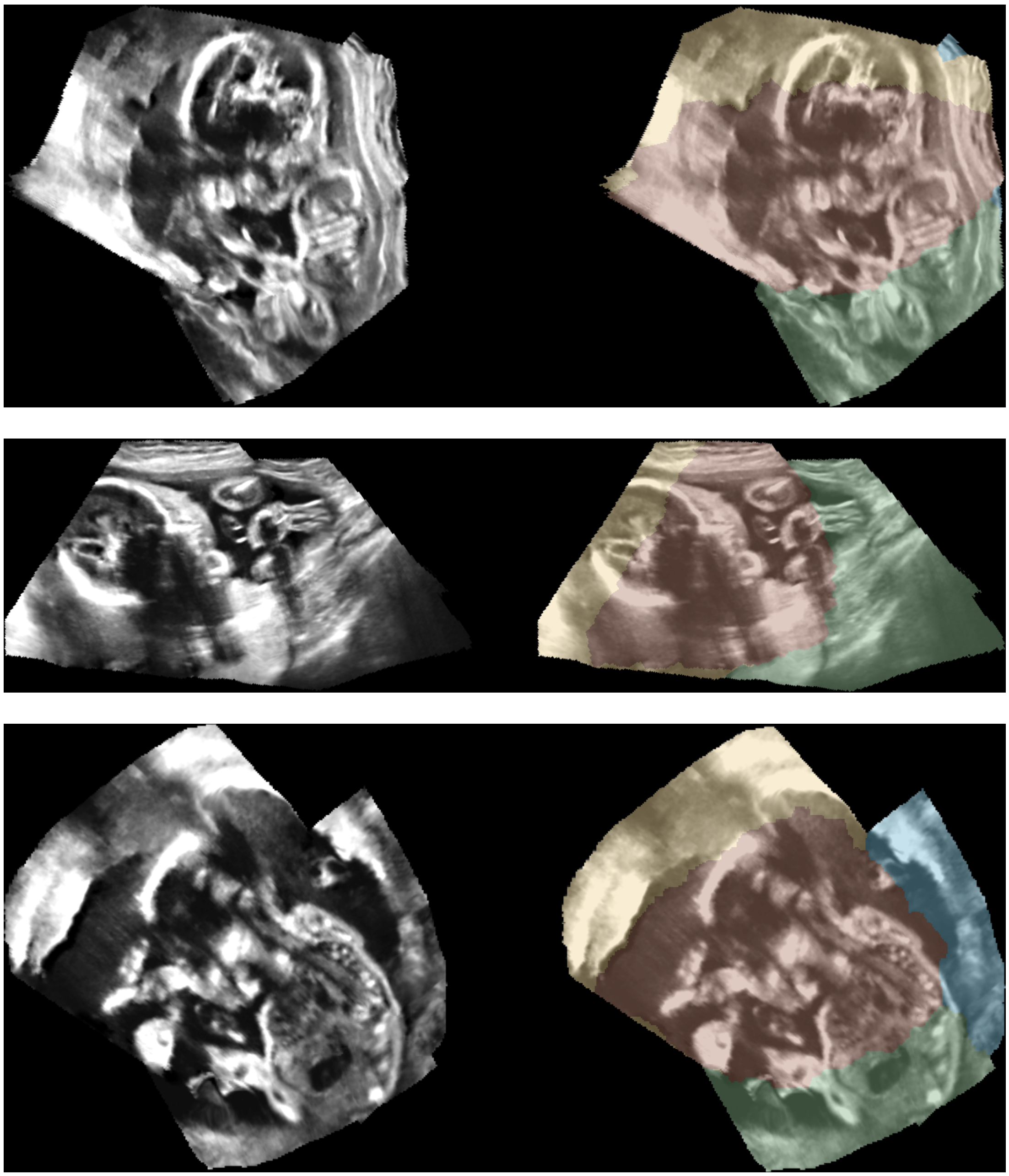


Figure 4-11: Re-slices of composite volume created from subject 2. Rightmost images show how final volume is partitioned into regions corresponding to overlapping source volumes.

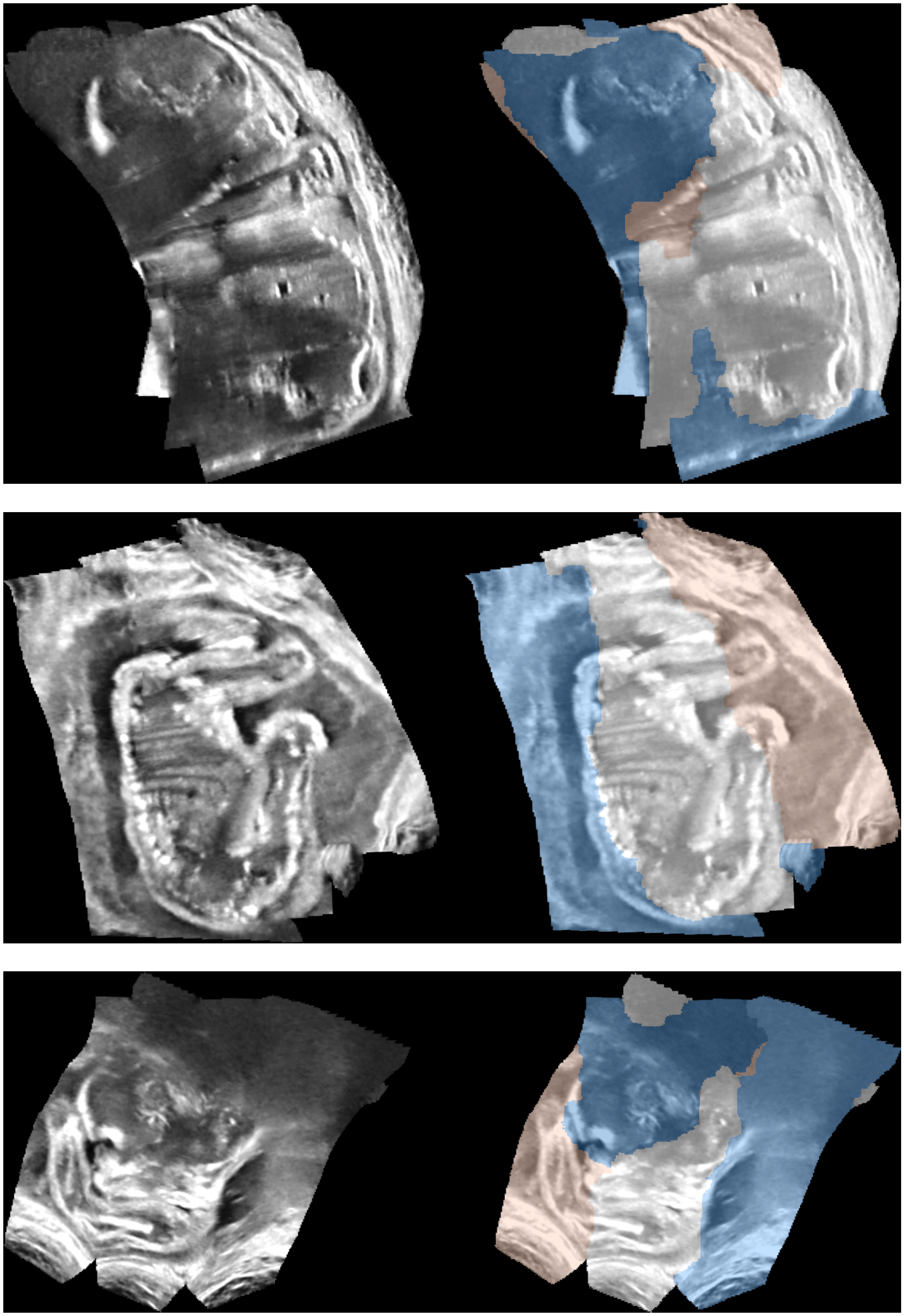


Figure 4-12: Re-slices of composite volume created from subject 3. Rightmost images show how final volume is partitioned into regions corresponding to overlapping source volumes.

both rigid and elastic components. Joint registration/segmentation would also be interesting to explore in this framework. Finally work towards efficient pre-computation of more sophisticated group-wise similarity metrics should be completed.

The qualitative evaluation by sonographers at UMass complements the quantitative results we found using a finite element model built from the Visible Human Project data. It demonstrates the ability of the mosaicing technique developed in this chapter, to overcome large deformations with only a small degree of overlap.

With the techniques for producing composite image volumes in place, additional simulator development work includes the addition of landmarks that will be used in educational modules providing training on the acquisition of common fetal measurements, which is a task clinicians must be proficient in. Also an evaluation of the simulators training benefit will be conducted by incorporating it into the curriculum of a small number of medical students.

Chapter 5

Optimizing spline surfaces using Particle Swarm Optimization

In this chapter we introduce a technique for the global optimization of spline surfaces and apply it to the task of finding the optimal stitching seam between two neighboring and overlapping 3D image volumes. Finding the optimal location of the control points that define the surface is a large scale constrained optimization problem and it is solved using a cooperatively coevolving particle swarm based approach. This was an early approach to our problem, which was abandoned when we realized the performance benefits of the graph based techniques described in Chapter 3. It is included in this dissertation for completeness and offers an alternative approach for optimizing spline surfaces.

5.1 Introduction

As discussed in Chapter 1, we believe that the choice of an optimal seam is an important step towards producing artifact-free joined image volumes in hard to stitch cases. This is analogous to 2D image panorama creation, where a good seam is one which divides the composite image, created from the source images, into regions such that few discontinuities occur along the boundary. This surface should avoid moving objects and regions of high information content because of the need to maintain optimal image continuity across the overlapping 3D image regions. We also do not want to non rigidly deform rigid objects like fetal arms and legs in obstetrics image volumes. 2D seam selection has been described extensively [48] and we have extended this concept to 3D. In this chapter, we have formulated the task of generating a 3D surface representing the ideal seam between volumes as a shape optimization problem.

5.2 Parametric seams with B-splines

In order to develop any shape optimization algorithm a suitable mathematical representation must be chosen. A number of models for deformable surfaces have been developed, each with unique properties. These models may be discrete or continuous, allow for or preclude topology changes, and differ in their evolution laws. We choose to implement the seam as a cubic B-spline surface because of their inherent smoothness constraints, i.e. first and second derivatives are continuous across control points. This property is very important for the implementation of an efficient global optimization scheme. There is no need to accommodate topology changes in our seam optimization application so other representations such as implicit surfaces were excluded. Discrete meshes, particle systems, and Fourier modes were also considered and rejected. A review of deformable models was presented in [49] which contains details on properties and methods of evolution.

B-splines are well known and a recursive formulation of the basis functions are used in our implementation. A surface can be constructed using the tensor product of these functions coupled with a control point grid. A properly formed knot vector guarantees that the surface terminates on the exterior control points and this property is used to ensure that the seam divides the overlapping volume into 2 sections. This is accomplished by constraining the exterior control points to reside on the boundary of the overlapping volumes.

For completeness the recursive definition will be given. Let $U = \{u_0, \cdots, u_{m+2k}\}$ be a knot vector where m is a positive integer, $k > 0$ and $u_0 \leq \cdots \leq u_i \leq \cdots \leq u_{m+2k}$. The B-spline basis functions of degree k are constructed using the follow recursive formula,

$$
\begin{aligned}
N_{i,0}(u) &= \begin{cases} 1, & \text{for } u \in [u_i, u_{i+1}) \\ 0, & \text{otherwise} \end{cases} \quad i = 0, 1, \cdots, m + 2k - 1 \\
N_{i,k}(u) &= \tfrac{u - u_i}{u_{i+k} - u_i} N_{i,k-1}(u) + \tfrac{u_{i+k+1} - u}{u_{i+k+1} - u_{i+1}} N_{i+1,k-1}(u),\ i = 0, 1, \cdots, m + k - 1
\end{aligned}
\tag{5.1}
$$

where we assume that $u_0 = u_1 = \cdots = u_k = 0, u_k < \cdots < u_{m+k}, u_{m+k} = \cdots = u_{m+2k} = 1$ and $\frac{0}{0} = 0$. The tensor product surface $\mathbf{x} \in \mathbb{R}^3$ is generated using the following equation,

$$
\mathbf{x}(u, v) = \sum_{i_1=0}^{m_1+k-1} \sum_{i_2=0}^{m_2+k-1} \mathbf{c}_{i_1,i_2} N_{i_1,k}(u) N_{i_2,k}(v),\ u, v \in [0, 1] \tag{5.2}
$$

where $\mathbf{c}_{i_1,i_2} \in \mathbb{R}^3$ is an individual control point in 3D space taken from an array indexed by the variables i_1 and i_2. In our application we take $k = 3$ which generates cubic basis functions. Also m_1 and m_2 must be positive integers. The B-spline basis function is represented by $N_{i,k}$.

The exterior control points of the grid define where the surface terminates, while the interior control points define the overall shape and the boundary between volumes. Figure 5-1 shows an example of how a tensor-product B-spline surface can be used to represent the seam. It also shows a region in the 3D volume corresponding to zero misalignment which is represented by a tetrahedral mesh. Using our shape optimization framework we would like to calculate the control point displacements which cause the B-spline seam to flow through this region. This process will be discussed in detail in the following section but this image serves to illustrate the b-spline model of the seam. The left image in figure shows the ideal seam region and the initial plane which bisects the volume. This volume was synthetically generated to test the algorithm on easy seams before advancing to anatomical data. The right image shows the surface after optimization. The algorithm was able to find the correct seam and by constraining the motion of the exterior control points the final result still bisects the volume. The B-spline model is well suited for this application.

Efficient implementations exist to compute points on a B-spline surface along with

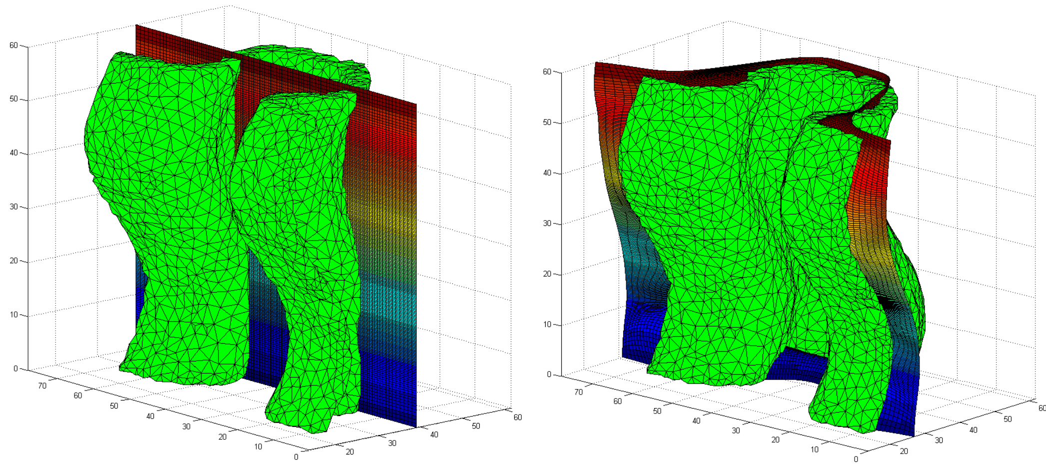


Figure 5-1: The left image shows the initial planar bisecting surface. The right image shows the surface converging to the correct region.

derivatives at their locations. This is important in global minimization because the fitness function for the surface will need to be evaluated numerous times during the optimization process. For example, digital filtering techniques are available for the processing and representation of signals (surfaces) in terms of continuous B-spline basis functions [80, 81]. Especially useful are the digital filters designed to evaluate derivatives along the surface. All B-spline operations can be implemented as convolutions in this framework, which may be employed in the future to speed up our algorithm.

The surface must be initialized before we can optimize it. A number of 3D volumes in different orientations overlap and the overlapping region between each pair of volumes is not necessarily aligned with the common coordinate axes. This requires care in order to initialize a planar surface, which bisects the overlapping region of two volumes that differ in orientation/position. Principal Component Analysis is used to generate a coordinate system aligned with the major axis of the overlapping region. A bisecting surface can be formed using the normal representation of a plane which is then transformed into its B-spline counterpart.

5.3 Fitness function for stitching seam surface

The fitness function evaluates the quality of a seam, which is parameterized by its control point locations. The total fitness is a weighted sum of an error term and a regularization term and we wish to minimize this sum. This is a common approach in medical image processing. We choose to use RMS error along the surface coupled with a regularization term based on elastic energy. The fitness function is shown in equation (5.3).

$$Fitness = \sqrt{\frac{1}{SurfArea}\iint V(\mathbf{x}(u,v))^2 \left|\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v}\right| dudv} \\ \lambda \iint \left[\kappa_1^2 + \kappa_2^2\right] \left|\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v}\right| dudv \quad (5.3)$$

Two principal curvatures are needed in order to calculate the regularization term. The letter $\mathbf{x}$ represents the surface parameterized by u and v. V represents the misalignment in the overlapping region we wish to divide. Both of these terms are surface integrals calculated along on the B-spline seam. The RMS error was chosen because it penalizes larger errors more severely. We can correct for small continuity violations between the volumes along the stitching surface but correcting for large concentrated errors are more difficult.

The second term in the fitness equation is arguably more important than the first because it is what makes global optimization of a B-spline surface possible. It is weighted by λ so its influence on the optimization can be adjusted. Low values of λ allow the surface to take on more complicated shapes while higher values will result in a smoother seam. Complete freedom of movement was needed to allow the seam to escape local minima during optimization. As will be discussed in the next section, most gradient based methods are inadequate because large steps are necessary. This results in a significant risk of self-intersection because each of the control points can freely move in any direction during optimization. When the B-spline surface described above self-intersects its elastic energy grows very large due to the flexing of the surface necessary to cause the intersection. Explicitly checking the surface for self-intersections is costly for an optimization method dependent on fitness function evaluations so elastic energy was the mechanism chosen to prevent this problem [20]. Before elastic energy was finally implemented we attempted to utilize curvilinear error to control the smoothness and prevent self intersection however this term didn't provide the regularization necessary. Curvilinear error has been used for B-spline regularization and also to control B-spline snake smoothness and is described in detail in [33, 84]. It is an error metric that increases when the control points aren't evenly spaced and thus penalizes shapes with regions of high curvature. The advantage of curvilinear error is that it only requires the computation of 1st derivatives however it failed to prevent self intersection during optimization. More research would be necessary to correctly implement this error metric and to determine if the computational savings is worth it. Curvilinear error for a curve is shown in (5.4) and can easily be extended to a surface.

$$Error = \int_0^1 \left[\left\| \frac{dx}{du} \right\| - CurveLength \right]^2 du \tag{5.4}$$

5.4 Optimization algorithm for seam surface selection

In order optimize the 3D stitching seam, in terms of minimizing RMS error, an algorithm capable of handing large scale non-separable problems was needed. The algorithm also had to be able to fully explore the solution space in order to avoid getting stuck in local minimums which is one of the disadvantages of using gradient based methods. This type of global optimization is well suited for evolutionary algorithms such as genetic and particle swarm based approaches. If we consider a B-spline surface with a 10x10 grid of control points located in a 3D coordinate system then there are now 300 variables to optimize, which poses large difficulties for traditional evolutionary algorithms. Since we also desire a surface with no self intersections this problem becomes a large scale constrained optimization problem. Many large scale optimization problems can be handled by grouping dependent variables together and solving several sub problems. This works well when the dependency is known and the problem is separable; however in the seam selection problem these characteristics are not present. The risk of self intersection adds dependencies between variables, which change based on the shape of the surface and are difficult to describe mathematically. The simplest method is to assume that the amount of variable dependence is a function of control point proximity; however every control point is somewhat dependent on all the others no matter what the distance between them. After investigating various methods such as differential evolution and modern genetic algorithms we implemented a particle swarms based algorithm described in [44] and developed a random grouping algorithm to partition the control point displacement variables into sub-swarms based on location.

5.4.1 Cooperatively Coevolving Particle Swarms for Large Scale Optimization

Particle swarm optimization (PSO) is modeled after the behavior of social animals in large groups. In nature it has been observed that these groups display a collective intelligence when problem solving. In PSO each particle maintains its current position as well as its personal best position in the solution space where the personal best position for the particles represents the highest fitness value it has found so far while traversing the solution space. In this application a position in the search space corresponds to a control point configuration which generates a unique surface. During each iteration of basic PSO, every particle updates its current position according to an update rule that takes into account the individual particle's best position found so far and also the best positions found by its neighboring particles. Early versions of PSO used a global best position in the update rule for each particle which meant that the particles all moved towards the same solution during each iteration, sometimes resulting in premature convergence. In this case all the particles in a swarm were considered neighbors, and communication between particles, i.e. exchanging personal best positions, is modeled as a fully connected graph. Here the global best position was the solution with the highest fitness value found by all the particles. In the optimization algorithm a Cauchy and Gaussian based update rule is used. Equation (5.5) shows the update rule.

$$x_{i,d}(t+1) = \begin{cases} y_{i,d}(t) + C(1)\left|y_{i,d}(t) - \hat{y}_{i,d}(t)\right|, & \text{if } rand \le p \\ \hat{y}_{i,d}(t) + N(0,1)\left|y_{i,d}(t) - \hat{y}_{i,d}(t)\right|, & \text{otherwise} \end{cases} \tag{5.5}$$

In eq, $x_{i,d}(t)$ represents particle i's position at time t. The subscript d indexes an individual variable for multi variable optimization problems. $C(1)$ is a Cauchy random variable and $N(0,1)$ is a normal random variable. $y_{i,d}(t)$ represents particle i's personal best position at time t and $\hat{y}_{i,d}(t)$ represents particle i's neighborhood best position. The variable p controls the likelihood of sampling around the particle's personal best position or its neighborhood best position. In our experiments p was set to .5 which used in [44]. The Gaussian random variable is used to sample around the particle's neighborhood best position and the Cauchy is used to sample around the particle's personal best position. An update equation which only uses a Gaussian distribution has a limited ability to search the solution space especially when its standard deviation becomes small. A Cauchy random variable's probability density function has larger spread, which means it is more likely to sample farther away from the particle's personal best thus being more exploratory. With p set to .5 half of the time we are exploring around a particle's personal best using a Cauchy r.v. and the other half we are converging on a plausible solution using a Gaussian random variable.

In PSO the neighborhood topology heavily influences the performance of the algorithm. A ring topology was used because it allows the greatest exploration of the global search space and prevents premature convergence. It accomplishes this by slowing down the spread of the swarms best fitness value, which may correspond to a local minima, to the distant particles where distance between particles is a function of the topology. One disadvantage of the ring topology is the slow convergence rate but

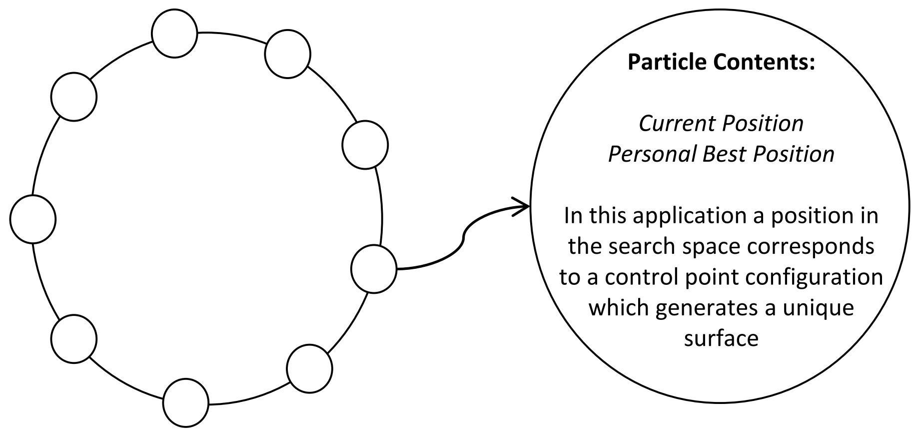


Figure 5-2: The left image shows a ring topology and the right image shows the contents of an individual particle.

for non separable problems increased exploration is important. Once topology has been defined population size must also be considered. Low connected topologies and a high number of particles results in slow propagation of information, and therefore a more parallel search is performed [19]. We choose a population of 64 which was shown to have good results. This number allowed us to implement a SIMD (single instruction multiple data) fitness function effectively utilizing all 32 CPU cores we had at our disposal.

Early versions of PSO only utilized one swarm. However this did not scale well as the number of variables to be optimized grew. For large scale optimization, the strategy of divide and conquer can be applied, where sub-swarms are used to optimized subcomponents of the problem and the final solution is constructed by concatenating the solutions of each sub-swarm. This approach is known as cooperatively coevolving optimization and has been applied to other evolutionary algorithms as well as particle swarm optimization [85]. This approach worked with separable problems but performed poorly on non separable problems which led to the development of random grouping. The reason for this was that the algorithms failed to capture the variable interdependencies and new strategies needed to be developed. This lead to the introduction of decomposition through random grouping [54]. The idea behind random grouping was to increase the chances that two dependent variables would be optimized together by a sub-swarm. Ideally all closely interacting variables would be grouped together in a subcomponent and the dependency between subcomponents would be minimized. Each subcomponent would then be input into a cooperatively coevolving optimization algorithm. Early work used genetic algorithms as the subcomponent optimizer but PSO was introduced later.

The algorithm cooperatively coevolving particle swarm optimization two (CCPSO2) [44] has been modified from its original version for the optimization of the B-spline surface described above. The modifications have been introduced to handle the non self-intersecting constraint and also to take into account the dependence that neighboring control points have on each other. The flow chart describing our algorithm is shown in Figure 5-3.

The concept of the algorithm is to split the seam into a number of random patches called subcomponents, each to be optimized individually by a sub-swarm. The control points which reside in the same patch are grouped together into a subcomponent which will be optimized by a sub-swarm. These patches are optimized in a round robin fashion as shown in Figure 5-3. In order to evaluate a particles fitness in each sub swarm a complete solution (seam surface) must be constructed. A context vector, $\hat{y}$, is used to hold the best control point locations found during optimization so far by all sub-swarms. To evaluate the fitness of each particle the variables in $\hat{y}$ corresponding to the current sub swarm are replaced by the particles values. For example we may be optimizing a surface with 100 control points and decide to split the surface so each patch or subcomponent contains 25 points, which then results in 4 sub-swarms optimizing 25 points each. In order to evaluate the fitness of a particle in each sub swarm we must construct the entire surface thus all 100 control point locations are required. In order to evaluate the i'th particle in the j'th sub-swarm the function $\mathbf{b}(j, P_j.x_i)$ is used which returns the n dimensional vector consisting of $\hat{y}$ with its jth component replaced by $P_j.x_i$. This is the complete solution needed for the fitness function which measures how well particle $P_j.x_i$ cooperates with the best solutions found so far in the other sub swarms. The for loops iterate through each sub swarm resulting in coevolution of the particles. Progress made in the search

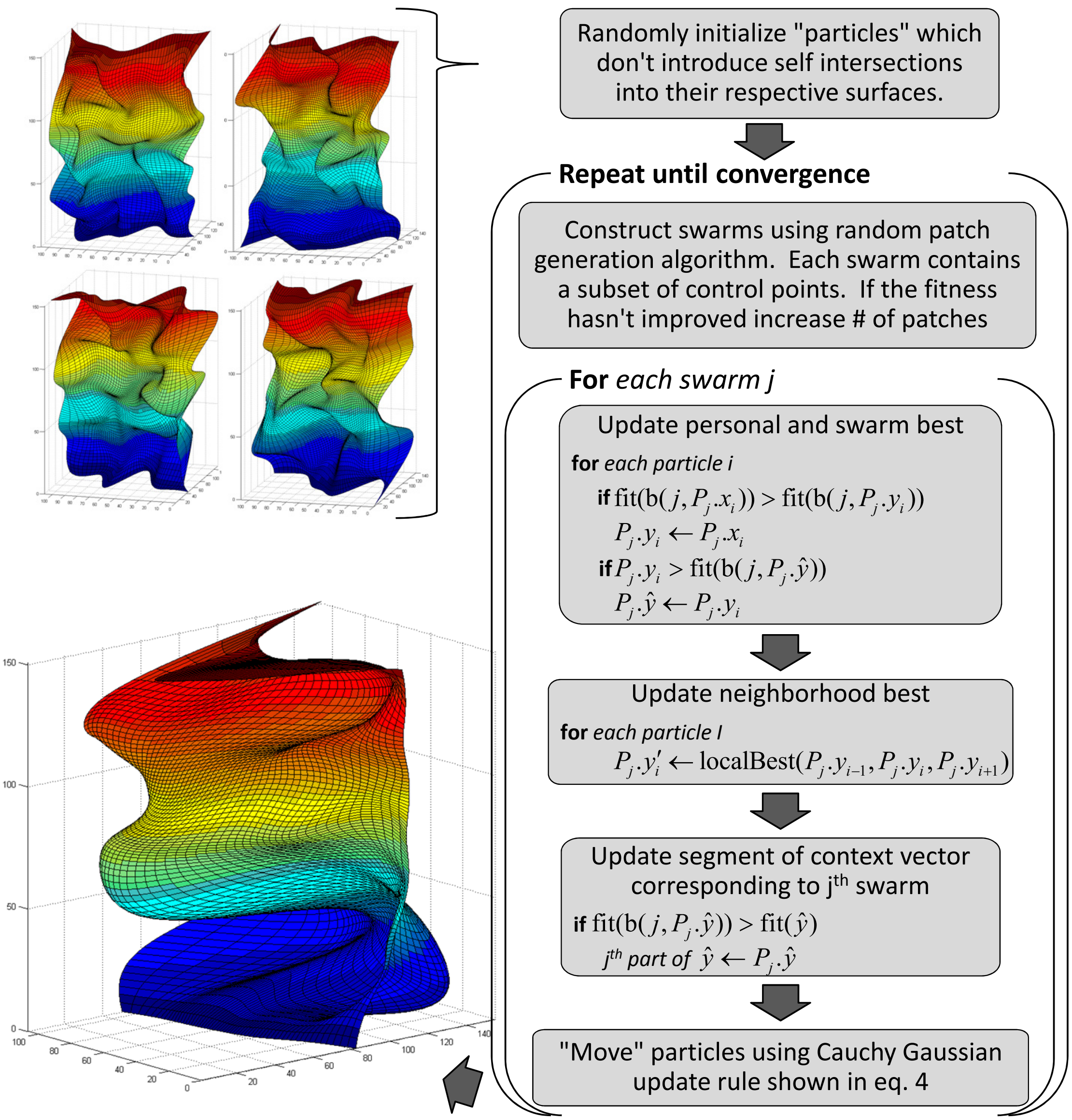


Figure 5-3: This flowchart illustrates the optimization process, beginning with the initial population of random surfaces or particles and concluding with the generation of a seam minimizing the RMS error calculated as the surface integral.

space during each iteration is used in subsequent iterations. In the first nested for loop we are updating the particle's personal best position, if their new position in the search space is an improvement. Also we are checking the jth swarm best for update, finding each particle's local best position using a ring topology, and finally updating the context vector $\hat{y}$ if the swarm best results in the highest fitness found so far. In the final nested for loop each particles new position is calculated using its personal best and neighborhood best values. As described in [44] matrix representations seem natural to hold the personal best and current position of each particle, as formulated in Equation representations.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{64,1} & x_{64,2} & \cdots & x_{64,n} \end{bmatrix} \quad Y = \begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,n} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{64,1} & y_{64,2} & \cdots & y_{64,n} \end{bmatrix} \tag{5.6}$$

Each row corresponds to a particle and each column is a unique variable where n depends on the number of control points we use to represent the surface. As stated before, we use 64 particles to efficiently utilize CPU power, and thus there are 64 rows. X holds each particles current position and Y holds each particles personal best position. Columns from these matrices are chosen using a random grouping method designed for a B-spline surface representation and put into sub swarms to be optimized using the CCPSO2 algorithm described above.

The grouping algorithm is applied before each iteration, prior to evolving the sub swarms. Each particle's personal best fitness value must be reevaluated at the beginning of each iteration as a result of being assigned to a new sub-swarm. Special care must be taken to ensure that the context vector $\hat{y}$ is reconstructed properly, its components corresponding to the newly formed groups. This ensures that optimal control points locations found in previous iterations can be used in future iterations. Also before calling the fitness function to evaluate a particle we must make sure that the variables in the input vector are rearranged and placed in their original locations in order to obtain the correct value.

Intelligent grouping strategy

The original CCPSO2 algorithm uses a simple random grouping strategy to create the sub-swarms. All variables have the same probability of being placed together which is good for optimization problems where dependency is unknown. Because of the frequent regrouping, dependent variables have a good chance of being optimized together at some point by CCPSO2. A high frequency of regrouping should benefit large scale non-separable problems since it is more likely that interacting variables will be placed in the same sub component. In our case however we have some prior knowledge of variable interaction in a B-spline surface. Using this fact we have designed a more intelligent grouping algorithm to improve performance. A cubic B-spline surface patch requires a 4x4 grid of control points to define it thus the optimal location of these control points are heavily dependent on each other. Following this we would like to randomly group control points based on location relative to each other. Each control point has 3 degrees of freedom defining its location relative to its initial position on the dividing plane, which was constructed using principal component analysis at the start of the optimization. Random regions on the surface are constructed and the control points are grouped according to which region they lie in. The 3 displacement variables belonging to a control point in a region are added to that region's subcomponent. All of the subcomponents are then used in one iteration of CCPSO. Regions are constructed by introducing randomly placed points in the control point grid. All of the spatial coordinates closer to point i than any other point belong in point i's region. These regions are equivalent to the Voronoi diagram of the randomly placed points which is shown in Figure by the solid lines. The points are generated to have a uniform spatial distribution within the grid. In order to calculate which region a control point belongs to a Delaunay triangulation is formed which is used as input into a nearest neighbor algorithm implemented by Matlab. The nearest randomly placed point to each control point is its subcomponent.

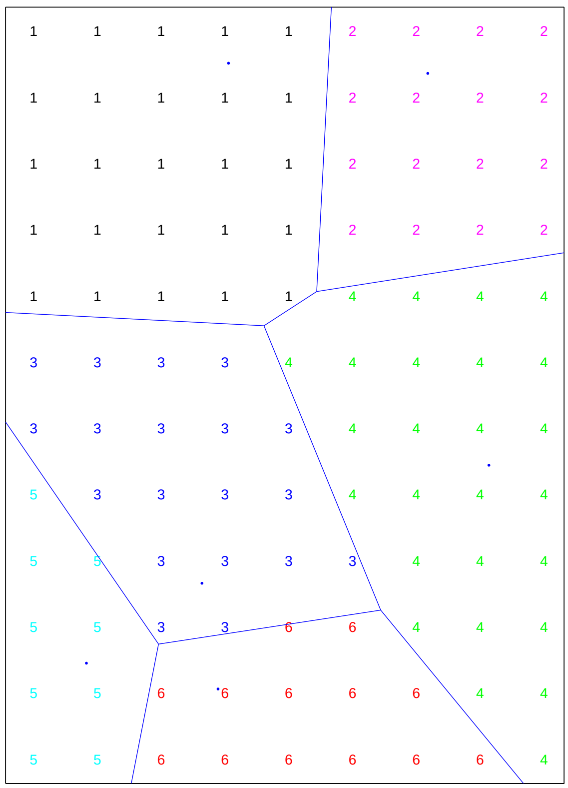


Figure 5-4: The control point grid is split into random regions before each iteration of CCPSO2 in order to improve performance. Numbers indicate which group the control point at this location is assigned to. The blue lines separate regions.

Generation of initial population

A key step of the optimization algorithm is the generation of the initial particle population where each particle represented a unique seam. As stated previously a swarm size of 64 was used which meant it was necessary to randomly generate 64 particles prior to optimization, each representing a unique seam. Ideally the initial population would cover enough of the search space to make optimization possible and efficient but would not contain particles which violated the non self-intersection constraint. Particles were generated by randomly perturbing the bisecting plane in 3D space. As long as the polygonal mesh formed by the control points does not contain any self-intersections than the resulting B-spline surface would not contain any self-intersections. To this end the control point mesh of each particle was tessellated and each tessellation was checked for self-intersections. Particles which passed this test were used in the initial population.

5.5 Experiments and results

Initial tests were completed on simple synthetic 3D volumes to ensure that the algorithm was capable of handling such a large number of optimization variables coupled with the non self intersection constraint. Volumes were created where the optimal seam was known a priori as shown in Figure 5-1. In this experiment the algorithm was easily able to find the correct seam. In subsequent tests the algorithm was tasked to find much more complicated seams where the desired result was not known a priori. The anatomical dataset from the Visible Human Female was used to construct a finite element model consisting of the major abdominal organs. This model was deformed by applying a force to the abdomen which was then reflected in the corresponding medical images producing deformed and undeformed volumes. Originally this model was constructed to evaluate registration methods. The surface optimization algorithm was then tasked to find a seam between the deformed and undeformed Visible Human Female volumes and the results are shown in Figure 5-5 for slices approximately 5 voxels apart. The RMS error along the seam before optimization was 21.39 and after

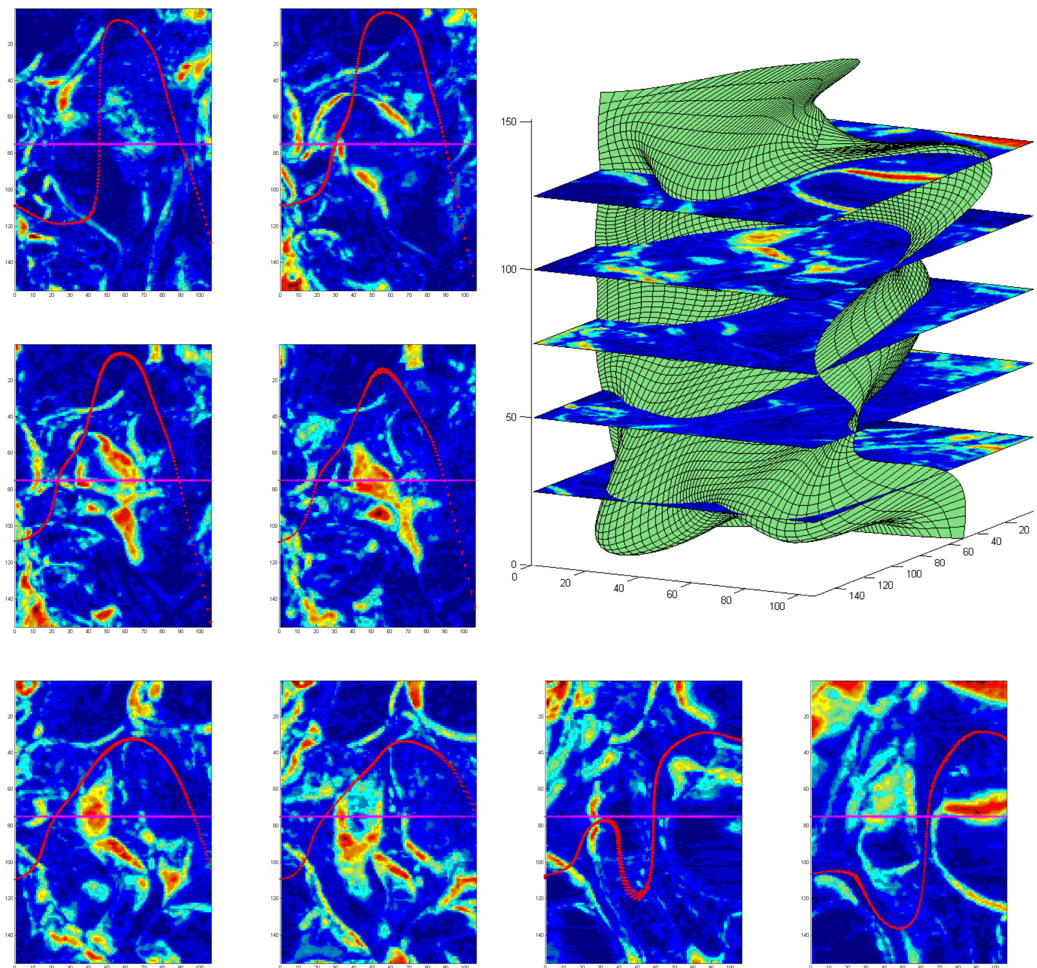


Figure 5-5: Slices through the Visible Human Female test volume show the final seam avoiding high intensity regions. The curve in each slice represents the surface's intersection with the image plane. The right image shows the final surface and 5 slices from the volume it traverses.

optimization this was reduced to 9.93 by the algorithm. The misalignment between adjacent volumes was significant and slices through the error volume changed fast. The control points were spaced around 10 voxels apart and the slices shown were 5 voxels apart so in order to reduce the RMS error further a refinement of the control point grid would be required. We feel that a control point spacing of 10 voxels is adequate for stitching clinical data. The Figure demonstrates the method's ability to avoid high intensity areas which would be more difficult to correct. The curve in each image represents the intersection of the optimized surface with that particular slice.

5.6 Conclusions

A global shape optimization algorithm was presented and applied to the task of 3D seam selection. The algorithm uses particle swarm optimization coupled with an intelligent grouping strategy which allows it to avoid local minima. Energy functions were evaluated to prevent self intersection from being introduced during evolution. It has potential to be extended to other medical image processing tasks such as segmentation or shape matching by slightly modifying the surface model and/or changing the fitness function. After experimenting with graph based methods this direction research was discontinued in favor of better performing approaches. The popularity of graph based techniques in all aspects of computer vision, ranging from registration to segmentation to denoising and many others, is impressive and demonstrates how effective they can be. Particle swarm optimization isn't as well suited to handle computer vision problems.

Chapter 6

Improving the mosaic: blending/filling

After the seam is calculated and the source volumes are non-rigidly registered there is one final step before the mosaiced training volume is complete and ready for integration into the simulation system. The trainee will have a better experience if the volumes are visually seamless and this usually requires some blending prior to stitching. Also, regions of the volume where no data was collected during scanning should have an ultrasound like appearance before being displayed in the simulator. This ensures the trainee's experience isn't degraded by unrealistic images. This chapter will elaborate on the approaches used to accomplish these tasks and provide examples of their application to the clinical datasets.

6.1 Blending of volumes using a discrete Poisson approach

Once the overlapping volumes have been brought into alignment minor discontinuities and intensity variations between adjacent volumes may still exist, necessitating the use of advanced blending techniques. The naive approach would be to use alpha blending, which simply calculates a weighted average of neighboring volumes near a seam. Employing this method would cause edges in the overlapping region, which may not be visible in both volumes due to shadowing artifacts, to appear blurry in the final blended volume. To overcome this issue we adapted a Poisson editing technique, which uses guided interpolation to perform image blending, for use with partially overlapping 3D data. This approach is a straightforward extension of [58], which was designed for photo editing. Other approaches for blending have been proposed, for example [87] is based on minimizing false edges and shows some improvement over [58] when using photos; however, ultrasound images are of too poor a quality so other works weren't considered in our application.

Before we review the theory and demonstrate how it can be applied to partially overlapping ultrasound volumes, some preliminary definitions will be helpful. Recall that in the group-wise registration chapter we defined the mosaicing function, $\Gamma(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}$, to map locations in the composite volume to a source ID, whose corresponding intensity value should be used to construct the mosaic. The composite volume is split into separate domains based on the mosaicing function where each domain, denoted by Ω_n, is linked to a specific source and represents its contribution to the composite volume. The left half of Figure 6-1 illustrates the problem using three partially overlapping images. We wish to blend the three images together in the region Ψ surrounding the seam. The boundaries between Ψ and the sources used in the mosaic are denoted by $\partial\Psi_1$, $\partial\Psi_2$ and $\partial\Psi_3$, which are shown in the right half of Figure 6-1. In this figure the image volumes are represented by the scalar functions I_1, I_2 and I_3 while the area between the volumes that we wish to fill is denoted by the scalar function F. The idea of the algorithm is to consider the calculation of image intensities F within the domain Ψ as an interpolation problem, where a specially constructed vector field is used for guidance. This vector field will be discussed shortly and ensures edge details are not lost during blending. The continuous formulation of

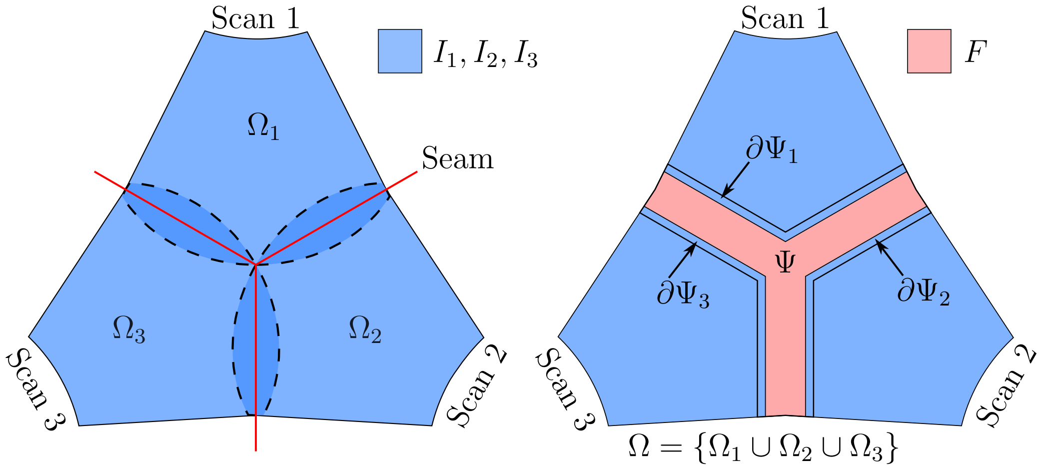


Figure 6-1: Regions/boundaries associated with 3 partially overlapping ultrasound volumes

this problem is shown below,

$$\min_F \int_\Psi |\nabla F - \mathbf{V}|^2 d\mathbf{x} \qquad \text{with} \qquad \begin{aligned} F|_{\partial\Psi_1} &= I_1|_{\partial\Psi_1} \\ F|_{\partial\Psi_2} &= I_2|_{\partial\Psi_2} \\ F|_{\partial\Psi_3} &= I_3|_{\partial\Psi_3} \end{aligned} \tag{6.1}$$

The boundary conditions force F to match the intensity values of the composite volume along the edges of Ψ, thus F is used to simply fill in the blended region. The solution of (6.1) is equal to the unique solution of the following Poisson equation using identical boundary conditions,

$$\Delta F = \nabla \cdot \mathbf{V} \qquad \text{over} \qquad \Psi \tag{6.2}$$

In (6.2) the symbol Δ represents the Laplacian operator and ∇ is the gradient operator. The vector field $\mathbf{V}$ is needed to include the features of each overlapping volume in the blended region and is defined as follows,

$$\text{for } \mathbf{x} \in \Psi, \qquad \mathbf{V}(\mathbf{x}) = \{\nabla I_n(\mathbf{x}) \, | \forall m : |\nabla I_n(\mathbf{x})| \geq |\nabla I_m(\mathbf{x})|\} \tag{6.3}$$

Equation (6.3) states that the guidance vector at $\mathbf{x}$ should be equal to the largest gradient of all the overlapping volumes at this location. This helps prevent shadow artifacts from diminishing the quality of edges in the blended region.

Next we will present the discrete version of (6.2) that we applied to our ultrasound mosaicing problem and then discuss how its solution must satisfy a particular system of linear equations, enabling the use of efficient solvers. This approach follows [58] but includes modifications for use with 3D data. A mathematical definition for the domain of the blended region is also presented and used in our implementation. Recall that $\mathcal{D}_n(\mathbf{x})$ was defined in Chapter 3 as a distance function which measures the minimum distance from point $\mathbf{x}$ to the region in the composite volume labeled as source n. The discrete set of voxels in the blended region can be expressed using the following equation,

$$\Psi = \left\{\mathbf{x} \in \mathbb{Z}^3 | \Gamma(\mathbf{x}) \neq 0, \exists n : 0 < \mathcal{D}_n(\mathbf{x}) < \varsigma\right\} \tag{6.4}$$

Here ς controls the distance from the seam where the blending will occur. The discrete counterpart for Ω_n is now defined as the set $\Omega_n = \{\mathbf{x} \in \mathbb{Z}^3 | \Gamma(\mathbf{x}) = n\}$. For each voxel $\mathbf{p} \in \mathbb{Z}^3$ in $\Omega = \Omega_1 \cap ... \cap \Omega_N$, let $\mathcal{N}_{\mathbf{p}}$ be the set of 8 connected neighbors in Ω. Also let $\langle \mathbf{p}, \mathbf{q} \rangle$ denote a voxel pair such that $\mathbf{q} \in \mathcal{N}_{\mathbf{p}}$. The boundaries of Ψ are discretized and can be expressed as $\partial\Psi_n = \{\mathbf{p} \in \Omega_n \backslash \Psi | \mathcal{N}_{\mathbf{p}} \cap \Psi \neq \varnothing\}$. Finally let $f_{\mathbf{p}} = F(\mathbf{p})$. Using the notation from [58] our task is to compute the set of intensities $f|_\Psi = \{f_{\mathbf{p}} | \mathbf{p} \in \Psi\}$. The finite difference discretization of (6.1) yields the following quadratic optimization problem,

$$\min_{f|_\Psi} \sum_{\langle \mathbf{p},\mathbf{q}\rangle \cap \Psi \neq \varnothing} (f_{\mathbf{p}} - f_{\mathbf{q}} - v_{\mathbf{pq}})^2 \text{ with } f_{\mathbf{p}} = I_n(\mathbf{p}) \text{ for all } \mathbf{p} \in \partial\Psi_n \tag{6.5}$$

where,

$$v_{\mathbf{pq}} = \{I_n(\mathbf{p}) - I_n(\mathbf{q}) | \forall m : |I_n(\mathbf{p}) - I_n(\mathbf{q})| \geq |I_m(\mathbf{p}) - I_m(\mathbf{q})|\} \tag{6.6}$$

It should be noted that if a particular image volume isn't defined at either $\mathbf{p}$ or $\mathbf{q}$ then it's gradient at this location is considered to be zero. The solution of this optimization problem also satisfies the following system of linear equations,

$$\text{for all } \mathbf{p} \in \Psi, \;\; |\mathcal{N}_{\mathbf{p}}| \, f_{\mathbf{p}} - \sum_{\mathbf{q} \in \mathcal{N}_{\mathbf{p}} \cap \Psi} f_{\mathbf{q}} = \sum_{n} \left(\sum_{\mathbf{q} \in \mathcal{N}_{\mathbf{p}} \cap \partial \Psi_n} I_n(\mathbf{q}) \right) + \sum_{\mathbf{q} \in \mathcal{N}_{\mathbf{p}}} v_{\mathbf{pq}} \qquad (6.7)$$

We used the preconditioned conjugate gradient method to solve (6.7) and obtained good blending results using our clinical datasets in a matter of minutes.

The overall goal of this work was to produce a seamless mosaiced image volume, removing any obvious transition between neighboring sources that might appear while the user was scanning with the simulator. Now we discuss the blending results we achieved using clinical volumes acquired at the University of Massachusetts Medical School. In the leftmost column of Figure 6-2 some of the irregularities between adjacent volumes, which remained after group-wise non-rigid registration, are identified. The rightmost column shows their remediation using the proposed blending technique. Elimination of the discontinuities in the composite volume makes for a much more realistic scanning experience.

6.2 Filling volumes using a discrete Poisson approach

This section will briefly describe the method we used to fill the regions of the composite volume where there was no source information, which appear black in Figure 6-2. The user experience with the simulator is improved if the scan plane they are currently viewing slowly transitions to gray instead of abruptly ending where no clinical data was acquired. The difficulty of acquiring and stitching ultrasound volumes, which encompass every possible view that the simulator may need, makes this step necessary. The idea is to set the exterior boundary voxels of the cubic volume to gray and then use the membrane interpolant based on the Poisson equation to fill in the black areas. This is a very well-known interpolation technique thus we won't discuss its implementation here. Next, multiplicative noise or speckle is added to make the

Slices of mosaiced volume before and after blending:

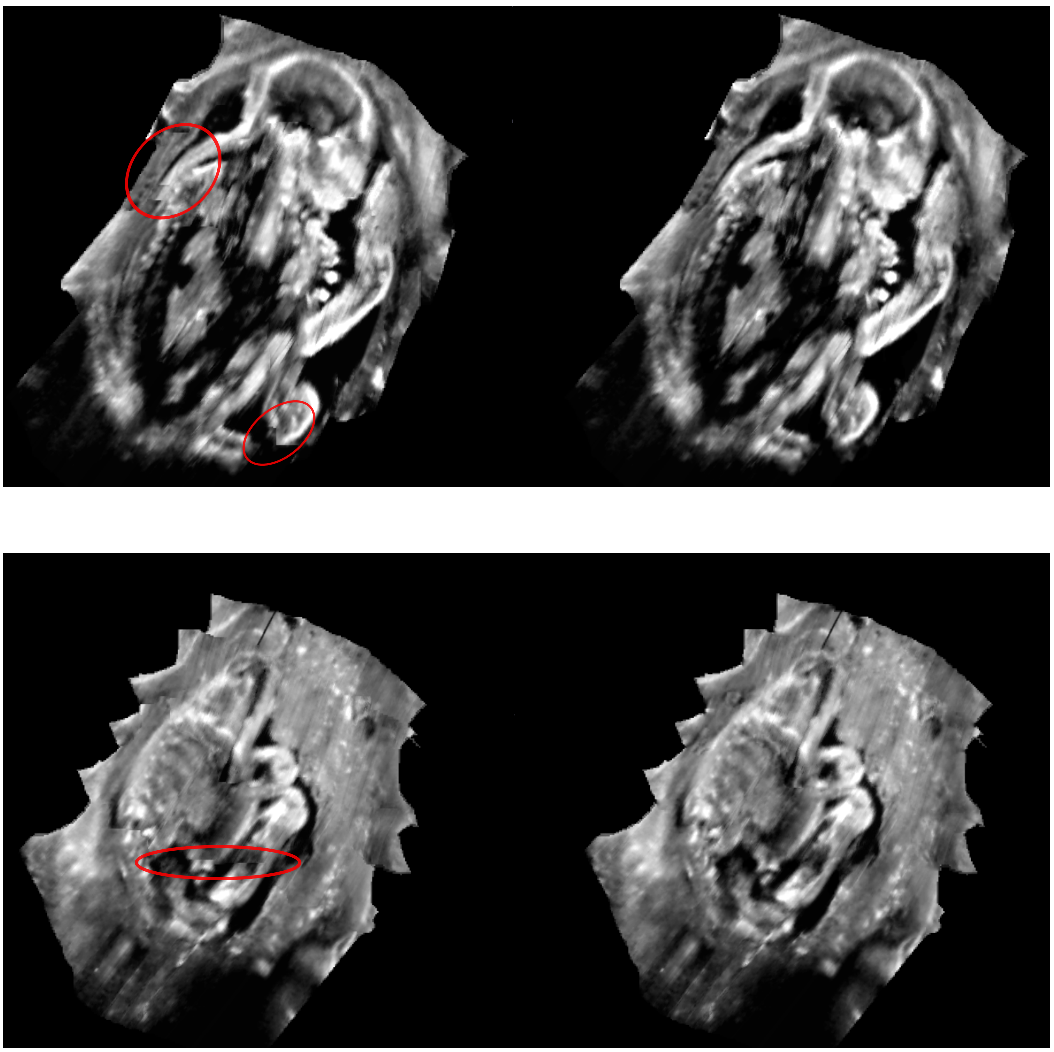


Figure 6-2: Results of blending algorithm. Jagged edges and noticeable transitions between sources are removed with circles spotlighting highly effected areas.

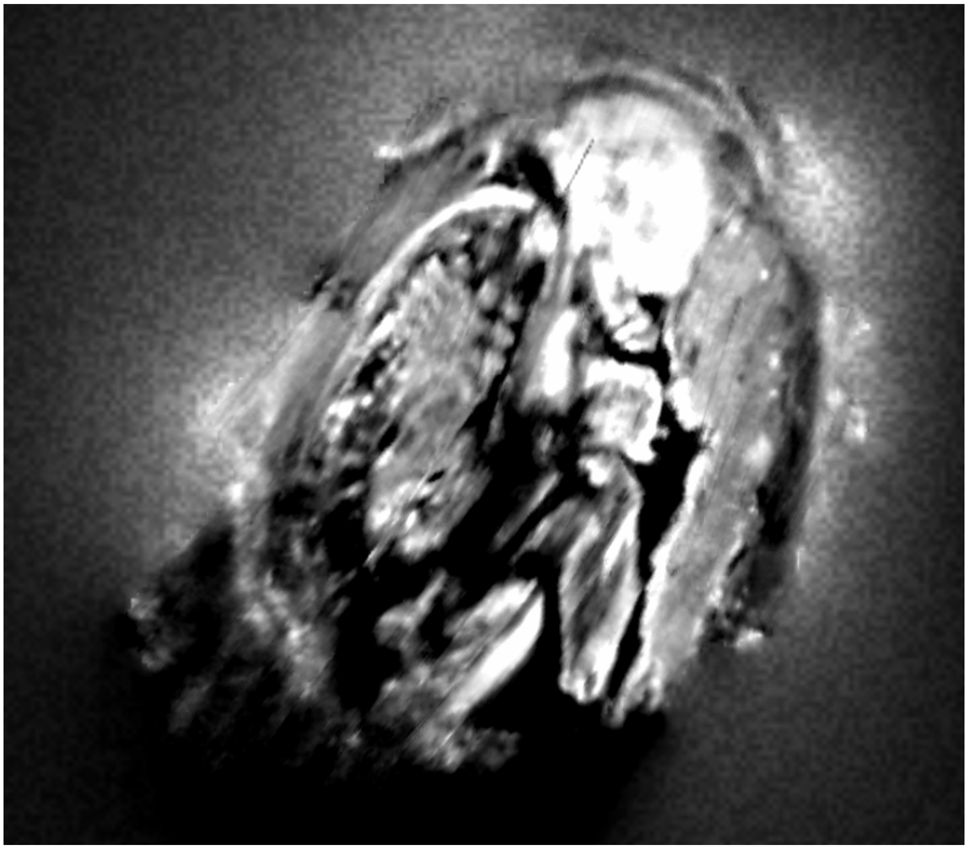


Figure 6-3: Results of filling technique applied to volume shown in top right of Figure 6-2.

filled in areas look more ultrasound like. The interpolation step can completed using (6.1) with the guidance vector **V** removed. The domain Ψ to be filled is defined as the black region surrounding the image information shown in Figure 6-2. The boundaries are now the exterior voxels of the cubic volume, which we set to gray, and the exterior voxels of the mosaiced data. Figure 6-3 shows the results of the filling procedure.

This simple approach achieves the desired result without putting an additional burden on the CPU/GPU; however generating the fill in real-time, based on the current simulator ultrasound image, would appear more realistic to the user. More complex algorithms have been proposed to perform image in-painting/blending, however our implementation of [37] resulted in undesirable texture patterns in the filled area. This method best suits our goal of developing a simulator which runs on most available laptops.

Chapter 7

Clinical results with fetal ultrasound

This chapter will describe the clinical component of the research of this dissertation. For the purpose of acquiring actual clinical ultrasound data, from which training datasets were produced, 11 pregnant subjects were scanned by experienced sonographers at the University of Massachusetts Medical School following an approved IRB protocol. First, we will discuss the software/hardware configuration that was used to conduct these scans, which included some custom mechanical design. With the data collected using this arrangement we generated seamless 3D fetal mosaics, that is, the training datasets, loaded them into our ultrasound training simulator, and then subsequently had them evaluated by the sonographers at University of Massachusetts Medical School (UMMS) for accuracy. These mosaics were constructed from the raw scan data using the techniques developed in Chapters 2 and 3. This chapter will also discuss the specific training objectives that were established based on the input from our collaborators in the obstetrics sonography group. Important fetal measurements are reviewed, which form the basis for training in obstetrics ultrasound. Next, we discuss how our subject scanning/mosaicing procedure was tailored with these measurements in mind. Finally clinical images demonstrating the sonographer making fetal measurements in practice, which were acquired by the Philips iU22 ultrasound machine from a live subject, are compared with screenshots of corresponding images produced with our simulator.

7.1 Motivation

Obstetrics sonography is a critical tool for monitoring the health of both mother and child during pregnancy. Using their skillset, sonographers trained in obstetrics ultrasound are able to determine gestational age and monitor growth to verify the proper development of the fetus. The sonographer would be the first to notice any abnormalities which may cause harm to the mother or child. Due to the variability between pregnancies this a challenging field in ultrasound medicine. The skill to correctly identify and measure the anatomy for a wide range of gestational ages and fetal positions requires a great deal of hands-on experience to master. The sonographer must not only be able to identity the anatomy in the images they are viewing; they must also develop the motor skills necessary to move the transducer to the correct location based on their current view. Normally this skill would be honed by practicing on patients; however, with the availability of a training simulator for obstetrics ultrasound, students can start to develop these motor skills by practicing on a simulator before ever scanning a live person. The generation of obstetrics training volumes for simulation purposes will be discussed right after the clinical acquisition configuration is presented in the next section.

7.2 Configuration of software/hardware

The main contribution of this dissertation is the development of a novel mosaicing algorithm used to stitch partially overlapping 3D volumes; however, we should review the techniques used to construct the individual 3D volumes since this is the first step in generating the mosaics that were incorporated into the simulator. As noted earlier we used the software package Stradwin [78] (to be described later in this section), which was developed at the University of Cambridge, for image acquisition and construction of the individual overlapping 3D volumes. In this section the basics of 3D swept ultrasound will be discussed along with our particular clinical configuration. For clinical scanning at UMMS, a laptop running the Stradwin software was connected to the video output of a Philips iU22, a high end ultrasound scanner often used in obstetrics ultrasound, using a common Belkin video capture device running through a USB port. Also connected to the laptop, via a SERIAL to USB adapter, was a 6 degree of freedom (DoF) Ascension Technologies trakSTAR system which was used track the ultrasound transducer. The three main components of the trakSTAR system include the position sensor, a DC magnetic transmitter, and the processing unit connected to our laptop. The interconnection of these hardware/software components is illustrated by Figure 7-1. Solid lines represent wired connections and arrows are used to denote the direction of data flow between components. A dashed arrow represents the DC magnetic field generated by the transmitter. This field is subsequently measured by the position sensor in order to determine the position and orientation of the transducer to which it has been attached. The left side of Figure 1 is associated with image acquisition while the right side is associated with position tracking. In order to accurately track the ultrasound transducer (a Philips C5-1 convex array transducer), the position sensor had to be fastened securely enough that the sonographer's grasp wouldn't shift it during prolonged scanning. To this end we designed a clamshell type bracket using Solidworks which fit snugly over the end of the transducer. In order to capture the non-uniform shape of the transducer's handle we placed it in a vice and utilized the trakSTAR's position sensor to capture a 3D point cloud that was subsequently turned into a NURBS (Non-Uniform Rational B-Spline) surface and imported into Solidworks. Both halves of this bracket are shown in Figure 7-2. A notch, visible in Figure 7-2, was also cut out for holding the position sensor. The bracket was constructed using a 3D printer, enabling the contours of the ultrasound transducer to be easily replicated. Nylon bolts were then used to fasten both sides together. Now that the position sensor was securely fastened to the transducer the next challenge was to build a structure with an arm that could hold the transmitter close to the subject's abdomen since its range is limited. We utilized the mid-range transmitter from Ascension Technologies which guarantees the position/orientation

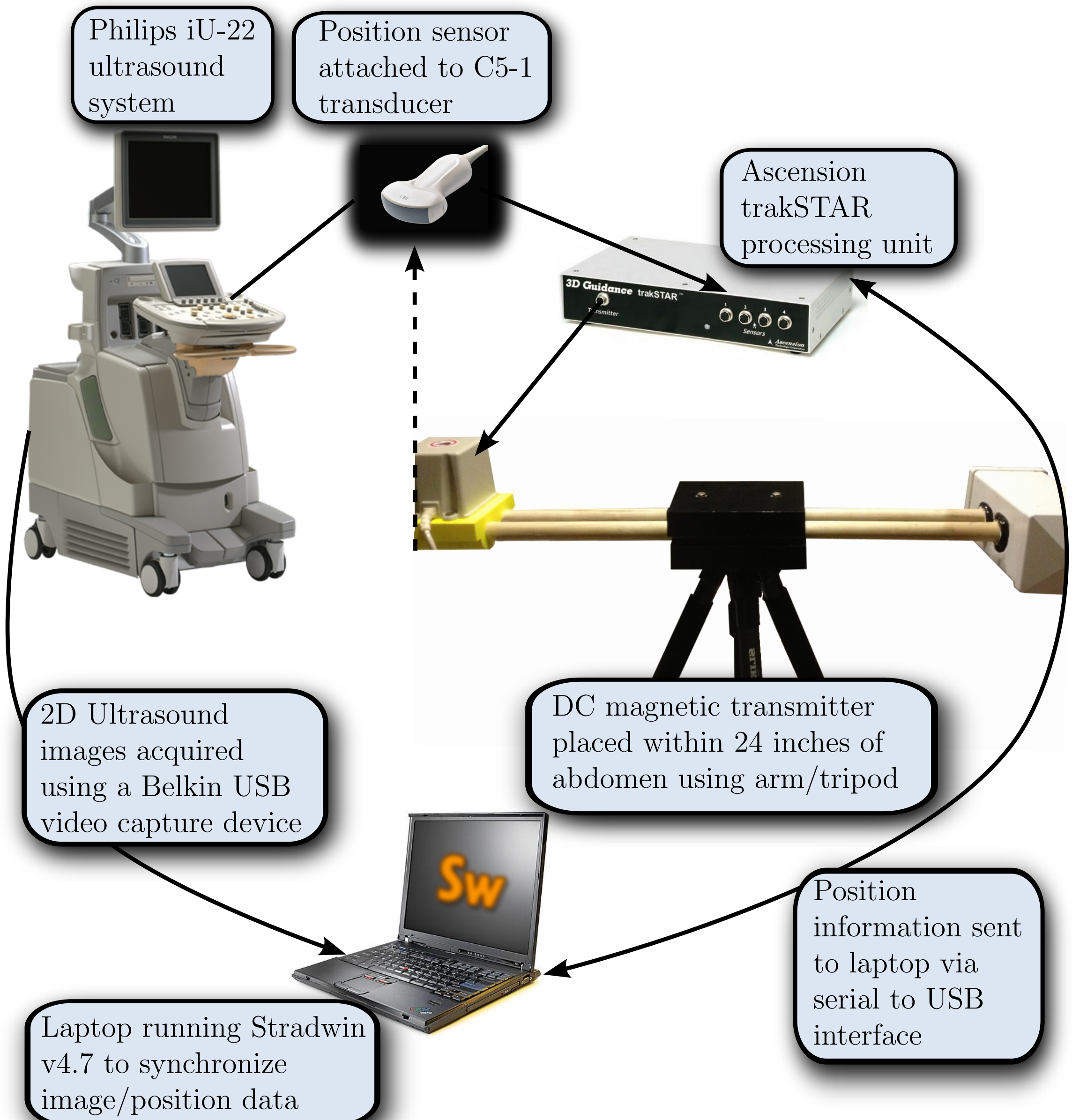


Figure 7-1: Diagram showing the interconnection of hardware/software components used for the freehand 3D scanning of patients

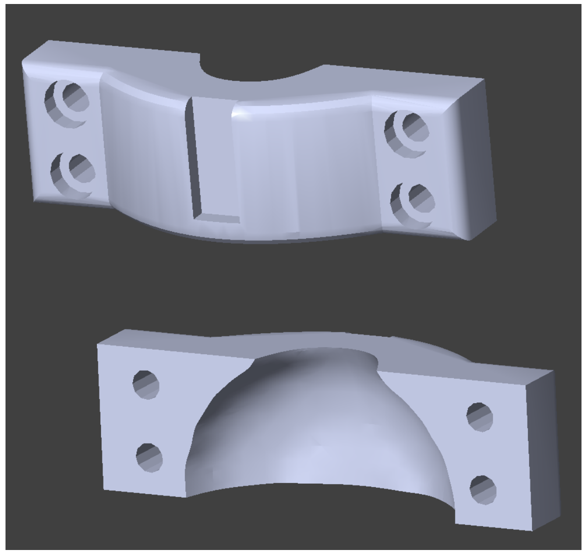


Figure 7-2: 3D rendering of bracket used to fasten position sensor to Philips C5-1 transducer

measurements to be accurate to within stated specifications [4] as long as the sensor remains inside a cube box with side dimensions of 31' directly in front of it. Outside interference can introduce noise and artifacts into the measurements and unshielded magnetic metals can cause distortion so the arm's materials were carefully chosen. We also wanted the option to swing the transmitter over the hospital bed so both sides of the abdomen could be accurately captured without requiring us to move the transmitter mid-scan, as the position sensor's accuracy decreases when its distance to the transmitter increases. The solution was to purchase a very sturdy tripod and design/build an attachment to hold the transmitter. The final design is shown in Figure 7-3. The tripod was made from carbon fiber so as to not have any metallic structures near the transmitter which could distort the magnetic field it generated. The arm itself was made out of two wooden dowels while the tripod/transmitter mounts were designed in Solidworks and constructed using a 3D printer. Since the transmitter is fairly heavy, the arm required a sliding counterweight which can be seen in the left side of Figure 7-3. The counterweight was constructed out of plastic and filled with sand, once again ensuring nothing interfered with the magnetic field generated by the transmitter. Also two plastic tubes pass through the main body of the counterweight, allowing it to slide along the wooden dowels. The Solidworks design of the mount

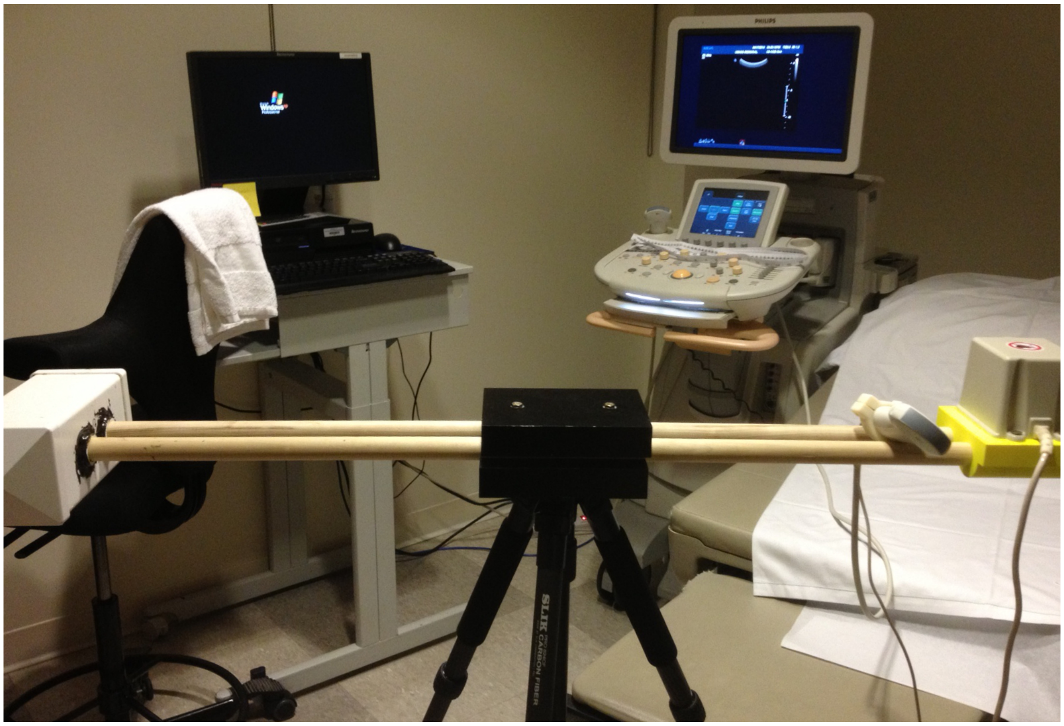


Figure 7-3: Photo showing our clinical setup at the University of Massachusetts Medical School

which connects the dowels to the tripod is shown in Figure 7-4. A threaded brass insert is pressed into it which allows the mount to be securely fastened to the tripod. It clamps the dowels in place using two additional bolts that can be seen in Figure 7-4. The right sides of Figure 7-3 and Figure 7-4 show the transmitter mount, which slides over the ends of the dowels and is held in place with glue. Two threaded brass inserts allow the transmitter to be screwed to it. Using this design we were able to capture the position/orientation information in a clinical setting accurately enough to produce realistic training volumes for our simulator.

Now that the hardware configuration for data collection has been presented we will discuss the software component in a bit more detail. Stradwin generates 3D ultrasound volumes by tagging each 2D image acquired by the video capture card with the corresponding position/orientation information given by the trakSTAR sensor. Next it places them appropriately in 3D space based on this information. It should be noted that since the sensor is arbitrarily mounted to the transducer a calibration procedure is required in order to calculate the transformation from the sensor to the

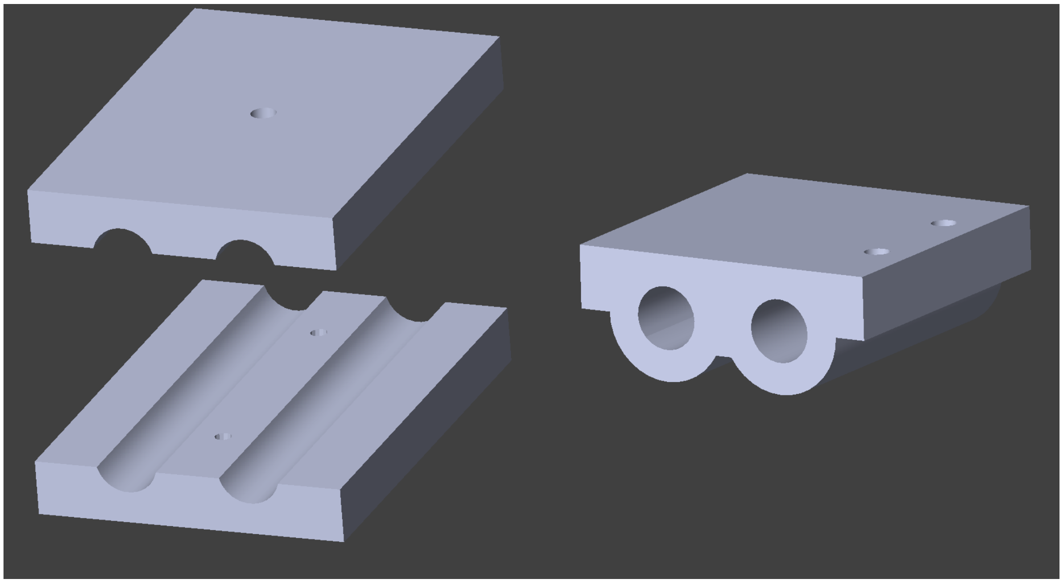


Figure 7-4: 3D rendering of parts used to construct the transmitter arm. The left-most images represent the tripod mount, which fixes the arm in place by clamping the wooden dowels. The right-most image is the transmitter mount which was fastened to the end of the dowels.

corner of the ultrasound image, which is the actual position/orientation we desire to store. The procedure we used is described in detail in [64]. The calibration only had to be performed once, when the position sensor was fixed to the transducer, so this method sufficed; however, much faster methods have been proposed recently [31]. A quicker calibration procedure could be performed prior to each scan, thus ensuring reconstruction accuracy even if the position sensor had been inadvertently shifted between scans. Stradwin generates re-slices of the 3D ultrasound volume in its viewing window directly from the acquired 2D images and thus doesn't allow 3D non-rigid registration between overlapping volumes. Interpolation methods [73] can be used to construct volumes with uniform voxel grids from the 2D images, which have been positioned in the transmitter's 3D coordinated system. Stradwin can export these volumes which can be used as input into our stitching algorithm.

7.3 Obstetrics image volumes for affordable ultrasound simulator

Each ultrasound mosaic was constructed with a number of obstetrics training tasks in mind. Firstly, the sonographer must be able to determine the fetal position, which becomes very important as the expecting mother nears delivery. One example would be occiput anterior, where the smallest part of the baby's head leads the way through the birth canal; thus is the most preferred. A C-section is usually recommended for difficult positions such as breech, making this a fundamental observation every sonographer needs to be comfortable with. Placental position is also visible in our training volumes. The bladder, lower uterine, and cervix are detectable in the majority of the cases.

We have implemented the most common fetal assessments and measurements in our training modules. After the mosaics were constructed and incorporated into the simulator landmarks were added to the composite volumes by Dr. Petra Belady, an assistant professor in the Dept. of Obstetrics and Gynecology at UMass Med. School. The students locate anatomical structure, determine fetal position and take fetal measurements within the simulator as they work their way through each learning objective. The identified structures and measured values are compared to the previously inserted landmarks to determine correctness. Currently implemented measurements include the amniotic fluid index, bi-parietal diameter, abdominal circumference, and femur length. The amniotic fluid index is calculated by measuring the amount of fluid in four separate quadrants within the uterus. Biparietal diameter is the diameter across the developing baby's skull, and is useful for estimating weight. Abdominal circumference is important in assessing size/growth during pregnancy and can be measured in the plane where the stomach is visualized. Finally the femur length is another mandatory measurement enabling the sonographer to exclude certain medical conditions such as dwarfism.

7.3.1 Scanning procedure for fetal ultrasound

This section will discuss the scanning protocol we used for freehand 3D swept ultrasound of a pregnant subject. The collected data was used in the generation of the composite training volumes, or fetal mosaics, described above. Based on our experience with the stitching process, using individual volumes created from ultrasound B-mode images acquired in both the sagittal and transverse planes of the fetus, we have determined that stitching was easier when B-mode images of the sagittal plane were used. These images were acquired along a set of parallel scan paths, with adjacent paths shifted about 2", formed by moving the transducer right to left across the subject's abdomen. The reasoning behind this is that the effects of the shadowing artifacts on the stitching process are amplified when stitching together individual volumes produced from images of the transverse plane, collected along scan paths which run superior to inferior. This may be due to the fact that ultrasound images of anatomical structures in the overlapping regions of adjacent volumes have been obtained from opposite sides of the abdomen.

A volume produced using images of the transverse plane on one side of the abdomen may display the fetus's arms and legs clearly while the vertebrae are completely absent. Conversely, scanning on the other side of the abdomen may give very clear images of the baby's back; however the arms and legs are missing from the image volume. This may be due to the attenuation of the ultrasound as at travels through the fetus or due to some shadowing effect. Because of this, stitching volumes acquired in this fashion is difficult because there are no or few common features which can be used to align the individual volumes and thus the registration (alignment) solution isn't well defined. We have found that stitching overlapping volumes together, which were produced with images of the fetus's sagittal plane, produced better results based on evaluation by clinicians. The reasoning is that the features captured by multiple ultrasound volumes acquired in this fashion are generally very similar in the overlapping region between volumes. The two different scanning directions are shown in Figure 7-5 with the preferred scan direction/orientation designated as the primary.

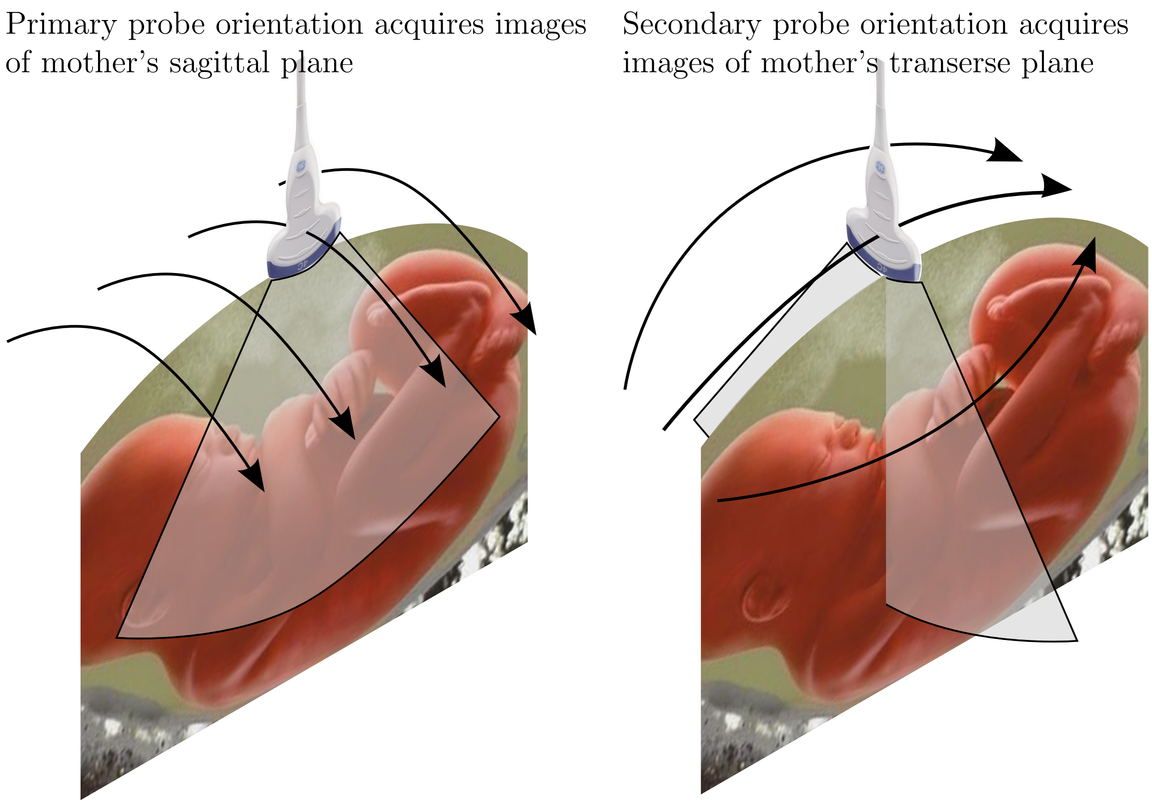


Figure 7-5: Illustration of the two possible scan paths along the abdomen which can be followed to obtain individual ultrasound volumes from a pregnant subject using freehand 3D techniques

To acquire the individual volumes necessary for the creation of a complete obstetrics training volume we propose to scan in the primary direction/orientation first, collecting 4-5 volumes encompassing the entire abdomen and fetus. These volumes should overlap minimally in order to limit the effects of fetal movement during the stitching process but also must contain enough shared structures to make registration possible. Also it is very important that the scan path used to produce the individual image volumes cover the full extent of the mother's abdomen because adding small volumes to an otherwise complete composite volume is difficult. Care must also be taken to completely capture the fetal head and bottom as well a subset of the mother's organs, which are needed to help the sonographer orient themselves. This procedure should be repeated a 2-3 times in order to minimize the likelihood that the baby was actively moving during the image collection which should result in 12-15 volumes. The sonographer should also scan along the secondary direction/orientation to acquire additional scans which may be used if the primary volumes prove too hard to stitch. Finally, fetal biometric measurements are taken which can be utilized after the mosaicing process in order to prove that the volume was constructed accurately. In addition to these measurements, the corresponding 2D ultrasound images are also saved because they identify the planes in which the trainee should complete the measurement tasks using the simulator, once the training volume has been generated of course.

SCAN PROTOCOL:

1. Scan so as to ascertain the baby's actual position
2. Determine the scan paths that will capture the baby's head and feet
3. Scan the abdomen of the pregnant subject with right to left sweeps acquiring images of the sagittal plane
4. Repeat the scanning in Step 3
5. Scan the abdomen of the pregnant subject with superior to inferior sweeps acquiring images of the transverse plane
6. Repeat the scanning in Step 5

Table 7.1 contains an overview of the clinical scans conducted at the University of Massachusetts Medical School. Each date corresponds to a different subject and provides information on the number/type of freehand 3D scans acquired. Because of fetal movement not all sessions produced a usable training volume for the simulator, which is noted in the last column.

Table 7.1: Scans of pregnant subjects conducted at UMass Medical School

Date conducted	# Scans comprised of transverse images.	# Scans comprised of sagittal images.	Suitable for training
30 April 2012	1 scan	0 scans	No, poor fetal resolution
5 May 2012	1 scan: 9 sweeps	0 scans	Yes
9 Jan 2013	1 scan: 10 sweeps	0 scans	Yes
14 Mar 2013	0 scans	2 scans: 12 sweeps/scan	No, too much movement
9 May 2013	2 scans: 10,11 sweeps/scan resp.	2 scans: 3,4 sweeps/scan resp.	Yes
25 June 2013	2 scans: 9,12 sweeps/scan resp.	0 scans	No, too much movement
26 June 2013	2 scans: 7,10 sweeps/scan resp.	0 scans	No, too much movement
21 Nov 2013	2 scans: 7,9 sweeps/scan resp.	2 scans: 8,9 sweeps/scan resp.	Yes
12 Dec 2013	2 scans: 6,11 sweeps/scan resp.	2 scans: 8,9 sweeps/scan resp.	No, too much movement
22 Jan 2014	1 scan: 11 sweeps	2 scans: 6,9 sweeps/scan resp.	Yes
17 April 2014	1 scan: 7 sweeps	3 scans: 5,5,6 sweeps/scan resp.	No, too much movement
14 May 2014	1 scan: 7 sweeps	3 scans: 10,7,8 sweeps/scan resp.	Yes

7.3.2 Using ITK-SNAP to isolate best view of anatomy for specific training objective

Due to the large amount of image overlap we acquired from each subject, using the scanning procedure outlined above, multiple views of the significant anatomical structures were captured. The position/orientation of the transducer during each sweep determines the region of the fetus that appears clearest; thus, concerning each fetal measurement/training objective, it is desirable to choose the volume that provides the most realistic image for the simulator. Using the mosaicing algorithm developed in Chapter 3, the sonographer has the ability to designate regions within the individual volumes that they feel best represent certain anatomical features. Each region which has been identified to contain the clearest view of a particular structure is used to generate that section of the composite volume. For example when considering a fetal measurement such as femur length, if an individual volume is identified as providing the most realistic view, then it should be used to construct the femur in the composite training volume. During production of the training volumes described in this dissertation special care was taken to select the best available images for the fetal head, femur, and abdomen, which are all crucial to the objectives described above.

The ability to preselect certain regions can be easily implemented in the framework proposed in Chapter 3 by introducing an energy term into the mosaicing algorithm which penalizes the seam if user designated regions are labeled with a different source ID than specified. For example if volume 1 contains the best view of the femur then a mask for volume 1 is created which segments this region. This mask penalizes the seam selection algorithm if it chooses to generate the femur in the composite volume using another source. Figure 7-6 demonstrates the construction of the mask using ITK-SNAP [86]. In this figure the leg is segmented because this particular volume provides the most realistic view for the simulator. The mosaicing algorithm will now use this source to generate the femur in the composite training volume.

Figure 7-6: ITK Snap is used to designate regions of individual volumes which must be included in the final mosaiced training volume. User interaction during the mosaicing process allows the clearest images to be selected from the group of individual 3D volumes, for each anatomical structure of interest. This example demonstrates the fetal leg being highlighted.

7.3.3 A catalog of simulator volumes

This section describes the library of fetal volumes generated during our research and subsequently evaluated at the UMMS for training value. The measure of quality for a particular composite volume was whether or not the sonographer could successfully complete each training objective using the simulator. The results of this evaluation are detailed in Table 7.2. Each date corresponds to a unique subject being scanned. An X mark indicates that the sonographer was unable to accurately perform the task listed, while a check mark indicates success. A check minus means that the image quality was marginal; however, the sonographer was still able to complete the particular training objective. The mosaics constructed using the datasets from 21 November 2013, and 14 May 2014 produced the highest quality training volumes. These are the only 2 datasets in which all 7 training objectives can be completed. Referring to the table we see that the capturing the bladder, lower uterine, and cervix was the most difficult. This could be due to the sonographer concentrating on capturing the fetus and neglecting to capture the surrounding structures completely.

Also, during stitching process we focused more on the fetal anatomy than on the mother's. Each volume listed in Table 7.2 adds training value to the simulator in some respect and combining the positive aspects of each provides a solid curriculum in obstetrics ultrasound. It should be noted that each volume is approximately 500 MB.

Table 7.2: Evaluation of training volumes by Dr. Petra Belady, Univerisity of Massachusetts Medical School

Date subject scanned	Bladder, lower uterine, cervix visible	Fetal position distinguish-able	Placental position visible	Amniotic fluid index measur-able	Biparietal diameter measur-able	Abdominal circumfer-ence measurable	Femur length measur-able
5 May 2012	✓	✓	✓	✓	✓	✓	✗
9 Jan 2013	✓	✓	✓	✓	✓	✓–	✓–
9 May 2013	✗	✓	✓	✓	✓	✓	✓
25 June 2013	✗	✓–	✓	✗	✓	✓	✓
21 Nov 2013	✓	✓	✓	✓	✓	✓	✓
12 Dec 2013	✗	✓	✓	✗	✗	✗	✗
22 Jan 2014	✗	✓	✓	✓	✓	✓	✓–
17 April 2014	✗	✗	✓	✗	✓	✓	✓
14 May 2014	✓	✓	✓	✓	✓	✓	✓

7.3.4 Comparison between clinically acquired and simulator generated images for fetal measurements

In this section we will compare the clinical ultrasound images with the simulator images for the abdominal circumference, parietal diameter, and femur length measurements. The initial positive assessment of the simulator by sonographers at the UMMS speak to the realism of the training volumes generated using the methodology presented in this dissertation. It is also encouraging that the fetal measurements obtained with the Philips iU22 ultrasound machine during initial scanning are in satisfactory agreement with those taken within the simulator, using the subsequently constructed 3D mosaic. For that particular evaluation, the sonographer who completed the initial fetal measurements on the live subject was different than the sonographer who operated the simulator. The clinical and simulated biometric measurements for two training volumes are presented in Table 7.3.

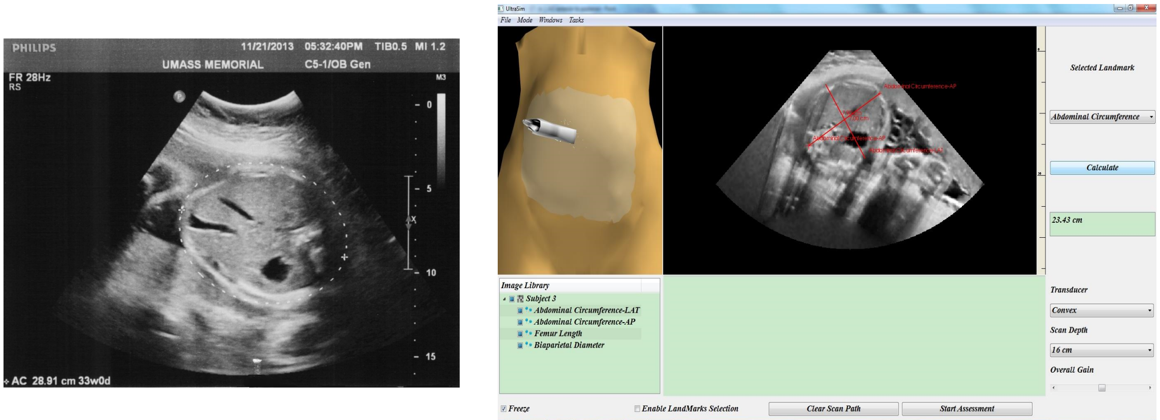
Table 7.3: Clinical vs. Simulated biometric measurements (cm)

Date subject scanned	Type of measurement	Abdominal circumference measurement	Biparietal diameter measurement	Femur length measurement
9 May 2013	Clinical	22.31	6.48	4.68
9 May 2013	Simulated	24.67	7.6	5.21
21 Nov 2013	Clinical	28.91	8.31	6.21
21 Nov 2013	Simulated	23.43	8.3	5.6

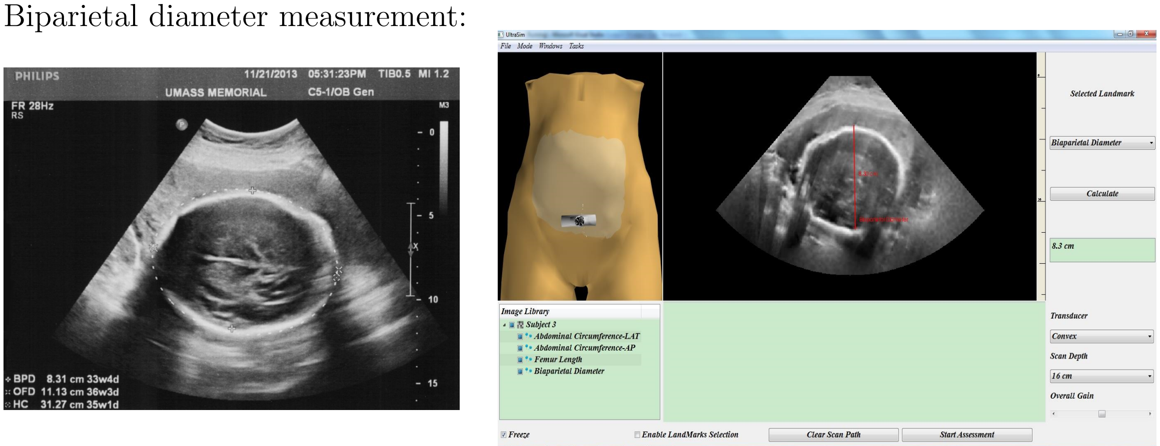
Using the dataset from 21 of November 2013 the sonographer measured 23.43 cm, 8.3 cm, and 5.6 cm for abdominal circumference, biparietal diameter, and femur length respectively within the simulator environment. These numbers can be directly compared with the measurements taken using the Philips iU22 on the same patient, which were 28.91 cm, 8.31 cm, and 6.21 cm respectively. We see that the biparietal diameter measured in the simulator is in good agreement with the clinical result. The femur length and abdominal circumference measurements vary but are within an acceptable level of error when considering the sensitivity of the measurements to factors like fetal position and differences between particular sonographers. Although the stomach bubble is distinct in both images where abdominal circumference is measured the umbilical vein doesn't have the same appearance, indicating that the sonographers measured the AC in different planes causing some deviation between the simulated and clinical result. Also, errors introduced by swept 3D ultrasound and non-rigid registration may affect these numbers. The simulated measurements using the dataset from 9 of May 2013 also agree fairly well with the clinical measurements. The abdominal circumference measurements were 22.31cm versus 24.67cm for the Philips iU22 and simulator respectively. The biparietal diameter measurements were 6.48cm versus 7.6cm for the Philips iU22 and simulator respectively. Finally, the femur length measurements were 4.68cm versus 5.21cm.

It is important to remember that the most important factor in this evaluation is how realistic the images look within the simulator. Figures 7-7 and 7-8 present a comparison of the clinical images with the simulator images, which are very similar in appearance. Because of fetal movement during prolonged scanning, the fetus was captured in different positions thus the images are not identical; however, it is hard to distinguish the image produced by the simulator from the image produced by the Philips iU22 for this volume. For each particular training objective, the sonographer needs to locate certain anatomical features in order to get accurate measurements. The first row of Figure 7-7 shows the abdominal circumference measurement, where the stomach cavity (or stomach bubble) is visible in both the simulator and clinical images. The second row shows the biparietal diameter, where the appearance of the cranium is very distinct. The last row shows the femur length being measured.

Abdominal circumference measurement:



Biparietal diameter measurement:



Femur length measurement:

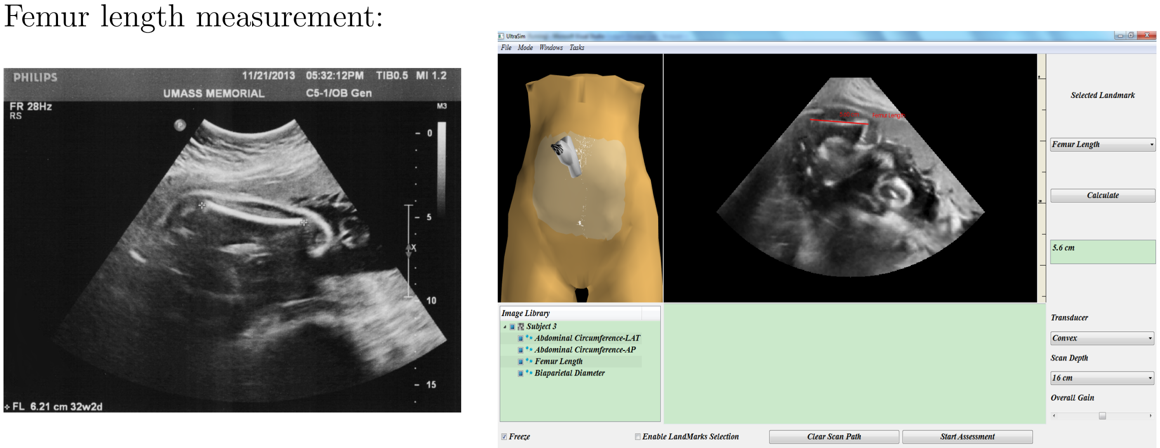
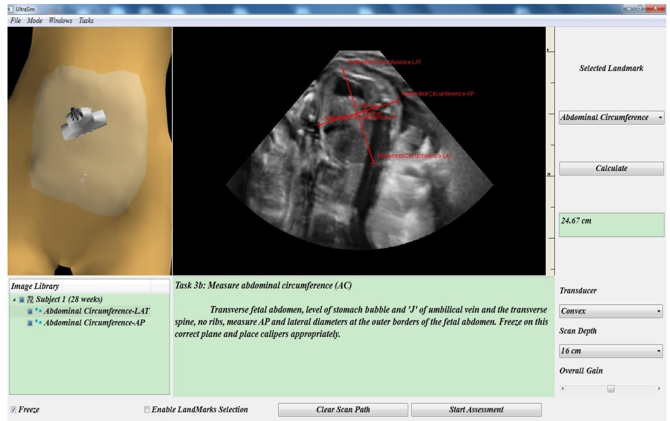
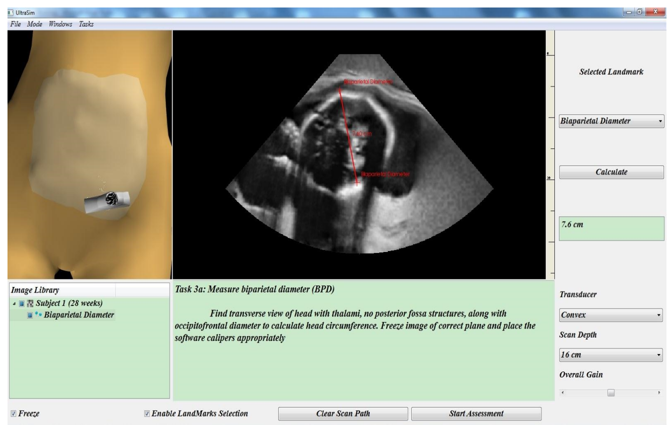


Figure 7-7: Clinical versus simulator images for subject scanned on 21 November 2013

Abdominal circumference measurement:



Biparietal diameter measurement:



Femur length measurement:

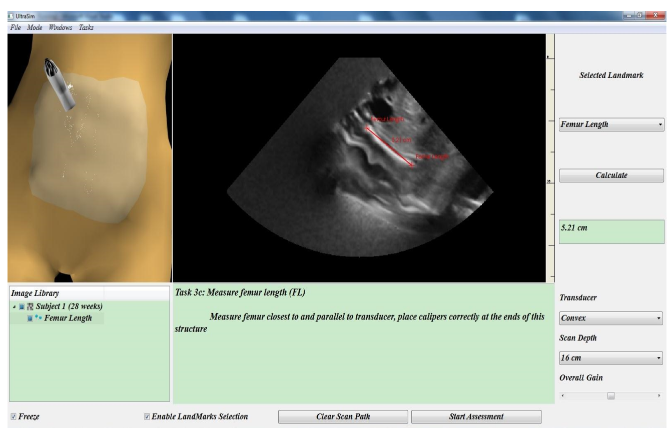


Figure 7-8: Clinical versus simulator images for subject scanned on 9 May 2013

7.4 Conclusions

We have successfully built a library of obstetrics training volumes using the novel stitching approach developed in Chapter 3. The composite volumes were evaluated by medical professionals to determine training value, since our ultrasound expertise is technical in nature. The results of this evaluation were very encouraging and the next step is to measure the effectiveness of the simulator in an obstetrics curriculum. The central question is whether the introduction of our ultrasound simulator can supplement the requirement for live patient scanning, possibly shortening the time necessary to train obstetrics sonographers.

Chapter 8

4D Fetal heart volume: construction from freehand sweep

This chapter will describe how a 4D fetal heart image volume can be constructed from swept 2D ultrasound. A 4D probe, such as the Philips X6-1 xMATRIX Array, would make this task simpler as it can acquire 3D ultrasound volumes of the fetal heart in realtime. In our approach we have to image multiple cardiac cycles in order to collect enough 2D slices to generate a 4D volume. The Philips X6-1 could image the entire 4D volume in 1 cycle with greater accuracy; However, probes such as these aren't widespread yet.

8.1 Introduction

The goal of our obstetrics ultrasound training simulator is to give the trainee a realistic scanning experience at an affordable cost. Anatomy such as the fetal heart, whose motion is highly during scanning, should not be overlooked despite the lack of specific training objectives developed for it. Properly displaying the heart is key to creating a more realistic experience for the sonographer, especially considering how fully fetal development is within the gestational age range of 24 to 36 weeks.

The goal is to capture 4D volumes of the fetal heart, integrate them with the 3D volumes discussed in Chapter 7, and then dynamically reslice the combined dataset during training; thus displaying the appropriate cross-section of the beating heart where the user's scan plane intersects it. Rendering the heart in the simulator is fairly simple once the 3D and 4D volumes are spatially aligned, as one just needs to track time and ensure effective blending along the boundaries between the stationary 3D volume and the dynamic 4D volume. The challenge was in capturing the 4D volumes themselves, using the tools that were available during the clinical scanning phase of this research.

An advanced scanner built on top of the Philips iU22 ultrasound system can utilize the X6-1 4D probe, however one wasn't available to us at the University of Massachusetts Medical School (UMMS). Even if the obstetrics department had such a system there would still likely be issues. The problem is that the raw 4D image data isn't usually accessible to the user because the clinical software on the system is designed for sonographers and not researchers like ourselves. Also, since this system would housed in a hospital and certified for use on patients, the software could not be modified to allow acquisition of raw 4D data. Constructing a 4D volume would have to be done using 2D slices acquired during a freehand sweep over the fetal heart.

8.2 Constructing the 4D volume

This section will describe how a 4D volume of the heart can be constructed using swept 2D ultrasound. All data was acquired at UMMS, from the same patients whose primary scans were used to produced the training volumes described in Chapter 7. The hardware configuration used for the generation of 3D training volumes was also used to collect data for the fetal heart; This system was also introduced in Chapter 7. When acquiring frames to build 3D volumes from swept ultrasound, Stradwin [78] is configured to only store the images if the transducer has moved a certain distance. This results in evenly spaced frames (usually 5 mm spacing) and prevents collection of redundant data if the sonographer pauses during the sweep. While the hardware is unchanged when transitioning from 3D to 4D collection, the video rate parameter is changed to continuously capture 25 frames per second (FPS). This is the only way for Stradwin to collect enough data to construct a 4D volume, as it no longer depends on transducer displacement.

The idea is to acquire ultrasound images, coupled with position/orientation information, at 25 FPS while slowly moving the transducer's image plane across the fetal heart. Since the cardiac cycle is periodic and the heart rate can be measured in beats per minute, we can use the fact that the capture rate is 25 FPS to group the acquired frames into sets, each corresponding to a particular point in the cycle. After this has been done, it is easy to generate a 3D volume for each of these sets using the available techniques described in Chapter 1. The individual 3D volumes, each corresponding to a different point in the cardiac cycle, are concatenated to form the 4D fetal heart dataset. The groups are formed by calculating the offset between frames that belong to the same point in the cycle. This is equivalent to finding the frames per beat and is done using the equation below,

$$\text{Frame offset} = \left(25\frac{\text{frames}}{\text{sec}}\right)\left(60\frac{\text{sec}}{\text{min}}\right)\left(\frac{1}{\beta}\frac{\text{min}}{\text{beat}}\right) \tag{8.1}$$

where β is the fetal heart rate in beats per minute. In one experiment the fetal heart rate was measured by the sonographer to be around 125 beats per minute, making the frame offset equal to 12 using 8.1. Using this offset, Figure 8-1 demonstrates how the individual frames of this sweep can be grouped into sets, which are used to generate the 3D volumes corresponding to unique points in the cycle. Let T_{cycle} be the time since the most recent cardiac cycle started. Because we know that a new cycle starts every 12 frames according to 8.1, every 12th frame corresponds to the same point in the periodic cycle. Even though we are grouping slices together from different cardiac cycles the image data should be the same since heart motion is periodic. This is shown for 3 volumes in Figure 8-1, where every 4th ultrasound frame is illustrated to make the diagram less cumbersome.

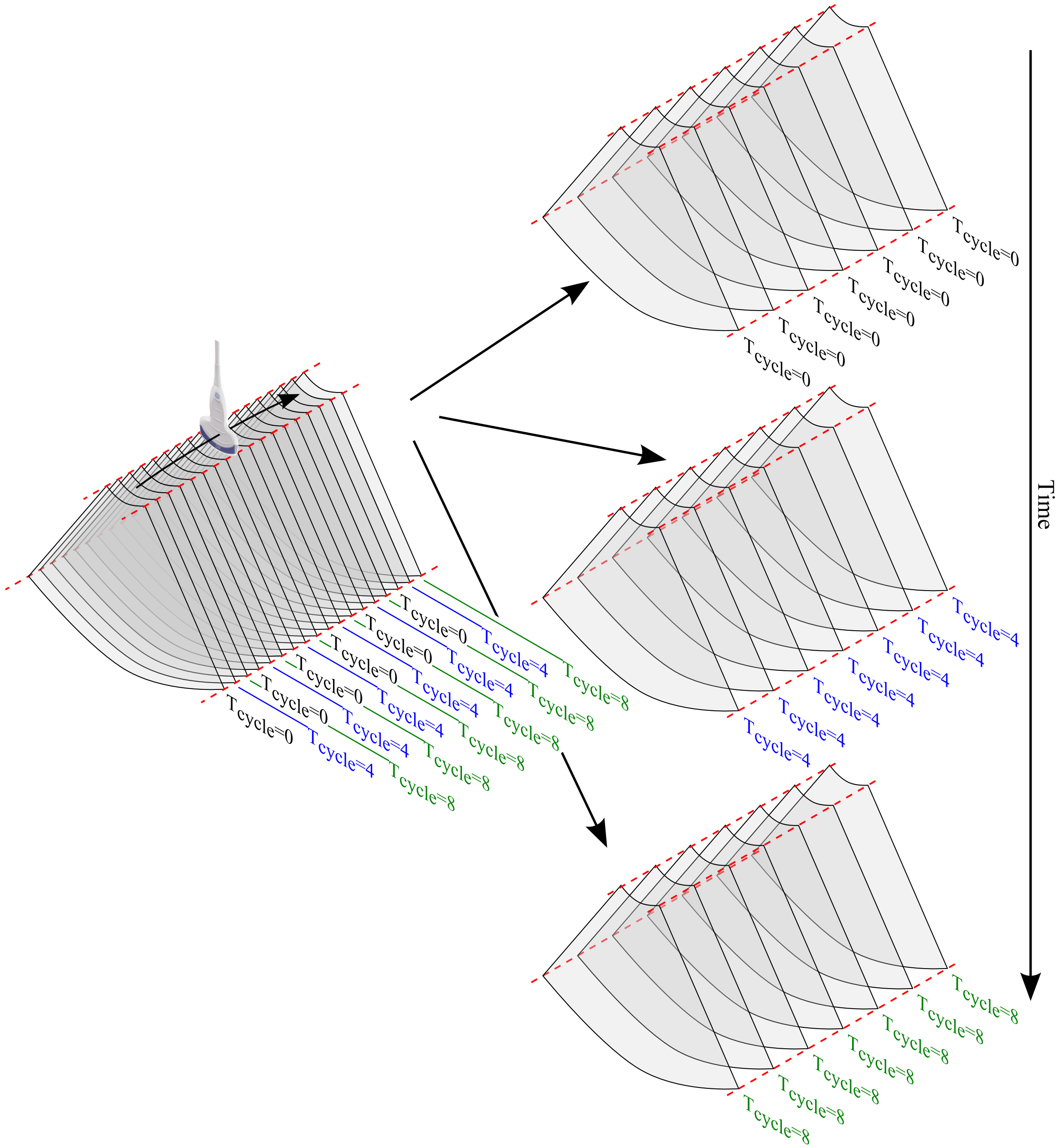


Figure 8-1: Construction of 4D fetal heart volume from 2D sweep. Ultrasound images were acquired at a constant 25 frames/second while the transducer was slowly moved across the fetus's heart.

8.3 Results

The results of this procedure can be seen in Figure 8-2, which shows time slices from a 4D fetal heart volume. This method was very sensitive to fetal movement during the scan and it was difficult for the sonographer to acquire such a small volume; however, the intention was to add realism to the simulator experience and not to capture a completely accurate fetal heart volume. The results were considered adequate for our purposes. In the future the fetal heart should be acquired using a true 4D ultrasound system, which would enable the development of a fetal heart training module.

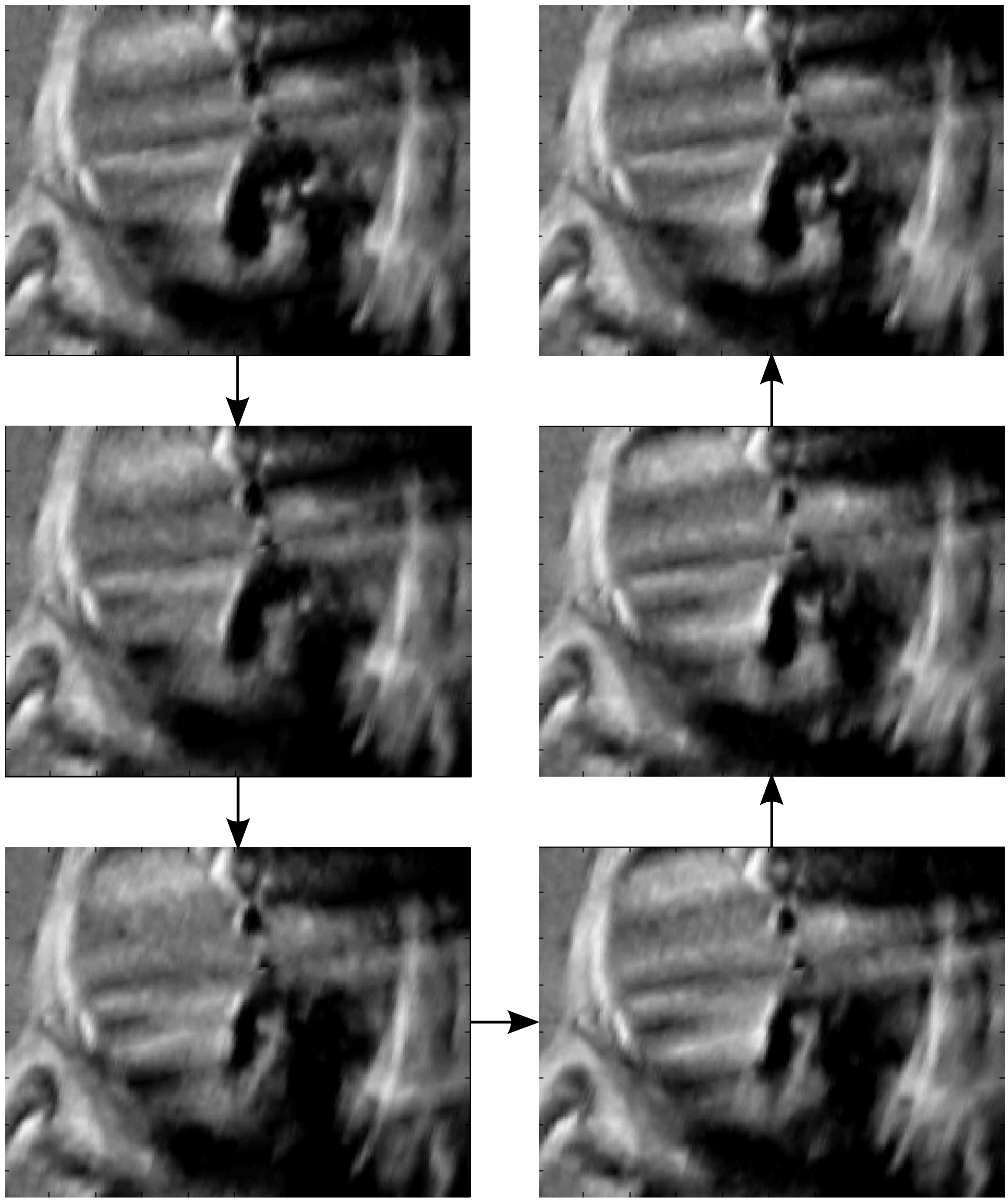


Figure 8-2: Slices from 4D fetal heart show progression through cardiac cycle

Chapter 9

Conclusions

In this work we have presented a novel method to mosaic many (≥ 3) partially overlapping ultrasound volumes collected in a clinical setting. Chapter 1 introduced the simulator design as well as the image processing pipeline which was used in the mosaicing process. Following the introduction, subsequent chapters concentrated on developing the theory for each individual step in the pipeline. In Chapter 2 a rigid registration algorithm, which accounts for shadowing, was developed to aid in the global alignment of image volumes. The bulk of our work was presented in Chapters 3 and 4, where a group-wise non-rigid registration algorithm based on Markov Random Fields was developed and evaluated. Chapter 3 explained the registration theory while Chapter 4 presented experimental results for a number of different scenarios. Chapter 5 discussed the optimization of spline surfaces using particle swarm methods, which was an early attempt at seam selection that was later dropped. A blending algorithm, adapted from Poisson image processing theory, and capable of producing seamless 3D mosaics, was presented in Chapter 6. The complete mosaicing pipeline was tested on image volumes collected from pregnant subjects at the University of Massachusetts Medical School (UMMS). The mosaiced volumes generated with the UMMS clinical data form the foundation of the simulator's obstetrics training modules and are presented in Chapter 8. Finally, the fetal heart was reconstructed from swept 2D ultrasound in order to make the simulator more realistic.

There are a number of directions for future work to take. With the advent of real-time 4D probes, the individual volume creation process could be greatly improved. As described in Chapter 1, we are using swept 3D ultrasound to generate the individual volumes that are stitched together using the process developed in this dissertation. This poses a significant problem in our fetal ultrasound application because the baby often moves during the sweep, thus ruining that particular volume. Our algorithm corrects for inter-sweep movement, which occurs between the partially overlapping volumes, but fetal motion can't be corrected during the sweep, before creation of the individual 3D volumes, when only 2D images are acquired. Due to fetal/maternal movement, many redundant volumes had to be collected from each patient so that a final composite training volume could be produced. Combining a 4D probe with a position sensor and an advanced motion tracking/registration algorithm could potentially correct fetal movement during a sweep, eliminating the need for so many redundant datasets. Since the position of the transducer is known and we are collecting 3D volumes, instead of 2D slices, as the transducer is moved along the abdomen it is possible to track fetal movement from one 3D volume to another and potentially undue it before generating a complete volume from the sweep.

Another direction could be to improve the calculation of the mosaicing function. Currently we use a discrete energy functional, which is optimized using alpha expansion. Due to the nature of the optimization there are no topology constraints, which means that it is possible for two disjoint regions of the final composite volume to be designated the same source ID. This result isn't intuitive, as one would expect all source volumes to be limited to one continuous region each in the final mosaic. Although this type of behavior hasn't posed a problem in our application it would be nice to have a mathematical solution guaranteeing certain topological properties of the mosaicing function. Another interesting direction for this work would be to combine the seam calculation with the registration optimization to produce an iterative joint solution. This might look similar to joint registration/segmentation algorithms.

If access to a true 4D probe and its accompanying data stream can be attained, the fetal heart should be re-imaged to provide a more realistic volume. A fetal heart learning module could then be created based on this newly improved dataset. Since a sonographic examination of the fetal heart needs to be conducted during the second-trimester in order to maximize the detection of heart anomalies [15], this would be a noteworthy addition to the simulator.

Finally, an ultrasound mosaicing toolbox could be created for 3D Slicer [22], which is an open source software package for visualization and medical image computing. This platform allows for easy integration of addition software modules using the Python programming interface provided to the user.

Bibliography

[1] M. J. Ackerman, V. M. Spitzer, A. L. Scherzinger, and D. G. Whitlock. The Visible Human data set: an image resource for anatomical visualization. *Medinfo*, 8 Pt 2:1195–1198, 1995.

[2] Aseem Agarwala, Mira Dontcheva, Maneesh Agrawala, Steven Drucker, Alex Colburn, Brian Curless, David Salesin, and Michael Cohen. Interactive digital photomontage. *ACM Trans. Graph.*, 23(3):294–302, August 2004.

[3] Myronenko Andriy. *Non-rigid Image Registration: Regularization, Algorithms and Applications.* PhD thesis, Oregon Health & Science University, 2010.

[4] Ascension Technology. *3D Guidance trakSTAR Installation and Operation Guide*, 9 2009. Rev. D.

[5] Serdar K. Balci, Polina Golland, and WilliamM Wells III. Non-rigid groupwise registration using b-spline deformation model. In *MICCAI*, volume 4791 of *Lecture Notes in Computer Science*, pages 105–121. Springer, 2007.

[6] R W Barnes, N P Lang, and M F Whiteside. Halstedian technique revisited. innovations in teaching surgical skills. *Annals of Surgery*, 210(1):118–121, 7 1989.

[7] S. Barry Issenberg, William C. Mcgaghie, Emil R. Petrusa, David Lee Gordon, and Ross J. Scalese. Features and uses of high-fidelity medical simulations that lead to effective learning: a beme systematic review. *Medical Teacher*, 27(1):10–28, 2005.

[8] Julian Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 192–236, 1974.

[9] K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert. Consistent groupwise non-rigid registration for atlas construction. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 908–911 Vol. 1, 2004.

[10] Tobias Blum, Andreas Rieger, and Nassir Navab. A review of computer-based simulators for ultrasound training. *Simulation in Healthcare*, 8(2):98–108, April 2013.

[11] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence*, 11(6):567–585, 1989.

[12] Ruma R. Bose, Robina Matyal, Haider J. Warraich, John Summers, Balachundher Subramaniam, John Mitchell, Peter J. Panzica, Sajid Shahul, and Feroze Mahmood. Utility of a transesophageal echocardiographic simulator as a teaching tool. *Journal of Cardiothoracic and Vascular Anesthesia*, 25(2):212–215, 2011.

[13] B. Burger, S. Bettinghausen, M. Radle, and J. Hesser. Real-time gpu-based ultrasound simulation using deformable mesh models. *Medical Imaging, IEEE Transactions on*, 32(3):609–618, March 2013.

[14] Peter J Burt and Edward H Adelson. The laplacian pyramid as a compact image code. *Communications, IEEE Transactions on*, 31(4):532–540, 1983.

[15] JS Carvalho, LD Allan, R Chaoui, JA Copel, GR DeVore, K Hecher, W Lee, H Munoz, D Paladini, B Tutschek, and S Yagel. Isuog practice guidelines (updated): sonographic screening examination of the fetal heart. *Ultrasound in Obstetrics & Gynecology*, 41(3):348–359, 2013.

[16] Samson Cheung. Proof of Hammersley-Clifford Theorem. Technical report, Indian Institute of Technology Bombay, February 2008.

[17] Boaz Cohen and Itshak Dinstein. New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition*, 35(2):455 – 463, 2002.

[18] Sara Damewood, Donald Jeanmonod, and Beth Cadigan. Comparison of a multimedia simulator to a human model for teaching fast exam image interpretation and image acquisition. *Academic Emergency Medicine*, 18(4):413–419, 2011.

[19] Marco A. Montes de Oca, Thomas Stützle, Mauro Birattari, and Marco Dorigo. Frankenstein's pso: a composite particle swarm optimization algorithm. *Trans. Evol. Comp*, 13:1120–1132, October 2009.

[20] Gershon Elber, Tom Grandine, and Myung-Soo Kim. Surface self-intersection computation via algebraic decomposition. *Computer-Aided Design*, 41(12):1060 – 1066, 2009.

[21] F. Essannouni, R. Oulad Haj Thami, D. Aboutajdine, and A. Salam. Simple noncircular correlation method for exhaustive sum square difference matching. *Optical Engineering*, 46(10):107004–107004–4, 2007.

[22] Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Cristophe C. Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona M Fennessy, Milan Sonka, John Buatti, Stephen R Aylward, James V Miller, Steve Pieper, and Ron Kikinis. 3d slicer as an image computing

platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30(9):1323–41, 11 2012.

[23] Roxane Gardner and Daniel B. Raemer. Simulation in obstetrics and gynecology. *Obstetrics and Gynecology Clinics of North America*, 35(1):97–127, 2008. Patient Safety in Obstetrics and Gynecology: Improving Outcomes, Reducing Risks.

[24] Ben Glocker, Nikos Komodakis, Georgios Tziritas, Nassir Navab, and Nikos Paragios. Dense image registration through mrfs and efficient linear programming. *Medical Image Analysis*, 12(6):731–741, 2008.

[25] Ben Glocker, Aristeidis Sotiras, Nikos Komodakis, and Nikos Paragios. Deformable medical image registration: Setting the state of the art with discrete methods*. *Annual Review of Biomedical Engineering*, 13(1):219–244, 2011.

[26] O. Goksel, K. Sapchuk, W.J. Morris, and S.E. Salcudean. Prostate brachytherapy training with simulated ultrasound and fluoroscopy images. *Biomedical Engineering, IEEE Transactions on*, 60(4):1002–1012, April 2013.

[27] Leo Grady. Random walks for image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(11):1768–1783, 2006.

[28] Pierre Hellier, Pierrick Coup, Xavier Morandi, and D. Louis Collins. An automatic geometrical and statistical method to detect acoustic shadows in intraoperative ultrasound brain images. *Medical Image Analysis*, 14(2):195 – 204, 2010.

[29] K Henriksen and E Dayton. Issues in the design of training for quality and safety. *Quality and Safety in Health Care*, 15(1):17–24, 2006.

[30] Richard A Hoppmann and Victor V Rao. An integrated ultrasound curriculum (iusc) for medical students: 4-year experience. *Critical Ultrasound Journal*, 3(1):1–12, 2011.

[31] Po-Wei Hsu, Richard W. Prager, Andrew H. Gee, and Graham M. Treece. Real-time freehand 3d ultrasound calibration. *Ultrasound in Medicine & Biology*, 34(2):239 – 251, 2008.

[32] L. Ibanez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide*. Kitware, Inc. ISBN 1-930934-15-7, http://www.itk.org/ItkSoftwareGuide.pdf, second edition, 2005.

[33] M. Jacob, T. Blu, and M. Unser. Efficient energies and algorithms for parametric snakes. *Image Processing, IEEE Transactions on*, 13(9):1231 –1244, Sept 2004.

[34] Athanasios Karamalis, Wolfgang Wein, Tassilo Klein, and Nassir Navab. Ultrasound confidence maps using random walks. *Medical Image Analysis*, 16(6):1101 – 1112, 2012.

[35] Neil Kirby, Cynthia Chuang, Utako Ueda, and Jean Pouliot. The need for application-based adaptation of deformable image registration. *Medical Physics*, 40(1):–, 2013.

[36] V. Kolmogorov and C. Rother. Minimizing nonsubmodular functions with graph cuts - a review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(7):1274–1279, 2007.

[37] N. Komodakis and G. Tziritas. Image completion using efficient belief propagation via priority scheduling and dynamic pruning. *Image Processing, IEEE Transactions on*, 16(11):2649–2661, Nov 2007.

[38] Dongjin Kwon, KyongJoon Lee, IlDong Yun, and SangUk Lee. Nonrigid image registration using dynamic higher-order mrf model. In David Forsyth, Philip Torr, and Andrew Zisserman, editors, *Computer Vision ECCV 2008*, volume 5302 of *Lecture Notes in Computer Science*, pages 373–386. Springer Berlin Heidelberg, 2008.

[39] Dongjin Kwon, Il Dong Yun, K.M. Pohl, C. Davatzikos, and Sang Uk Lee. Nonrigid volume registration using second-order mrf model. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 708–711, May 2012.

[40] A Lasso, T. Heffter, A Rankin, C. Pinter, T. Ungi, and G. Fichtinger. Plus: open-source toolkit for ultrasound-guided intervention systems. *Biomedical Engineering, IEEE Transactions on*, 2014.

[41] RK Latif, AF Bautista, SB Memon, EA Smith, C Wang, A Wadhwa, MB Carter, and O Akca. Teaching aseptic technique for central venous access under ultrasound guidance: a randomized trial comparing didactic training alone to didactic plus simulation-based training. *Anesthesia and analgesia*, 114(3):626–633, 2011.

[42] Victor Lempitsky, Carsten Rother, Stefan Roth, and Andrew Blake. Fusion moves for markov random field optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1392–1405, 2010.

[43] David C. Levin, Vijay M. Rao, Laurence Parker, and Andrea J. Frangos. Noncardiac point-of-care ultrasound by nonradiologist physicians: How widespread is it? *Journal of the American College of Radiology*, 8(11):772–775, 2011.

[44] X. Li and X. Yao. Cooperatively coevolving particle swarms for large scale optimization. *Evolutionary Computation, IEEE Transactions on*, in press.

[45] H. Maul, A. Scharf, P. Baier, M. Wstemann, H. H. Gnter, G. Gebauer, and C. Sohn. Ultrasound simulators: experience with the sonotrainer and comparative review of other training systems. *Ultrasound in Obstetrics and Gynecology*, 24(5):581–585, 2004.

[46] William C. McGaghie, S. Barry Issenberg, Elaine R. Cohen, Jeffrey H. Barsuk, and Diane B. Wayne. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Academic medicine : journal of the Association of American Medical Colleges*, 86(6):706–711, June 2011.

[47] Matthew Mellor and Michael Brady. Phase mutual information as a similarity measure for registration. *Medical Image Analysis*, 9(4):330–343, 2005.

[48] Alec Mills and Gregory Dudek. Image stitching with dynamic elements. *Image and Vision Computing*, 27(10):1593 – 1602, 2009.

[49] J. Montagnat, H. Delingette, and N. Ayache. A review of deformable surfaces: topology, geometry and deformation. *Image and Vision Computing*, 19(14):1023 – 1040, 2001.

[50] Christopher L. Moore and Joshua A. Copel. Point-of-care ultrasonography. *New England Journal of Medicine*, 364(8):749–757, 2011.

[51] Andriy Myronenko and Xubo B. Song. Intensity-based image registration by minimizing residual complexity. *IEEE Trans. Med. Imaging*, 29(11):1882–1891, 2010.

[52] Dong Ni, Yim-Pan Chui, Yingge Qu, Xuan S. Yang, Jing Qin, Tien-Tsin Wong, Simon S. H. Ho, and Pheng-Ann Heng. Reconstruction of volumetric ultrasound panorama based on improved 3d sift. *Comp. Med. Imag. and Graph.*, 33(7):559–566, 2009.

[53] Vincent Noblet, Christian Heinrich, Fabrice Heitz, and Jean-Paul Armspach. Symmetric nonrigid image registration: Application to average brain templates construction. In Dimitris N. Metaxas, Leon Axel, Gabor Fichtinger, and Gbor Szkely, editors, *MICCAI*, volume 5242 of *Lecture Notes in Computer Science*, pages 897–904. Springer, 2008.

[54] M.N. Omidvar, Xiaodong Li, Zhenyu Yang, and Xin Yao. Cooperative co-evolution for large scale optimization through more frequent random grouping. In *Evolutionary Computation (CEC), 2010 IEEE Congress on*, pages 1–8, July 2010.

[55] J. Orchard. Efficient least squares multimodal registration with a globally exhaustive alignment search. *Image Processing, IEEE Transactions on*, 16(10):2526–2534, Oct 2007.

[56] A. Papoulis and U. Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 4th edition, 2002.

[57] Peder C Pedersen and Daniel Skehan. Personal low-cost ultrasound training system. *Studies in health technology and informatics*, 173:344, 2012.

[58] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. *ACM Trans. Graph.*, 22(3):313–318, July 2003.

[59] K Petrinec, E Savitsky, and C Hein. Patient-specific cases for an ultrasound training simulator. In *Proc. Medicine Meets Virtual Reality 18*, volume 163 of *Studies in health technology and informatics*. IOS Press, 2011.

[60] R. Pittini, D. Oepkes, K. Macrury, R. Reznick, J. Beyene, and R. Windrim. Teaching invasive perinatal procedures: assessment of a high fidelity simulator-based curriculum. *Ultrasound in Obstetrics and Gynecology*, 19(5):478–483, 2002.

[61] David Gerard Platts, Julie Humphries, Darryl John Burstow, Bonita Anderson, Tony Forshaw, and Gregory M. Scalia. The use of computerised simulators for training of transthoracic and transoesophageal echocardiography. the future of echocardiographic training? *Heart, Lung and Circulation*, 21(5):267–274, 2012.

[62] J. P W Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.

[63] Tony C. Poon and Robert N. Rohling. Three-dimensional extended field-of-view ultrasound. *Ultrasound in Medicine & Biology*, 32(3):357 – 369, 2006.

[64] R.W. Prager, R.N. Rohling, A.H. Gee, and L. Berman. Rapid calibration for 3-d freehand ultrasound. *Ultrasound in Medicine & Biology*, 24(6):855 – 869, 1998.

[65] A. Roche, G. Malandain, and N. Ayache. Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology: Special issue on 3D imaging*, 11:71–80, 2000.

[66] A. Roche, X. Pennec, M. Rudolph, D. P. Auer, G. Malandain, S. Ourselin, L. M. Auer, and N. Ayache. Generalized correlation ratio for rigid registration of 3D ultrasound with MR images. In S. Delp, A.M. DiGioia, and B. Jaramaz, editors, *Proc. 3th MICCAI*, volume 1935 of *Lecture Notes in Computer Science*, pages 567–577, Pittsburgh, Oct. 11-14 2000. Springer Verlag.

[67] Andrew B. Ross, Kristen K. DeStigter, Matthew Rielly, Sonia Souza, Gabriel Eli Morey, Melissa Nelson, Eric Z. Silfen, Brian Garra, Alphonsus Matovu, and Michael Grace Kawooya. A low-cost ultrasound program leads to increased antenatal clinic visits and attended deliveries at a health care clinic in rural uganda. *PLoS ONE*, 8, 10 2013.

[68] Daniel Rueckert, Paul Aljabar, RolfA. Heckemann, JosephV. Hajnal, and Alexander Hammers. Diffeomorphic registration using b-splines. In Rasmus Larsen, Mads Nielsen, and Jon Sporring, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2006*, volume 4191 of *Lecture Notes in Computer Science*, pages 702–709. Springer Berlin Heidelberg, 2006.

[69] Richard M. Satava. Identification and reduction of surgical error using simulation. *Minimally Invasive Therapy &; Allied Technologies*, 14(4-5):257–261, 2005.

[70] Y Shen, P Fulgham, X Zhou, D Burke, and R Sweet. The wii transrectal ultrasonography simulator. *24th Engineering and Urology Society Annual Meeting*, page 31, April 2009.

[71] Dan Skehan. Virtual training system for diagnostic ultrasound. Master's thesis, Worcester Polytechnic Institute, 2011.

[72] Ronald W.K. So, Tommy W.H. Tang, and Albert C.S. Chung. Non-rigid image registration of brain magnetic resonance images using graph-cuts. *Pattern Recognition*, 44(1011):2450 – 2467, 2011.

[73] Ole Vegard Solberg, Frank Lindseth, Hans Torp, Richard E. Blake, and Toril A. Nagelhus Hernes. Freehand 3d ultrasound reconstruction algorithmsa review. *Ultrasound in Medicine & Biology*, 33(7):991 – 1009, 2007.

[74] Aristeidis Sotiras, Nikos Komodakis, Ben Glocker, Jean-François Deux, and Nikos Paragios. Graphical models and deformable diffeomorphic population registration using global and local metrics. In *MICCAI (1)*, pages 672–679, 2009.

[75] Aristeidis Sotiras, Yangming Ou, Ben Glocker, Christos Davatzikos, and Nikos Paragios. Simultaneous geometric - iconic registration. In *MICCAI (2)*, pages 676–683, 2010.

[76] Aristeidis Sotiras and Nikos Paragios. Discrete symmetric image registration. In *ISBI*, pages 342–345. IEEE, 2012.

[77] Thomas Szabo. *Diagnostic Ultrasound Imaging : Inside Out*. Academic Press, 2013.

[78] Graham M Treece, Andrew H Gee, Richard W Prager, Charlotte J.C Cash, and Laurence H Berman. High-definition freehand 3-d ultrasound. *Ultrasound in Medicine & Biology*, 29(4):529 – 546, 2003.

[79] T. Ungi, D. Sargent, E. Moult, A Lasso, C. Pinter, R.C. McGraw, and G. Fichtinger. Perk tutor: An open-source training platform for ultrasound-guided needle insertions. *Biomedical Engineering, IEEE Transactions on*, 59(12):3475–3481, Dec 2012.

[80] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing. i. theory. *Signal Processing, IEEE Transactions on*, 41(2):821 –833, Feb 1993.

[81] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing. ii. efficiency design and applications. *Signal Processing, IEEE Transactions on*, 41(2):834 –848, Feb 1993.

[82] Christian Wachinger. *Ultrasound Mosaicing and Motion Modeling - Applications in Medical Image Registration.* PhD thesis, University of Munich, 2011.

[83] Christian Wachinger, Wolfgang Wein, and Nassir Navab. Registration strategies and similarity measures for three-dimensional ultrasound mosaicing. *Academic Radiology*, 15(11):1404 – 1415, 2008.

[84] Guoliang Xu and Chandrajit Bajaj. Regularization of b-spline objects. *Computer Aided Geometric Design*, 28(1):38 – 49, 2011.

[85] Zhenyu Yang, Ke Tang, and Xin Yao. Large scale evolutionary optimization using cooperative coevolution. *Information Sciences*, 178(15):2985 – 2999, 2008. Nature Inspired Problem-Solving.

[86] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.

[87] A. Zomet, A. Levin, S. Peleg, and Y. Weiss. Seamless image stitching by minimizing false edges. *Trans. Img. Proc.*, 15(4):969–977, April 2006.

[88] Kelly H. Zou, Simon K. Warfield, Aditya Bharatha, Clare M.C. Tempany, Michael R. Kaus, Steven J. Haker, William M. Wells III, Ferenc A. Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic Radiology*, 11(2):178 – 189, 2004.