

**DETECTION OF CONGENITAL HEART DEFECTS IN
FETUSES USING FOUR-DIMENSIONAL
ULTRASOUND**

by

Olakunle Bolanle Eso

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

The University of Utah

August 2014

Copyright © Olakunle Bolanle Eso 2014

All Rights Reserved

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of Olakunle Bolanle Eso
has been approved by the following supervisory committee members:

V. John Mathews, Chair 06/12/2014
Date Approved

Neal Patwari, Member 12/18/2013
Date Approved

Tolga Tasdizen, Member 12/18/2013
Date Approved

Pieter Struijk, Member 12/18/2013
Date Approved

Michael Pulchalski, Member 12/18/2013
Date Approved

and by Gianluca Lazzi, Chair/Dean of

the Department/College/School of Electrical and Computer Engineering

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Congenital heart defects are classes of birth defects that affect the structure and function of the heart. These defects are attributed to the abnormal or incomplete development of a fetal heart during the first few weeks following conception. The overall detection rate of congenital heart defects during routine prenatal examination is low. This is attributed to the insufficient number of trained personnel in many local health centers where many cases of congenital heart defects go undetected. This dissertation presents a system to identify congenital heart defects to improve pregnancy outcomes and increase their detection rates. The system was developed and its performance assessed in identifying the presence of ventricular defects (congenital heart defects that affect the size of the ventricles) using four-dimensional fetal echocardiographic images.

The designed system consists of three components: 1) a fetal heart location estimation component, 2) a fetal heart chamber segmentation component, and 3) a detection component that detects congenital heart defects from the segmented chambers. The location estimation component is used to isolate a fetal heart in any four-dimensional fetal echocardiographic image. It uses a hybrid region of interest extraction method that is robust to speckle noise degradation inherent in all ultrasound images. The location estimation method's performance was analyzed on 130 four-dimensional fetal echocardiographic images by comparison with manually identified fetal heart region of interest. The location estimation method showed good agreement with the manually identified standard using four quantitative indexes: Jaccard index, Sørensen-Dice index, Sensitivity index and Specificity index. The average values of these indexes were measured at 80.70%, 89.19%, 91.04%, and 99.17%, respectively.

The fetal heart chamber segmentation component uses velocity vector field estimates computed on frames contained in a four-dimensional image to identify the fetal heart chambers. The velocity vector fields are computed using a histogram-based optical flow technique which is formulated on local image characteristics to reduce the effect of speckle noise and nonuniform echogenicity on the velocity vector field estimates. Features based on the velocity vector field estimates, voxel brightness/intensity values, and voxel Cartesian

coordinate positions were extracted and used with kernel k -means algorithm to identify the individual chambers. The segmentation method's performance was evaluated on 130 images from 31 patients by comparing the segmentation results with manually identified fetal heart chambers. Evaluation was based on the Sørensen-Dice index, the absolute volume difference and the Hausdorff distance, with each resulting in per patient average values of 69.92%, 22.08%, and 2.82 mm, respectively.

The detection component uses the volumes of the identified fetal heart chambers to flag the possible occurrence of hypoplastic left heart syndrome, a type of congenital heart defect. An empirical volume threshold defined on the relative ratio of adjacent fetal heart chamber volumes obtained manually is used in the detection process. The performance of the detection procedure was assessed by comparison with a set of images with confirmed diagnosis of hypoplastic left heart syndrome and a control group of normal fetal hearts. Of the 130 images considered 18 of 20 (90%) fetal hearts were correctly detected as having hypoplastic left heart syndrome and 84 of 110 (76.36%) fetal hearts were correctly detected as normal in the control group. The results show that the detection system performs better than the overall detection rate for congenital heart defect which is reported to be between 30% and 60%.

To Julius Kayode Eso, Cornelia Olabambo Eso, Oladapo Olufemi Eso, Diekolola Onaopemipo Eso and Olamide Folakemi Eso. Thank you for your love and support.

CONTENTS

| | |
|---|------------|
| ABSTRACT | iii |
| LIST OF FIGURES | ix |
| LIST OF TABLES | x |
| ACKNOWLEDGEMENTS | xii |
| CHAPTERS | |
| 1. INTRODUCTION | 1 |
| 1.1 Background | 1 |
| 1.2 Objectives and contributions | 2 |
| 1.3 The fetal heart | 5 |
| 1.4 Congenital heart defects | 7 |
| 1.5 Overview of dissertation | 7 |
| 2. ULTRASOUND PRINCIPLES | 9 |
| 2.1 Ultrasound physics | 9 |
| 2.2 Ultrasound systems | 16 |
| 2.2.1 Ultrasound transducer | 16 |
| 2.2.2 Operation of ultrasound systems | 17 |
| 2.2.3 Ultrasound imaging modes | 20 |
| 2.3 Three- and four-dimensional echocardiography | 21 |
| 2.4 Challenges of ultrasound image analysis | 22 |
| 2.4.1 Spatial resolution | 23 |
| 2.4.2 Temporal resolution | 24 |
| 2.4.3 Artifacts | 25 |
| 2.4.4 Speckle | 27 |
| 2.5 Summary | 29 |
| 3. LITERATURE SURVEY | 30 |
| 3.1 Segmentation in echocardiography | 30 |
| 3.1.1 Edge-based segmentation | 31 |
| 3.1.2 Region-based segmentation | 32 |
| 3.1.3 Deformable models | 33 |
| 3.1.4 Statistical models | 35 |
| 3.1.5 Level sets | 35 |
| 3.1.6 Three- and four-dimensional echocardiographic segmentation | 36 |
| 3.2 Motion estimation in echocardiography | 37 |
| 3.3 Summary | 38 |

| | |
|---|------------|
| 4. DATA | 39 |
| 5. FETAL HEART LOCATION ESTIMATION | 42 |
| 5.1 Abstract | 42 |
| 5.2 Introduction | 42 |
| 5.3 Edge detection | 44 |
| 5.3.1 Choice of f_v | 46 |
| 5.4 Deformable model | 48 |
| 5.4.1 Derivation of active contours in three dimensions | 51 |
| 5.5 Morphological filtering | 53 |
| 5.5.1 Morphological operations | 55 |
| 5.6 Experiments, results, and discussion | 56 |
| 5.6.1 Experiments | 56 |
| 5.6.2 Evaluation metrics | 57 |
| 5.6.3 Results and discussion | 59 |
| 5.7 Conclusion | 64 |
| 6. SEGMENTATION USING VELOCITY VECTOR FIELDS | 65 |
| 6.1 Abstract | 65 |
| 6.2 Introduction | 66 |
| 6.3 Motion estimation | 67 |
| 6.3.1 Horn-Schunck method | 69 |
| 6.3.2 Histogram-based Horn-Schunck method | 71 |
| 6.3.2.1 Bias in Horn-Schunck method | 72 |
| 6.3.2.2 Bias suppression with local statistics | 74 |
| 6.4 Segmentation | 78 |
| 6.4.1 Feature selection | 78 |
| 6.4.2 Choice of kernel function | 79 |
| 6.4.3 Clustering using kernel k -means algorithm | 81 |
| 6.5 Experiments, results, and discussion | 83 |
| 6.5.1 Experiments | 83 |
| 6.5.2 Evaluation metrics | 84 |
| 6.5.3 Results and discussion | 85 |
| 6.6 Conclusion | 92 |
| 7. DETECTION OF HYPOPLASTIC LEFT HEART SYNDROME | 93 |
| 7.1 Introduction | 93 |
| 7.2 Method | 94 |
| 7.3 Results and discussion | 95 |
| 7.4 Conclusion | 99 |
| 8. CONCLUSION AND FUTURE WORK | 100 |
| 8.1 Future work | 102 |
| APPENDICES | |
| A. CHARACTERISTICS OF THE FETAL HEART DATASET | 104 |

| | |
|---|------------|
| B. RESULTS FOR THE LOCALIZATION METHOD | 108 |
| C. RESULTS FOR THE SEGMENTATION METHOD | 112 |
| D. DETECTION RESULTS | 116 |
| REFERENCES | 120 |

LIST OF FIGURES

| | | |
|-----|---|----|
| 1.1 | A block diagram of the congenital heart defect detection system | 3 |
| 1.2 | A normal fetal heart blood circulation system | 6 |
| 2.1 | Longitudinal wave propagation depicting area of high pressure (compression) and low pressure (rarefaction) | 10 |
| 2.2 | Reflection and scattering of ultrasonic waves | 12 |
| 2.3 | Three common types of ultrasound transducers used in medical diagnostics . . | 17 |
| 2.4 | A block diagram depicting the mode of operation of an ultrasound system . . | 18 |
| 5.1 | A block diagram describing the process for isolating a fetal heart from a 4-D echocardiographic image | 44 |
| 5.2 | Approximate intensity regions of a fetal echocardiographic image | 47 |
| 5.3 | A fetal echocardiographic image showing two subimage regions | 49 |
| 5.4 | Comparison of different edge detectors showing their respective edge profile images | 50 |
| 5.5 | Evolution of the 3-D active contours at different iteration n | 54 |
| 5.6 | A venn diagram showing the statistical regions for comparing the location estimation region of interest with the manual region of interest | 58 |
| 5.7 | Box and whisker plot showing the statistics of the Jaccard, the Sørensen-Dice, the Sensitivity, and the Specificity indices for all 254 4-D fetal echocardiographic images | 60 |
| 5.8 | Four chamber view showing the region of interest boundary outline of the location estimation method (red) and the boundary outline of the manual region of interest (green) | 63 |
| 6.1 | A block diagram of the segmentation method | 68 |
| 6.2 | Two synthetic images used to evaluate the performance of Horn-Schunck method and Histogram Horn-Schunck method | 77 |
| 6.3 | Box and whisker plot showing the statistics of the evaluation metrics for all 130 4-D fetal echocardiographic images | 86 |
| 6.4 | Four chamber view showing the boundary outline comparison of the segmentation method with the boundary outline of the manually identified fetal heart chambers | 91 |

LIST OF TABLES

| | | |
|-----|--|----|
| 2.1 | Speed of sound in human tissues and liquids. Data source [55, 56] | 11 |
| 2.2 | The ratio of the reflected wave to the incident wave amplitude and the percentage energy reflected for perpendicular incidence (energy is proportional to amplitude squared). Data source [55, 56] | 13 |
| 4.1 | Breakdown of 4-D fetal echocardiographic dataset according to transducer type (X5-1 and X7-2), acquisition mode (full volume (FV) and live 3-D (L3D)), gestational age, number of frames, and frame time. | 41 |
| 5.1 | A table showing the location estimation results according to fetal gestational ages in weeks. The 25 th percentile, the average, and the 75 th values are shown for each evaluation metric. All values are percentages. | 62 |
| 5.2 | A table showing the location estimation results according to the transducer type used to acquire the images. The 25 th percentile, the average, and the 75 th values are shown for each evaluation metric. All values are percentages | 62 |
| 5.3 | A table showing the location estimation results according to the ultrasound system acquisition mode. The 25 th percentile, the average, and the 75 th values are shown for each evaluation metric. FV refers to full volume acquisition, and L3D refers to live 3-D acquisition. All values are percentages. | 62 |
| 6.1 | Average end point error comparison of the Horn-Schunck method and Histogram Horn-Schunck method. | 78 |
| 6.2 | Cross validation of different kernel function used in selecting the kernel with best classification error performance. | 81 |
| 6.3 | A table showing the average Sørensen-Dice index, absolute volume difference and Hausdorff distance values on a per patient basis | 89 |
| 6.4 | A table showing the fetal heart chamber segmentation results according to fetal gestational ages in weeks. The 25 th percentile, the average, and the 75 th values are shown for each evaluation metric. | 90 |
| 6.5 | A table showing the fetal heart chamber segmentation results according to the transducer type used to acquire the images. The 25 th percentile, the average, and the 75 th values are shown for each evaluation metric. | 90 |
| 6.6 | A table showing the fetal heart chamber segmentation results according to the ultrasound system acquisition mode. The 25 th percentile, the average, and the 75 th values are shown for each evaluation metric. FV refers to full volume acquisition, and L3D refers to live 3-D acquisition | 90 |
| 7.1 | A table showing the detection results for the 130 fetal echocardiographic images grouped according to patient number. | 96 |

| | | |
|-----|--|-----|
| 7.2 | A table showing the detection results according to fetal gestational ages. | 97 |
| 7.3 | A table showing the detection results according to transducer type. | 97 |
| 7.4 | A table showing the detection results according to acquisition mode. | 97 |
| 7.5 | A table showing the sensitivity and specificity analysis of the detection results for all 130 fetal echocardiographic images. | 98 |
| A.1 | A table showing the various characteristics of each image used in experiments. These include the transducer type, the acquisition mode, the gestational ages of the fetus, number of frames in each image, the frame time of each frame in each image, and the type of congenital heart defect. L3D represent live 3-D acquisition mode and FV represents full volume acquisition mode | 104 |
| B.1 | A table showing the region of interest evaluation results by applying Jaccard, Sørensen-Dice, Sensitivity, and Specificity metrics on each of the 130 4-D fetal echocardiographic images. | 108 |
| C.1 | A table showing the segmentation evaluation results by applying Sørensen-Dice index, absolute volume difference, and Hausdorff distance metrics on each of the 130 4-D fetal echocardiographic images. | 112 |
| D.1 | A table showing the results of applying an empirical threshold (0.85) to the chamber volume ratios of 130 4-D fetal echocardiographic images for detecting/flagging hypoplastic left heart syndrome. Chamber ratio values less than 0.85 are classified as hypoplastic left heart syndrome, while chamber ratios greater than or equal to 0.85 are classified as normal | 116 |

ACKNOWLEDGEMENTS

I thank God for his grace, mercies and favor throughout my academic journey. It would have been impossible to have completed this body of work without my faith in Him.

A big thank you to my advisor, Dr. V. John Mathews, for the invaluable guidance he provided. You were indeed a reason I enjoyed coming to work and staying motivated day to day. I appreciate all your advise and suggestions which contributed to my improvement and growth as a researcher. I appreciate your patience with me particularly when the research seemed to be moving at a snail pace. I am grateful for the freedom you gave me in defining the path of my research.

I thank Dr. Mike Puchalski for teaching me about fetal echocardiography and congenital heart defects. I am grateful for the clinical data you constantly provided during the course of the research. The time I spent analyzing ultrasound images with you provided valuable perspective of the importance of the work I was doing.

I thank Dr. Pieter Struijk for the many one-on-one meetings, Skype sessions and back and forth emails. Even though you live in the Netherlands, you constantly took time out to appraise my research and give suggestions for improvements.

I thank Dr. Neal Patwari and Dr. Tolga Tasdizen for the advice they provided during the course of my program. The suggestions and insights you both provided gave me a basis for many of the ideas I developed in this work.

I would like to thank the Signal Processing Group members, Shreyas Payal, Daimei Zhu, Jingru Zhou, Sabita Nahata, and Ahmad Zoubi for the countless discussions we had among each other concerning research problems. Our constant interaction and weekly meetings provided many ideas that proved valuable in the completion of this work.

I thank my girlfriend, Julie Taylor. You were instrumental in motivating me whenever I was stuck and frustrated with my research. Your unconditional love, prayers and words of encouragement are much appreciated. I love you.

I thank Jay, Linda, Jason, and Jaelyn Taylor for their support particularly during the final stages of my doctoral program. The conversations, advice, and guidance on how to

navigate the professional world will always be cherished. I am grateful that each of you treated me as a member of your family.

Finally, I thank my parents, Julius Kayode Eso and Cornelia Olabambo Eso, my brother, Oladapo Eso and his family, and my sisters, Diekolola and Olamide Eso for their continued prayers and unconditional love given to me throughout my educational journey. The countless phone calls about the progress of my research kept me focused. The words of encouragement and financial support proved essential in my completion of this work. I will not be where I am today without each of you. I love you all so much.

CHAPTER 1

INTRODUCTION

1.1 Background

The use of engineering in medicine has been around since the beginning of civilization. From the discovery of a 3,000 year old mummy with a wooden prosthetic to the inadvertent discovery of the stethoscope by Rene Laennec, a French physician, engineering has been an important component in the advancement of medicine. Pacemakers, dialysis machines, diagnostic equipment, imaging technologies of every kind, artificial organs, implants and advanced prosthetics are all evidence of how engineering has contributed to medicine. Moreover, diagnostic imaging has enhanced our knowledge of the human anatomy, and the etiology of many diseases. Magnetic resonance imaging (MRI), X-Ray, computed tomography (CT) and ultrasound have all changed the way doctors interpret and diagnose many diseases. Just as important are the analysis tools that accompany these imaging modalities, helping doctors to correctly interpret and analyzes the images obtained from them. Ultrasound, which is one of the most popular medical imaging technology, is at the forefront of diagnostic imaging in obstetrics. This is largely due to its noninvasive nature, its low cost, and its relative safety in imaging fetuses. Traditionally, two-dimensional (2-D) ultrasound accounted for the majority of fetal scans. However, with advances in ultrasound matrix transducer technology, three-dimensional (3-D) and four-dimensional (4-D) ultrasound imaging are becoming common. With the increase in popularity comes the added need for analysis tools that can perform 3-D and 4-D measurements and quantifications of anatomical structures and functions. In particular, the diagnosis of congenital heart defects (CHD) can benefit from the added dimensionality 3-D and 4-D ultrasound offers to provide more accurate characterization of these defects.

A congenital heart defect is a disease, present at birth, affecting the structure and function of the heart and major vessels. It is the most common type of birth defect occurring in 8 per 1000 births in the United States [70]. It is the leading cause of all infant deaths in the United States where twice as many children die from CHD each year than from all forms

of childhood cancer combined [61]. Moreover, the diagnosis, treatment, and management of the disease cost about \$2.6 billion annually [70].

A fetal heart begins to develop shortly after conception during which structural defects can occur. These defects can involve the walls and valves of the heart as well as the arteries and veins near the heart. Prenatal diagnosis of CHD is important for improved pregnancy outcome through timely changes in obstetric and neonatal care management. It also facilitates parental counseling regarding the diagnosis, prognosis, management and treatment of CHD.

The introduction of prenatal diagnostic ultrasound 50 years ago greatly enhanced the early diagnosis of many common fetal defects by pediatric cardiologists. As ultrasound technology progressed over the years, the images derived from them have improved both in spatial and temporal resolutions. The further development of 3-D ultrasound has also provided additional quantification for fetal anatomy. Despite these advances, the overall detection rate of CHD remain anemic; between 30% and 60% of CHD are undetected until after birth [172]. While tertiary institutions have high detection rates which can be ascribed to the availability of experienced care providers, the same is not true in many local prenatal care centers; a main reason for the low rates. Therefore, it is beneficial to have a system that can aid with diagnosis, bridging the detection gap between tertiary institutions and local prenatal centers.

The norm in the industry still relies on manual manipulation of echocardiographic images (ultrasound images of the cardiovascular system) to make inferences about the presence of CHD. While operator error could contribute to missed detection, it is also true that a computer aided system could also miss some subtleties that could be important in a potential diagnosis. As a consequence any system developed should be used as a flagging mechanism where specific measurements or metrics are considered to flag CHD. The flagged cases are then passed on for further analysis by a pediatric cardiologist. This has the potential of reducing human error, and increasing detection rate while preserving the integrity of the diagnosis. In addition, studies have shown that low-risk populations for CHD contribute substantially to missed detections [172]. Therefore a computer aided system could be deployed during the routine second trimester prenatal visit to concurrently test for CHD.

1.2 Objectives and contributions

The overall contribution of this dissertation is the development of a system that can detect the presence of CHD using 4-D fetal echocardiographic images. However, in achieving

this, some methods were developed to facilitate the detection process. Figure 1.1 shows a block diagram of the detection system with a fetal heart location estimation block, a fetal heart chamber segmentation block, and a detection block. The location estimation block is a preprocessing step used to isolate the fetal heart in any given 4-D fetal echocardiographic image. The method separates the fetal heart from other structures that may be present in an image. It is a hybrid region of interest approach where a combination of an edge detector and a deformable model is used to obtain the region of an image containing a fetal heart. The segmentation block uses velocity vector field estimates obtained from frames of a 4-D image to identify the individual chambers of a fetal heart. The velocity vector field estimates are computed using a histogram-based optical flow method by Tenbrinck *et al.* [192]. It uses a histogram constancy criteria to reduce the effect of speckle noise on motion estimates. Features based on these velocity vector field estimates and other image properties are used in a kernel based unsupervised learning framework to identify the fetal heart chambers. The kernel based unsupervised learning method transforms the features into a nonlinear kernel subspace where they can be grouped in a way that correspond to the individual fetal heart chambers. The novelty of the work presented in this dissertation include the development of a 4-D region of interest method for isolating a fetal heart in an image, a 4-D segmentation method that can be used to identify fetal heart chambers using velocity vector fields, and a system that can detect CHD using 4-D fetal echocardiographic images.

Current methods used for identifying CHD (prenatally) are mostly based on manual analysis of images by experts (cardiologists and sonographers). Although automatic image analysis tools used for analyzing the adult heart is an area of ongoing active research, little attention had been given to automatic analysis of fetal echocardiography until recently. Published work on fetal echocardiographic segmentation are limited. One of the few comprehensive work done on this is by Dindoyal *et al.* [47, 48, 49, 50], where an algorithm called Mumford Shah Sarti Collision Detection (MSSCD) is used to segment a fetal heart to determine blood volume measurements in its cardiac chambers. Other authors like

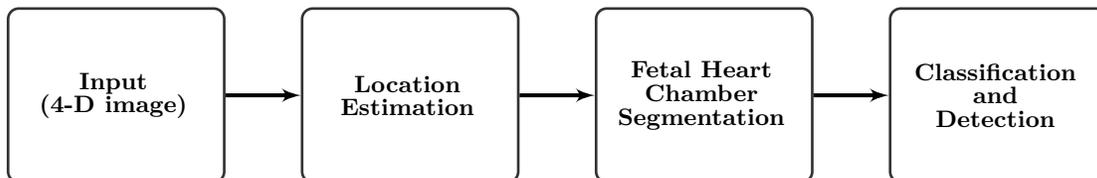


Figure 1.1: A block diagram of the congenital heart defect detection system.

Piccoli *et al.* [148] used neural networks to segment 2-D fetal echocardiographic images. Also, Siqueira *et al.* [179] used self-organizing maps and k -means clustering to perform segmentation. Active contour models have also been used by Lassige *et al.* [103] and Dindoyal *et al.* [48] to achieve segmentation. In all these methods, single (static) 2-D or 3-D images are used which means the results are dependent on the image quality and ability to resolve boundaries of the fetal heart chambers in any given image. This is a major drawback of the methods described in literature. Four-dimensional imaging is gaining foothold these days in cardiology, therefore a viable extension is to design analysis tools that take advantage of the added dimensionality to perform segmentation. The work described in this dissertation is one of the first to exploit 4-D fetal echocardiographic imaging for this purpose to the best of the author's knowledge. An advantage of 4-D segmentation is that in regions where boundaries cannot be resolved with a gradient-based segmentation algorithms like MSSCD or active contour model, the motion characteristics (obtained from the image sequence) can be used to fill in the missing boundary information thereby producing more robust segmentation results. Four-dimensional ultrasound technology and its advantages are described in Chapter 2. In summary, the contributions of this dissertation to the field of 4-D fetal echocardiographic image analysis and its use in detection of CHD are:

- A hybrid region of interest method for isolating fetal hearts in a 4-D ultrasound image is introduced in this dissertation. The method is robust to speckle noise and it is be used for preprocessing the image sequence to help eliminate unnecessary processing of regions of an image not used during detection.
- An automatic 4-D segmentation method that uses velocity vector field estimates of the fetal heart structures. Features extracted from the velocity vector field estimates, voxel brightness/intensity values, and voxel spatial positions are used in an unsupervised kernel learning framework to segment the individual chambers of a fetal heart.
- A system that uses the segmented fetal heart chambers to flag the presence of hypoplastic left heart syndrome in fetuses at +20 weeks gestational age using a binary volume classifier.

To motivate the discussions in this dissertation it is important to understand the structural and functional characteristics of a normal fetal heart and how deviations from this results in CHD. To this end, Section 1.3 describes the fetal heart blood circulation system and Section 1.4 describes hypoplastic left heart syndrome which is the CHD considered in this dissertation.

1.3 The fetal heart

The prenatal (fetal) heart has a slightly different structure and also functions differently than the postnatal heart. Structurally, a postnatal heart has four chambers: left and right atria, left and right ventricles. The left atrium and left ventricle are separated by the mitral valve (left atrioventricular valve), while the right atrium (RA) and right ventricle (RV) are separated by the tricuspid valve (right atrioventricular valve). The aortic valve lies between the left ventricle and aorta while the pulmonary valve lies between the right ventricle and the pulmonary artery. The right atrium receives deoxygenated (oxygen free) blood from the body and pumps it via the right ventricle into the lungs through the pulmonary artery. The tricuspid valve and pulmonary valve controls the direction of this flow through its periodic opening and closing. On the contrary, the left atrium receives oxygenated (oxygen rich) blood from the lungs and pumps it via the left ventricle in to the body through the aorta. The mitral valve and aortic valves control the direction of flow.

Figure 1.2 shows the blood circulation of a prenatal fetal heart. During a pregnancy, the placenta and umbilical cords acts as a conduit for sending oxygenated blood to the fetus and accepting deoxygenated blood from the fetus. Oxygenated blood from the placenta via the umbilical vein bypasses the fetal liver through a temporary opening called the ductus venosus. This allows most of the oxygenated blood to go directly to the fetal heart. The oxygenated blood enters the fetal heart through the foramen ovale, an interatrial opening (between the right atrium and left atrium). This opening allows blood to pass from the right atrium to the left atrium and then through the left ventricle and out of the aorta. In the opposite direction, deoxygenated blood also enters the right atrium and gets passed to the right ventricle. Most of the blood from the right ventricle bypasses the lungs through another opening called the ductus arteriosus, a short, muscular vessel that connects the pulmonary trunk to the aorta. Only enough blood reaches the fetal lungs to maintain the developing lung tissue. The blood through the ductus arteriosus is sent back to the placenta to pick up oxygen and the cycle is repeated. The main difference between the prenatal heart and the postnatal heart is that the lungs are not directly involved in circulation for the prenatal heart. Oxygen circulation between the placenta and the fetus is thus made possible by the ductus venosus, the foramen ovale and the ductus arteriosus. The ductus venosus, foramen ovale and ductus arteriosus begin to close right after birth as the baby takes over its own circulation. After birth the prenatal heart evolves towards the postnatal heart.

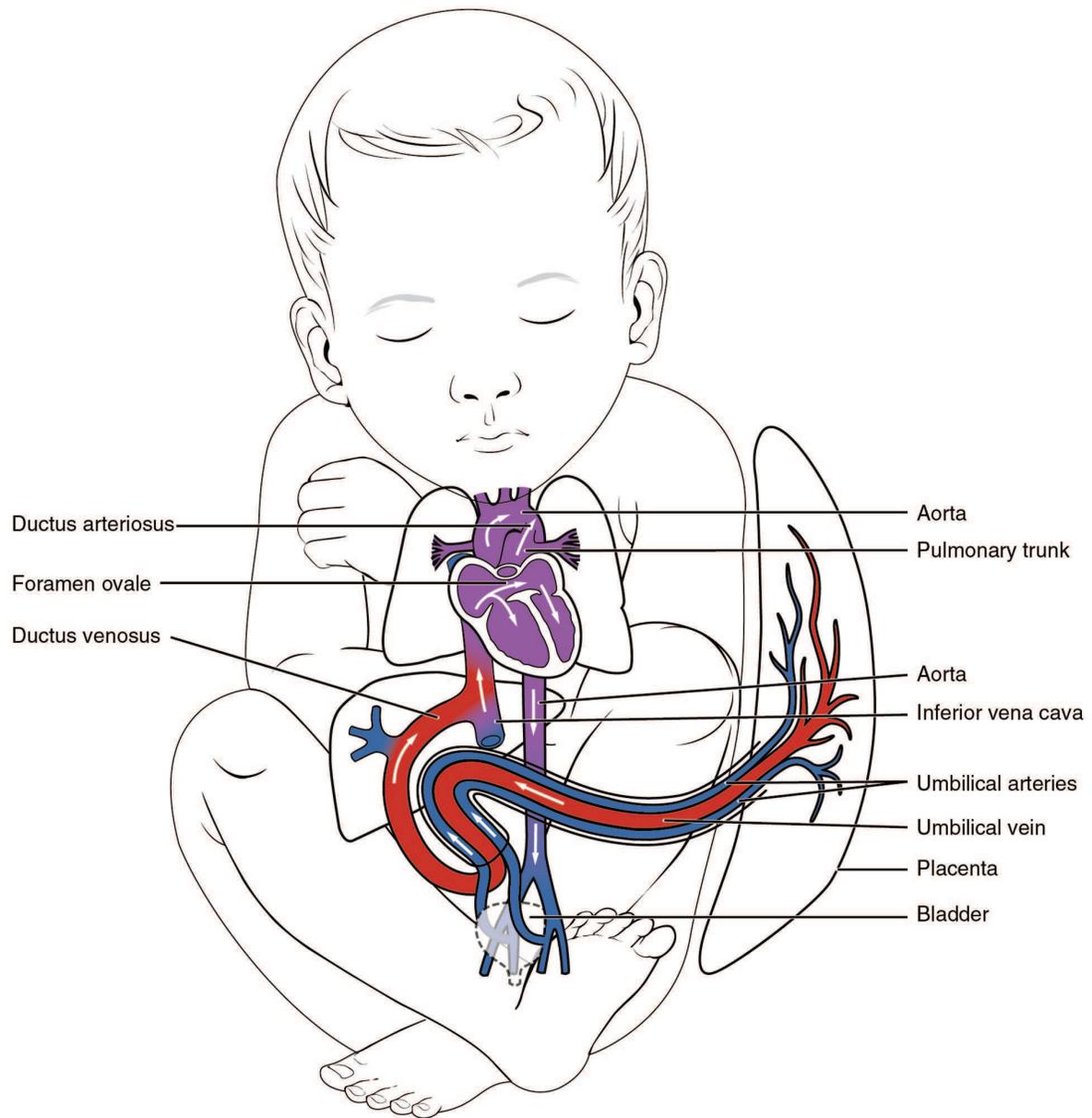


Figure 1.2: A normal fetal heart blood circulation system. The foramen ovale in the interatrial septum allows blood to flow from the right atrium to the left atrium. The ductus arteriosus is a temporary vessel, connecting the aorta to the pulmonary trunk. The ductus venosus links the umbilical vein to the inferior vena cava largely through the liver [33]. (This image is unedited and is used with permission under Creative Commons Attribution 3.0 License (<http://creativecommons.org/licenses/by/3.0/legalcode>)).

1.4 Congenital heart defects

Congenital heart defect is a broad term used for a myriad of structural or functional anomalies affecting the heart. There are over 40 types of CHD, each having different identification marker(s) for its diagnosis. Because of their variability, it is difficult to have a singular system that can diagnose all types of CHD. For the purposes of this dissertation only hypoplastic left heart syndrome (HLHS) is considered although the detection system can be applied to hypoplastic right heart syndrome (HRHS) or any other CHD characterized by underdeveloped chambers.

Hypoplastic left heart syndrome accounts for 4-8% of all cases of CHD. It is a combination of several abnormalities of the heart and great blood vessels. It affects the mitral valve, left ventricle, aortic valve and aorta. It is the underdevelopment of the left side of the heart, particularly the left ventricle and the ascending aorta. In addition, the mitral and the aortic valves are usually atretic (*i.e.*, small or absent). In newborns with hypoplastic left heart syndrome, the left side of the heart cannot pump oxygenated blood through the aorta to the body effectively, so the foramen ovale and the ductus arteriosus are used to bypass the left side of the heart. The right side then becomes the mechanism for pumping blood both to the body and to the lungs. However, after birth, as the foramen ovale and ductus arteriosus begin to close, this becomes increasingly difficult and the fetus receives less and less oxygen. This is usually referred to as the blue baby syndrome. If there is no surgical intervention after birth, the baby will die. To prevent this outcome, it is essential that hypoplastic left heart syndrome is diagnosed before birth.

1.5 Overview of dissertation

This dissertation is organized as follows: Chapter 2 describes the image formation model(s) behind ultrasound imaging. Ultrasound systems and 4-D ultrasound imaging are also briefly discussed. In addition, challenges encountered in the analysis of ultrasound images are highlighted. Chapter 3 is a literature review of segmentation and motion estimation techniques as they relate to echocardiography. Chapter 4 gives a description of the data used in all the evaluation experiments of this dissertation. Chapter 5 begins the description of the detection system with a discussion of the location estimation component. Chapter 6 discusses the segmentation component which includes a description of the mechanism used to identify individual fetal heart chambers based on velocity vector field estimates computed from frames of a 4-D fetal echocardiographic image. Chapter 7 introduces the detection component of the detection system. A binary classifier that is able to flag the presence of HLHS using the relative chamber volume ratios is described. Experiments, results and

the analysis of the overall detection system is also provided. Chapter 8 concludes this dissertation by summarizing the main contribution made by this work, the drawback of the detection system's components, and possible extension of the methods discussed throughout this dissertation. Appendix A, B, C, and D are included after Chapter 8 and they contain breakdown of the detection system's evaluation results.

CHAPTER 2

ULTRASOUND PRINCIPLES

Ultrasound imaging is a noninvasive, safe and inexpensive imaging technique. It is the most common form of imaging used in obstetrics. Traditionally, 2-D ultrasound imaging was the standard in prenatal care. However, advances in recent years has contributed to the increasing popularity of 3-D and 4-D ultrasound imaging. Despite the popularity of ultrasound imaging, there are some challenges it poses to the image analysis community. These includes but are not limited to multiplicative noise and low spatial resolution. Understanding these challenges will require an understanding of the principles behind ultrasound imaging modality. Therefore, the aim of this chapter is the following: 1) provide a background of the physics behind ultrasound image formation, 2) describe the basic system used in ultrasound imaging, 3) introduce 3-D and 4-D ultrasound imaging, and 4) introduce the challenges behind ultrasound image analysis, relating it to the image formation principles.

2.1 Ultrasound physics

Ultrasound images are derived from ultrasonic waves. They are acoustic waves, like those humans perceive, however, they operate at frequencies beyond the human hearing range (> 20 KHz). They are longitudinal waves which cause particles to oscillate back and forth producing a series of compressions and rarefactions (Figure 2.1). They are completely characterized by: 1) their speed in a medium, 2) their wavelength, and 3) their frequency. The speed is defined as how fast a wave travels through a medium. Table 2.1 gives typical speeds of ultrasonic waves in some media. These speeds are derived from (2.1) where c is the speed of sound in the medium, K is the coefficient of stiffness of the medium and ρ is the density of the medium.

$$c = \sqrt{\frac{K}{\rho}} \quad (2.1)$$

The frequency measures the rate of oscillation or vibration of the ultrasonic waves. In medical imaging applications, the frequency f of ultrasonic waves is between 1 MHz and

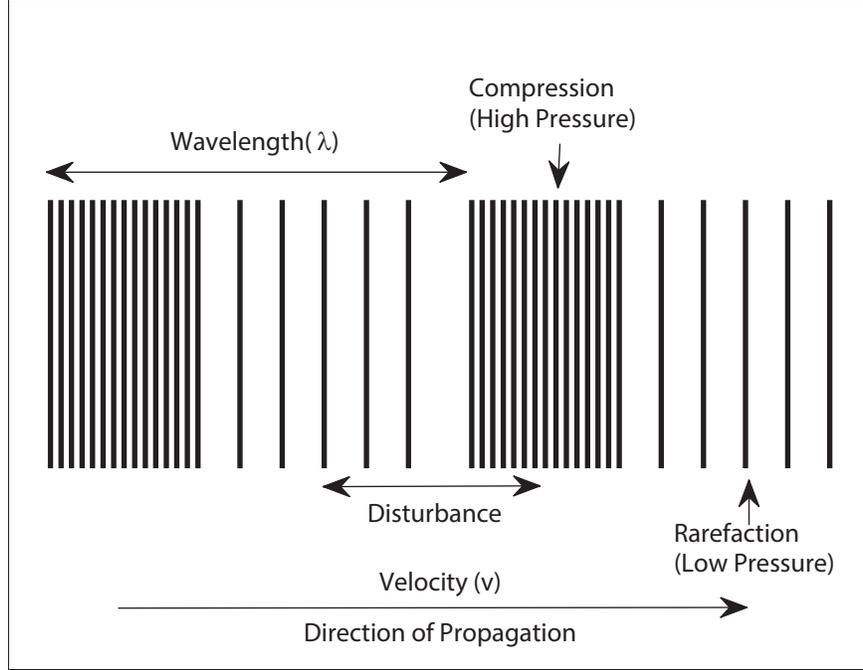


Figure 2.1: Longitudinal wave propagation depicting area of high pressure (compression) and low pressure (rarefaction).

20 MHz. This gives a corresponding wavelength λ between 0.1 mm and 1.0 mm from the relationship in (2.2), if $c = 1540 \text{ ms}^{-1}$ (see speed of sound in average tissue in Table 2.1).

$$c = f\lambda \quad (2.2)$$

The propagation of ultrasonic waves in a medium can be described by the 3-D acoustic wave equation [58]. In its simplest form it describes the evolution of the acoustic pressure p as a function of position \mathbf{r} and time t .

$$\nabla^2 p(\mathbf{r}, t) = \frac{1}{c^2} \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2} \quad (2.3)$$

where ∇^2 is the Laplacian operator (second order differential operator).

The general solution to (2.3) is given by:

$$p(\mathbf{r}, t, k) = \Re[Ae^{i\mathbf{k}\mathbf{r}} e^{i\omega t}] \quad (2.4)$$

where A is the amplitude, $\omega = 2\pi f$ is the angular frequency, $k = \omega/c$ is the wave number, and \Re is the real part of a complex quantity.

Table 2.1: Speed of sound in human tissues and liquids. Data source [55, 56]

| Materials | Speed (ms ⁻¹) |
|----------------|---------------------------|
| Brain | 1540 |
| Blood | 1570 |
| Muscle | 1580 |
| Skin | 1600 |
| Liver | 1578 |
| Kidney | 1560 |
| Amniotic fluid | 1534 |
| Fat | 1450 |
| Average tissue | 1540 |
| Water | 1480 |
| Bone | 3500 |
| Air | 330 |

Ultrasound images are formed by using reflections of ultrasonic waves off a target such as tissue boundaries, organs, bones and other irregularities in the human body. The relative positions of these reflections are used as a map to create a brightness image corresponding to various structures in the path of the wave. This brightness image is called a B-mode scan/image with the B in B-mode standing for brightness. Reflections occurring at tissue boundaries are attributed to the changes in the acoustic impedance of the medium. Acoustic impedance z of a medium is the measure of the response of its particles to an acoustic wave at a given pressure p . Mathematically, it is the ratio of the acoustic wave pressure p to the particle velocity v . It can also be expressed in terms of the density of the medium ρ and the speed of sound in the medium c .

$$z = \frac{p}{v} = \rho c \quad (2.5)$$

In medical diagnostics, the source of the ultrasound beams used for imaging is a transducer. It is a device that converts electrical pulses (energy) to ultrasonic pulses (energy) and vice versa. It can be described as performing the job of an actuator, producing ultrasonic waves, and a sensor, detecting reflected waves. The active element used in many transducers is piezoelectric ceramics. The wave emanating from the transducer travels towards the target (e.g., heart) during which the following processes occurs: reflection, scattering, refraction, absorption and attenuation [84, 191].

Reflection (sometimes called specular reflection, Figure 2.2) occurs at large interfaces such as boundaries between organs. The structures are usually several wavelengths larger

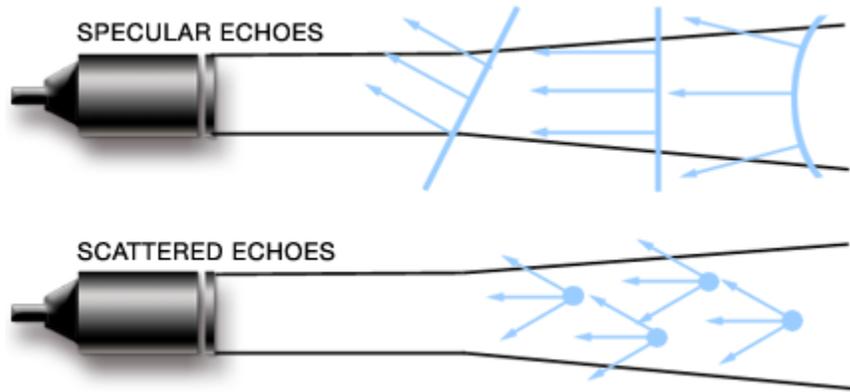


Figure 2.2: Reflection and scattering of ultrasonic waves. Specular reflection occurs when the incoming ultrasound echo is incident on a structure with size several wavelengths larger. Scattering occurs when the incoming ultrasound echo is incident on a structure with size several wavelengths smaller.

than that of the incident wave. Reflection can also be described as occurring when the incident wave encounters a change in acoustic impedance. Since the ultrasonic waves that form the ultrasound images are primarily from reflections, it is important to know how to measure the characteristics of the reflected waves. Let us assume medium 1 has an acoustic impedance z_1 and medium 2 has an acoustic impedance z_2 . From the conservation laws, the pressure and speed must be continuous and conserved at the boundary between these two media, that is the sum of the incident wave pressure p_i and the reflected wave pressure p_r must be equal to the transmitted wave pressure p_t . Likewise, the sum of the incident wave velocity v_i and the reflected wave velocity v_r must be equal to the transmitted wave velocity v_t .

$$\begin{aligned} p_i + p_r &= p_t, \\ v_i + v_r &= v_t \end{aligned} \tag{2.6}$$

The conservation of pressure and velocity can be used to determine how much of the incident wave power is contained in the reflected wave. This gives what is called the amplitude reflection coefficient and it is given as,

$$R_A = \frac{p_r}{p_i} = \frac{z_2 - z_1}{z_1 + z_2} \tag{2.7}$$

Some R_A values for selected human tissue pairs are shown in Table 2.2. A quick look

Table 2.2: The ratio of the reflected wave to the incident wave amplitude and the percentage energy reflected for perpendicular incidence (energy is proportional to amplitude squared). Data source [55, 56]

| Reflecting interface | Ratio of reflected to incident wave amplitude (R_A) | Percentage of energy reflected |
|-----------------------------|---|---------------------------------------|
| Fat/muscle | 0.10 | 1.08 |
| Fat/kidney | 0.08 | 0.64 |
| Muscle/blood | 0.03 | 0.07 |
| Bone/fat | 0.70 | 48.91 |
| Bone/muscle | 0.64 | 41.23 |
| Soft tissue/water | 0.05 | 0.23 |
| Soft tissue/air | 0.99 | 99.9 |
| Soft tissue/castor oil | 0.07 | 0.43 |
| Liver/fat | 0.11 | 1.21 |
| Liver/bone | 0.59 | 34.81 |
| Liver/kidney | 0.006 | 0 |
| Blood/kidney | 0.009 | 0 |

at the table gives the following inferences about the behavior of ultrasonic waves as they travel through a human body:

- There is almost no reflection between two soft tissue interfaces (liver/kidney). This means that almost all the energy (99.4%) is transmitted to produce further reflections at deeper depths in the ultrasonic wave path.
- There are some reflections that occur between a soft tissue and fat interface (liver/fat). This implies a significant proportion of the incident energy (89%) is transmitted to deeper depths in the ultrasonic wave path.
- A soft tissue and bone interface reflects over half of the energy (59%). This makes it difficult to get echoes beyond bony structures such as ribs.
- No useful echoes can be obtained beyond the interface of soft tissue and air because of the high reflection at their interface (99%). This means that ultrasound gives poor details about structures with air pockets such as the lungs. This is also why a coupling gel is used between the surface of a transducer and a patient's skin, so air is excluded from the ultrasonic wave path.

Unlike reflections which occur at structures much larger than the wavelength of the ultrasound wave, scattering or nonspecular reflection (Figure 2.2) occurs when the incoming

ultrasonic wave strikes a discontinuity with dimensions equal to or less than a wavelength. It was mentioned earlier that the wavelength of typical ultrasound beams are about 1 mm and within different organs in the human body, there are many structures with dimensions smaller than 1 mm. Scattering from these small structures provide important texture information about the interior regions of these organs. Although the scattered waves have much weaker power than those caused by reflection, the sensitivity of modern transducers make it possible to utilize information from scatterers for imaging. The total power of the scattered wave is proportional to the size of the structure and the wavelength of the wave. Scattering thus increases as the frequency of the ultrasonic beam is increased. The speckle noise pattern noticed in all ultrasound images is a consequence of destructive interference from these scattered waves.

Refraction occurs when the speed of the ultrasonic wave in the two media are different. This difference in speeds results in a change of direction of the transmitted wave. Refraction sometimes accounts for the artefacts seen in ultrasound images. In terms of the image formation process, the contribution of refraction is minimal. In fact refraction is not usually a problem since most structures can still be clearly depicted even with deviation of the transmitted wave by a few degrees [56]. Snell's law gives the mathematical relationship between the incident angle of the ultrasonic wave θ_i , the refracted angle of the ultrasonic wave θ_r , and the speeds in the two media c_1 and c_2 .

$$\frac{\sin \theta_i}{\sin \theta_r} = \frac{c_1}{c_2} \quad (2.8)$$

Absorption is the transference of ultrasonic beam energy to the propagating medium usually in the form of heat. The degree of absorption by a propagating medium is determined by the viscosity of the medium, its relaxation time and the frequency of the ultrasound beam through it. Viscosity is a measure of the frictional forces between particles of a medium as they move past one another. Alternatively it can be defined as the resistance of a medium to gradual deformation by shear stress or tensile stress. The higher the viscosity of a medium, the higher the friction between its particles, and the more the heat generated within the medium. Relaxation time is the time taken by particles of a medium to return to equilibrium from a perturbed state. When an ultrasound beam is incident on a medium, it causes vibration and displacement of its particles. A subsequent ultrasound pulse generates additional displacement which could be in opposing direction to the previous displacement. If the relaxation time of the particles is longer than the duration between pulses, collision of particles occurs resulting in additional dissipation of energy (heat). The beam frequency

is directly proportional to the speed at which the particles moves causing more viscous drag and reducing the probability that the particles will have reverted to their equilibrium position before the next disturbance. Therefore an increase in the beam frequency increases the absorption rate. Absorption is undesirable in diagnostic ultrasound because energy is lost by the ultrasound beam, leaving less energy available for examining tissues beyond the absorbing medium. Absorbent objects also tend to cast an acoustic shadow behind them making it more difficult to observe structures in their vicinity. Bones absorb ultrasound more strongly than soft tissues, which is why they tend to cast shadows on neighboring structures.

Attenuation is the loss of ultrasound beam energy as it passes through the tissues. Attenuation is different from absorption in that absorption involves conversion of ultrasonic energy into heat within the medium it is passing through, whereas attenuation accounts for the total propagation loss of the beam energy. Because of this, attenuation refers to all losses from reflection, scattering, refraction and absorption. Attenuation is controlled by three factors: the attenuation coefficient of the medium, the distance traveled by the ultrasonic waves, and the frequency of the ultrasonic waves. Attenuation coefficient is a measure of how easily a medium can be penetrated by acoustic waves. The higher the attenuation coefficient of a medium the quicker it attenuates an ultrasonic beam. Similarly, distance affects the attenuation; the farther an ultrasonic beam travels through a medium the more attenuated it becomes. Attenuation is inversely related to frequency; the higher the frequency of the ultrasound wave, the greater the attenuation. A general representation of amplitude attenuation is given in (2.9) where A_0 is the initial amplitude of the ultrasonic wave, α is the attenuation coefficient of the medium and r is the path length the ultrasonic beam traveled.

$$A(r) = A_0 e^{-\alpha r} \quad (2.9)$$

The general solution for the 3-D acoustic wave equation of (2.4) can be rewritten by substituting the wave amplitude with (2.9) to give,

$$p(\mathbf{r}, t, k) = \Re[A_0 e^{-\alpha \mathbf{r}} e^{i\mathbf{k}\mathbf{r}} e^{i\omega t}] \quad (2.10)$$

The attenuation coefficient α is frequency dependent and can be represented with (2.11) where f is the frequency and a and b are coefficients determined empirically.

$$\alpha \equiv \alpha(f) = a f^b \quad (2.11)$$

In the case of blood $b = 1.2$. Water on the other hand is quadratic in frequency, that is, $b = 2$. Soft tissues usually have $b > 1$.

2.2 Ultrasound systems

An ultrasound system is a device that transmits acoustic waves into the body, detects reflected acoustic waves from structures in the body, and processes the detected reflections. It generates images of internal organs and structures, maps blood flow and tissue motion and provides blood velocity information. The transmitter, receivers, and signal processing unit makes up an ultrasound system.

2.2.1 Ultrasound transducer

An important component of any ultrasound system is the ultrasound transducer. Transducers consist of vibrating elements that generate acoustic waves. These elements are usually made from piezoelectric ceramics or crystals. Piezoelectric materials are able to create electric charge in response to an applied mechanical force. A piezoelectric transducer contains an array of piezoelectric elements that transmit focused energy into the body and receive the resulting reflections. A transducer could have between 32 to 3500 elements depending on what it is used for. Transducers that are specifically used for 2-D imaging have smaller number of elements while more modern transducers capable of producing 3-D and 4-D images could have as many as 3500 elements.

Ultrasound transducers can be classified according to the arrangement of their elements. There are three common types, as shown in Figure 2.3, namely: 1) linear array transducers, 2) sector array transducers and, 3) curved array transducers. The elements of linear array transducers are arranged in a parallel configuration thereby producing ultrasonic waves parallel to each other. The images produced by such arrays are rectangular. The advantage of using linear arrays is that they produce good near field image resolution. A disadvantage is that they are not suitable for imaging curved surfaces because they produce air gaps between the skin and the transducer which causes severe attenuation of ultrasonic energy. An area of application is their use in imaging the neck and the upper and lower extremities. Sector array transducers produce a pie shaped image that is narrow near the transducer and increase in width with deeper penetration. The elements are arranged in a parallel or concentric ring configuration. Such arrays are useful for scanning abdominal areas. A disadvantage is their poor near field resolution. Curved array transducers are a compromise

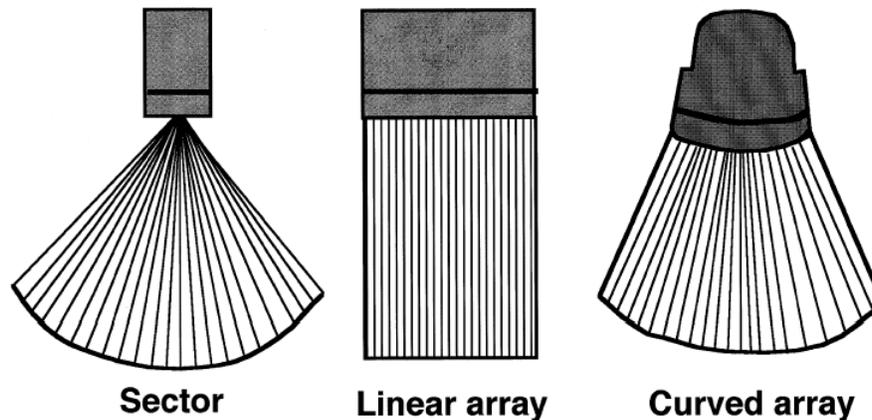


Figure 2.3: Three common types of ultrasound transducers used in medical diagnostics. Sector array transducers have poor near field resolution and are suitable for imaging abdominal regions. Linear array transducers have good near-field resolution and are suitable for imaging the extremities. Curved array transducers have good near field and good far field resolution and are suitable for imaging any part of the body.

between linear array transducers and sector array transducers. Their design is based on combining the advantages of the linear array and the sector array transducers, they therefore have good near field and better far field resolutions.

2.2.2 Operation of ultrasound systems

Figure 2.4 shows a block diagram of an ultrasound system. Its operation is built around transmitting an acoustic beam towards a target (transmitter), receiving reflected acoustic beam from a target (receiver), and signal processing of the received acoustic beam (signal processor).

At the transmitter, the generation of acoustic signals starts in the beamformer central control system block. This is used to control both the transmit beamformer and receive beamformer. Beamforming is a signal processing technique used to control the directionality of a signal. Beamforming can be achieved by selecting appropriate delays for signals in each transducer element such that the interference patterns from these signals produces a signal with majority of its energy is in a particular angular direction. The transmit beamformer block is thus used to generate digital transmit signals (through beamforming) with the appropriate timing and phase information. These digital signals are usually 8 bit to 12 bit signals at rates of about 40 MHz. A digital-to-analog converter (DAC) is used to convert the digital signal to an analog signal with its corresponding phase and timing information. The derived analog signals are used to excite the transmitter elements after they have been amplified by a high voltage amplifier (HV Tx). The transducer elements then emit the

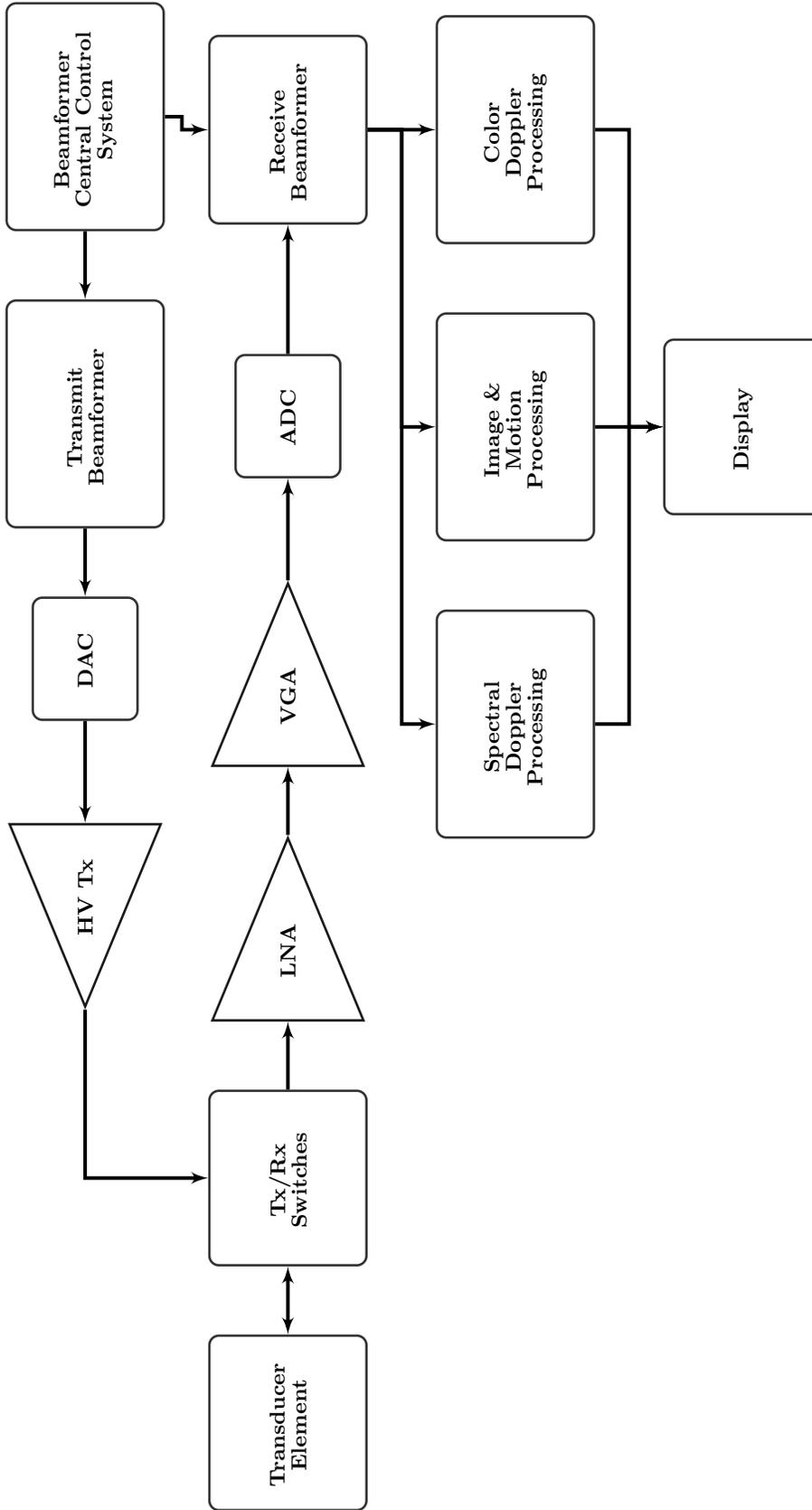


Figure 2.4: A block diagram depicting the mode of operation of an ultrasound system. HV Tx is the high voltage transmitter, DAC is the digital-to-analog converter, ADC is the analog-to-digital converter, LNA is the low noise amplifier, VGA is the variable gain amplifier, and Tx/Rx is the transmitter/receiver.

appropriate acoustic beam that is propagated through the body. The transmit/receive switches are used to switch between the transmitter and receiver so that the transducer can act as an actuator or a sensor. The Tx/Rx switches are implemented as multiplexers which help to dynamically change the active transducer aperture over the available transducer elements. The multiplexers must be able to transmit pulses with voltages as large as 200 Volts (peak-to-peak) and currents as high as 2 Amps. They must be able to switch rapidly to modify the configuration of the active aperture and maximize image frame rate. They must also have minimal charge injection¹ to avoid spurious transmissions and associated image artifacts.

At the receiver, the transducer collects the reflected acoustic waves which is passed to a Tx/Rx switch. The Tx/Rx switch at the receiver acts primarily as a protection mechanism for the low noise amplifier (LNA). It protects the low noise amplifier from the high voltage transmit pulses and isolates its input from the transmitter. It is implemented using an array of properly biased diodes which automatically turn on and off when they encounter high voltages. The Tx/Rx switch must have high recovery times to ensure that the receiver is on immediately after a transmit pulse is discharged. This is critical for imaging at shallow depths when the transmit time of the acoustic wave beam is very short. The output signal of the Tx/Rx switch passes through the LNA. The LNA is used to amplify the received signal power while reducing the effect of noise. It is designed to have a very good noise performance and sufficient gain. To achieve excellent noise performance, the connection between the transducer elements and the LNA must be terminated with low impedance. That is, the input impedance to the amplifier must be low. If this is not the case, capacitance effects and source impedance effects will limit the bandwidth of the received signal. The variable-gain amplifier provides the receiver with sufficient dynamic range over the full receive cycle. It is sometimes called time gain control. The main use of the variable-gain amplifier is to provide increased gain for signals from deeper in the body. These signals are the last to arrive and usually have lower energy levels. Therefore, a variable-gain amplifier dynamically varies the gain depending on the signal power level to maintain image uniformity. The analog-to-digital converter converts the received analog signal to a digital signal. The receive beamformer is then used to adjust the amplitude and delays of signals from each transducer element such that the signals at particular angles (or scan lines) can be retrieved.

¹Charge injection in analog switches and multiplexers refers to current leakage caused by stray capacitance associated with the transistors that make up the analog switch.

The final stage is the signal processor. The color doppler processing block is used to perform signal processing when color doppler images are desired, the image and motion processing block is used when B-mode or M-mode images are desired, and the spectral doppler processing block is used when spectral doppler images are desired. Color doppler, B-mode, M-mode and spectral doppler are discussed in Section 2.2.3.

The signal processing component of ultrasound systems can be broadly grouped into filtering, detection, and compression. It should be noted that there are other advanced signal processing techniques that can be applied to the received signal such as frequency compounding, time-frequency compensation, axial and lateral gain control, echo line averaging, and matched filtering. However, these will not be discussed here; Ali *et al.* [3], Evans *et al.* [56], and Hoskins *et al.* [84] provide detailed descriptions and can be used as references. The digitized (beamformed) received signal is first filtered to reduce noise outside the frequencies of interest. A filter is also used to select the type of imaging desired. For conventional imaging, a bandpass filter is used to select the fundamental frequency. For harmonic imaging a bandpass filter is used to select the second or higher harmonics. Conventional imaging, which uses the fundamental frequency of the transmit signal, has better penetration and is selected when imaging at deeper depths is required. Harmonic imaging on the other hand uses the higher order harmonic of the fundamental frequency. This produces images with better resolution due to the higher frequency (integer multiples of the fundamental frequency). It also gives better tissue distinguishing properties. After filtering, detection is performed. The detection process is essentially envelope detection which can be achieved by using Hilbert's transform. The detected signal is then compressed to allow for adequate display. Compression involves adjusting the dynamic range of the detected signal to a level appropriate for human perception. The dynamic range of the human eye is about 30dB. The detected signals have a dynamic range that is determined by the variable-gain amplifier and/or the ADC, which is much larger than 30dB, so Log compression is typically used for adjusting the dynamic range.

2.2.3 Ultrasound imaging modes

There are many imaging modes for ultrasound [191], but the most commonly used modes in echocardiography are: B-mode, M-mode and Doppler ultrasound. B-mode stands for brightness mode imaging. It is the image derived when the intensity or power of the reflected acoustic beam is resolved on a brightness scale. In these images the brightest regions show strongest reflections and the darker regions signifies weaker reflections. This is the most common type of images used by the image processing community for developing ultrasound

segmentation algorithms. The methods and evaluations performed in this dissertation are based on B-mode images.

M-mode stands for motion mode imaging. M-mode imaging is used to capture the movement of structures over time. It usually involves using a single scan line along an area of interest and observing how structures that intersect the scan line move towards and away from the ultrasound transducer over time. M-mode images are possible because ultrasound is able to produce images with good temporal resolution (discussed in Section 2.4). An M-mode image is commonly used for measuring chamber dimensions and calculating fractional shortening and ejection fraction of the ventricles.

Doppler imaging uses the doppler effect to measure the relative velocity of moving structures (usually blood). They are also used to determine whether these structures are moving towards or away from the transducer. Doppler imaging are of many kinds but the most used, particularly in echocardiography, are color doppler and spectral doppler. Color doppler is used to analyze the velocity and direction of blood flow using a color map superimposed on a B-mode image. In a color doppler image, blue represents blood flow away from the transducer, red represents blood flowing toward the transducer, black represents blood flowing perpendicular to the scanning plane and green or white represents areas of turbulent flow. Color doppler is essential in the diagnosis of many CHD. For example, in the diagnosis of septal defects, color doppler is used to pinpoint the location where mixing of blood between the ventricles or atria occur. Spectral doppler is used to display the spectrum of flow velocities.

2.3 Three- and four-dimensional echocardiography

The predominant technology in obstetrics, particularly fetal echocardiography, is 2-D ultrasound. Recent advances in 3-D imaging technology is making 3-D ultrasound systems an attractive and viable alternative. Three-dimensional ultrasound imaging attempts to address the limitations of 2-D ultrasound, particularly in its use as a diagnostic tool. The primary goal of 3-D ultrasound is to enhance image interpretation by providing a more intuitive spatial relationship of organs than 2-D images. The chances of finding a structural or functional anomaly is higher with 3-D than with 2-D scanning. Analyzing 2-D images of a 3-D structure does not usually provide the full diagnostic information required, and this could lead to variable or incorrect diagnosis. Four-dimensional ultrasound is an extension of 3-D ultrasound with a time component. The same transducers are used to obtain multiple volumes during a specific scan time. This helps to capture the 3-D motion of organs. Four-dimensional ultrasound is used predominately in fetal and adult echocardiography where the

motion of the heart and its structures are of particular interest. Three-dimensional and 4-D ultrasound imaging involves three steps: acquisition, volume rendering and visualization.

There are three different approaches used in acquiring 3-D or 4-D ultrasound images: freehand scanning, mechanical scanning and electronic scanning. Early approaches to 3-D ultrasound image acquisition were based on using a series of 2-D images in a process called freehand scanning. In freehand scanning the transducer is manually steered along the patient's skin to obtain a series of 2-D images. These 2-D images are then combined to form a 3-D volume. A mechanism that tracks the transducer's position is included so that a 3-D reconstruction can be done with the right 2-D image in the right spatial position. An advantage of this method is that an operator can obtain 2-D images from any angle or position that gives the most information. However, it is also possible to leave gaps in the volume if an adequate number 2-D images are not acquired. Mechanical scanning uses a controlled mechanical system to sweep for several 2-D images at constant spatial or angular intervals. Linear transducer arrays are usually used for this process. The images can be obtained either at uniform spatial intervals (parallel) or at uniform angular intervals (rotational) around a fixed axis. The uniform sampling ensures a more complete 3-D dataset of the structure of interest than freehand scanning. A newer 3-D acquisition method is electronic scanning. It is based on the work of Olaf Von Ramm and Stephen Smith [174]. They developed the first real-time 3-D echocardiographic scanner capable of acquiring volumetric data at frame rates sufficient to depict cardiac motion. Matrix array transducers are used to electronically steer the ultrasound beam to obtain the volume of interest. The volumes obtained are pyramidal in shape.

Volume rendering is achieved through interpolation of the 2-D image planes into a volume. The interpolation keeps track of the spatial positions of each image plane to ensure correct interpolation and representation of structures in the volume. Any resulting 3-D volume is only as good as its constituent 2-D planes.

Visualization is the presentation of a 3-D volume for interactive analysis. An advantage of 3-D volumes is that it enables arbitrary 2-D planes at arbitrary location and orientation to be displayed at the same. In 4-D imaging, a video of multiple 3-D volumes can be displayed to analyze motion characteristics of organs (like a heart).

2.4 Challenges of ultrasound image analysis

From an image analysis perspective, the challenges encountered in the postprocessing of ultrasound images are from the limitation of the ultrasound system and the imaging process. Limitations can be attributed to the spatial and temporal characteristics of the

acoustic beam and/or to assumptions made by the imaging process about the behavior of the acoustic beams (artifacts).

2.4.1 Spatial resolution

The ability of an acoustic beam, produced by an ultrasound system, to resolve two objects close together is referred to as the spatial resolution of the ultrasound system. Spatial resolution of an ultrasound system depends on the wavelength of the acoustic wave produced by the system. For example the wavelength of a 10 MHz acoustic beam is approximately 0.15 mm so it will be impossible to resolve objects less than 0.15 mm apart using a 10 MHz transducer. Spatial resolution can refer to either axial or lateral resolution. Axial resolution is a measure of the minimum distance required to differentiate two objects located parallel to the direction of ultrasound beam. In other words it is the resolution along the axis of the acoustic beam. Consider a short pulse of a few cycles generated by a transducer. The spatial pulse length is defined as the wavelength of the pulse multiplied by number of cycles in the pulse. This pulse is repeated each time the transducer elements are excited. We define the pulse repetition frequency as the number of pulses released by the transducer elements in one second. The axial resolution is directly related to this spatial pulse length. It is given as half the spatial pulse length. This means that two objects that are separated by over half the pulse length can be resolved distinctly. For example, if the pulse length is 2 mm then structures situated along the beam axis which are less than 1 mm apart will not be resolved. We can deduce that the shorter the pulse length, the better the axial resolution. Also, increasing the beam frequency shortens the beam wavelength which in turn shortens the spatial pulse length and gives improved axial resolution.

Lateral resolution is the ability to distinguish objects that are side by side in a direction perpendicular to the acoustic beam axis. It depends on the beam width, the beam frequency and the scan line density. If the distance between two objects is smaller than the beam width, the ultrasound system detects this as a single object. It assumes the reflections from these objects are emanating from only one object. Beam width varies with depth so the near field lateral resolution is better than the far field lateral resolution. The beam frequency is another factor that affects the lateral resolution. Higher frequency beams have a longer near field and a less divergent far field (better penetration) so the beam width is narrower when higher frequency beams are used as opposed to lower frequency beams which results in better lateral resolution. However, increasing the beam frequency also increases the attenuation effects which makes higher frequency imaging undesirable as a means of improving lateral resolution. Scan line density also contributes to the lateral resolution.

Ultrasound images are formed by sampling and combining echo information from scan lines generated by scanning an acoustic beam across a plane of interest. A large number of scan lines means that the echo samples forming the image are taken at closer intervals which improves the resolution. Therefore the higher the scan line density the better the lateral resolution.

The main challenge of spatial resolution to ultrasound image analysis can be posed as how the choice of beam frequency is a tradeoff between spatial resolution and imaging depth. Lower beam frequencies are able to image deeper into the body. Depth is essential when imaging a fetal heart where the ultrasound beam travels through extra layer of body tissue and amniotic fluid before arriving at the fetus. However, using high frequencies even though they improve spatial resolution are subject to larger attenuation coefficients limiting the depth penetration of the acoustic waves. In terms of image analysis, the ability to resolve objects is the cornerstone of segmentation algorithms, therefore poor spatial resolution makes analysis more challenging.

2.4.2 Temporal resolution

The ability of an acoustic beam, produced by an ultrasound system, to capture the movement of an object over time is referred to as the temporal resolution of the ultrasound system. Temporal resolution is determined by the frame rate, which is the number of images displayed per second. Temporal resolution is important for performing realtime imaging of rapidly moving structures (e.g., heart valves). It is the time taken to acquire a single frame of an ultrasound image sequence. For example a temporal resolution of 100 ms implies that one frame was acquired every 100 ms. Factors that affect the frame rate include scan depth, number of scan lines and pulse repetition frequency (number of pulses per unit time). Reducing the scan depth translates to reducing the time between sending and receiving an acoustic beam which reduces the time required to capture one frame (frame rate). The pulse repetition frequency is inversely related to the duration of each successive pulse therefore when scan depth is increased the temporal resolution is increased and the pulse repetition frequency is lowered. The number of scan lines also affects the frame rate. The number of scan line is a function of the sector angle and line density. An increase in either of these increases the number of lines per frame. The mathematical relationship between the frame rate, the pulse repetition frequency and the number of scan lines per image is given by the following:

$$\text{Frame Rate} = \frac{\text{Pulse Reptition Frequency}}{\text{Lines per Frame}} \quad (2.12)$$

Ideally it would be desirable to select each imaging parameter independently of the others so as to optimize image quality. However, this is impossible under all circumstances because of the interdependence of these parameters and as a result image quality measured by temporal or spatial resolution will always suffer. For example, if an improvement in temporal resolution is desired the number of scan lines per frame can be reduced, however, doing this will mean lower scan line density and lower lateral resolution since echo samples that form the image are farther apart.

2.4.3 Artifacts

Artifacts arise from assumptions made by the ultrasound system about the acoustic beam propagation in tissues [84]. These assumptions include:

- The ultrasound beam only travels in a straight line with a constant rate of attenuation
- The speed of sound in all body tissues is constant (1540m/s)
- The ultrasound beam is infinitely thin with all echoes originating from its central axis
- The pulse travels only to targets that are on the beam axis and back to the transducer

Deviations from these conditions introduce visible artefacts in the acquired images. There are many artefacts which arise from violation of these assumptions. We now consider some of these artefacts.

1. **Reverberation artifacts:** This occurs when acoustic beam is repeatedly reflected between two highly reflective surfaces. It occurs only for normal incidence of the acoustic beam. It usually occurs when there is a strong reflective interface parallel to the transducer surface at shallow depths. For example, the transducer could act as an additional reflective surface such that when returning acoustic beams strike it they are re-reflected back into the body. This creates an identical structure in the acquired image that is twice the distance away from the transducer. The transducer sees the resulting re-reflected wave as if it is being reflected from an object at twice the depth. This results in the reverberation artifact. The attenuation effect ensures that this process is not repeated indefinitely as the beam eventually dies out.
2. **Mirror image artifact:** This is a type of reverberation artefact caused by specular reflection of the beam at a large smooth interface with high reflection coefficient (tissue/air interfaces). If the reflected beam encounters a scattering object, echoes from the scatterer can be returned along a reciprocal path which gives the mirror-like

artifact. This effect can be observed posterior to the diaphragm when imaging the liver

3. **Side lobe artifacts:** They are generated by side lobe beams. Side lobe beams are produced at the edges of an ultrasound element. These beams are usually weaker than the main lobe beams, however, if a highly reflective surface is encountered, these side lobes may generate reflection that is seen back at the transducer. Since the transducer assumes all reflection are from the main lobes the reflection from the side lobe beams are incorrectly located on the acquired image.
4. **Multipath artifacts:** They are caused when the reflected beam has different paths. For example, when part of the original beam returns to the transducer and the other part is reflected off a second interface before returning.
5. **Beamwidth artifacts:** This is based on the lateral resolution. When the separation between two objects is smaller than the beam width, the objects appear as if they are a single object.
6. **Propagation speed errors:** This occur when the ultrasound beam speed does not propagate at 1540m/s. This results in objects appearing at incorrect depths in the acquired images.
7. **Acoustic shadowing:** This occurs when a highly reflective or attenuating object is in the path of the ultrasound beam. This results in poor beam penetration beyond the reflecting object causing signal dropout needed for effective imaging beyond it. Interfaces such as soft tissue/gas or soft tissue/bone typically display the effects of acoustic shadowing.
8. **Near field clutter:** This is caused by acoustic noise near the transducer. Acoustic noise is caused by high amplitude oscillation of the piezoelectric elements. This makes it difficult to identify the structures that are particularly close to the transducer.

Some of these artifacts can be avoided by proper scanning or acquisition techniques, however, others are generated by the physical limitations of the ultrasound system. In addition, proper recognition of artifacts can give clues to tissue composition and aid diagnosis. It is therefore important that image analysis techniques can identify artifacts of diagnostic significance, which is a nontrivial task.

2.4.4 Speckle

Speckle noise is a multiplicative noise that degrades all ultrasound images. It is an interference pattern arising from reflections from subresolution scatterers. These scatterers are small inhomogeneities in the path of the acoustic beam. Images of these subresolution scatterers are not random but deterministic, and they can be reproduced exactly if the transducer is returned to the same position and if all conditions stay constant [21, 191]. The deterministic nature of speckle enables it to be used to track tissue movement/displacement, and also to detect lesions, such as tumors. Speckle is detrimental because it reduces image contrast² and degrades boundaries of structures.

Most modern ultrasound systems have builtin speckle reduction mechanism in the form of online processes. Online processes make use of multiple scans at different frequencies or spatial locations in a process called compounding. Compounding is when speckle is reduced by adding images of the same region obtained from transducers using different frequencies and different spatial locations. There are also offline processes that are applied to images after they have been acquired. These are image filters applied to B-Mode ultrasound images to reduce speckle noise. The various image filters used for speckle reduction are based on the general speckle noise model [112] given by (2.13)

$$f(\mathbf{x}) = g(\mathbf{x}) + \eta_a \cdot \eta_m(\mathbf{x}) \quad (2.13)$$

where $f(\mathbf{x})$ is the noisy image, $g(\mathbf{x})$ is the unknown noise free image, η_m and η_a represent the multiplicative signal dependent noise (speckle) and additive noise functions, respectively.

There are two types of compounding used in reducing speckle in ultrasound systems namely: frequency compounding and spatial compounding. Frequency compounding [65, 118] involves averaging multiple images acquired with different acoustic beam frequency. A simple frequency compounding method uses subbanding where a wideband signal is subdivided into subbands which is then transmitted. These subband signals represent pulses at different frequencies. The received signals are then summed. This works because of the incoherent averaging of the different speckle characteristics obtained by changing the spectrum of the acoustic beam. Frequency compounding, however, degrades the axial resolution of the ultrasound image because the pulse widths are made smaller by the subbanding process

²Image contrast is the ability to see an object/structure against a background.

Spatial compounding [194] involves acquiring images of a region from different viewing angles and averaging them. The speckle in each image is different and independent because of the different path traveled by the acoustic beam. Summing the images from these different viewing angles averages the speckle. Spatial compounding can eliminate speckle if enough number of images can be averaged. In practice, however, the limited number of unique directions that an ultrasound scanner can redirect the ultrasound beam from a transducer of limited size means that only a moderate reduction in speckle can be achieved by this method.

Offline postprocessing is another way of performing speckle reduction of ultrasound images. It is a particularly attractive option because of the increasing power and speed of digital signal processing (DSP) chips. Many image filtering methods are used for speckle reduction, the most popular of which are median filter, wavelet based speckle reduction methods, and diffusion filter. Median filter [85] is a simple nonlinear operator that replaces a central pixel in a predefined rectangular window with the median value of all pixels within the window. The degree of smoothing of a median filter is controlled by the window size so in many applications the window size is chosen empirically depending on the noise level. An improved median filter called the adaptive weighted median filter (AWMF) [112] uses weighting coefficients on each pixel within a rectangle window such that the weight determine how many times a pixel is repeated in the computation of the median. For example, if a pixel has a weight of 5, then the pixel value will have 5 entries in the median computation. This helps to preserve edge details even with large window sizes. The weights are computed on the statistical properties of the current window so the weights vary depending on which region of the image is being considered.

Wavelet transform can be used for speckle noise reduction [74, 193]. Wavelet transform is a time-frequency tool used in digital signal processing to represent a signal. Wavelet denoising is based on applying a threshold on the wavelet coefficient. The basic procedure for wavelet speckle reduction is: *a*) compute the discrete wavelet transform (DWT) of an image, *b*) threshold its wavelet coefficients, *c*) compute the inverse discrete wavelet (IDWT) transform to obtain a noise free estimate of the image. The selection of the optimal threshold for wavelet speckle reduction is an ongoing area of active research.

Diffusion filtering uses the solutions to partial differential equations to reduce noise in an image. It involves the design of appropriate diffusion coefficients through the introduction of image features to achieve denoising. The most commonly used speckle reducing diffusion filter is the anisotropic diffusion filter proposed by Perona and Malik [146]. The Perona-

Malik filter is given by (2.14)

$$\frac{\partial f(\mathbf{x}, t)}{\partial t} = \text{div}(c(\mathbf{x}, t)\nabla f(\mathbf{x}, t)) \quad (2.14)$$

where $f(\mathbf{x}, t)$ denotes the noisy image to be denoised, ∇ denotes the gradient operator, $\text{div}(\cdot)$ is the divergence operator and $c(\mathbf{x}, t)$ is the diffusion coefficient that controls the rate of diffusion and it is usually a function of the image gradient. The time variable is introduced to signify that the original image is transformed over time towards an image that is the solution of the partial differential equation. Two functions that are sometimes used as the diffusion coefficient are given by (2.15), which is defined as,

$$c(\mathbf{x}, t) = \exp\left(-(|\nabla f(\mathbf{x}, t)|/K)^2\right)$$

$$c(\mathbf{x}, t) = \frac{1}{1 + \left(\frac{|\nabla f(\mathbf{x}, t)|}{K}\right)^2} \quad (2.15)$$

where K is a constant that controls the sensitivity to edges and is usually chosen experimentally or as a function of the noise in the image.

2.5 Summary

In this chapter, a necessary and simple description of the ultrasound imaging modality was presented. It is essential that the image formation process as well as the imaging systems used in acquiring ultrasound image is presented as a foundation for understanding the motivation and challenges of algorithms presented in subsequent chapters. It was shown that the source of ultrasound images, which are acoustic waves, are governed by electromagnetic laws which describe reflection, refraction, scattering, absorption and attenuation. Furthermore, properties such as spatial resolution, temporal resolution, artifacts and speckle that make ultrasound image analysis challenging were presented. In Chapter 3, a literature survey of methods that have been used to overcome some of these challenges for effective image analysis is presented. In particular, segmentation and motion estimation methods which are critical to the realization of the detection system described in this dissertation are reviewed.

CHAPTER 3

LITERATURE SURVEY

The overall goal of this dissertation is to present a system that can aid the prenatal detection of congenital heart defects (CHD) using 4-D ultrasound images. Three image analysis tools are critical to the realization of such a system. They are: 1) segmentation, 2) motion estimation, and 3) classification. Segmentation is essential to identifying specific structures that are affected by CHD. For example, a system to detect hypoplastic left heart syndrome (HLHS), a CHD that affects the left ventricle, must have a mechanism to isolate the left ventricle from an echocardiographic image; segmentation offers such a mechanism. Similarly, motion estimation is important in analyzing the motion characteristics of moving objects in an image. Since 4-D echocardiographic images include a time dimension, a mechanism to estimate the motion vectors is important if the added dimensionality is to be leveraged for detection of CHD. Motion analysis gives the means of extracting this information for a viable working system. Finally, classification gives the means for interpreting and grouping the extracted segmentation and motion information to adequately identify a CHD. The following is a presentation of some methods that have been proposed in literature to achieve segmentation and motion estimation. Because of the wide scope of these image analysis tools, the review is limited to only those used in echocardiographic image analysis. The segmentation and motion estimation method described in Chapter 5 and Chapter 6 incorporates some of the ideas discussed in this chapter.

3.1 Segmentation in echocardiography

Segmentation is one of the most critical aspects of digital image processing. It involves partitioning or grouping objects in an image either for identification or analysis. There are many methods in literature for achieving image segmentation, many of which are application-specific. Segmentation in medical image analysis has seen tremendous growth over the last 20-30 years. The need for better understanding of human anatomy through imaging has contributed to increased research activity in this area. The main challenge of medical image segmentation arises from the complexity and variability of the human

anatomy. This makes most medical image segmentation methods application-specific. For example, a segmentation method that produces good results with the liver may not do the same if it is applied to the heart or brain and vice versa. Also, a method that is robust to image noise may perform poorly when certain image artefacts are present. There are many instances where a successful segmentation task will involve multiple segmentation methods, applied one after the other, to give the best result.

Ultrasound image segmentation has proven to be a difficult task over the years because of the effects image quality has on segmentation results. Ultrasound images are characterized by artefacts such as speckle, shadows, attenuation and signal drop out. This makes ultrasound segmentation complicated. Chapter 2 of this dissertation was devoted to the principles behind these artefacts and to the reasons image quality is a significant impediment to analysis. In addition, low contrast between structures of interest makes detection of boundaries difficult. Advances in ultrasound technology has mitigated some of these complications. For example, frequency compounding [65, 118] and spatial compounding [194] in transducer systems gives reasonable reduction of speckle noise. A majority of the methods described in literature are based on 2-D rather than 3-D or 4-D ultrasound images. This is expected to change with the increase in diagnostic use of 3-D and 4-D ultrasound.

Echocardiography (ultrasound imaging of the heart) is the preferred imaging modality for fetal heart. This is because it is inexpensive, safe, and provides good temporal resolution for observing the motion of fetal heart structures [191]. Tracking of the structure and function of the left ventricle is important in diagnosing many heart diseases. For example, the ejection fraction, which is a measure of the amount of blood pumped out of the ventricles, cannot be measured without a method that can distinguish and identify the left ventricle from the other heart chambers. Segmentation methods can be broadly divided into edge-based methods, region-based methods, deformable methods, statistical methods and level sets.

3.1.1 Edge-based segmentation

Early echocardiographic segmentation techniques focused on identifying the boundary between blood region and tissue region using an edge detector. Edges are locations in an image with strong intensity contrast representing object or region boundaries. An edge detector is a tool that measures the degree at which image brightness changes or is discontinuous. Several edge detection methods for echocardiographic images are described in literature. Many publications [34, 37, 81, 93, 106, 182, 202] used 2-D gradient operators to find boundaries between blood and heart tissues (heart wall or septum). Because of

the widely differing ultrasonic backscatter property of blood and tissue, there is a strong spike in the gradient value at their interface. However, gradient edge detection methods are not robust to noise, particularly speckle noise. The use of gradient operators results in a poor ratio of true edges to false edges [10]. Rather than using only a gradient operator, some methods attempt to reduce detection of false edges with a combination of filtering, cross-correlation, neural networks etc.

Detmer *et al.* [40] used matched filtering to detect boundaries of the heart in echocardiographic images. Matched filtering is a method used in signal processing to obtain a known signal, called the template, contained in an unknown signal. It involves correlating the template to detect its presence in the unknown signal. Detmer *et al.* used a template derived from 322 images with manually selected boundaries points. They then cross-correlated this template with a new image. The maxima (in each direction) of the cross-correlation function is the expected boundary profile. A main drawback of this approach is the use of a fixed template. There is wide variability in tissue to blood interface under different scan conditions which their method does not account for. The template will have to be defined every time a scan is obtained which is not feasible under clinical conditions.

Hunter *et al.* [86, 87] used neural networks to identify edges. They used a training set of known blood-tissue boundary of the heart to train their neural network. To detect edge points in a new image, they used a 7×7 window around every pixel as input into the trained neural network. The neural network then classifies the pixel as either an edge pixel or a nonedge pixel. A neural network approach was also used by Sussner *et al.* [189]. Their approach involved using a neural network and cross-correlation to track edge points between frames of echocardiographic images. In the first frame, the neural network is used to detect the boundary between blood and tissue interface. In subsequent frames, cross correlation is used to track the boundary in an 11×11 window. Since ultrasound imaging modality produces images that exhibit wide variations in edge strength (due to speckle and artifacts), neural networks provide an advantage over edge detectors because the classification of pixels are based on their probability of being edges rather than purely by their edge strength.

3.1.2 Region-based segmentation

Region-based approaches to echocardiographic segmentation are based on the principle of thresholding. Like edge detection, it uses the differing back scatter properties of the blood and soft tissue to identify edges between interfaces. A threshold value is chosen based on the fact that the average gray-scale value of a pixel in a blood region is different from that of a soft tissue or cardiac wall region. The various region-based methods available in literature

are only different in the way each chooses the threshold value.

A common way of choosing a threshold value is manually. A human operator adjusts the threshold value while visually observing the resulting edge profile. This method is subjective and suboptimal. Another option of choosing threshold values is to use the image histogram. Consider an image with two distinct regions. The histogram for this image will be bimodal with two peaks representing the two regions. A suitable threshold will be the minimum between these peaks. Ultrasound images do not satisfy this bimodal distribution, which makes this method unreliable for echocardiographic applications. Nevertheless, some authors [182, 176] have used image histograms to select thresholds for segmentation of blood and tissue regions. They used manual inspection of the histogram to select the appropriate threshold. Zhang *et al.* [210] used temporal information from a sequence of images to derive the threshold automatically. They used a temporal co-occurrence matrix to identify the moving points between 2-D images. They reported that these points correspond to the tissue-blood boundary. They claimed their method was robust to low signal-to-noise ratio, echo dropout and gray level intensity variability. They conducted their experiments within a controlled acquisition environment which suggested limited clinical potential.

3.1.3 Deformable models

The newer echocardiographic segmentation methods have moved away from edge detection and thresholding techniques, instead segmentation is treated as a contour finding problem. Edge detection and thresholding, as previously discussed, are intensity based gradient methods which tend to have limited success because of intensity variations even in homogeneous regions. Contour based methods, on the other hand, are not as susceptible to intensity variations and it is easier to incorporate prior knowledge to constrain the contours. Contour methods are based on finding the equilibrium state of a closed system. In physics, equilibrium is achieved when an internal force balances an external force (*i.e.*, sum is zero). For example, a seesaw is at equilibrium when the force induced by gravity balances the forces induced by the weights at its two ends. In the same way, contour methods are based on an energy minimization scheme where the sum of an internal and external energy functional is minimized. The external energy is defined so that it controls the contour's attraction to image edges or features. The internal energy, on the other hand, is defined so that it controls the elasticity and curvature of the contour. The elasticity of an object measures its flexibility. It is the property of an object that enables it to return to its original shape or size when a distorting/deforming force is removed. Curvature measures the degree to which an object bends. A detailed description and mathematical formulation for how contours

can be defined in a closed force-balance system is given in Chapter 5.

The earliest use of deformable contours for segmentation was by Kass *et al.* [95]. They used an external energy term based on the edge profile of the image feature (or object) to be segmented. The internal energy term was defined to maintain a smoothness constraint on the curve. Kass' method suffers from the same drawback as edge detectors when used to segment echocardiographic images since the external energy is defined on the image gradient. There have been many variations of the original contour model by Kass *et al.*, each trying to tailor the energy functional for a particular application. Xu *et al.* [205] used a class of external forces that can increase the capture range and also deal with boundary concavities. Increasing the capture range implies that the initial contour need not be close to the feature(s) to be segmented. Mishra *et al.* [133] proposed a contour finding method for a sequence of images. They defined an initial curve in the first image by using lowpass filtering and morphological operations. They then minimized the active contour energy functional [36] using genetic algorithm under a nonlinear constraint on the image gradient. The nonlinear constraint was used to discourage contour evolution towards low gradient regions. In subsequent frames, they used the contour found in the previous frame to initialize the energy minimization function for the current frame. This provided a way of tracking contours across images. Their method used a computationally expensive genetic algorithm which may not extend well in clinical settings where fast/real-time computations are desired. Also, their method assumes that there is no relative movement between the transducer and the patient, which is almost always not the case in real-world applications. Mignotte *et al.* [126] used a statistical external energy term, claiming that it is well-suited for ultrasound images with missing boundaries. They modeled the gray level statistics of an ultrasound image as a Rayleigh distributed random variable. They performed energy minimization with a multiscale scheme proposed in [80]. Chen *et al.* [27, 28] incorporated prior information in their contour model. They formulated an energy optimization method that used intensity profiles and shape priors to constrain the evolution of a contour. The intensity profile guides the contour evolution based on the maximization of a mutual information criterion. Chalana *et al.* [26] developed a multiple-contour model to detect both the inner and outer walls of a heart in an echocardiographic image. They used the image intensity gradient as the attracting force for the contour. To maintain continuity between frames, they used an external energy term that constrained the motion between consecutive frames. They validated their method on 44 clinical datasets and compared their results with manual delineation of the inner and outer walls. They reported an average correlation coefficient

between manual delineation and their method as 0.95 for the outer walls and 0.91 for the inner walls. Kucera *et al.* [101] used a region based external force in the energy minimization scheme. Mikic *et al.* [128] described an approach for sequences of ultrasound images. They used a Gaussian smoothed intensity profile as the external force. For the propagation of a fitted contour from frame to frame, they used optical flow estimates. Optical flow is defined as the pattern of apparent motion of objects, edges or image brightness in an image sequence.

3.1.4 Statistical models

Another class of contour models is the statistical contour models. In these methods, the contour evolution is typically learned from a training dataset. The model derived from the training data is used to locate and/or detect a similar class of features present in other images. This is particularly useful for medical image segmentation because the variability within a particular class of features can be captured and used as a constraining criteria during segmentation. The performance of the statistical model is tied to the size of the training dataset and on how well the training dataset captures the variability within the feature class. Because of this, statistical models requires large databases for useful clinical applications. The pioneering work on statistical models was done by Cootes *et al.* [36]. Their method used principal component analysis (PCA) to model the variability within the training dataset. The parameters obtained through PCA were used in a custom search algorithm which ensures that the contour evolves in such a way that it does not deviate from the shape class represented by the training set. An extension of Cootes' work [35] included a combination of shape and texture information in defining the feature class. Bosch *et al.* [14] and Mitchell *et al.* [134] used an adaption of [35] to represent the shape, appearance and motion of the left ventricle. They used a nonlinear intensity normalization algorithm to accommodate ultrasound-specific intensity distribution. They reported correct left ventricle delineation in 97% of the images tested.

3.1.5 Level sets

Level sets method [144] is a numerical technique used for image segmentation that implicitly tracks the evolution of an interface, controlled by images forces, such that the zero level set is the desired boundary. Let us assume there is an interface represented by $\phi(t)$, then the zero level set of the interface is when $\phi(t) = 0$. For example, let us assume a 2-D boundary is to be tracked, then a 3-D $\phi(t)$ is defined such that under image forces, the surface evolution of $\phi(t)$ over time t yields a zero level set corresponding to the desired

boundary. The genius of this method is that $\phi(t)$ can be defined arbitrarily and the image forces will always drive it toward the desired zero level set. Level sets is sometimes used either as an alternative to contour models or in conjunction with them. The advantage of level set is that they can easily track shapes that change topology. For example, when a shape splits into two or develops a hole. Because of this, they are ideal for modeling time varying features. Yan *et al.* [207] applied level sets method to echocardiographic images using an adaptation of the fast marching method [169]. The fast marching method is a numerical method for solving boundary value problems. It is particularly used to compute solutions to the nonlinear Eikonal equation of which level sets is a special case [169]. In their fast marching computation, Yan *et al.* used an average intensity measure for the speed term in the level sets formulation rather than the local intensity gradient. They claimed that this reduced errors attributed to using local intensity gradient. Lin *et al.* [108] went further by using a multiresolution pyramidal approach (coarse to fine) to combine both edge and region information in the level sets formulation. A main drawback of their method was that it depended on being able to extract a contour at a high pyramid level (fine or high resolution level).

3.1.6 Three- and four-dimensional echocardiographic segmentation

One of the earliest work on 3-D echocardiographic segmentation was by Coppini *et al.* [37]. They considered segmentation and reconstruction of the left ventricle. They identified edge points in 2-D slices of the 3-D image by using the zero-crossings of a Laplacian-of-Gaussian edge detector. A neural network was used to classify the identified edge points as part of a boundary or not. Finally, an elastic surface model was used to fit the edge points that fall on a boundary to form a surface. Song *et al.* [185] formulated the segmentation of the heart as a surface fitting problem. Their goal was to find a 3-D surface that has the greatest probability given the 3-D image. They trained a Bayesian network with 20 images. Wolf *et al.* used a segmentation approached called restricted optimal path exploring segmentation (ROPES) on 2-D slices of a 3-D image. ROPES involves finding candidate edge points which minimizes a cost function based on a multiscale criterion. The identified edge points are then linked to form a closed contour. Angelini *et al.* described a segmentation method for 3-D echocardiographic images that used a wavelet de-speckling filter before using a deformable contour to identify contours in 2-D slices of a 3-D image. The 2-D contours are then interpolated to form a 3-D surface. Montagnat *et al.* [135] described a method that combines anisotropic diffusion filtering with a deformable model to segment

echocardiographic images. Anisotropic diffusion is a filtering methods formulated using partial differential equations. It applies the law of diffusion on pixel intensities to smooth textures in an image.

3.2 Motion estimation in echocardiography

Motion is often an important feature in medical images, particularly in cardiac imaging. Accurate estimation of motion has many diagnostic benefits in addition to improving certain image processing tasks such as segmentation. Since the images analyzed in this dissertation are 4-D echocardiographic images, estimating the motion of fetal heart structures will provide information that can be used to improve segmentation results. For example, motion estimates could be used to infer missing boundary information in one frame from another frame and vice versa. Motion estimation (also called optical flow) algorithms can be divided into the following types: 1) differential methods, 2) frequency-based methods, and 3) correlation based methods. Differential techniques [83, 115, 136, 139, 141, 152, 177, 198, 22, 29] compute image velocity from spatio-temporal derivatives of image intensities. The assumption is that the image domain is continuous and differentiable in space and time. Frequency based methods [2, 60, 78, 197] use velocity tuned filters. These filters are orientation sensitive and are applied in the Fourier domain. Correlation-based methods [9, 92, 100, 109, 110, 178, 190] are used when numerical differentiation is impractical because of small temporal support (only a few frames) or poor signal-to-noise ratio. Correlation based methods use feature matching that attempts to match features across sequences of images. A complete and detailed survey of optical flow algorithms that include the aforementioned differential-, frequency- and correlation-based methods can be found in [11].

In echocardiography, optical flow is used mostly for image enhancement or speckle reduction. Instances where optical flow is used in literature for echocardiographic image segmentation are limited. Differential methods for optical flow computation are most suited for ultrasound image sequences because the images have a spatio-temporal relationship. The two important papers on differential optical flow are by Horn and Schunck [83] and Lucas and Kanade [113]. Horn and Schunck described a global method of solving for image velocities. Their method used a variational framework to optimizing a functional (described in Chapter 6) based on the brightness constancy constraint. The brightness constancy constraint implies that for a small change in time, the intensity values of pixels in an image does not change. They also used a regularization term in the functional that controls the smoothness of the flow field. Lucas and Kanade described a local least-squares formulation for computing flow fields. Their method assumed that the velocity is essentially constant

in a local neighborhood of the pixel under consideration. Chapter 6 of this dissertation gives a detailed description of Horn and Schunck's method. In addition, a method [192] that is an adaptation of Horn and Schunck's method suitable for echocardiographic images is presented in Chapter 6.

3.3 Summary

This chapter presented a literature survey of segmentation and motion estimation methods for echocardiographic images. The progression of the state of the art from simple thresholding segmentation techniques to more sophisticated techniques that utilize ideas from machine learning, probability theory and statistics was highlighted. The methods described in this chapter are by no means an exhaustive list of methods available in literature, however, important methods with significant contribution to echocardiography was covered. The next chapter will commence the description of the detection system, starting with the preprocessing (location estimation) stage. The location estimation is used to extract the region of an echocardiographic image containing the fetal heart on which further analysis such as segmentation, and motion estimation (Chapter 6) is performed.

CHAPTER 4

DATA

A dataset containing 254 4-D fetal echocardiographic was obtained over a period of 5 years (2009-2013) from the Intermountain Healthcare Primary Children’s Hospital in Salt Lake City, Utah under the University of Utah Institutional Review Board (IRB) approval number 00030844. The 254 images were from a total of 43 patients with each patient signing a written informed consent form allowing their fetal echocardiographic scans to be used in the validation experiments. Out of the 254 images acquired, 19 images were from patients with confirmed diagnosis of hypoplastic left heart syndrome (HLHS) while the remaining 235 images were from patients without a congenital heart defect. The images were acquired via multiple Philips Healthcare *iE33 xMATRIX* ultrasound systems using Philips Healthcare *X5-1* matrix transducer and *X7-2* matrix transducer. The *X7-2* transducer is a matrix array transducer with 2500 elements operating at 2-7 MHz while the *X5-1* transducer has 3040 elements operating at 1-5 MHz. Two acquisition modes were used in acquiring the images and they are: 1) Live 3-D (L3D), 2) full volume (FV). Live 3-D implies acquisition at a narrow angle, about 30° in the elevation direction and about 65° in the longitudinal (azimuth) direction. The advantages of L3D are higher spatial resolution resulting from higher scan line density and higher temporal resolution because of narrow scan volume [20]. Full volume implies acquisition over a larger angle, up to 104° in elevation and longitudinal (azimuth) direction. Full volume images are derived from stitching together multiple L3D images using a technique called spatio-temporal image correlation (STIC) [20, 41].

An independent analysis of the images was performed by a fetal cardiologist, Dr. M. D. Puchalski, to identify which of the 254 images had the best quality and the best diagnostic information for use in the validation experiments in later chapters. Dr. Puchalski is the Director of Non-Invasive Imaging at the Primary Children’s Hospital in Salt Lake City, Utah and he has 14+ years of experience in diagnosing congenital heart defects. Of the 254 images, only 130 images were considered to provide good diagnostic information. Dr. Puchalski excluded images based on the following criteria:

1. Partial volume: Partial volume means the full volume of the fetal fetal heart was not acquired during a scan. The cardiac chambers are therefore truncated and since the detection system is based on the volume characteristics of the chambers, images displaying this characteristic were excluded.
2. Spatial resolution: This is based on visual inspection of the images to determine if small structures with diagnostic information could be identified.
3. Stitching artifacts: This occurs when images are acquired using the FV acquisition mode. As described earlier, FV acquisition is based on combining multiple L3D images, however, this combination can result in stitching artifacts which degrades the image and can contribute to error in analysis. Therefore, images exhibiting severe stitching artifacts were excluded.
4. Fetal movement: Images that exhibited fetal movement during image acquisition were excluded.

All validation experiments and analysis performed in this dissertation are based on the independently selected 130 images. The breakdown of the 130 images is as follows: 20 of 130 images are from 5 patients with HLHS, 110 of 130 images are from 26 patients without any congenital heart defect. There were 76 images acquired using the FV acquisition mode and 54 images acquired using the L3D acquisition mode. There were 46 images acquired using the $X5 - 1$ transducer and 84 using the $X7 - 2$ transducer. The 130 images were digitized at an average resolution of $176 \times 198 \times 211$ with an average physical voxel calibration factor of 0.519 mm/voxel . The gestational ages of the fetuses ranged from $19\frac{5}{7}$ to $35\frac{2}{7}$ weeks. The number of volumetric frames in the 130 4-D images ranged from 8 to 288 frames with acquisition times between 13.08 to 290.95 ms. Table 4.1 summarizes the characteristics of the 130 images according to gestational age, number of frames, frame time, acquisition mode and transducer type. Appendix A gives a comprehensive and extended breakdown of each image according to these characteristics.

Table 4.1: Breakdown of 4-D fetal echocardiographic dataset according to transducer type (X5-1 and X7-2), acquisition mode (full volume (FV) and live 3-D (L3D)), gestational age, number of frames, and frame time.

| Image Characteristics | Transducer | | Acquisition | | Gestational Age (weeks) | | | | | Diagnosis | | |
|-----------------------|------------|------|-------------|----|-------------------------|----------|----------|----------|----------|-----------------|------|--------|
| | X5-1 | X7-2 | L3D | FV | [0, 22) | [22, 24) | [24, 26) | [26, 28) | [28, 30) | [30, ∞) | HLHS | Normal |
| No. of Images | 46 | 84 | 54 | 76 | 19 | 23 | 54 | 15 | 8 | 11 | 20 | 110 |

| Image Characteristics | No. of Frames | | | Frame Time (milliseconds) | | | |
|-----------------------|---------------|----------|----------|---------------------------|---------|----------|-----------------|
| | [0, 15) | [15, 20) | [20, 25) | [25, ∞) | [0, 20) | [20, 50) | [50, ∞) |
| No. of Images | 13 | 51 | 30 | 36 | 42 | 52 | 36 |

CHAPTER 5

FETAL HEART LOCATION ESTIMATION

5.1 Abstract

In the analysis of fetal heart echocardiographic images, some structures other than the fetal heart are inadvertently captured during acquisition. These structures can either be from the mother or from the fetus which may complicate analysis by introducing errors through their misidentification as fetal heart structures. To limit such errors, a location estimation procedure is presented that extracts the region occupied by a fetal heart in an image. The location estimation procedure uses an edge detector and a deformable surface to extract the fetal heart region in a 4-D echocardiographic image. The edge detector estimates the epicardial (outer wall) surface of the fetal heart in each volumetric frame of a 4-D image. The deformable model uses the epicardial surface estimates as a gradient constraint to extract a region of interest. An aggregate region is computed by combining the regions of interest from each frame of the 4-D image which results in the largest volume or region occupied by the fetal heart during a cardiac cycle. The location estimation procedure has an added advantage of reducing computational cost of algorithms that analyze the fetal heart by only focusing on the relevant fetal heart region rather than the entire image with irrelevant information. The location estimation procedure was validated on 130 4-D fetal echocardiographic images and compared with manually identified fetal heart regions using four similarity indices: 1) Jaccard index, 2) Sørensen-Dice index, 3) Sensitivity index, and 4) Specificity index. The average values of these indices across the 130 images were measured as 80.70%, 89.19%, 91.04%, and 99.17%, respectively.

5.2 Introduction

Region of interest segmentation is a technique that classifies an image into a background and a foreground (binary image). The main purpose of extracting a region of interest is to isolate a structure from an image either for identification or for further processing. Thresholding is the most common form of region of interest segmentation where a threshold, chosen carefully, is applied to an image to divide it into a foreground containing the structure

of interest, and a background containing all other structures in the image. In medical images, thresholding is only effective when the structure of interest has a uniform intensity profile. For example, thresholding is effective for identifying the prostate or the liver [82] since these structures have uniform intensity distribution. In some cases they can also be used to identify lesions such as breast cancer lesions [170]. Since the structure of the heart is complex with nonuniform intensity distribution, thresholding is unreliable in determining the fetal heart region in an image. Another drawback of thresholding is that the spatial information of pixels are not considered in classifying them as background or foreground. For this reason images with high noise levels result in misclassified structures. Region growing [150] is another region of interest segmentation technique where a region is grown from a seed point by adding neighboring pixels with similar characteristic. The results of these methods are susceptible to speckle noise since speckle can change the characteristics of neighboring pixels that belong in the same region. Boundary methods [10, 34, 37, 40, 81, 93, 106, 145, 182, 202, 210] can also be used to localize structures by finding the boundary contour/surface of these structures. These methods use pixels with abrupt intensity changes (edges) to identify boundaries between regions. However, because of speckle noise, abrupt changes in intensities of echocardiographic images occur sporadically leading to many false edges. In addition, these boundary methods are not robust to echo shadowing or artifacts which are culprits for many missing boundary features in ultrasound images.

In this chapter a method that can be used to isolate a fetal heart in a 4-D echocardiographic image is presented. This method is a hybrid method that combines boundary information as well as region information. This method offers the following advantages over existing region of interest methods: 1) it can be used to find the region of interest of a complex structure such as the fetal heart with nonuniform intensity distribution, 2) it can be used with 4-D fetal echocardiographic images for which there are no existing method(s) to the author's best knowledge, 3) it takes advantage of all four dimensions available in a 4-D image to enhance missing boundary information that may arise from acoustic shadowing or ultrasound artefacts, and 4) spatial and intensity information are used to improve region of interest estimates. Figure 5.1 illustrates the procedure for the location estimation method. Each frame of an input 4-D fetal echocardiographic image is processed with an edge detection algorithm that estimates the epicardial surface. The estimates are passed to the deformable model block where a region of interest is extracted using the estimates as constraints. A cumulative region of interest is obtained by aggregating the regions from each frame. Finally, a morphological filtering block is used to remove noise

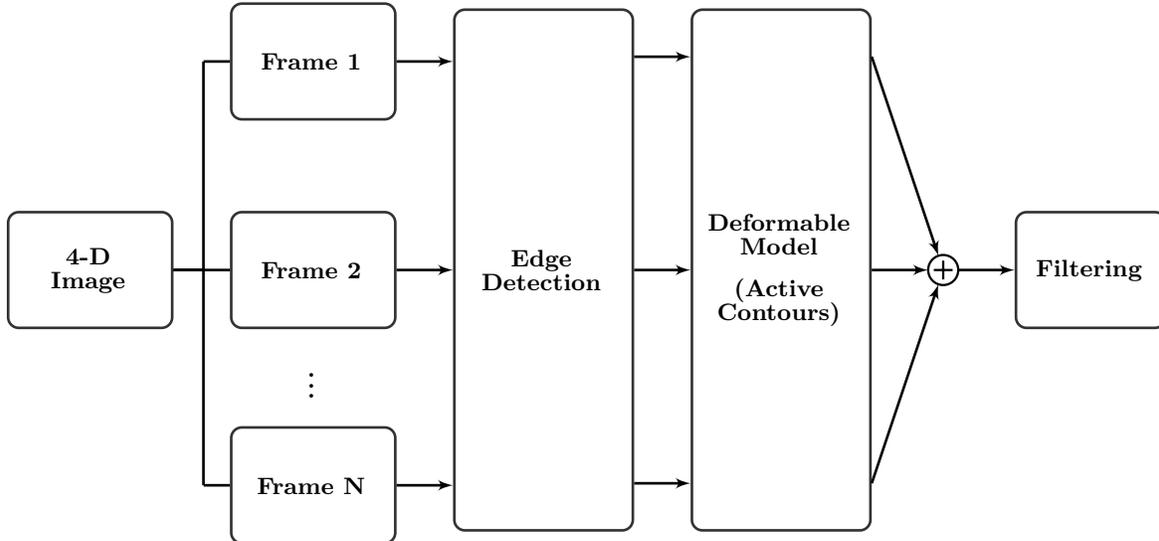


Figure 5.1: A block diagram describing the process for isolating a fetal heart from a 4-D echocardiographic image. Each frame in an input image is processed with an edge detector to estimate the epicardial surface. The region of interest is extracted for each frame using a deformable model. A cumulative region of interest is derived by aggregating the results from each frame which represents the largest volume occupied by a fetal heart during a cardiac cycle. The filtering block performs morphological filtering to remove noise and holes in the region of interest binary image.

and holes within the cumulative region of interest image.

The rest of this chapter is organized as follows: Section 5.3 describes the edge detector used to estimate the epicardial surface in a 4-D echocardiographic images. Section 5.4 discusses the deformable model used to extract the fetal heart region of interest from each frame. Section 5.5 describes the morphological filtering approach used. Section 5.6 describes the evaluation procedure and a discussion of the results obtained. Section 5.7 concludes this chapter.

5.3 Edge detection

The first step in the location estimation procedure is the use of an edge detector to find the epicardial surface estimate of a fetal heart. Edge detectors are used to identify points in an image where intensity values change abruptly or points where image intensities are discontinuous. These points are typically identified by computing the gradient of an image using a gradient operator such as the Laplacian-of-Gaussian (LoG) [77, 121] or derivative-of-Gaussian (DoG) [23]. Traditional edge detectors such as Sobel detector [183], Roberts detector [157], and Canny detector [23] are known to perform poorly at identifying boundaries in ultrasound images mainly because of the multiplicative speckle noise [18, 143].

Speckle noise introduces abrupt changes in the intensity values which mimics the boundary between two regions, so using a traditional edge detector results in many false edges. Filtering methods (see Chapter 2) such as median filtering [112] and anisotropic diffusion [146] are generally used to reduce the effects of multiplicative speckle noise in many ultrasound image analysis applications. However, a fetal heart at 20 weeks is very small in size, even smaller are its constituent structures, and applying these filtering techniques can smooth out these small structures which hamper their identification and analysis. Moreover, these filtering methods are highly sensitive to the window size (median filtering) and diffusion coefficients (anisotropic diffusion) which need to be chosen empirically for the best results. This makes automation of the region of interest evaluation almost impossible since these parameters cannot be optimized quantitatively.

A surface detection method is presented that uses a concept first proposed by Levoy [105] for displaying volumetric surfaces. While Levoy's method was used to visualize volumetric surfaces in computed tomography (CT) images by connecting surfaces with similar intensity values, the method described here is used to identify epicardial boundary voxels in fetal echocardiographic images. The method has the advantage of being robust to speckle while identifying the epicardial surface without the need for performing speckle reducing filtering. In addition the method does not suffer from the same drawback of false edges associated with traditional edge detectors in the absence of filtering. The procedure involved is based on using the local intensity and gradient information of each voxel to classify it as either a boundary voxel or a nonboundary voxel. Equation (5.1) represents the transformation equation $\alpha(\mathbf{x})$ for an image $f(\mathbf{x})$ where f_v is a threshold value (the choice of f_v is described in Section 5.3.1), $|\nabla f(\mathbf{x})|$ is the gradient magnitude of the image and r is a positive integer. The image is transformed to a new image with values between $[0, 1]$ according to three conditions. These conditions are based on the following:

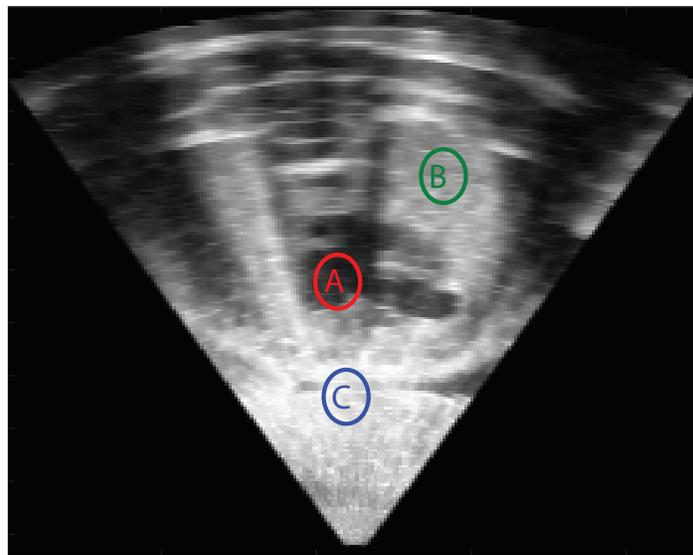
1. When a voxel has an intensity value equal to the threshold f_v *i.e.*, $f(\mathbf{x}) = f_v$.
2. When a voxel has a nonzero gradient magnitude and when the difference between its intensity value and f_v is within an integer multiple r of the gradient magnitude, *i.e.*, $|\nabla f(\mathbf{x})| \neq 0$ and $|f(\mathbf{x}) - f_v| \leq r |\nabla f(\mathbf{x})|$
3. When a voxel violates conditions (1) and (2) above.

$$\alpha(\mathbf{x}) = \begin{cases} 1 & f(\mathbf{x}) = f_v \\ 1 - \frac{1}{r} \left| \frac{f(\mathbf{x}) - f_v}{|\nabla f(\mathbf{x})|} \right| & |\nabla f(\mathbf{x})| \neq 0 \text{ and } |f(\mathbf{x}) - f_v| \leq r |\nabla f(\mathbf{x})| \\ 0 & \text{otherwise} \end{cases} \quad (5.1)$$

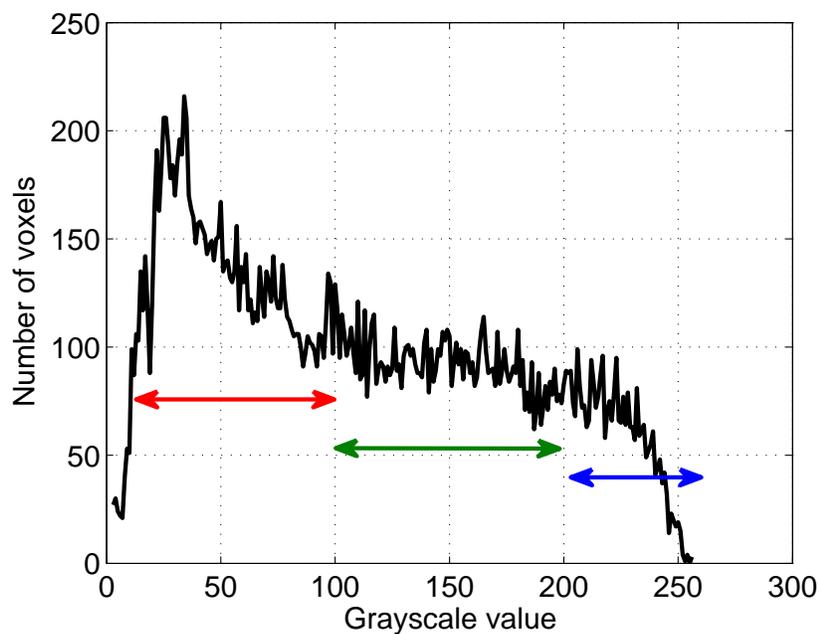
Condition 1 is special case of condition 2 and deals with the case when the intensity value of a voxel is equal to the value of the threshold f_v (*i.e.*, $f(\mathbf{x}) = f_v$). In such situations, a value of 1 is assigned as the transformation value. Condition 2 compares the difference between the threshold f_v and the intensity value of a voxel with its gradient magnitude. The significance of this condition is that a voxel with a high gradient magnitude is more likely to be classified as a boundary voxel (value close to 1) if it has an intensity value close to the threshold value and if voxels within r neighborhood exhibit similar characteristics. This reduces the occurrence of false edges resulting from speckle noise because not all high gradient voxels are classified as true boundaries. The integer r controls the size of a transition region around possible boundaries (high gradient magnitude). This transition region gives local information about whether a possible boundary point is a true edge or a false edge. In ultrasound images boundaries are not distinct [160], that is they are not usually represented by a single voxel, but rather they span a few voxels therefore voxels close to a true edge and within its transition region will have similar intensity and gradient characteristics. Condition 3 deals with cases not included in condition 1 or condition 2.

5.3.1 Choice of f_v

The choice of f_v is based on the characteristics of a fetal echocardiographic image. Consider the fetal heart image shown in Figure 5.2(a) and its histogram in Figure 5.2(b). The image can be roughly divided into three contrast regions. The brightest region (region C) corresponds to bony structures like the ribs and spine, the darkest region (region A) corresponds to cavities with blood flow like the heart chambers or the great vessels, and the intermediate region (region B) corresponds to soft tissues. The number of voxels belonging to each of these regions varies from image to image depending on the size of the fetal heart and the transducer sector angle. For example, a bigger fetal heart implies there will be a greater number of voxels in the darker regions because the fetal heart chambers will be bigger. Also a larger sector angle means that more soft tissue will be captured resulting in a higher number of voxels in the intermediate region. To reduce the impact of this variability, the threshold value is computed using a local mean criteria. The image is divided into



(a)



(b)

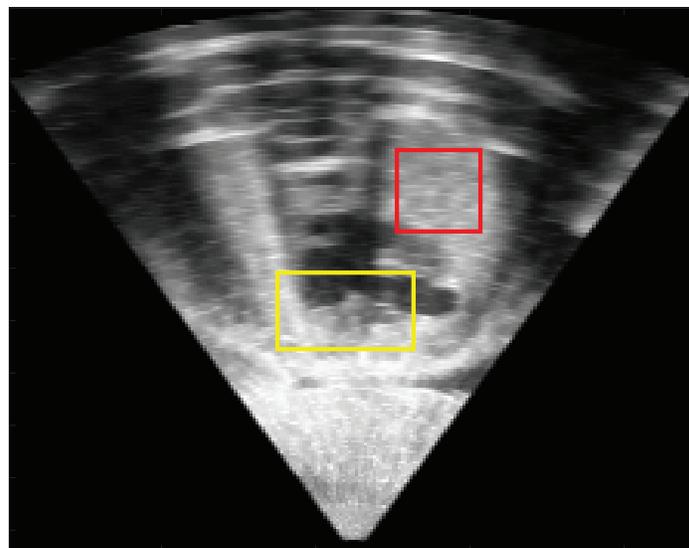
Figure 5.2: Approximate intensity regions of a fetal echocardiographic image. (a) Region A represent cavities (chambers) with blood flow, region B represents soft tissues, and region C represents bony structures. (b) Histogram of the image in (a) showing the gray value counts of the 3 intensity regions.

nonoverlapping subimages and the mean values of these subimages are used as f_v . The transformation of (5.1) is then applied to each of these subimages after which they are recombined to form the larger image. The advantage of using subimages is that they have either a uniform intensity distribution or a bimodal intensity distribution which helps with the identification of potential boundary voxels. Consider a subimage that is from a soft tissue region (red rectangle in Figures 5.3(a) and 5.3(b)). This subimage has an homogeneous or uniform intensity distribution with a mean value representative of this homogeneity. The gradient magnitude of voxels in this region will be approximately zero since their intensity values are close together. Applying condition 2 will give a transformation value close to zero which implies that these voxels are nonboundary voxels. Conversely, consider a subimage that contains a possible boundary between two regions (yellow rectangle in Figures 5.3(a) and 5.3(b)). This subimage has a bimodal distribution with a mean value close to the boundary voxels separating these two regions. Applying condition 2 gives a transformation value that is close to 1 for these boundary voxels and are classified as true boundary voxels.

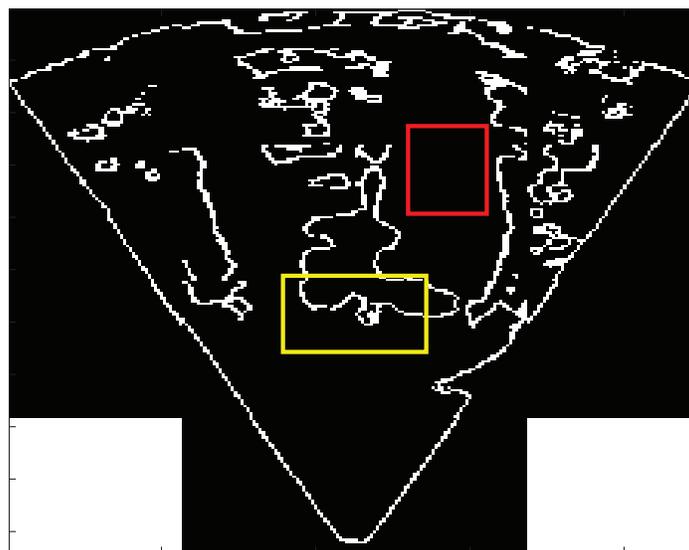
A comparison between the edge detection method using the transformation $\alpha(\mathbf{x})$ in (5.1) and traditional edge detectors such as Canny, Sobel, and Roberts edge detectors are shown in Figure 5.4. The advantage of $\alpha(\mathbf{x})$ over traditional edge detectors is clearly evident with the smaller number of false edges present in the edge profile image. It should also be noted that speckle reducing filtering was not applied to the image which further illustrates the robustness of $\alpha(\mathbf{x})$ to speckle noise.

5.4 Deformable model

The next block in the location estimation procedure is the deformable model block (Figure 5.1). The main use of the deformable model block is to extract the region of interest represented by the boundary edge profile obtained through the method of Section 5.2. Deformable contour models (active contours models) have been around for 25 years when Kass *et al.* [95] introduced an energy minimization approach for contour evolution to locate boundaries of objects in an image. It uses a balance between an internal and external force computed on image data to push a curve or surface towards pertinent image features or boundaries. For this specific application (*i.e.*, location estimation of a fetal heart), the pertinent image feature is the epicardial edge profile estimated in the previous section. Specifically, an initial ellipsoidal surface is deformed towards the identified epicardial edge points (from Section 5.3) such that the final surface represents an enclosure of the fetal heart. This enclosure is then converted into a binary mask where points inside the enclosure is the desired region (foreground) and points outside the enclosure are the irrelevant structures



(a)



(b)

Figure 5.3: A fetal echocardiographic image showing two subimage regions. (a) An homogeneous region (red rectangle) and an inhomogeneous region (yellow rectangle). (b) Results of applying the transformation $\alpha(\mathbf{x})$ to each subimage. The red subimage shows no boundary voxels and the yellow subimage shows boundary voxels.

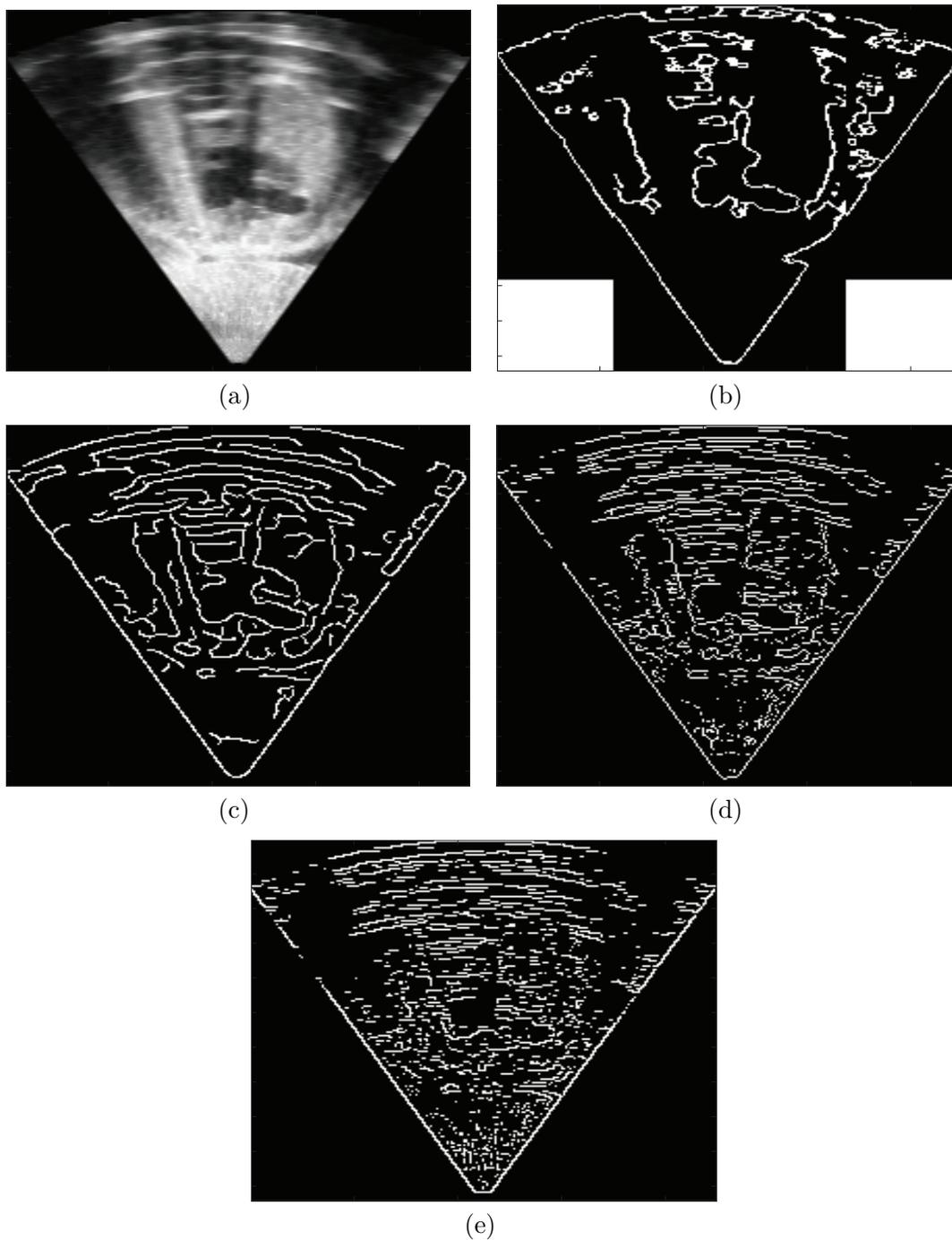


Figure 5.4: Comparison of different edge detectors showing their respective edge profile images. (a) A fetal echocardiographic image. (b) $\alpha(\mathbf{x})$ transformation. (c) Canny edge detector. (d) Sobel edge detector. (e) Roberts edge detector.

present in the echocardiographic image (background).

There is an extensive list of literature material devoted to the various applications and modifications of active contour models some of which are surveyed in [125, 142] and Chapter 3 of this dissertation and will not be repeated here. However, the mathematical formulation of the 3-D active contours used in the location estimation method (Figure 5.1) is derived in the next section.

5.4.1 Derivation of active contours in three dimensions

Definition 5.1. Suppose a 3-D image f to be analyzed is defined as a function that takes a 3-D coordinate in \mathbb{R}^3 and maps it to a single intensity value in \mathbb{R} :

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}$$

Definition 5.2. Let \mathbf{v} be a function (parametric surface) that uses two parameters (s, r) to parameterize a surface in \mathbb{R}^3 such that $(s, r) \subseteq \Omega$ for every $s \in [0, 1]$ and $r \in [0, 1]$. Ω is the parameter subspace in \mathbb{R}^2

$$\begin{aligned} \mathbf{v} : \Omega \rightarrow \mathbb{R}^3 &\equiv [0, 1] \times [0, 1] \rightarrow \mathbb{R}^3 \\ \mathbf{v}(s, r) &= (v_1(s, r), v_2(s, r), v_3(s, r)) \end{aligned}$$

Definition 5.3. Let E be an energy functional that maps a subspace \mathcal{A} containing all possible deformation of the parametric surface \mathbf{v} to \mathbb{R} :

$$E : \mathcal{A} \rightarrow \mathbb{R}$$

With these definitions, active contours is formulated as the minimization of the energy functional E , that is an active contour seeks the surface deformation in subspace \mathcal{A} defined by parameters (s, r) that minimizes an energy functional. According to Kass *et al.* [95], the energy functional E is formulated as:

$$\begin{aligned} E(\mathbf{v}) = \int_{\Omega} \omega_{10} \left| \frac{\partial \mathbf{v}}{\partial s} \right|^2 + \omega_{01} \left| \frac{\partial \mathbf{v}}{\partial r} \right|^2 + 2\omega_{11} \left| \frac{\partial^2 \mathbf{v}}{\partial s \partial r} \right|^2 + \\ \omega_{20} \left| \frac{\partial^2 \mathbf{v}}{\partial s^2} \right|^2 + \omega_{02} \left| \frac{\partial^2 \mathbf{v}}{\partial r^2} \right|^2 + P(\mathbf{v}) ds dr \end{aligned} \quad (5.2)$$

where w_{10} , w_{01} , w_{11} , w_{20} , and w_{02} are weights. $\left| \frac{\partial \mathbf{v}}{\partial s} \right|$ and $\left| \frac{\partial \mathbf{v}}{\partial r} \right|$ are the magnitudes of the first order partial derivatives of the parametric surface \mathbf{v} with respect to s and r while $\left| \frac{\partial^2 \mathbf{v}}{\partial s \partial r} \right|$,

$\left| \frac{\partial^2 \mathbf{v}}{\partial s^2} \right|$, and $\left| \frac{\partial^2 \mathbf{v}}{\partial r^2} \right|$ are the magnitudes of the second order partial derivatives of the parametric surface \mathbf{v} with respect to s and r . $P(\mathbf{v})$ is called the potential energy term. To understand the significance of the terms in (5.2), let us consider an object in a closed system that is acted upon by two forces, an internal force and external force. The equilibrium state for this closed system occurs when the internal forces and external forces are balanced or equal. Similarly, let us consider a parametric surface as the object in the closed system. We can safely say that in this closed system there are both internal and external forces acting on the surface. The internal forces are proportional to the magnitude of the first and second order partial derivatives of the parametric surface \mathbf{v} . Precisely, $\left| \frac{\partial \mathbf{v}}{\partial s} \right|$ and $\left| \frac{\partial \mathbf{v}}{\partial r} \right|$ are proportional to the internal forces that controls the elasticity of the surface. The elasticity of a surface is the ability of the surface to stay compact after deformation and this is modeled as a piecewise smooth regularization term (*i.e.*, the first order partial derivatives of the surface). $\left| \frac{\partial^2 \mathbf{v}}{\partial s \partial r} \right|$, $\left| \frac{\partial^2 \mathbf{v}}{\partial s^2} \right|$, and $\left| \frac{\partial^2 \mathbf{v}}{\partial r^2} \right|$ are proportional to the internal forces that controls the rigidity of the surface. The rigidity or stiffness is property of a surface to resist deformation and this is modeled as the surface curvature (*i.e.*, the second order partial derivatives). The external force in (5.2) is modeled as the force that attracts the parametric surface towards image features and it is proportional to the potential energy $P(\mathbf{v})$ of the parametric surface. In our application the image feature of choice is the boundary edge profile, so, the external force is modeled as the negative gradient of the image (5.3).

$$P(\mathbf{v}) = -|\nabla f(\mathbf{v})| \quad (5.3)$$

Finding the equilibrium state of the parametric surface is synonymous to finding a local minimum of E that fits the desired image feature. This is achieved by initializing the surface in the vicinity of the image feature and iteratively deforming the surface until E is minimized. According to calculus of variations, the surface \mathbf{v} that minimizes the energy functional E satisfies the Euler-Lagrange equation. The Euler-Lagrange equation is a necessary condition for finding a global/local minimum of E and it is given as follows:

$$\begin{aligned}
 & -\omega_{10} \frac{\partial}{\partial s} \left(\frac{\partial \mathbf{v}}{\partial s} \right) - \omega_{01} \frac{\partial}{\partial r} \left(\frac{\partial \mathbf{v}}{\partial r} \right) + 2\omega_{11} \frac{\partial^2}{\partial s \partial r} \left(\frac{\partial^2 \mathbf{v}}{\partial s \partial r} \right) + \omega_{20} \frac{\partial^2}{\partial s^2} \left(\frac{\partial^2 \mathbf{v}}{\partial s^2} \right) \dots \\
 & + \omega_{02} \frac{\partial^2}{\partial r^2} \left(\frac{\partial^2 \mathbf{v}}{\partial r^2} \right) = \nabla P(\mathbf{v})
 \end{aligned} \quad (5.4)$$

In general the energy functional is not a convex function but any local minimum will satisfy (5.4). Equation (5.4) can be solved numerically by constructing an iterative dynamic

system that allows the closed system to evolve toward equilibrium. This involves the introduction of a time variable which vanishes at equilibrium (*i.e.*, $\partial\mathbf{v}/\partial t = \partial^2\mathbf{v}/\partial t^2 = 0$). There are many iterative numerical solvers for solving the Euler-Lagrange equation for a deformable model such as [32, 124], however, an efficient multilevel adaptive finite difference solver proposed by Lurig *et al.* [117] is used because of its suitability to 3-D surfaces. Figure 5.5 shows a typical surface evolution of the numeric solver at different iterations. The input image is a boundary edge profile binary image typical of the image derived from Section 5.3. The deformable model is initialized with an ellipsoid centered in the middle of the image. The resulting regions of interest are shown for iterations $n = 0, 10, 50, 100$ and 200. The progression of the initial parametric surface deformation shown in Figure 5.5 illustrates how the parametric surface is attracted toward the epicardial boundary estimate under the influence of the internal and external forces.

5.5 Morphological filtering

The final step of the location estimation method is the filtering block (Figure 5.1). The input to this block is a cumulative region of interest obtained by aggregating the regions of interest from each frame of a 4-D fetal echocardiographic image. The significance of this is to improve the region of interest estimate by combining multiple boundary edge profiles thereby filling in missing boundary information. The filtering block is a simple morphological filter [99] that is used to remove noise from the aggregated region of interest. Morphological filtering is an image analysis technique that uses nonlinear operations to process an image based on the shape or morphology of a feature. The filtering process involves probing an image with a known shape (sphere, ellipsoid etc.) called a structuring element which is usually chosen according to some prior knowledge about the geometry of the relevant image feature. The filtering procedure involves placing a structuring element at each voxel location in an image and using the relative ordering of the overlapping neighboring voxels to construct an output image. Because morphological filtering relies only on the relative ordering of voxels and not on their numerical values they are especially suited to the processing of binary images. In our application, the morphological filter is used to remove noise arising from the aggregation of the different regions of interest. The noise are in the form of holes or disconnected voxels within the aggregated region of interest so the filter closes these holes by connecting all voxels within the region of interest.

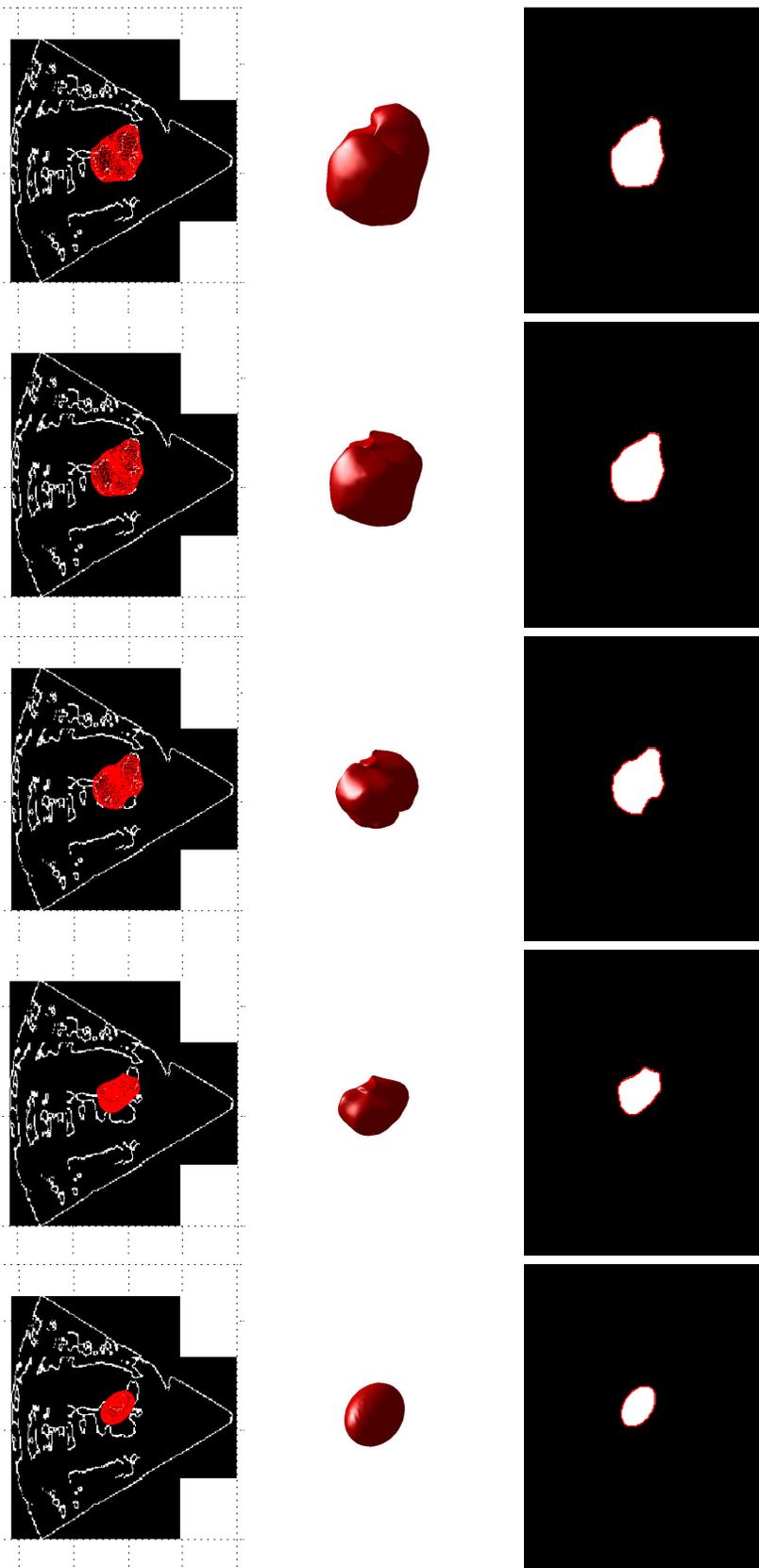


Figure 5.5: Evolution of the 3-D active contours at different iteration n . The iterations $n = 0, 10, 50, 100, 200$ are ordered from left to right, respectively. The top row images are the parametric surfaces superimposed on boundary edge profile images at each iteration n . The middle row images are 3-D view of the deformed parametric surface at each iteration n . The bottom row images are the extracted regions of interest at each iteration n .

5.5.1 Morphological operations

There are four basic operation in morphological filtering [175] and they are 1) erosion, 2) dilation, 3) opening, and 4) closing. There are other sophisticated morphological operations that are used in many image analysis application, however, only the four listed above will be briefly described below. Good references for the theory and applications of mathematical morphology in image analysis can be found in [52, 71, 166, 175, 184].

The morphological operations described next are formulated in term of set theory. A binary image f can thus be represented as the set of all voxel locations in the foreground:

$$f = \{\mathbf{x} | f(\mathbf{x}) = 1\}$$

Definition 5.4. Erosion of an image f by a structuring element h is given by the set operation

$$f \ominus h = \{\mathbf{x} \in \mathbb{Z}^3 | (\mathbf{x} + \mathbf{y}) \in f, \forall \mathbf{y} \in h\}$$

where \mathbb{Z}^3 is the set of all integers in 3-D. The erosion operation can be viewed as keeping only the voxels $\mathbf{x} \in f$ such that h centered at \mathbf{x} fits inside f . Erosion is usually used to shrink structures in an image.

Definition 5.5. Dilation of an image f by a structuring element h is given by the set operation

$$f \oplus h = \{(\mathbf{x} + \mathbf{y}) | \mathbf{x} \in f, \mathbf{y} \in h\}$$

The dilation operation can be viewed as taking the union of copies of the structuring element h , centered at every voxel location \mathbf{x} in the foreground of f . Dilation is usually used to enlarge structures in an image.

Definition 5.6. Opening of an image f by a structuring element h is an erosion of f by h followed by a dilation of f by h

$$f \circ h = (f \ominus h) \oplus h$$

The result of performing an opening operation on an image f is that foreground structures that are smaller than the structuring element h will disappear while larger structures will remain.

Definition 5.7. Closing of an image f by a structuring element h is a dilation of f by h followed by an erosion of f by h

$$f \bullet h = (f \oplus h) \ominus h$$

The result of performing a closing operation on an image f is that holes in the foreground that are smaller than the structuring element h will be filled

From the descriptions above, the morphological operation used for filtering in our application is the closing operation which performs the noise removal desired, that is it removes noise or holes in the aggregated region of interest by filling them.

5.6 Experiments, results, and discussion

To validate the location estimation method described in this chapter, experiments were performed on 254 4-D fetal echocardiographic images. The characteristics of these images are described in Chapter 4 and Appendix A and will not be repeated here.

5.6.1 Experiments

The location estimation method described in this chapter was implemented on a Microsoft[®] Windows 7, 12 Core, 16 GB RAM workstation running MATLAB[®] R2012b. In the edge detection step (Section 5.3), each frame of a 4-D image was divided into nonoverlapping subimages with a size that was one-fifth the original image size. The choice of one-fifth was determined empirically based on the observation that the best edge boundary profiles were generated when the subimage size is on the order of the fetal heart size in the image. In all the 130 images, the fetal heart on average occupied approximately one-fifth of the acoustic (echo) window from measurements performed. The mean of each subimage was then used as f_v , as described in Section 5.3.1, after which they are recombined to obtain the full image edge profile. The integer r describing the transition region around possible edge points was chosen as 3 voxels. From empirical observation, values lower than 3 voxels introduced additional false edges and values greater than 5 voxels introduced gaps in the edge profile estimate which are both undesirable effects. In the deformable model step (Section 5.4), the weights w_{10} , w_{01} , w_{11} , w_{20} , and w_{02} (5.4) were all assigned a value of 1. Also, the parametric curve was initialized as an ellipsoid according to the equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$ where a , b , and c are the semiprincipal axes with values 20, 15, and 30 voxels, respectively. The size of the ellipsoid was of little consequence as long as it was placed in the vicinity of the

epicardial surface to ensure convergence to the appropriate local minimum. Finally, in the morphological filtering step, a sphere of radius 3 voxels was used as the structuring element to fill holes and remove noise from the aggregated region of interest binary image.

5.6.2 Evaluation metrics

The region of interest obtained from the location estimation method was validated by comparing it with the region of interest obtained manually. The manual procedure involves selection of control points that correspond to the fetal heart epicardial boundary. This is done by observing the video of the 4-D images to pinpoint these control points. Four different metrics were used in the comparison and they are: 1) Jaccard index [154], 2) Sørensen–Dice index [46], 3) Sensitivity index [123], and 4) Specificity index [123]. To facilitate the following discussion about these indices let's consider a simple set Venn diagram showing two sets A and B according to Figure 5.6. Set A is the region of interest from the location estimation method and set B is the manual region of interest which is considered as the ground truth. TP is true positive and it represents correctly identified region. TN is true negative and it represents correctly rejected region. FP is false positive and it represents incorrectly identified region. Finally, FN is false negative and it represents incorrectly rejected region. The following definitions of the evaluation metrics can be formulated using TP , TN , FP , and FN .

1. **Jaccard similarity index:** Given two sets, A and B , the Jaccard similarity index $J(A, B)$, which measures the similarity or dissimilarity between the two sets, is defined as follows:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (5.5)$$

In terms of the statistical measures in Figure 5.6, the Jaccard similarity index is given as follows:

$$J(A, B) = \frac{TP}{FP + TP + FN} \quad (5.6)$$

A Jaccard index value of 1 (100%) implies perfect match between sets A and B while a value of 0 (0%) implies sets A and B are disjoint.

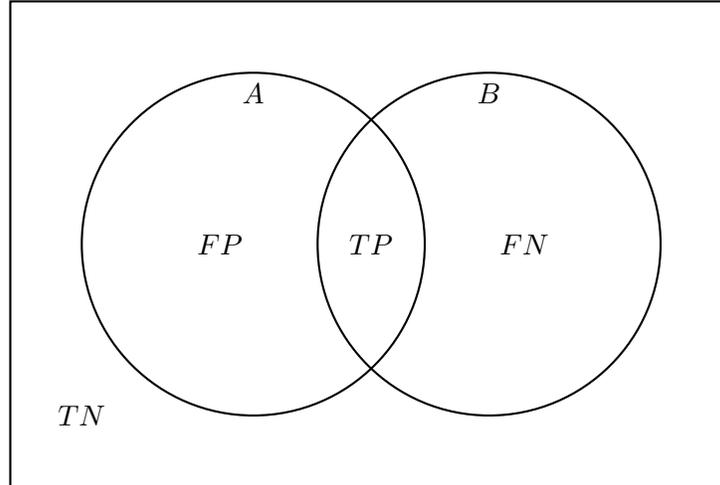


Figure 5.6: A venn diagram showing the statistical regions for comparing the location estimation region of interest with the manual region of interest. Set A is the location estimation region of interest and set B is the manual region of interest.

2. **Sørensen–Dice similarity index:** Sørensen–Dice index is similar to Jaccard index and it is used for measuring the similarity of two sample sets. It is defined as follows:

$$D(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (5.7)$$

where $|A|$ and $|B|$ are the sizes of sets A and B , respectively. In terms of TP , TN , FP , and FN in Figure 5.6, Sørensen–Dice index can also be expressed as follows:

$$D(A, B) = \frac{2 \times TP}{FP + 2 \times TP + FN} \quad (5.8)$$

A Sørensen–Dice index value of 1 (100%) implies perfect match between sets A and B while a value of 0 (0%) implies sets A and B are disjoint.

3. **Sensitivity index:** In a binary system with positive samples and negative samples, the Sensitivity index is the fraction of positive samples that are correctly detected. In terms of its use as a validation metric in the location estimation experiments, the positive samples are the samples of the location estimation method that falls within the ground truth region of interest, while the negative samples are the samples that fall outside the ground truth region of interest. The Sensitivity index can be formulated in terms of TP and FN (Figure 5.6) as follows:

$$ST = \frac{TP}{TP + FN} \quad (5.9)$$

The closer the Sensitivity index value is to 1 (100%) the better the performance of the location estimation method at identifying the fetal heart region of interest.

4. **Specificity index:** In a binary system with positive samples and negative samples, the Specificity index is the fraction of negative samples that are correctly detected. In terms of its use as a validation metric in the location estimation experiments, the positive samples are the samples of the location estimation method that falls within the ground truth region of interest, while the negative samples are the samples that fall outside the ground truth region of interest. The Specificity index can be formulated in terms of TN and FP (Figure 5.6) as follows:

$$SP = \frac{TN}{TN + FP} \quad (5.10)$$

The closer the Specificity index value is to 1 (100%) the better the location estimation method is at discriminating between the fetal heart region and other regions in a 4-D echocardiographic image.

5.6.3 Results and discussion

The average Jaccard index of the location estimation method, for all 130 images, with respect to the manual ground truth was measured at 80.70%. The Sørensen-Dice index produced an average similarity value of 89.19% while the average Sensitivity and Specificity indices were measured at 91.04%, and 99.17%, respectively.

Figure 5.7 summarizes the statistics for each evaluation index for all the images in a box and whisker plot. The Jaccard index had a minimum value of 68.41%, a maximum value of 93.20%, a 25th percentile value of 74.86%, and a 75th percentile value of 85.55%. The Sørensen-Dice index had a minimum value of 81.24%, a maximum value of 96.48%, a 25th percentile value of 85.62%, and a 75th percentile value of 92.21%. The Sensitivity index had a minimum value of 71.08%, a maximum value of 99.82%, a 25th percentile value of 88.25%, and a 75th percentile value of 95.81%. The Specificity index had a minimum value of 95.64%, a maximum value of 100%, a 25th percentile value of 98.79%, and a 75th percentile value of 99.82%. A complete breakdown of the evaluation indices for each image can be found in Appendix B.

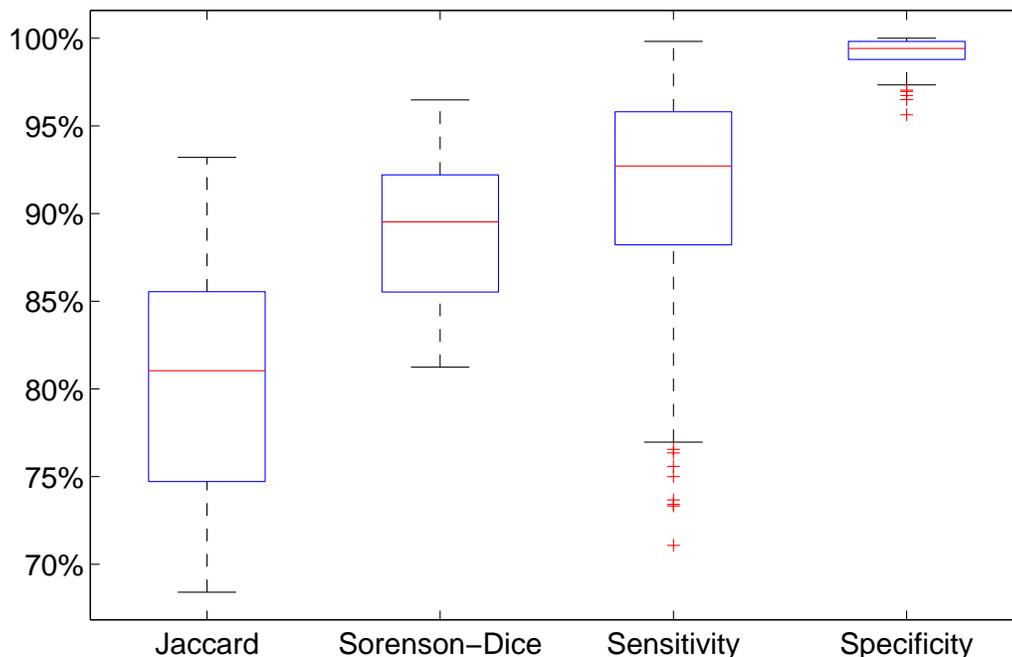


Figure 5.7: Box and whisker plot showing the statistics of the Jaccard, the Sørensen-Dice, the Sensitivity, and the Specificity indices for all 254 4-D fetal echocardiographic images.

The Jaccard index, the Sørensen-Dice index, and the Sensitivity index values show that on average the region obtained manually and the region obtained through the location estimation method have greater than 80% overlap. A reason why this is not closer to 100% is because in some images the false edge suppression of the transformation $\alpha(\mathbf{x})$ was inadequate. These erroneous edges caused the deformable model to overestimate or underestimate the expected boundary surface thereby making the region obtained either smaller or larger. The inadequate false edge suppression can be attributed to the choice of the threshold f_v . Currently, f_v is chosen adaptively by dividing an image into subimages with sizes equal to one-fifth the full image size. The one-fifth value was chosen empirically based on visual inspection of the 130 images. A more quantitative approach that seeks to measure the amount of false edges present in an image could be a better alternative, although how this can be achieved without some prior knowledge of the fetal heart location and size is not clear. Perhaps an optimization framework that seeks to find the optimal weights for a linear or nonlinear combination of voxels representing false edges and voxels representing true edges, with well defined constraints, could be explored. The Specificity index average value of 99.17% is an important result obtained from the experiments because

it shows how well the location estimation method can discriminate the fetal heart from other structures in a 4-D fetal echocardiographic image which is the main reason behind the location estimation component of the detection system.

The deformable model, which was used to extract a closed region bounded by the epicardial surface estimates, could be improved by automating the initial surface placement. In the experiments, this initialization was performed manually. This was necessitated by the fact that deformable models are very sensitive to initialization [205] and also because of the presence of some residual false edges. An initialization that is not close to the target will cause an underestimation or overestimation of the desired fetal heart region. Automation of the initial surface placement is a desired future extension of this method. Shan *et al.* [170] described a method for automatic seed point placement to extract the region of breast cancer lesions. Their method used the homogeneity of breast cancer lesion to define a threshold for placing a seed point in an image. This method could be explored as a way of automating the deformable model initialization although it is not immediately clear how their method can be extended to nonhomogenous structures like the fetal heart.

Additional analysis was performed on the location estimation method based on image and fetal heart characteristics such as the gestational age (Table 5.1), the transducer type used to acquire the images (Table 5.2) and the ultrasound image acquisition mode (Table 5.3). Table 5.1 shows the evaluation metric results based on the fetus' gestational age. The average values for each index at different gestational showed no significant variation which implies that the location estimation method is robust to gestational age. This is an important characteristic since a location estimation method that can work with different fetal heart sizes is desired.

Table 5.2 shows the evaluation results according to the transducer type used. The two transducers used to acquire the images are the *X5-1* and *X7-2* matrix transducer. The main difference between these two transducers is the number of piezoelectric elements; 2500 elements in *X7-2* and 3040 elements in *X5-1*. The number of elements controls the spatial resolution of the acquired images which helps in identifying small structures. The transducer type should provide an advantage if smaller structures (*e.g.*, heart valve or chambers) are being considered however the fetal heart region as a whole is several wavelengths larger than the incoming ultrasound beams so it is expected that the transducer type should not affect the results as confirmed by Table 5.2.

Table 5.3 show the evaluation results based on the full volume acquisition mode, and the Live 3D acquisition mode. In terms of image quality and spatial resolution, there is no significant difference between the two acquisition modes [20]. The main difference between

Table 5.1: A table showing the location estimation results according to fetal gestational ages in weeks. The 25th percentile, the average, and the 75th values are shown for each evaluation metric. All values are percentages.

| Gest. Age | Jaccard | | | Sørenson-Dice | | | Sensitivity | | | Specificity | | |
|-----------|------------------|------|------------------|------------------|------|------------------|------------------|------|------------------|------------------|------|------------------|
| | 25 th | Avg. | 75 th |
| [0, 22) | 78.3 | 81.9 | 85.4 | 87.8 | 89.9 | 92.1 | 90.4 | 92.9 | 97.0 | 99.2 | 99.4 | 99.8 |
| [22, 24) | 76.4 | 81.4 | 86.8 | 86.6 | 89.6 | 92.9 | 88.0 | 89.9 | 95.2 | 99.2 | 99.4 | 99.8 |
| [24, 26) | 73.9 | 79.7 | 84.3 | 85.0 | 88.6 | 91.5 | 88.4 | 91.3 | 95.2 | 98.6 | 99.0 | 99.7 |
| [26, 28) | 76.6 | 81.2 | 86.8 | 86.7 | 89.6 | 93.0 | 89.9 | 91.4 | 96.2 | 99.3 | 99.4 | 99.9 |
| [28, 30) | 76.6 | 80.6 | 84.3 | 86.8 | 89.2 | 91.5 | 81.2 | 85.9 | 90.5 | 99.4 | 99.3 | 100.0 |
| [30, ∞) | 77.5 | 81.4 | 87.0 | 87.3 | 89.6 | 93.1 | 92.5 | 92.3 | 95.7 | 98.4 | 98.9 | 99.5 |

Table 5.2: A table showing the location estimation results according to the transducer type used to acquire the images. The 25th percentile, the average, and the 75th values are shown for each evaluation metric. All values are percentages

| Trans. Type | Jaccard | | | Sørenson-Dice | | | Sensitivity | | | Specificity | | |
|-------------|------------------|------|------------------|------------------|------|------------------|------------------|------|------------------|------------------|------|------------------|
| | 25 th | Avg. | 75 th |
| X5-1 | 74.4 | 80.2 | 84.7 | 85.3 | 88.9 | 91.7 | 88.6 | 91.2 | 95.4 | 98.3 | 98.8 | 99.7 |
| X7-2 | 76.0 | 81.0 | 85.8 | 86.4 | 89.4 | 92.4 | 88.2 | 91.0 | 95.8 | 99.1 | 99.4 | 99.9 |

Table 5.3: A table showing the location estimation results according to the ultrasound system acquisition mode. The 25th percentile, the average, and the 75th values are shown for each evaluation metric. FV refers to full volume acquisition, and L3D refers to live 3-D acquisition. All values are percentages

| Acq. Mode | Jaccard | | | Sørenson-Dice | | | Sensitivity | | | Specificity | | |
|-----------|------------------|------|------------------|------------------|------|------------------|------------------|------|------------------|------------------|------|------------------|
| | 25 th | Avg. | 75 th |
| FV | 76.3 | 81.2 | 86.7 | 86.6 | 89.5 | 92.9 | 89.8 | 92.1 | 96.0 | 99.0 | 99.3 | 99.8 |
| L3D | 73.8 | 80.0 | 84.7 | 84.9 | 88.7 | 91.7 | 86.0 | 89.6 | 95.2 | 98.3 | 98.9 | 99.8 |

the modes is that the full volume mode is derived from multiple (4 or 7) live 3-D volumes. The evaluation metric values in Table 5.3 show that there is no significant mismatch in the performances of the two acquisition modes as expected.

Figure 5.8 shows location estimation results for some images from the dataset (Four chamber views shown). The green outline depicts the boundary from the manual region of interest and the red outline depicts the boundary from the location estimation method’s region of interest.

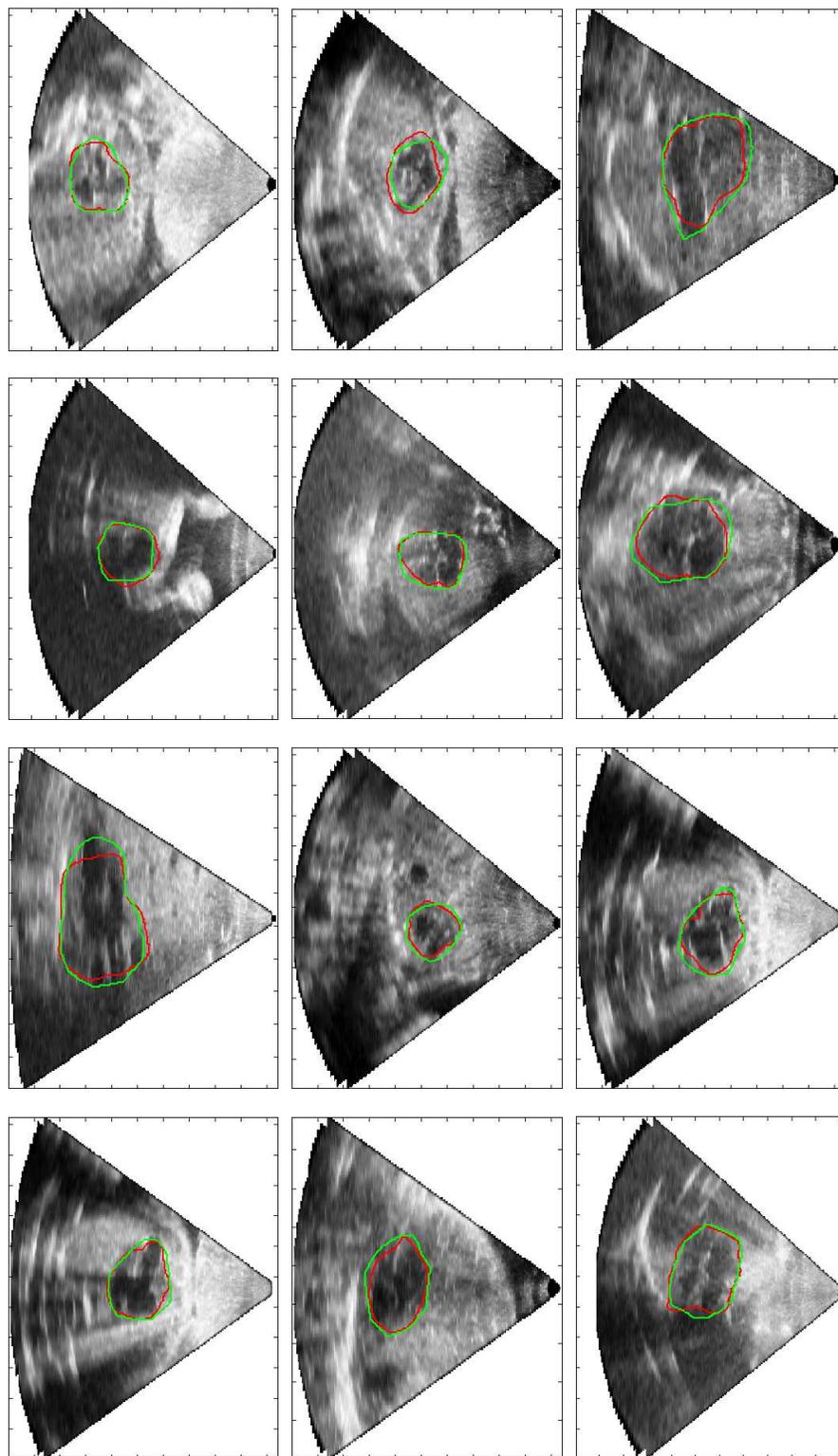


Figure 5.8: Four chamber view showing the region of interest boundary outline of the location estimation method (red) and the boundary outline of the manual region of interest (green).

5.7 Conclusion

This chapter presented a location estimation method that isolates the fetal heart region in a 4-D echocardiographic image. The process involves using an edge detector to estimate the epicardial surface of a fetal heart. This surface estimate is then used to guide a deformable model to obtain a region of interest containing the fetal heart. Results shows an average region agreement of greater than 80% when compared with manually obtained region of interest. The region of interest obtained in this chapter is used in the Chapter 6 where a segmentation method for identifying the individual chambers of a fetal heart is applied on the localized fetal heart (region).

CHAPTER 6

SEGMENTATION USING VELOCITY VECTOR FIELDS

6.1 Abstract

A motion based segmentation method for identifying individual fetal heart chambers in a 4-D echocardiographic image is presented. A developing fetal heart at 20 weeks gestational age is a small structure, even smaller are the individual fetal heart chambers which makes it difficult to identify them using boundary based segmentation techniques particularly in the presence of speckle noise. A segmentation method that uses both physical properties of the image and motion properties of the fetal heart can enhance the individual chamber identification process even in the presence of speckle noise and missing boundary information. The time dimension of a 4-D echocardiographic image contains motion information which is leveraged as a main driving force behind the described segmentation method. The method is composed of two steps: 1) the motion estimation step, 2) the segmentation step. The motion estimation step uses an optical flow method based on the local statistics around each image voxel. This help to reduce the fluctuation associated with speckle noise in the computation of the motion vectors. The segmentation step is based on the kernel k -means clustering algorithm in a high-dimensional nonlinear subspace using physical features extracted from the local image characteristics, and features extracted from the motion vector field estimates. The combination of these features helps with the identification of the small fetal heart chambers by providing a discriminative constraint for identifying them in the high-dimensional nonlinear subspace. The segmentation method was validated on 130 images obtained from 31 patients and compared with manually identified chambers. Validation experiments were based on three metrics, namely: 1) Sørensen-Dice index, 2) absolute volume difference, and 3) Hausdorff distance. The average values of these metrics across all 130 images were measured as 69.52%, 22.64%, and 2.89 mm, respectively. However, when the evaluation metrics are considered on a per patient basis, the average values improved to 79.40%, 17.48%, and 2.19 mm, respectively.

6.2 Introduction

Segmentation of medical images is an area of active and ongoing research. Various methods and techniques have been described in literature with many being specific to particular applications or imaging modality. Fetal echocardiography has not enjoyed the same level of interest as it pertains to segmentation. This is largely due to the small size of the fetal heart and because of speckle noise and acoustic artifacts inherent in ultrasound images which contribute to imaged structures having boundaries that are not always well defined. Recent advances in ultrasound imaging and transducer technologies have made the use of 4-D echocardiography for diagnosis commonplace, however, the image analysis tools accompanying the improved technology has not kept pace. In particular, the detection rate of congenital heart defects could be greatly improved if there are segmentation methods that can be applied to these 4-D images to quantify fetal heart chambers. One of the few published work on fetal echocardiographic segmentation is by Dindoyal *et al.* [51], where they described a level sets approach to identifying fetal heart chambers in 2-D and 3-D fetal echocardiographic images. They used a level set formulation by Sarti *et al.* [161] which used mean curvature and edge flow diffusion to incorporate missing boundaries in the level set evolution equation. They reported a volume error of 13% using a phantom of a fetal heart. Navaux *et al.* described a 2-D fetal segmentation method using neural networks and k -means clustering. They used neural network to select sample features that approximated the underlying probability distribution of the fetal heart chamber classes. K -means was then used to organize clusters according to this distribution. Lassige *et al.* [103] described a level set method for identifying septal defects in 2-D and 3-D echocardiographic images. They start by initializing their algorithm with manually defined seed points and allowing the level sets to converge to the borders of the septal defect. They reported 85% success rate in their trials. A drawback of their method is the potential for boundary overshoot because of the use of a constant speed term in the level set formulation. Tutschek *et al.* [195] described an approach for segmenting the endocardium (inner heart wall). They selected seed points manually and allowed it to grow towards the endocardium border by manually controlling the individual chamber volumes with a distance threshold. Their method is operator dependent and can only handle static volumes.

In this chapter a segmentation method is described that utilizes the time dimension of a 4-D echocardiographic image to identify the individual fetal heart chambers. During a typical cardiac cycle, the fetal heart structures have specific motion characteristics in the direction of the blood flow. During the diastolic cardiac phase, the motion is from the right

atrium to the right ventricle, and from the left atrium to the left ventricle. During the systolic cardiac phase, the motion is from the right ventricle to the pulmonary artery, and from the left ventricle to the aorta. The main structures that characterize the diastolic and systolic cardiac phases are the tricuspid valve, the mitral valve, the pulmonary valve and the aortic valves. These structures are periodically opening and closing throughout a cardiac cycle and voxels representing these structures will have motion mimicking the opening and closing of these valves. The combination of the motion characteristics of the valves and the cavities (chambers) gives a good discriminating feature than can be used to group voxels according to which chamber they belong. Figure 6.1 is a block diagram describing the procedures involved in the segmentation method. A 4-D fetal echocardiographic image $f(\mathbf{x})$ is passed into a motion estimation block where the motion vector fields (displacement of voxels) are calculated for each of the spatial dimensions using an histogram-based optical flow technique proposed by Tenbrinck *et al.* [192]. Their method uses a constraint based on the intensity histogram in a neighborhood around each voxel to reduce the effect of multiplicative speckle noise. The estimated motion vectors \mathbf{u}_i and the image $f(\mathbf{x})$ are then passed to the segmentation block. An expansion of the segmentation block in Figure 6.1(a) is shown in Figure 6.1(b) where the feature selection block is used to identify the features from the image $f(\mathbf{x})$ and the motion vectors \mathbf{u}_i that best discriminate the individual chambers. The feature matrix \mathbf{G} is then passed to the nonlinear transformation function (kernel function) $\Phi(\mathbf{G}) \cdot \Phi(\mathbf{G})$ producing a kernel matrix \mathbf{K} . The importance of this is to create better separability of the feature vectors by representing the data in a high-dimensional nonlinear subspace. Kernel k -means algorithm is then applied to the transformed feature matrix to obtain clusters which represent the fetal heart chambers.

The rest of this chapter is organized as follows: Section 6.3 describes the algorithm for obtaining the motion vector estimates. Section 6.4 covers the segmentation method describing how the features are selected, the choice of the nonlinear kernel function, and the kernel k -means algorithm. Section 6.5 describes the validation experiments performed and the results from these experiments. Section 6.6 concludes this chapter, summarizing the key finding of the motion segmentation method.

6.3 Motion estimation

Motion estimation or optical flow is defined as the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer (reference) and the scene. Alternatively, it can be described as the process of determining motion vectors that describe the transformation of an object in a sequence of images or

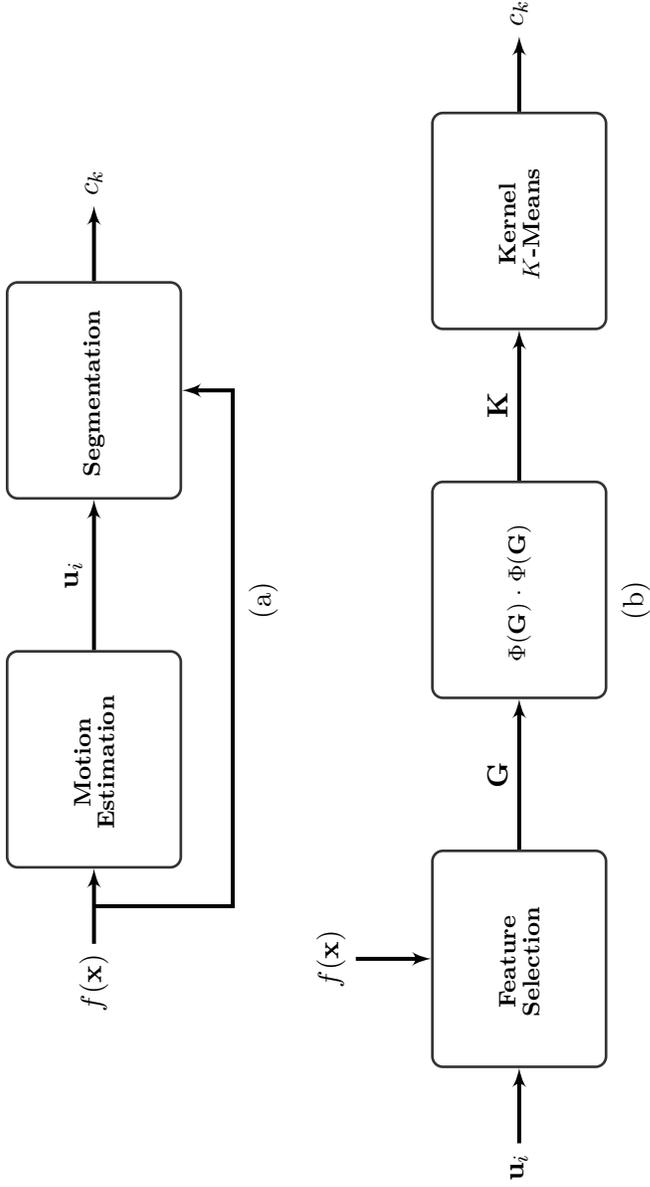


Figure 6.1: A block diagram of the segmentation method. (a) A block diagram of the complete segmentation method with an input 4-D image $f(\mathbf{x})$ into a motion estimation block that computes the motion vectors \mathbf{u}_i in direction i . The motion vectors and the input image are then passed into a segmentation block. (b) A block showing the components of the segmentation block in (a). The motion vectors \mathbf{u}_i and the input image $f(\mathbf{x})$ are passed into a feature selection block to get the feature matrix \mathbf{G} . The feature matrix undergoes a transformation $\Phi(\mathbf{G}) \cdot \Phi(\mathbf{G})$ into a high-dimensional nonlinear space producing a kernel matrix \mathbf{K} . Kernel k -means is then employed to extract the clusters c_k (in the nonlinear subspace) which correspond to the individual fetal heart chambers.

video. There are many ways of computing the optical flow field in literature, however, methods based on the variational approach are among the best performing techniques [9, 66]. Variational methods of computing optical flow fields are based on the minimization of a continuous energy functional $E(\mathbf{u})$ which is formulated as the sum of a data term and a regularization term (6.1). The idea of the energy functional formulation is to find the velocity estimate that best fits the motion characteristics of an object in an image at time t relative to its initial position or velocity in a reference image at $t = 0$. According to optimization theory [69], the solution to this energy optimization problem occurs when the energy of the system, defined by the motion estimates, is minimized. The general formulation of the energy functional is given as:

$$E(\mathbf{u}) = \int_{\Omega} \underbrace{M(D^k f, \mathbf{u})}_{\text{data term}} + \underbrace{\alpha S(\nabla \mathbf{u})}_{\text{regularization}} d\Omega \quad (6.1)$$

where the integration domain Ω is either a spatial domain or a spatiotemporal domain. ∇ is the partial derivative operator, f is the image containing the moving object, \mathbf{u} is the optical flow field (velocity vectors), and $D^k f$ is the partial (spatial and temporal) derivatives of f of order k . The data term $M(D^k f, \mathbf{u})$, which is a function of the k^{th} partial derivative of f and the optical flow field \mathbf{u} , is usually based on a brightness or intensity constancy assumption. More on this later. The regularization term $S(\nabla \mathbf{u})$ penalizes optical flow fields that deviates from this assumption (*i.e.*, it penalizes deviations from piecewise smoothness). A weight α is used to control the degree to which the deviations are penalized and it has a value greater than zero ($\alpha > 0$).

The optical flow algorithm which introduced the variational approach was first presented by Horn and Schunck [83]. Their algorithm is the foundation for many variational methods including the histogram based method used to perform fetal heart segmentation in this chapter. Because of its importance, Horn and Schunck's method is described next.

6.3.1 Horn-Schunck method

Definition 6.1. Suppose the motion field \mathbf{u} of an image f in an image subspace Ω is to be computed. The image f can be viewed as a function of the position $\mathbf{x} \in \mathbb{R}^n$ and time $t \in \mathbb{R}$ that gives an intensity value.

$$\begin{aligned} f &: \mathbb{R}^{n+1} \rightarrow \mathbb{R} \\ \mathbf{x} &\in \mathbb{R}^n, t \in \mathbb{R} \end{aligned}$$

The Horn-Schunck method for computing optical flow fields is based on two critical assumption [186], which are:

- **Brightness constancy constraint:** The observed brightness of any object point is constant over time.
- **Velocity smoothness constraint:** Nearby points in the image plane move in a similar manner.

The brightness constancy assumption can be interpreted mathematically as follows:

$$f(\mathbf{x}, t) = f(\mathbf{x} + \Delta\mathbf{x}, t + \Delta t) \quad (6.2)$$

where $\Delta\mathbf{x}$ is a small change in the position \mathbf{x} , and Δt is a small change in time t . The significance of (6.2) is that if a small enough change in position $\Delta\mathbf{x}$ and time Δt is considered, the intensity of the image can be assumed to be constant. A Taylor's series expansion of (6.2) gives,

$$f(\mathbf{x} + \Delta\mathbf{x}, t + \Delta t) = f(\mathbf{x}, t) + \nabla f(\mathbf{x}, t) \cdot \Delta\mathbf{x} + \frac{\partial}{\partial t} f(\mathbf{x}, t) \Delta t + \dots \quad (6.3)$$

where $\nabla f(\mathbf{x}, t) \cdot \Delta\mathbf{x}$ is the dot product of the partial derivatives (spatial) of the image $f(\mathbf{x}, t)$ with the small change in position $\Delta\mathbf{x}$. A further approximation can be made to (6.3) by only considering the linear terms, that is,

$$f(\mathbf{x} + \Delta\mathbf{x}, t + \Delta t) \approx f(\mathbf{x}, t) + \nabla f(\mathbf{x}, t) \cdot \Delta\mathbf{x} + \frac{\partial}{\partial t} f(\mathbf{x}, t) \Delta t \quad (6.4)$$

For (6.4) to satisfy the brightness constancy assumption of (6.2), the sum of terms involving the first order derivatives must be zero, *i.e.*,

$$\nabla f(\mathbf{x}, t) \cdot \Delta\mathbf{x} + \frac{\partial}{\partial t} f(\mathbf{x}, t) \Delta t = 0 \quad (6.5)$$

This can be rewritten as,

$$\nabla f(\mathbf{x}, t) \cdot \mathbf{u} = -f_t(\mathbf{x}, t) \quad (6.6)$$

where \mathbf{u} is the velocity vector, and f_t is equivalent to $\frac{\partial}{\partial t} f(\mathbf{x}, t)$, the partial derivative of the image $f(\mathbf{x}, t)$ with respected to time t . Equation (6.6) is not sufficient for solving

the velocity vectors completely because it only provides the normal component in the direction of the brightest gradient [186]. Horn-Schunck therefore proposed the velocity smoothness assumption which allows solutions to be crafted for both the normal and tangential component of the velocity vector. The smoothness constraint assumes that the velocity vector field is only changing slowly in a given neighborhood. The smoothness constraint S is given as,

$$S = \nabla \cdot \mathbf{u} \quad (6.7)$$

where $\nabla \cdot \mathbf{u}$ is the dot product of the partial derivative operator ∇ with the velocity vector \mathbf{u} . Based on the two assumptions derived, the energy functional for the Horn-Schunck method follows the general structure of (6.1) where the data term is represented as (6.6) and the regularization term as (6.7).

$$E(\mathbf{u}) = \int_{\Omega} |\nabla f(\mathbf{x}, t) \cdot \mathbf{u} + f_t(\mathbf{x}, t)|^2 + \alpha |\nabla \cdot \mathbf{u}|^2 d\Omega \quad (6.8)$$

where Ω is the image space, and α is a weight. Minimization of this functional can be performed iteratively using the Gauss-Siedel and Euler-Lagrange equations [83, 186].

6.3.2 Histogram-based Horn-Schunck method

The brightness constancy assumption, introduced in the previous section, is one of the two constraints used to formulate a solution for the optical flow problem. However, in the presence of noise this assumption may be violated. When an image is corrupted with additive noise, the brightness constancy constraint can still be used by choosing an appropriate regularization term that suppresses the influence of noise, however, when the noise model is signal dependent, as is the case with speckle noise, the correlation between pixel brightness values can lead to inaccurate velocity vector fields. Tenbrinck *et al* [192] proposed a modified Horn-Schunck method that can be used when an image is corrupted with speckle noise. They proved that a bias is introduced in the velocity vector field computation when Horn-Schunck's method is used in the presence of signal dependent noise, and they further showed that this bias is suppressed with their method. The general idea of their method is to replace the brightness constancy assumption with a local statistics constancy assumption, for which they used the histogram of a neighborhood around each point. Section 6.3.2.1 is a statement and proof of the theorem demonstrating the bias

of Horn-Schunck's method, and Section 6.3.2.2 shows the proof of bias suppression of Tenbrinck's method.

6.3.2.1 Bias in Horn-Schunck method

Definition 6.2. Consider an ultrasound image $f(\mathbf{x})$ formed according to an image formation model characterized by a signal dependent speckle noise η_m and an additive noise η_a . The speckle noise η_m depends on the acoustic signal $g(\mathbf{x})$ and is controlled by a parameter γ which defines the degree of signal dependence [112, 159]. The additive noise η_a is assumed to be Gaussian distributed with mean zero and variance σ^2 . Thus, the image formation model is,

$$\begin{aligned} f(\mathbf{x}) &= g(\mathbf{x}) + \eta_a \cdot \eta_m(\mathbf{x}) \\ \eta_m &= g(\mathbf{x})^{\frac{\gamma}{2}} \\ \eta_a &\sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

Theorem 6.1. Let $\gamma \geq 0$. Let $X, Y \in \mathbb{R}^n$ be random vectors with each component $X_i, Y_i, i = 1, 2, \dots, n$, i.i.d with constant (unbiased) image intensities μ_X and μ_Y , respectively. According to the noise model in Definition 6.2, the energy can be defined as the L_2 -norm of the error between X and Y . That is,

$$E = |X - Y|^2$$

The expected value of E attains its global minimum if and only if the following relationship holds,

$$\mu_X = \frac{\gamma}{2} \mu_Y^{\gamma-1} + \mu_Y$$

Proof. Without loss of generality, let the variance of the additive noise term be $\sigma^2 = 1$. According to the image formation model, the random variables X_i, Y_i are normally distributed with

$$\begin{aligned} X_i &\sim \mathcal{N}(\mu_X, (\mu_X)^\gamma) \\ Y_i &\sim \mathcal{N}(\mu_Y, (\mu_Y)^\gamma) \end{aligned}$$

Since the energy is defined as an L_2 -norm (convex function), a global minimum is guaranteed. Finding the expectation of the energy and using the identity, $\text{var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$, we have:

$$\begin{aligned}\mathbb{E}[E] &= \mathbb{E}[|X - Y|^2] \\ &= \sum_{i=1}^n \mathbb{E}[X_i^2] - 2\mathbb{E}[X_i Y_i] + \mathbb{E}[Y_i^2] \\ &= n \left((\mu_X)^\gamma + (\mu_X)^2 - 2\mu_X \mu_Y + (\mu_Y)^\gamma + (\mu_Y)^2 \right)\end{aligned}$$

The minimum of the expected value of the energy is found by fixing the parameter μ_X and finding the value of μ_Y that minimizes $\mathbb{E}[E]$, that is:

$$\arg \min_{\mu_Y \geq 0} n \left((\mu_X)^\gamma + (\mu_X)^2 - 2\mu_X \mu_Y + (\mu_Y)^\gamma + (\mu_Y)^2 \right)$$

To find the minimum of $\mathbb{E}[E]$, the partial derivative with respect to μ_Y is computed and set to zero.

$$\begin{aligned}\frac{\partial}{\partial \mu_Y} \mathbb{E}[E] &= \frac{\partial}{\partial \mu_Y} \left[n \left((\mu_X)^\gamma + (\mu_X)^2 - 2\mu_X \mu_Y + (\mu_Y)^\gamma + (\mu_Y)^2 \right) \right] \\ 0 &= -2n\mu_X + n\gamma(\mu_Y)^{\gamma-1} + 2n\mu_Y \\ \mu_X &= \frac{\gamma}{2}(\mu_Y)^{\gamma-1} + \mu_Y\end{aligned}$$

□

From the proof of this theorem, two things can be deduced

Corollary. *For two pixel patches X , Y with the same constant intensity values (i.e., $\mu_X = \mu_Y$) perturbed by noise according to the image formation model of Definition 6.2 the following condition holds:*

1. *The expected value of the L_2 -distance of X and Y , i.e., $|X - Y|^2$, is unbiased with a minimum of zero if and only if the data are perturbed by Gaussian noise, i.e., $\gamma = 0$.*
2. *For multiplicative speckle noise, i.e., $\gamma \geq 0$, the expected value of the L_2 -distance of X and Y is biased with a nonzero minimum which introduces errors in the computation of the velocity vector fields*

6.3.2.2 Bias suppression with local statistics

The bias associated with the brightness constancy assumption can be suppressed by using the local statistics in a neighborhood of a point. Tenbrinck *et al.* [192] proposed using the cumulative histogram of image patches and they called this the histogram constancy assumption. The idea stems from the fact that the speckle characteristic in an ultrasound image is dependent on the tissue properties. Though a single pixel can be altered by speckle across two images, the overall distribution within a local image region remains approximately constant because the tissue properties remain unchanged. Cumulative histogram is a good way of capturing this constancy in tissue properties even with speckle corruption. The histogram constancy assumption is defined in a similar way as the brightness constancy assumption of (6.2)

$$H(\mathbf{x}, t) = H(\mathbf{x} + \Delta\mathbf{x}, t + \Delta t) \quad (6.9)$$

where \mathbf{x} is the spatial position, t is time, $\Delta\mathbf{x}$ is a small change in position and Δt is a small change in time. H represents the cumulative histogram of the neighborhood around a central pixel \mathbf{x} . Mathematically, H is defined as:

$$H(\mathbf{x})[k] = \sum_{j=1}^n \mathbb{1}_{[f(\mathbf{x}_j) \leq k]} w(\mathbf{x}_j) \quad (6.10)$$

where $H(\mathbf{x})[k]$ represents the histogram value in bin k of an image patch centered at position \mathbf{x} . k represents the predetermined bin value such that all neighbors with an intensity value less than or equal to k are counted as a member of bin k . $f(\mathbf{x}_j)$ is the intensity value of the j^{th} neighbor, n is the size of the neighborhood around the central pixel inclusive and $w(\mathbf{x}_j)$ is the weight for the j^{th} neighbor such that $\sum_{j=1}^n w(\mathbf{x}_j) = 1$. $\mathbb{1}$ is called the indicator function. It is defined for a set $A(x)$ as follows:

$$\mathbb{1}_{[A(x)]} = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \quad (6.11)$$

Using Definition 6.2, the following is the theorem that justifies the use of cumulative histogram in combination with the L_2 -distance norm [192].

Theorem 6.2. *Let $\gamma \geq 0$. Let $X, Y \in \mathbb{R}^n$ be random vectors with each component $X_i, Y_i, i = 1, 2, \dots, n$, i.i.d with constant (unbiased) image intensities μ_X and μ_Y , respectively.*

According to the noise model in Definition 6.2, the energy can be defined as the L_2 -norm of the error between $H(X)$ and $H(Y)$.

$$E_n = |H(X) - H(Y)|^2$$

The expected value of E_n attains its global minimum for sufficiently large n if and only if $\mu_X = \mu_Y$, where n is the size of the image patch.

Proof. Let the variance of the additive noise term be $\sigma^2 = 1$. Let the number of histogram bins be k . According to the image formation model, the random variables X_i, Y_i are normally distributed with

$$\begin{aligned} X_i &\sim \mathcal{N}(\mu_X, (\mu_X)^\gamma) \\ Y_i &\sim \mathcal{N}(\mu_Y, (\mu_Y)^\gamma) \end{aligned}$$

For simplicity, let the spatial weights $w(\mathbf{x}_j)$, $j = 1, 2, \dots, n$ in (6.10) be $1/n$. This makes the cumulative histogram H a cumulative distribution function (CDF). Since the energy E_n is defined as an l^2 -norm (convex function), a global minimum is guaranteed. The expectation of the energy can be computed by using the identity, $\text{var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. Also, since cumulative histogram is represented as a CDF, the expected value of the indicator function is given by the identity $\mathbb{E}[\mathbb{1}_{[X \leq t]}] = \mathbb{P}(X \leq t)$,

$$\begin{aligned} \mathbb{E}[E_n] &= \mathbb{E}[|H(X) - H(Y)|^2] \\ &= \frac{1}{n^2} \sum_{i=1}^k \left[n(n-1) \left(\mathbb{P}(X_1 \leq i)^2 + \mathbb{P}(Y_1 \leq i)^2 \right) \right. \\ &\quad \left. - 2n^2 \mathbb{P}(X_1 \leq i) \mathbb{P}(Y_1 \leq i) \right. \\ &\quad \left. + n \left(\mathbb{P}(X_1 \leq i) + \mathbb{P}(Y_1 \leq i) \right) \right] \end{aligned}$$

For sufficiently large n , the limit of the expectation is given as,

$$\lim_{n \rightarrow \infty} \mathbb{E}[E_n] = \sum_{i=1}^k \left[\mathbb{P}(X_1 \leq i) - \mathbb{P}(Y_1 \leq i) \right]^2$$

E_n is a convex function with a global minimum. This global minimum occurs when E_n is zero which is guaranteed if and only if the following holds:

$$\mathbb{P}(X_1 \leq i) = \mathbb{P}(Y_1 \leq i)$$

This means that the probability distribution function of $X_i, Y_i, i = 1, 2, \dots, n$ have to be equal. If their distributions are equal, their means must also be equal, that is $\mu_X = \mu_Y$.

□

The implication of Theorem 6.2 is stated in the following corollary.

Corollary. *For two pixel patches X, Y with assumed constant intensity values within a sufficiently large histogram window and perturbed by signal dependent noise, the L_2 -distance of the respective cumulative histograms attains a minimum of zero independently of the noise characteristic γ , if and only if the unbiased intensity values correspond to each other, i.e., $\mu_X = \mu_Y$.*

In practice, the minimum of E_n is never zero because finding two image patches with the same unbiased intensity distribution is almost impossible with images acquired under clinical conditions. A formal representation of the Histogram Horn-Schunck energy functional can be written similar to (6.8) as follows:

$$E(\mathbf{u}) = \int_{\Omega} |\nabla H(\mathbf{x}, t) \cdot \mathbf{u} + H_t(\mathbf{x}, t)|^2 + \alpha |\nabla \cdot \mathbf{u}|^2 d\Omega \quad (6.12)$$

where Ω is the image space, $\nabla H(\mathbf{x}, t) \cdot \mathbf{u}$ is the dot product of the spatial partial derivatives of the image histogram $H(\mathbf{x}, t)$ with the velocity vector field \mathbf{u} . $H_t(\mathbf{x}, t)$ is the time derivative of the image histogram, α is a weight, and $\nabla \cdot \mathbf{u}$ is the dot product of the partial derivative operator (spatial) with the velocity vector field \mathbf{u} . Minimization of $E(\mathbf{u})$ gives the desired velocity vector field.

A comparison of Histogram Horn-Schunck and Horn-Schunck methods was performed using two synthetically generated ultrasound images representing two different temporal points of a cardiac phase as shown in Figure 6.2. The velocity vector field between these two images is known *a priori* and it serves as the ground truth for the comparison. Each image is corrupted with speckle noise according to Definition 6.2 with γ chosen as 1, and σ^2 in the range $(0, 2]$. γ value of 1 or 2 is typically used [112, 159] depending on the imaging system, and σ^2 controls the noise level. The average end point error (AEE) was computed between the ground truth field and the fields obtained from Histogram Horn-Schunck and Horn-Schunck methods, respectively. Average endpoint error is the mean squared error between the ground truth velocity vector field and the estimated velocity vector field and it is given as follows:

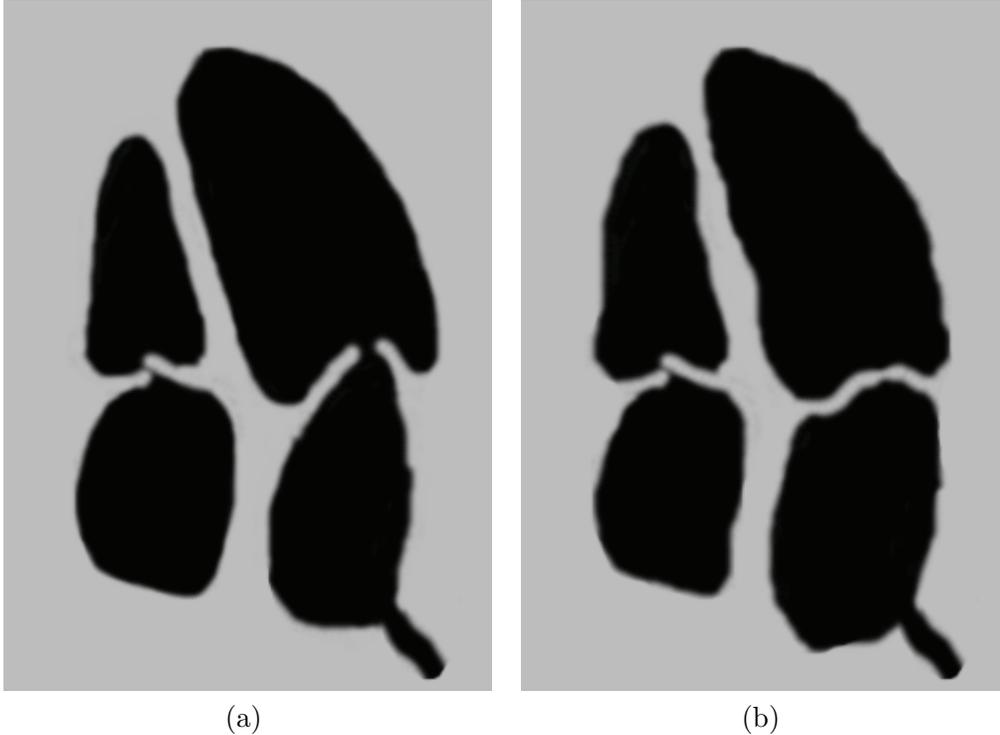


Figure 6.2: Two synthetic images used to evaluate the performance of Horn-Schunck method and Histogram Horn-Schunck method. (a) Frame 1 (b) Frame 2.

$$AEE = \sqrt{(\mathbf{u} - \mathbf{u}_{gt})(\mathbf{u} - \mathbf{u}_{gt})^T} \quad (6.13)$$

where \mathbf{u} is the estimated velocity vector field, \mathbf{u}_{gt} is the ground truth velocity vector field, and T is the matrix transpose operator. Table 6.1 summarizes the AEE for different σ^2 values. For smaller σ^2 values, Horn-Schunck's method performs better, however, as σ^2 increases the Histogram Horn-Schunck method outperforms the Horn-Schunck method. This is expected because as the noise level increases the local statistics is a better predictor of the image patch motion and it is not as susceptible to individual pixel fluctuations caused by speckle noise. Even though the performance gain derived from using Histogram Horn-Schunck method is small, it is still preferred in our application because it better models group motion rather than single voxel motion. The velocity vector fields derived from this section provide important features in the segmentation routine described in the next section.

Table 6.1: Average end point error comparison of the Horn-Schunck method and Histogram Horn-Schunck method.

| Noise level (σ^2) | Original Horn-Schunck | Histogram Horn-Schunck |
|-------------------------------|--------------------------|---------------------------|
| 0.05 | 0.7404 | 0.7525 |
| 0.15 | 0.7702 | 0.7807 |
| 0.25 | 0.9100 | 0.8308 |
| 0.50 | 1.0521 | 0.8825 |
| 0.75 | 1.1432 | 0.9437 |
| 1.00 | 1.1570 | 0.9483 |
| 1.50 | 1.6394 | 1.1253 |
| 2.00 | 1.7316 | 1.2625 |

6.4 Segmentation

The velocity vector field provides information about how the structures of a fetal heart are moving over time which can help in the improvement of fetal heart chambers identification. The block diagram of Figure 6.1(b) shows the three processing steps in the segmentation approach which are:

1. Features are selected from the 4-D image and the estimated velocity vector fields.
2. A nonlinear transformation (kernel method/function) is applied to the feature vectors to improve fetal heart chamber separability in a high-dimensional nonlinear subspace.
3. K -means clustering is applied in the high-dimensional space to separate the feature vectors into clusters representing the fetal heart chambers.

6.4.1 Feature selection

The features selected must have good discriminative properties such that the combination of these features can be used to identify the individual fetal heart chambers in a high-dimensional nonlinear subspace. There were seven features selected for the segmentation method that provided this discriminative property. The first three features are related to the position of each voxel in the 4-D ultrasound image. This provides distance based separability of the data such that points close together are more likely to be grouped together than points farther apart. The position features are particularly important since the chambers are side by side and there is a need for features that discourage points clusters overlapping two different chambers. The next three features are related to the estimated velocity vector field. The velocity vector field is represented in a spherical coordinate system

with three components: a radial magnitude, an elevation angle and an azimuth angle. The radial magnitude is a positive real number in the range $[0, \infty)$, the elevation angle is measured in radians with values in the range $[0, \pi]$, and the azimuth angle is measured in radians with values in the range $(-\pi, \pi]$. These three components completely define the magnitude and direction of the velocity vector field for each voxel. The velocity vector features are important in ensuring that voxels representing structures with similar motion characteristics, magnitude and direction, are grouped together. During the diastolic phase of a cardiac cycle, voxels representing structures in the right atrium have velocity vectors with direction pointing from the right atrium to the right ventricle which provides valuable information that can be used to group these voxels as emanating from the right atrium. Voxels representing the right ventricle, the left atrium and the left ventricle can also be grouped similarly through the magnitude and direction of their respective velocity vector features. The final feature is based on the intensity value of the voxels. The intensity value provides additional basis for grouping voxels belonging to each fetal heart chamber. Particularly, the intensity values are used to discriminate between chamber boundaries since the atrioventricular valves and the septum, which typically represent chamber boundaries, have higher intensity values than the chamber cavities.

In summary, the features selected provide physically measurable characteristics of a 4-D fetal echocardiographic image such that when they are combined, they provide an effective way of grouping voxels belonging to each fetal heart chamber. The list below summarizes the seven features used:

1. x coordinate position of a voxel
2. y coordinate position of a voxel
3. z coordinate position of a voxel
4. Radial magnitude of a velocity vector field
5. Elevation angle of a velocity vector field
6. Azimuth angle of a velocity vector field
7. Intensity/brightness value of a voxel

6.4.2 Choice of kernel function

The choice of nonlinear kernel function used to ensure separability of the features vectors was based on cross-validation tests performed on a random sample of 4-D fetal

echocardiographic images. The cross-validation approach is based on using a list of kernel functions and computing the error in classification based on each of these kernel functions. The following kernels were used in the cross-validation tests: 1) linear kernel, 2) polynomial kernel, 3) radial basis function kernel, and 4) sigmoid kernel. The classification results of each kernel was compared to manually identified fetal heart chambers by computing the classification error for each kernel function. The kernel function with the lowest average classification error produces the best separability of the fetal heart chambers. A comprehensive reference for kernel methods/functions in machine learning can be found in [122, 164, 173], however, brief descriptions of the above mentioned kernel functions are given below:

1. **Linear kernel:** The linear kernel is the simplest kernel function. It is given by the inner product $\mathbf{g}_1 \cdot \mathbf{g}_2 \equiv (\mathbf{g}_1^T \mathbf{g}_2)$ of two feature vectors, \mathbf{g}_1 and \mathbf{g}_2 , plus an optional constant c :

$$\mathbf{K}(\mathbf{g}_1, \mathbf{g}_2) = \mathbf{g}_1^T \mathbf{g}_2 + c \quad (6.14)$$

2. **Polynomial kernel:** The polynomial kernel is the linear kernel to a polynomial power d . In many applications, d is chosen as 2 because higher order polynomials tend to overfit the data in addition to increasing classification noise which leads to amplified classification error. The general polynomial kernel is given as:

$$\mathbf{K}(\mathbf{g}_1, \mathbf{g}_2) = \left(\mathbf{g}_1^T \mathbf{g}_2 + c \right)^d \quad (6.15)$$

3. **Radial basis function kernel:** The radial basis function (RBF) kernel is the most commonly used kernel because of its suitability for many naturally occurring dataset. Radial basis function kernel for two feature vectors \mathbf{g}_1 and \mathbf{g}_2 is defined as follows:

$$\mathbf{K}(\mathbf{g}_1, \mathbf{g}_2) = \exp \left(-\frac{|\mathbf{g}_1 - \mathbf{g}_2|^d}{2\sigma^2} \right) \quad (6.16)$$

where σ^2 is an adjustable parameter that controls the performance of the kernel and should be carefully tuned. If σ^2 is overestimated, the kernel starts to behave like a linear kernel and if it is underestimated it makes the classification decision boundary highly sensitive to noise. When $d = 2$ the RBF kernel is called a Gaussian kernel, when $d = 1$ it is called an exponential kernel or a laplacian kernel [173].

4. **Sigmoid kernel:** The use of sigmoid kernel in classification has its origin in multilayer perceptron and neural networks where it is often used as an activation function for artificial neurons. It is given as follows:

$$\mathbf{K}(\mathbf{g}_1, \mathbf{g}_2) = \tanh\left(\gamma \mathbf{g}_1^T \mathbf{g}_2 + c\right) \quad (6.17)$$

where γ is a scaling factor for the input data, and c is a shifting parameter that controls the threshold of the sigmoid mapping.

Table 6.2 shows the average classification error for a random sample of 20 4-D fetal echocardiographic image. The classification error is derived by comparing the segmentation results using each kernel with segmentation results from manually identified chambers. The parameters shown in the table are the best performing parameters for each kernel function from the cross-validation tests. The polynomial kernel of order two outperformed all the other kernels and it is therefore the kernel of choice in the segmentation method. The next step in the segmentation method is the grouping of these transformed feature vectors (using polynomial kernel of order two) such that the derived clusters represent the fetal heart chambers.

6.4.3 Clustering using kernel k -means algorithm

Clustering, in the context of image processing, is the unsupervised grouping of voxels with similar features for the purpose of segmentation. This works because of the assumption that certain structures have similar features which can be exploited to identify them. The transformed feature vectors in a high-dimensional nonlinear subspace are grouped into clusters representing the fetal heart chambers by using the kernel k -means algorithm. Kernel K -means [44, 165] is a generalization of the standard k -means algorithm [62] which can be used to handle data that are not linearly separable in the input space. It involves mapping

Table 6.2: Cross validation of different kernel function used in selecting the kernel with best classification error performance.

| Kernel function | Parameters | Classification error |
|-----------------|-----------------------|----------------------|
| Linear | $c = 0$ | 14.54% |
| Polynomial | $c = 10, d = 2$ | 10.22% |
| RBF | $\sigma^2 = 1, d = 2$ | 43.83% |
| Sigmoid | $c = 1, \gamma = 1$ | 78.77% |

points to a higher-dimensional nonlinear space and then finding hyperplanes that linearly separates the data in the high-dimensional nonlinear space. This mapping is done with a kernel function which in our case is a polynomial kernel of order 2 as described in Section 6.4.2. The formulation of Kernel k -means is given as follows: Given a set of feature vectors $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$, the standard k -means seeks to find clusters c_1, c_2, \dots, c_k that minimize the objective function:

$$\mathcal{D}(c_m) = \sum_{m=1}^k \sum_{\mathbf{g}_i \in c_m} |\mathbf{g}_i - \mu_m|^2, \quad (6.18)$$

where $\mu_m = \frac{\sum_{\mathbf{g}_i \in c_m} \mathbf{g}_i}{|c_m|}$

where the m^{th} cluster, for $m = 1, 2, \dots, k$, is denoted by c_m , $|c_m|$ is the number of points in cluster c_m , and the centroid or mean of cluster c_m is denoted by μ_m . Kernel k -means involves a transformation/mapping (represented by function Φ) of the feature vectors into a high-dimensional nonlinear space, so the corresponding objective function for kernel k -means algorithm follows from the standard k -means algorithm and it is given as:

$$\mathcal{D}(c_m) = \sum_{m=1}^k \sum_{\mathbf{g}_i \in c_m} |\Phi(\mathbf{g}_i) - \mu_m|^2, \quad (6.19)$$

where $\mu_m = \frac{\sum_{\mathbf{g}_i \in c_m} \Phi(\mathbf{g}_i)}{|c_m|}$

If the distance term $|\Phi(\mathbf{g}_i) - \mu_m|^2$ in the objective function is expanded we obtain the following

$$|\Phi(\mathbf{g}_i) - \mu_m|^2 = \Phi(\mathbf{g}_i) \cdot \Phi(\mathbf{g}_i) - \frac{2 \sum_{\mathbf{g}_j \in c_m} \Phi(\mathbf{g}_i) \cdot \Phi(\mathbf{g}_j)}{|c_m|} + \frac{\sum_{\mathbf{g}_j, \mathbf{g}_l \in c_m} \Phi(\mathbf{g}_j) \cdot \Phi(\mathbf{g}_l)}{|c_m|^2} \quad (6.20)$$

where the dot product terms $\Phi(\mathbf{g}_i) \cdot \Phi(\mathbf{g}_j)$ for $i, j = 1, 2, \dots, n$ represent elements of the kernel matrix \mathbf{K} (Section 6.4.2 and Figure 6.1(b)), that is $\mathbf{K}_{ij} = \Phi(\mathbf{g}_i) \cdot \Phi(\mathbf{g}_j)$. Note that the feature vectors $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$ are the observations (or rows) of the feature matrix \mathbf{G} in Figure 6.1(b). The minimization of the kernel k -means objective function is equivalent to grouping of feature vectors with similar characteristics by finding the best cluster center μ_m with minimum intracluster distance error. Algorithm 6.1 shows a procedure used to minimize (6.19) where an initial set of clusters $\{c_m^{(0)}\}_{m=1}^k$ are adapted iteratively until

Algorithm 6.1 Iterative algorithm for kernel k -means.

KERNEL_KMEANS (\mathbf{K} , k , t_{max} , $\{c_m\}_{m=1}^k$)

1: **Inputs:**

\mathbf{K} : kernel matrix,

k : number of clusters,

t_{max} : optional maximum number of iterations

2: **Outputs:**

$\{c_m\}_{m=1}^k$: final partitioning of the points

3: Randomly initialize the k clusters $c_1^{(0)}$, $c_2^{(0)}$, \dots , $c_k^{(0)}$.

4: Set $t = 0$

5: For each point \mathbf{g}_i , and every cluster m , using (6.20), compute

$$d(\mathbf{g}_i, \mu_m) = \mathbf{K}_{ii} - \frac{2 \sum_{\mathbf{g}_j \in c_m} \mathbf{K}_{ij}}{|c_m|} + \frac{\sum_{\mathbf{g}_j, \mathbf{g}_l \in c_m} \mathbf{K}_{jl}}{|c_m|^2}$$

6: Find $m^*(\mathbf{g}_i) = \underset{m}{\operatorname{argmin}} d(\mathbf{g}_i, \mu_m)$. That is find the cluster index m to which \mathbf{g}_i belongs such that the distance $d(\mathbf{g}_i, \mu_m)$ is minimized.

7: Compute the updated clusters as

$$c_m^{t+1} = \{\mathbf{g}_i : m^*(\mathbf{g}_i) = m\}$$

8: If not converged or $t_{max} > t$, set $t = t + 1$ and go to Step 5; Otherwise, stop and output final clusters $\{c_m^{(t+1)}\}_{m=1}^k$

convergence, that is until a cluster mean μ_m that gives minimum intracluster distance error is found.

6.5 Experiments, results, and discussion

The segmentation method presented in this chapter for identifying fetal heart chambers in 4-D echocardiographic images was validated on a dataset consisting of 130 images acquired from 31 patients. The characteristics of the images are described in Chapter 4 and Appendix A and will not be repeated here.

6.5.1 Experiments

The segmentation method described in this chapter was implemented on a Microsoft® Windows 7, 12 Core, 16 GB RAM workstation running MATLAB® R2012b. First, the motion estimation using Histogram Horn-Schunck method (Section 6.3.2) was performed on each of the 130 images to obtain the velocity vector fields. The fields obtained were in three perpendicular spatial directions (x , y , and z Cartesian directions). The weight α of (6.12) was chosen as 1 and remained constant for all experiments. Also, the energy

functional of (6.12) is minimized by solving the associated Gauss-Siedel and Euler-Lagrange equations [83, 186]. Next, features were selected for each image according to the description of Section 6.4.1. The features for each voxel in an image formed the feature matrix \mathbf{G} which was then passed to the nonlinear transformation using the polynomial kernel with $d = 2$ and $c = 10$ (Table 6.2), to obtain the kernel matrix \mathbf{K} . Finally the kernel k -means algorithm was applied to the transformed features to identify the individual fetal heart chambers. The initial clusters were chosen randomly and are iteratively adapted until convergence according to Algorithm 6.1. The number of clusters k was fixed at 4 representing the four chambers of a fetal heart. Convergence of the k -means algorithm was defined as when none of the points \mathbf{g}_i changed its cluster membership.

6.5.2 Evaluation metrics

The segmentation results were validated by comparing them with manually identified fetal heart chambers. The manual identification procedure is accomplished by observing the video of the 4-D images and selecting control points that correspond to the respective chamber boundaries. Three metrics were used to evaluate the correctness of the segmentation results compared to the manual annotation. They are: 1) Sørensen–Dice index [46], 2) absolute volume difference, and 3) Hausdorff distance [88]. These are briefly described below.

1. **Sørensen–Dice index:** Sørensen–Dice index is used to evaluate segmentation results and it measures the extent of spatial overlap between two segmentation results. Given two segmentation results (sets) A and B , the Sørensen–Dice index is defined as follows:

$$D(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (6.21)$$

where $|A|$ and $|B|$ are the sizes of sets A and B , respectively. The Sørensen–Dice index is in the range $[0, 1]$ where a value of 0 implies no overlap and a value of 1 implies perfect match.

2. **Absolute volume difference:** The absolute volume difference is used to measure the volume mismatch between two segmentation results. Suppose we want to compare a segmentation result A to a ground truth result B each with volumes V_A and V_B , respectively. The absolute volume difference is given by the following equation:

$$A_{VD}(V_A, V_B) = \frac{|V_A - V_B|}{V_B} \quad (6.22)$$

where V_A and V_B is given as the number of segmented voxels multiplied by the respective voxel dimension. Lower values of the absolute volume difference are desired because lower values implies lower volume error.

3. **Hausdorff distance:** Hausdorff distance is a metric used to measure the extent to which points in a set lies near some points of another set. Given two segmentation results A and B , the Hausdorff distance is defined as

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (6.23)$$

where $h(A, B)$ is given as,

$$h(A, B) = \max_{a \in A} \min_{b \in B} |a - b| \quad (6.24)$$

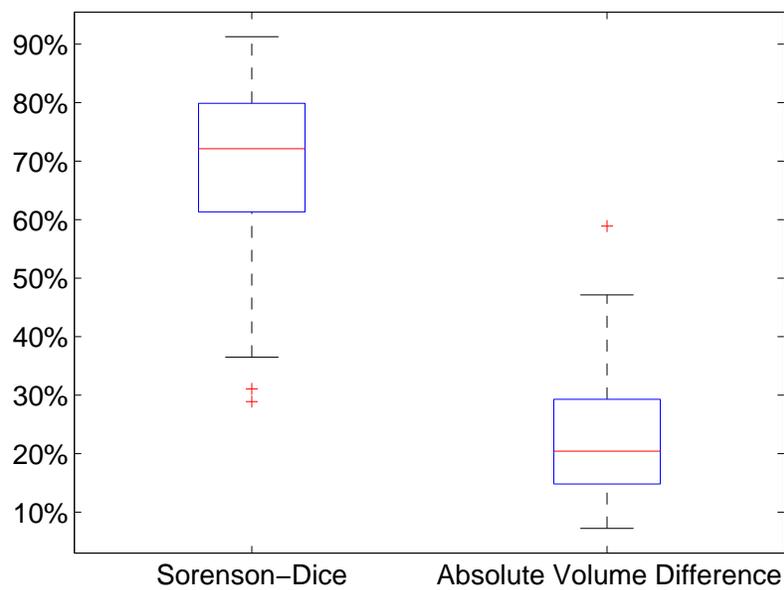
and $|\cdot|$ is the Euclidean norm or L_2 norm on the points of A and B . Because the Hausdorff distance as defined in (6.24) is very sensitive to outliers, a modified Hausdorff distance proposed by Dubuission *et al.* [54] is used, and it is defined as:

$$h(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} |a - b| \quad (6.25)$$

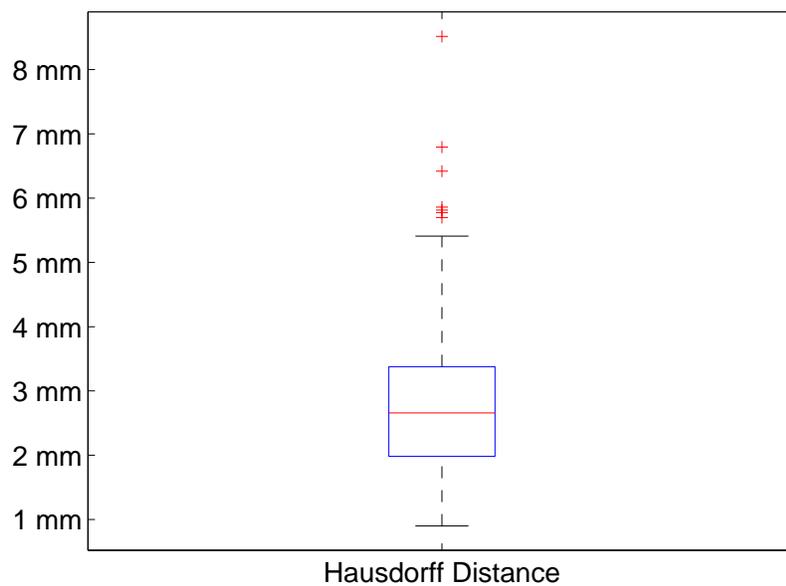
where $|A|$ is the number of points in A . Lower values of Hausdorff distance are desired because lower values implies better closeness of the two sets A and B .

6.5.3 Results and discussion

The evaluation metrics described in Section 6.5.2 were applied to each of the 130 images by comparing the segmented fetal heart chambers with manually identified fetal heart chambers at the end-diastole cardiac phase. End-diastolic cardiac phase is when the ventricles are filled with blood. It represents the largest volume of the heart obtained during the cardiac phase. The average of the Sørensen-Dice index, the absolute volume difference, and the Hausdorff distance for all 130 images was measured as 69.52%, 22.64%, and 2.89 mm, respectively. Figures 6.3(a) and 6.3(b) show the box and whisker plots for each of the metrics. Figure 6.3(a) shows that the Sørensen-Dice index values ranged from a minimum value of 28.90% to a maximum value of 91.27% with 25th percentile, median, and 75th percentile values of 61.38%, 72.11%, and 79.81%, respectively. Figure 6.3(a) also show that



(a)



(b)

Figure 6.3: Box and whisker plot showing the statistics of the evaluation metrics for all 130 4-D fetal echocardiographic images. (a) Sørensen-Dice and absolute volume difference. (b) Hausdorff distance.

the absolute volume difference values ranged from a minimum value of 7.23% to a maximum value of 58.88% with 25th percentile, median, and 75th percentile values of 14.85%, 20.42%, and 29.02%, respectively. Figure 6.3(b) shows that the Hausdorff distance values ranged from a minimum value of 0.90 mm to a maximum value of 8.52 mm with 25th percentile, median, and 75th percentile values of 1.98 mm, 2.66 mm, and 3.38 mm, respectively. An extended breakdown of the numbers for each image is given in Appendix C.

From Appendix C, images with Sorenson-Dice metric values less than 60%, or absolute volume difference values greater than 30%, or Hausdorff distance values greater than 4 mm were further analyzed to establish a pattern for their poor performances. There were a total of 26 images that satisfied two or more of the conditions above. The computed velocity vector fields for these 26 images did not exhibit the expected motion characteristics. The motion of a fetal heart's structures is similar to the blood flow through the chambers. The left atrium receives blood from the placenta through the foramen ovale which is then passed to the left ventricle and through the aorta. Similarly, the right atrium receives blood from the fetal body through the superior vena cava which then flows to the right ventricle and finally through the pulmonary artery. The motion of the voxels representing the fetal heart chamber should exhibit similar motion characteristics with their velocity vector fields pointing in the direction of blood flow. However, the observed velocity vector fields exhibited random pattern rather than the linear pattern synonymous with the blood flow. The random pattern can be described as when the velocity vectors of voxels close together and within an homogeneous region have random directions. The randomness of the velocity vector field is caused by speckle noise. Even though the histogram-based Horn-Schunck method uses local characteristics to reduce the effect of speckle noise, in these images the reduction is not sufficient. A larger window for computing the local histogram statistics could be explored as a means of reducing this randomness although doing this will also reduce the resolution of the velocity estimates and many critical velocity information associated with fetal heart motion will be lost. The consequence of the random motion pattern is that during clustering, the kernel k -means algorithm attempts to create a cluster in a nonlinear subspace that best represents these random velocity vector field which results in erroneous chambers in the wrong location and with the wrong sizes. Chambers in the wrong location caused the low Sorenson-Dice values while chambers with wrong sizes caused the high absolute volume difference and Hausdorff distance values. Since the motion estimation step is critical to the segmentation method, any misrepresentation of the true motion of cardiac structures will result in poor results and erroneous chambers. The occurrence of these random patterns can

be mitigated by using doppler velocity information directly from the ultrasound machines (radio frequency (RF) signal) which is more accurate than the motion estimates computed on post-processed B-mode image voxels. In general a future extension of the segmentation method should incorporate doppler velocity fields computed from the ultrasound system's RF signal for increased accuracy.

Another improvement that can be made to the segmentation method is in the selection of features. Features that could be explored include the image texture features and wavelet features. Image texture is a quantitative measure that describes the arrangement of the brightness or intensity in an image region. Because texture represents both positional information and brightness information, it could either enhance or be a replacement for the Cartesian coordinate position features (x , y , and z coordinate) and the brightness feature (grayscale values) used in our design. In addition, linear or nonlinear combinations of the seven features described in Section 6.4.1 can also be explored as a way of improving segmentation results. Wavelet features such as Haar wavelet have been used in a supervised learning framework to identify the best set of features that represents an objects position, orientation and scale. Wavelet features are attractive in object detection because they form a compact representation, encode edges of objects, capture information from multiresolution, and can be computed efficiently. Zheng *et. al* [211] described an algorithm called marginal space learning that uses Haar wavelet features to model position characteristics of an object in a medical image. This could be extended to the segmentation method as a way of selecting additional features that model the location of the fetal heart chambers.

Another set of analysis was performed on a per patient basis as shown in Table 6.3. The per patient evaluation is based on the average metric values for all images acquired for a patient. The clinical reason for acquiring multiple images per patient is to increase the probability of good image acquisition in case there is fetal movement, stitching artifacts, motion artifacts, or operator error. The advantage of the per patient analysis is that it helps to give a broad view of the evaluation metric performance for all the images as a whole as opposed to independent analysis of single images. From Table 6.3, the per patient averages are 69.92%, 22.08% and 2.82 mm for the Sørensen-Dice, the absolute volume difference, and the Hausdorff distance metrics, respectively.

Table 6.4 shows the results of the segmentation method on the 130 images based on fetal gestational ages. The table shows the 25th percentile, average, and 75th percentile values for each gestation age range. Considering the average values, the Sorenson-Dice metric for fetuses with gestational age greater than 30 weeks outperforms those with gestation age less

Table 6.3: A table showing the average Sørensen-Dice index, absolute volume difference and Hausdorff distance values on a per patient basis.

| Patient no. | Sørensen-Dice | Abs. Vol. Diff. | Hausdorff (mm) |
|-------------|---------------|-----------------|----------------|
| 1 | 66.09% | 25.07% | 0.33 |
| 2 | 66.02% | 21.34% | 0.41 |
| 3 | 69.08% | 24.64% | 0.42 |
| 4 | 79.50% | 14.60% | 0.12 |
| 5 | 66.71% | 18.25% | 0.38 |
| 6 | 71.17% | 25.18% | 0.25 |
| 7 | 68.25% | 27.56% | 0.32 |
| 8 | 67.09% | 29.57% | 0.23 |
| 9 | 62.77% | 17.02% | 0.23 |
| 10 | 87.32% | 11.30% | 0.16 |
| 11 | 74.64% | 16.28% | 0.18 |
| 12 | 58.23% | 19.66% | 0.32 |
| 13 | 63.06% | 29.34% | 0.26 |
| 14 | 79.09% | 15.78% | 0.18 |
| 15 | 71.89% | 22.96% | 0.32 |
| 16 | 72.82% | 25.30% | 0.20 |
| 17 | 72.79% | 20.65% | 0.18 |
| 18 | 68.35% | 37.03% | 0.22 |
| 19 | 49.05% | 26.12% | 0.40 |
| 20 | 80.00% | 17.06% | 0.25 |
| 21 | 67.55% | 29.65% | 0.30 |
| 22 | 84.34% | 19.05% | 0.28 |
| 23 | 78.16% | 19.87% | 0.26 |
| 24 | 66.14% | 38.05% | 0.34 |
| 25 | 67.26% | 20.50% | 0.37 |
| 26 | 65.80% | 21.99% | 0.21 |
| 27 | 76.36% | 18.80% | 0.29 |
| 28 | 53.05% | 25.03% | 0.61 |
| 29 | 69.24% | 15.70% | 0.20 |
| 30 | 74.39% | 16.76% | 0.27 |
| 31 | 71.22% | 14.51% | 0.25 |

than 30 weeks. This is expected because a higher gestational age means the structures of the heart are more developed and are bigger in size.

Table 6.5 shows the results of the segmentation method based on the transducer type used to acquire the images. From the table, the segmentation method shows no dependence on the type of transducer used since the metric values are only slightly different. Therefore, it can be concluded that the transducer type does not affect the segmentation method's results. Table 6.6 shows the segmentation method's results based on the ultrasound acqui-

Table 6.4: A table showing the fetal heart chamber segmentation results according to fetal gestational ages in weeks. The 25th percentile, the average, and the 75th values are shown for each evaluation metric.

| Gest. Age | Sørenson-Dice | | | Abs. Vol. Diff. | | | Hausdorff (mm) | | |
|-----------|------------------|--------|------------------|------------------|--------|------------------|------------------|------|------------------|
| | 25 th | Avg. | 75 th | 25 th | Avg. | 75 th | 25 th | Avg. | 75 th |
| [0, 22) | 59.00% | 68.82% | 77.19% | 15.05% | 23.63% | 29.55% | 1.51 | 2.16 | 2.70 |
| [22, 24) | 65.93% | 70.37% | 77.85% | 11.50% | 21.15% | 29.94% | 1.72 | 2.43 | 2.87 |
| [24, 26) | 57.84% | 66.69% | 78.59% | 15.03% | 24.04% | 31.52% | 2.34 | 3.31 | 3.93 |
| [26, 28) | 69.34% | 70.94% | 80.47% | 17.55% | 22.10% | 24.91% | 2.06 | 2.74 | 2.57 |
| [28, 30) | 68.23% | 69.94% | 78.86% | 16.91% | 22.72% | 26.50% | 2.23 | 3.11 | 3.58 |
| [30, ∞) | 74.29% | 77.68% | 84.00% | 15.43% | 19.89% | 23.65% | 2.80 | 3.26 | 3.27 |

Table 6.5: A table showing the fetal heart chamber segmentation results according to the transducer type used to acquire the images. The 25th percentile, the average, and the 75th values are shown for each evaluation metric.

| Trans. Type | Sørenson-Dice | | | Abs. Vol. Diff. | | | Hausdorff (mm) | | |
|-------------|------------------|--------|------------------|------------------|--------|------------------|------------------|------|------------------|
| | 25 th | Avg. | 75 th | 25 th | Avg. | 75 th | 25 th | Avg. | 75 th |
| X5-1 | 65.52% | 70.83% | 80.55% | 15.53% | 21.31% | 27.21% | 2.69 | 3.49 | 4.11 |
| X7-2 | 58.82% | 68.43% | 79.60% | 14.39% | 23.61% | 31.11% | 1.72 | 2.60 | 3.03 |

Table 6.6: A table showing the fetal heart chamber segmentation results according to the ultrasound system acquisition mode. The 25th percentile, the average, and the 75th values are shown for each evaluation metric. FV refers to full volume acquisition, and L3D refers to live 3-D acquisition

| Acq. Mode | Sørenson-Dice | | | Abs. Vol. Diff | | | Hausdorff | | |
|-----------|------------------|--------|------------------|------------------|--------|------------------|------------------|------|------------------|
| | 25 th | Avg. | 75 th | 25 th | Avg. | 75 th | 25 th | Avg. | 75 th |
| FV | 59.15% | 68.70% | 79.81% | 14.39% | 23.58% | 30.73% | 1.72 | 2.57 | 2.94 |
| L3D | 64.30% | 70.09% | 79.75% | 15.53% | 21.71% | 27.48% | 2.49 | 3.39 | 4.10 |

sition mode. From the table, the average values of the metrics show no dependence of the acquisition mode on the segmentation results.

Figure 6.4 shows sample results of the fetal heart chamber segmentation method. The first column shows the four chamber view of the images. The second column shows the clustering results from the segmentation method. The clusters are color coded to enhance visual comparison with the manually identified chambers and also note that the cluster boundaries are marked with a thick outline. The third column shows the manually identified chambers outlined with the same color code as the segmented chambers from the segmentation method (second column). The fourth column shows the boundary outline of

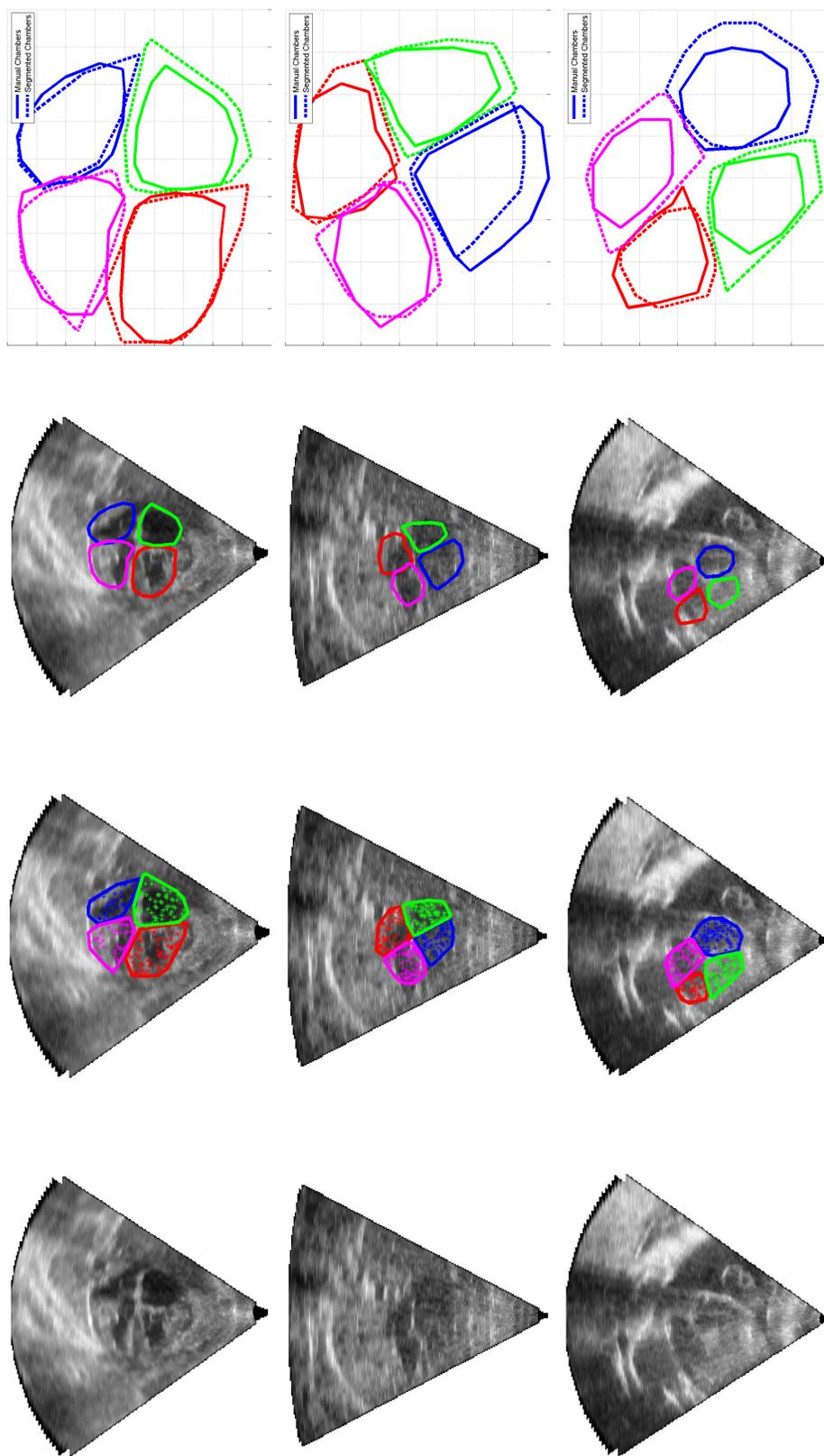


Figure 6.4: Four chamber view showing the boundary outline comparison of the segmentation method with the boundary outline of the manually identified fetal heart chambers. The first column shows the four chamber view of three images, the second column shows the segmented chambers using the segmentation method described in this chapter, the third column shows manually identified chambers, and the fourth column shows an overlap of the segmented chambers (dashed lines) and the manually identified chambers (solid lines).

the four chambers for both the segmentation method and the manual identification on the same plot. The segmented chambers are represented with dashed lines and the manually identified chambers are represented with solid lines with color codes similar to the second and third columns.

6.6 Conclusion

This chapter presented a segmentation method using velocity vector field. To the author's knowledge this is a first attempt at identifying individual fetal heart chambers using 4-D fetal echocardiographic images. Evaluation results produced average values of 69.52%, 22.64%, and 2.89 mm for the Sørensen-Dice index, absolute volume difference, and Hausdorff distance, respectively. When the images are considered on per patient basis, the averages were measured as 79.40%, 17.48%, and 2.19 mm, respectively. The method is an important step in helping care givers in underserved areas with the prenatal detection of hypoplastic left heart syndrome or any other congenital heart defect with underdeveloped chambers. In the next chapter, a procedure for detecting hypoplastic left heart syndrome that uses the identified fetal heart chambers of this chapter is described. This procedure is a binary classification method that uses an empirical threshold on the chamber volume measurements to decide whether the combination of the chambers belong to a control (normal) class of fetal hearts or fetal hearts with hypoplastic left heart syndrome.

CHAPTER 7

DETECTION OF HYPOPLASTIC LEFT HEART SYNDROME

7.1 Introduction

Hypoplastic left heart syndrome (HLHS) is a CHD characterized by an underdeveloped left ventricle (see Chapter 1, Section 1.4) and it accounts for 4-8% of all cases of CHD. The main diagnostic feature used in identifying HLHS by cardiologists is the size of the left ventricle relative to other chambers of a fetal heart, and the procedures developed in Chapter 5 and Chapter 6 of this dissertation are towards isolating and identifying this feature for the purpose of detecting HLHS. In particular, the segmentation of the individual fetal heart chambers of a 4-D echocardiographic image facilitates the quantification and classification of the chamber properties as exhibiting either normal fetal heart anatomy or HLHS anatomy.

To effectively identify HLHS using the relative size of the left ventricle to the other fetal heart chambers, we must first understand how the fetal heart chamber sizes vary with gestational age. Numerous studies [42, 59, 171, 200] have evaluated the fetal heart chamber sizes in 2-D. Luewan *et al.* [116] conducted a comprehensive study of fetal heart structures in 4-D images. Their study was conducted using volume datasets acquired using spatiotemporal image correlation with M-mode display (4-D cardio STIC-M). They evaluated 657 normal fetal heart volumes over a period of 2 years, with gestational ages between 14 to 40 weeks. They found the ratio of the left ventricle (LV) to the right ventricle (RV) to be approximately one-to-one (1 : 1) throughout the gestational ages considered. This ratio is consistent with the value obtained by other researchers [59, 171] using 2-D ultrasound. In addition, the ratio of the right atrium (RA) to the left atrium (LA) showed a one-to-one (1 : 1) size correspondence. In fetal hearts with HLHS, the LV:RV ratio is much smaller ($\ll 1$) [153], which is expected since the LV is underdeveloped.

In this Chapter, a simple volume-based binary classification method is presented which uses an empirical threshold on the relative sizes of the fetal heart chambers to flag/detect HLHS in a 4-D fetal echocardiographic image. This chapter is organized as follows: Section

7.2 describes the binary volume classification method. Section 7.3 discusses the results obtained from applying the classification on a dataset of 4-D fetal echocardiographic images consisting of both HLHS and normal fetal heart images. Section 7.4 concludes this chapter.

7.2 Method

The volume classification procedure for identifying HLHS from 4-D images involved two steps. The first step is the computation of an empirical threshold (training), and the second step is the application of the empirical threshold (detection). The empirical threshold is computed using the 130 image dataset (see Chapter 4) and it involves using volumes of manually identified chambers as a basis for discriminating between normal fetal heart and fetal hearts with HLHS. The image dataset consists of 20 images with confirmed diagnosis of HLHS and 110 images of normal fetal hearts. The 20 images were obtained from a total of 5 patients and the 110 images were obtained from a total of 26 patients. To compute the threshold, the fetal heart chambers are first identified manually. The manual identification involves observing the video of each 4-D image and selecting control points that define the boundaries for each chamber. The ratios of the sum of adjacent chamber volumes are then computed for each image. There are two possible combinations of adjacent chamber volume ratio and they are:

1. Left Atrium + Right Atrium, Left Ventricle + Right Ventricle (LA+RA : LV+RV)
2. Right Atrium + Right Ventricle, Left Atrium + Left Ventricle (RA+RV : LA+LV)

Because the specific chamber that corresponds to the RA, LA, RV, and LV are unknown, the mean of the two ratios above are computed. This ensures that any adjacent chamber combination is included in the volume ratio computation. Next, the images are divided into two sets: the first set contained the 20 images with confirmed HLHS diagnosis, and the second set contained the 110 images of normal fetal hearts. The purpose of this is to determine the range of values of the volume ratio for these sets of images. For the HLHS images, the volume ratio was in the range $[0.57, 0.93]$ with average and median values of 0.75 and 0.76, respectively, while for the normal fetal heart images the range was $[0.75, 1]$ with average and median values of 0.95 and 0.96, respectively. The empirical threshold was computed by finding the mean of the average ratios of the two sets of images, *i.e.*, $(0.75 + 0.95)/2 = 0.85$. The empirical threshold as computed implies that any image with volume ratio less than 0.85 is flagged as HLHS and any volume ratio above 0.85 is flagged as normal. Using this threshold on the manually identified fetal heart chambers, 19 of 20

(95%) fetal hearts with HLHS are detected correctly, and 106 of 110 (96.36%) normal fetal heart images are detected correctly. The primary reason for an imperfect manual detection is because of a few outliers that deviated from the average chamber volume ratios.

The second step in the volume classification procedure is the application of the empirical threshold to the segmentation results of Chapter 6. This is the detection process where the results of the segmented chamber volume ratios are flagged based on the empirical threshold computed using manual chamber identification. In the next section the result of applying the empirical threshold to the segmented fetal heart chambers of Chapter 6 is reported.

7.3 Results and discussion

The volume ratio of the segmentation results of Chapter 6 were computed and the results were classified as either HLHS or normal fetal heart by using the empirical volume threshold derived in Section 7.2. For the 20 images with confirmed HLHS diagnosis, the segmented volume ratio values ranged between $[0.47, 0.94]$ with mean and median values of 0.67 and 0.68, respectively. For the 110 normal fetal heart images, the segmented volume ratio values ranged between $[0.56, 1]$ with mean and median values of 0.89 and 0.92, respectively. Using the empirical threshold of 0.85 on each image, 18 of 20 (90%) fetal hearts with HLHS were correctly detected and 84 of 110 (76.37%) normal fetal hearts were correctly detected.

Table 7.1 shows the detection results for the 130 images grouped according to patients. The number of images with correct detection per patient is shown (see Appendix D for more details). From the table, there are three patients (patients 16, 19, and 24) where a majority of images had incorrect detection. These incorrect detections can be attributed to the segmented chambers derived in Chapter 6. A closer look at these images revealed high absolute volume difference values (*i.e.*, $> 30\%$) between manually identified chambers and the segmented chambers. Because the detection process uses the volume ratios, the large absolute volume difference errors resulted in the failed cases.

Table 7.2, Table 7.3 and Table 7.4 shows the detection results according to gestational age, transducer type, and acquisition mode, respectively. In Table 7.2 the detection system performed well for fetuses with gestational ages in the range $[0, 22)$, $[28, 30)$, and $[30, \infty)$ with success rates greater than 80%. The poor performance of the detection system for fetuses in the range $[26, 28)$ can be attributed to images from patient 16 (described above) which accounted for all 5 incorrect detections. In Table 7.3, the *X5-1* transducer outperformed the *X7-2* transducer in terms of detection accuracy. This is expected because the *X5-1* transducer has a higher number of piezoelectric elements (3,040) than the *X7-2* transducer

Table 7.1: A table showing the detection results for the 130 fetal echocardiographic images grouped according to patient number.

| Patient no. | Confirmed Diagnosis | No. of Images with Correct Detection | % of Images with Correct Detection |
|--------------------|----------------------------|---|---|
| 1 | HLHS | 4 of 5 | 80.00% |
| 2 | HLHS | 3 of 4 | 75.00% |
| 3 | HLHS | 2 of 2 | 100.00% |
| 4 | Normal | 1 of 1 | 100.00% |
| 5 | Normal | 5 of 5 | 100.00% |
| 6 | Normal | 5 of 7 | 71.43% |
| 7 | Normal | 3 of 4 | 75.00% |
| 8 | Normal | 3 of 4 | 75.00% |
| 9 | Normal | 2 of 2 | 100.00% |
| 10 | Normal | 2 of 2 | 100.00% |
| 11 | Normal | 4 of 5 | 80.00% |
| 12 | Normal | 2 of 2 | 100.00% |
| 13 | Normal | 3 of 6 | 50.00% |
| 14 | Normal | 5 of 5 | 100.00% |
| 15 | HLHS | 6 of 6 | 100.00% |
| 16 | Normal | 2 of 7 | 28.57% |
| 17 | Normal | 8 of 9 | 88.89% |
| 18 | Normal | 2 of 4 | 50.00% |
| 19 | Normal | 3 of 7 | 42.86% |
| 20 | Normal | 4 of 4 | 100.00% |
| 21 | Normal | 3 of 4 | 75.00% |
| 22 | Normal | 7 of 7 | 100.00% |
| 23 | Normal | 4 of 5 | 80.00% |
| 24 | Normal | 0 of 1 | 0.00% |
| 25 | Normal | 4 of 5 | 80.00% |
| 26 | HLHS | 3 of 3 | 100.00% |
| 27 | Normal | 3 of 3 | 100.00% |
| 28 | Normal | 5 of 6 | 83.33% |
| 29 | Normal | 1 of 1 | 100.00% |
| 30 | Normal | 1 of 1 | 100.00% |
| 31 | Normal | 2 of 3 | 66.67% |

with 2500 elements. More elements means images can be acquired with higher scan line density which improves spatial resolution. In Table 7.4, the detection accuracy of the live 3-D acquisition mode is better than the full volume acquisition mode. This is expected since the spatial resolution of the live 3-D mode is better than the full volume mode because the acquisition angle is narrower which implies a higher scan line density. The main importance of the results reported in Tables 7.2 - 7.4 is that the detection system can be used for fetuses

Table 7.2: A table showing the detection results according to fetal gestational ages.

| Gest. Age | No. of Normal | No. of HLHS | No. of Correct Detection | % of Correct Detection |
|------------------|----------------------|--------------------|---------------------------------|-------------------------------|
| [0, 22) | 19 | 0 | 16 of 19 | 84.21% |
| [22, 24) | 18 | 5 | 17 of 23 | 73.91% |
| [24, 26) | 38 | 6 | 42 of 54 | 77.78% |
| [26, 28) | 15 | 0 | 10 of 15 | 66.67% |
| [28, 30) | 3 | 5 | 7 of 8 | 87.50% |
| [30, ∞) | 7 | 4 | 10 of 11 | 90.90% |

Table 7.3: A table showing the detection results according to transducer type.

| Trans. Type | No. of Normal | No. of HLHS | No. of Correct Detection | % of Correct Detection |
|--------------------|----------------------|--------------------|---------------------------------|-------------------------------|
| <i>X5-1</i> | 40 | 6 | 39 of 46 | 84.46% |
| <i>X7-2</i> | 70 | 14 | 63 of 84 | 75.00% |

Table 7.4: A table showing the detection results according to acquisition mode.

| Acq. Mode | No. of Normal | No. of HLHS | No. of Correct Detection | % of Correct Detection |
|------------------|----------------------|--------------------|---------------------------------|-------------------------------|
| FV | 70 | 6 | 56 of 76 | 73.68% |
| L3D | 40 | 14 | 46 of 54 | 85.19% |

with gestational ages below 20 weeks with 84.21% accuracy. This early detection helps with the management of HLHS by providing necessary information that can be used to plan surgical intervention to correct the defect after birth.

The analysis of the detection system's results show that its performance is closely tied to the performance of the segmentation method described in Chapter 6. However, some steps could be taken to improve the volume classification process discussed in Section 7.2. First, a larger database of images are need to determined the threshold value. Currently, the threshold value computed in Section 7.2 is sensitive to outliers. If there are 20 images and 5 of those images are outliers, then the mean computation will be skewed towards the outliers. Therefore, a larger image database is needed to reduces the effect of these outliers. In addition, a future improvement could include specific chamber identification, that is identifying which of the segmented chamber is the right atrium, the right ventricle, the left

atrium, or the left atrium. This will involve a mechanism for identifying the fetal spine which is usually used by doctors to determine the chamber positioning. The advantage of the specific chamber identification is that for fetuses close to birth, the ventricle sizes become bigger than the atrial sizes and the 1 : 1 atrioventricular correspondence described in Section 7.1 does not hold [102, 140]. Therefore, having the ability to compare specific chambers rather than adjacent chambers is desirable and could potentially improve detection accuracy.

Table 7.5 summarizes the results of the detection system by considering the sensitivity and specificity statistical measures. Sensitivity measures the proportion of HLHS that are correctly detected and specificity measures the proportion of normal fetal hearts that are correctly detected. In the table, true positive (TP) represents the number of HLHS correctly detected as HLHS, false positive (FP) represents number of normal fetal hearts incorrectly detected as HLHS, true negative (TN) represents number of normal fetal hearts correctly detected as normal, and false negative (FN) represents number of HLHS incorrectly detected as normal.

Finally, the detection process described in this chapter was only applied to HLHS, however, the same procedures can be extended to the detection of hypoplastic right heart syndrome (HRHS). Hypoplastic right heart syndrome is similar to HLHS expect the right ventricle is underdeveloped instead of the left ventricle. The use of relative volume ratio to perform the detection process allows for its flexibility in detecting HRHS. In fact, any CHD that is characterized by a small chamber can be detected using the methods discussed in this dissertation.

Table 7.5: A table showing the sensitivity and specificity analysis of the detection results for all 130 fetal echocardiographic images.

| | | |
|--------|---|--|
| HLHS | TP = 18 | FP = 26 |
| Normal | FN = 2 | TN = 84 |
| | Sensitivity = $\frac{TP}{TP+FN} = 90\%$ | Specificity = $\frac{TN}{FP+TN} = 76.36\%$ |

7.4 Conclusion

A simple volume classification method for detection hypoplastic left heart syndrome was presented. The results showed that the methods described in this dissertation provide a foundation towards an automated detection system for CHD that can help care givers in underserved regions. The results also show that the detection system of this dissertation improves on the current clinical detection rate which is between 30% and 60% [172]. The purpose of this dissertation was met by showing the feasibility of a system that uses 4-D fetal echocardiographic image for identifying CHD. Even though only hypoplastic left heart syndrome was considered, it is my hope that other researchers can build on this system and extend it to other forms of CHD.

CHAPTER 8

CONCLUSION AND FUTURE WORK

The goal of this dissertation was to develop a diagnostic tool using 4-D ultrasound to identify congenital heart defects in fetuses as early as 20 weeks gestational age. To this end, two contributions were made to the current state of the art as it relates to 4-D fetal echocardiographic image analysis. These are: 1) a method to estimate the location of a fetal heart in a 4-D fetal echocardiographic image, 2) a method to identify the individual fetal heart chambers using velocity vector field estimates and kernel-based unsupervised learning. The predominant method for detecting congenital heart defects is through manual analysis of fetal images by a cardiologist to infer the function and characteristics of the fetal heart. There are many areas without access to the services of these experts which contributes to higher missed detection rates. Therefore, a system that can automate fetal heart analysis is important for improving detection and pregnancy outcomes. Currently, no system exists, to the best of the author's knowledge, that utilizes 4-D ultrasound images for this purpose. This body of work has contributed to the establishment of a framework that can be followed to design similar systems using 4-D ultrasound. The described detection system consisted of three components: 1) a fetal heart location estimation component using 4-D ultrasound, 2) a fetal heart chamber segmentation component using velocity vector field estimate and kernel learning, and 3) a detection component using a volume based binary classifier.

The location estimation method was used to isolate a fetal heart in a 4-D echocardiographic image for the purpose of reducing errors that are related to misidentification of the fetal heart structures. It also provided computational complexity reduction whereby computations and analysis are focused on a smaller region of the image. Evaluation of the location estimation method on 130 images produced average values of 80.70%, 89.19%, 91.04%, and 99.17% for the Jaccard index, the Sørensen-Dice index, the Sensitivity index and the Specificity index, respectively. Of particular note is the Sensitivity and Specificity indexes which shows that the location estimation method is capable of identifying the fetal heart with 91.04% accuracy, and also able to distinguish the fetal heart from other structures

with 99.17% accuracy. The isolated fetal heart region was used in subsequent analysis for identifying the individual fetal heart chambers. Although the location estimation method as implemented achieved high accuracy values, there were some drawbacks to the algorithm as presented. First, the edge classification function $\alpha(\mathbf{x})$ (Chapter 5) showed promising false edge reduction in the estimation of the epicardial boundary surface, however, the reduction was found to be insufficient in some of the images considered. This caused underestimation or overestimation of the fetal heart region in these images. This region error was propagated in subsequent analysis in the form of wrong volume measurements of the fetal heart chambers. The parameter that controls the false edge suppression, f_v is currently chosen adaptively by dividing the original image into subimages with sizes chosen empirically which is suboptimal. An optimization method that seeks to minimize false edge suppression based on the subimage sizes is a viable alternative that can be explored in the future iterations of this work. Second, the deformable model, used to extract a closed region bounded by the epicardial surface estimates, could be improved by automating the initial surface placement. In the location estimation experiments, this initialization was performed manually because of the sensitivity of active contours to initialization [205]. Residual false edges that were not suppressed using $\alpha(\mathbf{x})$ interfered with convergences to the desired epicardial surface. A future automated initialization using ideas described by Shan *et al.* [170] could be modified for our specific purposes.

The segmentation method used features derived from motion estimates of the 4-D image. The advantage of this approach was that fetal heart structures were identified not only by their voxel intensity values or voxel positions but also by their motion characteristics. For a moving structure like a fetal heart, this approach provided a way of identifying the fetal heart structures that are usually small and difficult to segment. To the author's knowledge, this was the first attempt at segmenting the fetal heart chambers using 4-D fetal echocardiographic images. Evaluation experiments were performed and compared with manually identified fetal heart chambers showing per patient average values of 69.92% for the Sørensen-Dice index, 22.08% for the absolute volume difference, and 2.82 mm for the Hausdorff distance. In some images considered during segmentation experiments, the velocity vector field estimates had random patterns as opposed to the linear pattern that characterizes blood flow through the fetal heart chambers. This randomness is attributed to speckle noise. Although the histogram-based Horn-Schunck methods uses local image characteristics to reduce the speckle noise effect on velocity vector field estimates, this reduction is not sufficient. Using a larger local statistics window can further reduce the

speckle noise effect, however, this also reduces the velocity vector field resolution. A reduction in resolution means that critical velocity information are lost and the performance of the segmentation is further degraded. In the kernel learning step of the segmentation method, only a few kernel functions were considered. A future extension should include consideration of a wider array of kernel functions or perhaps a custom kernel function that best matches the feature set. Moreover, the effect of additional features like wavelet features [211] and texture-based features should be considered as a way of improving the segmentation method.

The detection method used the segmented fetal heart chambers to detect hypoplastic left heart syndrome. An empirical volume ratio, computed on a dataset of manually identified fetal heart chambers, was used on the segmented chambers to flag the presence of hypoplastic left heart syndrome. The empirical threshold was calculated as 0.85, and 18 of 20 fetal hearts were correctly identified with hypoplastic left heart syndrome while 84 of 110 were correctly identified as normal fetal hearts.

The combination of the location estimation, segmentation, and detection methods formed the detection system and it showed the viability of an automated system that can be used to identify congenital heart defects. A main limitation of the detection system is that it can only be used for congenital defects that affect the chamber sizes such as hypoplastic left heart syndrome, hypoplastic right syndrome and any other defect(s) characterized by mismatch in chamber sizes.

8.1 Future work

The location estimation method was based on the image gradient and region characteristics of the fetal heart, however, using statistical models such as active shape model [36] or active appearance model [35] could improve the location estimation. Investigating this approach could be considered as an alternative to the location estimation method of this dissertation. It is important to note that the quality of the fetal heart location estimate using statistical models will depend on how much variability can be captured by a training data.

A main component of the 4-D segmentation method is the velocity vector field estimation during a cardiac cycle. Computing motion estimates on speckle corrupted B-mode images is suboptimal; a better approach will be to use doppler velocity estimates directly from the ultrasound transducer since it computes velocity vector fields from the received ultrasound beam which is a purer source of the signal than postprocessed B-mode images. A more accurate velocity field will improve the fetal heart chamber segmentation results. Also,

rather than using kernel k -means, which is an unsupervised clustering method, a supervised method such as support vector machines or boosting techniques could be used to provide improved separability of the fetal heart chamber features. Supervised learning methods require a training dataset that learns the desired feature classifier so that an unseen feature set can be classified appropriately based on the learned classifier. A drawback is that the training set will need to capture as much variability in the fetal heart chamber properties as possible which will require a large image dataset.

Finally, only hypoplastic left heart syndrome was considered for this dissertation, however, there exists many more forms of congenital heart defects. The main drawback of the detection system described in this dissertation is that it is only applicable to congenital heart defects with chamber size variability. Extending this work to other forms of congenital heart defects will require methods that can identify particular causal features. For example, a system that can detect transposition of the great arteries should have a mechanism for identifying the relative orientation of the two great vessels, the aorta and the pulmonary artery. The intention of this work was to provide a framework on which other approaches for identifying different forms of congenital heart defects can build on. Even with its limited scope, hypoplastic left heart syndrome, hypoplastic right heart syndrome, and other defects that are characterized by underdeveloped fetal heart chambers can be detected with this system.

APPENDIX A

CHARACTERISTICS OF THE FETAL HEART DATASET

Table A.1: A table showing the various characteristics of each image used in experiments. These include the transducer type, the acquisition mode, the gestational ages of the fetus, number of frames in each image, the frame time of each frame in each image, and the type of congenital heart defect. L3D represent live 3-D acquisition mode and FV represents full volume acquisition mode.

| Image no. | Patient no. | Trans. Type | Acq. Mode | Gest. Age (weeks) | No. of Frames | Frame Time (ms) | Type of CHD |
|-----------|-------------|-------------|-----------|-------------------|---------------|-----------------|-------------|
| 1 | 1 | X7-2 | L3D | 28 3/7 | 17 | 182.74 | HLHS |
| 2 | 1 | X7-2 | L3D | 28 3/7 | 17 | 182.74 | HLHS |
| 3 | 1 | X7-2 | L3D | 28 3/7 | 54 | 55.94 | HLHS |
| 4 | 1 | X7-2 | L3D | 28 3/7 | 17 | 182.74 | HLHS |
| 5 | 1 | X7-2 | L3D | 28 3/7 | 17 | 182.74 | HLHS |
| 6 | 2 | X5-1 | L3D | 34 | 80 | 50.04 | HLHS |
| 7 | 2 | X5-1 | L3D | 34 | 26 | 155.12 | HLHS |
| 8 | 2 | X5-1 | L3D | 34 | 26 | 155.12 | HLHS |
| 9 | 2 | X5-1 | L3D | 34 | 26 | 155.12 | HLHS |
| 10 | 3 | X5-1 | L3D | 22 1/7 | 112 | 18.00 | HLHS |
| 11 | 3 | X5-1 | L3D | 22 1/7 | 89 | 22.49 | HLHS |
| 12 | 4 | X7-2 | FV | 24 3/7 | 25 | 16.07 | Normal |
| 13 | 5 | X7-2 | FV | 27 4/7 | 9 | 39.82 | Normal |
| 14 | 5 | X7-2 | FV | 27 4/7 | 16 | 22.76 | Normal |
| 15 | 5 | X7-2 | FV | 27 4/7 | 16 | 22.75 | Normal |
| 16 | 5 | X7-2 | FV | 27 4/7 | 16 | 22.75 | Normal |
| 17 | 5 | X7-2 | FV | 27 4/7 | 16 | 22.75 | Normal |
| 18 | 6 | X7-2 | FV | 25 6/7 | 21 | 19.53 | Normal |
| 19 | 6 | X7-2 | FV | 25 6/7 | 21 | 19.53 | Normal |
| 20 | 6 | X7-2 | FV | 25 6/7 | 21 | 19.53 | Normal |
| 21 | 6 | X7-2 | FV | 25 6/7 | 21 | 19.53 | Normal |
| 22 | 6 | X7-2 | FV | 25 6/7 | 21 | 19.53 | Normal |
| 23 | 6 | X7-2 | FV | 25 6/7 | 21 | 19.53 | Normal |
| 24 | 6 | X7-2 | FV | 25 6/7 | 21 | 19.53 | Normal |
| 25 | 7 | X7-2 | FV | 24 6/7 | 17 | 22.75 | Normal |

Table A.1 – continued

| Image no. | Patient no. | Trans. Type | Acq. Mode | Gest. Age (weeks) | No. of Frames | Frame Time (ms) | Type of CHD |
|-----------|-------------|-------------|-----------|-------------------|---------------|-----------------|-------------|
| 26 | 7 | X7-2 | FV | 24 6/7 | 17 | 22.75 | Normal |
| 27 | 7 | X7-2 | FV | 24 6/7 | 17 | 22.75 | Normal |
| 28 | 7 | X7-2 | FV | 24 6/7 | 20 | 19.53 | Normal |
| 29 | 8 | X7-2 | FV | 24 1/7 | 23 | 17.79 | Normal |
| 30 | 8 | X7-2 | FV | 24 1/7 | 28 | 14.51 | Normal |
| 31 | 8 | X7-2 | FV | 24 1/7 | 22 | 17.79 | Normal |
| 32 | 8 | X7-2 | FV | 24 1/7 | 22 | 17.79 | Normal |
| 33 | 9 | X7-2 | FV | 20 6/7 | 26 | 14.51 | Normal |
| 34 | 9 | X7-2 | FV | 20 6/7 | 24 | 16.24 | Normal |
| 35 | 10 | X7-2 | FV | 23 2/7 | 22 | 17.80 | Normal |
| 36 | 10 | X7-2 | FV | 23 2/7 | 22 | 17.79 | Normal |
| 37 | 11 | X7-2 | FV | 22 6/7 | 19 | 19.53 | Normal |
| 38 | 11 | X7-2 | FV | 22 6/7 | 19 | 19.53 | Normal |
| 39 | 11 | X7-2 | FV | 22 6/7 | 19 | 19.53 | Normal |
| 40 | 11 | X7-2 | FV | 22 6/7 | 23 | 16.24 | Normal |
| 41 | 11 | X7-2 | FV | 22 6/7 | 23 | 16.24 | Normal |
| 42 | 12 | X7-2 | FV | 21 4/7 | 22 | 17.80 | Normal |
| 43 | 12 | X7-2 | FV | 21 4/7 | 22 | 17.79 | Normal |
| 44 | 13 | X7-2 | FV | 23 2/7 | 17 | 21.02 | Normal |
| 45 | 13 | X7-2 | FV | 23 2/7 | 17 | 21.02 | Normal |
| 46 | 13 | X7-2 | FV | 23 2/7 | 17 | 21.02 | Normal |
| 47 | 13 | X7-2 | FV | 23 2/7 | 17 | 21.02 | Normal |
| 48 | 13 | X7-2 | FV | 23 2/7 | 17 | 21.02 | Normal |
| 49 | 13 | X7-2 | FV | 23 2/7 | 17 | 21.02 | Normal |
| 50 | 14 | X7-2 | FV | 24 | 24 | 16.07 | Normal |
| 51 | 14 | X7-2 | FV | 24 | 24 | 16.07 | Normal |
| 52 | 14 | X7-2 | FV | 24 | 22 | 17.64 | Normal |
| 53 | 14 | X7-2 | FV | 24 | 22 | 17.64 | Normal |
| 54 | 14 | X7-2 | FV | 24 | 17 | 21.95 | Normal |
| 55 | 15 | X7-2 | FV | 25 2/7 | 15 | 26.04 | HLHS |
| 56 | 15 | X7-2 | FV | 25 2/7 | 14 | 26.04 | HLHS |
| 57 | 15 | X7-2 | FV | 25 2/7 | 16 | 24.30 | HLHS |
| 58 | 15 | X7-2 | FV | 25 2/7 | 12 | 31.06 | HLHS |
| 59 | 15 | X7-2 | FV | 25 2/7 | 12 | 31.06 | HLHS |
| 60 | 15 | X7-2 | FV | 25 2/7 | 14 | 27.78 | HLHS |
| 61 | 16 | X7-2 | FV | 27 5/7 | 16 | 22.76 | Normal |
| 62 | 16 | X7-2 | FV | 27 5/7 | 18 | 21.02 | Normal |
| 63 | 16 | X7-2 | FV | 27 5/7 | 21 | 17.79 | Normal |
| 64 | 16 | X7-2 | FV | 27 5/7 | 21 | 17.80 | Normal |
| 65 | 16 | X7-2 | FV | 27 5/7 | 21 | 17.79 | Normal |
| 66 | 16 | X7-2 | FV | 27 5/7 | 21 | 17.80 | Normal |
| 67 | 16 | X7-2 | FV | 27 5/7 | 21 | 17.79 | Normal |
| 68 | 17 | X7-2 | FV | 20 | 18 | 21.02 | Normal |

Table A.1 – continued

| Image no. | Patient no. | Trans. Type | Acq. Mode | Gest. Age (weeks) | No. of Frames | Frame Time (ms) | Type of CHD |
|-----------|-------------|-------------|-----------|-------------------|---------------|-----------------|-------------|
| 69 | 17 | X7-2 | FV | 20 | 18 | 21.02 | Normal |
| 70 | 17 | X7-2 | FV | 20 | 19 | 19.53 | Normal |
| 71 | 17 | X7-2 | FV | 20 | 19 | 19.53 | Normal |
| 72 | 17 | X7-2 | FV | 20 | 19 | 19.53 | Normal |
| 73 | 17 | X7-2 | FV | 20 | 19 | 19.53 | Normal |
| 74 | 17 | X7-2 | FV | 20 | 19 | 19.53 | Normal |
| 75 | 17 | X7-2 | FV | 20 | 19 | 19.53 | Normal |
| 76 | 17 | X7-2 | FV | 20 | 19 | 19.53 | Normal |
| 77 | 18 | X7-2 | FV | 21 | 17 | 22.75 | Normal |
| 78 | 18 | X7-2 | FV | 21 | 16 | 22.75 | Normal |
| 79 | 18 | X7-2 | FV | 21 | 17 | 22.76 | Normal |
| 80 | 18 | X7-2 | FV | 21 | 19 | 19.53 | Normal |
| 81 | 19 | X7-2 | FV | 25 1/7 | 19 | 21.02 | Normal |
| 82 | 19 | X7-2 | FV | 25 1/7 | 19 | 21.02 | Normal |
| 83 | 19 | X7-2 | FV | 25 1/7 | 19 | 21.02 | Normal |
| 84 | 19 | X7-2 | FV | 25 1/7 | 19 | 21.02 | Normal |
| 85 | 19 | X7-2 | FV | 25 1/7 | 19 | 21.02 | Normal |
| 86 | 19 | X7-2 | FV | 25 1/7 | 19 | 21.02 | Normal |
| 87 | 19 | X7-2 | FV | 25 1/7 | 19 | 21.02 | Normal |
| 88 | 20 | X5-1 | L3D | 26 1/7 | 17 | 185.25 | Normal |
| 89 | 20 | X5-1 | L3D | 26 1/7 | 38 | 80.17 | Normal |
| 90 | 20 | X5-1 | L3D | 26 1/7 | 15 | 185.25 | Normal |
| 91 | 20 | X5-1 | L3D | 26 1/7 | 17 | 185.26 | Normal |
| 92 | 21 | X5-1 | L3D | 24 5/7 | 28 | 146.20 | Normal |
| 93 | 21 | X5-1 | L3D | 24 5/7 | 28 | 146.20 | Normal |
| 94 | 21 | X5-1 | L3D | 24 5/7 | 24 | 135.41 | Normal |
| 95 | 21 | X5-1 | L3D | 24 5/7 | 30 | 135.41 | Normal |
| 96 | 22 | X5-1 | L3D | 35 4/7 | 13 | 155.12 | Normal |
| 97 | 22 | X5-1 | L3D | 35 4/7 | 13 | 155.12 | Normal |
| 98 | 22 | X5-1 | L3D | 35 4/7 | 13 | 155.12 | Normal |
| 99 | 22 | X5-1 | L3D | 35 4/7 | 13 | 155.13 | Normal |
| 100 | 22 | X5-1 | L3D | 35 4/7 | 13 | 155.13 | Normal |
| 101 | 22 | X5-1 | L3D | 35 4/7 | 15 | 135.41 | Normal |
| 102 | 22 | X5-1 | L3D | 35 4/7 | 15 | 135.41 | Normal |
| 103 | 23 | X5-1 | L3D | 24 6/7 | 15 | 66.98 | Normal |
| 104 | 23 | X5-1 | L3D | 24 6/7 | 30 | 66.98 | Normal |
| 105 | 23 | X5-1 | L3D | 24 6/7 | 30 | 66.98 | Normal |
| 106 | 23 | X5-1 | L3D | 24 6/7 | 20 | 100.47 | Normal |
| 107 | 23 | X5-1 | L3D | 24 6/7 | 20 | 100.46 | Normal |
| 108 | 24 | X5-1 | L3D | 22 | 161 | 24.91 | Normal |
| 109 | 25 | X5-1 | L3D | 25 6/7 | 40 | 100.46 | Normal |
| 110 | 25 | X5-1 | L3D | 25 6/7 | 193 | 20.75 | Normal |
| 111 | 25 | X5-1 | L3D | 25 6/7 | 28 | 146.20 | Normal |

Table A.1 – continued

| Image no. | Patient no. | Trans. Type | Acq. Mode | Gest. Age (weeks) | No. of Frames | Frame Time (ms) | Type of CHD |
|-----------|-------------|-------------|-----------|-------------------|---------------|-----------------|-------------|
| 112 | 25 | X5-1 | L3D | 25 6/7 | 189 | 21.22 | Normal |
| 113 | 25 | X5-1 | L3D | 25 6/7 | 116 | 21.22 | Normal |
| 114 | 26 | X7-2 | L3D | 22 | 14 | 290.95 | HLHS |
| 115 | 26 | X7-2 | L3D | 22 | 37 | 107.53 | HLHS |
| 116 | 26 | X7-2 | L3D | 22 | 46 | 87.71 | HLHS |
| 117 | 27 | X5-1 | L3D | 28 4/7 | 13 | 155.13 | Normal |
| 118 | 27 | X5-1 | L3D | 28 4/7 | 46 | 43.68 | Normal |
| 119 | 27 | X5-1 | L3D | 28 4/7 | 52 | 38.78 | Normal |
| 120 | 28 | X5-1 | L3D | 24 1/7 | 42 | 48.32 | Normal |
| 121 | 28 | X5-1 | L3D | 24 1/7 | 42 | 48.32 | Normal |
| 122 | 28 | X5-1 | L3D | 24 1/7 | 42 | 47.85 | Normal |
| 123 | 28 | X5-1 | L3D | 24 1/7 | 42 | 47.85 | Normal |
| 124 | 28 | X5-1 | L3D | 24 1/7 | 43 | 47.39 | Normal |
| 125 | 28 | X5-1 | L3D | 24 1/7 | 43 | 47.38 | Normal |
| 126 | 29 | X5-1 | L3D | 19 5/7 | 57 | 35.49 | Normal |
| 127 | 30 | X5-1 | L3D | 20 | 51 | 39.31 | Normal |
| 128 | 31 | X5-1 | L3D | 22 | 13 | 155.13 | Normal |
| 129 | 31 | X5-1 | L3D | 22 | 66 | 30.58 | Normal |
| 130 | 31 | X5-1 | L3D | 22 | 53 | 37.86 | Normal |

APPENDIX B

RESULTS FOR THE LOCALIZATION METHOD

Table B.1: A table showing the region of interest evaluation results by applying Jaccard, Sørensen-Dice, Sensitivity, and Specificity metrics on each of the 130 4-D fetal echocardiographic images.

| Image no. | Patient no. | Jaccard | Sørensen-Dice | Sensitivity | Specificity |
|-----------|-------------|---------|---------------|-------------|-------------|
| 1 | 1 | 87.67% | 93.43% | 88.62% | 99.95% |
| 2 | 1 | 83.16% | 90.81% | 83.71% | 99.97% |
| 3 | 1 | 81.70% | 89.93% | 86.16% | 99.39% |
| 4 | 1 | 72.09% | 83.78% | 73.42% | 99.93% |
| 5 | 1 | 73.67% | 84.84% | 73.67% | 100.00% |
| 6 | 2 | 88.13% | 93.69% | 89.86% | 99.91% |
| 7 | 2 | 73.33% | 84.61% | 73.33% | 100.00% |
| 8 | 2 | 74.31% | 85.26% | 92.47% | 98.31% |
| 9 | 2 | 83.54% | 91.03% | 96.31% | 98.99% |
| 10 | 3 | 81.22% | 89.64% | 83.32% | 99.62% |
| 11 | 3 | 81.94% | 90.07% | 95.17% | 98.27% |
| 12 | 4 | 86.78% | 92.92% | 87.91% | 99.91% |
| 13 | 5 | 75.52% | 86.05% | 76.97% | 99.91% |
| 14 | 5 | 84.53% | 91.62% | 95.14% | 99.23% |
| 15 | 5 | 87.44% | 93.30% | 92.28% | 99.68% |
| 16 | 5 | 86.83% | 92.95% | 95.47% | 99.46% |
| 17 | 5 | 77.69% | 87.45% | 96.29% | 98.85% |
| 18 | 6 | 89.31% | 94.35% | 93.77% | 99.67% |
| 19 | 6 | 78.66% | 88.05% | 96.86% | 98.73% |
| 20 | 6 | 69.59% | 82.07% | 94.99% | 98.61% |
| 21 | 6 | 79.13% | 88.35% | 98.18% | 98.78% |
| 22 | 6 | 81.31% | 89.69% | 99.00% | 98.71% |
| 23 | 6 | 72.50% | 84.06% | 87.98% | 98.90% |
| 24 | 6 | 79.57% | 88.62% | 81.09% | 99.91% |
| 25 | 7 | 74.45% | 85.35% | 94.49% | 99.04% |
| 26 | 7 | 77.62% | 87.40% | 94.31% | 99.02% |
| 27 | 7 | 73.99% | 85.05% | 99.11% | 98.32% |
| 28 | 7 | 78.62% | 88.03% | 91.69% | 99.11% |
| 29 | 8 | 73.84% | 84.95% | 87.02% | 99.01% |

Table B.1 – continued

| Image no. | Patient no. | Jaccard | Sørensen-Dice | Sensitivity | Specificity |
|-----------|-------------|---------|---------------|-------------|-------------|
| 30 | 8 | 82.40% | 90.35% | 95.96% | 98.43% |
| 31 | 8 | 79.26% | 88.43% | 80.73% | 99.87% |
| 32 | 8 | 69.65% | 82.11% | 95.91% | 98.55% |
| 33 | 9 | 70.66% | 82.81% | 97.14% | 97.95% |
| 34 | 9 | 73.61% | 84.80% | 99.12% | 98.51% |
| 35 | 10 | 80.85% | 89.41% | 96.89% | 98.96% |
| 36 | 10 | 90.85% | 95.21% | 96.69% | 99.63% |
| 37 | 11 | 73.18% | 84.51% | 75.59% | 99.91% |
| 38 | 11 | 74.24% | 85.22% | 87.85% | 99.39% |
| 39 | 11 | 76.47% | 86.67% | 91.93% | 99.28% |
| 40 | 11 | 83.59% | 91.06% | 95.79% | 99.08% |
| 41 | 11 | 70.66% | 82.81% | 97.13% | 98.05% |
| 42 | 12 | 80.66% | 89.30% | 90.98% | 99.17% |
| 43 | 12 | 85.21% | 92.01% | 94.12% | 99.46% |
| 44 | 13 | 85.04% | 91.91% | 88.77% | 99.84% |
| 45 | 13 | 89.18% | 94.28% | 90.59% | 99.96% |
| 46 | 13 | 90.47% | 95.00% | 94.67% | 99.87% |
| 47 | 13 | 87.99% | 93.61% | 89.35% | 99.94% |
| 48 | 13 | 85.55% | 92.21% | 95.30% | 99.61% |
| 49 | 13 | 89.17% | 94.27% | 93.47% | 99.82% |
| 50 | 14 | 72.97% | 84.37% | 90.99% | 96.73% |
| 51 | 14 | 93.20% | 96.48% | 94.58% | 99.80% |
| 52 | 14 | 76.35% | 86.59% | 80.68% | 99.56% |
| 53 | 14 | 91.30% | 95.45% | 91.79% | 99.95% |
| 54 | 14 | 73.87% | 84.97% | 91.61% | 98.56% |
| 55 | 15 | 78.13% | 87.72% | 91.80% | 99.23% |
| 56 | 15 | 72.19% | 83.85% | 88.60% | 99.10% |
| 57 | 15 | 83.33% | 90.91% | 92.94% | 99.40% |
| 58 | 15 | 87.27% | 93.21% | 96.10% | 99.69% |
| 59 | 15 | 85.54% | 92.21% | 88.32% | 99.91% |
| 60 | 15 | 71.05% | 83.07% | 71.08% | 100.00% |
| 61 | 16 | 80.16% | 88.99% | 97.03% | 99.39% |
| 62 | 16 | 81.45% | 89.78% | 96.19% | 98.98% |
| 63 | 16 | 79.40% | 88.52% | 93.01% | 99.49% |
| 64 | 16 | 89.92% | 94.69% | 93.21% | 99.86% |
| 65 | 16 | 80.06% | 88.92% | 95.81% | 99.52% |
| 66 | 16 | 86.83% | 92.95% | 87.59% | 99.96% |
| 67 | 16 | 87.21% | 93.17% | 98.25% | 99.41% |
| 68 | 17 | 76.51% | 86.69% | 92.70% | 99.37% |
| 69 | 17 | 88.62% | 93.96% | 89.72% | 99.97% |
| 70 | 17 | 81.44% | 89.77% | 83.77% | 99.92% |
| 71 | 17 | 85.06% | 91.93% | 98.94% | 99.49% |
| 72 | 17 | 86.68% | 92.87% | 91.25% | 99.81% |
| 73 | 17 | 83.17% | 90.81% | 96.91% | 99.38% |
| 74 | 17 | 88.00% | 93.62% | 89.78% | 99.92% |

Table B.1 – continued

| Image no. | Patient no. | Jaccard | Sørensen-Dice | Sensitivity | Specificity |
|-----------|-------------|---------|---------------|-------------|-------------|
| 75 | 17 | 85.36% | 92.10% | 93.54% | 99.70% |
| 76 | 17 | 83.11% | 90.77% | 91.77% | 99.70% |
| 77 | 18 | 85.34% | 92.09% | 98.57% | 99.56% |
| 78 | 18 | 76.18% | 86.48% | 94.42% | 99.30% |
| 79 | 18 | 80.06% | 88.93% | 95.95% | 99.68% |
| 80 | 18 | 82.45% | 90.38% | 87.29% | 99.85% |
| 81 | 19 | 73.90% | 84.99% | 92.95% | 98.92% |
| 82 | 19 | 83.15% | 90.80% | 97.02% | 99.19% |
| 83 | 19 | 82.79% | 90.58% | 93.13% | 99.53% |
| 84 | 19 | 74.72% | 85.53% | 90.07% | 98.80% |
| 85 | 19 | 90.47% | 94.99% | 92.73% | 99.85% |
| 86 | 19 | 87.99% | 93.61% | 91.85% | 99.74% |
| 87 | 19 | 81.80% | 89.99% | 84.18% | 99.83% |
| 88 | 20 | 77.33% | 87.22% | 90.00% | 99.75% |
| 89 | 20 | 73.67% | 84.84% | 97.25% | 97.65% |
| 90 | 20 | 74.53% | 85.40% | 79.79% | 99.86% |
| 91 | 20 | 73.39% | 84.65% | 76.55% | 99.91% |
| 92 | 21 | 80.04% | 88.91% | 89.71% | 99.42% |
| 93 | 21 | 84.64% | 91.68% | 91.06% | 99.73% |
| 94 | 21 | 73.22% | 84.54% | 95.29% | 98.38% |
| 95 | 21 | 87.48% | 93.32% | 99.04% | 99.29% |
| 96 | 22 | 82.11% | 90.18% | 97.16% | 98.90% |
| 97 | 22 | 68.52% | 81.32% | 92.59% | 97.66% |
| 98 | 22 | 80.78% | 89.37% | 92.66% | 98.79% |
| 99 | 22 | 88.59% | 93.95% | 95.11% | 99.52% |
| 100 | 22 | 87.28% | 93.21% | 92.89% | 99.55% |
| 101 | 22 | 86.79% | 92.93% | 98.39% | 98.51% |
| 102 | 22 | 82.01% | 90.12% | 94.71% | 98.21% |
| 103 | 23 | 80.86% | 89.42% | 94.31% | 98.81% |
| 104 | 23 | 91.08% | 95.33% | 92.91% | 99.88% |
| 105 | 23 | 90.80% | 95.18% | 93.85% | 99.75% |
| 106 | 23 | 80.54% | 89.22% | 93.43% | 98.58% |
| 107 | 23 | 80.67% | 89.30% | 99.35% | 97.88% |
| 108 | 24 | 78.24% | 87.79% | 91.07% | 98.90% |
| 109 | 25 | 78.74% | 88.10% | 97.96% | 98.25% |
| 110 | 25 | 86.89% | 92.98% | 96.36% | 98.75% |
| 111 | 25 | 82.09% | 90.17% | 95.50% | 99.25% |
| 112 | 25 | 68.41% | 81.24% | 94.45% | 96.97% |
| 113 | 25 | 72.38% | 83.98% | 91.95% | 97.42% |
| 114 | 26 | 71.00% | 83.04% | 75.00% | 99.48% |
| 115 | 26 | 84.15% | 91.39% | 86.74% | 99.77% |
| 116 | 26 | 76.35% | 86.59% | 76.35% | 100.00% |
| 117 | 27 | 84.76% | 91.75% | 85.95% | 99.95% |
| 118 | 27 | 77.64% | 87.41% | 99.82% | 95.64% |
| 119 | 27 | 84.09% | 91.36% | 96.25% | 99.30% |

Table B.1 – continued

| Image no. | Patient no. | Jaccard | Sørensen-Dice | Sensitivity | Specificity |
|-----------|-------------|---------|---------------|-------------|-------------|
| 120 | 28 | 75.29% | 85.90% | 89.60% | 97.34% |
| 121 | 28 | 72.66% | 84.17% | 87.87% | 96.51% |
| 122 | 28 | 76.13% | 86.45% | 77.24% | 99.70% |
| 123 | 28 | 80.06% | 88.93% | 82.26% | 99.46% |
| 124 | 28 | 70.89% | 82.96% | 81.12% | 97.45% |
| 125 | 28 | 86.40% | 92.71% | 86.51% | 99.98% |
| 126 | 29 | 91.80% | 95.73% | 98.77% | 98.58% |
| 127 | 30 | 71.39% | 83.30% | 80.39% | 98.41% |
| 128 | 31 | 84.71% | 91.72% | 95.22% | 99.82% |
| 129 | 31 | 89.48% | 94.45% | 92.41% | 99.49% |
| 130 | 31 | 69.79% | 82.21% | 88.22% | 97.03% |

APPENDIX C

RESULTS FOR THE SEGMENTATION METHOD

Table C.1: A table showing the segmentation evaluation results by applying Sørensen-Dice index, absolute volume difference, and Hausdorff distance metrics on each of the 130 4-D fetal echocardiographic images.

| Image no. | Patient no. | Sørensen-Dice | Abs. Vol. Diff. | Hausdorff (mm) |
|-----------|-------------|---------------|-----------------|----------------|
| 1 | 1 | 76.58% | 18.20% | 0.23 |
| 2 | 1 | 78.35% | 32.48% | 0.20 |
| 3 | 1 | 82.01% | 13.06% | 0.21 |
| 4 | 1 | 38.28% | 20.15% | 0.41 |
| 5 | 1 | 55.22% | 41.44% | 0.58 |
| 6 | 2 | 67.55% | 27.12% | 0.32 |
| 7 | 2 | 61.81% | 13.20% | 0.41 |
| 8 | 2 | 81.02% | 28.24% | 0.34 |
| 9 | 2 | 53.70% | 16.82% | 0.58 |
| 10 | 3 | 66.56% | 18.68% | 0.42 |
| 11 | 3 | 71.60% | 30.61% | 0.42 |
| 12 | 4 | 79.50% | 14.60% | 0.12 |
| 13 | 5 | 39.23% | 19.73% | 0.64 |
| 14 | 5 | 50.24% | 17.09% | 0.50 |
| 15 | 5 | 80.67% | 20.24% | 0.25 |
| 16 | 5 | 80.54% | 22.44% | 0.26 |
| 17 | 5 | 82.87% | 11.74% | 0.25 |
| 18 | 6 | 67.77% | 30.51% | 0.28 |
| 19 | 6 | 74.86% | 8.40% | 0.26 |
| 20 | 6 | 80.89% | 22.83% | 0.18 |
| 21 | 6 | 78.57% | 20.60% | 0.21 |
| 22 | 6 | 74.52% | 25.03% | 0.28 |
| 23 | 6 | 62.81% | 36.37% | 0.31 |
| 24 | 6 | 58.76% | 32.50% | 0.24 |
| 25 | 7 | 70.20% | 15.22% | 0.18 |
| 26 | 7 | 81.34% | 21.36% | 0.22 |
| 27 | 7 | 70.78% | 26.52% | 0.36 |
| 28 | 7 | 50.67% | 47.13% | 0.51 |
| 29 | 8 | 61.62% | 44.37% | 0.27 |

Table C.1 – continued

| Image no. | Patient no. | Sørensen-Dice | Abs. Vol. Diff. | Hausdorff (mm) |
|-----------|-------------|---------------|-----------------|----------------|
| 30 | 8 | 85.49% | 15.90% | 0.14 |
| 31 | 8 | 69.69% | 14.32% | 0.23 |
| 32 | 8 | 51.56% | 43.69% | 0.30 |
| 33 | 9 | 48.14% | 20.04% | 0.29 |
| 34 | 9 | 77.39% | 14.01% | 0.16 |
| 35 | 10 | 86.63% | 11.66% | 0.18 |
| 36 | 10 | 88.00% | 10.93% | 0.14 |
| 37 | 11 | 76.96% | 10.70% | 0.13 |
| 38 | 11 | 61.54% | 29.28% | 0.25 |
| 39 | 11 | 73.42% | 20.23% | 0.18 |
| 40 | 11 | 79.63% | 9.86% | 0.16 |
| 41 | 11 | 81.63% | 11.34% | 0.17 |
| 42 | 12 | 55.33% | 27.04% | 0.38 |
| 43 | 12 | 61.13% | 12.29% | 0.27 |
| 44 | 13 | 42.67% | 41.90% | 0.42 |
| 45 | 13 | 49.70% | 34.69% | 0.31 |
| 46 | 13 | 68.78% | 23.16% | 0.22 |
| 47 | 13 | 78.75% | 13.57% | 0.17 |
| 48 | 13 | 79.50% | 17.98% | 0.17 |
| 49 | 13 | 58.99% | 44.76% | 0.26 |
| 50 | 14 | 74.18% | 18.93% | 0.22 |
| 51 | 14 | 84.32% | 9.03% | 0.14 |
| 52 | 14 | 81.13% | 12.07% | 0.15 |
| 53 | 14 | 89.45% | 12.53% | 0.14 |
| 54 | 14 | 66.35% | 26.34% | 0.28 |
| 55 | 15 | 66.54% | 24.46% | 0.38 |
| 56 | 15 | 78.59% | 24.20% | 0.28 |
| 57 | 15 | 59.64% | 35.63% | 0.42 |
| 58 | 15 | 79.87% | 18.89% | 0.26 |
| 59 | 15 | 86.81% | 12.55% | 0.21 |
| 60 | 15 | 59.88% | 22.01% | 0.36 |
| 61 | 16 | 76.52% | 24.91% | 0.22 |
| 62 | 16 | 77.41% | 31.21% | 0.24 |
| 63 | 16 | 61.78% | 37.44% | 0.21 |
| 64 | 16 | 69.34% | 30.81% | 0.21 |
| 65 | 16 | 72.14% | 16.72% | 0.17 |
| 66 | 16 | 80.47% | 18.45% | 0.13 |
| 67 | 16 | 72.06% | 17.55% | 0.20 |
| 68 | 17 | 58.29% | 30.43% | 0.28 |
| 69 | 17 | 75.32% | 14.50% | 0.21 |
| 70 | 17 | 46.42% | 11.89% | 0.26 |
| 71 | 17 | 85.50% | 14.83% | 0.14 |
| 72 | 17 | 80.25% | 9.65% | 0.09 |
| 73 | 17 | 76.61% | 25.39% | 0.18 |
| 74 | 17 | 80.43% | 25.52% | 0.15 |

Table C.1 – continued

| Image no. | Patient no. | Sørenson-Dice | Abs. Vol. Diff. | Hausdorff (mm) |
|-----------|-------------|---------------|-----------------|----------------|
| 75 | 17 | 76.57% | 23.22% | 0.17 |
| 76 | 17 | 75.70% | 30.39% | 0.15 |
| 77 | 18 | 76.09% | 32.65% | 0.19 |
| 78 | 18 | 47.53% | 58.88% | 0.45 |
| 79 | 18 | 83.43% | 15.76% | 0.12 |
| 80 | 18 | 66.37% | 40.83% | 0.14 |
| 81 | 19 | 28.90% | 32.80% | 0.54 |
| 82 | 19 | 69.89% | 40.23% | 0.28 |
| 83 | 19 | 39.58% | 22.52% | 0.44 |
| 84 | 19 | 36.49% | 12.96% | 0.52 |
| 85 | 19 | 54.82% | 34.05% | 0.39 |
| 86 | 19 | 57.84% | 31.52% | 0.32 |
| 87 | 19 | 55.81% | 8.78% | 0.32 |
| 88 | 20 | 76.95% | 9.57% | 0.25 |
| 89 | 20 | 78.97% | 19.04% | 0.28 |
| 90 | 20 | 85.11% | 18.27% | 0.25 |
| 91 | 20 | 78.96% | 21.37% | 0.22 |
| 92 | 21 | 70.13% | 21.18% | 0.31 |
| 93 | 21 | 79.54% | 27.48% | 0.19 |
| 94 | 21 | 67.37% | 29.99% | 0.30 |
| 95 | 21 | 53.16% | 39.94% | 0.38 |
| 96 | 22 | 83.98% | 20.05% | 0.27 |
| 97 | 22 | 83.00% | 22.16% | 0.30 |
| 98 | 22 | 82.94% | 25.15% | 0.28 |
| 99 | 22 | 84.02% | 14.04% | 0.28 |
| 100 | 22 | 85.25% | 17.11% | 0.23 |
| 101 | 22 | 88.05% | 13.34% | 0.29 |
| 102 | 22 | 83.16% | 21.54% | 0.29 |
| 103 | 23 | 86.67% | 18.75% | 0.21 |
| 104 | 23 | 91.27% | 7.23% | 0.15 |
| 105 | 23 | 68.85% | 14.92% | 0.32 |
| 106 | 23 | 66.02% | 36.25% | 0.33 |
| 107 | 23 | 78.00% | 22.19% | 0.30 |
| 108 | 24 | 66.14% | 38.05% | 0.34 |
| 109 | 25 | 71.11% | 18.77% | 0.30 |
| 110 | 25 | 55.68% | 27.98% | 0.44 |
| 111 | 25 | 61.33% | 21.80% | 0.39 |
| 112 | 25 | 64.93% | 19.38% | 0.44 |
| 113 | 25 | 83.27% | 14.59% | 0.29 |
| 114 | 26 | 52.52% | 32.49% | 0.34 |
| 115 | 26 | 72.93% | 13.81% | 0.14 |
| 116 | 26 | 71.96% | 19.68% | 0.15 |
| 117 | 27 | 80.39% | 10.68% | 0.23 |
| 118 | 27 | 76.14% | 21.21% | 0.29 |
| 119 | 27 | 72.56% | 24.50% | 0.34 |

Table C.1 – continued

| Image no. | Patient no. | Sørenson-Dice | Abs. Vol. Diff. | Hausdorff (mm) |
|-----------|-------------|---------------|-----------------|----------------|
| 120 | 28 | 55.92% | 34.87% | 0.59 |
| 121 | 28 | 62.42% | 15.03% | 0.53 |
| 122 | 28 | 54.99% | 27.49% | 0.57 |
| 123 | 28 | 31.08% | 40.90% | 0.85 |
| 124 | 28 | 67.34% | 10.80% | 0.42 |
| 125 | 28 | 46.57% | 21.10% | 0.68 |
| 126 | 29 | 69.24% | 15.70% | 0.20 |
| 127 | 30 | 74.39% | 16.76% | 0.27 |
| 128 | 31 | 75.87% | 15.88% | 0.23 |
| 129 | 31 | 72.08% | 16.72% | 0.27 |
| 130 | 31 | 65.72% | 10.91% | 0.25 |

APPENDIX D

DETECTION RESULTS

Table D.1: A table showing the results of applying an empirical threshold (0.85) to the chamber volume ratios of 130 4-D fetal echocardiographic images for detecting/flagging hypoplastic left heart syndrome. Chamber ratio values less than 0.85 are classified as hypoplastic left heart syndrome, while chamber ratios greater than or equal to 0.85 are classified as normal.

| Image no. | Patient no. | Chamber Vol. Ratio | Confirmed Diagnosis | Detection Result |
|-----------|-------------|--------------------|---------------------|------------------|
| 1 | 1 | 0.83 | HLHS | HLHS |
| 2 | 1 | 0.70 | HLHS | HLHS |
| 3 | 1 | 0.86 | HLHS | Normal |
| 4 | 1 | 0.55 | HLHS | HLHS |
| 5 | 1 | 0.47 | HLHS | HLHS |
| 6 | 2 | 0.72 | HLHS | HLHS |
| 7 | 2 | 0.66 | HLHS | HLHS |
| 8 | 2 | 0.94 | HLHS | Normal |
| 9 | 2 | 0.73 | HLHS | HLHS |
| 10 | 3 | 0.59 | HLHS | HLHS |
| 11 | 3 | 0.56 | HLHS | HLHS |
| 12 | 4 | 0.96 | Normal | Normal |
| 13 | 5 | 0.93 | Normal | Normal |
| 14 | 5 | 0.92 | Normal | Normal |
| 15 | 5 | 1.00 | Normal | Normal |
| 16 | 5 | 0.93 | Normal | Normal |
| 17 | 5 | 0.97 | Normal | Normal |
| 18 | 6 | 0.74 | Normal | HLHS |
| 19 | 6 | 0.86 | Normal | Normal |
| 20 | 6 | 0.90 | Normal | Normal |
| 21 | 6 | 0.85 | Normal | Normal |
| 22 | 6 | 0.97 | Normal | Normal |
| 23 | 6 | 0.81 | Normal | HLHS |
| 24 | 6 | 0.99 | Normal | Normal |
| 25 | 7 | 0.89 | Normal | Normal |
| 26 | 7 | 0.98 | Normal | Normal |
| 27 | 7 | 0.92 | Normal | Normal |

Table D.1 – continued

| Image no. | Patient no. | Chamber Vol. Ratio | Confirmed Diagnosis | Detection Result |
|-----------|-------------|--------------------|---------------------|------------------|
| 28 | 7 | 0.61 | Normal | HLHS |
| 29 | 8 | 0.70 | Normal | HLHS |
| 30 | 8 | 0.99 | Normal | Normal |
| 31 | 8 | 0.98 | Normal | Normal |
| 32 | 8 | 0.86 | Normal | Normal |
| 33 | 9 | 0.91 | Normal | Normal |
| 34 | 9 | 0.94 | Normal | Normal |
| 35 | 10 | 0.98 | Normal | Normal |
| 36 | 10 | 0.96 | Normal | Normal |
| 37 | 11 | 0.95 | Normal | Normal |
| 38 | 11 | 0.66 | Normal | HLHS |
| 39 | 11 | 0.92 | Normal | Normal |
| 40 | 11 | 0.92 | Normal | Normal |
| 41 | 11 | 0.96 | Normal | Normal |
| 42 | 12 | 0.92 | Normal | Normal |
| 43 | 12 | 0.85 | Normal | Normal |
| 44 | 13 | 0.98 | Normal | Normal |
| 45 | 13 | 0.66 | Normal | HLHS |
| 46 | 13 | 0.81 | Normal | HLHS |
| 47 | 13 | 1.00 | Normal | Normal |
| 48 | 13 | 0.88 | Normal | Normal |
| 49 | 13 | 0.82 | Normal | HLHS |
| 50 | 14 | 0.96 | Normal | Normal |
| 51 | 14 | 0.86 | Normal | Normal |
| 52 | 14 | 0.94 | Normal | Normal |
| 53 | 14 | 0.99 | Normal | Normal |
| 54 | 14 | 0.98 | Normal | Normal |
| 55 | 15 | 0.58 | HLHS | HLHS |
| 56 | 15 | 0.63 | HLHS | HLHS |
| 57 | 15 | 0.56 | HLHS | HLHS |
| 58 | 15 | 0.77 | HLHS | HLHS |
| 59 | 15 | 0.84 | HLHS | HLHS |
| 60 | 15 | 0.49 | HLHS | HLHS |
| 61 | 16 | 0.84 | Normal | HLHS |
| 62 | 16 | 0.71 | Normal | HLHS |
| 63 | 16 | 0.68 | Normal | HLHS |
| 64 | 16 | 0.84 | Normal | HLHS |
| 65 | 16 | 0.97 | Normal | Normal |
| 66 | 16 | 0.94 | Normal | Normal |
| 67 | 16 | 0.80 | Normal | HLHS |
| 68 | 17 | 0.91 | Normal | Normal |
| 69 | 17 | 0.89 | Normal | Normal |
| 70 | 17 | 0.96 | Normal | Normal |

Table D.1 – continued

| Image no. | Patient no. | Chamber Vol. Ratio | Confirmed Diagnosis | Detection Result |
|-----------|-------------|--------------------|---------------------|------------------|
| 71 | 17 | 0.87 | Normal | Normal |
| 72 | 17 | 0.95 | Normal | Normal |
| 73 | 17 | 0.97 | Normal | Normal |
| 74 | 17 | 0.95 | Normal | Normal |
| 75 | 17 | 0.97 | Normal | Normal |
| 76 | 17 | 0.80 | Normal | HLHS |
| 77 | 18 | 0.95 | Normal | Normal |
| 78 | 18 | 0.58 | Normal | HLHS |
| 79 | 18 | 0.88 | Normal | Normal |
| 80 | 18 | 0.78 | Normal | HLHS |
| 81 | 19 | 0.98 | Normal | Normal |
| 82 | 19 | 0.86 | Normal | Normal |
| 83 | 19 | 0.78 | Normal | HLHS |
| 84 | 19 | 0.93 | Normal | Normal |
| 85 | 19 | 0.74 | Normal | HLHS |
| 86 | 19 | 0.58 | Normal | HLHS |
| 87 | 19 | 0.80 | Normal | HLHS |
| 88 | 20 | 0.85 | Normal | Normal |
| 89 | 20 | 0.96 | Normal | Normal |
| 90 | 20 | 0.95 | Normal | Normal |
| 91 | 20 | 0.96 | Normal | Normal |
| 92 | 21 | 0.95 | Normal | Normal |
| 93 | 21 | 0.98 | Normal | Normal |
| 94 | 21 | 0.92 | Normal | Normal |
| 95 | 21 | 0.56 | Normal | HLHS |
| 96 | 22 | 0.97 | Normal | Normal |
| 97 | 22 | 0.91 | Normal | Normal |
| 98 | 22 | 0.91 | Normal | Normal |
| 99 | 22 | 0.98 | Normal | Normal |
| 100 | 22 | 0.99 | Normal | Normal |
| 101 | 22 | 0.99 | Normal | Normal |
| 102 | 22 | 0.90 | Normal | Normal |
| 103 | 23 | 0.92 | Normal | Normal |
| 104 | 23 | 0.99 | Normal | Normal |
| 105 | 23 | 0.98 | Normal | Normal |
| 106 | 23 | 0.69 | Normal | HLHS |
| 107 | 23 | 0.99 | Normal | Normal |
| 108 | 24 | 0.67 | Normal | HLHS |
| 109 | 25 | 0.82 | Normal | HLHS |
| 110 | 25 | 0.93 | Normal | Normal |
| 111 | 25 | 0.86 | Normal | Normal |
| 112 | 25 | 0.91 | Normal | Normal |
| 113 | 25 | 0.97 | Normal | Normal |

Table D.1 – continued

| Image no. | Patient no. | Chamber Vol. Ratio | Confirmed Diagnosis | Detection Result |
|------------------|--------------------|---------------------------|----------------------------|-------------------------|
| 114 | 26 | 0.47 | HLHS | HLHS |
| 115 | 26 | 0.77 | HLHS | HLHS |
| 116 | 26 | 0.69 | HLHS | HLHS |
| 117 | 27 | 0.99 | Normal | Normal |
| 118 | 27 | 0.89 | Normal | Normal |
| 119 | 27 | 0.97 | Normal | Normal |
| 120 | 28 | 0.83 | Normal | HLHS |
| 121 | 28 | 0.96 | Normal | Normal |
| 122 | 28 | 0.86 | Normal | Normal |
| 123 | 28 | 0.92 | Normal | Normal |
| 124 | 28 | 0.90 | Normal | Normal |
| 125 | 28 | 0.91 | Normal | Normal |
| 126 | 29 | 0.88 | Normal | Normal |
| 127 | 30 | 0.97 | Normal | Normal |
| 128 | 31 | 0.91 | Normal | Normal |
| 129 | 31 | 0.78 | Normal | HLHS |
| 130 | 31 | 1.00 | Normal | Normal |

REFERENCES

- [1] M. Abdellaoui, E. Petit, and J. Lemoine. Edge dynamical model applied to ventricular boundary detection on echocardiographic images. *Computers in Cardiology*, 18:485–488, 1991.
- [2] E. Adelson and J. Bergen. The early detection of motion boundaries. In *IEEE Proc of Workshop on Visual Motion*, pages 151–156, Charleston, 1986.
- [3] M. Ali, D. Magee, and U. Dasgupta. Signal processing overview of ultrasound systems for medical imaging. *Texas Instruments, White Paper*, 2008.
- [4] E. Allgower and K. Georg. *Introduction to Numerical Continuation Methods*. Springer, Berlin, Heidelberg, 1990.
- [5] A. Amini, T. Weymouth, and R. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Trans Patt Anal Mach Intell*, 12(9):855–867, 1990.
- [6] E. Angelini, A. Laine, S. Takuma, J. Holmes, and S. Homma. Lv volume quantification via spatiotemporal analysis of real-time 3-d echocardiography. *IEEE Trans Med Imag*, 20(6):457–469, 2001.
- [7] E. Ashton and K. Parker. Multiple resolution bayesian segmentation of ultrasound images. *Ultrason Imag*, 17(4):291–304, 1995.
- [8] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *Int J Comput Vis*, 92(1):1–31, 2011.
- [9] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *Int J Comput Vis*, 12(1):43–77, 1994.
- [10] M. Basu. Gaussian-based edge-detection methods-a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions*, 32(3):252–260, 2002.
- [11] S. Beauchemin and J. Barron. The computation of optical flow. *ACM Surveys*, 27(3):433–467, 1995.
- [12] M. Bertero and T. Poggio. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76(8):869–889, 1988.
- [13] T. Binder, M. Sussner, D. Moertl, H. Strohmer, T. Baumgartner, G. Maurer, and G. Porenta. Artificial neural networks and spatial temporal contour linking for automated endocardial contour detection on echocardiograms: A novel approach to determine left ventricular contractile function. *Ultrasound Med Biol*, 25(7):1069–1076, 1999.

- [14] J. Bosch, S. Mitchell, B. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. Reiber. Automatic segmentation of echocardiographic sequences by active appearance motion models. *IEEE Trans Med Imag*, 21(11):1374–1383, 2002.
- [15] D. Boukerroui, A. Baskurt, J. Noble, and O. Basset. Segmentation of ultrasound images - multiresolution 2-d and 3-d algorithm based on global and local statistics. *Pattern Recognit Lett*, 24(4-5):779–790, 2003.
- [16] D. Boukerroui, J. Noble, and M. Brady. Velocity estimation in ultrasound images: A block-matching approach. In *Information Processing in Medical Imaging, ser Lect Note Comput Sci*, pages 586–598, Berlin, 2003. Springer-Verlag.
- [17] D. Boukerroui, J. Noble, M. Robini, and M. Brady. Enhancement of contrast regions in suboptimal ultrasound images with application to echocardiography. *Ultrasound Med Biol*, 27(12):1583–1594, 2001.
- [18] A. Bovik. On detecting edges in speckle imagery. *IEEE Trans. Acoust., Speech, Signal Processing*, 36(4-5):1618–1627, 1988.
- [19] P. Bradley, O. L. Mangasarian, and W. N. Street. Clustering via concave minimization. In *Advances in Neural Information Processing Systems -9*, pages 368–374. MIT Press, 1997.
- [20] T. Buck, A. Franke, and M. Monaghan. *Three-dimensional Echocardiography*. Springer, Berlin, Heidelberg, 2011.
- [21] C. Burckhardt. Speckle in ultrasound b-mode scans. *Sonics and Ultrasonics, IEEE Transactions*, 25(1):1–6, 1978.
- [22] M. Campani and A. Verri. Computing optical flow from an overconstrained system of linear equations. In *Proc IEEE Int Conf Comput Vis (ICCV'90)*, pages 22–26, 1990.
- [23] J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, PAMI-8(6):679–698, 1986.
- [24] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours. *Numer Math*, 66:1–31, 1993.
- [25] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proc IEEE Int Conf Comput Vis (ICCV'95)*, pages 694–699, 1995.
- [26] V. Chalana, D. Linker, D. Haynor, and Y. Kim. A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Trans Med Imag*, 15(3):290–298, 1996.
- [27] Y. Chen, F. Huang, H. Tagare, M. Rao, D. Wilson, and E. Geiser. Using prior shape and intensity profile in medical image segmentation. In *Proc IEEE Int Conf Comput Vis (ICCV'03)*, pages 1117–1125, Nice, France, 2003.
- [28] Y. Chen, F. Huang, H. Tagare, M. Rao, D. Wilson, and E. Geiser. A coupled minimization problem for medical image segmentation with priors. *Int J Comput Vis*, 71(3):259–272, 2007.
- [29] C. Chu and E. Delp. Estimating displacement vectors from an image sequence. *J Opt Soc Am*, 6(6):871–878, 1989.

- [30] C. Chu and D. Razi. A nonlinear multiresolution approach to echocardiographic image segmentation. *Computers in Cardiology*, 17:431–434, 1990.
- [31] L. Cohen. On active contour models and balloons. *CVGIP: Image Understand*, 53(2):211–218, 1991.
- [32] L. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-d and 3-d images. *IEEE Trans Pattern Anal Machine Intell*, 15:1131–1147, 1991.
- [33] OpenStax College. Development of blood vessels and fetal circulation, June 2013. [Online; accessed 26-February-2014].
- [34] S. Collins, D. Skorton, and E. Geiser. Computer-assisted edge detection in two-dimensional echocardiography: Comparison with anatomic data. *Am J Cardiol*, 53:1380–1387, 1984.
- [35] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. In H. Burkhardt and B. Neumann, editors, *5th Eur Conf Comput Vis*, volume 2, pages 484–498, 1998.
- [36] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models: Their training and application. *Comput Vis Imag Und*, 61(1):38–59, 1995.
- [37] G. Coppini, R. Poli, and G. Valli. Recovery of the 3d shape of the left ventricle from echocardiographic images. *IEEE Trans Med Imag*, 14:301–317, 1995.
- [38] C. Corsi, G. Saracino, A. Sarti, and C. Lamberti. Left ventricular volume estimation for real-time three-dimensional echocardiography. *IEEE Trans Med Imag*, 21(9):1202–1208, 2002.
- [39] R. Davies, C. Twining, and C. Taylor. *Statistical Models of Shape: Optimisation and Evaluation*. Springer, Manchester, UK, 2008.
- [40] P. Detmer, G. Bashein, and R. Martin. Matched filter identification of left-ventricular endocardial borders in transesophageal echocardiograms. *IEEE Trans Med Imag*, 9:396–404, 1990.
- [41] G. DeVore, P. Falkensammer, M. Sklansky, and L. Platt. Spatio-temporal image correlation (stic): New technology for evaluation of the fetal heart. *Ultrasound in Obstetrics and Gynecology*, 22(4):380–387, 2003.
- [42] G. DeVore, B. Siassi, and L. Platt. Fetal echocardiography. iv. m-mode assessment of ventricular size and contractility during the second and third trimesters of pregnancy in the normal fetus. *Am J Obstet Gynecol*, 150:981–988, 1984.
- [43] I. Dhillon and D. Dharmendra. Concept decompositions for large sparse text data using clustering. *Machine Learning*, 42(1-2):143–175, 2001.
- [44] I. Dhillon, Y. Guan, and B. Kulis. Kernel k-means: Spectral clustering and normalized cuts. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 551–556, New York, NY, USA, 2004. ACM.
- [45] J. Dias and J. Leitão. Wall position and thickness estimation from sequences of echocardiograms images. *IEEE Trans Med Imag*, 15(1):25–38, 1996.

- [46] L. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
- [47] I. Dindoyal. *Foetal Echocardiographic Segmentation*. Phd dissertation, University College London, London, UK, 2009.
- [48] I. Dindoyal, T. Lambrou, J. Deng, C. Ruff, A. Linney, and A. Todd-Pokropek. An active contour model to segment foetal cardiac ultrasound data. *Proceedings of Medical Image Understanding and Analysis, (University of Sheffield, UK, 2003)*, pages 77–80, 2003.
- [49] I. Dindoyal, T. Lambrou, J. Deng, and A. Todd-pokropek. Automatic segmentation of low resolution fetal cardiac data using snakes with shape priors. In *Image and Signal Processing and Analysis, 2007. ISPA 2007. 5th International Symposium*, pages 538–543, Sept 2007.
- [50] I. Dindoyal, T. Lambrou, J. Deng, and A. Todd-pokropek. Level set snake algorithms on the fetal heart. In *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium*, pages 864–867, April 2007.
- [51] I. Dindoyal, T. Lambrou, J. Deng, and A. Todd-Pokropek. 2d/3d fetal cardiac dataset segmentation using a deformable model. *Med Phys*, 38(7):4338–4349, 2011.
- [52] E. Dougherty. *An Introduction to Morphological Image Processing*. SPIE Optical Engineering Press, Bellingham, Wash., USA, 1992.
- [53] D. Downey, A. Fenster, and J. Williams. Clinical utility of three-dimensional ultrasound. *Radiographics*, 20(2):559–571, 2000.
- [54] M-P. Dubuisson and A. Jain. A modified hausdorff distance for object matching. In *Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing, Proceedings of the 12th IAPR International Conference*, volume 1, pages 566–568. IEEE, Oct 1994.
- [55] F. Duck. *Physical Properties of Tissue, A Comprehensive Reference Book*. Academic Press, London, 1990.
- [56] D. Evans and W. McDickens. *Doppler Ultrasound, Physics, Instrumentation, and Signal Processing*. Wiley, West Sussex, second edition, 2000.
- [57] J. Feng, W. Lin, and C. Chen. Epicardial boundary detection using fuzzy reasoning. *IEEE Trans Med Imag*, 10(2):187–199, 1991.
- [58] R. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures on Physics*. Addison-Wesley, 1963. 3 volumes.
- [59] C. Firpo, J. Hoffman, and N. Silverman. Evaluation of fetal heart dimensions from 12 weeks to term. *Am J Cardiol*, 87:594–600, 2001.
- [60] D. Fleet and A. Jepson. Computation of component image velocity from local phase information. *Int J Comput Vis*, 5(1):77–104, 1990.
- [61] Centers for Disease Control and Prevention. Congenital heart defects, 2013. [Online; accessed 2-September-2013].

- [62] E. Forgy. Cluster analysis of multivariate data: Efficiency versus interpretability of classifications. *Biometrics*, 21:768–769, 1965.
- [63] A. Frangi, W. Niessen, and M. Viergever. Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE Trans Med Imag*, 20(1), 2001.
- [64] N. Friedland and D. Adam. Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated anneal. *IEEE Trans Med Imag*, 8(4):344–353, 1989.
- [65] R. Galloway, B. McDermott, and F. Thurstone. A frequency diversity process for speckle reduction in real-time ultrasonic images. *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, 35:45–49, 1988.
- [66] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills. Recovering motion fields: An analysis of eight optical flow algorithms. In *British Machine Vision Conference*, Southampton, England, September 1998.
- [67] E. Garcia, A. Ezekiel, and R. Levy. Automated computer enhancement and analysis of left ventricular two-dimensional echocardiograms. *Computers in Cardiology*, 9:399–402, 1982.
- [68] E. Garcia, P. Gueret, and M. Bennett. Real-time computerization of two-dimensional echocardiography. *Am Heart J*, 101:783–792, 1981.
- [69] E. Gill, W. Murrat, and M. Wright. *Practical Optimization*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1981.
- [70] A. Go, D. Mozaffarian, V. Roger, E. Benjamin, J. Berry, W. Borden, D. Bravata, S. Dai, E. Ford, C. Fox, S. Franco, H. Fullerton, C. Gillespie, S. Hailpern, J. Heit, V. Howard, M. Huffman, B. Kissela, S. Kittner, D. Lackland, J. Lichtman, L. Lisabeth, D. Magid, G. Marcus, A. Marelli, D. Matchar, D. McGuire, E. Mohler, C. Moy, M. Mussolino, G. Nichol, N. Paynter, P. Schreiner, P. Sorlie, J. Stein, T. Turan, S. Virani, N. Wong, D. Woo, and M. Turner. Heart disease and stroke statistics—2013 update: A report from the american heart association. *Circulation*, 127(1):e6–e245, 2013.
- [71] R. Gonzalez and R. Woods. *Digital Image Processing*. Prentice Hall, Upper Saddle River, N.J., third edition, 2007.
- [72] A. Gopal, M. Schnellbaecher, Z. Shen, L. Boxt, J. Katz, and D. King. Freehand three-dimensional echocardiography for determination of left ventricular volume and mass in patients with abnormal ventricles: comparison with magnetic resonance imaging. *J Am Soc Echocardiogr*, 10:853–861, 1997.
- [73] F. Guichard. A morphological, affine, and galilean invariant scalespace for movies. *IEEE Trans Med Imag*, 7(3):444–456, 1998.
- [74] S. Gupta, R. Chauhan, and S. Saxena. Robust non-homomorphic approach for speckle reduction in medical ultrasound images. *Medical and Biological Engineering and Computing*, 43(2):189–195, 2005.
- [75] C. Han, K. Lin, W. Wee, R. Mintz, and D. Porembka. Knowledge-based image analysis for automated boundary extraction of transesophageal echocardiographic left-ventricular images. *IEEE Trans Med Imag*, 10:602–610, 1991.

- [76] M. Handschumacher, J. Lethor, S. Siu, D. Mele, J. Rivera, and M. Picard. A new integrated system for three-dimensional echocardiographic reconstruction: Development and validation for ventricular volume with application in human subjects. *J Am Coll Cardiol*, 21:743–753, 1993.
- [77] R. Haralick. Zero crossing of second directional derivative edge operator. *SPIE Symp. Robot Vision*, 336:91–99, May 1982.
- [78] D. Heeger. Optical flow using spatiotemporal filters. *Int J Comput Vis*, 1:279–302, 1988.
- [79] T. Heimann and H. Meinzer. Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis*, 13:543–563, 2009.
- [80] F. Heitz, P. Perez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. *CVGIP: Image Understanding*, 59:125–134, 1994.
- [81] I. Herlin and N. Ayache. Features extraction and analysis methods for sequences of ultrasound images. *Image and Vision Computing*, 10:673–682, 1992.
- [82] N. Hiransakolwong, K. Hua, Vu Khanh, and P. Windyga. Segmentation of ultrasound liver images: An automatic approach. In *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference*, volume 1, pages I–573–6, July 2003.
- [83] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.
- [84] P. Hoskins, K. Martin, and A. Thrush. *Diagnostic Ultrasound, Physics, and Equipment*. Cambridge University Press, Cambridge, second edition, 2010.
- [85] T. Huang, G. Yang, and G. Tang. A fast two-dimensional median filtering algorithm. *IEEE Trans. Acoust. Speech Signal Processing*, 27(1):13–18, 1979.
- [86] I. Hunter, J. Soraghan, J. Christie, and T. Durrani. Detection of echocardiographic left ventricle boundaries using neural networks. *Computers in Cardiology*, 20:201–204, 1993.
- [87] I. Hunter, J. Soraghan, and T. McDonagh. Fully automatic left ventricular boundary extraction in echocardiographic images. *Computers in Cardiology*, 22:741–744, 1995.
- [88] D. Huttenlocher, G. Klanderman, and W. Rucklidge. Comparing images using the hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(9):850–863, 1993.
- [89] G. Jacob, J. Noble, C. Behrenbruch, A. Kelion, and A. Banning. A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography. *IEEE Trans Med Imag*, 21(3):226–238, 2002.
- [90] G. Jacob, J. Noble, A. Kelion, and A. Banning. Quantitative regional analysis of myocardial wall motion. *Ultrasound Med Biol*, 27(6):773–784, 2001.
- [91] G. Jacob, J. Noble, M. Mulet-Parada, and A. Blake. Evaluating a robust contour tracker on echocardiographic sequences. *Med Image Anal*, 3:63–75, 1999.
- [92] M. Jenkin, A. Jepson, and J. Tsotsos. Techniques for disparity measurement. *Comput Vis Graph Imag Process*, 53(1):14–30, 1991.

- [93] J. Jenkins, G. Qian, M. Besozzi, E. Delp, and A. Buda. Computer processing of echocardiographic images for automated edge detection of left ventricular boundaries. *Computers in Cardiology*, 8:391–394, 1981.
- [94] L. Jiang, J. Vazquez de Prada, M. Handschumacher, J. Guererro, G. Vlahakes, and M. King. Three-dimensional echocardiography: In vivo validation for right ventricular free wall mass as an index of hypertrophy. *J Am Coll Cardiol*, 23:1715–1722, 1994.
- [95] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int J Comput Vis*, 1(4):321–331, 1988.
- [96] E. Kerut, M. Given, E. McIlwain, G. Allen, C. Espinoza, and T. Giles. Echocardiographic texture analysis using the wavelet transform: Differentiation of early heart muscle disease. *Ultrasound Med Biol*, 26(9):1445–1453, 2000.
- [97] S. Khor, F. Szaboki, J. Nieberl, M. Khor, and E. Kekes. A new method of echocardiographic edge detection using velocity maps. *Computers in Cardiology*, 20:623–625, 1993.
- [98] D. King, M. Harrison, D. King Jr., A. Gopal, R. Martin, and A. DeMaria. Improved reproducibility of left atrial and left ventricular measurements by guided three-dimensional echocardiography. *J Am Coll Cardiol*, 20:1238–1245, 1992.
- [99] J. Klingler, C. Vaughan, T. Fraker, and L. Andrews. Segmentation of echocardiographic images using mathematical morphology. *IEEE Trans Biomed Eng*, 35(11):925–934, 1988.
- [100] R. Kories and G. Zimmerman. Versatile method for the estimation of displacement vector fields from image sequences. In *IEEE Proc of Workshop on Motion: Representation and Analysis*, pages 101–106, 1986.
- [101] D. Kucera and R. Martin. Segmentation of sequences of echocardiographic images using a simplified 3-d active contour model with region-based external forces. *Comput Med Imag Graph*, 21(1):1–21, 1997.
- [102] R. Lang, M. Bierig, R. Devereux, F. Flachskampf, E. Foster, P. Pellikka, M. Picard, M. Roman, J. Seward, J. Shanewise, S. Solomon, K. Spencer, M. St. John Sutton, and W. Stewart. Recommendations for chamber quantification. *European Journal of Echocardiography*, 7(2):79–108, 2006.
- [103] T. Lassige, P. Benkeser, D. Fyfe, and S. Sharma. Comparison of septal defects in 2d and 3d echocardiography using active contour models. *Comput. Med. Imaging Graph.*, 24(6):377–388, 2000.
- [104] B. Leroy, I. Herlin, and L. Cohen. Multi-resolution algorithms for active contour models. In *Proc Int Conf Analy and Optimiz of Sys*, pages 58–65, 1996.
- [105] M. Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3):29–37, 1988.
- [106] W. Liang, S. DeJong, C. McKay, and M. Sonka. Automated detection of left ventricular borders from intracardiac ultrasound images. *Computers in Cardiology*, 9:265–268, 1996.

- [107] C-J. Lin, C-Y. Lin, C-H. Chen, B. Zhou, and C-P. Chang. Partitioning the heart: Mechanisms of cardiac septation and valve development. *Development*, 139(18):3277–3299, 2012.
- [108] N. Lin, W. Yu, and J. Duncan. Combinative multi-scale level set framework for echocardiographic image segmentation. *Med Image Anal*, 7(4):529–537, 2003.
- [109] J. Little. Accurate early detection of discontinuities. In *Vision Interface*, pages 2–7, 1992.
- [110] J. Little, H. Butlhoff, and T. Poggio. Parallel optical flow using local voting. In *Proc Int Conf Comput Vis*, pages 454–459, 1988.
- [111] C. Loizou and C. Pattichis. *Despeckle Filtering Algorithms and Software for Ultrasound Imaging*. Morgan & Claypool, 2008.
- [112] T. Loupas, W. McDicken, and P. Allan. An adaptive weighted median filter for speckle suppression in medical ultrasound images. *IEEE Transactions on Circuits and Systems*, 36(1):129–135, 1989.
- [113] B. Lucas. *Generalized Image Matching By the Method of Differences*. Phd dissertation, Carnegie-Mellon University, Pittsburgh, PA, 1984.
- [114] B. Lucas and T. Kanade. An iterative image-registration technique with an application to stereo vision. In *DARPA Proc of Image Understanding Workshop*, pages 121–130, 1981.
- [115] M. Luetttgen, W. Karl, and A. Willsky. Efficient multiscale regularization with applications to the computation of optical flow. *IEEE Trans on Imag Proc*, 3(1):41–64, 1994.
- [116] S. Luewan, Y. Yanase, F. Tongprasert, K. Srisupundit, and T. Tongsong. Fetal cardiac dimensions at 14-40 weeks gestation obtained using cardio-stic-m. *Ultrasound Obstet Gynecol*, 37:416–422, 2011.
- [117] C. Lurig, L. Kobbelt, and T. Ertl. Hierarchical solutions for the deformable surface problem in visualization. *Graphical Models*, 62(1):2–18, 2000.
- [118] P. Magnin, O. von Ramm, and F. Thurstone. Frequency compounding for speckle contrast reduction in phased array images. *Ultrasonic Imaging*, 4:267–281, 1982.
- [119] S. Malassiotis and M. Strintzis. Tracking the left ventricle in echocardiographic images by learning heart dynamics. *IEEE Trans Med Imag*, 18(3):282–290, 1999.
- [120] R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans Pattern Anal Mach Intell*, 17(2):158–174, 1995.
- [121] D. Marr and E. Hildreth. Theory of edge detection. *Proc. Roy. Soc.*, B207:187–217, 1980.
- [122] S. Marsland. *Machine Learning: An Algorithmic Perspective*. Chapman & Hall/CRC, 1st edition, 2009.
- [123] D. Matthews and V. Farewell. *Using and Understanding Medical Statistics*. S. Karger Pub., 4 edition, 2007.

- [124] T. McInerney and D. Terzopoulos. T-snakes: Topology adaptive snakes. *Med Imag Anal*, 4:73–91, 2000.
- [125] T. McInerney and D. Terzopoulos. Deformable models. In I. Bankman, editor, *Handbook of Medical Image Processing and Analysis*, pages 145–166. Academic Press, San Diego, 2nd edition, 2008.
- [126] M. Mignotte and J. Meunier. A multiscale optimization approach for the dynamic contour-based boundary detection issue. *Comput Med Imag Graph*, 25(3):265–275, 2001.
- [127] M. Mignotte, J. Meunier, and J. Tardif. Endocardial boundary estimation and tracking in echocardiographic images using deformable template and markov random fields. *Pattern Anal Appl*, 4(4):256–271, 2001.
- [128] I. Mikic, S. Krucinski, and J. D. Thomas. Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates. *IEEE Trans Med Imag*, 17(2):274–284, 1998.
- [129] K. Mikula, T. Preuber, M. Rumpf, and F. Sgallari. On anisotropic geometric diffusion in 3-d image processing and image sequence analysis. *Trends Nonlinear Anal*, pages 305–319, 2000.
- [130] K. Mikula, T. Preuber, M. Rumpf, and F. Sgallari. Morphological image sequence processing. *Comput Vis Sci*, 6(4):197–209, 2004.
- [131] K. Mikula, A. Sarti, and F. Sgallari. *Handbook of Medical Image Analysis: Segmentation and Registration Models*. Marcel Dekker Inc, New York, 2008.
- [132] A. Milkowski, Y. Li, D. Becker, and S. Ishrak. Speckle reduction imaging. *GE Medical Systems, Ultrasound*, 2003.
- [133] A. Mishra, P. Dutta, and M. Ghosh. A ga based approach for boundary detection of left ventricle with echocardiographic image sequences. *Image Vis Comput*, 21:967–976, 2003.
- [134] S. Mitchell, J. Bosch, B. Lelieveldt, R. van der Geest, J. Reiber, and M. Sonka. 3-d active appearance models: Segmentation of cardiac mr and ultrasound images. *IEEE Trans Med Imag*, 21(9):1167–1178, 2002.
- [135] J. Montagnat, M. Sermesant, H. Delingette, G. Malandain, and N. Ayache. Anisotropic filtering for model-based segmentation of 4-d cylindrical echocardiographic images. *Pattern Recognit Lett*, 24(4-5):815–828, 2003.
- [136] N. Mukawa. Estimation of shape reflection coefficients and illuminant direction from image sequences. In *Int Conf on Comput Vis (ICCV'93)*, 1993.
- [137] M. Mulet-Parada and J. Noble. 2d+t acoustic boundary detection in echocardiography. *Med Image Anal*, 4(1):21–30, 2000.
- [138] J. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, Menlo Park, CA, 1984.
- [139] H. Nagel. On a constraint equation for the estimation of displacement rates in image sequences. *IEEE Pattern Anal and Mach Intell*, 11(1):13–30, 1989.

- [140] S. Nagueh, C. Appleton, T Gillebert, P. Marino, J. Oh, O. Smiseth, A. Waggoner, F. Flachskampf, P. Pellikka, and A. Evangelista. Recommendations for the evaluation of left ventricular diastolic function by echocardiography. *Journal of the American Society of Echocardiography*, 22(2):107–133, 2009.
- [141] S. Negahdaripour and C. Yu. A generalized brightness change model for computing optical flow. In *Int Conf on Comput Vis (ICCV'93)*, pages 2–11, 1993.
- [142] J. Noble and D. Boukerroui. Ultrasound image segmentation: A survey. *IEEE Trans Med Imag*, 25(8):987–1010, 2006.
- [143] C. Oliver, I. McConnell, D. Blacknell, and R. White. Optimal edge detection in sar. in *Proc. Int. Conf. Remote Sens.*, 2584:152–163, 1995.
- [144] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulation. *J Comput Phys*, 79:12–49, 1988.
- [145] T. Pavlidis and Y. Liow. Integrating region growing and edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 12(3):225–233, 1990.
- [146] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Machine Intell.*, 2:629–639, 1990.
- [147] C. Perreault and M-F. Auclair-Fortier. Speckle simulation based on b-mode echographic image acquisition model. In *Computer and Robot Vision, 2007. CRV '07. Fourth Canadian Conference on*, pages 379–386, 2007.
- [148] L. Piccoli, A. Dahmer, J. Scharcanski, and P. Navaux. Fetal echocardiographic image segmentation using neural networks. In *7th Int Conf on Imag Proc Appl*, volume 2, pages 507–511, Manchester, 1999.
- [149] C. Poon, M. Braun, R. Fahrig, A. Ginige, and A. Dorrell. Segmentation of medical images using an active contour model incorporating region-based images features. In *3rd Conf Vis in Biomed Comp (VBC '94)*, volume 2359, Rochester, MN, 1994.
- [150] S. Poonguzhali and G. Ravindran. A complete automatic region growing method for segmentation of masses on ultrasound images. In *Biomedical and Pharmaceutical Engineering, 2006. ICBPE 2006. International Conference*, pages 88–92, Dec 2006.
- [151] T. Preuber and M. Rumpf. A level set method for anisotropic geometric diffusion in 3-d image processing. *SIAM J Appl Math*, 62(5):1772–1793, 2002.
- [152] J. Prince and E. McVeigh. Motion estimation from tagged mr image sequences. *IEEE Trans Med Imag*, 11(2):238–249, 1992.
- [153] P. Rajiah, C. Mak, T. Dubinsky, and M. Dighe. Ultrasound of fetal cardiac anomalies. *American Journal of Roentgenology*, 197(4):747–760, 2011.
- [154] R. Real and J. Vargas. The probabilistic basis of jaccard's index of similarity. *Syst. Biol.*, 45(3):380–385, 1996.
- [155] C. Rekeczky, A. Tahy, Z. Vegh, and T. Roska. Cnn-based spatiotemporal nonlinear filtering and endocardial boundary detection in echocardiography. *Int J Circuit Theory Appl*, 27(1):171–207, 1999.

- [156] T. Sørensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. *Kongelige Danske Videnskabernes Selskab*, 5(4):1–34, 1948.
- [157] L. G. Roberts. Machine perception of three-dimensional solids. In *Optical and Electro-Optical Information Processing*, Cambridge, MA, 1965. MIT Press.
- [158] R. Rockafellar and R. Wets. Variational analysis. In *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1998.
- [159] L. Rudin, P. Lions, and S. Osher. Multiplicative denoising and deblurring: Theory and algorithms. In S. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision, and Graphics*, volume 93, pages 103–122, Berlin, 2003. Springer.
- [160] G. Sakas and S. Walter. Extracting surfaces from fuzzy 3d-ultrasound data. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '95, pages 465–474, New York, NY, USA, 1995. ACM.
- [161] A. Sarti, R. Malladi, and J. Sethian. Subjective surfaces: A geometric model for boundary completion. *Int J Comput Vis*, 46(3):201–221, 2002.
- [162] A. Sarti, E. Mazzini, C. Corsi, and C. Lamberti. Maximum likelihood segmentation of ultrasound images with rayleigh distribution. *IEEE Trans Ultrason Ferroelectr Freq Control*, 52(6):974–960, 2005.
- [163] A. Sarti, K. Mikula, and F. Sgallari. Nonlinear multiscale analysis of 3-d echocardiographic sequences. *IEEE Trans Med Imag*, 18(6):453–466, 1999.
- [164] B. Scholkopf and A. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA, 2001.
- [165] B. Schölkopf, A. Smola, and K-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.*, 10(5):1299–1319, 1998.
- [166] J. Serra. *Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances*. Academic Press, London New York, second edition, 1988.
- [167] S. Setarehdan and J. Soraghan. Automatic left ventricular feature extraction and visualisation from echocardiographic images. *Computers in Cardiology*, 23:9–12, 1996.
- [168] S. Setarehdan and J. Soraghan. Automatic cardiac lv boundary detection and tracking using hybrid fuzzy temporal and fuzzy multiscale edge detection. *IEEE Trans Biomed Eng*, 46(11):1364–1378, 1999.
- [169] J. Sethian. A fast marching level set method for monotonically advancing fronts. In *Proc Nat Acad Sci*, volume 93, pages 1591–1595, San Diego, CA, 1996.
- [170] J. Shan, H. Cheng, and Y. Wang. A novel automatic seed point selection algorithm for breast ultrasound images. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference*, pages 1–4, Dec 2008.
- [171] I. Shapiro, S. Degani, Z. Leibovitz, G. Ohel, Y. Tal, and E. Abinader. Fetal cardiac measurements derived by transvaginal and transabdominal cross-sectional echocardiography from 14 weeks of gestation to term. *Ultrasound Obstet Gynecol*, 12:404–418, 1998.

- [172] G. Sharland. Fetal cardiac screening and variation in prenatal detection rates of congenital heart disease: Why bother with screening at all? *Future Cardiol*, 8(2):189–202, 2012.
- [173] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, New York, NY, USA, 2004.
- [174] K. Sheikh, S. Smith, O. Von Ramm, and J. Kisslo. Real-time three-dimensional echocardiography: Feasibility and initial use. *Echocardiography*, 8:119–125, 1991.
- [175] F. Y. Shih. *Image Processing and Mathematical Morphology: Fundamentals and Applications*. CRC Press, FL, 2009.
- [176] W. Siler, T. Cooper, and R. Hess. A cell automation algorithm for region consolidation and edge detection on sector scans. *Computers in Cardiology*, 9:403–406, 1982.
- [177] E. Simoncelli, E. Adelson, and D. Heeger. Probability distributions of optical flow. In *IEEE Proc of Conf on Comput Vis and Patt Recog (CVPR'91)*, pages 310–315, 1991.
- [178] A. Singh and P. Allen. Image-flow computation: An estimation theoretic framework and a unified perspective. *Computer Vision, Graphics Image Process.: Image Understanding*, 56(2):152–177, 1982.
- [179] M. Siqueira, J. Scharcanski, and P. Navaux. Echocardiographic image sequence segmentation and analysis using self-organizing maps. *J. VLSI Signal Proc. Syst. Signal, Image, Video Technol.*, 32:135–145, 2002.
- [180] S. Siu, J. Rivera, J. Guerrero, M. Handschumacher, J. Lethor, and A. Weyman. Three-dimensional echocardiography: In vivo validation for left ventricular volume and function. *Circulation*, 88:1715–1723, 1993.
- [181] B. Sklar. *Digital Communications, Fundamentals and Applications*. Prentice-Hall, NJ, second edition, 2009.
- [182] D. Skorton, C. McNary, J. Child, F. Newton, and P. Shah. Digital image processing of two-dimensional echocardiograms: Identification of the endocardium. *Am J Cardiol*, 48:479–486, 1981.
- [183] I. E. Sobel. *Camera Models and Machine Perception*. Phd dissertation, Stanford University, Palo Alto, CA, 1970.
- [184] P. Soille. *Morphological Image Analysis : Principles and Applications*. Springer, Berlin, NY, 1999.
- [185] M. Song, R. Haralick, F. Sheehan, and R. Johnson. Integrated surface model optimization for freehand three-dimensional echocardiography. *IEEE Trans Med Imag*, 21(9):1077–1090, 2002.
- [186] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. Thomson-Engineering, 2007.
- [187] L. Staib and J. Duncan. Model-based deformable surface finding for medical images. *IEEE Trans Med Imag*, 15(5):720–731, 1996.
- [188] Sonotech Medical Equipment Supplier. Reflection and refraction, 2013. [Online; accessed 7-October-2013].

- [189] M. Sussner, M. Budil, T. Strohmer, M. Greher, G. Porenta, and T. Binder. Contour detection using artificial neuronal network pre-segmentation. *Computers in Cardiology*, 22:737–740, 1995.
- [190] M. Sutton, W. Walters, W. Peters, W. Ranson, and S. McNeil. Determination of displacement using an improved digital correlation method. *Imag and Vis Comput*, 1(3):133–139, 1983.
- [191] T. Szabo. *Diagnostic Ultrasound Imaging: Inside Out*. Elsevier Academic Press, Burlington, MA, second edition, 2004.
- [192] D. Tenbrinck, S. Schmid, X. Jiang, K. Schäfers, and J. Stypmann. Histogram-based optical flow for motion estimation in ultrasound imaging. *J of Math Imag and Vis*, 47(1-2):138–150, 2013.
- [193] A. Thakur and R. Anand. Image quality based comparative evaluation of wavelet filters in ultrasound speckle reduction. *Digital Signal Processing*, 15(5):455–465, 2005.
- [194] G. Trahey, S. Smith, and O. Von Ramm. Speckle pattern correlation with lateral aperture translation: Experimental results and implications for spatial compounding. *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, 33:257–264, 1986.
- [195] B. Tutschek and D. Sahn. Semi-automatic segmentation of fetal cardiac cavities: Progress towards an automated fetal echocardiogram. *Ultrasound Obstet. Gynecol.*, 32:176–180, 2008.
- [196] S. Pildner von Steinburg, A-L. Boulesteix, C. Lederer, S. Grunow, S. Schiermeier, W. Hatzmann, K-T. Schneider, and M. Daumer. What is the “normal” fetal heart rate? *PeerJ*, 1:e82, 2013.
- [197] A. Watson and A. Ahumada. Model of human visual motion sensing. *J Opt Soc Am*, 2(2):322–341, 1985.
- [198] J. Weber and J. Malik. Robust computation of optical flow in a multi-scale differential framework. In *Proc Int Conf on Comput Vis (ICCV’93)*, pages 12–20, 1993.
- [199] J. Weickert, A. Bruhn, T. Brox, and N. Papenberg. A survey on variational optic flow methods for small displacements. In O. Scherzer, editor, *Mathematical Models for Registration and Applications to Medical Imaging*, volume 10 of *Mathematics in Industry*, pages 103–136. Springer Berlin Heidelberg, Berlin, 2006.
- [200] J. Wladimiroff, R. Vosters, and J McGhie. Normal cardiac ventricular geometry and function during the last trimester of pregnancy and early neonatal period. *Br J Obstet Gynaecol*, 89:839–844, 1982.
- [201] I. Wolf, M. Hastenteufel, R. De Simone, M. Vetter, G. Glombitza, S. Mottl-Link, C. Vahl, and H. Meinzer. Ropes: A semiautomated segmentation method for accelerated analysis of threedimensional echocardiographic data. *IEEE Trans Med Imag*, 21(9):1091–1104, 2002.
- [202] E. Wolfe, E. Delp, C. Meyer, F. Bookstein, and A. Buda. Accuracy of automatically determined borders in digital two-dimensional echocardiography using a cardiac phantom. *IEEE Trans Med Imag*, 6:292–296, 1987.

- [203] P. Woodward, A. Kennedy, R. Sohaey, J. Bryne, K. Oh, M. Puchalski, T. Winter, and L. McLean. *Diagnostic Imaging: Obstetrics*. Amirsys, Canada, second edition, 2011.
- [204] G. Xiao, M. Brady, J. Noble, and Y. Zhang. Segmentation of ultrasound b-mode images with intensity inhomogeneity correction. *IEEE Trans Med Imag*, 21(1):48–57, 2002.
- [205] C. Xu and J. Prince. Snakes, shapes and gradient vector flow. *IEEE Trans Imag Process*, 7(3):359–369, 1998.
- [206] R. Xu and D. Wunsch II. Survey of clustering algorithms. *IEEE Trans Neural Networks*, 16(3):645–678, 2005.
- [207] J. Yan and T. Zhuang. Applying improved fast marching method to endocardial boundary detection in echocardiographic images. *Pattern Recognit Lett*, 24(15):2777–2784, 2003.
- [208] A. Yuille, D. Cohen, and P. Hallinan. Feature extraction from faces using deformable templates. In *Proc IEEE Comput Vis Patt Recog (CVPR'89)*, pages 104–109, San Diego, CA, 1989.
- [209] A. Yuille and T. Poggio. Scaling theorems for zero crossings. *IEEE trans on Pattern Anal and Mach Intell*, 8(1):15–25, 1986.
- [210] L. Zhang and E. Geiser. An effective algorithm for extracting serial endocardial borders from two-dimensional echocardiograms. *IEEE Trans Biomed Eng*, 31:441–447, 1984.
- [211] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. Four-chamber heart modeling and automatic segmentation for 3-d cardiac ct volumes using marginal space learning and steerable features. *Medical Imaging, IEEE Transactions*, 27(11):1668–1681, 2008.
- [212] X. Zhuang, C. Yao, Y. Ma, D. Hawkes, G. Penney, and S. Ourselin. Registration-based propagation for whole heart segmentation from compounded 3d echocardiography. In *IEEE Int Sym Biomed Imag*, pages 1093–1096, Rotterdam, 2010.
- [213] X. Zoon, A. Laine, and E. Geiser. Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. *Medical Imaging, IEEE Transactions*, 17(4):532–540, 1998.