FOETAL ECHOCARDIOGRAPHIC SEGMENTATION

Irving Dindoyal

Department of Medical Physics and Bioengineering
University College London

This thesis is submitted as a requirement for the Doctorate of Philosophy

July 2009

DECLARATION

I, Irving Dindoyal confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Congenital heart disease affects just under one percentage of all live births [1]. Those defects that manifest themselves as changes to the cardiac chamber volumes are the motivation for the research presented in this thesis.

Blood volume measurements in vivo require delineation of the cardiac chambers and manual tracing of foetal cardiac chambers is very time consuming and operator dependent. This thesis presents a multi region based level set snake deformable model applied in both 2D and 3D which can automatically adapt to some extent towards ultrasound noise such as attenuation, speckle and partial occlusion artefacts. The algorithm presented is named Mumford Shah Sarti Collision Detection (MSSCD). The level set methods presented in this thesis have an optional shape prior term for constraining the segmentation by a template registered to the image in the presence of shadowing and heavy noise.

When applied to real data in the absence of the template the MSSCD algorithm is initialised from seed primitives placed at the centre of each cardiac chamber. The voxel statistics inside the chamber is determined before evolution. The MSSCD stops at open boundaries between two chambers as the two approaching level set fronts meet. This has significance when determining volumes for all cardiac compartments since cardiac indices assume that each chamber is treated in isolation. Comparison of the segmentation results from the implemented snakes including a previous level set method in the foetal cardiac literature show that in both 2D and 3D on both real and synthetic data, the MSSCD formulation is better suited to these types of data. All the algorithms tested in this thesis are within 2mm error to manually traced segmentation of the foetal cardiac datasets. This corresponds to less than 10% of the length of a foetal heart. In addition to comparison with manual tracings all the amorphous deformable model segmentations in this thesis are validated using a physical phantom. The volume estimation of the phantom by the MSSCD segmentation is to within 13% of the physically determined volume.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ACRONYMS USED IN THIS THESIS

3D+t	Three Dimensional space and time
4CV	Four Chamber View
ASM	Active Shape Model
bpm	Beats Per Minute
CHD	Congenital Heart Disease
CT	Computed Tommography
ECG	ElectroCardioGram
ED	End Diastole
EF	Ejection Fraction
EM	Electromagnetic
EPCA	Edge-Penalised Constant Advection
ES	End Systole
fps/FPS	Frames Per Second
GVF	Gradient Vector Flow
GTC	Generalised Tanimoto Coefficient
ICP	Iterative Closest Point
kHz	Kilohertz
LA	Left Atrium
LS	Level Set
LV	Left Ventricle
MHz	Megahertz
ml	Millilitres
mm	Millimetres
MR/MRI	Magnetic Resonance/Magnetic Resonance Imaging
MS	Mumford-Shah
MSSCD	Mumford-Shah Sarti Collision Detection
*n*D	Number of Dimensions where *n* is a positive integer
PZLS	Pseudo Zero Level Set (in explicit form)
rms/RMS	Root Mean Square
RA	Right Atrium
RV	Right Ventricle
SDM	Signed Distance Map
SP	Shape Prior
TC	Tanimoto Coefficient
TGC	Time Gain Compensation
TIMS	Template Initialised Mumford-Shah
US	Ultrasound

ACKNOWLEDGEMENTS

I would like to thank the many members of both the NHS and UCL Medical Physics & Bioengineering Department who have supported and encouraged me over the years I have been working in this project.

I would like to give my special thanks to Dr Jing Deng without whose help this project would not have been possible. As well as being the Clinical supervisor for the project and providing all of the clinical datasets, he was instrumental in the manual segmentation, and visual inspection of the resultant images.

This work was funded by the EPSRC and MRC as part of the MIAS IRC.

1 INTRODUCTION

Accurate structural acquisition by ultrasound (US) is needed to monitor the health of the myocardium. One particular important pathological case in the foetal heart is the septal defect in which a hole develops in the septal wall allowing blood to leak between atria or between ventricles.

Cardiac volume indices are useful in assessment of functional pathologies. Foetal anaemia can cause the heart to grow larger to compensate and this may be detected by measurement of the chamber volumes. Volumes can be estimated by fitting geometric primitives of the heart chambers to cross sections. However, geometric primitives are a simplification of the chamber shapes and 3D manual segmentation provides more accurate volumes since it makes no assumptions about shapes [2].

The heart is difficult to mentally reconstruct from viewing 2D slices and in the foetal case the walls are thin and are not always resolved by the finite width of the US beam. There has been relatively little work performed in the area of 3D foetal cardiac acquisition due to the lack of reliable foetal ECG gating and unpredictable body movements. Volume acquisition can be made by sweeping out a volume from conventional single slice acquisition probes or recent advances in matrix ultrasound transducer technology allow imaging of entire volumes at video frame rates [3]. The former requires Electrocardiogram (ECG) gating and tracking of the probe to obtain volumetric data without motion artefacts.

From the datasets acquired by the Medical Imaging group at University College London (UCL) the complex structure of the endocardium and its contents, and the severe dropout artefacts, make it ambiguous and time consuming to segment the cardiac chambers manually in a repeatable manner. This thesis presents a method to automatically segment the cardiac chambers by use of a high level deformable model that can flex in a controlled manner to delineate the blood pool under in the presence of partially resolved walls.

1.1 Objectives of Thesis

In the adult heart literature cardiac segmentation is a very active area of ongoing research [4]. The foetal heart on the other hand due to the difficulty in acquiring volumetric data during pregnancy has less development in this area. This thesis provides the following objectives to the foetal cardiac segmentation field:

- Deformable model segmentation of the foetal cardiac chambers with minimal user interaction.
- Ejection Fraction (EF) measurements from automatically determined volumes of the foetal heart and compare them with the manually determined ones and other measured values in the literature.
- The automatic segmentation algorithm must be compared on both real and foetal cardiac synthetic data with other similar algorithms for performance evaluation.
- Quantitative evaluation of algorithms to manual tracings in terms of physical distances.
- Assessment of clinical accuracy.
- Comparison to ground truth by validation of algorithms using a physical phantom.
- Compare performance of the segmentation in both 2D and 3D.
- The ability to provide surface rendering of the foetal cardiac chambers.

1.2 Hypotheses

The hypotheses for the content in this thesis are listed below although a lot of the theory will be explained later. These statements are assumed to hold true for foetal cardiac ultrasound data.

H1. Region based segmentation is more robust than edge based methods.
H2. Segmentation accuracy will be better in 2D compared to 3D due to the reduced resolution of matrix probes.
H3. Shape priors improve segmentation accuracy.
H4. Manual tracing of the endocardium is more accurate than automatic means.

1.3 Structure of Thesis

This thesis presents some of the cardiac pathologies that require reliable 3D segmentation to obtain functional measurements for diagnosis. A brief introduction to foetal echocardiography is then presented. The dominant artefacts that arise from foetal cardiac acquisition are discussed with regard to how they will influence an automatic segmentation algorithm. Next some image preprocessing techniques are described which are required to be run on the data before most computer vision algorithms can analyse the useful content. Deformable models are then introduced in their various forms from the computer vision literature, presented in chronological order. These are presented in two chapters one being unconstrained by shape and the other shape aware. Endocardium wall tracking is then described although it is not implemented in my algorithm to produce the main results of this thesis; it is

presented since tracking is intended for an area of future work. A literature review of the work that has been established in the field of foetal echocardiac segmentation is presented before the algorithm proposed in this thesis is then outlined. This thesis describes a snake (constrained deformable model) algorithm to segment foetal cardiac data. A variety of flavours of snake algorithms is presented along my proposed version, each with its own merits and pitfalls. These algorithms are tested on synthetic data and on nine foetal heart datasets (seven 3D and two 2D). The synthetic data are used to test the behaviour of each snake algorithm under different levels and types of noise. The ground truth for the real data was assumed to be the manual segmentation by a foetal cardiologist expert and this was used to compare the different automated approaches. Validation of the algorithms using a physical phantom is then presented. Finally the results are discussed and proposals for future directions of research are outlined to increase the accuracy and reliability of the segmentation by the proposed algorithm.

1.4 The Foetal Heart

1.4.1 Introduction

The foetal heart has characteristics that are both functionally and structurally different to the adult case. Since the foetus obtains oxygenated blood from the maternal environment both atria share a common blood pool. In the foetal heart both ventricles are of equal importance and size until after birth where the left ventricle (LV) develops dominance over the right ventricle (RV) and the atria septal defect diminishes. Figure 1 shows a long axial 2D slice through the foetal heart so that all 4 cardiac compartments are visible in 4 chamber view (4CV). LA and RA refer to the left and right atrium respectively.

1.4.1.1 Foetal Heart Anatomy

The anatomy of the foetal heart showing each cardiac chamber is presented in Figure 1.

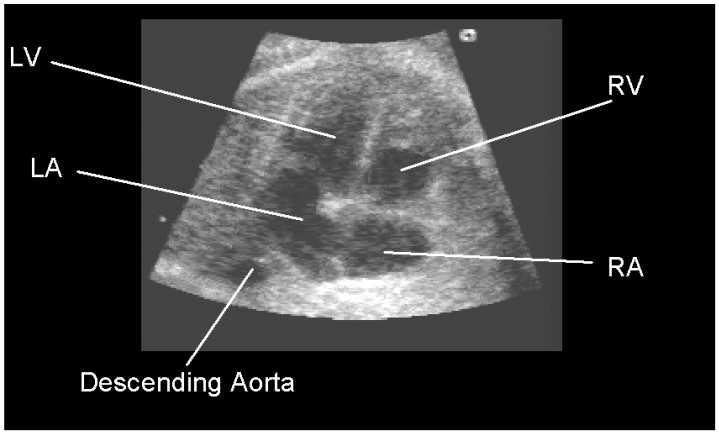


Figure 1 A slice showing the foetal heart in 4CV. The ventricles appear at the top of the image due to the orientation of the foetus in the uterus. The ultrasound probe would be at the top of this image in order to produce this view.

1.4.1.2 Cardiac Function Measurables

The table below outlines some of the indices that are useful in cardiology. Cardiac contractility refers to the ability of the cardiac fibre muscle to become shorter at a given fibre length in response to stimuli and is a performance measure of the cardiac muscle [5], [6].

Table 1 Table of Cardiac Function Measurables

Measurable	Computation	Significance
Heart Rate (bpm)	Measurement made directly from ECG	An indication of the ventricular fill rate which can be an indication of the health of the myocardium. The heart rate is around 140±20 bpm for a resting foetus, although small deviations outside this range are still considered normal.
Stroke Volume (ml)	(End Diastole Volume) - (End Systole Volume)	This index is a measure of the contractility of the heart muscle.
Ejection Fraction or EF	(Stroke Volume)/ (End Diastole Volume)	Another index of contractility of the heart muscle.
Cardiac Output (ml min^{-1})	(Stroke Volume) x (Heart Rate)	Another index of contractility of the

		myocardium but averaged over a period of time.
Cardiac Index (ml min^{-1} cm^{-2})	(Cardiac Output) / (Body Surface Area)	Relates heart performance to size of body. This index is fairly consistent over a wide range of individuals.

1.4.2 Foetal Heart Pathologies

Congenital heart disease occurs in 8 out of 1000 live births [7]. A variety of pathologies occur in foetal cardiology that are congenital in nature and these can be either structural or functional or both. Structural anomalies do not require time lapse images to be recorded for diagnosis. Functional pathologies however will be more apparent from viewing a time series of images of the foetal heart and knowledge of cardiac measurables is useful to quantify the anomalies. Volumetric cardiac indices require 3D data to be measured accurately as opposed to the use of geometrical primitives to estimate volumes from cross sectional slices [8]. Since the foetal LA and RA share the same blood pool functional anomalies may be less frequent in occurrence than structural ones.

The ventricular septal defect is a one of the most common types of congenital heart disease. This pathology is an abnormality in the ventricular septum that allows blood to pass between two ventricles. Small holes can close naturally but in most cases surgery becomes necessary. Early diagnosis can lead to better treatment planning. Since the myocardium is continuously maturing after surgery, follow-up observations will be required to assess the stability of the affected area.

This thesis focuses on the ability to measure the volumes of cardiac chambers accurately with little user interaction. There are various clinical reasons for obtaining these measurements; apart from establishing functional volumes of the foetal heart at various stages during pregnancy for research purposes, the cardiac indices may provide early signs of foetal anaemia. Immunohaemolysis is a condition where the maternal immune system attacks the red blood cells of the foetus so that there is insufficient oxygen supplied to the foetal tissues and organs. A result of this deficiency is foetal anaemia. To compensate for reduced blood oxygen concentration the foetus can develop some form of heart abnormality: thickened cardiac muscle (hypertrophy) and/or dilated cardiac chambers can develop which can lead to dysfunction of the valves, strain on the heart muscles by increased cardiac output, stroke volumes and heart rates [9]. Hydrops is another condition resulting from

anaemia where un-drained fluid can build up in the pericardial cavity. Treatments for foetal anaemia during the pregnancy can be performed by introduction of drugs to suppress the mother's antibodies from invading the foetal blood or a blood transfusion via the umbilical cord.

1.4.3 Imaging the Foetal Heart

Echocardiography is a common non-invasive imaging technique that is considered safe for clinical obstetrics [10]. Ultrasound is considered to be a very cost effective and versatile imaging technique with high temporal and moderate spatial resolution compared to other non-invasive modalities. The high temporal sampling possible with ultrasound is important since the foetal heart rate is around twice that of a normal adult. When compared to other modalities for structural imaging of the heart, the main drawbacks in echocardiography are the high speckle noise content (due to the simple ray tracing algorithms in the scanner's image reconstruction), and the low spatial resolution perpendicular to the beam (due to the dispersive nature of the wavefront as it spreads out from the transducer elements).

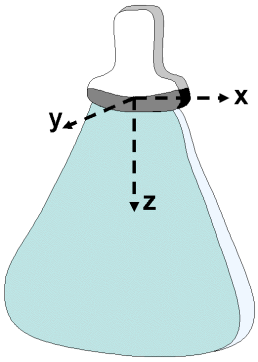


Figure 2 Ultrasound imaging geometry for 2D acquisition. The tissue imaged lies in the x-z plane.

Dynamic beam focussing has made it possible to obtain a relatively evenly sampled image in the x-z plane of the transducer array (refer to Figure 2) for imaging geometry. However the beam thickness in the y direction is considerably worse than in plane as with many volumetric image modalities acquired by stacking a series of slices. True volume acquisition is possible using a 2D array of transducer elements within the probe (matrix probes) to reduce the in-homogeneity of the spatial resolution perpendicular to the beam axis. Matrix probes usually have more transducer elements than those contained in linear arrays. The time taken to fire and receive US waves from all these elements usually results in reduced frame rate and noticeably lower spatial resolution than from the conventional 1D probes.

Echocardiography research is undergoing rapid development due to preferences to visualise the anatomy in three-dimensional (3D) for pathologies such as congenital

heart disease [11], [12]. Once the foetal heart is acquired in 3D it can be sliced in any plane to permit views that cannot be obtained on routine scans [13], or produce a "textbook view" to facilitate foetal examination [14], [15]. An added advantage of the 3D ultrasound dataset is that volume measurements can be made directly from the data by counting the number of voxels within the volume of interest and then multiplying by the volume of each voxel. It is also easier to consistently identify structures from a 3D representation of the data instead of the clinician mentally reconstructing the heart from 2D planar slices. In addition to foetal heart morphology from volumetric data, functional measurements can aid in diagnosis of pathologies that affect the motion of the heart.

Unlike adult echocardiography control of the foetus movement and breath holds is not reliable. Motion gating is required for a volumetric dataset that is acquired by sweeping the probe in the y direction. The foetal ECG is swamped by the maternal environment so that reliable gating by ECG signal is a challenging task and continues to be an active area of research [16]. The persistence of human vision requires a minimum frame rate of about 18-30Hz to perceive smooth video playback [3]. This frame rate is related to the length of time the image stays on the retainer and is adequate for imaging the foetal heartbeat, but if the clinician desires slow motion visualisation of the cardiac cycle to inspect various cardiac phases then a higher temporal sampling rate is needed to acquire the images. Since the foetal heart beats around 120-170bpm good sampling in the time domain is essential to capture all of the cardiac phases of the heartbeat. Assuming that there are 8 phases that are equally spaced within the heartbeat of 170bpm this would require a minimum of 16 volumes to be acquired per cardiac cycle (Nyquist criterion) at a rate of 45Hz. The existing Acuson scanner available to the department is able to acquire slices at 25Hz non-interlaced or 50Hz interlaced so provides a trade off between structural and temporal sampling. The group have access to a newer Phillips Live 3D scanner which can sample at a rate of about 24 volumes per second.

In adult echocardiography the probe can be orientated to capture a long axial view of the left ventricle. A breath hold can be asked from the adult to reduce motion artefacts whilst the acquisition is made. In the prenatal heart this is more difficult since the foetus is moving and can not be asked to do a breath hold. The foetus starts practice breathing in the third trimester for neonatal life; it can be infrequent and irregular. Fortunately the third trimester occurs after about 23 weeks which is the cut off period to scan the foetal heart without heavy shadowing artefacts from the calcification in the foetal ribs. After 27 weeks the amniotic fluid volume decreases relatively and the foetal heart moves further away from the transducer [17]. In addition to these unfavourable imaging conditions unpredictable body

movements of the foetus result in heavy motion artefacts. To scan the prenatal heart it is necessary to scan when the foetus is not restless. This means that axial orientated acquisitions of the prenatal heart are not a priority. To produce the 4CVs that are shown in this thesis the datasets have to be resliced offline in software after acquisition.

A representative example of 2D images in foetal echocardiography from the Acuson scanner is shown in Figure 3. It is clear that there is a narrow range of intensity values that dominate the blood cavity, unfortunately shadowed regions also have similar grey level values to the cardiac chambers. The partial volume effect that is most noticeable occurs mainly at the apices of the ventricles where the papillary muscles and valves have similar intensity ranges with the cavity and myocardium. The myocardium has some overlap with these regions and covers a large range. It is apparent that the histogram is not smooth inside the various regions but comprises of significant speckle noise. It appears that simple thresholding is not adequate to reliably segment the endocardium. Extensive shadowing can occur as a result of calcification of the foetal ribs after about 23 weeks of gestation. Datasets with this phenomenon can cause occlusion to parts of the heart rendering the intensity values in this region to similar levels as the cardiac chambers. The image histograms in Figure 3 show a differential non-linearity from the analogue to digital conversion process. This is evident from the sudden drops in the population that occur at regular intervals. There are zero counts at even greyscale values therefore the map of echoes is not sampled at 8 bit resolution. Regular drops can also be seen occurring at greyscale levels of 27, 43, 59 i.e. at an interval of 16. This refers to the 5^{th} bit of the 8 bit word but unlike the even bins the count is not zero. On either side of the drops there is not a large difference in the number of counts so the histogram does not show any indication of systematic errors of either rounding up or down.

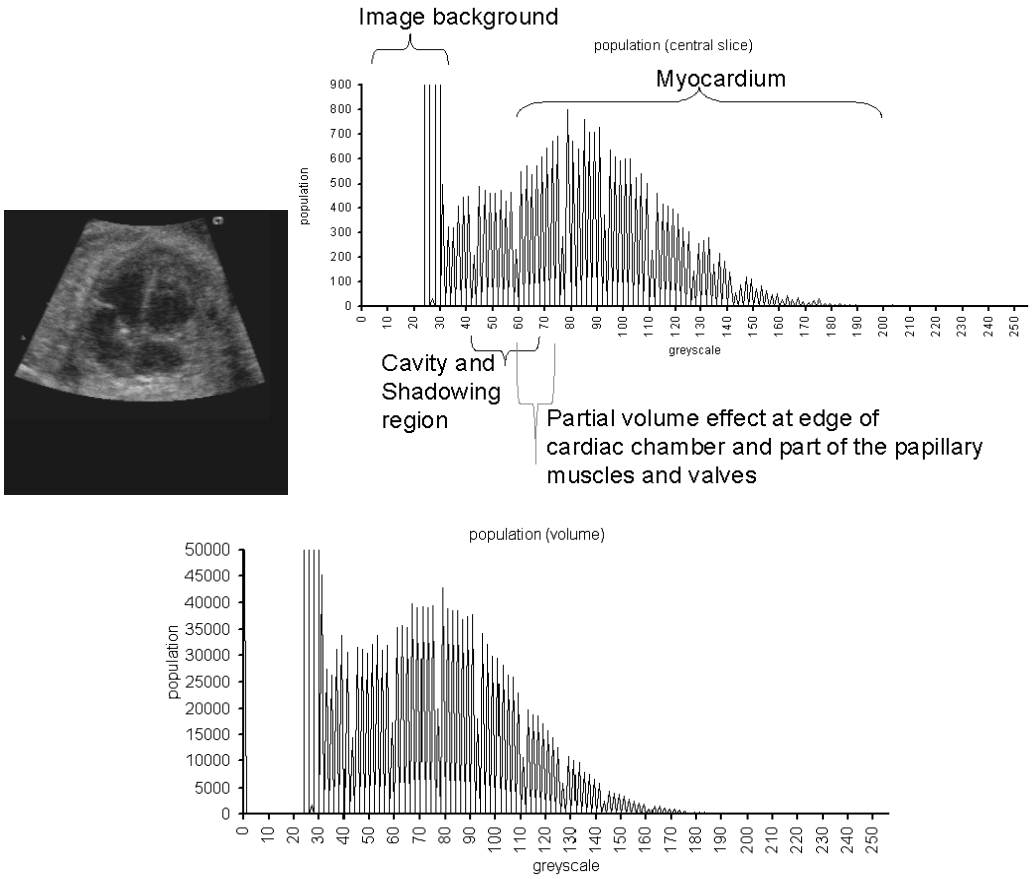


Figure 3 (*top left*) Typical 4-chamber foetal cardiac image (long axis view) and the corresponding histogram for the slice shown (*top right*) and for 3D volume of this frame (*bottom*). This histogram shows differential non-linearity from the analogue to digital conversion process. Note the histogram will change slightly depending on the dataset, scanner and transducer used during the acquisition, TGC parameters, slice through the heart and the amount of shadowing/enhancement effects present in the field of view. The histograms have been clipped at low populations to emphasise the cardiac structures.

Figure 4 shows an example histogram from a volume acquired by the Phillips Live 3D scanner. There is less overlap in the grey levels between cavity and papillary muscles which may be due to the low speckle content in the images. There is no sign of differential non linearity from the Live 3D image histogram as in the data obtained from the Acuson scanner.

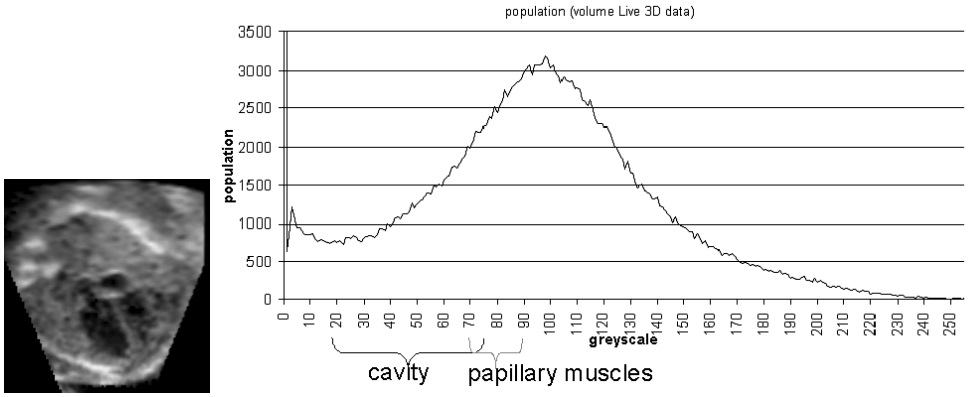


Figure 4 Central image slice from Live 3D dataset (*left*) and corresponding histogram of all the slices in the first frame from the dataset (*right*). The histogram was clipped to emphasise the cardiac structures.

A selection of 5 easiest to segment manually datasets was chosen from the sample. From these the value of the mean was estimated for both inside and outside the chambers.

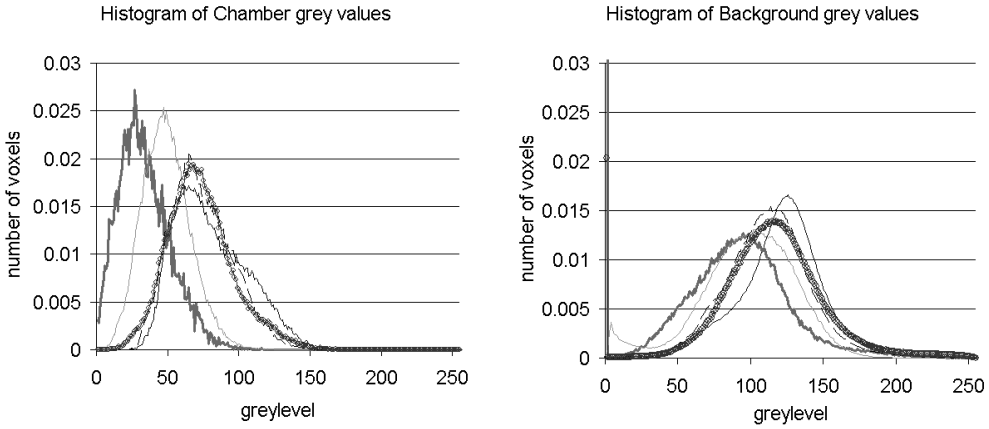


Figure 5 Intensity distributions of voxels in a sample of 4 Live 3D datasets (normalised). The corresponding means of the 5 datasets are 40 and 91 for the interior and exterior of the chambers respectively. The means are rounded to the nearest whole integer.

2 THEORY

2.1 Ultrasound Imaging

2.1.1 Introduction

Diagnostic ultrasound is becoming more widespread in all medical fields. It offers a low cost, relatively high speed and resolution for cardiovascular imaging compared with other common modalities like Computed Tommography (CT) and Magnetic Resonance Imaging (MRI). Ultrasonic image formation has improved significantly over the past 10 years with advances in digital signal processing, transducer developments, digital beam focussing, higher bandwidths, a range of scanning frequencies, high dynamic resolutions of ~120dB, multi-focal zones and contrast agents [18].

Acquisition of foetal heart datasets in three and four dimensions is a challenging research area in echocardiography. There are problems of the high speckle content (a random interference pattern in an image formed with scattering of coherent wavefronts by a medium containing many sub-resolution scatterers) in three-dimensional adult echocardiograms. In addition to this the corresponding foetal datasets suffer from significant motion artefacts (due to rapid heartbeat and unpredictable foetal and mother movement and breathing). Signal dropout due to shadowing from calcified foetal ribs is another major artefact that can cause partial or total occlusion of myocardiac structures [16]. Since the foetal electrocardiogram is difficult to obtain accurately due to interference from the maternal ECG [14], alternate methods of gating the acquisition have to be used [15], [19], [20]. These gating protocols may take place online which greatly reduces the post processing time [20] or offline [19], [21].

2.1.2 Ultrasound Principles

The following is a brief overview of ultrasound physical principles, further details can be found in [22]. Ultrasonic waves are well above the human audible range (20kHz) but medical imaging applications make use of ultrasound in the MHz region to achieve narrow beams and short pulses to accurately locate structures within the body. The depth of the imaged structure and the attenuation properties of the tissues along the path of the beam determine the choice of frequency used in the

clinical examination. Muscle is a stronger attenuator of ultrasound than fat and attenuation increases with frequency; therefore a compromise is made between the spatial resolution (high frequency) and penetration (low frequency) of the beam. Thus higher frequencies tend to be used for structures closer to the skin and lower frequencies for deeper structures. The speed of ultrasound waves is media dependent and in the human body propagation occurs at around 1500 metres per second (ms^{-1}), although this will vary according to the exact tissues in the path of the beam. Ultrasonic waves are pressure waves and hence cause disturbance within the medium. The particles oscillate with amplitude that diminishes with time due to energy conversion into heat via molecular vibration within the anatomical structures. This process is called absorption and can be modelled to a first approximation for thin beams passing through homogeneous media by an exponential decay relationship as in equation (2.1) where $I(x)$ is the intensity along the beam at position x, μ is the attenuation coefficient in nepers as given by $\mu x = -\ln\left(I\left(x\right)/I_0\right)$. μ is measured in decibels per centimetre ($dBcm^{-1}$) and I_0 is the intensity at the air skin interface.

$$I(x) = I_0 e^{-\mu x} \tag{2.1}$$

Attenuation is frequency dependent with stronger attenuation occurring at higher frequencies than for lower ones. Attenuation limits the working frequency range for scanning the heart and abdomen at 2-5 MHz. When μ is expressed in $dBcm^{-1}$ it is then found to increase linearly with frequency for many tissues and can be taken as a value of 0.7 on average. For muscle this value is around 1.5; blood has value of 0.2 and bone gives a much higher value of 10.0 [23].

The ultrasound disturbance moves at a fixed velocity in a given medium where the medium is characterised by its elastic modulus K measured in kg m^{-1} s^{-2} as given by:

$$K = \frac{stress}{strain}$$
$$c = \sqrt{\frac{K}{\rho}} \tag{2.2}$$

Where c is the speed of ultrasound in the medium and ρ is the density of the medium in kgm^{-3}. From this equation the wave propagation is higher for denser materials such as bone when compared to soft tissue structures such as muscle.

An ultrasound impulse fired from the probe at the skin surface will pass through several layers of tissue and multiple anatomical structures. Acoustic reflectors are interfaces between two acoustic media and the acoustic dissimilarity properties of the

adjacent media will determine the strength and phase shift of the reflected wave. Scattering interfaces occur at changes in density or compressibility of the medium. By the conservation of energy principle the reflected and transmitted wave energy must equal the energy contained within the incident wave stream. In specular reflection only a single reflected ray is created upon interaction of the incident wave with the interface and the amount of intensity reflected depends on the reflection (*R*) and transmission (*T*) coefficients at the interface:

$$\begin{aligned} R &= \frac{I_r}{I_i} = \frac{(Z_2 - Z_1)^2}{(Z_1 + Z_2)^2} \\ T &= \frac{I_t}{I_i} = \frac{4Z_1 Z_2}{(Z_1 + Z_2)^2} \\ T + R &= 1 \end{aligned} \tag{2.3}$$

Where *Z* is the acoustic impedance of the medium (measured in kg m^{-2} s^{-1}) given by:

$$Z = \rho c = \sqrt{\rho K} \tag{2.4}$$

Snell's law of reflection states that the reflected wave leaves the interface at the same angle to the normal of the interface as the incident wave for structures larger than the wavelength of the incident beam. However, waves incident on a rough surface or on particles (refer to Figure 6) that have dimensions comparable to the wavelength of the ray, scattering of the beam occurs and the reflection becomes diffuse and much more difficult to locate a single point on the structure accurately from the reflected rays received by the probe. The amount of scattering is dependent on frequency by a non-linear relationship.

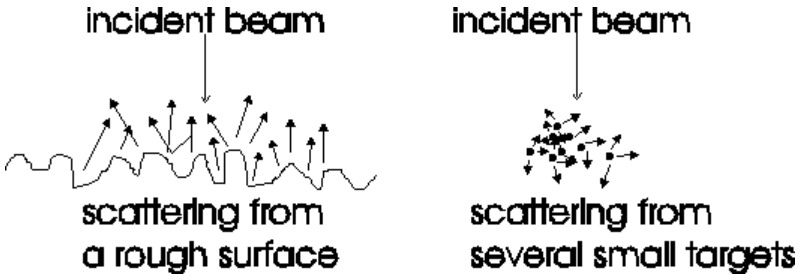


Figure 6 Two examples of scattering (a) by a rough surface and (b) by several small particles such as blood cells.

Each point on the US wavefront will give rise to both reflected and transmitted sources at acoustic interfaces. The sub-resolution scatterers can cause multiple reflected wavelets that travel in various directions and paths to reach the probe head. This interference pattern of wavetrains leads to speckle in the image. It is non-trivial to decorrelate speckle noise from the data since it is produced by the underlying structures that are imaged. An example of modelling diffuse scattering from many scatterers in US images is governed by the Rayleigh distribution [24]. The Rayleigh distribution is illustrated in Figure 7 and is stated in equation (2.5).

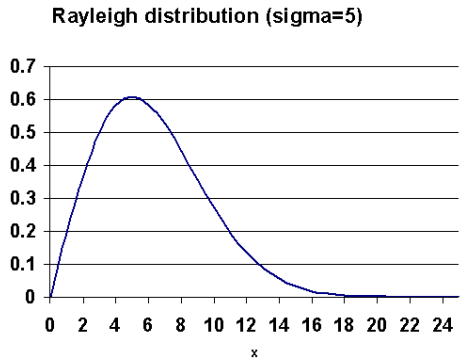


Figure 7 Rayleigh Distribution in 1D for positive values of x.

$$p(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{2.5}$$

Particles with dimensions much smaller than the wavelength of the US wave are known as Rayleigh scatterers and their intensity distribution follows a quartic law [25]. The intensity dropoff as a result of Rayleigh scattering aids the myocardial imaging process. A probe operating at 1-5MHz images blood at a level 20 decibels below the signal for tissue enabling good blood-tissue contrast. At higher ranges 20-30MHz the blood cells become Rayleigh scatterers and their echo signal is as bright as for tissue. The Rayleigh distribution assumes isotropic scatterers. During the cardiac cycle the US wavefront will be scattered in a time varying anisotropic manner due to contracting fibre orientations in the myocardium.

The scanner measures the two way travel time of the wave (incident and echo) to determine the depth of the reflector. Due to increasing attenuation with depth the echoes from deeper reflectors are weaker, and so Time Gain Compensation (TGC) is applied on the scanner to boost the strength of echoes that arrive later so that the signal appears uniform in brightness along the axis of the scanning beam. Usually the TGC algorithm is manufacturer specific and can be difficult to remove in order to recover the scanner's unprocessed acqusition. There are several reasons for ultrasound attenuation in the body; these fall into the following categories, absorption, scattering, beam divergence, reflection and refraction. The first two cases are tissue specific and each media will have its own coefficient of absorption and scattering. Geometric attenuation occurs due to beam divergence since this reduces the intensity of the beam (and its echoes) due to an increase in cross sectional area with beam propagation. In the case of planar wavefronts this effect is lessened and principal attenuation factors are absorption and scattering. The amount of energy attenuated by reflection is dependent on the nature of the interface between two media. If the transmission coefficient is small compared to the reflection coefficient then structures beyond this interface appear dark since very

little energy reaches them. This effect is known as shadowing and some of its effects on the foetal heart can be seen in section 2.1.2.5.

Harmonic imaging uses the higher frequency multiples of the fundamental echogenic signal than just the fundamental. Tissue harmonics are generated as the beam passes through tissue. Since the speed of the wave depends on the density of the medium the compression part of the US wave increases the tissue density fractionally and the rarefractions decreases it slightly. Therefore the density changes due to the beam cause part of the wave to speed up or slow down which causes distortion to the waveform and its angular components represent the harmonics [26]. Conventional US uses only the fundamental and discounts the remaining weaker components as noise whereas these harmonics can offer better contrast and spatial resolution to reconstruct the image especially in areas of focalisation where the acoustic energy is high enough to allow harmonics to be generated. Imaging the second harmonic instead of just the first one reduces wavefront aberration, attenuation on the forward path, causes narrowing of the beam and suppresses sidelobes [27] since they are produced by the centre portion of the main beam. However, the harmonics are 20dB less than the fundamental [26] and this can reduce the signal to noise ratio of the images. Therefore harmonic imaging is better suited for scanners with high dynamic ranges. At the moment only second harmonic tissue imaging is commercially available but this may change in the foreseeable future.

2.1.2.1 Types of probe to acquire structural information

There are five main modes of ultrasound scanning (M, A, Doppler, B, and real time) and a variety of probe designs. Although there are five, each of the first three modes could be considered real time since they can produce updates at video frame rates. B mode is a compounded image resulting from manual changes to the angle of the probe during acquisition and so is not generated in real time. Each scanning mode or a combination of modes can be implemented into a single scanner design. TGC can be incorporated for each mode so that echoes from deeper reflectors appear with similar amplitude to those from the shallowest reflector. The simplest ultrasonic instrument uses A-mode which fires and receives ultrasound pulses down a stationary scan line (acquired parallel to the probe long axis) so that a graph of amplitude against time can be plotted of the retuning echoes. This is achievable with single or multiple element transducers. To change the angle of the scan line the probe has to be physically moved by the operator. B-scans take this principle further and automatically oscillate the direction of the beam after receiving the echoes for the current angle so that a fan of lines is swept out (see Figure 8 which has been reproduced from [23]).

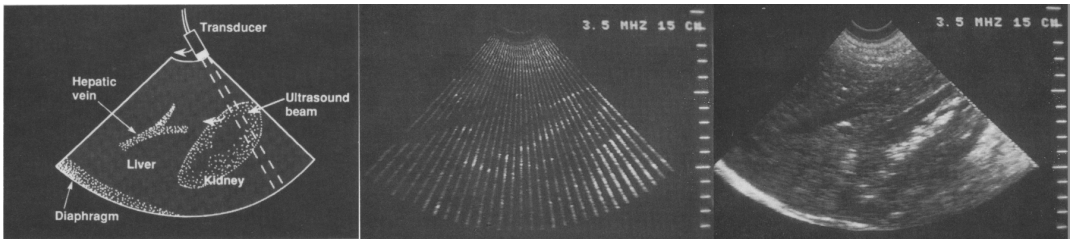
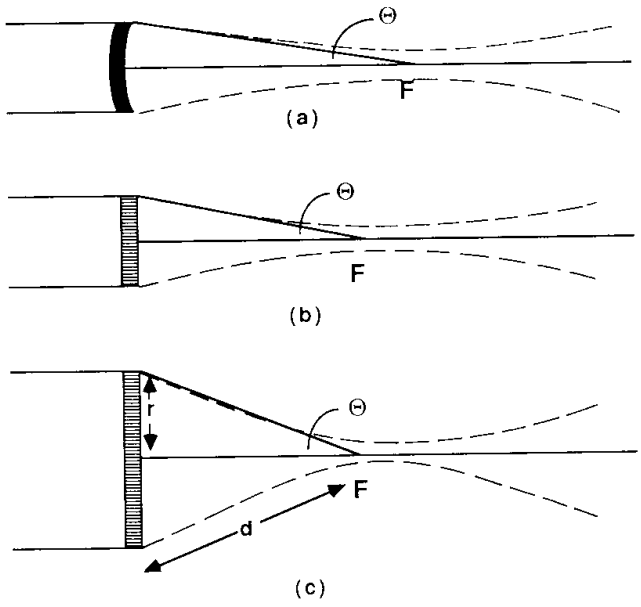


Figure 8 The sweep of an ultrasound beam through tissues (*left*) Sample lines of echo information (*middle*), Ultrasound B-image created from sweeping of many scan lines (*right*). (Taken from [23]).

Since the spatial extents are related to the time and velocity of the pulse through the body, a spatial map of the reflectors can be built up with every sweep of the sampling beam. To perform this scan the transducers can be physically moved either by the operator (by oscillating the entire probe) or by electromechanical action via an electric motor. The third way of sweeping an A-line is to use a phased array of transducers within the probe so that a beam is generated and steered by incremental delays of several neighbouring elements. This method of beam steering is based on Huygen's superposition principle of wavefront where each composite wavefront acts as a point source for secondary spherical wavefronts. The same principle is used in beam focussing.

Phased array probes are the preferred choice of modern transducers since beam focussing is possible at any depth by increasing the number of active elements in the probe for deeper structures (see Figure 9 which has been reproduced from [23]). The numerical aperture is a measure of the focussing of the beam and this can be maintained at various depths through the relation given in Figure 9.



$$\text{Numerical aperture} = \sin\Theta = \frac{r}{d}$$

Figure 9 Definition of Numerical aperture. (Taken from [23]).

Mechanically driven transducers often require sophisticated electronics and shielding methods to reduce the electromagnetic interference from the motors. However, mechanically driven probes are still used today in cases where the multi-element phased arrays are too large for example in inter-caustic scanning. In this application mechanical probes are used to rotate the imaging plane about the beam axis to scan the heart through the ribs.

2.1.2.2 Volumetric scans

Several probe types can be used to acquire volumetric data. The most common varieties of phase array transducers are shown in Figure 10. For 1D arrays the volume is acquired by sweeping the probe perpendicular to the imaging plane. Tracking of the probe becomes an issue as soon as it is moved to maintain registration of the image slices but this is often handled by the scanner or its built in reconstruction software. Motion artefacts that arise become very prominent and can usually be compensated for by ECG gating. However in the foetal heart case these triggering ECG signals are often difficult to obtain reliably due to the interference from the maternal environment.

Phased array probes are convenient for imaging through narrow gaps (compared to linear arrays) because of their small transducer heads but this feature also limits their field of view in acquisition to a range of 45°-70°. Mechanical oscillation

transducers that oscillate in the image plane are generally better at capturing data from broader angles and can sweep up to a 90° sector. Linear arrays have the widest field of view and can easily obtain 180°. In addition to these in-plane viewing angles, the probe can be moved either mechanically or by freehand perpendicular to the slice plane such that it sweeps out a volume of 2D slices.

Freehand scanning requires a steady movement of the probe in order to minimise the gaps between slices (refer to Figure 10). In the case of the 2D phased array, no mechanical movement of the probe is required and so many of the problems with the sweeping volume acquisition are reduced. However, the drawback of the volumetric probe is its more bulky design (since it holds more transducer elements than the 1D arrays), its smaller field of view and reduced temporal and spatial resolution when operating in volumetric mode compared to its conventional planar B-mode.

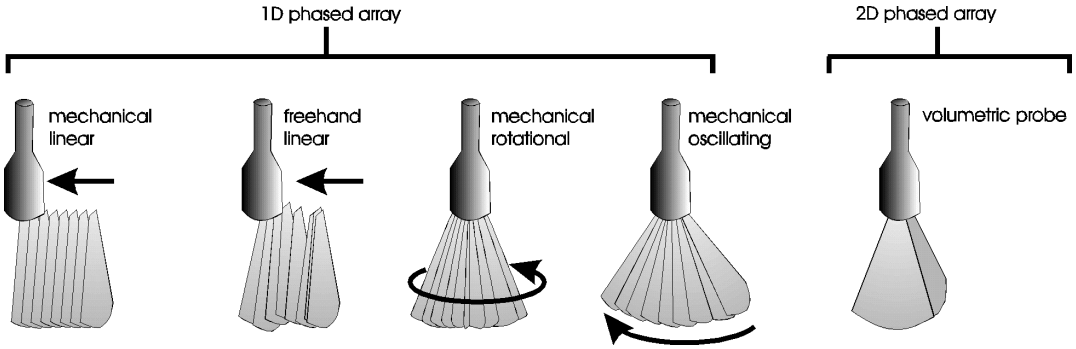


Figure 10 Methods of acquiring volumetric ultrasound data.

2.1.2.3 Spatio Temporal Image Correlation (STIC)

STIC is a component of the basic 3D/4D software built into the GE Voluson sonography systems (GE Healthcare Technologies, Waukesha, WI). This method is a way of compounding successive images that are repetitions in order to build a target with less noise than any individual image. The correlation is a matching process that can be applied to sort through the successive images in order to find the repetition rate regardless of orientation of the heart. STIC allows the acquisition of a 4D cardiac sequence and presents it as a cine sequence. Acquisition is performed with an automated slow sweep that collects sequential frames at a rapid rate. The longer the acquisition time, the better the reconstructed volume becomes. Since the foetal heart beats during acquisition the cardiac diameter changes periodically and this time period can be detected by STIC and used to gate the images [28] (see Figure 11). The method assumes that the foetal heart rate does not change during the acquisition and that there is no foetal or maternal body movement. Some voxel averaging takes place during STIC and this may blur certain structures such as septal defects.

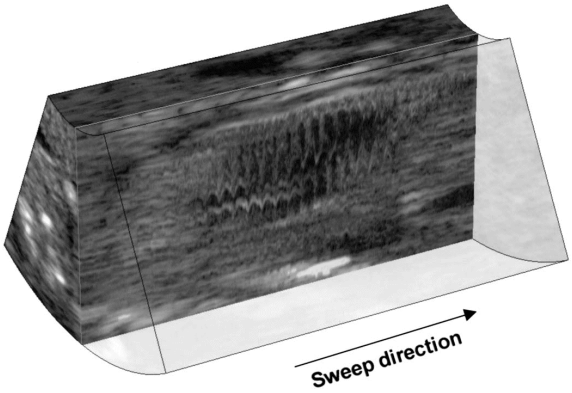


Figure 11 Raw data image used to calculate the foetal heart rate. This image is generated after the STIC acquisition is performed as a single, slow sweep. Information from this raw data is used to rearrange the two-dimensional frames. This particular image is orthogonal to the original two-dimensional frame. Because of the long acquisition time (7.5 to 15 seconds from the left to the right end of this image), the beating heart draws a motion pattern. This pattern is analyzed in terms of periodical changes of greyscale information, and the foetal heart rate is calculated. Beat-to-beat changes of the heart rate would appear as shortening or elongation of the above motion pattern. This image is not visible on the system during STIC acquisition, but helps to understand the technique. (Taken from [28]).

2.1.2.4 Physical Limiting Factors in the Acquisition

The velocity of sound waves in dense material (such as bony structures) is much higher than in soft tissue (3500 ms^{-1} compared with 1540 ms^{-1}). The main limiting factor in the rate of B-mode ultrasound imaging is the pulse repetition frequency due to the finite propagation time of the pulse and echo in the body. For a depth of 15cm, the two-way travel time in soft tissue is about 0.2 milliseconds for the scanline of a single transducer element. This corresponds to a frequency of ~5000Hz. However to allow the complete image to be built up from 128 elements restricts the pulse repetition frequency to about ~40Hz. Higher frame rates can be applied to image shallow located objects than for deeper structures. Typical acquisition rates of ~25 images per second are achievable using modern scanners although interlacing can raise this to 50 images per second. Around ~25 volumes per second is achievable with the Phillips Live 3D scanner which incorporates a 2D phased array transducer. Deng et al [29] comments that this volume acquisition rate is sufficient for most foetal wall movement imaging if detailed information about the valve motion is not required and the spatial resolution is not compromised. The frame rate of the volumetric probe is less than that of the 1D probe because the extra dimension increases the number of transducer elements to fire wavefronts and hence more echoes to receive. Usually probes with 2D arrays are used for volume acquisition and so have fewer elements arranged in a straight line to acquire single slice images than for a conventional 1D probe hence matrix transducers have lower spatial resolutions than 1D probes.

The near field (Fresnel zone) of the ultrasound beam is extremely non-linear due to the side lobes but fortunately these effects start to diminish for structures located deep within the maternal body such as the foetal heart (10-15cm). An illustration of side lobes is shown in Figure 12. The near field boundary (NFB) can be approximated (for high frequencies) by

$$NFB = \frac{D^2}{4\lambda} \tag{2.6}$$

where D is the diameter of the transducer and λ is the wavelength. In multi element transducers D is the diameter of the set of active elements in the array. Beyond this boundary the wavefronts are more planar and the pressure falls off regularly with distance. The average speed of sound (v) and the relation between the centre operating frequency (f) and wavelength (λ) of the transducer is given by

$$v = f\lambda \tag{2.7}$$

Assuming that v =1540ms^{-1} in soft tissue then for a 3MHz 13mm unfocussed transducer the NFB will be approximately 8.5cm; a 5MHz version will be approximately planar beyond 14cm [30]. US beam focussing will reduce these values of the NFB slightly but will still be close to the depth of the foetal heart (~100mm) and the strength of the effect influencing the acquired images will depend on the size of the patient.

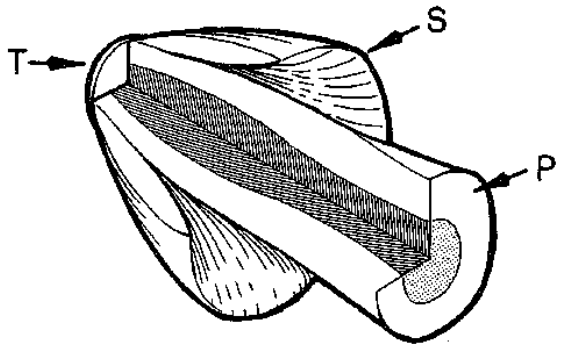


Figure 12 Illustration showing the non-uniformity of the ultrasound beam [30]. The diffraction process of the transducer *T* produces a primary beam *P* and side lobes *S* that extend outward at an angle. Most of the energy is localised in the primary beam centre.

Side lobes are unwanted energy emission that arise from the sinc^2 response of a rectangular aperture and diverge in a direction away from the main beam (see Figure 12). In multi-element transducers each element generates its own side lobes. The resulting unwanted lobes are called grating lobes. The side lobes contribute to the partial volume effect even though most of the energy is contained within the primary

lobe. Since the main lobe carries most of the energy in the total beam emission both side and grating lobes are not normally apparent from in US images unless they fall upon a particular echogenic region. For reflectors normal to the beam on axis sensitivity is enhanced even in the presence of side lobes as shown in Figure 13.

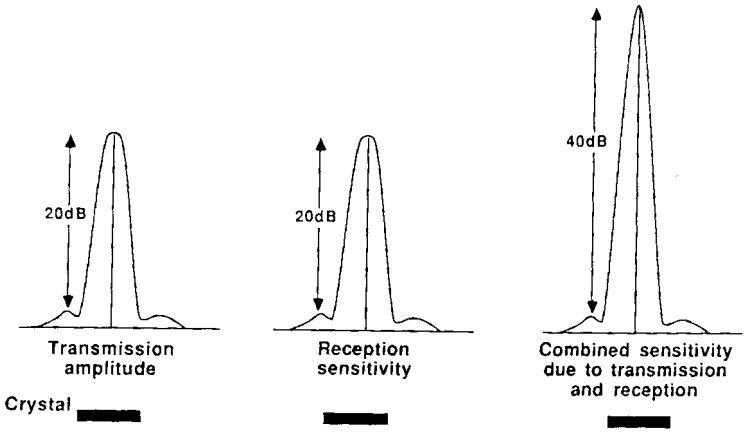


Figure 13 Side lobes in a transmission field and a reception zone [23].

To reduce side lobe emission the transducer crystals are designed with a width of at least ten times the ultrasound wavelength [31]. Apodization is the weighting of the contributions from different transducer elements before combining them to form an ultrasound field or reception zone. During beam formation adjusting the apodization can be used to modify the beam shape and size of the side lobes. Usually this is accomplished by providing more power to the transducer elements nearer to the centre of the probe.

Dispersion of the wave through the media allows many acoustic paths to any given point in the body. As a result the travel time to and from a point will depend on the acoustic path taken; this is not necessarily an Euclidean distance since passage through denser structures may be quicker than sound transmitted through a shorter Euclidean path of acoustically slower media. Wave dispersion causes the beam width to diverge as it propagates through the body and this limits the off-axis spatial resolution. Using phased arrays and non-planar transducers allows dynamic beam focussing during the acquisition for a given angle so that echoes at various depths are imaged with comparable lateral resolution. Lateral in this case refers to the direction perpendicular to the beam axis. The axial spatial resolution of the beam is often better than in the lateral case. Axial resolution is defined here as the separation distance between two reflectors that are just distinguished by the scanner along the beam whereas lateral resolution measures the ability to resolve structures off axis. Plane wavefronts undergo diffraction on reflection from a non-planar object (such as an organ). If the surface is rough with features smaller in length than the wavelength of the beam this information cannot be resolved. From the Nyquist

criterion the beam width should be less than half the size of the feature that it is sampling in order to resolve it correctly.

The frequency of the ultrasound wave affects the spatial sampling along the beam axis; higher frequency allows the ability to resolve finer structures. It is common for the scanner output to provide a uniformly sampled image of the raw data to reduce the emphasis of under sampling of the deeper structures compared to the shallower ones. The diagnostic ultrasound imaging window for the human body is between 2-50MHz. Eye imaging is possible using 50MHz transducers to achieve 100μm resolution to look at the eye's blood flow over the cardiac cycle [32]. Ultrasound microscopy uses frequencies of 100MHz to resolve at 0.1μm for cellular structures such as nerves. Ultrasound microscopy can look at the cellular level of tissue and be used to examine blockages in arteries using intravascular ultrasound [33]. The energy contained within the sound wave is more attenuated at a given depth with higher frequencies and so in practice higher frequencies (>5MHz) may be used to image structures closer to the surface of the body such as eyeballs, and lower frequencies (2-5MHz) are more useful for deeper structures such as liver, heart, and foetus. Higher frequencies can allow finer spatial resolution along the beam axis of structures near the surface of the body. However, the probe frequency chosen can depend on the size of the patient since the heart may be closer to the chest in a smaller body compared to within a larger person. Thus there is a trade-off between imaging depth and resolution since higher frequencies can only be used for shallow organs and lower frequencies for deeper ones. Generally the frequency used for the acquisition will be determined by examining the clarity of the structures within the ultrasound images during a rough pre-scan of the patient.

Contrast resolution is a measure of the ability of the scanner to display different tissue types as different grey levels. The contrast is often corrupted by the partial volume effect which produces speckle so that it is difficult to distinguish between small changes in grey levels.

2.1.2.5 Dominant Image Artefacts

The ultrasound instrument assumes that both the incident and reflected waves travel in straight lines along narrow beams and at constant speed. The scanner ignores refraction effects of the beam and multiple reflections of the wave front between acoustic boundaries are assumed not to take place. However, if situations arise which deviate from these assumptions then areas of the images reconstructed by the scanner will become distorted or incorrectly located. The following presents a summary of the effects of imaging structures with real ultrasound scanners [22].

- A reflection occurs when there is an impedance mismatch at the interface between two media. When a wavefront passes through two interfaces especially if the boundaries are highly reflective, multiple reflections or reverberations can arise. The transducer receives multiple echoes (reverberations); each successive echo will be at lower amplitude due to energy attenuation within the tissue and at each reflection (refer to Figure 14a). The echoes will arrive at integer multiples of the travel time to the reflector. This reverberation effect is very noticeable on wave propagation through bowel gas in the colon or through the bladder. Reverberation can occur if there is airspace between the probe head and the skin of the patient. To reduce this effect impedance matching gel is first placed on the area of contact of the transducer before scanning the patient. In the case of an off-axis reflector, the echo time is not simply twice the single travel time from transducer to this reflector as illustrated in Figure 14b.

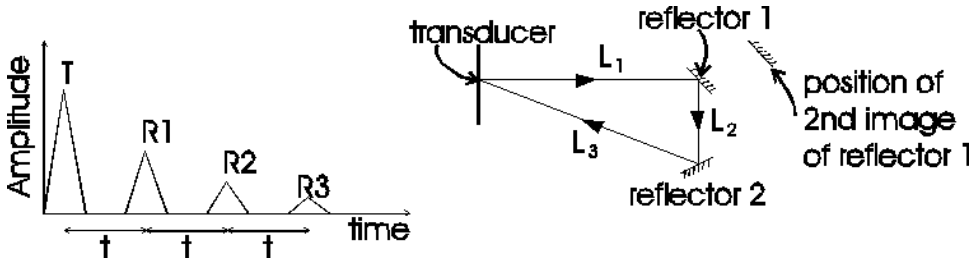


Figure 14 (a) Reverberations arriving at the receiver transducers on an A-mode display. T is the time taken for the first echo to the reflector and R1, R2, R3 the time taken for the first, second and third reverberation respectively. (b) Multiple reflections caused by an off-axis reflector. The depth of the second image is given by $(L_1+L_2+L_3)/2$.

- Echo amplitudes can be affected by the following sources [23]:
  - Differences in acoustic impedance changes at reflecting and scattering targets
  - Size and shape of reflectors and scatterers
  - Spatial variations in intensity within the ultrasound beam
  - Interference that produces a speckle pattern
  - Inaccurate TGC correction

- Since the energy in the ultrasound wave dissipates as it passes through subsequent structures and the intensity drops with increasing distance, TGC is needed to correct for the echo amplitudes from deeper reflectors. Although the automatic TGC algorithm is manufacturer specific it corrects for direct reflection from the acoustic interface, but often neglects refraction mechanisms as a means for wave propagation both from and to the transducer elements. The brightness of deeper reflectors along the beam axis can be boosted manually to override the TGC by the user and this introduces

operator dependence into the image acquisition. Manual TGC parameters are not recorded in the acquisition.

- Shadowing is when ultrasound is strongly reflected from dense structures such as an air/bone interface and so little energy will remain for imaging the layers below. Since the transmitted wave will be greatly attenuated, scatterers behind this structure that causes the shadowing will appear very low in amplitude and hence dark in the image. The amount of energy reflected from bone/muscle interfaces can be as high as 40% and 0.07% at muscle/blood boundaries [23]. Shadowing can occur from structures that contain airspace such as the lungs and bony material like as ribs. Figure 15 shows an example of shadowing that causes dropout of structures in the foetal heart. To avoid this artefact the operator must angle the probe such that the bony objects do not obscure the path of the beam.

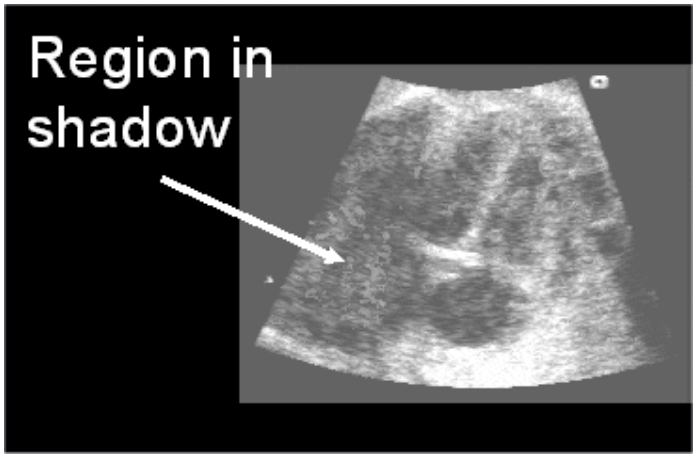


Figure 15 Shadowing and its impact on the foetal heart. Part of the left ventricle and atrium are occluded in this long axis slice. The orientation of the foetus is such that the ventricles appear above the atria in this image.

- Ultrasound datasets are notorious for their speckle content compared to other imaging modalities. The speckle size is not fixed since it is strongly affected by resolution. The spreading of the beam increases with depth and this limits the resolution. Therefore the speckle tends to take on a banana shaped appearance whose length grows with increasing depth (see Figure 16 which has been reproduced from [34]). Figure 17 shows an example of the banana shaped speckle appearance in the foetal heart dataset where the apparent dark region contains elongated curved structures that lie horizontally in this scanning geometry. To some extent dynamic beam focusing compensates for this to provide a more homogeneous distribution of speckle sizes. However the lateral resolution is worse than in the axial direction so the speckle noise appear as elongated structures (refer to Figure 17). Lateral speckle size is

highly dependent on the distance to the transducer and not a direct image of any structure. To get an idea of the average size of speckle in a homogeneous region the autocovariance function can be used

$$ACVF(x,z) = \iint_{ROI} \left\{ I(\theta+\theta', r+r') - \bar{I} \right\} \left\{ I(\theta, r) - \bar{I} \right\} d\theta' dr' \tag{2.8}$$

where *ROI* is the region of interest, I is the image intensity, $\bar{I}$ is the intensity within the ROI, θ and r are the lateral and axial coordinates respectively [34].

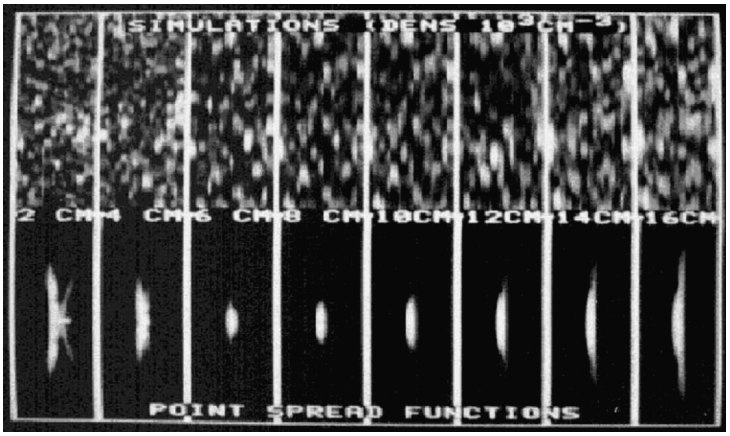


Figure 16 *Top row*: Echographic speckle images from homogeneously scattering medium. Depth increases from zero (transducer) to 16cm, from left to right. *Bottom row*: PSF obtained by scanning a point reflector at depths from 1 to 15cm.

Figure 17 Speckle orientation in an apparently homogeneous medium. *Left* shows image and *right* shows the blow up of the rectangular region.

- The US beam naturally diverges with increased depth however; beam focussing by phased array can narrow its spread at the current depth that is being imaged. Since the beam has a finite width structures that are within the beam but off-axis are integrated with those located along the scan line. This results in the image of the structures appearing elongated perpendicular to the beam scan lines by the width of the ultrasound beam (see Figure 17). The beam divergence is not confined to within the imaging plane and so the sampling of the image slice will have a finite width (Figure 18). Off axial reflectors within the scanning beam will be imaged as if they are part of the reflectors lying along the scan plane producing the speckle phenomenon. The

artefact shows up as a reduction in image contrast because of the integration of lower intensity echoes from points not located on the scan plane. Other effects from this noise phenomenon are contribution to the partial volume effect, grainy appearance to image, corrupted boundary information and may cause two closely located structures to be visualised as one.

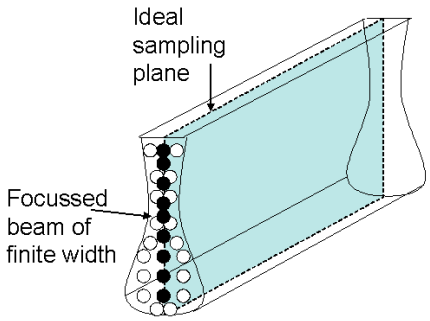


Figure 18 Illustration showing finite beam width for an image slice with ideal sampling (black) and off plane sampling (white) that contributes to the partial volume effect.

- Signal dropout can occur in the absence of shadowing if any of the dimensions of the structure are not resolvable by the beam. Due to the inhomogeneous sampling nature of the beam structures perpendicular to the axis are more prone to the partial volume effect than structures parallel to the beam. In the heart atrial septal walls particularly in the foetal case may be sufficiently narrow that they appear absent in the acquired image. This can cause leakage between chambers for segmentation algorithms that rely on region growing from initial seed points.

- 3D volume reconstruction can be misleading if a moving object is imaged by acquisition on a slice by slice basis. If the frequency and path of the moving object is known the imaging plane can swept so that the slices are acquired at the right place and time to yield a 3D representation by stacking the slices alongside each other similar to a loaf of sliced bread. Unpredictable body movements of either the mother or foetus during the US scan could lead to motion artefacts in reconstruction of the 3D volume if the position of the structures is not explicitly tracked. Consequentially the slices acquired by this method may require rearrangement so that they do not stack in a parallel manner to represent the heart in 3D and often irregular gaps arise in the reconstruction. The problems of 3D reconstruction by 2D slice sweeping can be overcome by full volumetric acquisition using a 2D array of transducers in the probe head. However, these devices usually trade off volume acquisition speed for better inter slice continuity over sweeping planar acquisition methods.

2.2 Salient Features in echocardiography

The US image is derived from acoustic reconstruction of reflectors along a series of scan lines. Along each scan line the reflectors are assumed to be point like. The Euclidean path of the beam and its echo are assumed to follow Snell's law of reflection. Most scanners do not account for refraction in the reflector reconstruction and so display a map of assumed reflected rays between the source and acoustic boundaries within the path of the scanning beam. The strength of the received echoes is a function of the differences of acoustic impedances of the media on either side of the boundary. The acoustic impedance is a property of the density and elasticity of the medium. In the heart the blood-endocardium interface is one of the more prominent features in echocardiography. Since the myocardium contains layers of fibrous structures it gives relatively stronger echogenic signals than the blood pool region. Therefore several echocardiographic segmentation algorithms in the literature look for edge features to delineate the blood pool. The epicardium is usually more difficult to visualise and detect in US images because the acoustic impedance from blood-tissue boundaries is higher than from two different tissue types. Also the cardiac mode on US scanners can be set specifically to enable high blood-wall contrast. The segmentation task is to isolate these blood-wall boundaries from those produced by echoes from the fibrous layers in the myocardium. Within the blood pool papillary muscles and other thin structures can be considered to belong to the myocardium since they are not blood. Different research groups choose to include these structures as part of the myocardium and others classify them as part of the blood pool however, this distinction depends on what is to be measured. If the size and shape of the walls are to be measured over time (tracked) then these structures can be treated as blood since they do not contribute to the functional movement of the myocardium. If the blood volumes are important for example measurement of cardiac output then the intra-chamber structures occupy space and must not be considered as blood. This thesis aims to quantify cardiac output and stroke volume which are functional measurements of blood flow within the chambers. This implies that the intra-chamber structures such as papillary muscles are treated as part of the myocardium.

2.3 Low Level Ultrasound Image Processing

Images will always contain some degree of noise since no real world acquisition system is perfectly noise free. The process of digitising images introduces quantisation noise in addition to the artefacts of the imaging system. In sonography the main noise sources are due to reconstruction of the image from the reflected ultrasound waves. Artefacts from the reconstruction include shadowing, enhancement, partial volume effects and speckle. The speckle effect is caused by

scattering from non smooth or grainy reflectors within the path of the ultrasound beam (section 2.1.2.5). Usually speckle reduction filtering is required before applying a segmentation algorithm to detect either boundaries or regions within the image.

2.3.1 Speckle reduction techniques during acquisition

Ultrasound scanners typically have some speckle reduction filters integrated in the acquisition process. These algorithms are often not disclosed by the manufacturer and have very few parameters to set since in clinical use a simple repeatable method of imaging a specified organ is desirable. Often speckle reduction is present in the scanner's user interface with an option to adjust it to control the amount of filtering. Tomtec representatives for the Phillips Live 3D scanner claim that the filtering is frequency based but can not provide the details of its operation. There is a control on the scanner to adjust the amount of low level signals in the image but does not specify the nature of these signals or how the filtering is done. Some of the new scanners (which may include the Phillips Live 3D) are capable of speckle reduction by compounding. Two of the common methods are described below.

Pulsed ultrasound tends to have a correlation length of ~1 microsecond [25]. Correlation length refers to the time length beyond which two events are uncorrelated. The degree of correlation can be defined by the normalised cross correlation coefficient ρ. For two random signals x and y the time expression of ρ is [35]:

$$\rho_{x,y} = \frac{\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y)dt}{\sqrt{\int_{-\infty}^{\infty} (x-\mu_x)^2 dt \int_{-\infty}^{\infty} (y-\mu_y)^2 dt}} \tag{2.9}$$

For a 2D image (2.9) would have double integrals and triple integrals for 3D.

If the image is captured during the correlation length then a single speckle pattern is observed. When several images (acquired from different orientations or from the same direction when using various transducer performance characteristics) are averaged the temporally uncorrelated speckle phenomena will not constructively interfere and so will be greatly reduced in magnitude. If the speckle components are randomly distributed in time then stacking reduces the contrast of the speckle pattern by a factor of $n^{1/2}$ (where n is the number of signals that are stacked) [36]. If

spatial compounding is performed from acquisition at different aperture positions the geometry will have to be precomputed by the scanner so that the images to be averaged are correctly aligned. The drawback of speckle compounding is that the final image can only be output after stacking multiple images and this reduces the acquisition rate. For this reason speckle averaging is usually disabled for foetal heart acquisition.

Frequency compounding allows averaging of speckle patterns from different frequency bands in k-space. Each frequency band has different support in k-space (k-space refers to the image in the Fourier domain). The signals in each band require some degree of correlation to those in other bands to be useful in this method of speckle reduction. The signal to noise ratio will vary in each band and so levelling may be required but this may raise the noise in some bands. One of the main improvements is signal to noise ratio but this is balanced by a decrease of spatial resolution [34]. This can be seen in equation (2.10) where the axial speckle size (FWHM) is inversely proportional to the frequency halfwidth that corresponds to the transmitted pulse (Δf).

$$FWHM = \frac{0.61}{\Delta f} \tag{2.10}$$

All spatial filtering methods degrade the image information to some extent; in most cases this involves low pass filtering of spatial frequencies and this broadens edges. This will lead to poorer delineation of the myocardium and small components such as papillary muscles and valves. The fine intra cavity components may broaden and appear to merge with surrounding structures. Speckle reduction algorithms work best in homogeneous regions where the speckle is unrepresentative of the underlying structures and offer little useful information. Since the algorithms are routinely used for acquisition protocols but their internal workings are not disclosed, their smoothing effects will be treated as part of the noise in the image formation. The 3D dataset obtained by 2D slice sweeping from the Acuson scanners has a higher speckle content and grainy appearance than the images from Phillips Live 3D. Whilst this may be due to the smaller pixel size in 2D, it may also be due to little or no speckle filtering applied by the scanning protocol in the Acuson scanner whilst some pre-filtering was present during acquisition with the Live 3D scanner. It is possible that this phenomenon is due to the finite width of the beam so that speckle originating from adjacent beam planes are superimposed onto the image. This effect is worsened if the structures imaged are moving perpendicular to the image plane.

2.3.1.1 Speckle reduction algorithms after acquisition

In the absence of a theoretical model of the noise distribution the classic Gaussian kernel is often employed as the low pass filter since in the Fourier domain its transform is conveniently another Gaussian. In Marr and Hildreth's paper [37] they show how useful Gaussian convolution can be across scale when combined with zero crossings of Laplacian. They argue that the Gaussian filtering limits the rate of intensity variation and the Laplace operator applied to the result will pick out the strongest and most likely edges using the following equation.

$$\nabla^2 G(\underline{r}) = -\frac{1}{\pi\sigma^4}\left(1-\frac{\underline{r}^2}{2\sigma^2}\right)e^{-\frac{\underline{r}^2}{2\sigma^2}} \tag{2.11}$$

Where $G(\underline{r})$ denotes the Gaussian operator in 2D and σ is the standard deviation. The filtering operation is separable to allow sequential application in each dimension.

If this method is used in at least two scales to find the loci of common zero crossings the false negatives can be weeded out leaving the most probable candidate edges. Scale space involves first low pass filtering the image and then sub-sampling it so that a hierarchal image pyramid is constructed. If the filtering used in each stage is Gaussian, the structure is then said to be a Gaussian pyramid (Figure 19). Note this method assumes that the desired edges in the image are more prominent than the underlying noise.

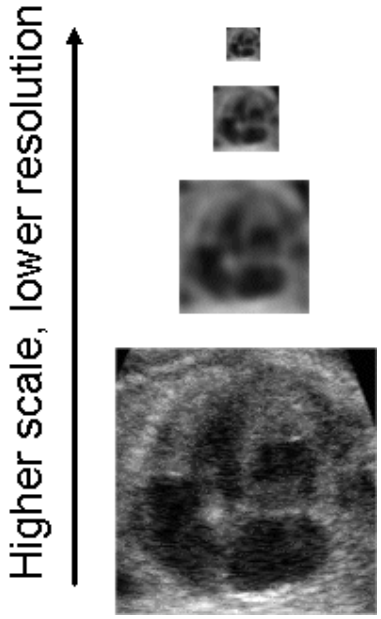


Figure 19 Gaussian image pyramid. At each level of scale the image is Gaussian filtered and then sub-sampled.

Low level segmentation techniques work directly on the image pixels and have little knowledge of the image content other than local neighbourhood information. These

algorithms usually rely on salient image features to be present in the image for example homogeneous regions or strong edges in comparison to the level of image noise. By themselves data driven approaches can make oversimplified assumptions about the structure of interest within real noisy images and can lead to inaccurate and spurious segmentation. Edge detection operations are a key example of this since they accentuate local changes in intensity gradient at region interfaces but also give high responses with speckle in US images. The traditional method to reduce the noise is to pre-process an image by low spatially frequency filtering so that the image is at a higher scale [37;37], [38] before attempting to extract features from it. There is a trade of between the amount of high frequency noise filtering and good edge localisation. At a particular spatial scale most of the noise is removed from the image whilst leaving intact, various regions that correspond to different materials. These regions can be detected and labelled by region growing methods, the edges between them identified by performing edge detection. Morphological Operations can be run on regions to shave misidentified pixels on the region boundaries or to fill gaps between two adjacent similar regions. The problem with these algorithms is that they are applied equally to the entire image unless manual intervention is used to restrict the set of pixels to a particular region.

2.4 Deformable Model Based Segmentation

This class of segmentation methods uses a higher-level approach to examine the structure within the image by imposing a priori knowledge of the parameters that describe the object of interest. These methods are widely recognised for their ability to deal with the variability of biological structures and allow user interaction in the segmentation process. Deformable models can be parameterised in a compact representation to allow them to be applied to the image repeatedly in a manner that is computationally efficient and with small storage requirements. The biological organ has elasticity and has some minimum energy associated with its shape. This constraint is incorporated in the modelling by representing the organ as a surface or volume with only allowable deformations in the image. Deformable models come from the theory of elastic structures under the action of physical forces. The general deformable model is initialised within the image and is deformed by a combination of image forces and regularisation constraints to lock on to salient image features. The energy of the model can be minimised using a variational calculus approach such that the final smooth shape incurs a minimum degree of bending and potential energy. A variety of methods have been combined with deformable models the most common approach is to use edges and homogeneous region information, however other more complex approaches may involve use of texture models of the image or information

from training images that are stored in addition to the shape of the deformable models.

2.4.1 Explicit Parameterisation

Kass et al. created the snake model in 1988 [39]. It is an energy minimising deformable contour that detects image features and was designed for semi-automatic segmentation problems. The concept of a snake deformable model has been applied to feature segmentation and tracking in images in several applications including medical data [40], [41], [42], [43], [44], [45], [46]. Since its introduction in computer vision the model has been modified over the years by researchers for their specific images. The basic snake equation as proposed in 1988 is as follows

$$E^*_{snake} = \int_0^1 E_{snake}(\underline{v}(s))ds = \int_0^1 E_{\text{int}}(\underline{v}(s))ds + \int_0^1 E_{ext}(\underline{v}(s))ds + \int_0^1 E_{con}(\underline{v}(s))ds \qquad (2.12)$$

where $\underline{v}(s)$ is a vector of coordinates of the snaxel (snake-element or snake-pixel) in parametric form. E_{snake} is the energy of the snaxel, E_{int} is the internal energy due to bending, E_{ext} arises from image forces and E_{con} is given by external constraint forces. The first term on the right hand side is a regularisation term that controls the smoothness of the shape of the contour and can be written as

$$\begin{aligned} E_{\text{int}} &= E_{tension} + E_{curvature} \\ E_{\text{int}} &= \left(\alpha(s)\left|v_s(s)\right|^2 + \beta(s)\left|v_{ss}(s)\right|^2\right)\Big/2 \end{aligned} \qquad (2.13)$$

The first term in equation (2.13) controls the tension between snaxels and the second regularises the curvature. α(s) and β(s) are weighting coefficients for the respective terms and whose values are selected on the basis of their relative importance. Usually the values of the weighting terms are image dependent and may be chosen as constants. In the original Kass method [39] these terms were approximated as finite differences.

The E_{ext} term in (2.13) is an interaction of the image on the snake and usually incorporates some form of edge detector across scale space in medical images and the constraint term can be developed for interactive use.

To lock on to salient image features whilst requiring a smooth contour, the snake integral must be stationary and is minimised in variational form to give rise to the following two independent Euler equations:

$$\alpha x_{ss} + \beta x_{ssss} + \frac{\partial E_{ext}}{\partial x} = 0$$
$$\alpha y_{ss} + \beta y_{ssss} + \frac{\partial E_{ext}}{\partial y} = 0 \qquad (2.14)$$

These coupled pair of equations can be solved iteratively in implicit form by a gradient descent method.

In its current form the basic snake equation (2.12) can shrink to a point or a line depending on the boundary conditions and this requires the initial curve to be manually placed close to the image feature of interest. To avoid manual initialisation various additions to the snake equation were introduced such as the balloon model by Cohen LD (1991) [43]. In this method the following balloon term is added to the snake model

$$E_{balloon} = k\underline{\hat{n}}(s) \qquad (2.15)$$

where $\underline{\hat{n}}$ is the unit local outward normal to the snake and k is the weighting coefficient for this term. Its sign controls whether the curve shrinks or expands. In practice the weighting coefficient for this term is chosen such that it does not dominate the salient image points. This term allows the user to select a single point within the region of interest and the snake can grow or shrink to align itself against the prominent edges of the structure. The balloon model allows the contour to grow and pass over weak edges reducing the need for a manual initialisation of the contour close to the feature of interest. However, its main weakness is its failure to handle sparse data such as images with missing edges [47]. Such situations can arise in the case of ultrasound imaging sequences where shadowing can cause black streaks to hide features behind bony structures or dropout artefacts of unresolved thin structures (section 2.1.2.5). In the heart there are naturally open boundaries such as the valves opening and closing repeatedly within the cardiac cycle and in the foetal heart, septal defects which are holes that exist in the septum in pathological cases. These scenarios provide the balloon term with an opportunity to punch through and invade the space outside the chamber of interest. It is possible to adjust k so that the contour is penalised for leaving a chamber or by imposing additional constraining terms but these issues are still ongoing areas of research.

An alternative model to the balloon term is the Gradient Vector Flow field (GVF field) [47] which is the vector field $\underline{v}$ that points towards the local maximum of the field. The vector field is derived from diffusion of an edge map using the Laplace equation or steady state diffusion equation. Expressed in a vector notation the GVF field becomes:

$$\underline{F}_{GVF} = \mu \underline{v} + \left|\nabla f\right|^2 \left(\underline{v} - \nabla f\right) \tag{2.16}$$

Where $\nabla f(x, y, z)$ is an edge map derived from the image. For homogeneous intensity regions, ∇f is zero and therefore the first term in the equation dominates by smoothing the $\underline{F}_{GVF}$ solution. When it is large the second term dominates and draws the contour towards local edges. Figure 20 shows the evolution of the GVF snake to the edges in an image by edge flow diffusion.

Figure 20 GVF diffusion edgemap (first image) and stages of evolution of the contour to segment a non convex shape by the GVF snake.

This formulation replaces the E_{ext} term in the basic snake equation with an edge detector when edges are in the vicinity of a particular snaxel and a vector field that points in the direction of the nearest edges in the absence of any local edges. The vector field is provided by the diffusion of either an edge map or of the image itself. The weighting coefficient μ is to adjust the relative importance of the first term in relation to the second; in the presence of more noise the diffusion field is given more significance when local edges are not present or reliable. Two advantages over the balloon model are that the GVF snake can be initialised as a contour partially occupying the structure to be segmented; and holes in this structure do not lead to punch through (although if there are strong edges on the other side of the hole the snake can be drawn towards them). Xu and Prince [47] have demonstrated that the GVF field can allow the snake to segment deep concave structures more reliably then a snake driven by a balloon or chamfer edge map.

Figure 21 illustrates some drawbacks of the GVF field in the presence of a complete edgemap; it has a finite effective range from the edges in the edge map. At a certain distance from the edges the influence from the GVF field may become dwarfed by the other terms in the snake equation. This threshold level of adequate GVF contribution is dependent on the distance to the edge, the strength and width of the edge, i.e. the integration of edge magnitude over the edge profile. The level will also be affected by the contributions from the other terms. Another problem with the GVF field is an apparently undocumented rare convergence of the snake to false minimum in the centre of a GVF crater. Since the contour is pulled in two opposing directions it will have difficulty in reaching the desired edge (Figure 21). To overcome a weak or undefined GVF field a balloon term was added to the GVF snake equation by Sanchez et al 2000 [48]. Tao and Tagare, 2005 [49] proposed a method of tunnelling descent

which can be used to overcome local minima. In their method the snake can climb out of local minima even if all paths from this point increase its energy further. It is able to escape this false convergence with a constrained expansion until the snake starts lowering its energy again. This could be applied in the optimisation of the GVF snake to reduce convergence to false minimum in the centre of the GVF crater. An alternative method of avoiding local minimia would be to use a stochastic optimiser such as a genetic algorithm or simulated annealing to minimise the energy of the snake equation.

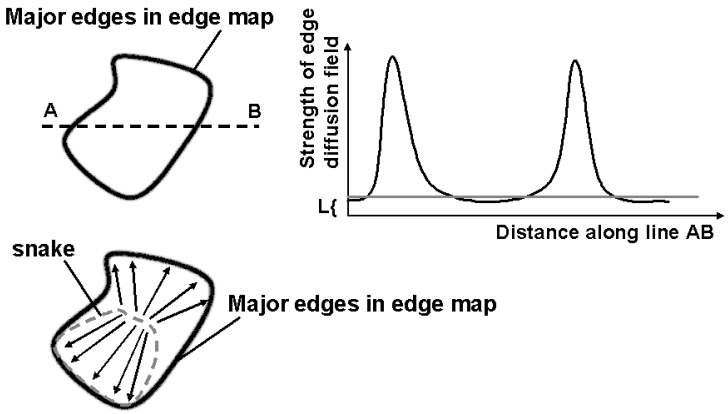


Figure 21 Illustration showing some of the potential problems with the diffusion field of an edge. Top row: Horizontal profile through the edge map where L is a threshold dependent which may give negligible contribution from the diffusion term. Bottom row: An example where the snake could become trapped in a false minimum.

The deformable models presented in this chapter process single frames of the cardiac cycle. Temporal tracking was not implemented in any of the snakes since it was observed that initialisation from the previous frame would lead to propagation of errors in the segmentation as each frame in the cardiac cycle is processed. In [50] I implemented a rigid body shape constraint to fit the current frame based on the previous frame but this requires a good segmentation of the previous frame to be effective. Thus this section is aimed at providing different algorithms to segment isolated frames as accurately as possible; tracking can be implemented at a later stage when the initial frame is satisfactorily segmented.

2.4.2 Explicit Boundary Parameterisation

The snake was assumed to be fully closed both as a contour 2D or a surface in 3D. The explicit contour parameterisation is trivial since each node is connected to two nearest neighbours by a straight rod. The rod would become shorter by placing more vertices in the contour until at the resolution of the image, the rods become short enough that there is more than one vertex in a pixel. In practice this limit would be too severe since explicit parameterisation is prone to self intersection in places where

many snaxels are close together. A solution to this problem is to have a threshold on how close vertices could be before decimation of the mesh nodes. The large gaps between vertices could be interpolated.

In 3D the contours can be stacked on top of each other in the z direction. There are two approaches that could be taken in the construction of a closed surface by extension of the contour parameterisation. The first and most simple case was to enforce smoothness in the z direction by an inter-slice internal force that would let the contours evolve in 2D in their own slice planes. The main drawbacks of this method are that motion is not allowed in between the slices. Hence sub-voxel resolution can not be achieved and therefore the cap slices of the chamber would be difficult to seal off. This creates a pseudo 3D snake surface as shown in Figure 22.

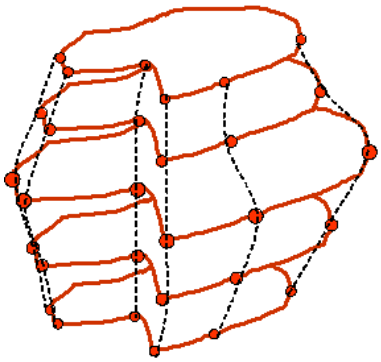


Figure 22 Illustration of connectivity (dashed lines) between adjacent snake planes to create a pseudo 3D mesh.

A more accurate approach would be to allow the contours to evolve in all three dimensions so that they do not remain planar. This was implemented to produce a 3D enclosed surface as represented by Figure 23. This approach to describe a closed deformable surface of the endocardium as in cartography of the Earth was adapted from Bosnajak et al [51]. The advantage of this anisotropic surface description is that the cardiac chambers are likely to have a principal axis, and this can be aligned with the cap snaxels to provide better parameterisation of the higher curvature at these extrema. This method was used in this thesis to describe the 3D explicit snake surfaces.

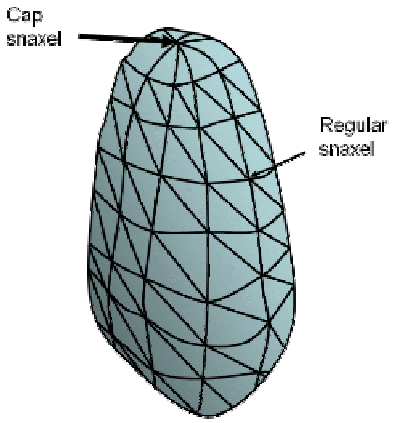


Figure 23 A 3D enclosed surface of snaxels constructed from rows of tessellated triangles. Regular or body snaxels have 6 nearest neighbours and cap (polar) snaxels have N nearest neighbours where N is the number of snaxels per line of latitude in this cartographic geometry.

A variety of snake approaches were applied to the dataset each snake force equation is presented together with its assumptions and limitations. The explicit snake parameterisations are implemented in both 2D and 3D and the implicit formulations are able to handle 2D and 3D with no changes. The explicit snakes are presented here in chronological order that they I implemented them. They will be compared to the level set snake proposed in this thesis on 3D synthetic and both 2D and 3D real data to show their limitations.

The snaxel positions in the explicit snakes were updated by resolving forces on every iteration. The evolution was treated as a constant acceleration system of particles whose state was computed by an explicit Euler update as follows

$$\begin{aligned} \underline{x}_{t+1}^{i} &= \underline{x}_{t}^{i} - \underline{v}_{t}^{i} dt \\ \underline{v}_{t+1}^{i} &= \underline{v}_{t}^{i} - dt \sum_{n=0}^{N} \underline{F}_{t,n}^{i} \end{aligned} \tag{2.17}$$

where $\underline{x}$ is the position of a snaxel i, $\underline{v}_t^i$ is its velocity at time t and $\underline{F}_{t,n}^{i}$ refers to the nth force term acting on snaxel i at time t. Note the resultant force acting on the snaxel was normalised before adding its contribution to the position of the snaxel to reduce instabilities in the Euler update equation. The forces involved in (2.17) can be turned into acceleration by dividing by the mass of the snaxel. In this implementation the mass was arbitrary set to unitary. The Euler numerical method of updating the position of snaxels is used in many snake algorithms based on the classical snake model [39] because of its low computational complexity but requires small timesteps to avoid numerical instability. The time update equation for the level set snake was defined in a similar way as described in section 2.5. The timestep was set to a very small value often much smaller than the pixel dimensions.

It was discovered that the update order for the snaxels can affect the stability of the snake evolution. The snake appeared to rotate slowly as it evolves if the snaxels are updated in the same order on subsequent iterations. This effect is due to bias in the time update of the numerical solution and was compensated for by reversing the order of snaxel updates on even iteration numbers. This constraint was also enforced for the level set snake and any iterative solver used in this thesis such as in the generation of the edge maps by iterative solution of the diffusion equation.

All explicit snake implementations presented in this thesis have bilinear or trilinear image interpolation as appropriate. Since the implicit snake was created on a regular grid to match the image size point wise image access without interpolation was used where required.
All edge detection algorithms use centre difference evaluations to reduce spatial offsets.

All the snake implementations presented in this thesis were solved iteratively using gradient descent methods for speed. This was of particular importance with the level set snake which takes much longer to evolve than the explicit versions. Stochastic methods are expected to converge to more accurate solutions but it was expected that future implementation of a shape prior into the snake equation will provide the behaviour of convergence towards global minima with a computational complexity of a gradient descent based method. Since the current snakes rely on image gradient and region information with regularisers, drop out artefacts and shadowing will not be reliably segmented using any type of snake without the presence of a shape prior term regardless of the optimisation method.

The main driving force is defined as the term that allows the snake to migrate towards the salient image features that are to be segmented.

2.4.2.1 Scale Space Snake

This snake equation is analogous to the original snake developed by Kass in 1988 [39]. However, the force equation for a single snaxel at position $\underline{x}$ resembles unit masses on springs obeying Newton's second law of motion.

$$\dot{\underline{x}}_i = \alpha \sum_{n=0}^{N-1} \left(\underline{x}_i - \underline{x}_n \right) + \beta \sum_{n=0}^{N-1} \left(\underline{M}(\underline{x}_i) - \underline{x}_n \right) + \gamma \nabla G\left(I(\underline{x}) \cdot \hat{\underline{n}}(\underline{x}) \right) \tag{2.18}$$

The first term on the right hand side refers to local continuity (tension), the second resists bending and the third a local interaction with the image. G is a Gaussian filtered image I, and $\hat{\underline{n}}$ is the unit vertex normal vector to the snake boundary. $\underline{x}_i$

is the position of the snaxel i and $\underline{x}_n$ is the position of the nth nearest neighbour snaxel to i. To make use of edge polarity in ultrasound images it was assumed that the outside edge pixel/voxel belongs to the myocardium and so has a greater intensity value than the inner one (blood chamber). Typical greyscale values in the image are shown in section 1.4.3. N refers to the number of nearest neighbours around snaxel i and $\underline{M}$ is the midpoint of the convex hull defined by the nearest neighbours and snaxel i. In some instances $\underline{M}$ may lie on a saddle (in 3D) but will still correspond to the local minimum of the nearest neighbours. α, β, γ are weighting coefficients for the corresponding terms.

The first term shrinks the snake in the absence of any external forces and is similar to the curve shortening effect in the level set literature and causes curve to smoothen until a circle is reached before the snake collapses to a point. Its main use is however to locally redistribute the snaxels on the snake to reduce vertex clustering.

Under steady state conditions the left hand side of equation (2.18) reduces to zero. There is no damping present in this equation since it was found by experiment that sufficiently small time steps proved adequate to substantially reduce them. This was set as a constant and was usually much smaller than the pixel size. The equation was solved numerically as an initial value problem using explicit Euler update for the time step. Although this is a force equation the "energy" for the snake can be computed on every iteration. One iteration is defined as an update to all snaxels in the snake. The energy state of the snake can be obtained by adding the vector magnitudes of the first two terms and subtracting the vector magnitude of the third for each snaxel after each iteration integrated for all snaxels. In this way increases in stretching and bending of the snake lead to higher energy and edges from the image have the effect of minimising the energy.

The snake was allowed to evolve in Gaussian scale space, at each level the number of allowed iterations was reduced proportionally to the change in image resolution. Some C++ code for scale space snake evolution framework is shown below:

```
initialfiltersize = 7;  //KERNEL WIDTH OF FILTER
tmax = 3;   //SET TO EITHER 3/4 DEPENDING ON IMAGE RESOLUTION
numSnakes = 4;    //4 CHAMBERS IN HEART
for(t = tmax - 1; t >= 0; t--)
{
      fac = pow(2, -t);
      fraction = (filtersize-t)/initialfiltersize;
```

```
        multiResImg = &image;
        multiResImg.resampleImage(fac);
        for(n = 0; n < numSnakes; n++)
        {
              allSnakes[n].calcSnakeCOM(com);
              com *= 2;
              allSnakes[n].moveCentroidTo(com);
              allSnakes[n].scaleTheSnakeFactor(2);
        }//n
        for(i = 0; i < maxIterations; i++)
        {

              for(n = 0; n < numSnakes; n++)
                    allSnakes[n].advanceSnake(&multiResImg);

        }//i
}//t
```

The scale space snake uses the local edge term as the main driving force.

2.4.2.2 Gradient Vector Flow (GVF) Snake

The scale space snake is notorious for restricted influence from the local region. To overcome this Xu and Prince [47] proposed a new external force for snakes to replace the image term in equation (2.18). The equation of the GVF snake when cast in vector form was implemented as the following:

$$\dot{\underline{x}}_i = \alpha \sum_{n=0}^{N-1} \left(\underline{x}_i - \underline{x}_n \right) + \beta \sum_{n=0}^{N-1} \left(\underline{M}(\underline{x}_i) - \underline{x}_n \right) + \mu \nabla u + \nabla f \left| \underline{v} - \nabla f \right|^2 \tag{2.19}$$

The main driving force in equation (2.19) is the GVF term which is composed of the diffusion and edgemap terms. To create the edgemap for the GVF field Gaussian was used to reduce the speckle content because it is a low-pass filtering technique. The optimum filter size for use on the images to allow adequate segmentation varied with the speckle content of the individual dataset. Images from the Phillips Live 3D scanner contained less speckle noise than those acquired from the 2D slice method. Edge detection using centre difference was performed on the filtered image. The following anisotropic diffusion equation is solved iteratively to propagate the edge influence throughout the image u

$$\dot{u} = \exp\left(-\frac{\nabla^2 u}{\kappa} \right) \tag{2.20}$$

where κ is the diffusion coefficient. Since this equation is solved iteratively varying degrees of diffusion can be applied depending on the number of iterations.

One of the problems with edge driven deformable models (such as those evolving by GVF and scale space) are that the snake requires relatively strong and complete edges to be attracted to. In some stages of the cardiac cycle the ventricular valves are open and so these images present an open chamber to the snake surface. The atrial septum is very thin and is often suffers from severe dropout due to the poor resolution of the scanning beam perpendicular to its axis. The propagating surface tends to shrink away from the incomplete boundary as it evolves towards the remaining complete edges under the influence of the regularisation terms (Figure 24). This problem is more prominent in 3D than in 2D because of the increased connectivity surrounding each snaxel and so regularisation forces have more influence over the snaxel's motion for identical weighting coefficients. To some extent the addition of a balloon term to the GVF snake overcomes this effect but it is still present. The main driving force behind this deformable model is the GVF term although it is aided by a weak balloon term.

Figure 24 Some examples images where the 3D segmentation has been pulled away from a weak or missing boundary due to the internal regularisation forces. The chambers with partial boundaries are denoted by a white square. (Only central slices are shown).

2.4.2.3 Pseudo Zero Level Set (PZLS) Snake in Explicit Form

The main problems with the GVF field are that it relies on relatively strong, homogeneous edges to attract the snake. If there is an intensity inhomogeneity that affects the whole image then the edge diffusion field from areas of lower contrast will be weaker and may not attract the snake to sufficiently overcome the internal self-regularisation forces. As a result the snakes can completely collapse to a minimum size in the absence of any substantial image forces. The following snake equation (2.21) is adapted from the standard level set form similar to Lassige et al [12] and Corsi et al [52] to produce a balloon (advection term) that moves the snaxels along the local outer normals. **This is an original contribution that I made to the explicit snake field**. The speed function from the level set formulation is used to slow the balloon at image edges. Unlike in the fast marching method used by Lassige et al [12] the speed function used here is a vector quantity and so can reduce overshooting of edge boundaries.

$$\dot{\underline{x}} = F(\underline{x}) k \hat{\underline{n}}(\underline{x}) \tag{2.21}$$

Where $\hat{\underline{n}}$ is the local unit normal, k an empirical constant and F is a problem dependent speed function that modulates the balloon term. The true level set snake

is implicit in its parameterisation and so does not require interpolation if snaxels become unevenly spaced, therefore in this explicit Pseudo Zero Level Set (PZLS) form the continuity force in equations (2.18), (2.19) was added to equation (2.21) in order to maintain equidistant vertices so that the force on a given snaxel becomes

$$\dot{\underline{x}}_i = \alpha \sum_{n=0}^{N-1} \left(\underline{x}_i - \underline{x}_n \right) + F(\underline{x}_i)\hat{\underline{n}}(\underline{x}_i) \tag{2.22}$$

where k has been absorbed into the speed function. **Equation (2.22) was an original contribution that I devised as an intermediate step towards implementation of the level set snake proposed in section 2.8**. It has been included in this thesis for direct comparisons between the previous explicit snakes that rely on edge information. Equation (2.22) is similar to the form derived in [53] but uses an equidistant vertices constraint instead of Euclidean curvature. The speed function presented here relies on regional information instead of edge detection and edgeflow presented in [53].

Unlike the GVF snake which is data-driven (it is attracted to edges if their influence is strong enough), the balloon is the model term that propels the snake whose sign and magnitude can be controlled by the data. F returns a value within the range [-1, 1] and is defined by the image part of the Mumford-Shah (MS) energy functional (see section 2.5).

$$F = \lambda_1 \left[I - \mu_i \right]^2 - \lambda_2 \left[I - \mu_o \right]^2 \tag{2.23}$$

where I is the current voxel intensity under investigation. μ_i, and μ_o are the means of the internal and outside regions of the dataset defined by the snake boundary and λ_1, λ_2 are weighting coefficients. F is the MS term which is the main driving force behind this deformable model.

This is an improvement over Cohen's original balloon term [43] since the balloon term was fixed both in sign and magnitude and so could overcome weak edges very easily unless a threshold edge strength was set for example in [54].

The diagram in Figure 25 shows where this type of snake and all the proposed versions before it in this section will fail to produce correct segmentation. This can be corrected by local contour/surface interpolation in places where there are large distances between snaxels. Correction of this is trivial in 2D since the snake can be treated as a linked list and vertices can be inserted or removed easily. However, in 3D this is not as straightforward since the method of interpolation depends strongly on the type of connectivity between the snaxels. Triangular surface patches are well suited to this application but meshes that can be adaptively refined usually require a

reestablishment of connectivity whenever a snaxel is added or deleted from the surface. Local mesh refinement is not treated in this thesis since the proposed explicit snake model was intended to be replaced by an implicit one (see section 2.8) because of its elegance in the segmentation of complex shapes with no change in the parameterisation during its evolution.

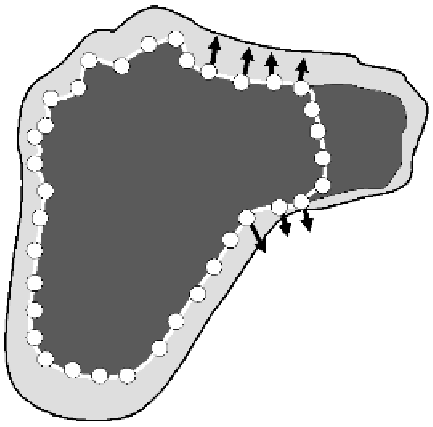


Figure 25 2D Illustration showing an example where local snake interpolation is needed. The arrows show the direction of the local contour/surface normals and so would be unable to propagate the snake into the boundary concavity since this will be actively discouraged by the continuity term. Local snake interpolation in areas where there are few snaxels will permit the snake to move into the boundary concavity.

In the preceding models intersections between neighbouring elements were penalised by the curvature and continuity terms. In traditional explicit snake implementations there is often no penalisation for an element to intersect with another non neighbouring element and so self intersections within the snake can still occur. This is just as likely to occur in 3D to cause cusps. To an extent reducing the number of snaxels can prevent self intersections from forming but it grows increasingly likely as the shape of the boundary becomes more complex. In this implementation the contour was regularised by repeatedly resampling in 2D and to some extent this reduced the phenomenon.

Table 2 lists the properties of each explicit snake. Scale space uses multiscale edges, PZLS exploits the MS term and the GVF relies on edge advection and edge detection to interact with the image. All three snakes use the same primitive initialisation which is a cylindrical mesh in 3D or a circle in 2D manually placed at the centre of each chamber.

Table 2 Explicit snakes properties

	Scale Space	GVF	PZLS
Edge term	yes	yes	No
Edge advection	no	yes	No

Region term	no	no	yes
Multiscale	yes	no	No
Main drive force	edge	edge advection	MS
Initialisation	primitive	primitive	primitive

2.5 Implicit Parameterisation and the level set

The snake theory up until this point deals with snakes represented in a Lagrangian framework. Whilst this approach can be considered to be more computationally efficient since the contour is explicitly parameterised without the need to impose a grid on the image, it has issues handling automatic topological changes and segmentation of long thin concavities such as a vascular network.

A good paper by Montagnat and Delingette [55] compares the approaches of the two snake parameterisations. The Lagrangian formulation of deformable models was implemented in this thesis since cardiac chambers can be considered smooth structures with simple topologies as reflected by the explicit parameterisation approaches in the literature [11], [41], [42], [13], [56], [7], [45]. It also defines explicit point wise correspondence between different contours/surfaces and can ease motion tracking for example in a temporal snake [11]. Shape registration and statistical analysis can be simplified by use of point wise correspondence (for example in active shape models [57]) and it is aimed that shape priors will be combined with the current work.

Some algorithms such as the "T-snake" proposed by McInerney [58] have elements of both explicit parameterisation and a simplical superimposed grid and are able to manage topological changes with ease. Contours are free to merge and split according to the topology present in the image structures. An alternative method to the T-snake model was the level set approach in which the motion of the contour takes place by curve evolution theory of partial differential equations instead of energy minimisation; however, energy minimisation practices can still be applied to the contour. The evolution equation of the snake is solved in an $n+1$ dimension space where n is the number of dimensions of the dataset. Since the approach is solved implicitly topological changes can occur naturally as illustrated by Figure 26 where the cones represent the evolution history of the level set function. The zero levels track the location of the front in image space.

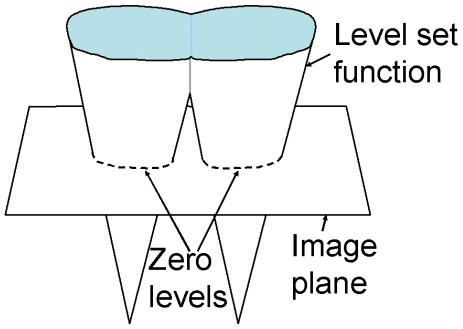


Figure 26 Illustration of topological change occurring in the level set function. The two fronts were initiated as isolated seed points that grow and can merge if not stopped by internal or external forces.

Although level set (LS) snakes have been used in adult cardiac segmentation [12], [59], [52], it is not as common an approach as the explicit parameterisation. The level set snake is implemented in this thesis however, the chambers are assumed to be isolated (even though the blood eventually passes through all chambers) so that their volumes can be measured independently over time. This is because volume estimation of the chambers in isolation is more useful clinically then the joint volume of the shared blood pool.

The generic level set equation for image processing can be written in the form

$$\dot{\varphi} = \Omega |\nabla \varphi| \tag{2.24}$$

where φ is the level set function and Ω is a problem dependent speed function.

A formulation of the temporal evolution of the LS equation in foetal cardiac images [12] is given by

$$\frac{\partial \varphi}{\partial t} = \phi |\nabla \psi| \left(div\left(\frac{\nabla \varphi}{|\nabla \varphi|} \right) + \underline{F} \right) \tag{2.25}$$

where $\varphi(x,y,t)$ is a multidimensional function of the curve in time, the sign of F controls whether the curve contracts or expands. In this equation ϕ is an image-based speed function that slows the curve at salient edges and is given by

$$\phi = \frac{1}{1 + |\nabla G_\sigma I|} \tag{2.26}$$

where G_σ is a Gaussian operator and I the image grey level intensity under investigation.

Echoes from the fibrous wall structures and thin papillary muscles structures within the cardiac chambers give rise to a speckle interference pattern around the endocardium. During the acquisition the partial volume effect spreads these structures to reduce the image contrast between the myocardium and cavity. The

endocardium boundary becomes convoluted with closely located intracavity structures and so edge detection algorithms will find many spurious edges around the chamber walls that will require filtering to reduce below a desired threshold.

The regional mean intensity information of the interior pixels/voxels and of background can be computed instead of edge detection since it is less sensitive to image noise within a homogeneous region. The velocity function $\underline{\phi}$ is controlled by the image part of the Mumford-Shah energy functional cast in force form [60], [61].

$$\underline{\phi} = \gamma_i \left(I - \langle I_i \rangle \right)^2 - \gamma_o \left(I - \langle I_o \rangle \right)^2 \tag{2.27}$$

In (2.27) i refers to the interior region and o the outside region, I is the local pixel/voxel intensity. γ_i, γ_o are two arbitrary constants that will have some dataset dependence. The difference in sign between the two terms can force the balloon term to change direction according to whether the front is closer to the background or the mean $\langle I_n \rangle$ of region n. A steady state is reached as the snake tries to separate the chamber from the background and was shown in [61] to be useful in images without clear boundaries. This function assumes that the regions inside and outside the propagating fronts are piecewise constant in intensity. A better way of modelling the distribution of the regions is found in [62] where the intensity greylevels are modelled by fitting a Rayleigh distribution using maximum likelihood expectation maximisation for 2D ultrasound images. However the difference in segmentation of automatic contours between piecewise constant and Raleigh distribution region terms was not substantial, although the spread in errors was greater using the piecewise constant case.

To initialise the level set from small seed regions a signed distance transform of the seed region boundary is used. This can often be an Euclidean distance but this can be computationally expensive especially in 3D since the distance of each voxel would be checked in order to find the minimum distance to the current voxel. A common speed up is to use the chamfer distance transform [63] or a variant which involves non Euclidean distances but often a suitable enough approximation. The advantage of this algorithm is that the entire signed distance transform can be computed in two passes - first forwards and then in reverse. In the implementation of the level set snakes in this thesis both the standard Euclidean and standard chamfer distance transform was used. It seems that during the course of the evolution the level set surface becomes highly non Euclidean as it moves towards the chamber edges. Chamfer DT was preferred for speed.

Narrow banding [64] is an adaptation of the LS method to reduce the number of active elements involved in the evolution of the front. When the front reaches the edge of the band the old band is set inactive and a new band is initialised. This technique allows the level set algorithm be locally influenced like the explicit snake. It is useful to reduce nucleation of new fronts if multiple objects (say cardiac chambers) are within the same vicinity and a different front is evolving in each. The width of the narrow band is application dependent and could be set to less than the thickness of the chamber wall. Narrow banding is also much faster than updating φ over the entire level set surface. Despite its speed advantages narrow banding was not used because of variable septum thickness in different images. It can be difficult to set beforehand a threshold that is suitable for the width of the narrow band since the septal width is not constant within the image due to the partial volume effect. However, apart from speed disadvantages lack of narrow banding can increase under image leakage through thick walls given sufficient number of iterations to the level set equation.

The LS used in this thesis was solved using a first order iterative scheme because of its low computational complexity

$$\varphi_{n+1} = \varphi_n + \Delta N F \varphi_n \tag{2.28}$$

where n is the iteration number and ΔN is the timestep. Upwind differencing schemes (see equation (2.37)) were used where appropriate to maintain numerical stability around the propagating front as well as a small timestep.

2.5.1 Non-Amorphous Deformable Models

Deformable models such as the snake allow a flexible structure to be fit to the image boundaries to segment the chambers of interest. However these models can take on any shape that fit the constraints of elasticity and curvature imposed by the model. If the boundary is not well defined, partially occluded or saturated by noise, as is often the case in ultrasound images then conventional image driven deformable models fail to converge to the desired boundary. Shape templates allow spatial constraints that no elasticity and curvature penalisation can match since it comes from the accepted segmentation of similar images.

2.6 Integrating Shape Prior Information in the Level Set Equation

There are various approaches to incorporating shape influence into the level set framework e.g. [65], [66]. One of the simplest methods is the signed difference between the current level set and the shape prior level set

$$\varphi_{SP} = \Phi - \varphi \tag{2.29}$$

where φ is the current level set segmenting the image and Φ is the shape prior level set. For this contribution to be meaningful Φ must be registered to φ, therefore as the snake evolves updating the registration parameters is required. However, in Cremers et al [67] changes in pose and orientation between the dataset and model were not accounted for.

2.7 Fitting a Template to the Image

The task of aligning a template to an image usually involves solving a general transformation matrix of the form

$$\begin{pmatrix} x' \\ y' \\ z' \\ w' \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ T_{21} & T_{22} & T_{23} & T_{24} \\ T_{31} & T_{32} & T_{33} & T_{34} \\ T_{41} & T_{42} & T_{43} & T_{44} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \tag{2.30}$$

$$x' = Tx$$

Unlike the perspective transform in equation (2.30), rigid body transformations are such that the shape is preserved and parallel lines remain parallel. These transformations involve rotation, translation and scaling. To allow for size changes scaling can be also incorporated into the matrix. In 2D the global transformations look like this

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{pmatrix} 1 & 0 & tx \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} sx & 0 & 0 \\ 0 & sy & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{2.31}$$

$$x' = Tx$$

where t represents translation and s scaling. To determine the optimum transformation as cost function is employed to measure the quality of the fit of the template to the image. The cost function used was taken from Tsai et al [68]

$$E_{align} = \frac{\int_{\Omega} (I_T - I_F)^2 \, d\Omega}{\int_{\Omega} (I_T + I_F)^2 \, d\Omega} \tag{2.32}$$

where Ω is the image domain, T and F denote the transformed and fixed images respectively. The denominator penalises the transformed image from shrinking to zero.

Differential Evolution [69] was chosen for minimisation of (2.32) because of its global optimisation properties. It is a population based stochastic optimiser that operates similar to genetic algorithms in the way a selection from a population is randomly mutated with random crossovers to obtain new trials for subsequent generations. In addition to this Differential Evolution builds a weighted difference vector (Figure 27) from a random selection of the population. This difference vector is added to another randomly selected population member and spliced with the target vector taken from the current population to build a new trial solution. This is evaluated by the cost function and if it is small then it is retained for the next population level. The number of generations is a variable of the application as with the number in each population.

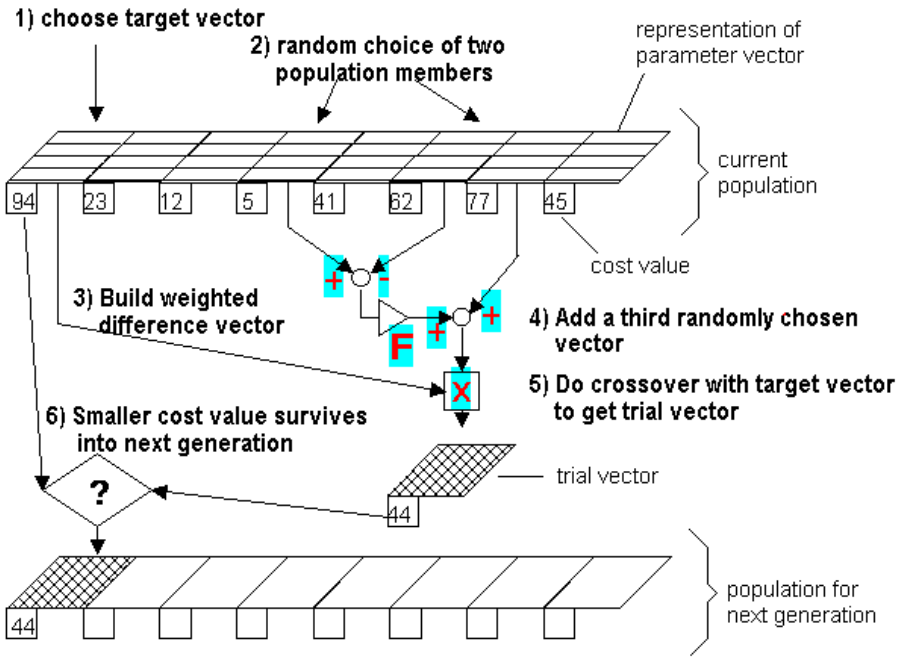


Figure 27 Differential Evolution flowchart [69].

This method of functional minimisation was used in my publications [70], [71]. However, for the results in this thesis the registration algorithm used was still rigid body and the optimisation algorithm was multi-resolution gradient descent (for speed in 3D) written by in house developers for use in both 2D and 3D. The software used was "regtool" written by CMIC (Centre for Medical Image Computing, UCL). Since the regtool algorithm did not provide a user defined cost function, the nearest method to equation (2.32) was used which was simple sum of square differences. In

practice if the source image was close enough (in translation, rotation and scale) to the target image satisfactorily registrations results were easily obtainable. It was found that to obtain reliable registrations both images were manually cropped to similar regions of interest, scaled to approximately the same size and were centred prior to running the algorithm.

2.8 Level set snakes implemented in this thesis

The material below is taken from my publication [71]. In contrast to explicit models, level set approaches use a dense contour parameterization usually at the native resolution of the pixels in the original image. The implicit evolution avoids the need to track surface markers in relation to each other and can allow for topological changes.

In this section I outline the basic equations behind the explicit snakes and LS approaches implemented in the foetal heart literature. For each snake I define the term main driving force as the force that propels the snake outwards from a small initial circular region within the chamber. For the remainder of the thesis the acronym SP denotes Shape Prior term.

In each level set unless otherwise stated the front was manually initialized as a circle or sphere in each chamber. Each chamber contained its own snake which was stored in separate memory space from those in other chambers to prevent the common level set merging of neighbouring fronts. The distance transform for each snake was defined as a cone with negative values inside the front and positive elsewhere. This can be computed very quickly for such simple geometry in a single pass by testing voxel position against radius of the front. For non primitive initializations it may be necessary to use more general efficient distance transforms such as chamfering. The usual criterion of normalising the distance transform was enforced.

The LS equations are based on the general formulation in (2.24) and were solved using an explicit Euler time integration method for speed. The timestep was set to the smallest of the pixel/voxel dimensions.

2.8.1.1 Edge-based level set

In 2000 Lassige et al [12] applied the following level set to foetal echocardiography:

$$\dot{\varphi} = g_2 \left(e\nabla \cdot \frac{\nabla \varphi}{|\nabla \varphi|} + v \right) + g_2 w. \tag{2.33}$$

In this equation e is a weighting constant, v is a constant advection term which is the main driving force and g_2 is an edge detection stopping term given by $g_2 = \left(1 + \left|\nabla G I^2\right|\right)^{-1}$ where G is the Gaussian operator applied to image I. The LS function φ in (2.33) is edge based and the constant negative advection term provides an outward growing force that is slowed by the presence of a prominent edge in the image such as at a strong blood-endocardium interface. If the image gradient is not infinite in magnitude the constant advection term may push the propagating front through weaker edges but at a reduced rate. Throughout this thesis I will denote this algorithm as Edge-Penalized Constant Advection (EPCA).

2.8.1.2 Region-based level set

Based on the Mumford-Shah (MS) energy functional for active regions [60] and edge flow using Sarti's work [72], Dindoyal 2005 [73] presented a LS deformable model with region competition. **This is an original contribution that I made to the level set literature**. This formulation can be found in equation (2.34) where $g = \left(1 + \left|\nabla G Q\right|\right)^{-1}$ is an edge detector that returns a value between 0 and 1, with G denoting Gaussian filtering and Q is the image. φ is the LS function, I is the current voxel intensity under investigation. μ_i and μ_o are the means of the internal and external regions of the dataset defined by the LS front.

ξ is a function that tests if any of the enclosed regions from individual snakes overlap. If there is overlap ξ returns 1 if true and 0 otherwise. The collision detection component is heavily penalizing and tends to stop two intersecting fronts immediately upon contact so that a steady wall is formed where the two meet (an example of this is shown in Figure 28). This allows blood volumes to be measured for each of the 4 cardiac compartments in isolation. From preliminary experiments it was discovered that for the collision to occur at the right place (where part of the chamber wall has suffered signal dropout due to the beam resolution); the two snakes should be started from as close to the centres of their respective chambers as possible. This prevents one snake from invading the adjacent cavity due to its arrival at the missing boundary first.

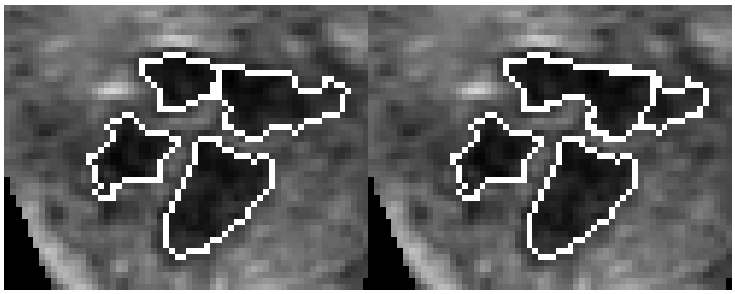


Figure 28 Effect of the inter snake collision penalisation term (enabled on the left image and disabled on the right image).

This equation uses the MS term which is the main driving force and is less prone to changes in intensity than simple edge based methods since it uses the average intensities of the regions inside and outside the snake. Its main assumption is that the regions are piecewise constant and tries to achieve segmentation by separating the background mean from foreground mean. The exponential curvature dependence is to reduce leakage into shadowed regions of the image but this is dependent on the influence of the weighting coefficient κ. The weighting coefficients α, β, λ_1, λ_2, κ for the respective terms were determined experimentally as for all the deformable models presented in this thesis. Justification for the weighting parameter choices is determined in 5.2.

$$\begin{aligned}\dot{\varphi} &= \left\{ \alpha g \nabla \cdot \left[\frac{\nabla \varphi}{|\varphi|} \right] + \beta (\nabla g) \cdot (\nabla \varphi) \right\} \|\nabla \varphi\| \\ &+ \left(\lambda_1 \left[I - \mu_i \right]^2 - \lambda_2 \left[I - \mu_o \right]^2 \right) \exp\left(-\kappa \nabla \cdot \left[\frac{\nabla \varphi}{|\varphi|} \right] \right) (1 - \xi) \|\nabla \varphi\| .\end{aligned} \tag{2.34}$$

The part of equation (2.34) inside curly braces is Sarti's term taken from [72] and its first term is standard mean curvature flow weighted by an edge stopping coefficient. It serves to regularize the curve where the data is sparse and propagation can be further reduced by the presence of edges. The advection term drives the front towards image edges from a pre-computed edge diffusion field of the Gaussian filtered image.

Sarti et al [72] proposed an edge flow advection term combined with the simple mean curvature flow LS snake:

$$\dot{\varphi} = \alpha \phi div\left(\frac{\nabla \varphi}{|\varphi|} \right) + \beta (\nabla \phi) \cdot (\nabla \varphi) \tag{2.35}$$

Where ϕ is given by equation (2.26) repeated below for convenience.

$$\phi = \frac{1}{1 + |\nabla G_\sigma I|} \tag{2.36}$$

The second term in equation (2.35) is evaluated from a pre-computed edgemap of the image and behaves similarly to Gradient Vector Flow. This algorithm was able to segment objects with partial boundaries and was applied to a cardiac ultrasound dataset [72]. The derivatives of φ can be approximated using a simple approach to entropy satisfying upwind schemes [74] as shown in equation (2.37) for two dimensions

$$\nabla T = \left[\max^2\left(D_{ij}^{-x}T,0\right) + \min^2\left(D_{ij}^{+x}T,0\right) + \max^2\left(D_{ij}^{-y}T,0\right) + \min^2\left(D_{ij}^{+y}T,0\right) \right]^{1/2} \qquad (2.37)$$

where $D_{ij}^{-n}T$ is a backwards and $D_{ij}^{+n}T$ represents forward difference operator on T .

Sarti's second term in equation (2.35) was updated using upwind differencing schemes (e.g. as in equation (2.37)) to maintain numerical stability in the solution. The snake algorithm (equation (2.34)) will be refered as MS Sarti Collision Detection (MSSCD) in the remainder of this thesis.

This snake utilises some terms in the PZLS and GVF snake whilst retaining the elegance of boundary parameterisation in 2D and 3D. A LS model based on the snake equation from Sarti et al 2002 [72] uses the edge flow diffusion properties of the GVF snake. Sarti [72] segmented datasets with signal dropout using the properties of the mean curvature and edge flow terms. In this thesis a new term was added to Sarti's evolution equation to incorporate region growing based on local deviations from the interior and exterior regions using the image part of the MS functional. In this implementation the MS force is heavily penalised by curvature and inter-snake collision detection to reduce inter-chamber leakage. This is shown in equation (2.37) where Sarti's geometric model for boundary completion is enclosed in curly braces. The main weakness of Sarti's term is the presence of many edges at various strengths as is often found in sonography. Edge flow by advection is heavily dependent on the quality of the edgemap and so may fail to propagate the front towards the edges sufficiently to overcome the mean curvature flow. Without the presence of a clear edgemap from the data Sarti's snake fails to propagate appreciably towards the desired boundary (refer to Figure 29). The proposed term aims to provide some expansion or contraction forces dependent on the local tissue type in the absence of a strong edge field, e.g. when the front is in homogeneous regions. Unlike in Lassige's algorithm [12] this force can propagate the front in either direction according to the boundaries and so would be less prone to overshoot. To overcome this restriction I implemented the bidirectional MS term in conjunction with Sarti's algorithm which yielded a closer segmentation to the expert's delineation. The MS term minimises its energy by arriving at a segmentation that separates the background from the foreground and has strongest contribution in homogeneous

regions. As with the PZLS model the foreground was estimated from a small circle/sphere placed inside the chamber prior to evolution and the background was assumed to be the remainder of the dataset. Since the appropriate λ_1 and λ_2 could potentially vary significantly between datasets, the MS term was normalised to 1 by dividing by the maximum number of grey levels in the images to reduce the dependence on these coefficients. This same precaution is implemented in the PZLS snake.

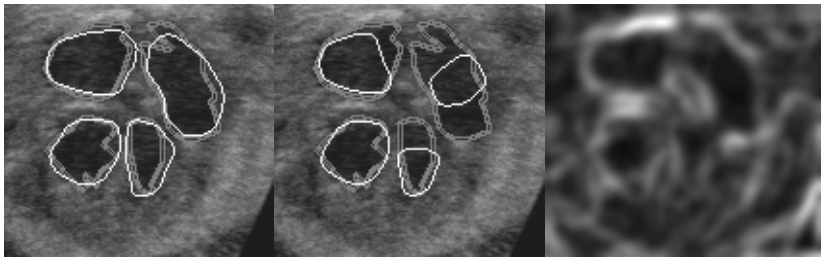


Figure 29 Segmentation by the algorithm proposed in this paper (left) and segmentation by Sarti's algorithm (middle). The white contours are automatically generated and grey denotes manual tracings. Atria appear at the top of the image and ventricles at the bottom with LV on the *right hand side.* The right image shows the edgemap.

Unlike in several applications of LSs where there multiple small interacting fronts that merge to form the segmented curve [75],[76] the foetal heart snakes are defined on separate Signed Distance Maps (SDMs). The basic idea is that there is an array of 4 snake objects (classes in C++ terminology). Each snake is given access to the image to be segmented. Since a LS is essentially an image superimposed on top of the image to be segmented it can share the same pixel pointwise correspondences. In this manner all 4 level sets are defined at the same spatial resolution as each other and the main image. As the snakes evolve when superimposed they can overlap but by examining any particular voxel address different snakes can be tested for occupancy. This is the basis of the implementation of the collision detection term in (2.34).

This is similar in principle to the work done by Vese in 2002 [76] where multiphase level sets can segment simultaneously objects of different grey levels on the same image; but the notation used in this thesis is much more simplified. In their case each phase (or organ) had a distinct average intensity for each region say as grey, white matter and cerebrospinal fluid in a Magnetic Resonance Image scan. In the application of the foetal heart each chamber snake sees a very similar foreground and background level, so the work in [76] may not be as successful in segmenting distinct chambers since it would merge different level sets as they share the same phases (chamber and background).

2.8.1.3 Shape-based level set

Another original contribution that I made to the in the foetal heart literature [70] is given by

$$\begin{aligned}\dot{\varphi} &= \alpha\nabla\cdot\left[\frac{\nabla\varphi}{|\varphi|}\right]\|\nabla\varphi\| \\ &+\lambda\left(\left[I-\mu_i\right]^2-\left[I-\mu_o\right]^2\right)\exp\left(-\kappa\nabla\cdot\left[\frac{\nabla\varphi}{|\varphi|}\right]\right)\|\nabla\varphi\| \\ &+\gamma\left\{\varphi_{sp}-\varphi\right\}\|\nabla\varphi\|.\end{aligned} \tag{2.38}$$

In this equation the first term is mean curvature flow and the second term incorporates the same MS driving force as used in Dindoyal 2005 [73] with the exponential curvature dependence. The last term incorporates the vector difference between the signed distance transforms of the shape prior and the current LS.

The image from which the template was created consisted of binary filled tracings of foetal hearts in long axial four chamber view. Overall 26 2D images were taken across a mixture of cardiac phases from 3 datasets due to lack of a large library of foetal cardiac data. The resampled pixel size varied from 0.3 to 1.46mm across the collection of images. As more datasets are acquired and manually traced the template becomes more representative of the population. Although different cardiac phases were used to construct the prior, care was taken to ensure that all four cardiac chambers were of similar size to ease the registration task. To align the images I use a multi-resolution method similar to Tsai et al [7] where the binary images were registered using the following cost function in an affine sense

$$E_{align}=\frac{\int_{\Omega}\left(I_T-I_F\right)^2 d\Omega}{\int_{\Omega}\left(I_T+I_F\right)^2 d\Omega} \tag{2.39}$$

where Ω is the image domain, T and F denote the transformed and fixed images respectively. The denominator penalises the transformed image from shrinking to zero. Differential evolution was used as the global optimisation method [8] unlike in Tsai [7] whose registration method is based on local gradient descent.

Multi-resolution rigid registration to the first image in the set was performed on all four cardiac chambers simultaneously to maintain their relative positions in the image. This step may be modified in future to allow individual chambers to be aligned as a refinement of the registration process. The results of the registration can be seen in Figure 30. Since the idea was not to find local corresponding points

but to align the shapes globally, rigid registration was sufficient for this task. The rigid registration error is evaluated in [126].

Figure 30 Stages through the registration process illustrated by superposition of manually segmented images. The *first* image shows preregistration. The *second* and *third* images show the intermediate registration at low resolution, the *fourth* at intermediate resolution. The *final* registered image is shown on the *far right* at the native resolution. Atria are shown above ventricles in this orientation with the *left* side of the image corresponding to the left part of the heart and similarly for *right* hand side.

Each of the registered binary images was transformed into a Signed Distance Map of the chamber edges of magnitude 1 with negative values inside the shape and positive elsewhere. From each of the SDMs the mean was computed using equation (2.40).

$$\mu = \frac{1}{n}\sum_{i=1}^{n} u_i \tag{2.40}$$

A binary image was formed from the mean of the manually segmented images of the foetal heart. This binary image would be the shape template that is registered to unseen images in order to segment them. From the template, SDMs would be created (as shown in Figure 31) so that each chamber of the heart is isolated in memory space to the rest of the compartments (see also [70] and Figure 31).

The main advantage of this snake over the previous two is that it constrains the propagating front to a predefined shape in areas of ambiguity such as missing chamber walls or several spurious edges. From Figure 31 the chambers in the mean template are distinct. Unlike MSSCD this algorithm does not require the presence of a collision detection term to be present since the relative boundary positions of the individual snakes are constrained by the shape prior and so any inter-chamber leakage will be minimal. This snake algorithm will be referred to as Template Initialized MS + Shape Prior term (TIMS+SP) for the rest of the thesis. The main driving forces in this snake are the shape prior and MS terms.

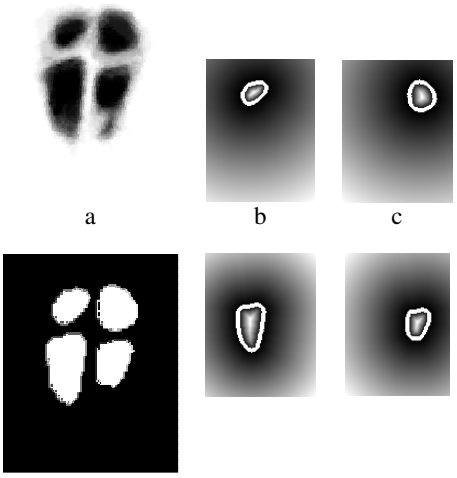


Figure 31 Registered manually segmented images illustrated as a superposition (a). Binary image of the mean of registered manual data (d). In this orientation the ventricles appear *below* the atria. Individual chamber shape prior SDMs for each snake are displayed as distance transform images with overlaid zero levels as *white* contours. The SDMs are shown in the following order: left atrium (b), right atrium (c), left ventricle (e) and right ventricle (f).

Table 3 shows the components of each level set snake implemented in this thesis. The TIMS snake has a shape prior and since there is no change in parameterisation of the snake surface between different level set snake types, EPCA and MSSCD automatically have the ability to use shape priors as well.

Table 3 LS snakes with properties listed

	EPCA	MSSCD	TIMS
Edge term	yes	yes	no
Region term	no	yes	yes
Collision Detection	no	yes	no
Shape Prior	optional	optional	yes
Initialisation	seed	seed	template
Main driving force	constant balloon	MS	MS, SP

2.9 Summary of my contribution to level sets

- Unification of Sarti and MS snake in MSSCD
- Exponential curvature penalisation in MSSCD
- Collision detection to reconstruct missing inbetween chamber boundaries in MSSCD also enables simultaneous region detection of the same tissue type without inter-region boundaries present
- Shape prior incorporated within existing level set snakes that did not feature a shape prior (EPCA, MSSCD)

- Comparison of existing LS foetal heart segmentation in literature with MSSCD and TIMS both with and without SP

2.10 Snake Overlap Correction

In the presence of partially resolved thin walls such as in a septal defect a snake may leak into the next chamber. This is illustrated in Figure 32B. The example in Figure 32 shows the snakes in the left chambers leaking into the right side of the heart more predominantly than the right snakes leaking into the left. If the snakes are rastered in rank order of volume (Figure 32E) then the meeting point of left and right snakes is closer to the septum than if rastered in order of volume (Figure 32D).

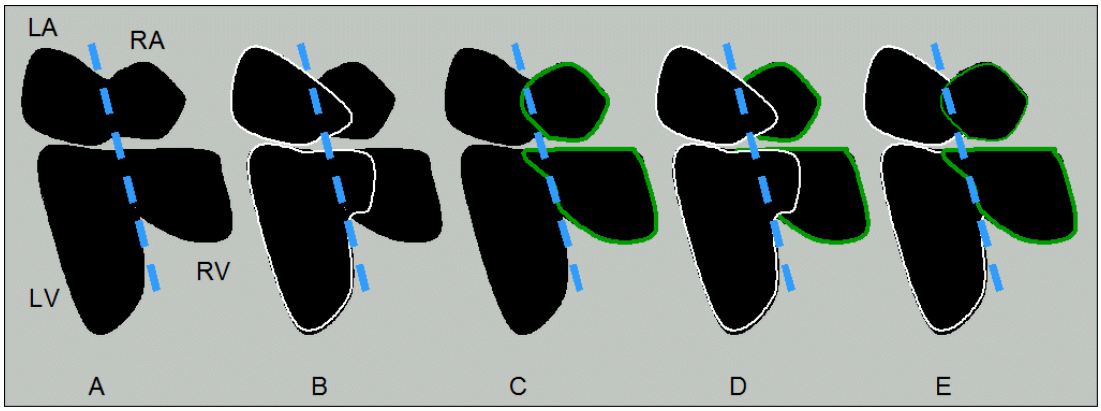


Figure 32 Illustration of snake overlap correction. In this figure there are four snakes (one per chamber). *A* shows the chambers before segmentation. The *dashed line* illustrates the assumed position of the septum. *B* shows the snakes in the left part of the heart leaking into the right. *C* shows the snakes in the right side of the heart leaking into the left. *D* shows left snakes superimposed on right. *E* shows right snakes superimposed on left.

This very simple correction that I designed was applied to all snakes in this thesis to create reconstruction of the septal wall when it is poorly resolved and not sufficient to stop a snake from expanding. It is an ad hoc assumption that the chambers on either side of the septal wall are closer in volume than with the chambers on either side of the valve plane. An example dataset where the snake overlap correction is turned off and then on is shown in Figure 33. Note this correction is very minor for the MSSCD algorithm since it already incorporates collision detection within the evolution equation. In other snakes this correction is more effective.

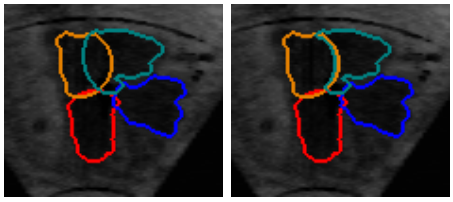


Figure 33 Before and after snake overlap correction algorithm is applied to the atria. Ventricles are below atria with LV on *left hand side* of image.

2.11 Contour Tracking

Although no tracking algorithms are presented for the main algorithms in this thesis the theory here offers some framework for the future work of the snakes in segmenting the foetal heart.

2.11.1 Segmentation Using Tracking

A motion model is necessary to track the heart in situations where the cardiac cycle cannot be followed accurately due to aliasing in the time domain (inter-frame acquisition time) and motion-blur artefacts due to the finite time required to acquire a single frame. The heart walls undergo several complex non rigid motions during the course of a single cardiac cycle and investigation into the deformations is an active area of research [77] for understanding of cardiac biomechanics and diagnosis of pathologies involving the myocardium.

Since the shape of the ventricles vary widely throughout the cardiac cycle an ordinary shape model for a single cardiac phase that can identify a particular shape in the sequence may struggle to satisfactorily segment the remaining frames. Two approaches can be used to overcome this: Either to include temporal sequences into the training set or to impose some biomechanical model to constrain the shape as it segments the chamber during the cardiac cycle. The first case is time consuming for a clinician to manually segment the entire cardiac cycle and can lead to more subjective bias than segmentation of a single frame alone unless the training set includes statistical agreement between the manual tracings of several clinicians. Automatic tracking of subsequent frames is more difficult to design algorithmically since the heart undergoes complex deformations which are dependent on various factors such as the natural variation between individuals, the health of the myocardium and valves. However, there are similarities that can be used to generate a generic motion model from the data. This approach makes key assumptions about the response time of the muscles. Deformable models that incorporate resistance to tension and bending forces make such assumptions that the motion is smooth and if successful tracking is achieved, these models can be used to interpolate the missing cardiac time points between the acquired frames.

The "aperture problem" in feature tracking applications relates to the size of the kernel window. If a straight line feature longer than the kernel (aperture) moves perpendicular to the line then the movement can be tracked, however if this feature was to move along the axis of the line then the feature will appear stationary (refer

to Figure 34 for a visual explanation). Thus useful search windows for tracking edges in images with fairly homogeneous regions will be along the edge normals.

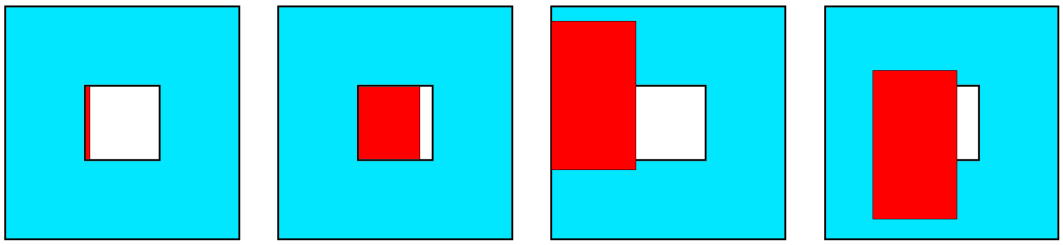


Figure 34 The aperture problem. Line behind aperture (*a*). Line appears to move to the right in direction normal to the line (*b*). Line is now in front of aperture (*c*). Line is in front of aperture and moved to the right and downwards (*d*) and gives the illusion as seen in (*b*). This shows that parallel motion to the line can not be seen when line is behind the aperture and the straight line feature is longer than the kernel window [78].

This section describes some of the approaches of tracking the ventricles in the heart. Most of them operate in a Lagrangian frame of reference where markers are explicitly tracked via a motion model. These algorithms usually assume that the temporal data acquisition is such that the deformation between frames is small. Tracking algorithms rely on temporal as well as spatial differentials and are hence even more sensitive to high frequency noise, thus some kind of filtering is often performed on the images before attempting to track the motion of the structures. The deformable models above can be applied to ventricle tracking but in the case of temporally under-sampled data or a noisy acquisition process, additional filters can be incorporated as a preprocessing stage to ensure that the tracking process is more robust. To ensure that contour points are constrained to move smoothly between frames for temporally undersampled data the contour can be rigid body deformed and registered from the previously segmented frame to the current one. The cost function used involved matching the vertex normal profiles in both the source and target image (Figure 35). The rigid registration then serves to initialise the snake in the current frame. This was implemented in my publication [50] for GVF snakes in 2D applied to the foetal heart.

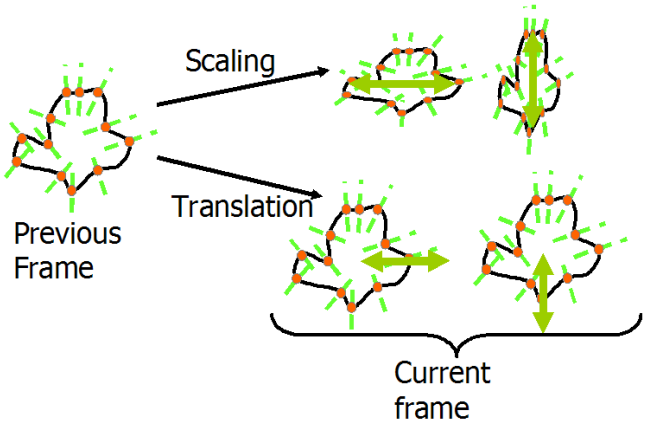


Figure 35 Scaling and translating the contour in the previous frame. The dashed lines indicate the normal profile to the vertices.

2.12 Assessing Segmentation of Medical Images

Segmentation in medical images is the delineation of the structures of interest (e.g. anatomical, physiological) within the image from the background or other structures. The segmentation operation is required to obtain a compact geometrical representation of the structure (either in the form of a curve, surface, or a 2D/3D region) so that its area/volume, shape properties can be measured or produce a rendered model for visualisation. It is possible to achieve adequate segmentation both with and without manual intervention but many algorithms require some combination of both. Usually manual aid is used to provide an initialisation to identify a rough location of the structure of interest. In a volumetric dataset with many slices complete manual segmentation is time consuming, tedious and operator dependent. Automatic, repeatable and reliable segmentation of structures are thus welcomed in the medical community since it enables faster diagnosis of pathologies and therefore a higher throughput of patients.

2.12.1 Performance and Accuracy of a Segmentation Algorithm

There are various factors that determine whether automated segmentation algorithm is the preferred choice to manual methods. The algorithm should require little manual intervention so that many images can be segmented without operator bias. The algorithm should be fast enough to generate acceptable results within reasonable time limits. A reasonable time for clinical use could be a fraction of the time required to perform manual tracing of the structures. For medical images segmented offline

this is not a critical requirement since these images are often obtained for research and not diagnostic purposes. Accuracy of the segmentation is of vital importance since clinical measurements will be made on the basis of these results. The algorithm should be able to match or deemed acceptable to manual tracings by several expert cardiologists. This can be verified by computing distance errors (as defined below) between the computer and manual curves. Validation can be determined by evaluation of the segmentation algorithm on imaged objects of known volumes; and only then errors in volume measurements can be quantified. A known synthetic heart corrupted in a controlled manner can be used to assess the precision and accuracy of the algorithm in various scenarios such as under different levels of noise.

2.12.2 Error Metrics

The root mean square (rms) and mean errors are distance metrics for computing the differences between two contours or surfaces. Figure 36 shows the projected point of one contour onto another.

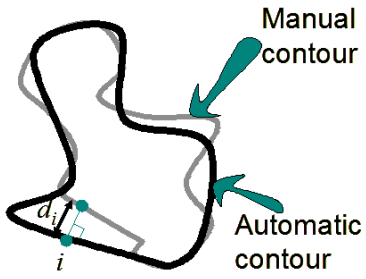


Figure 36 Projected point from one contour onto another.

The rms error is then given by (2.41)

$$E_{rms} = \sqrt{1/n \sum_{i=1}^{n} d_i^2} \tag{2.41}$$

whilst the mean error is given by (2.42)

$$E_{rms} = 1/n \sum_{i=1}^{n} d_i \tag{2.42}$$

To compute either distance metric over 4 cardiac chambers the following equation was used whilst normalising over chamber volume

$$\mathrm{E}_{\mathrm{Heart}} = \frac{1}{\mathrm{N}_{\mathrm{Heart}}} \sum_{\mathrm{chambers,c}} \bar{\mathrm{E}}_{\mathrm{c}} \mathrm{N}_{\mathrm{c}} \tag{2.43}$$

where N_{c} is the number of points on the chamber surface, $\mathrm{N}_{\mathrm{Heart}}$ is the number of points on the entire heart surface and E represents the distance error metric used. The Tanimoto Coefficient (TC) was used to compare overlapping regions

$$TC = \frac{\sum_{voxels,i} MIN\left(A_i, B_i\right)}{\sum_{voxels,i} MAX\left(A_i, B_i\right)} \tag{2.44}$$

A method of comparing two multi region segmentations by overlap methods can be computed by using the Generalised Tannnimoto Coefficient (GTC) defined in [79] using equation (2.45)

$$GTC = \frac{\sum_{labels,l} \alpha_l \sum_{voxels,i} MIN\left(A_{li}, B_{li}\right)}{\sum_{labels,l} \alpha_l \sum_{voxels,i} MAX\left(A_{li}, B_{li}\right)} \tag{2.45}$$

where A is a region in the first image and B in the second. The label at a voxel *i* is either 0 or 1. α_l represents the inverse mean volume of the two labels l i.e $1/V_l$.

This weighting reduces the contribution of relatively larger chambers in the overall overlap. I found the GTC to be more penalising than the averaged distance errors over the entire heart. It seems that the GTC value is lower than the standard Tanimoto Coefficient for a single chamber.

2.13 Summary

This chapter outlined the segmentation problem faced with basic image processing methods and outlined the need for various deformable model approaches. Each of the snakes were introduced in the order of progression of this project with the exception of tracking which is placed last because it leads up to the future work. Finally metrics of comparing the segmentation algorithms quantitatively in this thesis were stated. In the next chapter I will outline the literature review that is relevant to foetal echocardiography segmentation.

3 LITERATURE REVIEW

3.1 Acquisition of volumetric datasets of the foetal heart

There have been a fair number of 3D foetal heart acquisition publications in the literature. These are summarised below. The first few entries use conventional 2D B-mode transducers swept through a volume to acquire 3D data. Towards the end direct volumetric acquisition becomes more accessible to foetal cardiologists and gating becomes less necessary.

Using a 2D US transducer orientated such that it swept through an angle of (60-90)° in three orthogonal planes, Kuo H-C et al 1992 [80] were able to generate a 3D foetal heart dataset. Their imaging software then allowed visualisation of anatomical structures from three orthogonal planes simultaneously. Combined viewing of these three planes greatly increases examination of structures such as septal defects and cardiac lesions.

In the study by Nelson TR and co-workers (1996) [15] a method of gating foetal heart acquisition was undertaken offline. An Acuson 128/XP, Mountain View, CA scanner was used together with an electromagnetic (EM) position sensor transducer with six degrees of freedom (x, y, z, θ, ϕ, Ω) to acquire sonographic data at a rate 30 frames per second. A maximum of 30s was required per dataset. The volume scan data are acquired using a slow sweeping motion through the foetal heart and major vessels. They drew a line manually on the valve in one of the images of the foetal heart at the mid-chamber for use in gating the motion. Pixel intensity along this line is then plotted against time for each image in the slice. The information presented from this plot is similar to M-mode cardiac motion curves and is a direct reflection of the mechanical action of the valves. The periodic behaviour can be analysed using a temporal Fourier transform to reveal the periodic frequency components of the valve motion. The fundamental frequency from this analysis is used to gate the motion. The images were then assigned their proper location in the cardiac cycle from the fundamental frequency. The sliced data was re-projected into a common coordinate system. Due to fluctuations in rate of freehand sweeping missing planes in the volume data can result. These gaps are filled in using a nearest neighbour interpolation technique. A median filter is then used to reduce some of the noise but caused minimal degradation to the image data. They proposed that the quality of

the results are due to improved signal to noise ratio and reduced uncorrelated speckle from the signal compounding with the EM positioning system and freehand probe. They suggested that foetal echocardiography shows potential to provide a more unified and comprehensive assessment of cardiac anatomy when compared to conventional 2D sonography techniques. A limitation of the offline gating method used here requires that the foetal heart is regular during the acquisition period and ectopic beats from the volume data were excluded in the analysis. Optimal foetal 3D data for visualisation is dependent on (a) the overall image quality of the 2D slices, (b) the plane of acquisition, (c) the overall foetal orientation and the amount of acoustic shadowing that can result from overlaying structures such as calcified foetal ribs, and (d) the amount of foetal movement. They claimed a need to resolve the small details within the foetal heart that require new scanning technology capable of sub-millimetre resolution.

A feasibility study of foetal echocardiography with 3D US was undertaken by Zosmer et al 1996 [17] in 54 cases where the gestation age varied between 17 and 37 weeks. In this investigation the transducer was mechanically swept through 45° to scan volumetrically the heart in 3D. The acquired volume had a maximum of 65,536 US lines each with 512 samples. The maximum number of lines per B-mode scan was 1024 but typically between 80 to 350, and an upper limit to the number of scans per volume set between 50-250. In each patient the quality of the 3D sonographic volumes is assessed qualitatively by the appearance of the 4CV. 85.2% of the cases showed clear structure within this view and are used for further investigation. Restless foetal movement during scanning prevented the acquisition of clear 3D volumes for 4 patients. It turns out that the best visibility of anatomical areas of interest within the heart was observed at a gestation age of between 22 and 27 weeks. After this period the amniotic fluid volume is relatively decreased and the heart moved farther away from the transducer. At this stage in the pregnancy the quality of the images are degraded and the number of cardiac views on reformatted planes of the dataset that could be clearly identified are reduced. Some limitations of the current 3D acquisition technology are due to an un-gated mechanical transducer that is used to sweep over an area of interest. This produces a series of nearly parallel sonographic sections over 3-4 seconds (~6 foetal cardiac cycles) while the foetus was resting. Owing to intracardiac motion, the individual sections within the volume did not fit together as well as they would if a stationary organ was scanned instead. They proposed that a 3D US system based on multiple transducers could overcome this limitation. Severe problems in 3D echocardiography of the foetus were due to foetal breathing and gross body movements. They recommended that the patient should be asked to perform a breath hold during the acquisition to reduce artefacts due to maternal breathing.

Deng J and co-workers 1996 [19] developed a 3D acquisition method using a freehand EM position transducer to capture 2D foetal echocardiograms at various positions and orientations. By simultaneously using real-time 2D directed M-mode, cardiac motion curves could be recorded with the 3D heart volumes. The transducer is placed at the required position on the heart for 1 second before moving it to the next position. The inter-slice separation is about 1mm and 15-25 sections are acquired per time sequence. The set up required the foetus to remain stationary for up to 30 seconds. The M-mode curves were analysed offline to gate 20-40 frames/dataset; these curves are used to manually reconstruct phased volumes for 3D processing. Some advantages of using the M-mode gating are recordings of clear cardiac phase motion by utilising both the systolic and diasystolic information of both the cardiac valves and ventricular walls. A limitation at the time (1996) only allowed the acquisition rate to capture 8-12fps during the simultaneous M-mode and structural US scanning, and this does not meet the desired frame rate of 25fps (i.e. video rate) for visualising smooth playback of a dynamic 3D foetal heart. The M-mode beam will generate periodic waveforms for each sampled slice. However, each imaging plane will contain different frequency spectra due to the rate at which the structures inside the slice are moving. Since the entire heart does not beat in unison (e.g. the valves move faster than the walls), the number of physiologically meaningful time points per cardiac cycle will be greater for rapidly moving structures and less for the slower ones. This can cause ambiguity when identifying the corresponding phase of the cardiac cycle for a particular slice. Thus matching adjacent short-axial slices so that they beat co-operatively requires manual clinical intervention. At the myocardium apices the M-mode sampling tends to pick up movement arising from structures surrounding the heart leading to ambiguity in identifying the correct cardiac motion curve.

A study into assessing the volume measurement of the foetal heart was made by Chang F-M et al 1997 [8]. In the study a 5MHz trans-abdominal Voluson transducer was used and scanning was performed when the foetus was at rest with little foetal respiratory movement. The 4CV is identified by the high resolution 2D US and the Voluson sector transducer is then used to scan the foetal heart by sweeping through an angle of 40° within 3 seconds to generate a 3D dynamic volume. At this acquisition speed the frame rate is higher than the foetal heart rate and so it is possible to obtain a clear outline of the heart without blurred contours. They found that in the cases where this rate was not adequate, a faster scanning mode needs to be selected to acquire images at higher temporal resolution sacrificing spatial quality. However, they found that setting is not sufficient to compromise the volume assessment since the borders were still clear. The determination of the volume from

2D US was estimated from assuming an ellipsoid model of the heart and measurements of three orthogonal axis multiplied by a scaling constant. The results show that 3D US has a better reproducibility than 2D US in heart volume assessment. They also found that the scaling constant leads to an overestimate from the 2D US views and was modified from the 3D volume measurements.

Sklansky MS et al 1997 [21] investigated the interpretation of 3D foetal echocardiography compared with 2D imaging. In both cases a 2D freehand EM tracked transducer was used to acquire the data, a single slice for the 2D imaging and multiple slices for the 3D acquisition. Steady sweeps lasting between 13 and 37 seconds were carried out in either the sagittal or transverse orientations with respect to the prenatal heart. To perform the offline gating a plot of the pixel intensity as a function of time was made from a square region of interest encompassing the heart. A temporal Fourier Transform is calculated of each plot and from this the fundamental frequency of the cardiac motion and average heart rate could be determined. The 2D images were then mapped onto the appropriate part of the cardiac cycle and the result was a series of 12-16 volumes per heartbeat. Missing volume data are filled in using a nearest neighbour technique and a median filter is used to reduce some of the noise. Post processing per subject required around 20-30 minutes. The results suggest that 3D foetal echocardiography was superior in providing consistent anatomical structural information when compared to the 2D imaging especially in the cases where more than one view is required to examine pathology.

Levental et al 1998 [81] presented a paper that compared 2D with 3D foetal echocardiography. Their 2D imaging was performed with a conventional Acuson 128XP with a 3.5 or 5MHz transducer. Carefully orientation of the probe during acquisition allowed the traditional 4CV to be clearly visible. For 3D US they used a Combison 530 (Kretztechnik, Zipf, Austria) with a 5MHz mechanically driven annular-array volume transducer with adjustable focal length. The 3D acquisition was un-gated and swept the volume in 4s. In the 3D cases they scanned during times of minimal foetal movement in the transverse and longitudinal planes of the heart. To minimise acquisition time and the amount of data collected, each dataset is acquired using normal speed (medium line density) with the smallest possible ROI box. The foetal heart views that were acquired were not significantly affected by the gestation age. They found that the un-gated 3D data is inferior to the un-gated 2D data in the ability to provide adequate quality of the basic cardiac views of important anatomical structures and in some cases did not present sufficient information for diagnosis. It was apparent from the un-gated 3D US that false negatives could be obtained such as a missing atrioventricular lesion when compared to the 2D US images. Since the

volumetric scan is performed using a mechanical device to move the transducer, this acquisition method produced several shadowing artefacts from the limbs or ribs and drop out artefacts of the membranous ventricular septum. As a result, repeated repositioning of the 3D transducer assembly is required to acquire satisfactory visualisation of the structures. If a foetal heart beats at 140bpm then for an acquisition lasting 4s, over nine heartbeats will take place during this time. The resulting 3D un-gated data will consist of a series of temporally mismatched 2D planer images, which will be combined to form a 3D dynamic volume. Their opinion of 3D echocardiography is that it shows potential in visualisation of specific cardiac structures from multiple views, but when analysed un-gated it presents the clinician with more ambiguity in distinguishing normal from abnormal cases compared with conventional 2D US.

A review paper on 3D US imaging by Nelson TR and Pretorius DH in 1998 [81] presents work from a selection of authors who acquired either gated or un-gated 3D foetal echocardiography and some of their visualisation methods. They also review their own progress from their 1995 and 1996 papers.

Sklansky MS et al 1999 [14] describe their use of real-time pyramidal volume foetal echocardiography. The 3D echocardiography system (Volumetrics Medical Imaging, Durham, NC) was equipped with a transducer with a 3.5 MHz central frequency and an acquisition rate of 20 volumes per second. They found the 50ms inter-volume time satisfactory to capture the dynamics of the foetal heart, even when played back at a lower rate. Since the entire heart was acquired in a single volume there was no need to gate the motion however, random motion artefacts were present when there was sufficient maternal or foetal movement. Acquiring 3D+t data without moving the probe is less time consuming and less dependent on operator experience. In some foetuses it was possible to see structures as small as the pulmonary veins. It was found that the spatial resolution of the images limited the number of abnormal cardiac structures that could be identified reliably when compared to conventional 2D US, but future improvements in scanner design may provide adequate image quality for diagnostic purposes.

Deng et al 2000 [20] demonstrated the simultaneous use of two US scanners for providing online foetal cardiac gating to acquire 3D+t heart data. Since two probes may interfere when driven at similar frequencies, a study was made using various paired transducers on a phantom. The factors affecting interference were believed to be the separation distance and angles between the two probe heads. As a qualitative metric, they classified the strength of the interference into three visual cases: slight, moderate and severe. By wrapping rubber gloves around the probe heads they

concluded that the interference is not caused by direct electrical conduction of the two probe heads with the phantom. A layer of metal foil wrapped around the transducers greatly attenuated part of the interference pattern suggesting that some EM coupling took place between the two probes. They proposed that the remaining majority of the interference is caused by acoustic cross talk and by increasing the probe separation distance to at least 55mm significantly decreased the interference. They found that the separation distance between the umbilical cord and foetal heart (the regions of interest for the simultaneous acquisition) is not linearly related to the gestation age. Angling of the probes towards each other seemed to intensify the noise and contrary for the diverging case. The second scenario is useful if the position of the regions of interest is such that the two probes were closer than 55mm. They concluded that the interfering signal is inversely proportional to the distance between the two sources and receivers. A greater overlap of the sector planes from both transducers would increase the interference and contrary otherwise. Two 1D transducer array probes were used one to provide 3D structural information of the foetal heart as it was swept over the volume and the other to detect the Umbilical Arterial Doppler Waveforms (UADW). Only 15-30 seconds are required to acquire the data for each patient. This set up with two transducers operating simultaneously enabled offline motion gating of the foetal heart dataset with the UADW. The UADW are pseudo ECG signals identified of careful filtering of the Doppler signals from the umbilical cord, see [20] for more details. Experimentally they observed that the optimum level of interference was due to the pairing of specific transducers in the study and only the cases with severe noise (3 out of 9) could not be 3D+t processed for reconstruction. They also comment on their previous method [19] where they used an M-mode probe to acquire waveforms as well as structural information. However this method is unsuccessful in the clarity of the waveforms produced compared to the UADW in this paper since it is difficult to hold a bulky rotating probe steady when recording the M-mode time curves.

Guerra FA et al 2000 [82] proposed a new approach to studying the 3D foetal heart using conventional freehand ultrasound transducers. Their approach takes conventional volume acquisition and feeds it directly into commercially available freehand ultrasound 3D-View software. Although un-gated, the acquired images are formatted in a simultaneous multi-planar and B-mode display and provided sufficient information required to show some foetal heart pathologies. However, the current version of the commercial software does not allow manual measurements to be made on the data and serves more as a visualisation aid than a diagnostic tool.

Deng and co-workers [16] used online motion-gated foetal echocardiography to acquire 3D+t data to generate a dynamic 3D visualisation of the heart. From the

optimal conditions discovered in Deng 2000 [20], interference from two ultrasound probes operating simultaneously on two scanners could be minimised. On the first scanner a Doppler probe is used to locate the umbilical artery and thus provide the UADW signals required to gate the second scanner. The second probe is operated in B-mode and positioned so that it could capture volumetric data as its imaging plane swept through the entire foetal heart. Three types of B-mode probe were used in conjunction with the Doppler transducer: freehand EM tracked or a parallel mechanic or rotational mechanic. The conversion of the UADW into pseudo ECG signals [83] allowed successful gating of the datasets when the triggering efficiency was above 50%. The triggering efficiency is defined as the percentage of valid triggering signals out of a total number of consecutive UADW signals that were necessary to complete a 4D scan. A valid triggering signal is one that could trigger the 4D system to record serial 2D images at one cardiac section and move the probe to the next section. For the cases in which this figure dropped below 50%, longer acquisition time was required and considerable foetal movement may degrade with the quality of the acquired data. In the freehand transducer case, they found that the magnetic tracking was interrupted by the EM noise generated by the Doppler scanner and it was suggested to replace the mechanical Doppler probe with a phased array one. Extensive shadowing from bony structures (the mechanically driven probes were more subject to this effect due to their relatively limited freedom of movement compared to freehand transducer), and random motion artefacts due to foetal movement were the main causes for reduction in the quality of the data acquisition. Nevertheless the gating precision was able to allow fine structure such as chordea tendinae and the trabecular muscles to be seen in the reconstruction.

In [83] Deng et al constructed an electronic conversion box to convert noisy UADW signals from lambs to mimic the R-wave of the ECG in the form of Cardiac Cycle Triggering Signals (CCTS). These online gating signals are used to trigger another US system, which then collected 3D sheep foetal echocardiograms. Using sheep foetal twins to test the conversion performance, a transit-time ultrasonic flow-meter probe was placed around the main arteries to provide a gold standard for measurement of the cardiac cycle. The conversion efficiency is assessed by the percentage of successful UADW to ECG conversions. In the cases where acoustic interference are not considered severe the conversion efficiency was in the range of (83-100)% but fell to (0-71)% for cases where the interference was extreme. They define conversion efficiency as the percentage of 20 CCTS that could be obtained from the UADWs. Since there is a delay between the heartbeat and the corresponding UADW they implanted probes to obtain the Pulmonar Arterial Flow Waveforms (PAFW). The time lag was defined as the time delay between the onsets of a PAFW and a corresponding CCTS. Triggering accuracy was defined as:

Triggering accuracy (%) = 100×[(mean time lag)- (99% confidence level)]÷(mean time lag)

They measured the triggering accuracy to be between (90-96)% in the cases where interference and strength of the UADW signals were not a problem and (0-90)% otherwise. Although the study was performed on sheep foetuses, they expect that the lower heart rate of a human foetus will mean that each waveform will contain more cardiac phase information. They proposed that integration of these two scanning systems should contribute to reducing the acoustic interference problems.

Deng performed volumetric acquisition of foetal cardiac data using the Phillips Live 3D matrix transducer [3]. In this paper Deng outlines some of the desired acquisition parameters that are necessary to obtain informative 3D foetal echocardiographic datasets. No temporal tracking of the probe is necessary since the probe is moved at a comparatively slow speed compared to the heart. The volumetric acquisition is capable of about video frame rates which are enough for sampling most of the cardiac phases at least once per imaging volume cardiac cycle. This paper was introduced to inform clinicians of the technology and terminology involved in making 3D acquisitions of moving organs.

In 2003 Gonçalves scanned 69 foetuses using a new cardiac gating technique Spatio Temporal Image Correlation (STIC) [28] (refer to section 2.1.2.3 for more details about STIC). The 4D volumetric ultrasound scanning is done on foetuses aged between 17 and 35 weeks using a VOLUSON 730 Expert system (General Electric Medical Systems, Kretztechnik, Zipf, Austria). The acquisitions took between 7 and 13 seconds during absence of foetal movement where possible. The probe was moved to collect either sagittal or transverse sweeps. Multiplanar reslicing is then done offline to visualise various cardiac structures including atrial septal defect. The STIC correction correlates the various 2D images based on the repetition of the heartbeat and produces a single visualisation volume. This process of rearranging the images into a series of volumes, each representing a snapshot of the foetal heart during a single beat takes under 10 seconds.

3.2 Speckle reduction

There are many image processing methods in the literature to reduce speckle. The brief selection presented here deal with cardiac ultrasound data in ways that are not covered by section 2.3 in the theory above. Speckle was not the dominant noise artefact in the datasets under investigation in this thesis and the length of this section reflects this.

Although not strictly speckle reduction Tomasi and Manduchi [84] produced a method that simultaneously filters images in both domain and range. The combined filtering effect is termed Bilateral filtering. The domain is the pixel location and the range refers to the intensity at each pixel. Both the domain and range component of the filter are Gaussian based with increased blurring with larger variance size. The domain filtering takes into account the Euclidean distance between the kernel centre and the position in the kernel. In range filtering the photometric closeness of the intensity values within the centre of the kernel and of the intensity at a position in the image are used to smooth the greyscale differences in the image without sacrificing edge detail. Pixels that are very different in intensity from the central pixel have lower weighting even if they are phyisically closer to the centre of the kernel.

A method of automatically fusing two 3D ultrasound transthoracic acquisitions taken at different orientations was proposed by Ye et al [85] using a rotational probe. In this paper the LV is automatically detected by a feature asymmetry algorithm based on phase congruency [86]. The feature that is detected by this algorithm is phase. A step intensity edge can be decomposed into Fourier components. Phase congruency is at a maximum when all Fourier components at the point of an edge transition are exactly in phase. Since the Fourier components represent different scales of signals enforcing this maximum tends to be more robust than simple intensity based edge detection. The algorithm has been shown to be intensity amplitude independent when used in a 2D+t approach [86]. This intensity invariant method computes a measure for feature asymmetry using log-Gabor wavelet filters. The features are orientation and edge polarity dependent. The feature detection produces several spurious candidate points. To pick those that correspond to the endocardium numerous threshold values such as grouping of candidate feature points that are in a proximity to within $[c_{min}, c_{max}]$ are chosen without justification for the parameters in the algorithm. In 2D a cubic B-spline curve is used to fit the sparse LV features that remain after eliminating the outliers. To fuse image features from different 3D views, knowledge of the spatial position of each feature is computed from the probe tracking during the acquisition. Instead of combining the different views using the maximum or average of the pixels a weighted feature compounding method was used. The compounding weight is defined in the paper by

$$\alpha = \left|\cos(\theta)\right| \tag{3.1}$$

where θ is the angle between the beam axis and the 3D normal of a given feature point. Less emphasis is placed on feature points located further off-axis to the main beam. The calculation of the 3D normals of each feature point is computed by

principal component analysis of a k-neighbourhood of features around this point. The resulting vector will be orthogonal to the k-neighbourhood set. Finally an ICP based surface fitting method (initialised by a manually placed ellipsoid) is then applied to the reconstructed feature volume. Local fitting involved deformating the surface to the image features by minimising the following energy function

$$d\left(\underline{M},\underline{N}\right)^2 = \beta\left(\left\|\underline{M}\,\underline{N}\right\|^2 + \left\|\underline{n}_M - \underline{n}_N\right\|^2\right) \tag{3.2}$$

where $\underline{M}$, $\underline{N}$ are points in the image feature dataset and the model surface respectively; $\underline{n}_M$, $\underline{n}_N$ are the 3D normals of the two points and β is a weighting coefficient. They found that the reconstruction from multiple acoustic windows is able to recover from shadowed regions and yields better estimation of ejection fractions that are closer to those obtained from MRI scans compared to US data acquired from a single window.

Speckle Reducing Anisotropic Diffusion (SRAD) is a method developed by Yu and Acton 2002 [87] to suppress speckle noise for both ultrasound and radar images. The algorithm preserves edges and smoothes noise in homogeneous regions. The SRAD PDE is of the form

$$\frac{\partial I}{\partial t} = div\left[c(q)\nabla I(t)\right] \tag{3.3}$$

where I is the intensity, and $c(q)$ is the instantaneous coefficient of variation. The amount of smoothing at iteration t applied by SRAD is related to the speckle scale function which is given by

$$q_0(t) = \frac{\sqrt{\mathrm{var}\left[z(t)\right]}}{\overline{z(t)}} \tag{3.4}$$

where $\mathrm{var}\left[z(t)\right]$ and $\overline{z(t)}$ are the intensity variance and mean over a homogeneous region. In this paper they approximate this smoothing as an exponential decay in (3.5) which was deduced by experimental means

$$q_0(t) \approx q_0 \exp(-t) \tag{3.5}$$

The edge preserving features are sensitive to edge magnitude and Laplacian operators. Although they claim that the nonlinear diffusion proceeds along edges, there is still some broadening of edges present after filtering. They tested the filtering method on an ultrasound scan of a carotid artery.

Montagnat et al [7] implemented 4D anisotropic filtering in cylindrical coordinates of a cardiac dataset acquired by a rotating probe. They found that 4D anisotropic diffusion yielded more continuous boundaries compared to results from filtering in 3D. The diffusion function is given by

$$g(x,a) = 1 - e^{-\left(\frac{3.315}{\left[\frac{x}{a}\right]^4}\right)} \quad (3.6)$$

where a is a gradient threshold for a boundary point and x is the intensity of the US image. Diffusion is higher in the direction parallel to the boundary than orthogonal to it. This function leads to an enhancement of boundaries. Edge detection of boundaries is performed by a Deriche filter in cylindrical geometry. Cylindrical coordinates allowed reliable edge detection and a high reduction of speckle near the rotation axis of the probe where the data is better sampled.

3.3 Attenuation correction

The methods presented here are suitable for raw data from the acquisition. The scanners available tend to apply their own TGC algorithms which can not be easily deconvolved from the data without prior knowledge of the methodology used by the manufacturers. Many of these algorithms are simplistic and tend to be position dependent (and independent of tissue type) so that different tissues in the same compensating range will be corrected by the same amount [88]. In addition the data available have been resampled from the A-lines by the scanner to produce a regular Cartesian image grid of pixel/voxel intensities at 8-bit greyscale. This output format is fixed by the manufacturer and so it is difficult to post process the original echogenic signal in its current form to correct for ultrasound attenuation. Some methods below correct for attenuation in video images not raw data but are done without the intervention of TGC. I did not incorporate any of these approaches since raw data from the scanner was not available to the group nor was the TGC algorithm or details of how to remove it.

Hughes and Duck [89] developed an attenuation compensation algorithm that models the received echogenic signal in terms of integration of local attenuation and backscatter effects. A linear relationship between backscatter and attenuation is assumed to hold for all tissues on a particular scan line in the US images. The method is able to provide some visual recovery to the areas of attenuation although some shadowed regions are filled with noise after amplification. Little improvement could be achieved for shadowing due to ribs but diagnostic images could still be visualised after the algorithm was applied. It may provide additional boundaries for surface rendering algorithms although they will have higher noise content than those not in shadow.

Xiao et al 2002 [88] developed a method of correcting the low frequency image intensity inhomogenity in ultrasound video images using a combination of maximum

a posteriori (MAP) and Markov random field (MRF) methods by means of a multiplicative model of the image formation. A Gaussian probability distribution models the image intensities. They justify this by claiming the algorithm applied in a scale space framework tends towards a Gaussian at coarse scales. Initially the distortion field is assumed to be zero everywhere in the image. The pixel intensities are assumed to be statistically independent. The prior probability of the distortion field is modelled as a Gaussian with zero mean. This was estimated using the MAP and MRF methods whilst simultaneously labelling the image regions based on the corrected local intensity statistics. Experiments on synthetic and real data show that the inhomogeneous field can be successfully removed and segmented provided the class information and distribution is known. The method does not correct for shadowing or frequency dependent attenuation.

3.4 Segmentation of LV in adult heart using low level techniques

Han et al [90] presented a system that extracts the LV endocardium boundary from transesophageal echocardiography images. This method uses a combination of low level image processing operations combined with and high level knowledge of the LV anatomy to automatically segment the endocardium. The images are low pass filtered and reduced to one ninth of the original size. Statistical analysis of 140 images from 7 patients produced an automatic threshold level that corresponded to the pixels outside the LV region. The threshold level retains the bottom 74% of the intensity values of the image histogram for further analysis. Following the thresholding stage overlapping testing circles are used to determine the extent of the LV by considering whether or not the circle fits inside the dark regions entirely. The largest connected labelled region is chosen as the LV region. The labelled LV region is then compared to the original image and then subdivided into a 3x3 grid to determine the optimum local threshold for each sub-region. This step allows better connectivity of the muscle regions. Morphological dilation of the LV region constrained by the muscle regions is then applied to the images. After dilation a contour linking procedure is applied in polar coordinates to interpolate and correct for incomplete and noise corrupted boundaries. This step used the assumption that the chamber is roughly circular in short axis view, that the papillary muscles appear roughly 120° apart with respect to the centre of the LV. However this algorithm has problems segmenting datasets satisfactorily when special cases arise such as when the viewpoint does not show the LV in the centre of the image or the papillary muscles in the assumed orientation.

A fuzzy connected approach to segmentation of both the LV endocardium and epicardium was presented by Furuie et al 1997 [91]. This algorithm uses a manually predefined region within the LV to estimate Gaussian statistics of connected voxel intensities within the LV. All connected voxels to this region are tagged as part of the region. A cost function that incorporates texture and strength of Sobel edges is applied to all the voxels to determine which voxels had a higher cost than a threshold and should therefore be dropped from this region. The second stage of the algorithm involves minimising the differences in shape between the epi and endocardium. This is incorporated into the cost function under the assumption that the blood pool will be more clearly defined than the muscle in the MR images. The algorithm is iterated until the cost is greater than the threshold for all voxels. Results indicate that the algorithm is able to achieve segmentation within 6% error in phantom studies.

In 1999 Friedland et al managed to segment 2D LV in long axis echocardiograms [92]. The assumption made is that the LV was roughly elliptical and that an ellipse could be fitted to the LV using a radial search of a threshold edge from the centre of the chamber combined with a 1D Hough Transform along the radial direction. A simulated annealing algorithm is run to optimise edge detection along the radii, radial value smoothness, maximisation of cavity volume and temporal continuity with previous frame. They claim that this method is the first relaxation algorithm for edge detection in echocardiograms. A B-Spline is used to join the connecting points on the endocardium boundary at the end of the optimisation.

Sanchez-Ortiz et al [93] developed a fuzzy clustering based algorithm that interpreted the global intensity distribution of the image with prior knowledge of the LV as a constraint in cases where low signal to noise ratio and image contrast are present. The prior model also reduces outliers detected by phase-based edge detection. The method iteratively uses anisotropic diffusion and fuzzy c-means clustering in scale space to reduce the problematic effect of intensity fluctuations in the US images. The clustering algorithm produces a fuzzy partition in which every pixel is classified with a probability of membership to each of the regions of the image according to its attribute vectors in feature space. A cost function is iterated to approach the local minimum by grouping pixels with similar attributes. This is performed across scale space until convergence is reached below some predefined threshold. The LV cavity was identified by the centre of mass from the darkest cluster closest to the centre of the image. Spurious points are filtered out based on their position relative to two thresholds D_{min} and D_{max} which are related to the width of the ventricle. These parameters were estimated from the length and width of the cluster over the cardiac cycle. The cluster analysis provides a starting point for a surface fitting method to refine the segmentation of 3D rotation echocardiography

images. This surface is deformed locally using B-splines and globally using affine transformations to track the LV surface over the cardiac cycle. The segmentation method is automated and although preliminary it can produce similar tracking results from manual surface tracing on good data.

In 2006 Lynch et al [94] used a k-means classifier to segment the LV cavity and LV wall in 3D MRI data after smoothing the images using an edge preserving filter. The clustering parameters used for the k-means classifier were mean and variance of the pixels. They assumed that the LV cavity was circular in shape on each short axial slice. By overlapping regions in adjacent slices and imposing tolerance on the greyscale values they could ensure continuity between slices in order to segment the LV volume. A spline fitted to the segmented slices enabled a mesh to be formed. The method is fully automated and was compared against a basic fast marching level set edge based method. Their proposed clustering algorithm outperformed the level set approach in terms of distance errors to manual tracings.

3.5 Amorphous deformable models

The literature review of deformable models without shape in this section covers the adult heart only. Deformable models used to segment the foetal heart can be found in section 3.8). Note some of the algorithms in this section (3.5) could be described as tracking algorithms, but are presented here since they are limited to direct temporal regularising of corresponding points between adjacent frames. In section 3.7 the tracking algorithms involve matching or prediction of the next frame and so are considered separate from section 3.5.

Richens et al [13] produced a semi-automated method to segment cardiac MR images based on a modified g-snake. The model is applied to 2D short axis slices through the left ventricle. In addition to the bending and stretching components of the internal energy in the classical snake model, a term that is controlled by the area enclosed by the contour allows the curve to shrink and expand. Successive Gaussian filtering is used to pre-process the images to reduce noise. A term coupling contours from neighbouring frames was added to the snake equation:

$$E_{couple}^{i,j} = \int_0^1 \alpha \left\| \underline{x}^i - \underline{x}^j \right\|^2 ds \qquad (3.7)$$

where i, j are adjacent frames, α is a weighting coefficient, $\underline{x}$ is the position of the snaxel. This allowed stronger edges from neighbouring frames to guide the contour in frames from the diastolic phase where the lower blood velocity in ventricular filling often produces a weaker signal at the endocardial border.

A dynamic deformable surface model was created by McInerney et al [45]. The model used a Newtonian Physics based approach to deform the surface and incorporated a balloon model to expand the surface to meet non local edges. An edge image was created using a 3D Canny-Deriche edge detector before application of the algorithm to track a cardiac sequence of CT images. The surface model allows regularisation of missing data from the edge images. Tracking of non rigid cardiac motion was also performed using the model by initialisation of the surface in the next frame with the currently segmented one.

In 1995 Ranganath [41] applied 2D snakes to cardiac MR images to segment the LV endocardium with the aims of minimum user interaction required and to overcome the poor temporal resolution of the data. Salient image features were found by Sobel edge detection of the Gaussian filtered image. The user initialised the seed contour in the middle slice within the dataset at diastole. To propagate the contour between slices the intensity profiles normal to each snaxel of the snake in the previous slice are compared to the normal profiles of the snake in the current slice. The best match is found for a particular snaxel in the current snake by computing the correlation coefficient of the current and corresponding profile in previous slice. Temporal constraints are present by forcing the contour to shrink on migration from diastole to systole by greatly increasing the weighting coefficients of the internal energy. The image is then thresholded (based on the expression $T=m+a\sigma$ where T is the threshold, a is a factor (chosen experimentally), m is the mean, and σ is the standard deviation of the interior pixel intensities within the contour. Canny filtering of the threshold image is used to provide the image energy for the snake so that the contour is maintained at the edge of the blood pool during the cardiac cycle. Good segmentation performance was achieved - errors in segmented areas were less than 7% when compared to manual tracings from one expert clinician. However, they found that local snaxels would behave unpredictably in response to papillary muscles since no prior knowledge of these structures are incorporated in the snake model.

Kucera and co-workers 1997 [11] produced a snake algorithm that makes use of 2D region based internal forces and inter frame constraints to both spatially and temporally regularise the contours in the echocardiographic sequence. The temporal information guided the segmentation of the ventricles when heavy signal dropout or high speckle content was present in a frame but easier to segment in adjacent ones. The snake model is solved by finite elements in space and finite difference in time. They found that 20 snaxels were sufficient to describe the shape of the ventricular chamber. To initialise the contour in the image they placed a circle manually close to the endocardium so that the papillary muscles are enclosed. The region-based external forces consisted of local statistics of image intensity extracted from the

image along normal profiles between each snaxel. An error function measures how well the local intensities (*f*) in the region *R* fit a statistical model of mean intensity values $\langle f \rangle_R$

$$Error(R) = \int_R \left\| f(x, y) - \langle f \rangle_R \right\|^2 dxdy \tag{3.8}$$

The region-based intensity forces are calculated by a decision whether or not the snaxel should be moved along the inside or outside normal directions based on the degree of dissimilarity of the two regions R_1 and R_2. To estimate this dissimilarity the classical Ward distance was used:

$$Dist(R_1, R_2) = Error(R_1 \cup R_2) - Error(R_1) - Error(R_2) \tag{3.9}$$

They found that their algorithm underestimated the required area when compared to manual outlining of contours in a training set of end-diastole and end-systole images. A discrepancy in the radial distance of 3.02mm between the manual and computed contours existed however; no fundamental cause for this value was given. They corrected for this bias by incorporating an outward displacement into the algorithm to match the clinical findings in the training data. After the correction mean error between the average manual outline and the computed contours was 0.4mm.

A 3D GVF snake was applied to segment cardiac MR data by Santarelli et al [42]. The algorithm pre-computed a non-linear anisotropic filtered edgemap to preserve edges in the MRI dataset. The GVF snake followed the edgemap to avoid local minima and detect both the endo and epicardium for the LV. The segmentation provided volume measurements and estimation of myocardial mass.

Bosnjak compared the volume obtained from segmentation of the LV in echocardiography by a LS snake with fitting of a deformable superquadric [51]. The main idea was to reduce the sequence of volumes to a small group of parameters to characterise the movement of the LV. The superquadric model is able to approximate the volumes found by the LS snake to within an error of 2%. The LS snake used edge detection to stop the propagating front at the endocardium and seemed to only segment the right ventricle reliably when low speckle noise was present.

Sánchez et al 2000 [48] used an active contour algorithm to track the boundaries of the endocardium in echocardiac images. Their approach combined the GVF and balloon terms into a new one called the hybrid force. They use the hybrid term when its individual components fail to detect any prominent image forces. By exploiting the GVF term a snake can be attracted to a predefined edge map from far away. Unlike in the conventional balloon model by [43], inflation or deflation of the contour

does not have to be specified a priori to running the algorithm in every frame. The balloon in this algorithm can initialise the snake when the GVF field is not defined in areas such as in the first frame. To instantiate the contour in subsequent frames the contour is derived from the previous image as an approximation, together with a reduced number of iterations of the GVF field. To avoid clustering a rigidity term penalises points on the basis of their separation from nearest neighbours being greater or less than the average spacing of snaxels in the snake. To ensure that the snake is attracted to the endocardium they make use of region information by exploiting the fact that the myocardium appears brighter in intensity than the ventricular chambers. Their algorithm obtains enclosed areas in agreement with that of hand drawn contours by an expert cardiologist to within 90%.

A 3D LS snake was implemented by Corsi et al [52] to segment left ventricular volume in 3D echocardiography. The snake has no inflation force and so requires a few manual points to define a rough contour placed close to the endocardium in five slices of the dataset in short axis view. The snake interpolates these contours in 3D and is able to segment the volume as well as compensate for gaps in the data. The algorithm is able to segment pathological cases such as LV aneurysms. Only the zero order level set was used so topological changes could not take place. Bland-Altman analysis was performed to assess the accuracy of the algorithm's volume estimation on in vitro balloon phantoms. Clinical application of the snake to measure ejection fraction and LV volumes was estimated for 20 patient cases.

Paragios [95] developed a GVF-like diffusion term in the level set snake to segment single slice cardiac MRI LV data. Region information of the various tissue types was determined from the image histogram by modelling the distributions by a combination of Gaussian functions. Expectation maximisation was used to model the region intensity properties of the epicardium and myocardium. The approach uses coupled active contours searching along normal vectors to the curve to maintain an equidistant spacing between the epi and endocardium snakes. The spacing was determined from a predetermined acceptable minimum and maximum thickness limits of the endocardium and was found to be dependent on the slice position through the heart. The parameterisation for the snake was determined by experimental means.

Recently Sarti published work on ultrasound segmentation using Rayleigh intensity priors [62]. In his paper he models the pixels both inside and outside of the LS as Rayleigh distributions and fits them to the data using expectation maximisation. He applies the algorithm to the adult echocardiography and foetal whole body segmentation. Manual tracings are used as his gold standard. As a comparison he

uses the Chan and Vese algorithm [60] and claims that the two approaches obtain similar results but the Rayleigh fitting allows better distance errors compared to manual segmentation.

In 2007 Bernard developed a segmentation using level sets applied to the radio frequency ultrasound signal [96]. They use the Chan and Vese method [76] for multiphase images and apply Generalised Gaussian modelling of the intensity distribution. The LS is able to separate the image signal into tissue and blood pool and show the resulting segmentation on 2 images. It appears that the images are not reconstructed for missing boundaries and so would be unsuitable for chamber blood volume estimation.

A 2D GVF snake was applied to segment MRI cardiac short axial data in 2007 by El Berbari and co-workers [97]. Before this step the image was filtered using a connected operator algorithm which is similar to region growing followed by an area closing method to cover the papillary muscles in a short axial slice of the LV. A Canny edge detector is then used to find the local edges from the filtered image and this was derived the GVF field for the GVF snake, which in turn is initialised close to the Canny edges. The method was evaluated on 29 images taken from the End Diastolic cardiac phase. Overlapping area between manual and automatic segmentation is used to determine the algorithm's accuracy as well as mean and max distances between the two contours.

3.6 Non-amorphous cardiac deformable models

This section covers the literature review for the adult heart only since there are very few shape aware deformable model entries in the foetal heart literature they will be treated separately in section 3.8.

Hill and Taylor [98] applied a flexible contour template of a left ventricle to segment single frames from 20 temporal sequences of 2D long-axis echocardiographic data. The model is constructed from 6 parameters that describe properties of the shape of the left ventricle boundary. In addition 4 transformation parameters specified in polar coordinates control how the model is fitted to the image. To instantiate the model the six control parameters are generated, and from these further points on the boundary can be defined by spline interpolation. At each point on the contour a 20-pixel normal profile is used to extract the local grey level values from the image. The cost function implemented favours solutions with strong edges of equal magnitude

that are close to the boundary as predicted by the model. To minimise the objective function two stochastic optimisations were employed and they found that application of a genetic algorithm produced closer matches to expert tracings of the boundaries than use of a simulated annealing approach.

Point Distribution Models were applied to adult transoesophageal echocardiography in 1994 by Parker, Hill, Taylor, Cootes and Graham [99]. The model is constructed by manually landmarking points on the ventricle and surrounding structures in the long axis view at both end diastole and end systole (from 33 patients), and for the short axis view only 63 images from 7 patients are used. To further improve the specificity of the model grey level profiles at each point are used to construct an Active Shape Model (ASM) (which was originally developed and applied to photographs in [57]). They found only 12 significant modes of variation for the long axis case and 6 in the short axis view. To automatically segment the images an instance of the mean shape was placed in the image and a multi-resolution ASM search is performed using a genetic algorithm. The pose of the current shape is modified to minimise a weighted sum of the distances between the current and new instances of the shape. The iterative refinement of the shape is adjusted by constraints represented by the PDM so that only solutions consistent with the training set are allowed. These two stages are then iterated until convergence is reached.

Hamarneh and Gustavsson 2000 [100], [101] created an algorithm that combined the allowed variations of shape within a training set to constrain a snake model in LV echocardiography. Discrete Cosine Transform (DCT) coefficients are used as shape descriptors to capture ventricular shape variation instead of landmark points placed on the contour as in the original ASM. DCT parameterises the shape in the frequency domain because it produces real coefficients and does not require explicit point correspondence on the contours. Principal components analysis of the coefficients can then identify the main modes of variation. The snake is initialised in the dataset and allowed to segment the LV. DCT analysis is then performed to project the contour into the same subspace as the principal components so that the statistical model could be applied and constrain the contour to the reduced modes within the training set. The evolution of the contour is subsequently iterated between the snake and model until convergence.

A simplex mesh was used by Montagnat et al 2000, [56] to segment 4D MR, SPECT and US cardiac images. A second order Newtonian evolution scheme deformed the mesh during the segmentation. The algorithm searches along surface normals for image edges and incorporated both temporal and spatial regularisation as part of the internal forces. The method can work either with or without *a priori* shape

knowledge of the dynamic ventricle. The prior model can be provided by either 3D segmentation of individual frames or by 4D segmentation of a reference image. With knowledge of *a priori* reference segmentation the vertex position would be compared to a stored LV shape and trajectories of the myocardium. When no prior knowledge used the temporal constraint attracts a vertex towards the mid point of temporally adjacent vertices. They found that temporal regularisation led to less accumulation of errors in calculation of ventricular volume during the cardiac cycle.

Song et al [102] developed an integrated approach to optimisation of a surface model applied to LV cardiac US data. The optimisation is performed in a Bayesian framework using the probability of pixel classes, feature vectors of pixel appearance and shape knowledge of the surface from a training set (86 manually segmented images to create meshes) to maximise the posterior probability. The pixel class probabilities were calculated from both training sequences and simulated ultrasound backscatter images. When the likelihood profile (pixel class) and the predicted class from the surface model are well matched, the surface model is assumed to be a good explanation of the images. In a set of 25 test studies both the epicardial and endocardial surfaces agreed well with the ground truth measurements.

Fully automated 3D Active Appearance Models (AAM) were constructed from 55 cardiac MR and 65 US images by Mitchell and co-workers 2002 [103]. In the MR case data were acquired volumetrically. However from the ultrasound images only 2D echocardiograms were collected and so time was converted into the appropriate units using a speed factor to construct the third dimension in the model. The training set is manually segmented from the images to generate 3D shapes which were automatically landmarked with aid of representing the ventricles in a normalised cylindrical coordinate system with the primary axis lined up with the long axis of the heart. To generate the landmark points a fixed number of slicing planes are placed perpendicular to the long axis of the ventricles in the defined coordinate system and points populate the surface of the chamber evenly distributed around the ventricle on the slicing planes. 3D Delaunay triangulation is used to generate a mesh of the ventricle. Alignment of each training shape is accomplished by standard Procrustes analysis and least squares minimisation. Modelling of the volume appearance voxel-wise correspondence is achieved by warping one image to another using piecewise affine or thin-plate splines. To match the 3DAAM to the image data required minimisation of the rms intensity difference between image and model instance by modifying the affine transformation, global intensity parameters and appearance coefficients. Their analysis shows that the segmentation resulting from the model performs better in the MR images than in echocardiograms. As with many segmentation models based on training data, a limited number of pathological

datasets are included in the training set that so the model may have difficulties segmenting abnormal data.

Montagnat et al updated their segmentation of 4D echocardiographic work from 2000 [7]. The dataset was pre-processed by a 4D anisotropic diffusion filter to create homogeneous regions in the image whilst not corrupting edges. A manually placed ellipsoid initialises the mesh in the first frame of a high resolution single frame dataset, which is then guided to the endocardium by a region-based data attraction force applied along surface normals. The normal intensity profile at each point on the surface is scanned from the inside to the outside to find a series of voxels that are above a predetermined threshold. A voxel with the correct edge polarity in this neighbourhood whose gradient was above a given threshold is then located to identify the boundary. Some prior knowledge of the ventricular shape from the segmented higher quality dataset serves as a template to initialise the mesh in a sparse dataset. The shape is fitted to the dataset globally and afterwards allowed to refine locally. To track the ventricle in a cardiac sequence both global and local deformations were applied to the surface of the previous frame.

A 3D cardiac LV active shape model was created by van Assen et al [104]. This method applies a fuzzy inference model to determine candidate points from the images to be segmented. The iterative closest point algorithm [105] registered manual contours drawn by experts. To fit the data using the model, the contours are resampled and triangulated into a 3D mesh. Where the mesh points intersect short axial image planes a patch centred on this point was sampled normal to the boundary and the tissue edge is found using fuzzy c-means classification. If a pixel can not be classified to a certain tolerance it is not considered for fuzzy inference. In this case neighbouring patches are used to update the model point. The main results show that fuzzy inference is able to better delineate structures like papillary muscles and trabeculations in the images compared to using features from edge detection. There is also a substantial reduction in blood pool volume error by using the new fuzzy inference method but epicardial volume errors remain comparable to those obtained edge detection.

Fritscher et al [106] recently developed a method of automatically creating shape models of presegmented images using rigid registration by mutual information. The presegmented images are turned into signed distance maps and registered. Rigid body registration only provides rough correspondence and was sufficient for building a model to initialise the segmentation. Leventon's algorithm [107] is used in the shape modelling process. Principal Components Analysis is performed on the registered signed distance maps to create a mean shape. Shape templates are

created of both the whole endocardium and the chambers. A multiresolution mutual information approach guided the alignment of the mean shape to the target image. Shape variation within the training set is not used in fitting the model to the image or for segmentation. After the template fitting processes segmentation of the images is achieved by a LS snake with curvature, constant advection and edgeflow terms. Since the shape template is used the initial front is very close to the desired boundaries and the balloon force can be set very low to reduce unwanted leakage. The first frame is used as a template for the entire cardiac cycle. Satisfactory segmentation results were produced by the shape template.

MR cardiac LV images were segmented using a LS snake by Lynch and co-workers in 2006 [108]. They incorporated standard fast marching to detect the blood-endocardium boundary and from this projected a LS for the epicardium which has less contrast against surrounding structures. The external LS is coupled to the first by a prior model based on probability maps of binary manually segmented images. This prior then steered an edge advection term towards the boundary. They achieve a correlation coefficient of 0.86 of the blood pools when compared with manual segmentation.

In Wang's paper [109] an atlas template was created from a single healthy volunteer from MRI data and "validated" by an expert. The volumetric template mesh is constructed using Delaunay triangulation and then affine registered to tagged MRI cardiac data followed by thin plate splines to fit the model locally. The mesh is deformed to track the heart during the cardiac cycle using Lagrangian dynamics, Newton's 2nd law of motion and Hooke's law applied locally. Strain analysis of the left ventricle was then computed.

Szilágyi et al [110] managed to segment LV echocardiographic images using a 4D AAM from 35 patients. ECG traces were acquired simultaneously with the US images. The US training images were grouped according to ECG events such as clustering around the QRS complex. This allowed the datasets to be catalogued according to normal and abnormal cardiac cycles. The algorithm was tested on 4 patients to analyse the recognition rate of the patient specific QRS beat of the cardiac cycle.

A whole heart explicit mesh atlas developed by Phillips Research Labs for cardiac CT data was applied to static MRI cardiac volume data in 2007 [111]. Peters and co-workers transformed the atlas to fit the atlas to 42 MRI cardiac datasets using the generalised hough transform. Edge detection along profiles from the centres of the triangles are used in the cost function. The edge detection is normalised to the

gradient magnitude to reduce the image intensity inhomogeneities. Some linear rescaling of the image histogram is needed to bring the images into the calibrated intensity framework of a reference image used during the training phase. To further refine the fit the individual cardiac regions are piecewise affine deformed independently. They measure a mean distance error of 0.76mm.

In 2007 Zheng et al [112] as part of Siemens Corporate Research group constructed a 3D explicit mesh of the all four chambers of the adult heart derived from manually segmented data. The mesh is open at the valve plane to allow for open valves. The mesh was applied to CT cardiac data and aligned to the data using Marginal Space Learning with steerable filters which cuts down the optimisation space dramatically. This method works by allowing the different degrees of freedom to be evaluated independently with different number of optimisation steps for each one. The steerable filters sample a few points from the volume under a special pattern. To evaluate the steerable filters under a specific orientation the sampling pattern is orientated without the need to rotate the volume. This combination of methods allows full 4 chamber fitting to the data within 8 seconds on a 3GHz PC. Once the mesh is fitted to the data a thin plate spline is used to refine it locally together with the shape constraint. The study was done on 137 patients.

Recently Zhuang and co-workers [113] managed to segment 19 3D MRI cardiac images using an image atlas of the whole heart and a combination of affine and non-rigid registrations. This approach does not use a deformable model but the image atlas is allowed to deform. In the paper the mean of 18 of the manually segmented datasets was used to segment the remaining one in a leave one out approach. After global affine registration incorporating normalised mutual information the multiple chambers are allowed to move independently in a local affine sense. Penalisation for overlap and a threshold distance between chambers is incorporated into the cost function and this is checked on every stage of the evolution. After local affine the segmentation is refined using non rigid registration.

3.7 Segmentation of the heart using tracking algorithms

All the entries in this section with the exception of Dindoyal 2003 [50] refer to the adult heart.

Chalana et al 1996 [40] used a multiple active contour model to segment both the epicardium and endocardium of the left ventricle. The model is an active surface that

evolves within the dataset to minimise its energy by means of a cost function. In contrast to the original snake model the external energy term incorporates temporal information which monotonically constrains the contour from the previous frame of the sequence into contraction or dilation depending on the phase within the cardiac cycle. The surface is initialised by a rough drawing of the epicardial border on the end-diastole image by the user. The contour initialised by the user is then attracted to the edge map created by Canny's edge detector applied in scale space. The weighting coefficients for each energy term in the snake were determined empirically for the intended application of echocardiography. They claim that their segmentation results of the epicardium and endocardium are relatively insensitive to the initial curve and are comparable to inter-observer variability.

In 1998 Jacob and co-workers [114] devised a tracking algorithm for a LV planar shape based on learned training set. The manually presegmented images are modelled as B-Splines and PCA analysis is made on the coefficients to extract 95% of the variation. Any shape in the training set can be expressed as an initial template and a multiple of the shape matrix. In the training set the contours are not registered so the dominant modes are translation and scaling and a combination of these two transformations. To apply the tracking to an unseen long axial echocardiac sequence following manual segmentation of the first frame, they used a Kalman filter to iteratively update the tracking algorithm using a prediction-update strategy. The prediction is based on the model applied to the shape from the previous frame and is then updated using the measurement process. Feature detection is done along contour normals and the maximum response is assumed to be the desired feature.

Declerck et al [115] segmented a gated SPECT cardiac dataset following image resampling of the LV in 3D-planispheric coordinates (a combination of both spherical and cylindrical coordinates). In this geometry the heart appears as a thick bowl and allows easier identification of cardiac boundaries to surrounding structures. To detect both endocardium and epicardium a Canny-Deriche recursive filter is used to find the edges. Starting from the centre of the chamber radial intensity profiles are evaluated on the image and the endocardium is detected as the first edge along the line and epicardium is the second. From this dense set of feature points the algorithm performs inter-frame matches using their non-rigid registration technique which incorporates template matching to a truncated ellipsoid by an adaptation of the Iterative Closest Point algorithm. The results of the matches are used to estimate a 4D polar transformation to best fit the list of points. This transformation allows for LV centripetal contraction, rotation around the long axis, and elevation. Temporal continuity and periodicity of the transformation is enforced as part of the constraints.

The resulting matches form smooth periodic trajectories of the points throughout the cardiac cycle and can be used to analyse wall motion in detail.

Malassiotis et al [116] proposed a temporal learning filter to refine the segmentation of the epicardium by an active contour model. The snake is initialised by a Hough transform ellipse in the first frame and then left to evolve and capture the coarse segmentation of the entire sequence. Principal component analysis is then used on these segmented contours to analyse the principal modes of the wall motion. A statistical observation model was built based on the coarse segmentation curves. It was updated by either a snake whose energy was explicitly constrained along the reduced set of eigenvectors or a Kalman filter that incorporated the principal eigenvectors in the state vector. The statistical model that incorporated knowledge of heart dynamics was found to produce temporally coherent segmentation compared to results using the active contour model alone.

A deformable surface model that incorporated *a priori* knowledge of heart motion was applied to segmentation of 3D+t cardiac LV US data by Gerard et al [117]. The 4D motion model was acquired from 12 healthy volunteers using a CSPAMM (Complementary Spatial Modulation of Magnetisation) protocol in MRI which makes it possible to tag slices and follow them in a complete cardiac cycle. Transformations imposed from the model included a contraction/dilation and rotation around the barycentre in the short-axis plane. A homogeneous affine deformation is modelled in the long-axis plane. Cylinderical coordinates are used to express the transformation of a vertex whilst maintaining the topology and one-to-one correspondence over the cardiac cycle. To initialise the model tracking procedure a simplex mesh is manually placed and orientated in the end-diastole frame of the cardiac sequence. The mesh is deformed to fit the dataset by searching for edges along the normals to the surface patches. In some cases manual correction to the surface was required. The segmentation is propagated through the cardiac cycle using the motion model from MRI data. At all times user control of the surface is permitted. It was found that the endocardium could be detected in cases where severe echo dropout occurred on the LV free wall. Wall motion indices can be generated automatically by the algorithm.

Boukerroui et al [118] presented a method of estimating velocities in ultrasound images using a block matching approach. The method relies on optical flow using Singh's algorithm where the velocity estimation is defined by the shift in two image features/regions at different times by maximisation of a similarity measure. The similarity measure is defined by maximum likelihood estimation of the image where the speckle multiplicative noise model is modelled by a Rayleigh probability distribution. To account for attenuation in ultrasound images the blocks are

normalised to the same mean and variance. A square search window is created around the corresponding pixel in the next frame of the image sequence. Velocity estimates are obtained for the candidate pixels in the neighbourhood from Singh's algorithm. Gaussian weights are assigned to the velocities of the surrounding pixels since they are likely to be related to the pixel in the neighbourhood centre. Modifications were made to the probability mass function in Singh's algorithm to improve its velocity estimation in low signal to noise conditions. After computation of the velocity field all frames were registered to the first frame and a registration error was calculated. The registration error is minimised for determination of the optimal parameter set in Singh's algorithm. Velocity estimates for the current were computed using intensity information from the preceding two frames and in the process propagates errors to subsequent images in the time sequence. The method appears to be satisfactory in tracking low frame rate data in the absence of a global motion model of deformation.

I have published work [50] based on segmentation by a 2D GVF snake that is initialised from tracking grey level intensity profiles along normals to a contour by rigid body deformations between frames during the cardiac cycle. This paper is reviewed in section 3.8 since it refers to the foetal heart.

In 2005 Cho et al developed a velocity-aided active contour model to overcome problems associated with flow-related signal loss, heterogeneity of the myocardium signal and papillary muscles in cardiac phase contrast 2D short axial MRI [119]. They use the orientation of the velocity vector field gradients to determine the direction of maximum change in local vector fields. This term is used in addition to the normal intensity edge detection on the snake to guide the evolution from frame to frame. Since this velocity vector term is derived from the motion of the phase contrast it reduces propagation of segmentation errors by placing the initial seed contour closer to the desired boundary in the current frame. The method was sensitive to velocity wrap-around artefacts in phase contrast MRI.

In an attempt to combine Level sets with a shape prior of the adult heart deformation in 2008 Lynch et al [120] published a tracking LS method that fitted an inverted single Gaussian mixture model to the cardiac cycle volume-time graph using the expectation maximisation algorithm. The Gaussian fit is used to constrain the fast marching LS segmentation as a shape prior term in the LS equation. Their 3D+t segmentation produced smooth results over the entire cardiac cycle and achieved segmentation regression value of 0.76 compared to manual segmentation areas. The epicardium is estimated as a thickness constraint on another LS outside the endocardium zero levels.

3.8 Segmentation of foetal cardiac data

There is relatively little published research on segmentation of foetal cardiac chambers compared to of the adult heart. This chapter presents a review of previous work on segmentation of foetal cardiac data.

Piccoli et al 1999 of the Navaux's group, Brazil [121], proposed an unsupervised self-organising map (SOM) neural network approach to 2D foetal echocardiographic segmentation. To approximate the probability density function of the underlying classes without the need of human intervention required training of the SOM network; i.e. by acquiring samples from random locations within the image. At each step the algorithm extracted a pattern from the set of samples and selected the nearest neuron in feature space as the "winner neuron". By updating the neuron's weights and those from neighbouring neurons during training enabled convergence to the extracted pattern. At convergence the neurons group themselves into regions of higher probability. K-means post processed the SOM clustering step after the training stage to group the pixels into k clusters according to the similarity of the neuron weights. The modified Hubert index provides a criterion to select k based on the Euclidean distance between the centres of the clusters. Experimental results show that classification with more than two clusters tends to better preserve the heart structures than segmentation of myocardium and blood pool.

Lassige et al. 2000 [12] presented a method of detecting the presence of a septal defect between two chambers of the foetal heart using a geometric snake based on the level set method and was attracted to endocardiac edges in the images. Gaussian filtering was used to reduce image noise before edge detection. Their algorithm can readily adapt to splitting and merging contour topologies and form corners. To initialise the model they manually placed several seed points on either side of the septum. The algorithm performed most reliably when a straight line that passed through the septal defect could join the centres of these two clusters of seed points. The algorithm was tested on a training set for different initialisation points on raw volumetric data acquired from a Hewlett Packard HPSONOS Ultrasound system, which contained 60 images acquired by rotational scan and represented cross-sectional slices of the heart at three-degree intervals about a central axis. The test results showed that the algorithm could measure the size of the septal defect with a maximum variation of 3.7 pixels. The algorithm failed to achieve reliable segmentation by frequently overshooting the cardiac walls since the stopping term only slowed the contours at edges; but given sufficient number of iterations it can grow beyond the boundaries. This was due to a constant advection term that controlled inflation of the contour. Absence of edges around the contour can cause

the snake to evolve into long straight regions that fail to grow at an appreciable rate due to the curvature term in the speed equation.

Fernades and co-workers of Navaux's group [122] updated their algorithm from 1999 to include in addition to the self-organising map neural network and k-means classification; a LEGION (Local Excitatory Global Inhibitory Oscillator Network) neural network for extraction of features within the image. The LEGION network is an array of Terman-Wang oscillators [123] with local coupling that allows phase synchronism between neighbouring oscillators with similar external inputs, and a function to achieve asynchronism between distant oscillators. A group of oscillators (or "legion") with similar inputs can be led by a leader oscillator in network. On the other hand isolated oscillators are due to noisy fragments and can not be characterised as leaders. The LEGION network generated a binary image as a result of the classification and the medical specialists used this to obtain measurements of the cardiac structures. The feature space required to satisfactorily classify and segment the images consisted of mean and variance of pixel intensities. They applied the LEGION network to the data so that it could use more than one frame and enabled a spatio-temporal segmentation of the input signal. The algorithm can not provide adequate segmentation of chamber areas in the case of strong corruption of the boundary between the LV and LA.

Navaux's group segmented the 2D US LV chamber widths as a function of time in the cardiac cycle [124]. In this paper median filtering reduced speckle to an acceptable level after they applied histogram equalisation to the images. The self organising maps approximate the probability density of the mean and variances in the image sequence. The map is post processed by the k-means clustering algorithm to identify groups of neurons with similar weights. They found that SOM followed by k-means better represent the dimensions of the cavities than with k-means classification alone. A least squares contour fitting technique is applied to the pixelated borders of the classified image to better approximate the endocardium.

Foetal cardiac tracking was performed using a coarse to fine wavelet approach by Hsieh et al [125] from a manually traced contour of the first frame in the sequence. The motion vectors are obtained by minimising the sum of square differences (SSD) of the locations of similar pixel intensities in adjacent images. In the model the 2D wavelet basis functions represent the motion vectors. The current image is warped to the previous image based on the estimates of the motion vectors and to update the SSD. This process of warping the image is done from coarse to fine scale with progression in refinement taking place once the SSDs falls below a given threshold. The method is able to generate dense motion vectors from the tracked contours to

enable correspondence for PCA analysis of the contour shape once tracking has been achieved. This study is to investigate Sudden Infant Death Syndrome (SIDS). Since the first frame has to be traced manually this method would not easily extend to three dimensions without automating the initialisation.

In my entry publication into the foetal heart literature (Dindoyal et al 2003 [50]) a GVF snake segments and tracks the cardiac walls by modelling both their rigid and non-rigid deformation. The assumptions made by the method are that luminance is dependent on the structures that are imaged (i.e. a structure will remain at the same luminance during the cardiac cycle) and that the TGC parameters remained constant for acquisition of each frame of the cardiac cycle. Although speckle will corrupt the boundary information it is assumed that the intensities of pixels inside the chamber will appear darker than those on the myocardium in all cardiac phases. Tracking of the myocardium from the first automatically segmented frame in the sequence is performed by recognising the intensity normal profile across the myocardium-blood boundary. This method of tracking is used to reduce sensitivity to drop out artefacts of the cardiac walls that were present in the dataset which would cause the snake to leak out of its respective chamber into the neighbouring ones. The first frame (end-diastole) is segmented by the GVF snake (once initial seed points were placed), since this cardiac phase seems to have minimum drop out artefacts. Simple Gaussian filtering was applied to the image to reduce speckle before application of the snake. The image profiles normal to the contour are stored for comparison with the snake in the next frame of the sequence. After initialisation of the this contour into the next frame it was globally scaled and translated until the edge profiles in this frame matched those in the previous in a sum of absolute differences sense. Once this optimum rigid body deformation is complete the snake can then evolve briefly to capture the non-rigid motion of the cardiac chambers. This is repeated for the remainder of the frames of the cardiac cycle. The main drawback with this method is that it is sensitive to initialisation. If the first frame can be initialised reasonably accurately with the GVF snake then the remaining frames are more likely to be tracked correctly. However as this thesis demonstrates, there are limitations with the GVF method in foetal echocardiography, but if the GVF snake was replaced by another more accurate method, the remainder of the algorithm is capable of providing reasonable tracking of the myocardium. To compensate for global image intensity in-homogeneities the elements of the local edge profile can be normalised to the sum of the kernel intensities.

In my publication during 2005 [73] a level set snake algorithm based on the MS term was applied to all four cardiac chambers simultaneously. The level set snake is able to handle partially missing intra chamber walls with the presence of a collision

detection term. The algorithm was applied to one 3D and one 2D dataset. The entire cardiac cycle was segmented and EF was computed and compared to other's findings in the literature [2].

To extend my work further I incorporated a shape prior into my previous level set algorithm from 2005 on 2D images [70]. In this work the training data comprised of both 2D and 3D datasets (3 in total) from which 26 slices were treated as independent images for segmentation. The images are hand segmented with each chamber filled. The filled manual images are then registered using a rigid body global optimisation method in order to construct a mean template from the registered SDMs. This is done by Differential Evolution [69]. This template is used to initialise and constrain the snakes in unseen images in a leave one out approach. As expected the segmentation results with shape prior were much better than with the prior disabled.

In order to make comparisons fairer I retrofitted the shape prior term from [70] into existing level set snakes that have been applied to the foetal heart. I published this work in 2007 [71] and in this study algorithms from [12], [73], [70] were directly compared on the same datasets both with and without the shape prior enabled. All algorithms benefited from the shape prior but [73] turned out to be more conforming to manual segmentations than [70] due to the presence of collision detection and edge advection terms as well as a mechanism for reducing the constraints of the shape prior.

In [126] I replaced the construction of mean template from the training SDMs (in [70]) with mean template from the registered training binary images. This provided a large speed up in the computation of the mean template but in order to compute four distinct chambers offsets to the mean, -2 standard deviations was needed. This factor was determined experimentally and would probably change as more training data becomes available. I then took the best snake from the head to head comparison in [71] (which was [73] + SP defined in [70]) and applied it to 53 2D images taken from only 8 datasets. Most of the segmentation results were to within 3 pixels of manually traced data. The snake with and without prior was then used to segment a real foetal heart with synthetically generated noise added.

Navaux and team [127] have published their latest paper in 2008 where they hierarchically found the heart automatically in 2D B-mode images. In contrast to segmentation algorithms they do not try to delineate the chambers but instead to roughly characterise the centre and size of the chamber with a fitted circle. To search for the heart and chambers they constructed a circular mould centred on the

heart in an example image. They created a similar one for the RA chamber. In their paper their images are not manually rotated to present the ventricles below the atria (as in this thesis) so they have a more challenging issue of identifying each chamber since they claim their method is rotation invariant. The intensity distribution of the pixels within the mould is stored and matched to the mould initialised automatically in another image. The circular mould is repeatedly scanned around the image and its size is refined to match the structure that it found. In their results they state that the heart was localised correctly but the RA was only found automatically in ~71% of the images on average.

Recently Tutschek and Sahn published segmentation of six foetal cardiac datasets in 2008 [128]. They analysed 3D echocardiographic datasets obtained using spatiotemporal image correlation. They analysed the datasets offline on a standard PC using a commercially available segmentation algorithm designed for ovarian folliculometry (RAB 4-8 transducer, Voluson 730 Expert ultrasound machine; SonoAVC, 4DView version 7.0 software, GE Healthcare, Zipf, Austria. The algorithm allowed them to define volumetric seeds both automatically and manually within a region of interest that would grow and approach the borders. They used manual control over the volume of each seed and a distance threshold between adjacent chambers separated by weak echoes. If the grown volumes failed to recognise the true chamber boundaries manual correction was applied to the segmentation. A complete good quality dataset could be interactively segmented by this method in less than 10 minutes. They claim that the algorithm has its limitations including operator dependence and can only handle static volumes.

I recently published the main results of this thesis in a clinical paper [129]. In this paper 9 datasets of which 7 were 3D and 2 in 2D were automatically segmented by the MSSCD algorithm after being initialised by user defined seed points in the centre of each cardiac chamber. The segmentation accuracy was comparable to manually traced contours.

3.9 Summary

This chapter presented a literature review of foetal cardiac segmentation as well as some techniques that researchers applied for preprocessing and segmenting adult heart data. It shows that previous efforts have been made on 2D foetal cardiac datasets using a variety of algorithms from low level classification of pixels to deformable models both explicit and implicit. My previous work focussed on tracking the foetal heart using a GVF snake in 2D images. Developments in US transducers have enabled direct volume acquisition of cardiac structures and so there is a need to

develop methods that will segment 3D foetal heart datasets to enable blood volumes to be estimated. The next chapter continues the approach of deformable models to aid in reconstruction of partial boundaries during segmentation of the endocardium.

4 MATERIALS AND METHODS

4.1 Data Acquisition

Volumetric data were acquired using two methodologies, the first from sweeping the imaging plane along an axis perpendicular to the images and the second by using a probe housing a 2D array of transducers (matrix probe) which is capable of capturing 3D volumes directly. The first method will be labelled as "2D slice sweeping" whilst the second as "Live 3D". In both methods tissue second harmonic imaging was not enabled on the scanner since this feature tends to reduce the fundamental ultrasound beam by a factor of two and this results in poorer spatial resolution. For example a fundamental frequency of 4 MHz would provide a good depth resolution and adequate spatial sampling of a deep structure such as the foetal heart; the second harmonic imaging will reduce the fundamental to 2 MHz and provide reception for non-linear echoes at 4 MHz. Currently second harmonic echo reception at 8 MHz is not provided on the scanners available to the group.

4.1.1 2D slice sweeping

Deng J et al [20], [16], [83] acquired data with the simultaneous use of two US scanners, one was used for recording structural greyscale images and the other for acquisition of umbilical artery Doppler waveforms which were then converted to provide pseudo foetal ECG gating. The dataset was acquired with B-mode scanning using a parallel mechanic probe operating at a frame rate of 25Hz. All frames for each slice are recorded before the probe is moved into position to acquire the next slice. At least one cardiac cycle is acquired at each slice position. Manual reformatting of the images from several cardiac cycles is required by the clinician to produce a 4D dataset that shows corresponding frames for each slice and lasts for a single heartbeat. The dataset was acquired with spatial resolution of 0.26mm in the fan beam plane and 0.5mm between slices. Although the 2D images from the paired scanners were stacked in 3D with motion gating, there were parts of the volume with noticeable motion artefacts that caused misalignment between slices (refer to Figure 37 and Figure 38). For this reason the volume dataset from the Acuson scanners is treated as separate 2D images in this thesis.

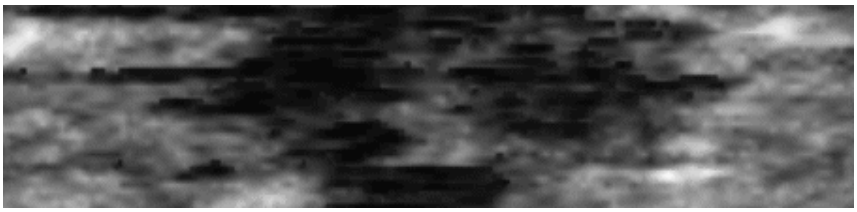


Figure 37 Reconstructed image showing short axis view. The black streaks indicate that the acquired 2D slices are not properly registered.

There is some in plane movement of the heart during acquisition of each image slice and this shows up as a wiggle in the surface rendering of the cardiac boundaries as displayed in Figure 38. This artefact needs to be corrected before any cardiac indices are derived from the segmentation.

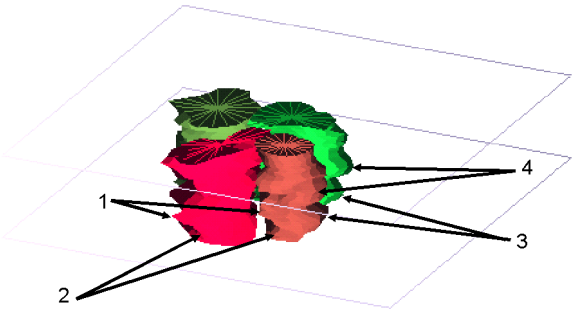


Figure 38 Surface rendering of the four chambers by 2D segmentation showing that the movement of the heart is apparent between slices. It is apparent that the slices are misaligned from this figure because of curvature correspondence between the individual chambers as indicated by the arrows. The chambers farthest away are the atria and those closer to the camera, the ventricles.

4.1.2 Live 3D

The Live 3D system is a trademark of Philips Medical technologies. The scanner is a Philips Sonos 7500 with the X4 matrix probe capable of imaging about 20-25 volumes per second with square pixels in the plane perpendicular to the beam axis. The resolution along the axis is lower than in the lateral plane but datasets acquired with this scanner can be automatically resampled to produce cubic voxels. Over 3600 elements make up the matrix probe. The operating frequency is between 2-5 MHz. If the entire heart fits within the volume of acquisition then the imaging volume does not have to be moved to accommodate the part of the heart that are not currently imaged and so gating is not required. If the heart is particularly large then two options are available; either to fix the position of the probe and allow the phased array to sweep the heart by steering the beam or to move the probe very slowly to sweep out a longer imaging volume (refer to Figure 39). In the 2D slice

sweeping method motion gating is required to trigger the acquisition when the heart is in the right place for imaging as the probe is moved steadily. In the Live 3D volumetric acquisition this practice was not used since most of the heart fits inside a static imaging volume and so the motion artefact will be less than movement of the conventional slice acquisition probe.

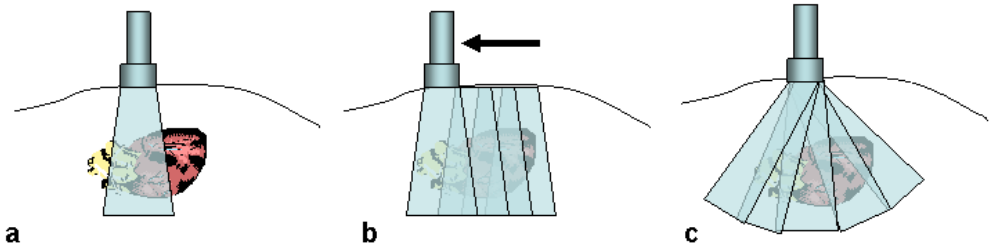


Figure 39 Illustrations to show how the entire heart can be acquired if it does not fit completely within the imaging volume. (a) Shows a static probe and beam, in (b) the probe is moved slowly and (c) shows the static probe but with a sweeping beam.

4.1.3 Dataset Characteristics

The datasets are resampled into a Cartesian grid by the scanner and output as 8-bit greyscale. The ones presented for this project are listed in Table 4. Although some of the datasets appear to have high resolution this apparent small voxel size is not the true width of the US beam that samples the reflectors since the A-lines are superimposed on a Cartesian grid and the unfilled regions are interpolated. The lateral beam width is several voxels therefore the partial volume effect and speckle size are larger than the voxel dimensions. The effect is more noticeable in 3D than 2D since there are more transducers per imaging plane in conventional 2D probes. Although the smaller 3D datasets with 40 and 43 voxels across the length of the heart appear low resolution they are in fact amongst the best in terms of complete boundaries for segmentation by both manual and automatic methods.

Table 4 dataset characteristics

Dataset	Approx num of pixels across entire heart length (*zoomed in and resampled by sinc interpolation)	Voxel Size (mm)	2D/3D
A	80	1.03x0.36x1.03	3D
B	66	0.62x0.32x0.62	3D
C	40	0.49x0.49x0.49	3D
D	43	0.68x0.68x0.68	3D
E	95	1.00x0.31x1.00	3D
F	108*	0.31x0.31x0.31	3D
G	30	0.32x0.61x0.61	3D
H	73	0.46x0.46x0.91	2D
I	109	0.26x0.26x0.5	2D

Although there are more datasets in 3D rather than in 2D there are not enough to test for statistical significance. This is because the datasets are acquired from volunteer patients that are not part of routine foetal pregnancy scanning protocol. Since the largest artefact for 3D scanning is foetal movement, acquisition occurs during the resting phase of the foetus. This means that the acquisition parameters are not necessarily at optimum and the foetal heart may be not fully within the scanning volume or the acquisition angle may not be suitable due to heavy shadowing. Another common problem is that the atria are not easily seen in some datasets since they may be too small in the image due to incorrect zooming set on the scanner controls or the intra cardiac boundaries have very little contrast to the chambers themselves. The clinical research project ended in December 2006 so no further datasets are being acquired.

4.2 Snakes

The foetal heart is a rapidly beating hollow soft tissue organ which comprises of a comparatively smooth exterior surface (epicardium) and a fibrous inside wall (endocardium). A non-rigid deformable model approach is used to segment the heart since it is capable of mimicking the motion of the heart walls through the cardiac cycle and variation in shape between individuals. Snakes have some attractive properties in this application. The first is that the snake is a series of connected elements that form a high level contour which can evolve to fit the data through minimisation of energy. The contour can deform to segment the data with little user interaction based on salient features that it extracts from the image. Since

the model has both constrained stretching and bending components it can provide a smooth interpolation where missing boundaries are present in the data.

The explicitly parameterised snake algorithms in this thesis constrain image feature extraction to the 1D search space along the local contour normals to reduce vertex clustering. The implementation of a snake algorithm in this thesis is to sub-pixel/sub-voxel precision and computes local forces on each snaxel instead of energy. Force unlike energy is a vector quantity that comprises both magnitude and direction and so provides additional information to determine how to move a system of connected elements. Other force based deformable models in the literature include the work of McInerney et al [58], [46], Lobregt and Viergever [130], Montagnat et al [7].

The snakes are initialised as the same single sphere (level set) or cylindrical chamber primitive (explicit snake aligned with long axis of chamber) within each chamber to ensure fairness (or from a template as in TIMS+SP) and were allowed to evolve for a fixed number of iterations sufficient for them to reach past the edge of the endocardium borders if necessary. Each type of snake required a different maximum number of iterations to fulfill this due to the weighting parameters, contour description and timestep restrictions. To assess the accuracy of segmentation, the rms projected distances of the snaxels onto the uncorrupted model image were computed.

The MSSCD and EPCA were started as large spheres with diameters at roughly the three quarters of the width of the chamber. Since MSSCD can both expand and contract according to the image this is not a major problem. In the case of EPCA the constant advection term transports the zero levels in a pure expansion fashion. For this reason the primitive spheres initialisation were set to radii smaller than the chamber width in the EPCA snake. From experimental findings the initial seed point was required to be at least half the radius of the chamber for sufficient numbers of pixels to contribute to the statistics of the grey levels within the inside region when using the MS term.

It was anticipated that the TIMS+SP with the SP term would be more robust across many images compared to the amorphous snakes. Therefore we implemented the same shape prior term from (2.29) in each of the amorphous level set snakes described in section 2.8 to make the test fairer. However, instead of the initialization from the template as in the TIMS snake I used the seed points to start the snakes with the added shape term. The justification for this is that the amorphous snakes in section 2.8 were designed without a shape prior term, and if they start from seed

points instead of a template and can still recover the required shape they are more robust to an incorrect template fit initialization to the image. To distinguish these enhanced snakes from the originals we shall use the same terminology defined in section 2.8 but with the term +SP added on.

For the TIMS level set snake I used the same method of starting from the template since this is the way it was implemented in [70] and acts as a benchmark for the other snake equations started from seed points. I used manually traced boundaries as a gold standard to evaluate errors. TIMS can segment both with and without the +SP term but was initialised in both cases from the template image instead of the seed points.

The weighting coefficients for the snakes were determined experimentally on a typical image taken from those used in this paper. The parameters were then held fixed for the remainder of the images under investigation. It was found by visually inspecting the automatic segmentation that the order of magnitude was important in setting the coefficients and not their exact values themselves. When the shape prior term is added to a particular snake the weight is set to an arbitrary value of 0.5. In this thesis the shape prior analysis was only performed in 2D in order to obtain several slices from the limited number of 3D datasets.

The segmentation pipeline for each snake is shown in Figure 40.

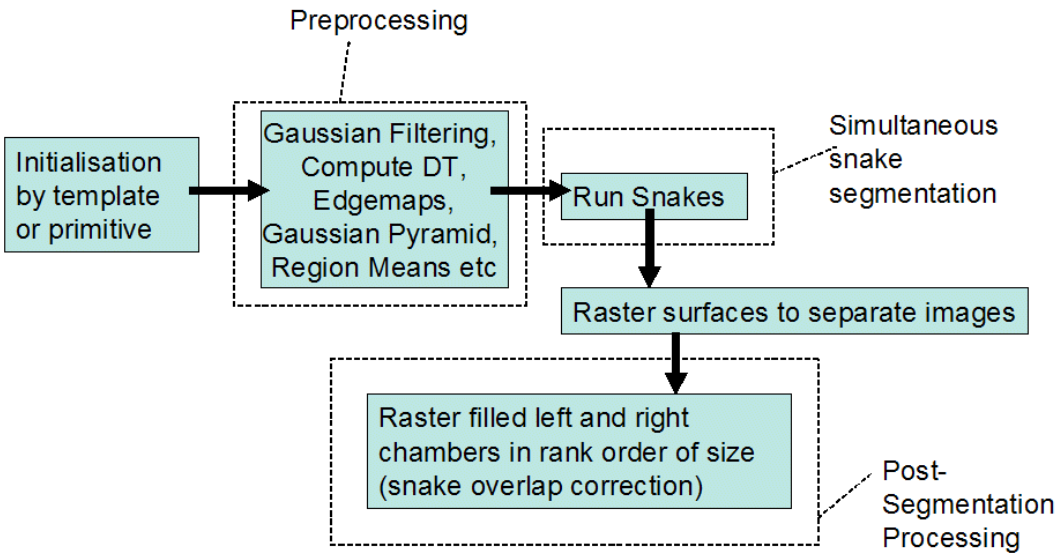


Figure 40 Segmentation pipeline

After the surfaces are rastered to separate images they can be flood filled and the post-segmentation processing stage compensates for partial leakage between chambers. This algorithm is outlined in section 2.10.

4.3 Synthetic Dataset Construction

One of the easiest to visualise datasets was chosen to act as the model for the image corruption tests. This dataset was manually segmented to reveal a binary image. From a selection of 5 foetal heart datasets the histogram of both inside the chambers and outside was computed and is displayed in Figure 41.

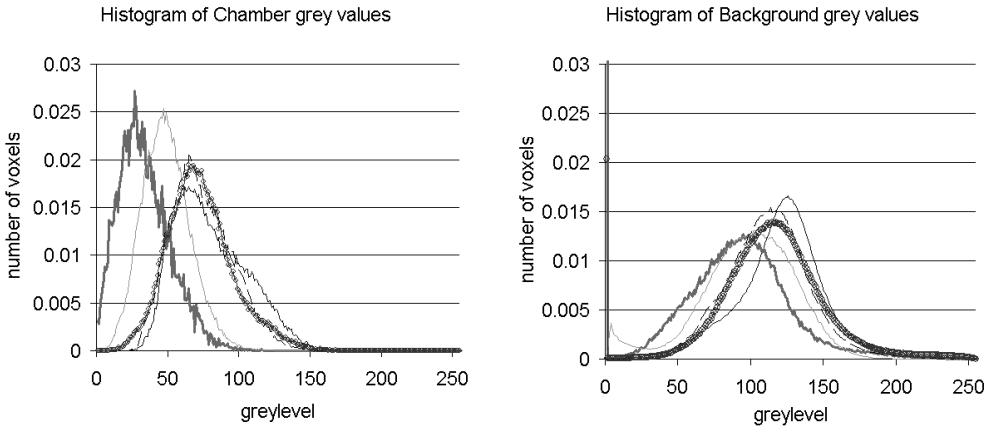


Figure 41 Histograms of 5 selected datasets showing inside the chambers (*left*) and background outside (*right*). Each line corresponds to a separate dataset. The histograms are normalised to the area under the graph to show them at comparable heights (number of voxels). The image on the *right* is clipped at *0.03* to emphasise the greyscale data rather than the volume outside the field of view of the probe.

This figure shows that there is quite a substantial overlap between the mean of grey levels distributed both inside and outside of the chambers. From these two graphs a typical region intensity for a model image was computed for both inside and outside the chambers as shown in Table 5. The mean intensity inside the chambers is then computed by averaging the histograms of chamber regions over all selected images. To calculate the average background the histograms of background grey values are averaged over the selected images.

Table 5 Mean and variance of the foetal heart dataset, rounded to the nearest whole number

Region	Region Mean	Region Variance
Chamber	40	~300
Background (outside chamber)	91	~1300

The model test data generated from the region means can be seen in Figure 42 with the appropriate piecewise constant region intensities from Table 5.

Figure 42 Synthetic image (only central slice shown) where the region intensities are 40 inside and 91 outside. The LV is on the *right hand side* of the image and the ventricles are below the atria.

This image is then be corrupted by an appropriate noise type to test the robustness of each snake to imaging noise. Each noise type is applied separately since it will serve to identify which type dominated the failure rate of the snakes.

4.4 Validation of Snake Algorithms using a Physical Phantom of Known Volume

A grape is used to model the heart since it is approximately the same size (~2cm length). Although simpler in shape compared to foetal cardiac morphology and it can only represent a single phase of the cardiac cycle; this was deemed an adequate static model for the entire foetal heart for ultrasound scanning by an expert foetal cardiologist. There is only a single compartment to segment in the grape unlike in the foetal heart so it is much easier to constrain the snake in terms of leakage outside the boundary. Also since there are no motion artefacts due to foetal and maternal movement, this phantom is a much easier structure for the deformable models to segment than an in vivo foetal heart. To achieve ultrasound penetration though the skin the grape is microwaved for about 4 minutes on low power to soften the tissue. The grape is then scanned in a cylinder of ultrasound scanning gel (as shown in Figure 43). By experiment the grape was positioned in the tube to achieve optimum visibility of the entire skin and minimum shadowing under US acquisition. The narrow tube allowed for multiple reflections of the US beam but to some extent this can be minimised by adjusting the gain and depth of the grape in the gel. Since the grape is immersed in ultrasound gel with no bony structures there are less shadowing artefacts for the segmentation algorithms to adapt to. The contrast between the ultrasound gel and the grape skin makes it easier to manually trace the boundary compared to foetal cardiac structures.

Figure 43 Scanning the grape in a cylinder of ultrasound gel.

4.5 Summary

This chapter introduced the dataset characteristics acquired from both 2D and 3D scanners along with scanning protocols. Next the initialisations were described for each snake type as well as the segmentation pipeline. The method by which the generation of the synthetic dataset was explained was followed by the setup and imaging of the physical phantom by the Live 3D scanner. In the next chapter each of the snakes will be run on all the introduced datasets from this section and analysis of the results will be presented.

5 RESULTS AND DISCUSSIONS

5.1 Sensitivity of the Snakes to Imaging Noise

The synthetic images were generated as in section 4.3 and four ultrasound noise types were added to simulate the acquisition. The noise types are added separately and increased in a controlled manner to find the breaking point of the snakes.

5.1.1 Rayleigh noise

Rayleigh distributed additive noise was used to corrupt the test image. Each deformable model with the exception of the scale space snake required the image to be Gaussian prefiltered to reduce the speckle before segmentation. The Rayleigh equation is shown in equation (5.1) and the graph shown in Figure 45. Note this graph is similar in shape to the image histograms in Figure 41.

$$p(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{5.1}$$

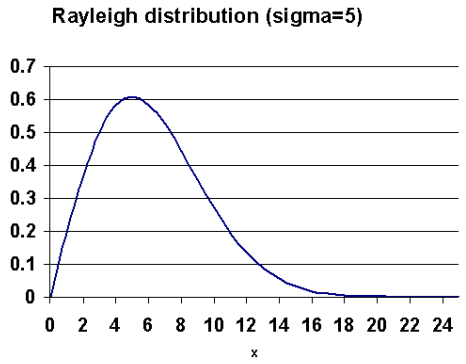


Figure 44 Rayleigh Distribution in 1D for positive values of x.

Figure 45 shows the segmentation results for each of the 2D snakes under progressive levels of Rayleigh distributed additive noise. The value of ~300 is chosen for the nominal noise variance since this was found in the chambers of the datasets under investigation. The noisy images appear darker than the uncorrupted version because the noise is added to the test image in the form:

$$I_{corrupted} = 0.5I_{original} + 0.5I_{Rayleigh} \tag{5.2}$$

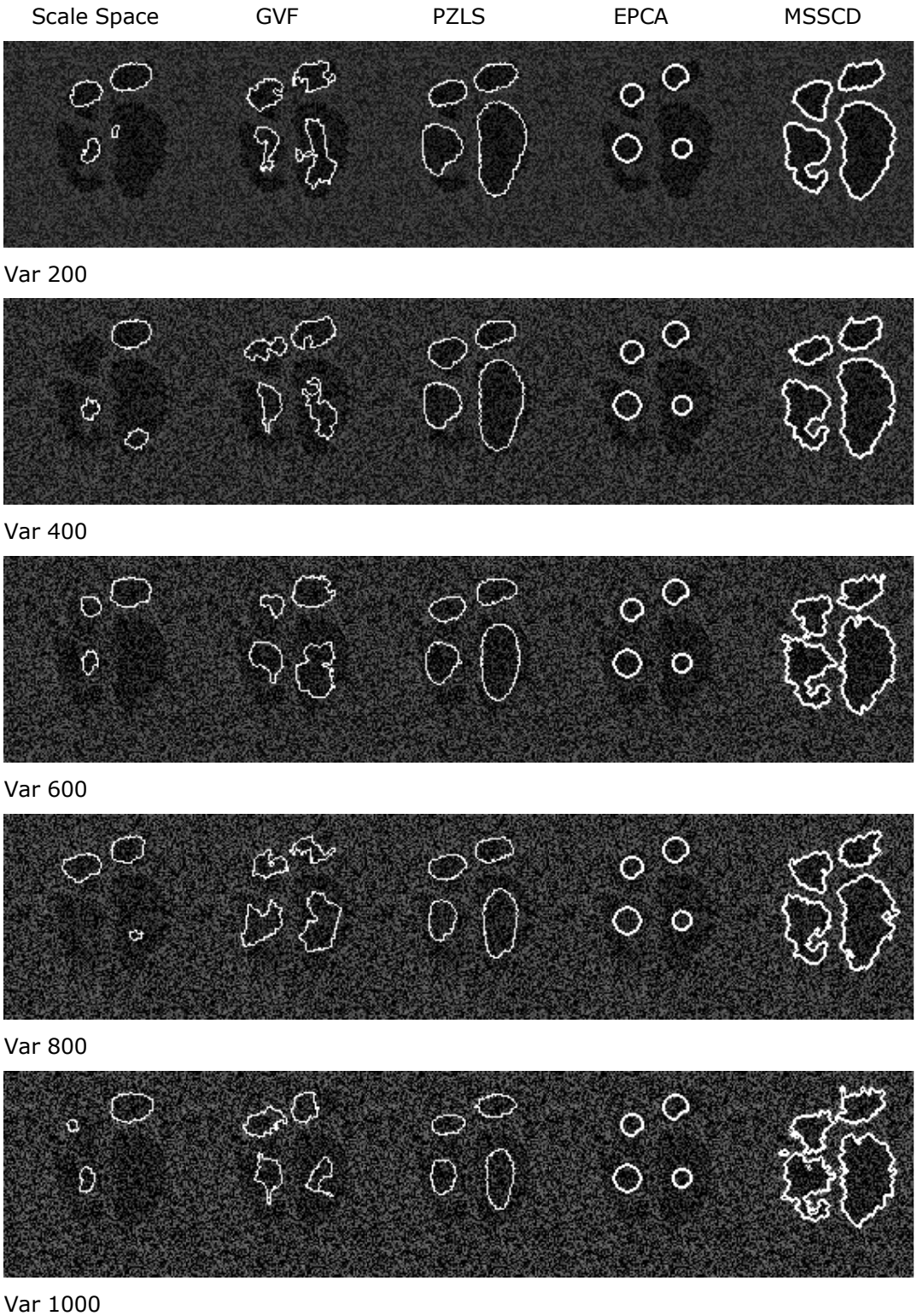


Figure 45 Rayleigh noise corrupted images test and their effect on the snakes without shape prior (only central slices shown). The numbers below the images refer to the Rayleigh variance (σ^2).

The modelling of speckle in these images is a simplification of real speckle found in US images. Equation (5.1) is a random process and has no dependence on the underlying US scatterers that are being imaged.

It is apparent that region based snakes are more effective at segmenting noisy endocardium boundaries in the presence of speckle noise over snakes that use edge information. This also shows up in the quantitative analysis in Figure 53. A balloon term seems necessary to provide a balance force to maintain a non collapsed shape in the presence of such noise. This is apparent for both PZLS and Lassige. Ordinary edge detection guided contours seem to collapse upon themselves as in the case of GVF and scale space. The GVF is also problematic in this area but has a very strongly weighted GVF diffusion term to avoid collapse. For all of these synthetic noise tests a very weak balloon term was added to the GVF snake since in practice the GVF field appears to be overpowered easily by the local curvature and continuity terms. Since the mesh does not reparameterise itself to account for growing in volume the continuity forces can dominate as the vertices get further apart during evolution. This is more a problem with the ventricles than the atria on the synthetic dataset noise tests as they appear larger in this cardiac phase.

Out of four chambers that the scale space snake appears to converge upon the correct edges of the LA and sometimes the RA are segmented with no detection of the ventricles. This is most likely due to the close proximity of the initial primitive to the chamber boundaries in the atria. The ventricles are larger and their boundaries especially at the apices are therefore further away from the initial primitive than in the atria. The LA of the scale space seems fairly consistent visually as noise is increased; the PZLS seems to do fairly well until about variance 400, after this level the chambers shrink away from the blood pool boundary. Since MSSCD has both edge and region terms it can reach the boundary even beyond this level of noise. However, the delineation becomes very jagged since at high levels the speckle dominates comparably to the contrasting chamber-background regions.

It is clear that the region snakes deteriorate rapidly as the level of Rayleigh noise increases, more so in PZLS than in MSSCD. This is because they are modelling homogeneous regions which are highly speckled. Oddly enough the accuracy of the edge based snakes does not change much with speckle when the chamber edge is detected. Perhaps this is due to the large presence of local noise at all levels in this test. Edge based segmentation accuracy remains low with a Generalised Tannimoto Coefficient (GTC) of below 0.3. The MSSCD on the other hand starts off at 0.9 and drops to 0.75 at a variance of 1000 which is still very good compared to segmentation of real data. I measured chamber variance to be ~300 in a sample of real images and the background variance to be ~1300 (refer to Table 5). Taking this into account the MSSCD handles speckle noise in the chambers well.

5.1.2 Attenuation

Attenuation was simulated using an exponential decay function

$$I = I_0 \exp(-\lambda y) \tag{5.3}$$

Where I_0 is the intensity at the top of the image, λ is the attenuation coefficient or decay constant and y the vertical vector along the simulated beam with origin at the top of the image.

Gaussian prefiltering served to homogenise the intensities in the image and is used before application of each snake algorithm, with the exception of the scale space snake where hierarchical scale space filtering is handled internally. Figure 48 shows the segmentation results of the central slice through the dataset for each snake. The attenuation test is harsher than the attenuation found in the real datasets presented in this thesis since the TGC does a fairly good job in the absence of shadowing (see Figure 46 for an example). Shadowing is modelled in the next section by occlusion to simulate strongly attenuating structures.

Figure 46 Imaging outside the heart structures. The bright areas are bony structures in the foetal chest. Some shadowing can be seen under these bony regions.

In the attenuation test the weighting of the internal forces was increased more than usual for each snake to compensate for possible drop out artefacts from occurring.

The mean chamber intensities after attenuation are shown in Figure 47. The most severe falloff comes from the background which at the greatest value of the decay constant starts to approach the mean chamber intensities within the atria. At this level of attenuation the mean intensity of the LV and RV are below half of their uncorrupted values. The atria are less affected since they are located nearer the simulated probe at the top of the image.

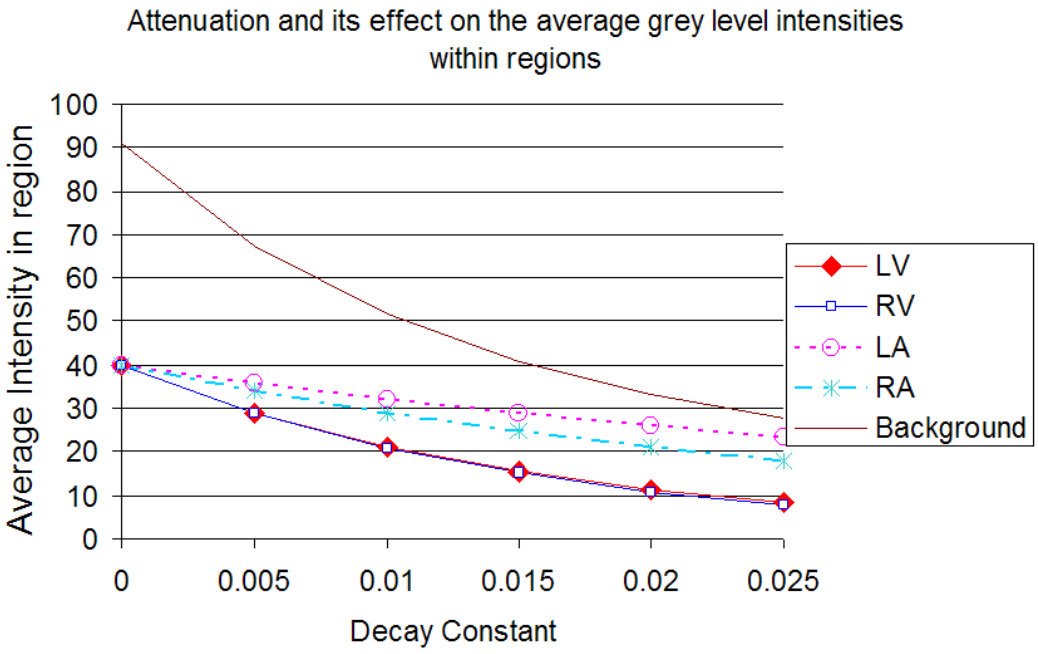


Figure 47 Mean of region intensities after attenuation.

In Figure 48 the scale space snake has trouble locating edges even with mild attenuation. Once again a balloon term could be useful in preventing collapse of the contours due to the self regularisation internal forces. On a positive note the snake maintains a fairly uniform level of GTC with increasing attenuation even though the degree of overlap is very low between 0.1-0.2 (refer to Figure 53).

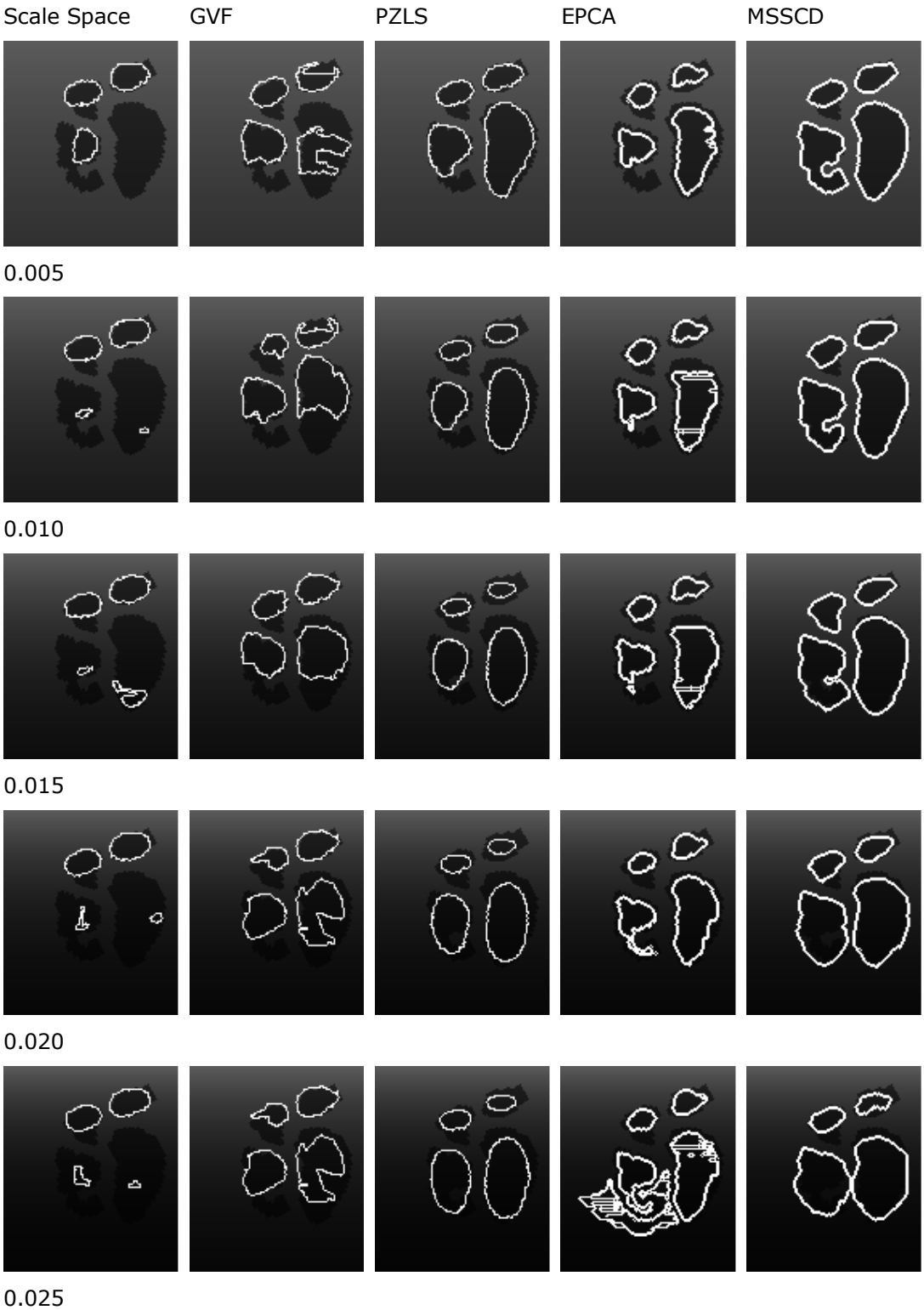


Figure 48 Attenuation test for no shape prior (only central slices shown). The numbers *below* the image *row*s refer to the decay constant.

In both the GVF and scale space algorithms the atria are better segmented than the ventricles and this is evident from Figure 48. It is likely that since the atria are smaller than the ventricles that they are easier for the longer range edge forces to pick up a significant contribution to counteract the mesh regularisation forces. Also

the attenuation effect is more pronounced in the bottom of the image and this leads to weaker edges in the ventricles. In both Rayleigh and attenuation graphs in Figure 53 the GVF and scale space GTCs seem to be steady but low throughout the range compared to the region based snakes.

In the PZLS snake it is clear that this algorithm's delineation is becoming more rounded as the attenuation coefficient is increased. This seems to have a stronger influence on the top part of each chamber. A possible explanation is that as the attenuation is increased the mean intensity inside the zero levels decreases and the balancing between the inside and outside MS forces occurs at a lower intensity i.e. before the snake reaches the myocardium edge at the top of the chamber. This effect does not appear to occur in MSSCD which uses the same MS term unless the initial primitive for the LS MS snake is larger than the explicit MS one. If this occurs then the region statistics would be better sampled for the LS snake and it has a better chance of reaching the boundary at the top of the chamber.

From Figure 48 it is clear that EPCA maintains a consistent shape over the entire series of increasing attenuation noise. At the extreme case of 0.025 value of decay constant this snake leaks outside of the chambers. Given sufficient number of iterations the snake overcomes weak boundaries due to the constant advection term. The graph in Figure 53 shows the similar trend with only a relatively small drop in the GTC due to the leakage. On close inspection of the central slice images it appears that the contours do not quite reach the endocardium boundary. In the third dimension this would also be the case and the volume segmented would be underestimated. To some extent this is caused by Gaussian prefiltering before running the EPCA algorithm. There is some stratification shown in the segmentation of the chambers interior particular the larger LV in some of the images. The vertical attenuation profile has created a false edge for the edge detector in EPCA to latch onto.

The MSSCD snake manages to remain approximately within the chambers throughout all stages of the attenuation test. Higher attenuation causes the detected boundaries to become more rounded and more bloated causing some leakage to occur. This is probably due to the low contrast of the background compared to the blood pool with increasing attenuation. Since the MS term tries to achieve segmentation between these two piecewise constant regions some contrast in the mean intensities of both regions is needed. In Figure 53 the falloff is more pronounced in the Attenuation test for MSSCD compared to Rayleigh and is probably due to leakage and rounding of the chamber delineations as the attenuation coefficient is increased. There is less deviation from the chamber shape in the

Rayleigh noise model. Also the attenuation coefficient is multiplied by the image and so has a greater effect on the uncorrupted image whereas the Rayleigh component is only added.

5.1.3 Occlusion

Shadowing due to calcified foetal ribs can cause occlusion of lower structures and an attempt to simulate this effect on the snake segmentations is shown below. A vertical beam of fixed width 15 pixels (arbitrarily chosen) was placed at the left side of the foetal heart. The attenuation provided in this simulated shadowing is uniform instead unlike in the previous section. The intensities within the chambers that fall within the shadowed region are unaffected since they should contain no echogenic structures apart from blood which is not detected by the scanner protocols. Wall structures that fall within the beam region are attenuated to match the shadowed area. The regions used in this simulation are indicated in Figure 49 where I refers to intensity of a region, C, B and S refers to the chamber, background and shadowed regions respectively. The intensity I_S of region S can be varied such that

$$I_C \leq I_S \tag{5.4}$$

The test is to determine at what degree of shadowing causes snake leakage from region C into region S. When I_S is approximately equal to I_C leakage from chamber to background is expected to occur.

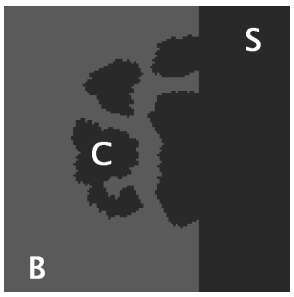


Figure 49 Simulated constant intensity shadowing region (*S*). *C* refers to the endocardiac chamber region. *B* is the myocardium or background intensity.

Figure 50 shows the effect of the segmentation of each snake algorithm. It is clear that the region based snakes are at a disadvantage here since the shadowed region is of a similar intensity to the chamber. In the absence of speckle noise the edge based algorithms do no leak into the void but still pick out the weak edge. Therefore it is a trade off whether to use an edge detection snake for a dropout boundary which is poor at speckle rejection or a region based snake that may leak out of the chamber but pass over the speckle noise. In shadowed regions it is difficult to manually trace cardiac chambers without some informed guesswork since the low intensity areas are often plagued with noise which can corrupt edges.

From the graph in Figure 53 the edge based snakes are the most robust to the occlusion noise but the region based ones are not much worse than their falloff results in the Rayleigh speckle test. The PZLS falloff is similar to the MSSCD since it uses the same MS term but the LS framework allows for segmentation into concave boundaries such as the right papillary muscle (see Figure 50) which is not allowed by the non interpolating explicit mesh of the PZLS.

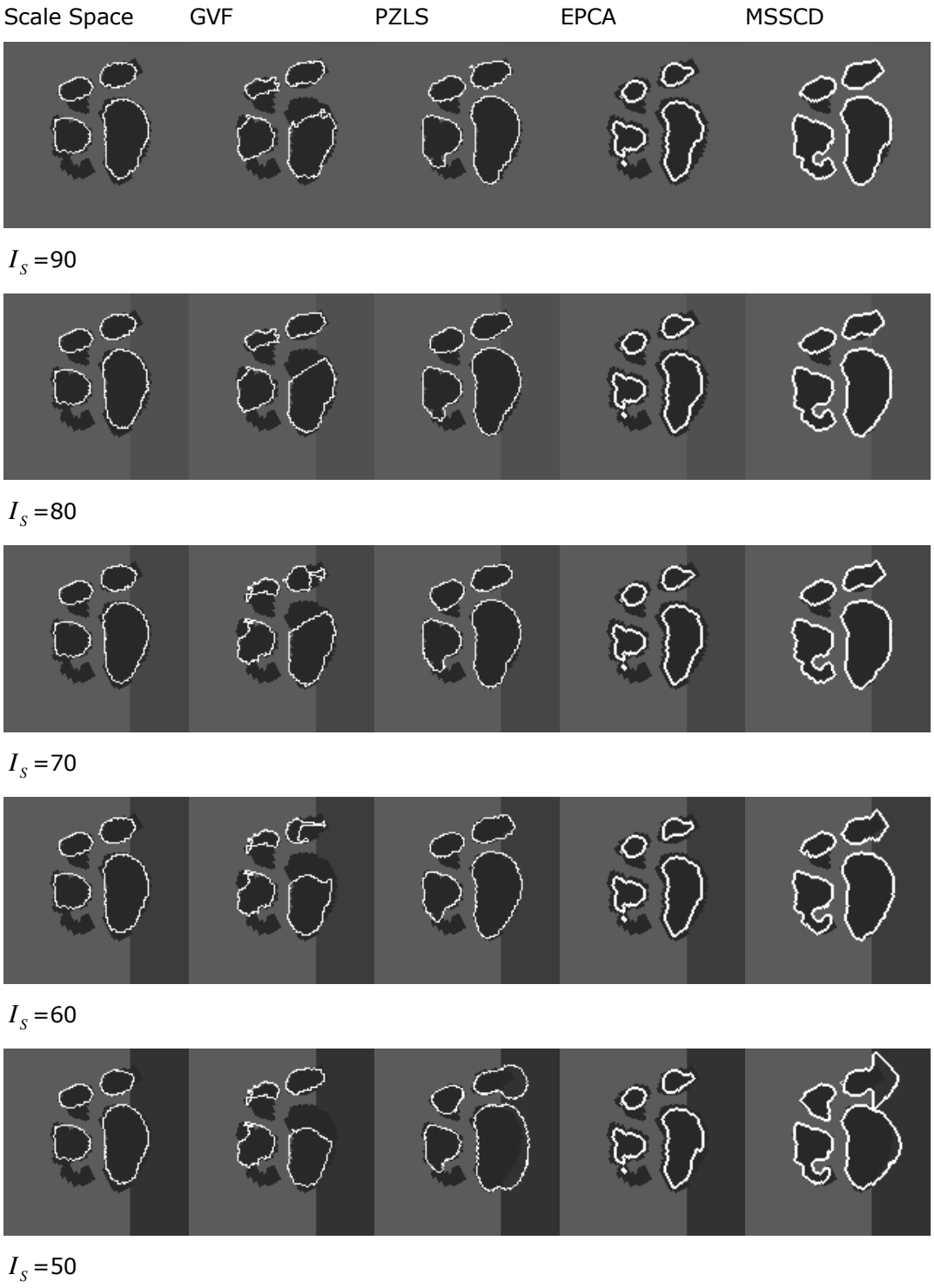


Figure 50 Occlusion test results, no shape prior, only central slice shown.

5.1.4 Partial Volume

All medical images are prone to the partial volume effect due to the finite size and resolution of detectors. Most medical modalities have cuboid voxels with weaker resolution in between the image acquisition plane direction. Using 2D US matrix probes to acquire volume datasets the sampling region plane perpendicular to the beam can be symmetric in resolution. However, along the axis sampling is often better owing to wave dispersion perpendicular to the beam. Imaging deeper organs usually leads to more noticeable asymmetric sampling than for structures closer to the probe (see section 2.1) that take on a curved approach due to the radial propagation of the wavefront.

This section presents the results of simulating the partial volume effect on the segmentation results. Since the change in sampling was restricted along the horizontal axis it does not fully mimic the effect of increasing partial volume with depth. This could be modelled for a given field of view as an extension to this work. To create the reduced resolution in the horizontal direction due to beam dispersion the uncorrupted image was Gaussian filtered then subsampled using linear interpolation. It was assumed that scanners use unsophisticated models of interpolation to make the pixels square. For this part of the test nearest neighbour interpolation was used to simulate the scanner's reconstruction.

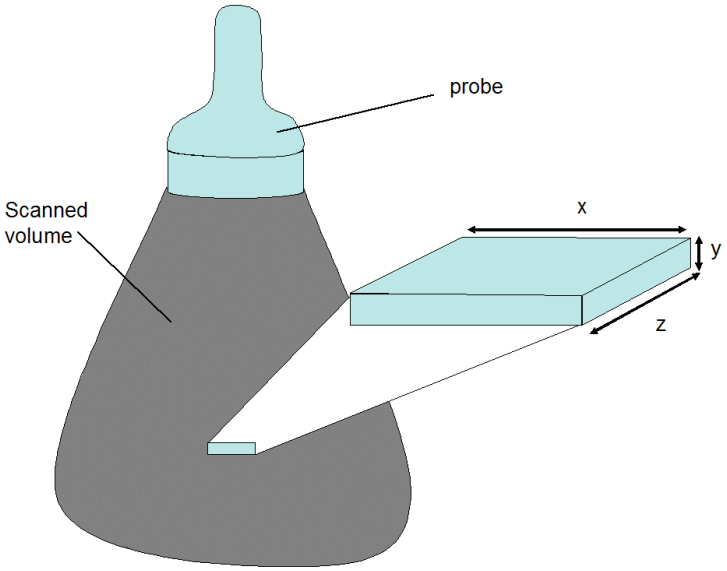


Figure 51 Modelling the partial volume effect in the x-z plane. The lengths x and z are equal. The voxel dimension in the y direction parallel to the beam axis remains fixed.

Figure 52 shows the results of the algorithms on the partial volume noise corrupted images. The tests involved varying the voxel aspect ratio (ratio of the lateral beam

voxel width to the axial beam voxel width as illustrated in Figure 51). The test stopped at a voxel aspect ratio of 3.5:1 since this was enough to be comparable to the voxel aspect ratios in the real data (refer to Table 4 in section 4.1). The GVF seems to be non consistent in the quality of the segmentation of each chamber as the partial volume noise increases but suddenly at 3.5 level each chamber is segmented satisfactorily. This places some concerns about the stability of the GVF algorithm. For the other algorithms there are no major changes in the quality of the segmentation visually. The snakes are not sensitive to this type of noise. This is also apparent in Figure 53 since there is hardly any falloff with the exception of the GVF which seems to get better with increasing the degree of partial volume.

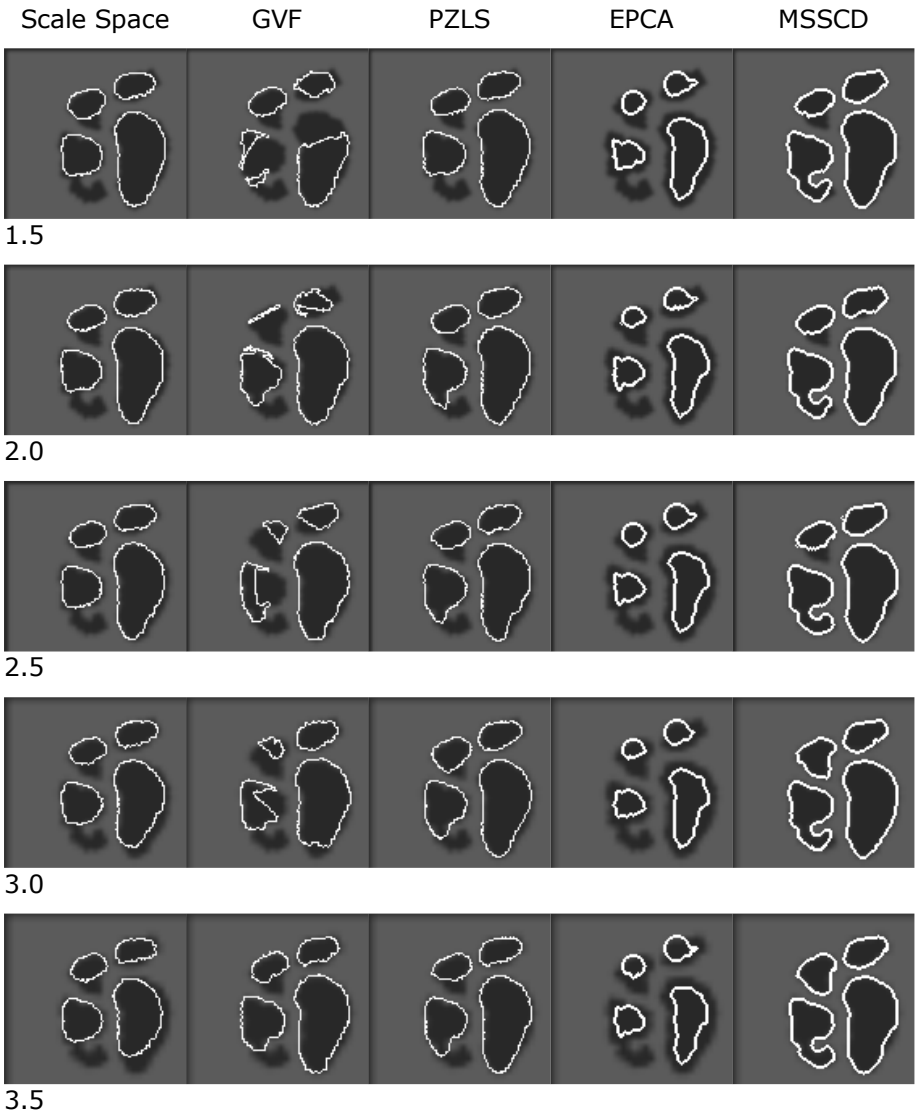


Figure 52 Partial volume test results. The number below each row refers to the ratio of the voxel dimensions beam-lateral to beam-axial.

To summarise the partial volume noise tests were the least challenging for all snakes, the region snakes are most affected by attenuation and occlusion noise and the edge based snakes are best kept away from the Rayleigh noise test. In each test the MSSCD had the highest overall GTC but this was because it could provide the most accurate segmentation in the absence of noise. The edge based snakes were towards the bottom end of the GTC scale except in the attenuation test where the excess rounding of chamber delineations in the PZLS causes its GTC to dip below the EPCA and GVF snakes.

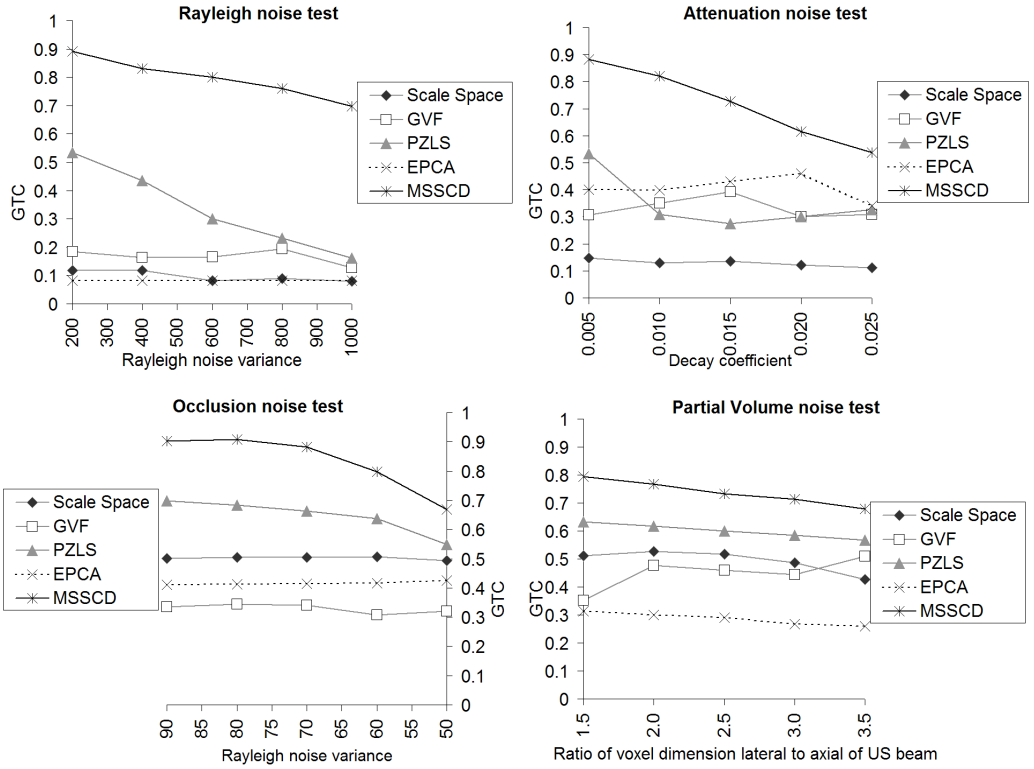


Figure 53 All noise tests quantitative analysis, no shape prior.

5.2 Sensitivity of the Snakes to Different Weighting Coefficients

In this test nominal values for each weighting coefficient were chosen empirically and two were varied at a time between ±80% of the chosen value whilst the other parameters remained fixed. The snake segmentation was compared to manual tracings to compute a GTC for each parameter set. When GTC was plotted against a pair of parameters this allowed surface plots to be generated for each snake. The snakes were applied to one of the higher contrast datasets in 3D. To generate the optimisation maps took 300 hours using six Pentium 4 PCs for the single high

contrast dataset. Hence there are no error bars on the surface plots. When analysing the optimisation maps it should be noted that the TC of the individual chambers is higher than the final GTC value. In these optimisation maps the aim is to maximise the GTC values rather than to minimise in usual optimisation applications.

5.2.1.1 Scale Space Snake

The possible combinations of different paired parameters for the scale space snake optimisation graphs are shown in Table 6. Y means tested, N indicates no test was done and N/A refers to absence of this combination of terms in this algorithm.

Table 6 Parameter combinations on optimisation graph plots for the scale space snake

Balloon	N/A	N/A	Y	N
Edge	N/A	N/A	N	N/A
Continuity	N/A	N	Y	Y
Curvature	N	Y	Y	Y
	Curvature	Continuity	Edge	Balloon

The GTC results for this snake type are shown in Figure 54 and remain low in every case of the scale space optimisation with minor fluctuations within the parameter surfaces. In these optimisation plots higher values are better since the objective is to maximise the GTC by varying the parameters. The GTC never rose above 0.45 which makes comparison with clinical tracings poor compared to the other snakes. Using scale space the snake locked itself on the strongest sets of edges which were not necessarily part of the heart, often in US enhancement areas. It approaches the edges at multiscale so it can take larger steps then the other snakes, allowing it to jump across weaker edges that would have stopped the non multiscale ones. The multiscale framework is relatively insensitive to parameter change as indicated by the uniformity in the colour maps and could be useful to implement in the other snakes. Only blues and light greens are shown in the parameter maps with no local hot spots above the halfway mark on the colour bar, so this snake cannot achieve a high GTC over the selected parameter range.

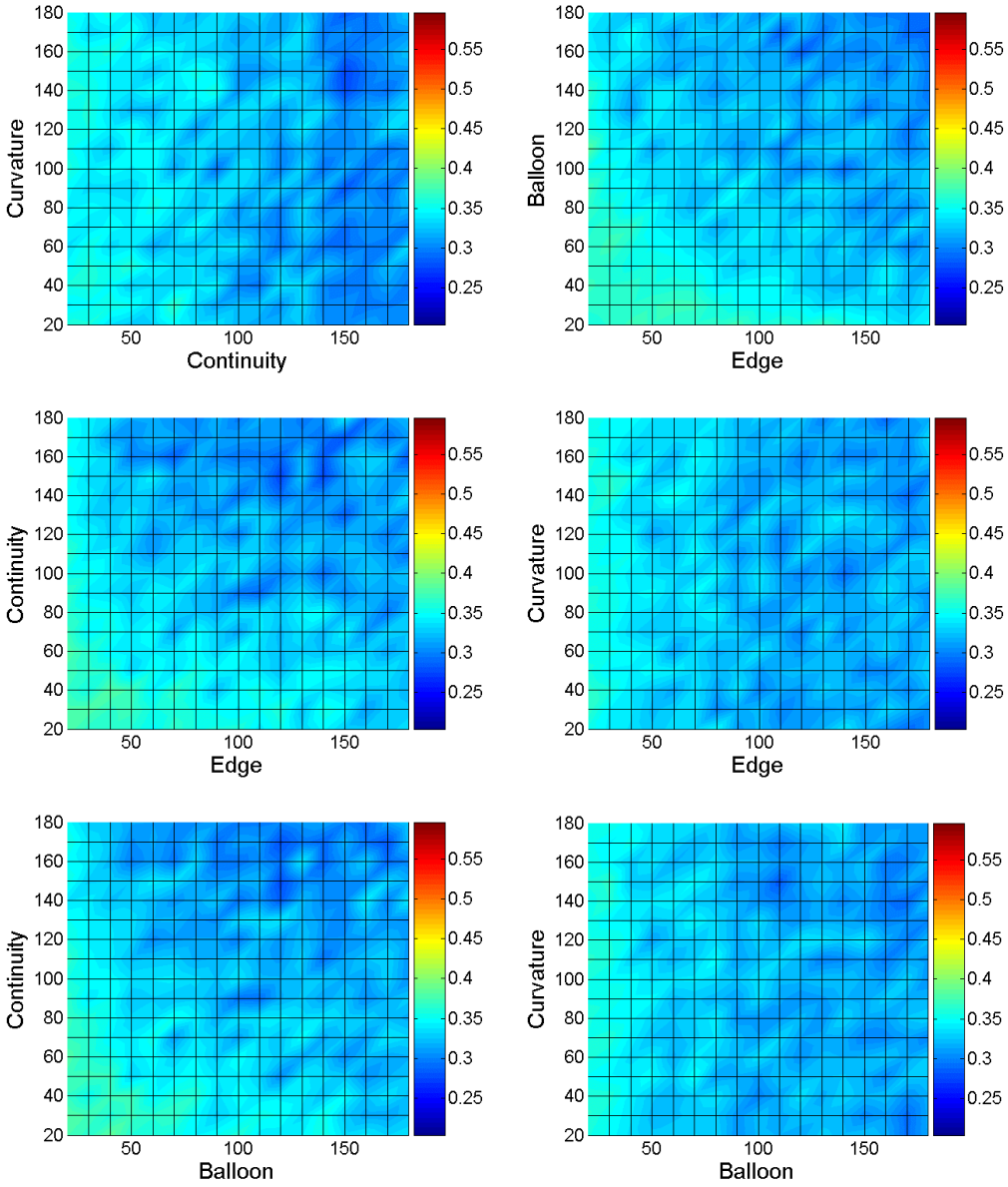


Figure 54 Sensitivity analysis of the weighting coefficients used in the scale space snake model over 80% variation. The parameters are expressed as percentages and the colourbar shows GTC used as the measure of quality of segmentation.

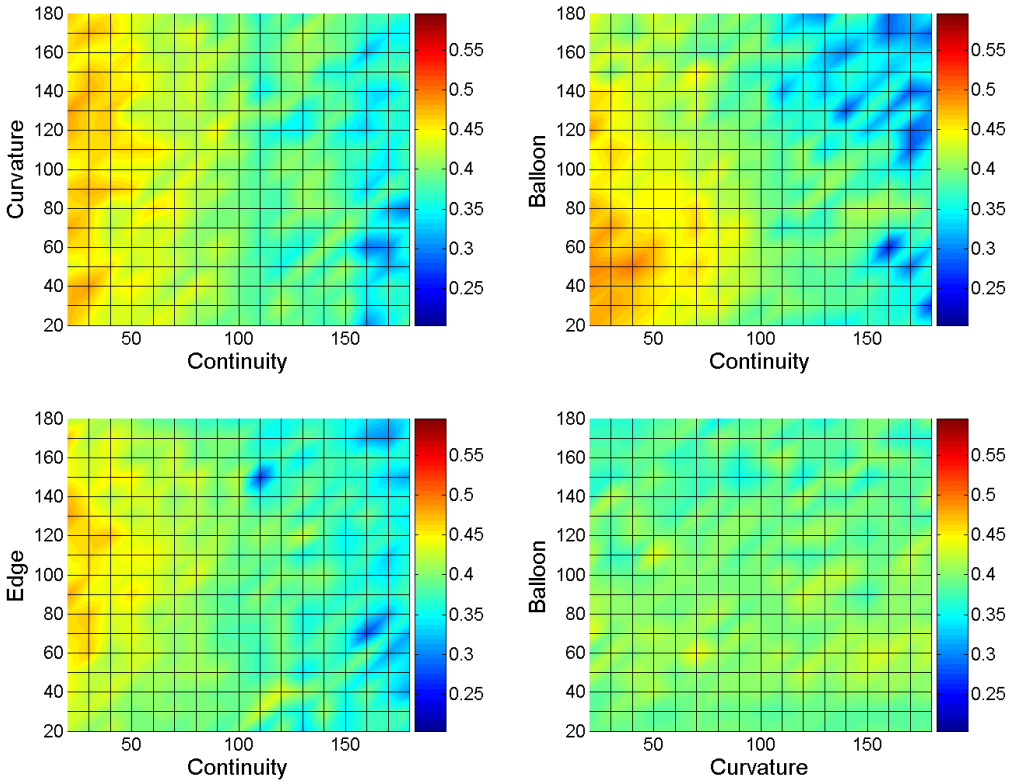
5.2.1.2 GVF Snake

A pairwise grouping of the different terms present in the GVF snake can be found in Table 7.

Table 7 Parameter combinations on optimisation graph plots for the GVF snake

GVF	N/A	N/A	N/A	N/A	N
Balloon	Y	Y	Y	N	Y
Edge	Y	Y	N	N/A	Y
Continuity	N/A	N	N/A	N/A	Y
Curvature	N	Y	N/A	N/A	Y
	Curvature	Continuity	Edge	Balloon	GVF

By increasing the GVF contribution the diffusion term gets more weighting and this does not necessarily yield a better GTC (see Figure 55). As has been previously mentioned the GVF field is a relatively weak force and the parameter optimisation shows that low values of continuity is the greatest contributing factor towards a high GTC. It is worth noting that the original parameter values of 100% were suboptimal and could have yielded a closer match to the manual tracings by lowering the continuity weighting by a factor of 2. At high continuity values vertex clustering dominates and the mesh shrinks causing low values of GTC. In each parameter optimised surface the colour map is 'hotter' than in the scale space snake which implies that over this parameter range the GVF is superior to the scale space snake.



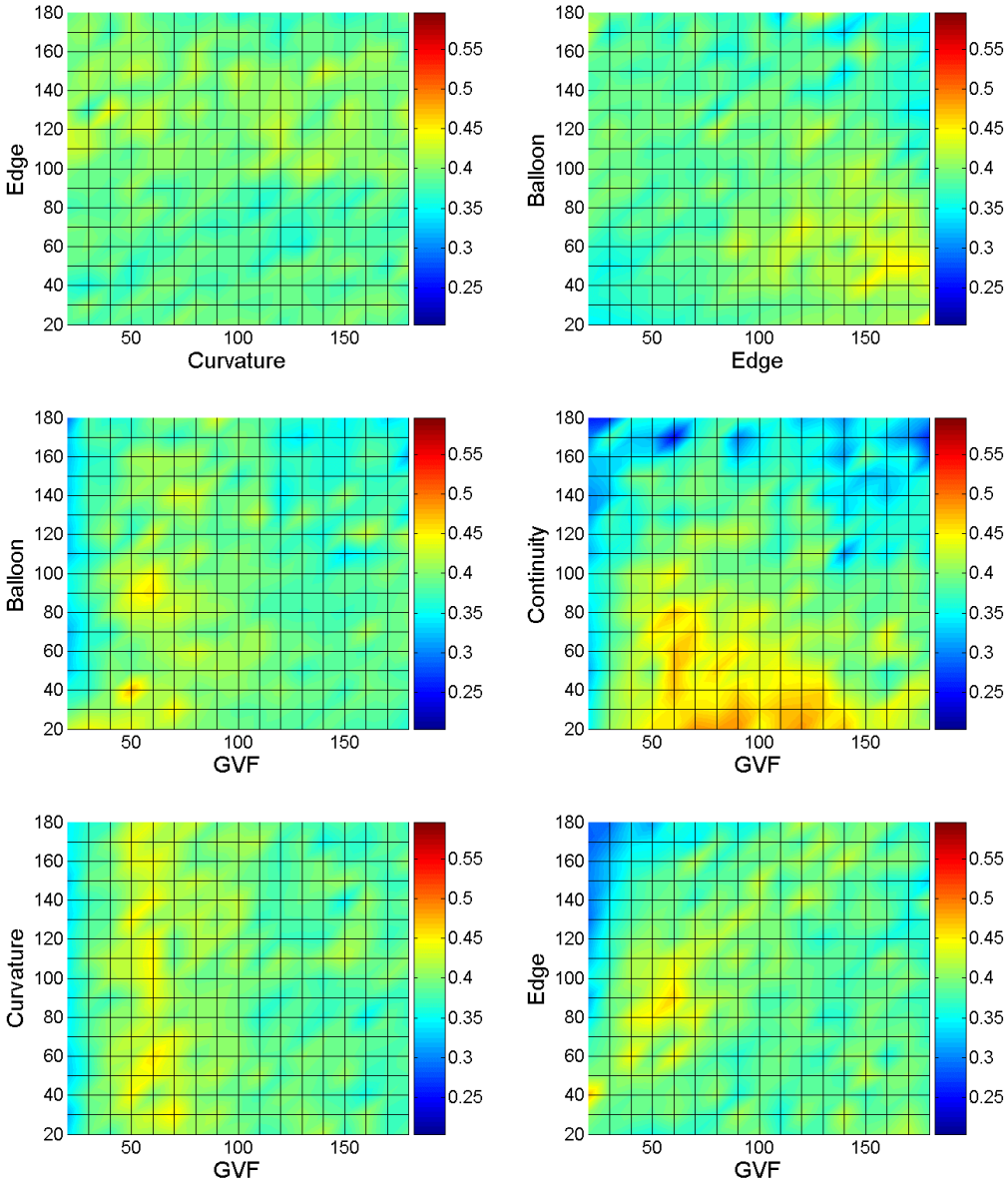


Figure 55 Sensitivity analysis of the weighting coefficients used in the GVF snake model over 80% variation. The parameters are expressed as percentages and the colourbar shows GTC used as the measure of quality of segmentation.

5.2.1.3 PZLS Snake

A pairwise grouping of the different terms present in the PZLS snake can be found in Table 8.

Table 8 Parameter combinations on optimisation graph plots for the PZLS snake

MS	Y	Y	Y	N
Balloon	N/A	N/A	N	N/A
Continuity	Y	N	Y	N/A
Curvature	N	N/A	Y	N/A
	Curvature	Continuity	Balloon	MS

Figure 56 shows an 80% variation in the parameters for the PZLS snake. Continuity is the most sensitive parameter allowing the map to go from dark blue to light red in the parameter optimisation surfaces. In the absence of any mesh interpolation as the mesh expands the vertices get further away and the continuity forces grow significantly.

The MS term is the main driving force and is the dominating parameter that affects the overall GTC. Although this is parameter is not very sensitive compared to continuity it does make the GTC larger than what is possible with the other parameters. At the nominal value of 100% MS term, the colour map shows a slightly unoptimised snake when coupled with the other parameters. A better GTC value can be obtained by lowering the MS term contribution slightly. When compared to Figure 58 (the MS-curvature surface plot) a common trend is a steady improvement in GTC with higher weighting for the MS coefficient. In PZLS case the surface rises more steeply but this could be due to the lack of interpolation in the explicit mesh so the curvature has a greater effect than it does for MSSCD. Also any increase in MS will have a greater influence on the GTC since in PZLS since the mesh vertices become further apart and more able to fill the chamber. In PZLS the Curvature-MS plot, the surface is flatter towards the low end of the MS parameters when compared to the MSSCD case. This is probably because vertex clustering is occurring at lower values of MS weighting in the explicit mesh but this does not occur within the implicit snake. In terms of reducing the quality of the segmentation, the curvature term is very weak within the PZLS snake. At low to medium weighting of the MS term the curvature has very little influence on segmentation and the GTC remains at 0.55 or higher. As the weighting coefficient of the MS term is allowed to increase further above 100% of the nominal value and with the other parameters remaining fixed the segmentation accuracy drops slightly (by a maximum of 0.1). For a given value of the MS term, changes to the curvature over the tested range did not significantly affect the GTC in this snake.

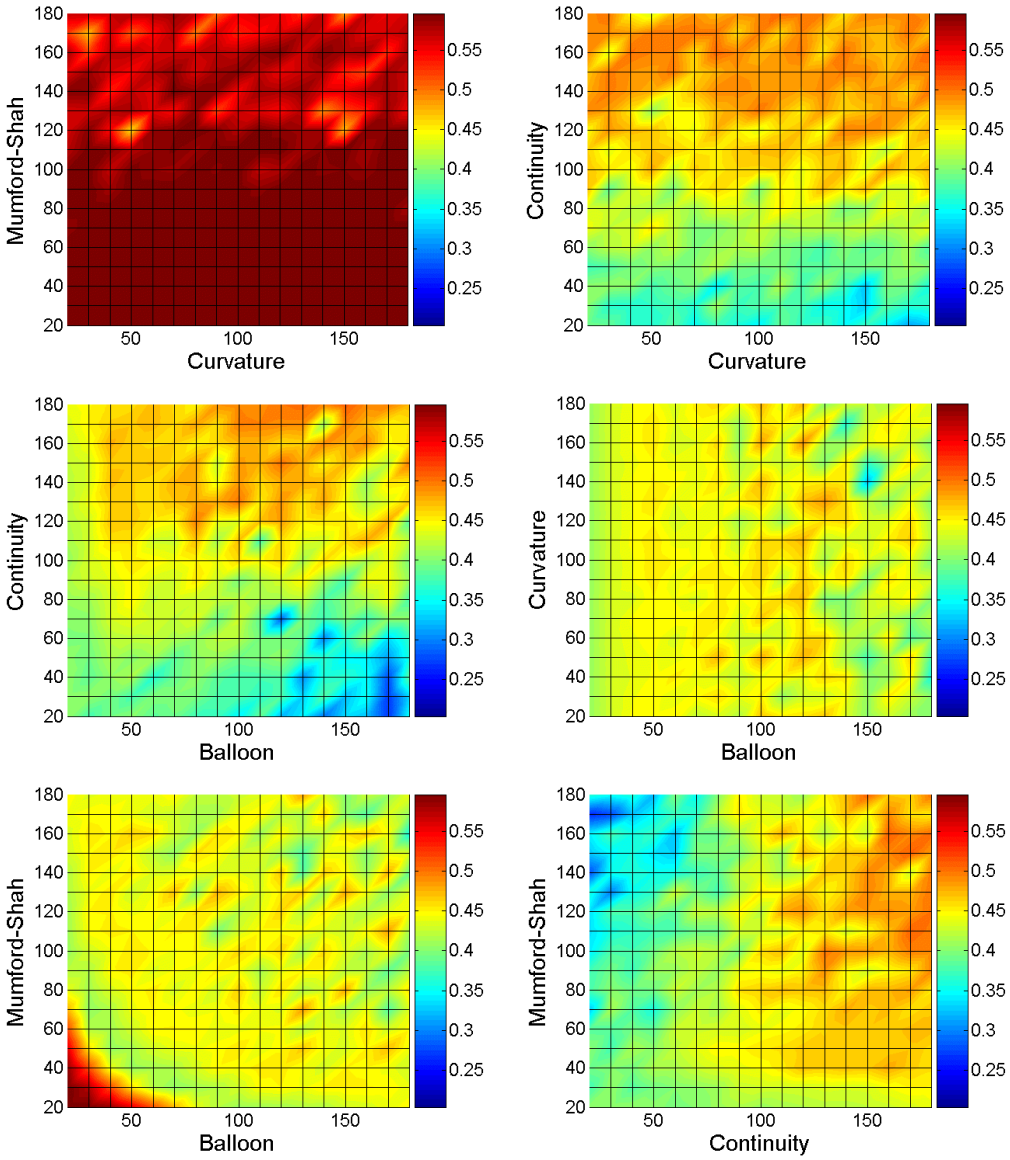


Figure 56 Sensitivity analysis of the weighting coefficients used in the PZLS snake model over 80% variation. The parameters are expressed as percentages and the colourbar shows GTC used as the measure of quality of segmentation.

5.2.1.4 EPCA Snake

The possible combinations of different paired parameters for the EPCA snake optimisation graphs are shown in Table 9.

Table 9 Parameter combinations on optimisation graph plots for the EPCA snake

Balloon	Y	Y	N
Edge	N/A	N	N/A
Curvature	N	Y	N/A
	Curvature	Edge	Balloon

In Figure 57 it would come as no surprise that the balloon term is the most sensitive term in the snake since it is the main driving force. Unlike in MSSCD it is possible for the GTC to approach zero by setting a very small weighting for the balloon coefficient. When this trough occurs in the optimisation surface at low balloon weighting the propagating front of the level set is collapsing to a point under the influence of the internal forces. So over the choice of parameters for the balloon term the GTC ranges from 0 to just over 0.4 which makes this snake the most sensitive on test. Selecting parameters for this snake to run on the datasets was challenging for this snake. Unlike the explicit snakes the surface plots are relatively smooth. This could be due to the presence of self intersections in the explicit mesh that do not occur in the level set framework. Also repositioning mesh vertices around after the mesh has settled may cause some fluctuations to these surface plots.

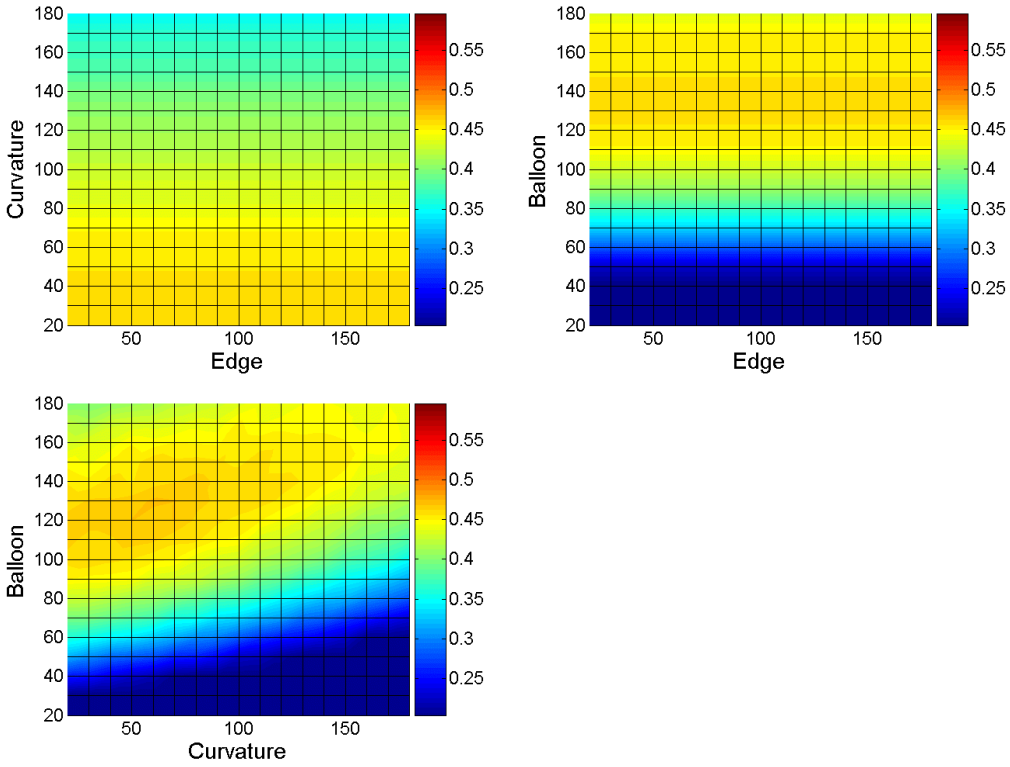


Figure 57 Sensitivity analysis of the weighting coefficients used in the EPCA snake model over 80% variation. The parameters are expressed as percentages and the colourbar shows GTC used as the measure of quality of segmentation.

5.2.1.5 MSSCD Snake

The possible combinations of different paired parameters for the MSSCD algorithm optimisation graphs are shown in Table 10.

Table 10 Parameter combinations on optimisation graph plots for the MSSCD snake

MS	Y	Y	N
Sarti	Y	N	N/A
Curvature	N	N/A	N/A
	Curvature	Sarti	MS

The surface plots for this snake can be found in Figure 58 where it is clear that the GTCs are towards the top end of the colourbar scale, even surpassing the PZLS snake. At this GTC plateau the MS term seems to dominate since it is the main driving force of the snake. The edge advection and curvature terms do little to adversely shift the GTCs making this snake relatively insensitive to a lot of parameters. The main parameter to adjust is just the MS term provided the edge advection and curvature terms are in broad acceptable range. It seems that for this

dataset the MSSCD snake is fairly tolerant of the MS term over this 80% range. Weakening this term by a lot does not adversely affect the GTCs too much since the colours remain warm unlike in the other snakes. It seems that the MS term can be optimised further by increasing it from the chosen starting value on this dataset, although this seems to level off at around a value of 0.65 for the GTC.

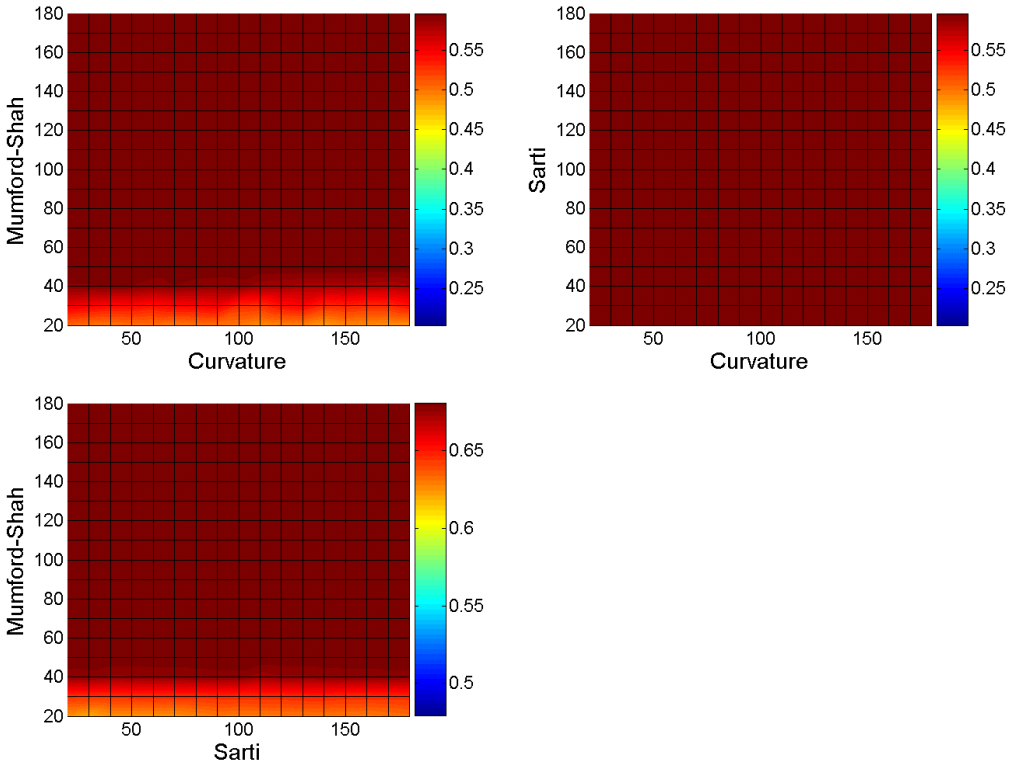


Figure 58 Sensitivity analysis of the weighting coefficients used in the MSSCD snake model over 80% variation. The parameters are expressed as percentages and the colourbar shows GTC used as the measure of quality of segmentation. Sarti's term is edge advection based.

From the results of the surface plots parameter optimisation does not change the overall level of the snake accuracy, that is still determined by the snake type and terms themselves. The results of the surface plots are summarised in Table 11. For example MSSCD operates on GTC of 0.5 or above for this range of chosen parameters, where as the edge based snakes operate below this upper limit. The EPCA snake had the worst performance in parameter optimisation and it was possible for the GTC to approach zero for a large part of the optimisation space as the balloon contribution was weakened sufficiently. Since there are no self regularisation forces as in the explicit snake the volume of the snake defined by EPCA could reach zero if no outward force was strong enough. Scale space because of its multiresolution approach did not vary much with the change in parameters, but it often failed to get a good enough segmentation compared to the other snakes. The dominant edges do not necessarily coincide with the manual delineation. The GVF and scale space

snakes optimisation spaces had blue in them if continuity was too high and this caused very low minimum GTC values in Table 11.

Table 11 GTC dependence on snake over specified parameter range, ranked according to colour temperature

Snake	GTC obtained from parameters	Colour (hotter colours are better, cooler colours are worse)
MSSCD	>0.5	Dark red, red, orange
PZLS	>0.2, <0.6	Red, orange, yellow, green, blue
EPCA	<0.45	Light green, yellow, blue
GVF	>0.3, <0.5	Orange, yellow, green, blue
scale space	<0.4	Blue, green

Whilst the numerical values of the GTCs were not very informative the localised hotspots were useful in deciding how to adjust the parameters for a particular dataset in order to improve the segmentation. The segmented data is presented in the next section based on the findings from these surface plots.

5.3 Foetal Cardiac Segmentation using Amorphous Deformable Models

5.3.1 Visual Inspection of Segmentation

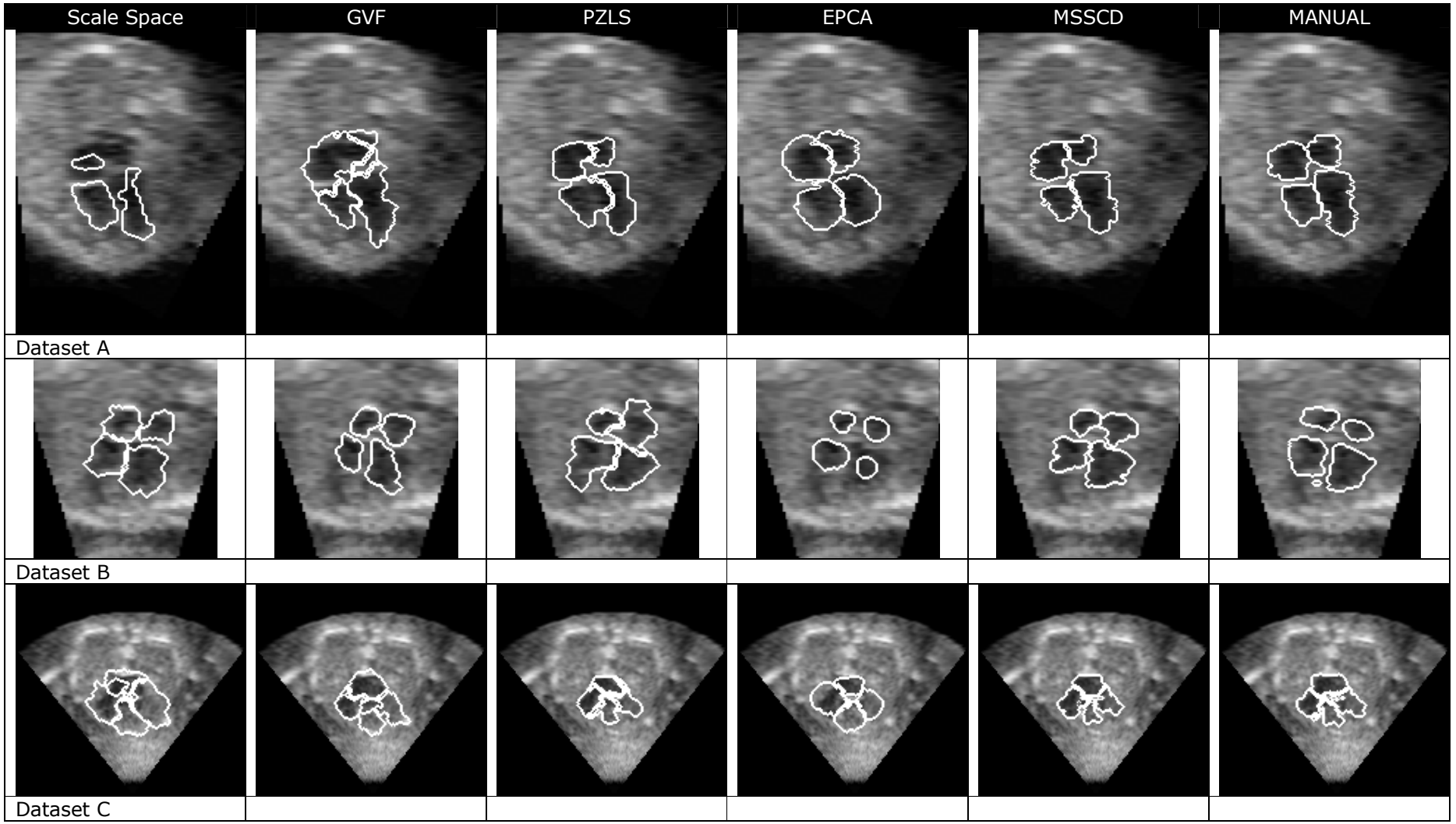
Figure 59 shows some example central slice segmentations for each snake on all datasets (both 3D and 2D). Manual tracings are also shown for comparison. All these images have been reorientated so that LV appears in the bottom left corner. Even though there are only 9 datasets there is a lot of variation between them. No interchamber snakes appear to be overlapping and this is because of the snake overlap correction method as described in section 2.10. The individual contours are filled then rastered onto separate blank images, from the blank images the boundaries of the filled chambers were extracted and were rastered to the greyscale images shown in Figure 59. This crude correction reduces chamber overlap from appearing in the final segmentation and is most beneficial for snakes that did not have a collision detection term.

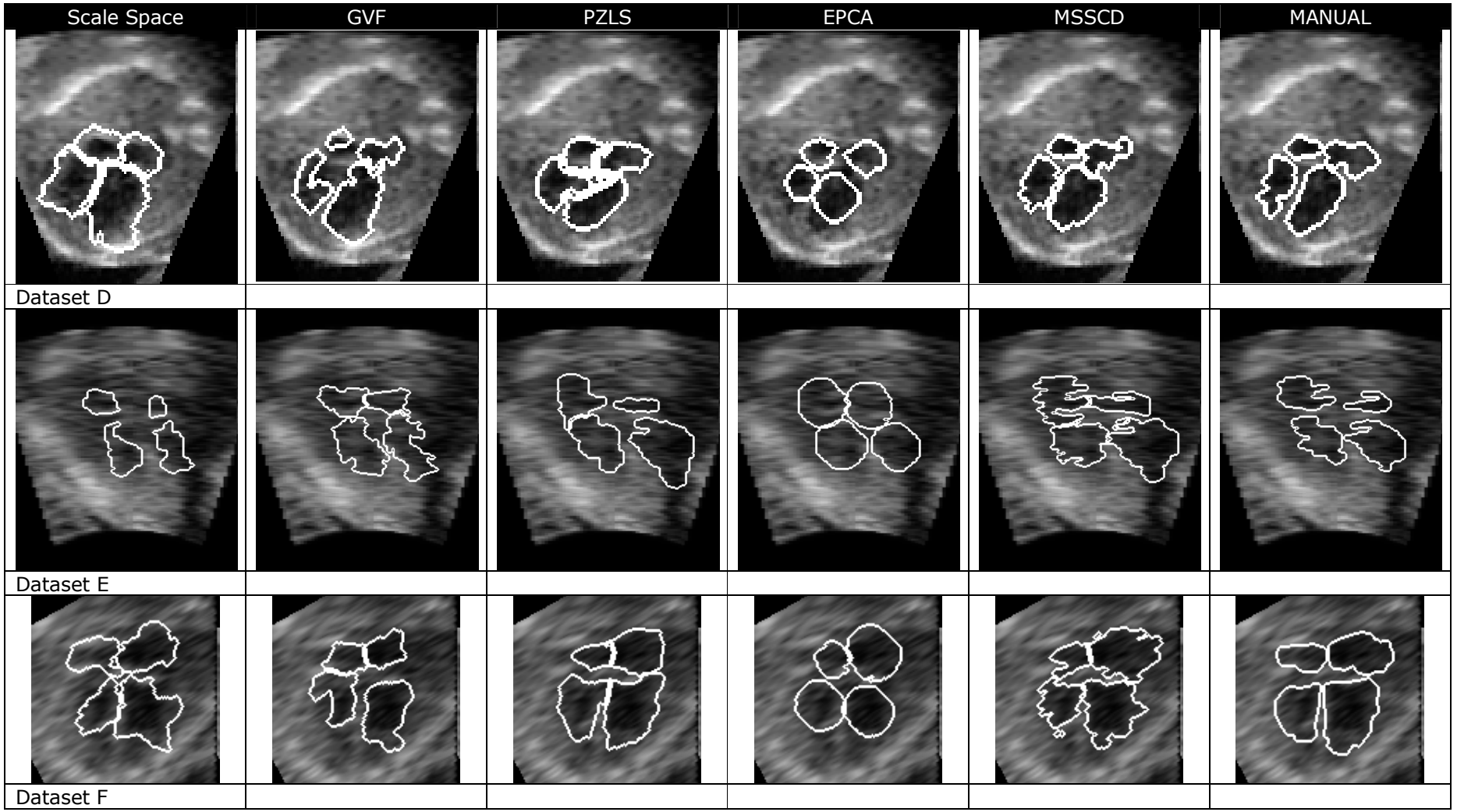
Given sufficient number of iterations EPCA would have grown out of the chambers and across the entire image. It seems that the real cardiac US data do not provide sufficiently strong edges to stop the constant advection term as in the synthetic data. To make the segmentation results more comparable EPCA's evolution was stopped prematurely by manual intervention when the propagating fronts started to spill out of the chambers. This explains the rounded appearance of the EPCA segmentation as compared to the other snake types. In some datasets it could provide very low distance errors to the chamber boundary as a result of this manual constraint.

In dataset H the myocardium is very dark and the the valves are weakly visible to provide an edge for scale space and GVF to be attracted to. As a result their segmentation stops short in the centre of the heart but is defined better at myocardium boundaries. PZLS in dataset H had intra chamber boundary leakage and this is apparent only by the slight bulge of the RA snake into the LA chamber. In dataset A this is more apparent with the EPCA, PZLS and GVF snakes.

The scale space deformable model is strongly intensity dependent and can be seen to be attracted to higher intensities such as above the RA in datasets C, D, G, H, I where it grows out of the chamber. This happens to a lesser extent with the GVF snake for the RA in datasets A and I. If the edge detection term was normalised to the edge magnitude or average edge magnitude in the snake this may reduce the intensity dependence.

Dataset E has shadowing on the right hand side of the dataset and this affects the boundary on the left part of the heart. It is questionable where the manual tracing should be drawn in this region but some attempt was made to follow the curvature of the heart's visible boundary. Datasets A, D and I appear to have the most complete boundary with high blood pool myocardium contrast and should prove the easiest for all snakes to segment.


Scale Space
GVF
PZLS
EPCA
MSSCD
MANUAL
Dataset A
Dataset B
Dataset C


Scale Space
GVF
PZLS
EPCA
MSSCD
MANUAL
Dataset D
Dataset E
Dataset F

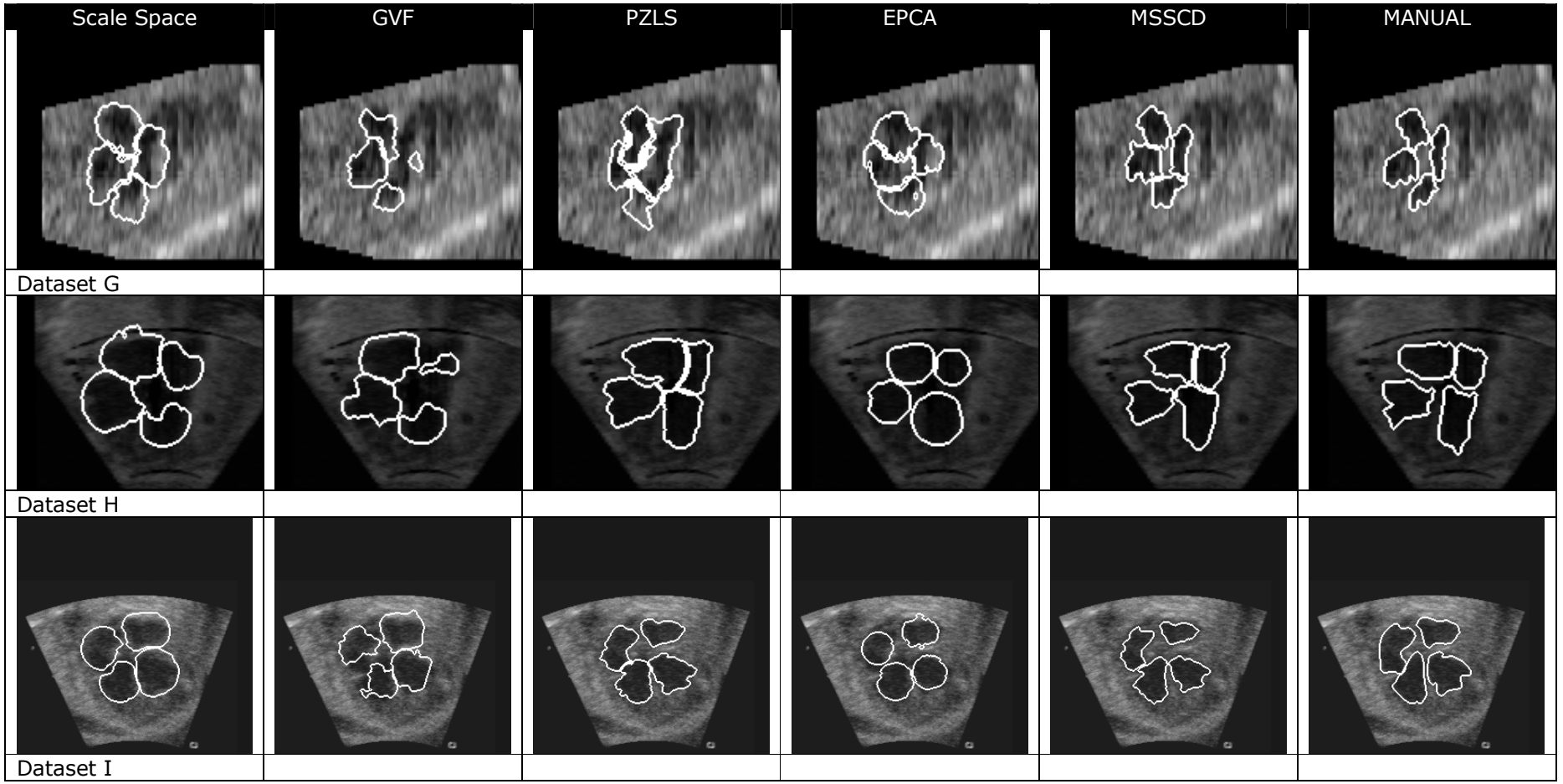


Figure 59 Central slices of datasets segmented with automatic snakes and manual tracings. Datasets A-G are 3D and H-I are 2D. The datasets have been rotated so that the atria are *above* and ventricles *below* with the LV on the *right hand side* of the image.

Two of the 3D datasets (one good and one bad quality) are shown in Figure 60 and Figure 61 respectively. In the better quality dataset it is clear that scale space algorithm has a positive bias on the segmentation. This may be easily corrected if sufficient training datasets were available. The GVF tends towards a positive bias when the edges are strong such as around the endocardium closest to the myocardium where the wall is thicker and better resolved. It also under bias when the walls are thinner and so less resolved hence there are weaker edges such as the atria, valve plane and septum. The PZLS algorithm since it is region based does not rely on edge strength therefore can pass through weak edges if the greylevel on the other side of the edge is similar to that inside the snake (an example of this is the edge between neighbouring chambers). Thus the PZLS leaks between chambers and this is indicated by the heavy white region at the centre of the heart in Figure 60. EPCA was stopped prematurely to prevent overspilling out of the cardiac chambers hence the difference image for this snake appears with very rounded black structures. It seems MSSCD provides the best segmentation since the thickness of the discrepancy is less than for the other snakes apart from the LA where part of the vascular tubes enter the atrium. It appears that MSSCD has a negative bias closer to the ventricle apices and slightly positive elsewhere. It crosses the endocardium where the septum lies but is met by the snake on the neighbouring chamber causing the collision detection to halt both snakes.

On the poorer quality dataset in Figure 61 there is more noise inside the chambers and this makes it difficult to distinguish the atria in this slice. This is a common problem for each of the snakes and a good manual segmentation is also very difficult to achieve. The edge based snakes seem to not reach the shadowed region on the LHS of the LV and this is indicated by blackness on their difference images. MSSCD and PZLS do a better job than the explicit edge driven snakes and is difficult to assess the quality of their segmentations visually since both have regions of positive and negative bias. EPCA was stopped prematurely but its difference image does not look as bad as in Figure 60.

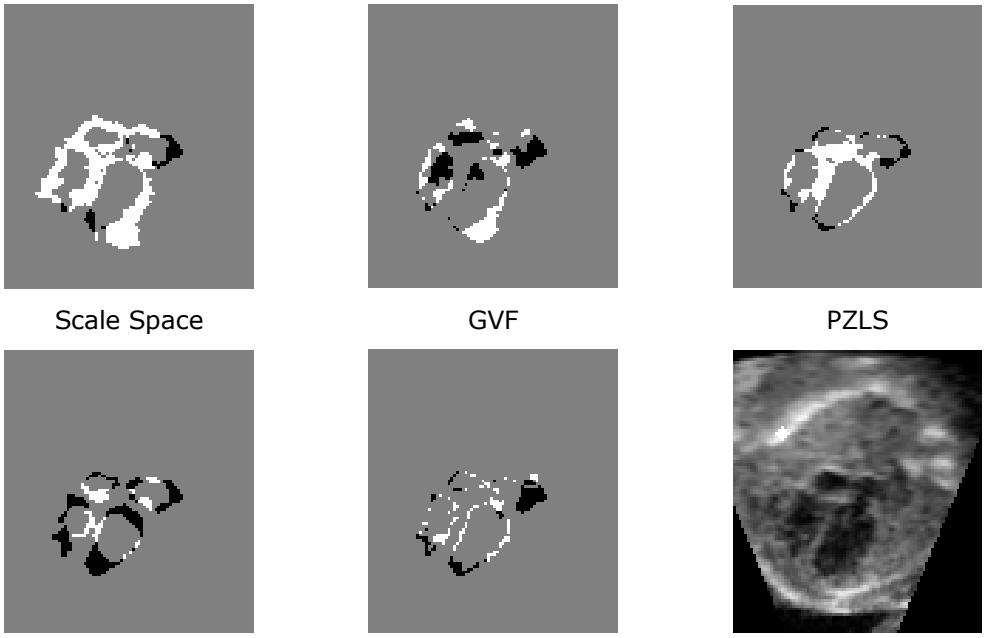


Figure 60 Automatic-Manual segmentation difference images for each snake on one of the datasets. *White* is positive, *black* negative and *grey* is zero. The quality of this dataset is very good since there is minimal shadowing.

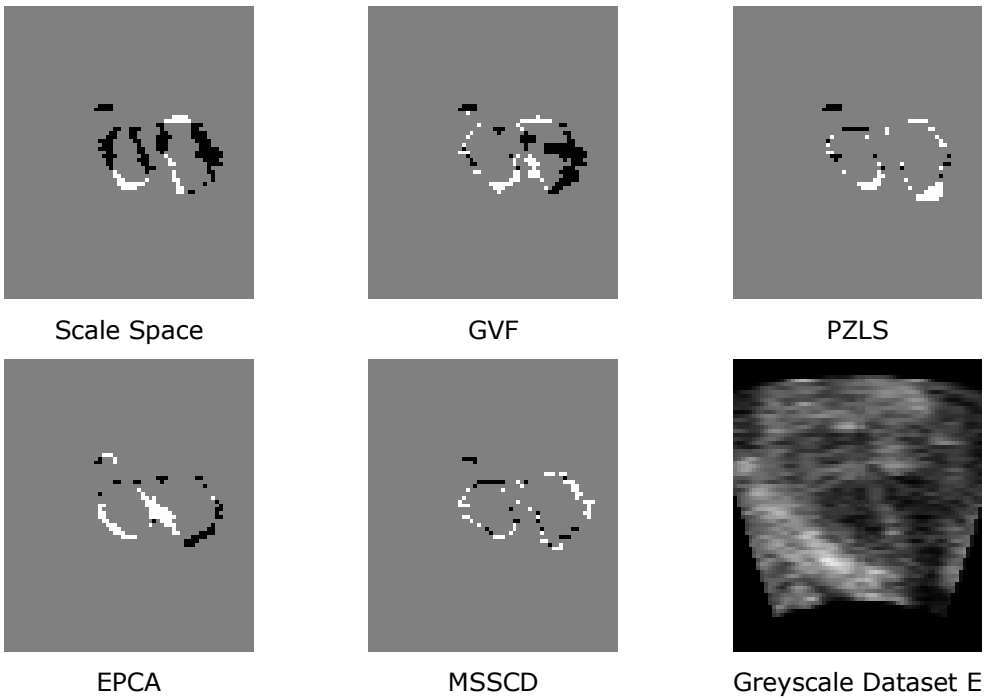
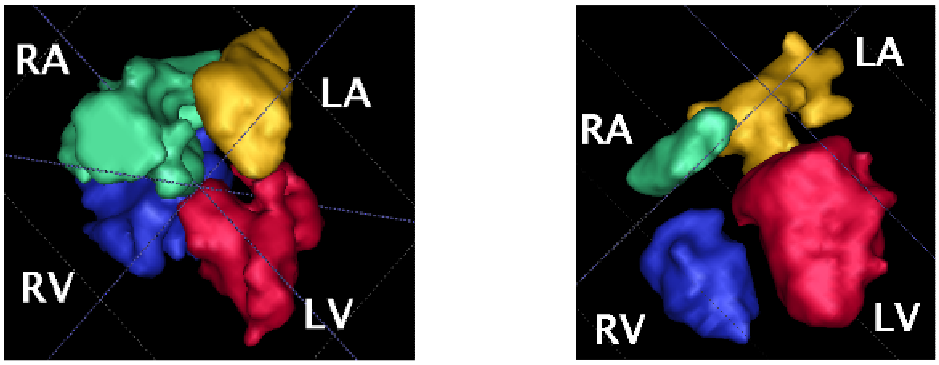
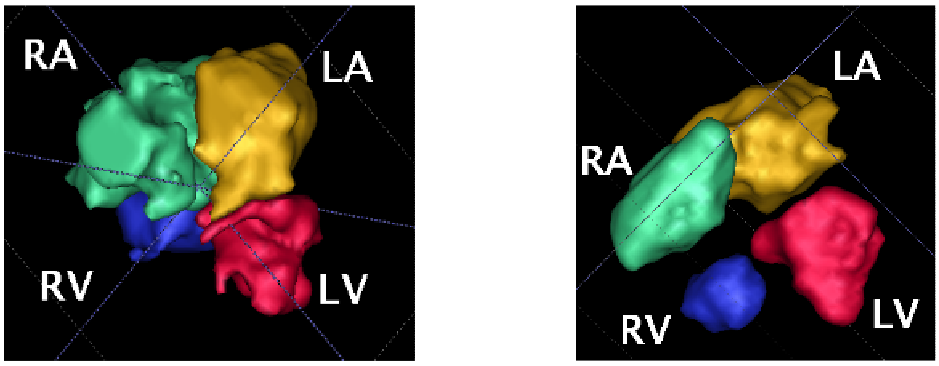


Figure 61 Automatic-Manual segmentation difference images for each snake on one of the datasets. *White* is positive, *black* negative and *grey* is zero. The quality of this dataset is poor since there is shadowing coinciding with part of the chambers on the left part of the heart on the *right side* of image.

Example surface renderings of two automatically segmented datasets by the MSSCD algorithm are shown in Figure 62. The rendering was provided by itk-SNAP [131]. ED and ES refers to End Diastole and End Systole respectively.



Example 1 ED Example 2 ED



Example 1 ES Example 2 ES

Figure 62 Example surface renderings of two segmented 3D datasets during End Diastole (ES) and End Systole (ES).

5.3.2 Distance Errors

The rms and mean distance errors of the automatic delineations averaged over the entire heart are shown in Figure 63 and Figure 64 respectively. The error bars represent the standard error. It is assumed that all segmentation results are independent when computing the error bars to show the standard error over the sample e.g. speckle does not influence one dataset more than another. From these two figures the errors are larger in 3D than for 2D by a discrepancy of 0.5mm or less. ES does slightly better than ED probably because in ED the valves are open and it is questionable where the atrium-ventricle boundary really lies. The trends in both figures are very similar since they are both distance metrics computed from the same point correspondence. The rms errors are slightly higher than for the corresponding mean errors but the difference is less than 0.5mm. The only exception to this is for PZLS and GVF in the 2D ED graph where PZLS is better than GVF in the mean distance metric but not according to rms.

The errors are less than 2mm for all snakes and given that the length of a foetal heart is about 20-30mm in the datasets measured these errors are under 10%.

Overall the worst snakes in terms of these metrics are the GVF and scale space and the best deformable model is the MSSCD which is followed by the PZLS algorithm. Since PZLS lacks the collision detection, leakage between mainly atrial chambers limit this algorithm's accuracy in some of the datasets. The EPCA does a very good job at keeping the errors low compared to GVF and scale space but the iterations in this algorithm were truncated to prevent it over spilling out of the chambers due to the constant advection term.

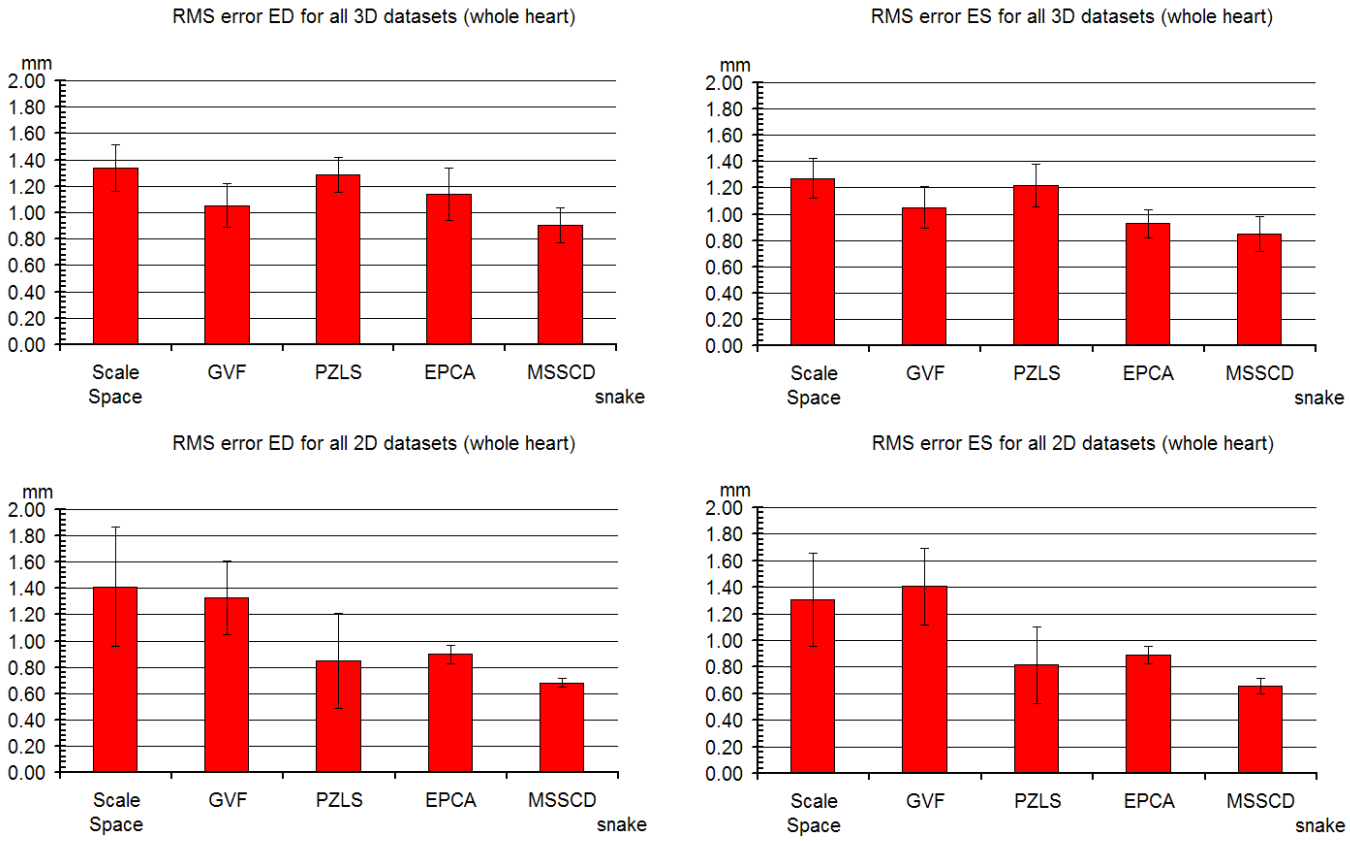


Figure 63 RMS errors for 3D (*top row*) and 2D snakes (*bottom row*) averaged over all the datasets. Results for ED are shown in graphs on *left* hand side and ES on *right* hand side. *Lower values are better.*

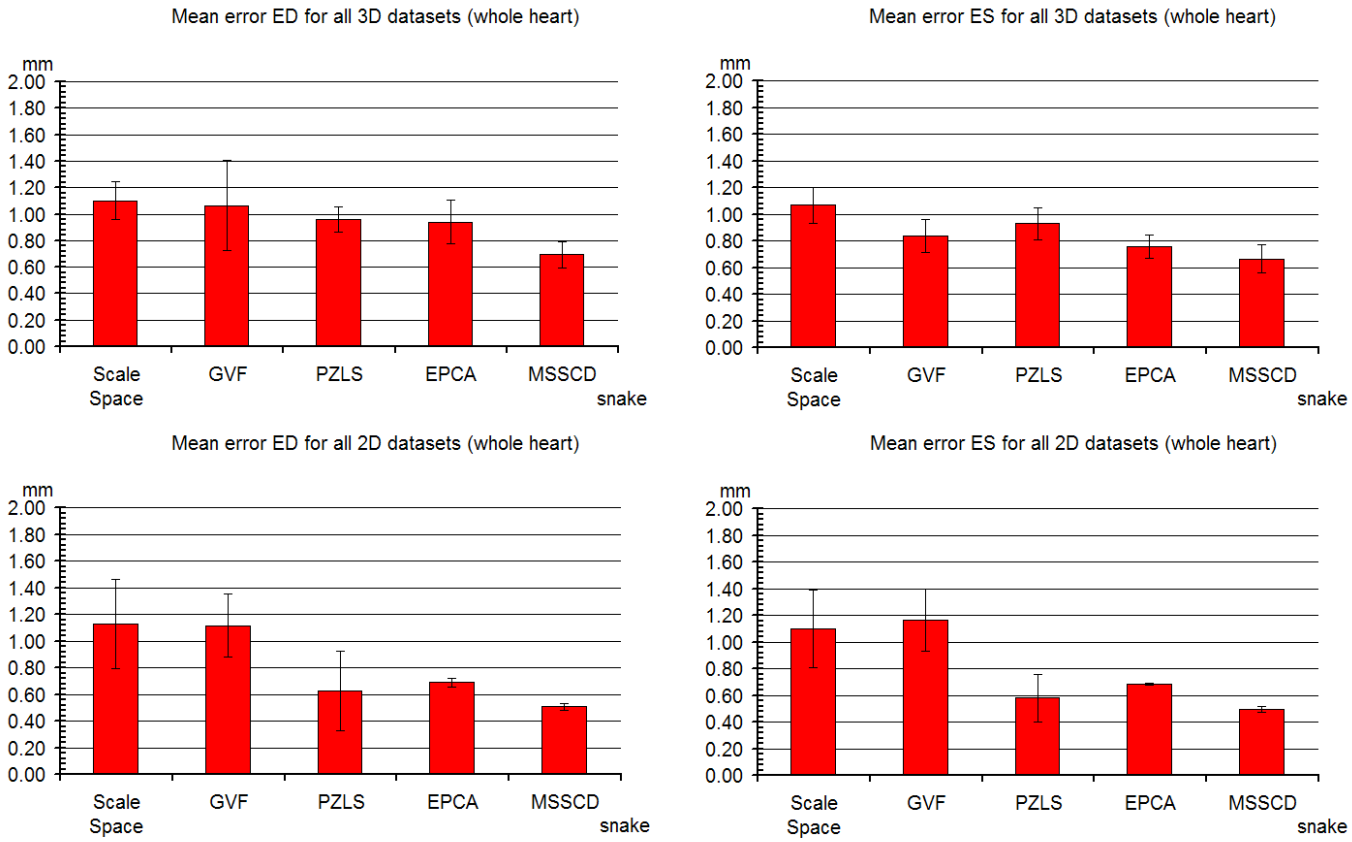


Figure 64 Mean errors for 3D (*top row*) and 2D snakes (*bottom row*) averaged over all the datasets. Results for ED are shown in graphs on *left* hand side and ES on *right* hand side. *Lower values are better.*

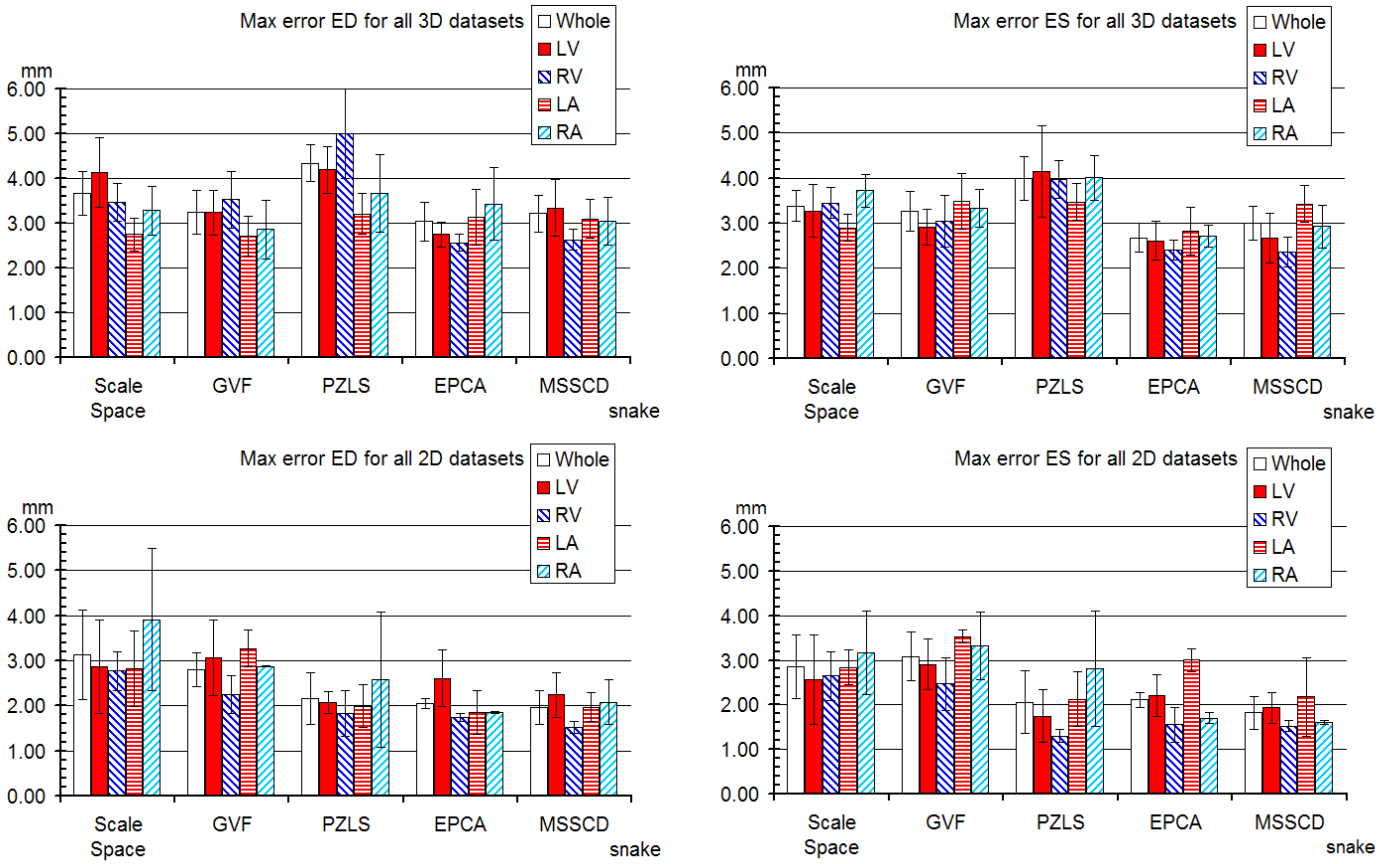


Figure 65 Max errors for 3D (*top row*) and 2D snakes (*bottom row*) averaged over all the datasets. Results for ED are shown in graphs on *left* hand side and ES on *right* hand side. *Lower values are better.*

Figure 65 shows the maximum distance errors from each algorithm's delineation to the manual tracings for individual chambers and averaged over the entire heart for both ED and ES cardiac phases. In rms and mean errors ED had higher errors than for ES. In Figure 65 which show graphs of maximum errors this distinction is more difficult to spot. In 3D the PZLS, EPCA and MSSCD have the lowest errors but in 2D where the spatial resolution is much better, GVF and scale space are just as good. In Figure 65 for 2D datasets the RA shows the longest error bar. PZLS has the worst maximum error bar in 2D (~1.5mm) where the MS term causes leakage across atrial chambers due to lack of collision detection (see dataset H in Figure 59). The error bar lengths in scale space are almost as large in 2D as those in 2D PZLS. In dataset H there is some local enhancement above the RA and this attracts mainly the edge based snakes. If there were more than only 2 datasets available for the 2D segmentation then this error bar would be smaller since it is dependent on the standard error. In the 3D datasets there was no obvious long error bar in most of the snakes, although since shadowing was present in the left hand side for at least one dataset the LV has a fairly large error bar in both PZLS and scale space snakes. In 3D there are more datasets and so the standard error is better defined than in 2D. If there were more available 3D datasets there may be a chamber that shows up in dominating the maximum errors. In both 2D and 3D the maximum error of EPCA algorithm is low due to forced truncation of iterations. The MSSCD is comparable to EPCA in both 2D and 3D. The absolute maximum error is around 4 or 5mm and this corresponds to about 20-25% error for a 20mm foetal heart. For a larger foetal heart of about 30mm the percentage error would be smaller.

5.3.3 Cardiac Cycle

One of the datasets was fully manually segmented to reveal the cardiac cycle. The snake algorithms were applied to each frame and the volume time curves can be seen in Figure 66. It appears that the cardiac cycle is truncated in this sequence since the volume does not rise sufficiently to meet the diastolic volume at frame 0. It is obvious that PZLS and scale space have the most positive bias in the segmentation volumes. In PZLS this was predominantly due to inter atrial and inter ventricle leakage for the partially resolved septum. Scale space overspills in the majority of datasets including B, C, D, F, G, H and I. If the volume bias is consistent over a large number of training data these systematic errors can be corrected for each snake individually. EPCA has the most negative bias but this was because of premature evolution termination. GVF and MSSCD seem to closely follow the trend of the manually defined volumes over time.

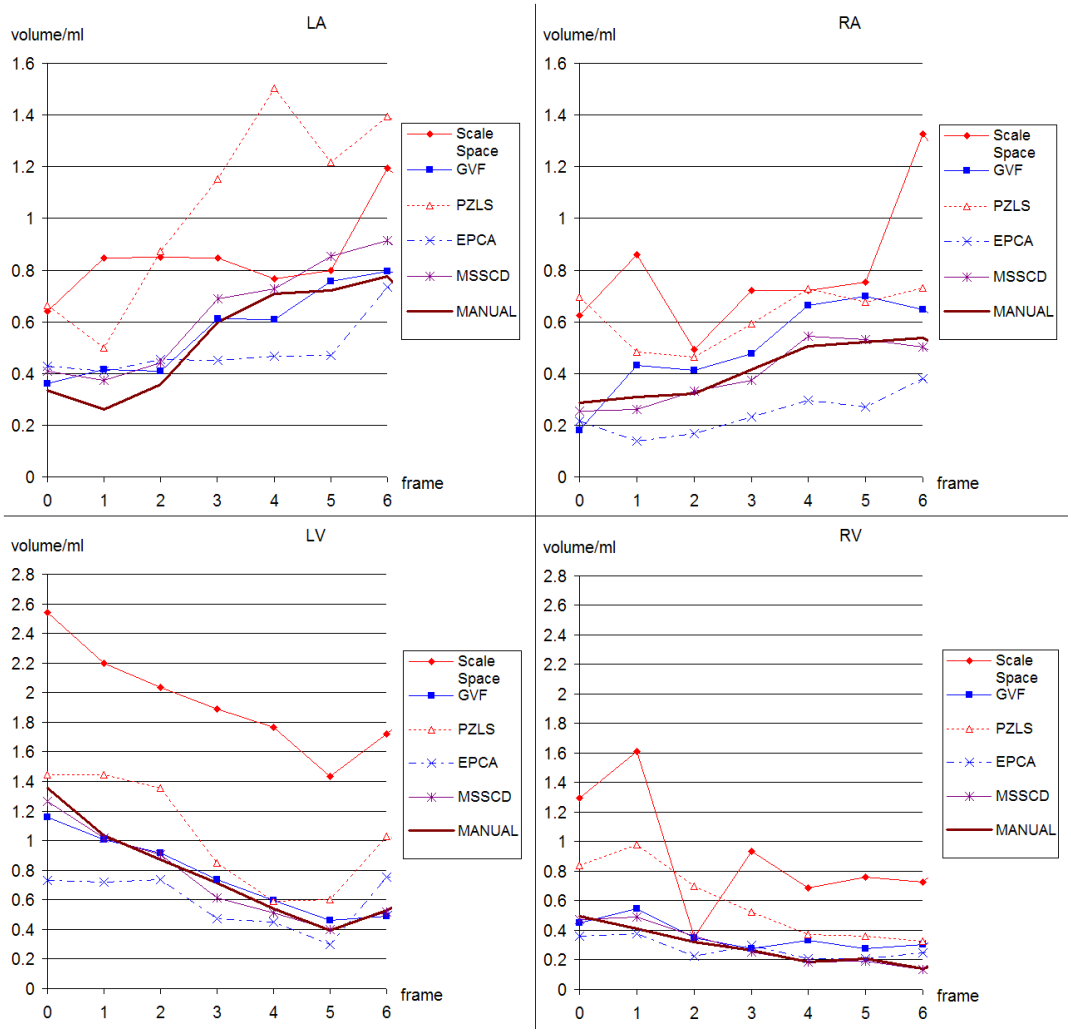


Figure 66 Cardiac cycle shown for one of the 3D datasets.

The change in ventricular volumes seems to be in antiphase to the fluctuation in atrial volumes during the cardiac cycle and so show the expected trend. The next section numerically compares the ejection fraction as a measure of the cardiac cycle for each snake.

5.3.4 Ejection Fraction

The LV and RV EFs were computed for both manual and automated delineations using the 3D datasets. The mean and standard error in the mean are presented in Table 12 and for comparison Esh-Broder's measurements using manual tracings for 20 datasets are presented in Table 13. Esh-Broder used a SonoREAL 3D ultrasound image reconstruction system from tracked 2D slices on both HDI 5000 and HDI 3500 (ATL, Seattle, WA, USA) scanners [2]. The obvious observation is that the EF results presented in this thesis (including manual tracings) are smaller than those in [2].

Even allowing for error bar overlap the manual tracings from two sources do not agree or are just within the bounds. However, the protocol for the segmentation results presented in this thesis excluded the space behind the papillary muscles (less visible in ES) since the partial volume effect made it difficult to distinguish chamber from papillary muscle in some cases. This would cause a reduction in the EF for the datasets presented in this thesis. To some extent this is valid since the blood can not occupy the physical space taken up by the papillary muscles. In [132] foetal cardiac ventricular mass was determined from the segmentation and this involved including the papillary muscles as part of the myocardium and separate from the blood pool. For comparison to the results in this thesis the four normal hearts segmented by using an interactive method in [128] using RAB 4-8 transducer, Voluson 730 Expert ultrasound machine (SonoAVC, 4DView version 7.0 software; GE Healthcare, Zipf, Austria), produced ejection fractions that were between 49% and 58%. These values are between those in Table 12 and Table 13.

Table 12 Mean Ejection Fraction averaged over all the 3D datasets. The standard errors in the mean are also shown.

Chamber	Scale Space	GVF	PZLS	EPCA	MSSCD	Manual
average LV	41.0 ± 3.3	39.6 ± 5.5	35.0 ± 4.1	26.6 ± 3.5	42.9 ± 2.8	43.0 ± 3.3
average RV	22.6 ± 4.5	39.0 ± 5.5	42.5 ± 3.7	17.0 ± 4.4	47.2 ± 2.6	49.4 ± 3.1

Table 13 Mean Ejection Fraction across all datasets presented in [2]

Chamber	mean	min	max
LV	57.5	49.1	65.9
RV	54	48.0	60.0

The MSSCD snake was closest in matching the manual tracings and associated errors with PZLS just behind in the case of the RV and slightly lower EF for the LV. In some of the datasets the shadowing from the foetal ribs was predominantly on the left hand side and so it was difficult to segment the LV in some cases for both automated and manual methods. EPCA is not representative of the boundaries since in a lot of the 3D cases the iterations were halted prematurely to prevent over spillage out of the boundaries.

The GVF offers the next most consistent set of results with both LV and RV sharing the same errors and mean although the means are noticeably lower than the manual values. This is quite surprising since the GVF struggled to occupy half the chamber in some of the segmentations due to weak boundaries and open valves, although in the others with stronger edges the GVF seemed to grow out of the chamber. There may be sufficient overestimation of volume in the datasets with strong edges to raise the

EFs sufficiently to oppose the cases where the GVF was unable to fill the chamber, and thus produce comparable values with the manually determined EF. Scale space forms consistent LV EF to the manual segmentation but for the RV it is about only half the value.

5.4 Inter Observer Validation

Two observers (one of which is a clinician specialised in foetal hearts) were given the task to segment several slices and cardiac phases on 3D datasets from those presented in this thesis. The total number of image slices segmented by either observer was 36. From these manually defined sets of contours the rms, mean and max distance errors are shown in Figure 67. The GTCs of the regions from both observers were computed and are presented in Figure 68. There is a lot of variability between the two observers, mainly because there were inconstancies adhering to either inclusion or exclusion of the partially resolved papillary muscles, cardiac vessels, atrial inlets, open valve areas and interpreting the missing wall between atria. It appears that the strongest anatomical features were followed for manual segmentation which may not be consistent in all datasets. The LV shows the most variability most likely due to some shadowing in some of the datasets on this side and the partial volume effect amongst the papillary muscles. This is also reflected in the graphs in Figure 67 and Figure 68. The max averaged discrepancy is 2.5mm or below, which is 13% or less error for a 20mm length heart. The mean distance error over all hearts corresponds to within half of this percentage error.

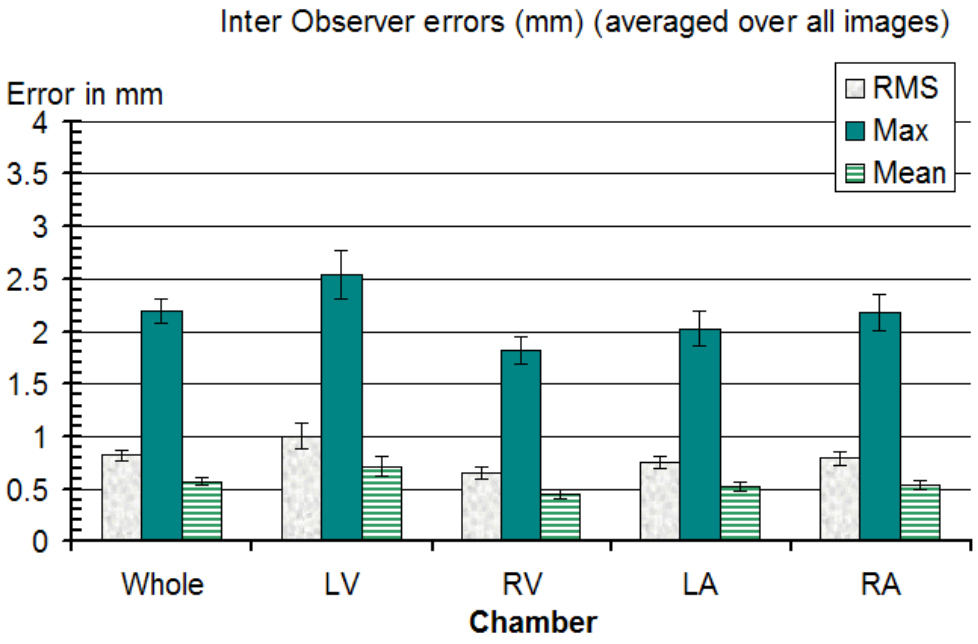


Figure 67 Inter observer distance errors measured on the same datasets. The error bars show standard errors.

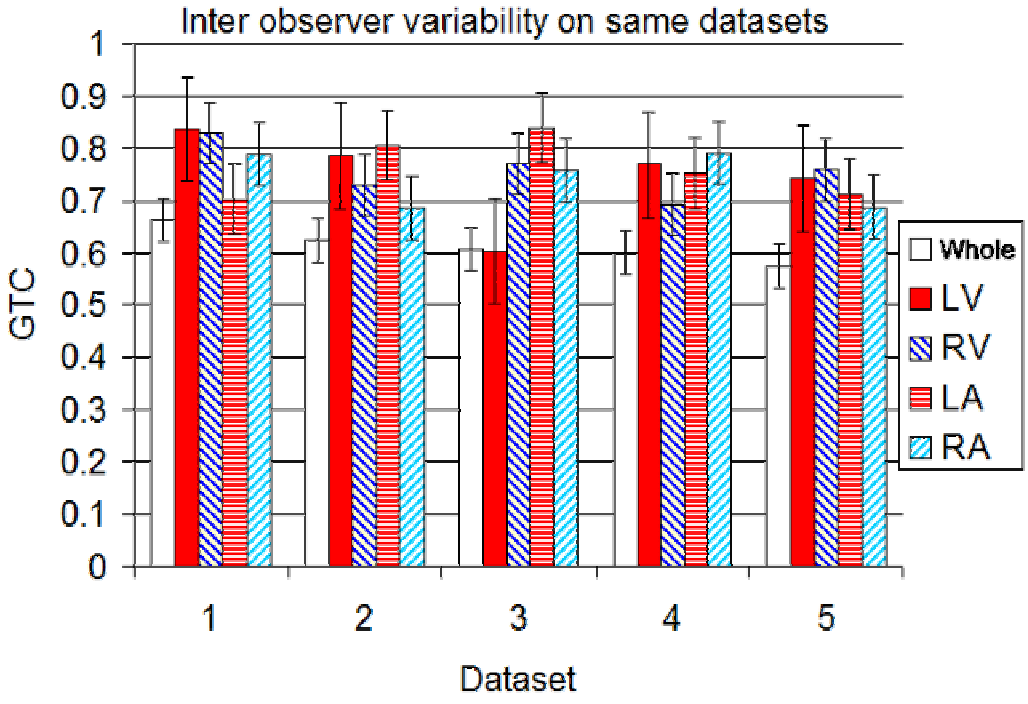


Figure 68 Inter observer variability on the same datasets. The columns refer to the average GTC over several slices and the error bars refer to the standard error.

Figure 69 shows the various GTCs for each snake in 2D and 3D datasets and with the inter observer segmentation results for comparison. In both 2D and 3D the MSSCD snake comes out with the highest GTC although in 2D the PZLS is not far behind.

MSSCD has amongst the lowest standard error in the GTC. This could be due to the stability of the collision detection term since the PZLS although it shares the same region term and has a good GTC, has a relatively high standard error.

On the whole the error bars are lower in 3D since there were more available 3D datasets compared to 2D ones and since the error bars show the standard error of the mean, increasing the number of datasets will reduce the standard error by $\sqrt{n}$ where n is the number of datasets. This follows from the definition of standard error of the mean as given by

$$SE = \frac{s}{\sqrt{n}} \tag{5.5}$$

where s is the sample standard deviation.

It seems that the scale space and GVF swap in the rankings from 2D to 3D. On the 2D dataset both segmentations are quite far from the manual delineation but the GVF collapses more. It is possible that modification of the parameters would permit less collapse and a superior segmentation from the GVF over the scale space snake. Since there are fewer 2D datasets used in this study few conclusions can be made on direct comparison of 2D with 3D.

It is worth noting that the inter observer variability is better than the 3D but not 2D segmentation; this could be due to the presences of less artefacts in the 2D images than in 3D (the inter observer tests were made on slices taken from the 3D datasets only).

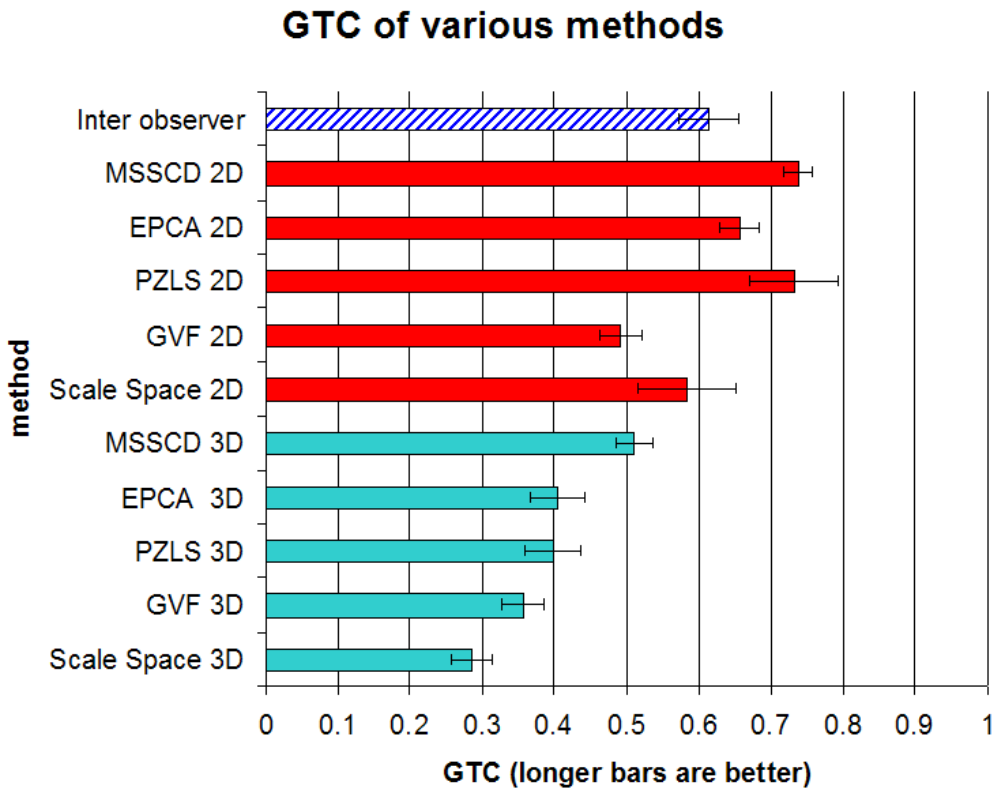


Figure 69 Average GTC of various methods compared for both 3D and 2D. The inter observer error is computed from 36 slices taken from 5 datasets. On this graph longer bars are better. The error bars represent standard errors.

5.5 Validation Results of Snake Algorithms using a Physical Phantom

A grape of comparable size to the foetal heart was scanned in ultrasound gel (details of the experiment can be found in section 4.4). Its physical volume was determined using Archimedes' principle. The deformable model algorithms were applied to determine its volume from the scanned images.

The phantom appeared to have more attenuation than the cardiac datasets possibly due to the high reflectivity of the grape skin. Although microwaving reduced its smoothness some attenuation is visible in the lower right corner of the grape in Figure 70. In this region the snakes particularly GVF and MSSCD showed difficulty in latching onto a boundary. EPCA managed to find this boundary but only after a lot of iterations. This snake had difficulty distinguishing between the interior and exterior of the grape. The balloon force in EPCA could be set so that both walls were picked

out by the zero levels (as in Figure 70) or grow out of the grape into the background since it was difficult to adjust the weighting coefficients to make EPCA stop at the grape outer surface without growing past it. The GVF and MSSCD algorithms had less trouble with this task; all snakes were initialized at slightly larger than in the previous datasets so that they could grow out of the grape's dark interior. It is surprising that PZLS is not able to overcome the dark region of the grape interior like the MSSCD snake even though they share the same MS term. It is possible that the sphere initialisation for the MSSCD samples higher intensities than the cylindrical primitive initialisation for the PZLS algorithm and this may provide a higher starting interior mean for MSSCD than in PZLS.

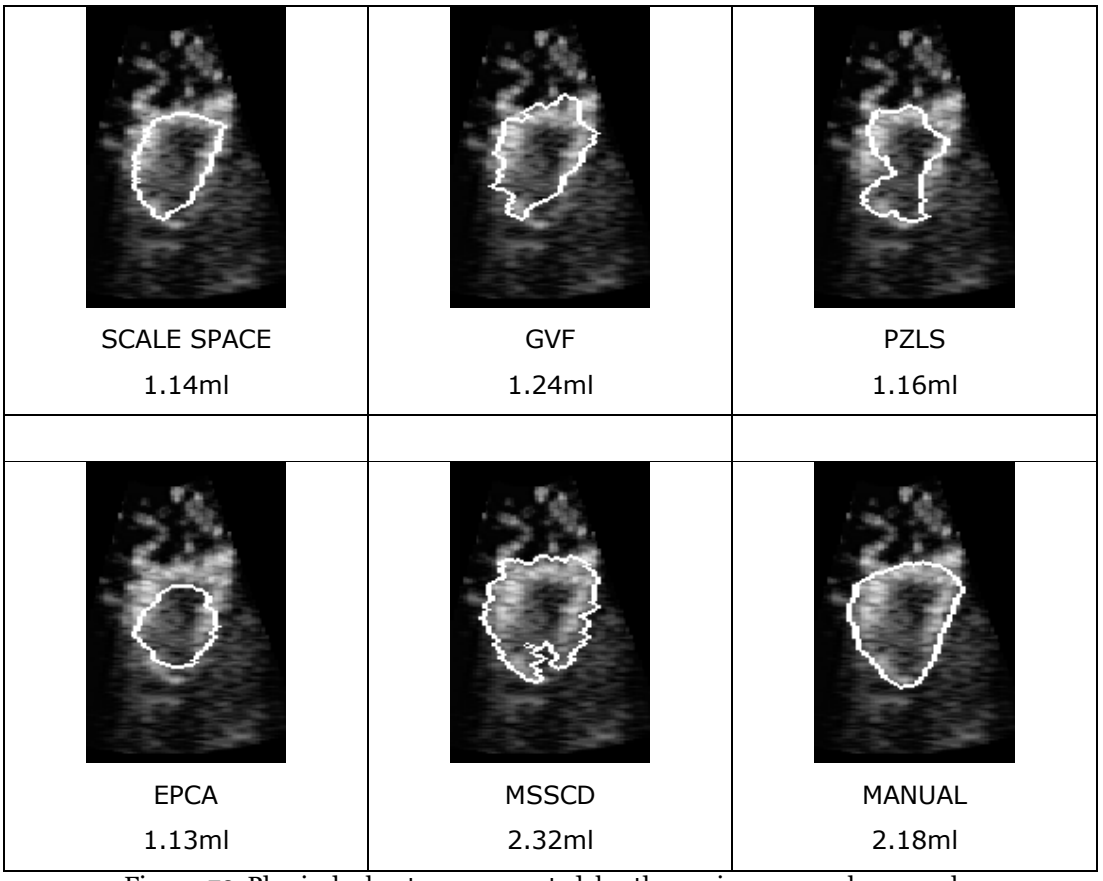


Figure 70 Physical phantom segmented by the various amorphous snake algorithms and by manual segmentation.

The physical volume of the grape was measured using a measuring cylinder to be 2.5±0.5ml. Only the MSSCD estimation of volume agrees with the physically measured one within the boundaries of experimental error. The mean rms error between the MSSCD and manual tracings is 2.1mm. All the snakes underestimate the required volume of the grape but since the grape's volume was measured after the scanning it is possible that some ultrasound gel residue remained on the skin before submerging into the measuring cylinder. Without this residue the displaced volume of water would probably be closer to that estimated by the MSSCD snake.

The hand segmented snake contours give a value of 2.18ml which is consistent with the physically measured volume so there is some validity made towards the manual tracings.

5.6 Foetal Cardiac Segmentation using Non-Amorphous Deformable Models

This section presents some of the results from my publications [70], [71] in which the mean image from manual segmented data was used to initialise and constrain the level set snakes. The main results from this work are presented below.

Inclusion of a shape prior term is bound to be more robust to noisy data; so to make the comparisons between the different snakes fairer, the same shape prior term is included as part of the snake equation for all level set snakes that have been applied to the foetal heart. The results presented below are taken from my publication [71] where the shape prior (SP) term (implemented as described in sections 2.6, 2.7, 2.8) was applied to EPCA, MSSCD and TIMS in turn. The shape prior was constructed using 3 datasets, but from these a total of 26 2D images were used. The pixel size in the datasets varied from 0.3-1.46mm. A leave one out method on the library of images is used to ensure that the template did not include the image that would be segmented by the snake. Some example segmentation results of the three snakes are shown in Figure 71. Without the shape prior the EPCA snake completely overshoots the boundaries due to its constant advection term. The MSSCD snake appears to perform similarly both with and without the shape prior term. Its collision detection abilities appear to handle the weak boundaries in-between cardiac chambers well. With the shape prior term all three snakes appear to respond to boundary drop outs adequately. When the prior is turned off in the TIMS algorithm the cardiac boundaries can be segmented but the snake punches through the weak walls between chambers due to its lack of a collision detection term. For comparison the shape prior is registered to the cardiac chambers in Figure 71. From this template fit it is clear that the individual chamber contours appear distinct. However, the rounded appearance of the template due to its construction from the mean does not allow it to capture the shape of the chambers. Therefore the snake algorithms need to be evaluated with the template used as a shape prior to better conform the shape of the hearts that are dissimilar to the mean.

Even with the presence of the SP constraint, the constant advection force in the EPCA snake overshoots most of the cardiac boundaries but broadly retains the shape of the prior. The MSSCD snake was started from a small seed circle in the centre of the chambers and not from the template as in the TIMS snake. It was probably less

constrained by the template to take the shape of the chambers more readily than the TIMS snake. The TIMS+SP segmentation appears to be more rounded than for MSSCD+SP term since it lacks edge flow contribution from Sarti's term.

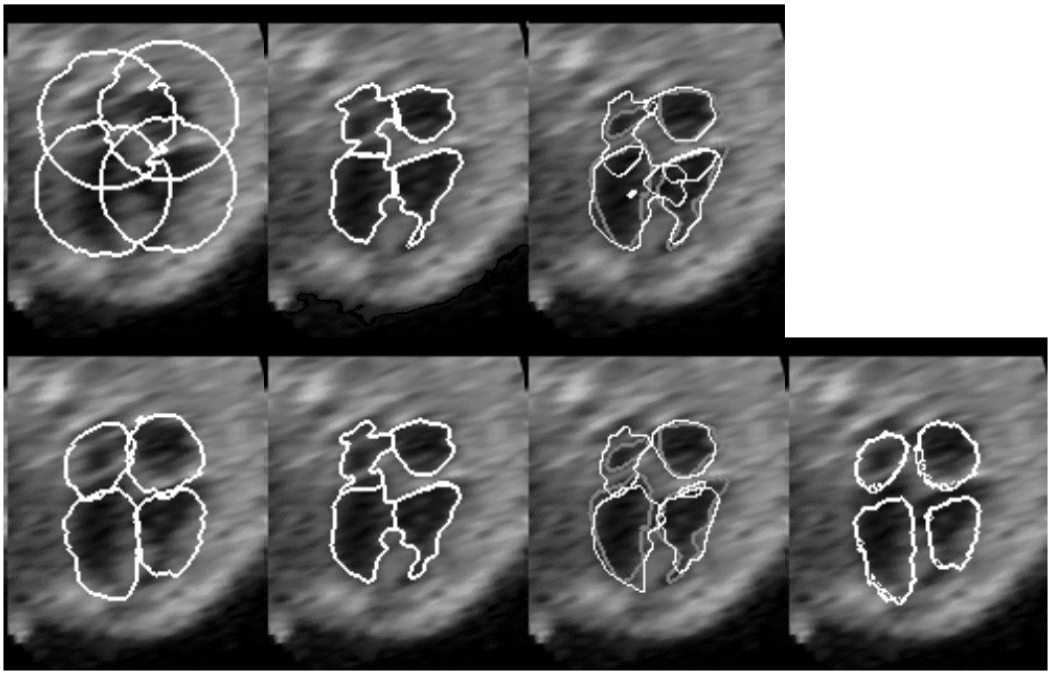


Figure 71 Example segmentation results. Without SP term (*top row*) and with +SP term enabled on the *bottom row*. *First column*: EPCA, *second column*: MSSCD and the *third column* shows TIMS level set snakes. Manual tracings are in *grey*. The *far right* shows the affine registered prior to the image. *Left hand side* of each image corresponds to the left part of the heart.

In the graphs shown in Figure 72 the errors are rounded up so a value of 4.1 will appear as 5 by the frequency analysis. In each snake type, addition of the prior shape term yields much better segmentation agreement with manual tracing with the EPCA snake showing greatest improvement. Examination of the histograms for each deformable model algorithm shows that the peak frequency counts appear at lower rms errors when the shape prior term is enabled. Although the peak rms error in the histogram distributions appears lowest in the TIMS level set, it is clear that the histogram with the narrowest spread in errors is due to the MSSCD+SP snake. This algorithm has the most number of adaptive terms of all the snakes tested here. The Sarti edge flow and collision detection in (2.34) are unique amongst the other snakes tested in this paper and are needed to obtain more conforming delineations.

The constant advection term of the EPCA snake caused the contour to frequently overshoot the boundaries even with the use of the SP term. When the SP term was disabled the collision detection term in MSSCD was a worthy competitor to the TIMS+SP in some images (e.g. in Figure 71).

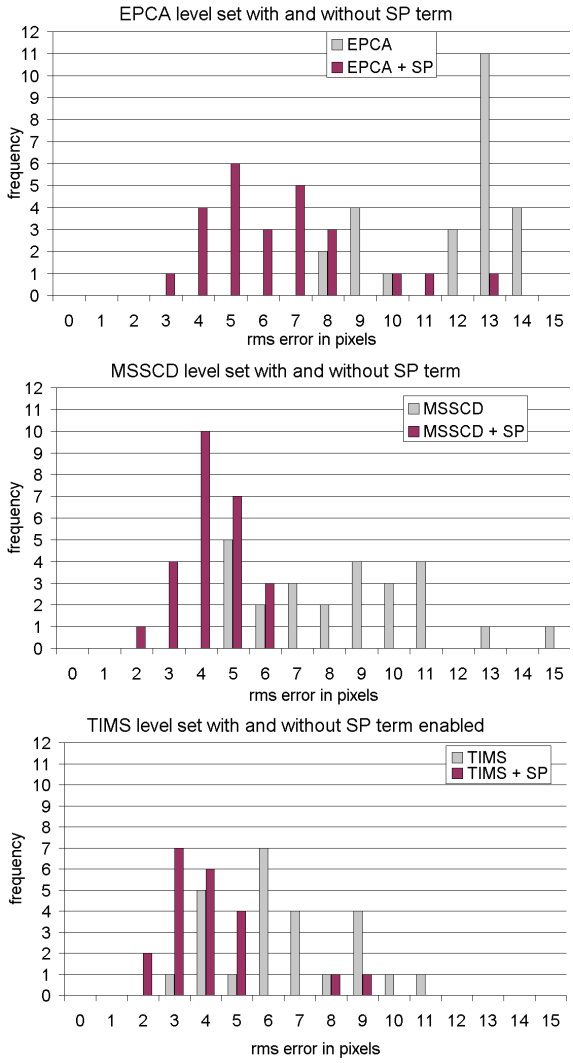


Figure 72 Average rms errors (over all cardiac chambers) for each snake. Peaks towards the lower end of the *horizontal* scale indicate better performance.

Each pixel in the image shown in Figure 73 was corrupted with 90% white noise in the range [0, 255] and the template was registered to the image. The results of the segmentation both with and without the shape prior term can be seen in Figure 73. When the prior is enabled the snake remains faithful to the template and retains most of the corrupted thin septal walls.

Figure 73 Segmentation of foetal heart image with 90% random noise added. *First* image shows initial position of the template before registration, segmentation by the snake with shape prior in the *second* image and the *third* shows delineation using the snake only. Automatic segmentations appear in *white* and manual delineation in *grey*. Average rms errors are 3.0 pixels in the second image and 5.2 in the third. The LV is on *left hand side*.

5.7 Early Tracking of the Foetal Heart

The proposed algorithms in this thesis currently has no tracking capabilities but this section is presented in isolation for completeness since tracking is part of the scope of future work for this project.

Tracking algorithms are usually applied to a succession of temporal frames. One of their main assumptions is that the image frame rate is sufficiently high that the future course of the tracked object can be predicted from the last few frames. In my previous paper [50] I modelled the deformation between frames in a global optimisation framework using constrained rigid body transforms to partially compensate for low frame rate acquisitions. The heart is a dynamic organ that deforms in a 3D non-rigid manner. Consequentially temporal tracking algorithms that are applied to 2D datasets cannot be expected to capture the entire functional information since it is likely that the heart walls will move out of the acquisition plane. Nevertheless I applied temporal wall tracking in 2D (see section 2.11 and [50]) using the GVF algorithm as a precursor of the 3D analysis to be completed at a later stage. The main results from this paper are presented below.

Figure 74 shows the segmentation from both GVF + tracking and manual methods where in the Bland-Altman plots the bias is positive signifying over-segmentation for both left and right ventricles. The spread is roughly centred about the mean and most if not all of the points fall within the 95% confidence interval. These plots shows that by comparing areas defined by the manual and automatic curves, the algorithm produces less fluctuation and systematic bias in segmentation of the left ventricle when compared to the segmentation of the right (mean 26 with standard deviation 15 and mean 84 with standard deviation 46 for the left and right ventricles respectively).

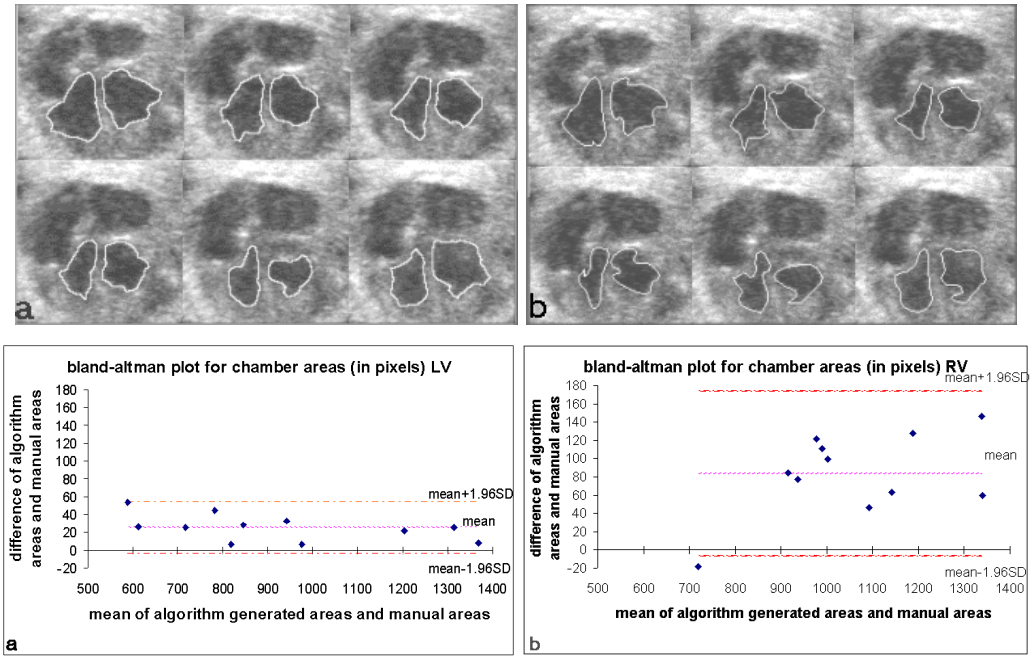


Figure 74 Automatic segmentation (*a*) and manual segmentation (*b*) as a result of GVF + tracking algorithm. LV appears on *left hand side* of images.

Although these were early results and shape tracking was not switched off for comparison, the correlation coefficient between manual and automatic methods was over 0.9 for an 11 frame sequence with rms distance errors varied between 1 and 4.5 pixels. The pixel size of this dataset was 0.26mm which is comparable to the results of the 2D and 3D segmentation algorithms presented before in section 5.3. From these limited set of results it is not clear if the extra implementation of rigid body shape constraints from frame to frame is necessary to achieve superior segmentation results to non tracking algorithms; further analysis on more datasets and in 3D are required to verify.

5.8 Summary

This chapter presented the results and discussions of all the deformable models implemented in this project. First the sensitivity of the models was tested on a synthetic image constructed from one of the foetal cardiac datasets and corrupted with controlled levels of different types of ultrasonic noise. Rayleigh speckle, attenuation, occlusion and partial volume noise types were simulated. The MSSCD algorithm obtained the highest overall GTC compared to the uncorrupted synthetic image in each noise test.

The next stage of testing was to modify the weighting coefficients of the terms in each snake equation in order to determine the sensitivity to suboptimal parameters. The parameter optimisation was run on a single representative foetal heart dataset

with good contrast primarily because of the lengthy time taken to compute all the segmentations. Once again the MSSCD shone through with the fewest number of parameters to optimise, and its main dependence was on only one of them. As expected from this test the main driving force was the most sensitive parameter in each snake.

Following parameter optimisation was segmentation of all the foetal cardiac images using each snake in turn applied to the same data for direct comparisons. MSSCD was overall superior to the other snakes, both in terms of volumetric and distance errors when compared to manual segmentation. The inter observer validation on 2D slices was performed with the result that the GTC of one observer to the other was comparable to the GTCs of some of the snakes in 2D.

Validation of the algorithms proved successful using a physical phantom to simulate the heart's appearance in the ultrasound scanner. Out of all the deformable models MSSCD provided the closest estimation of the true physically measured volume and minimum distance errors to the manual segmentation.

The level set snakes were compared both with and without a shape prior atlas in 2D datasets. The MSSCD + SP algorithm was found to be superior to TIMS + SP. In the absence of any shape prior term, MSSCD obtained the best segmentation due to its collision detection term. Segmentation with the shape prior alone did not yield a proper delineation of the heart due to its rounded appearance.

The final sets of results indicated that functional shape constrained tracking in the GVF snake was possible. This methodology could also be incorporated in each of the deformable models presented in this thesis.

The next chapter concludes the thesis and outlines future work.

6 CONCLUSIONS AND FUTURE WORK

6.1 Summary of Progress

This thesis presents two deformable model boundary descriptions in both 2D and 3D. Both types of representation have their individual merits and problems which are treated in application. From the implicit method the segmentation appears to have better control for wrapping around irregular shapes with less self intersections and aliasing effects. The explicit models are faster to evolve and would be easier to establish correspondence for registration of two contours/surfaces for tracking purposes. Their sparse nature allows a coarse to fine resolution framework to be adapted in evolution however this was not implemented for this thesis.

The PZLS was developed as the precursor for the level set algorithm implemented for this thesis. Since it was implemented after experimentation with the GVF and scale space snakes on the datasets, it is more adaptive in segmentation of missing foetal cardiac structures in the absence of strong edges and image inhomogeneity. However, since it has no penalisation for inter snake overlap it is less resilient to collapse under internal forces due to open valves and membranous septum dropouts than the MSSCD snake.

A summary of all the snake properties and findings is presented in Table 14.

Table 14 Summary of snake properties and findings (+VE and –VE) refer to positive and negative respectively

	Scale Space	GVF	PZLS	EPCA	MSSCD	TIMS
+VE	-Multiresolution -Fast evolution	-Fast evolution	-Similar results to TIMS without SP term -More tolerant to local noise than edge based explicit snakes	-Fast Marching is quickest LS algorithm tested here	-Smooth segmentation results -Collision Detection	-Collision Detection not needed when +SP term present
-VE	-Some self intersections -No Collision Detection -Strongest edge is not necessarily desired edge	-Some self intersections -No Collision Detection -GVF field very weak without balloon term	-Some self intersections -No Collision Detection	-Over spills easily across most boundaries in real foetal cardiac data. -Can leak through vessels connecting to chambers	-Can leak through vessels connecting to chambers	-Can show leakage between chambers without SP term since no Collision Detection term is present.

Despite the model based segmentation approach of a snake, further improvement involved additional constraints in the form of a shape prior to provide guidance in the absence of structural information such as during attenuation and partial volume imaging of thin structures.

All of the tested algorithms work on 2 scanners in both 2D and 3D, and on US datasets of variable quality in the image and/or acquisition parameters. In terms of approximating the clinical manually drawn tracings the MSSCD approach was most successful overall on both real and synthetic data as well as the phantom. An existing level set segmentation algorithm on the foetal heart [12] was compared with the proposed ones both with and without shape awareness and was outperformed by the ones proposed in this thesis.

The main limitation with the shape prior algorithm is the lack of variance during the fitting of the shape knowledge to the image and the future work in section 6.5 attempts to address these issues by drawing on strengths from other methods for incorporation into the algorithm.

6.2 Evaluation of objectives

The following objectives to the foetal cardiac segmentation field are repeated from section 1.1 for the MSSCD and TIMS algorithms:

Table 15 Evaluation of objectives at end of thesis

Objective from thesis	Objective met?
Deformable model segmentation of the foetal cardiac chambers with minimal user interaction.	Met, with manual placing of seed points in centre of chambers, no user interaction needed with +SP term.
Ejection Fraction (EF) measurements from automatically determined volumes of the foetal heart and compare them with the manually determined ones and other measured values in the literature.	Objective met.
The automatic segmentation algorithm must be compared on both real and foetal cardiac synthetic data with other similar algorithms for performance evaluation.	Objective met.

Quantitative evaluation of algorithms to manual tracings in terms of physical distances.	Objective met.
Assessment of clinical accuracy.	Objective met, could be tested on more data and repeatability testing for each observer.
Comparison to ground truth by validation of algorithms using a physical phantom.	Objective met. Beating contractile phantom could be used for tracking algorithm.
Compare performance of the segmentation in both 2D and 3D.	Objective met.
The ability to provide surface rendering of the foetal cardiac chambers.	Objective met but rendering is done from filled segmentation using Marching Cubes algorithm in itk-SNAP [131].

6.3 Evaluation of Hypotheses

The following hypotheses for the content in this thesis are repeated from section 1.2 and are evaluated for foetal cardiac data.

Table 16 Evaluation of Hypotheses at end of thesis

Hypothesis	Outcome
Region based segmentation is more robust than edge based methods	Proven
Segmentation accuracy will be better in 2D compared to 3D due to the reduced resolution of matrix probes	Proven
Shape priors improve segmentation accuracy	Proven

6.4 Original Contributions of Thesis

According to my research in the foetal cardiac segmentation field this thesis is the first to

- Compare different deformable model segmentation algorithms on the foetal heart, both edge and region based
- Propose the PZLS algorithm although similar to [53] retrofits the MS term in the explicit snake formulation
- Use modulating data driven term for the PZLS algorithm and applied this methodology in the edge based explicit snakes

- Segment automatically from more than one scanner with datasets in both 2D and 3D
- Provide reconstruction for missing walls at partially resolved ventricular/atrial septal defects and at valve planes
- Provide automated EFs for both manual and automated methods
- Provide the entire volume time graphs of the cardiac cycle for all four chambers
- Provide a framework for template segmentation from captured expert knowledge of the chamber shapes
- Apply registration techniques to the foetal heart for segmentation
- Use shape prior in level set snakes that were designed without shape awareness and benchmarked their improvement in the segmentation
- Quantify the distance errors of automatic segmentation against manual tracings
- Compare more than one observer's manual tracings of the chambers
- Specify a protocol when segmenting the papillary muscles by automated means
- Validate the automatic and manual segmentation tracings with a physical phantom of known volume
- Attempt to track the foetal heart during the cardiac cycle using global and non-rigid deformation
- Model the US corruption (involving shadowing, partial volume, attenuation and speckle) on a synthetic image derived from a real foetal heart and its impact on the segmentation algorithms
- Unification of Sarti and MS snake in MSSCD
- Provide an exponential curvature penalisation such as in MSSCD
- Application of collision detection to reconstruct missing inbetween chamber boundaries in MSSCD also enables simultaneous region detection of the same tissue type without inter-region boundaries present

6.5 Future Work

6.5.1 Further improvements to snake algorithms

The algorithm works well on the datasets available for this study and as a result 5 conference publications came out of the work [50], [70], [71], [126], [129] and a Lecture Notes in Computer Science paper [73], with additional scope for journal papers. The steps below are an outline of what needs to be done to perform more accurate segmentation with the presented deformable model algorithms applied to the available foetal heart data. If more datasets were being acquired the algorithms could be tested against a full Active Shape Model/Active Appearance Model approach.

6.5.1.1 Shape prior applied to all frames in cardiac cycle

The current shape prior for the level sets implemented in this thesis is able to handle the following

- Chambers with approximately same cardiac phase as the mean template
- Little variation in the relative positions of individual chambers.

The shape prior fitting algorithm needs to be able to handle different cardiac phases and this may be improved by allowing joint deformation of the two ventricles and likewise for the atria during the registration. Since the two atria contract together and similarly for the ventricles during the cardiac cycle this asymmetric deformation of the multichamber template has some similarities to the cardiac function. This methodology is similar to my previous work on foetal cardiac tracking presented in [50] where I used the rigid deformation independently for each chamber segmented contours from the previous frame as a shape prior for the current one.

Xiahai et al [113] have used a similar approach to segment adult cardiac MRI data with satisfactory results. In his method a mean template was affine registered to the unseen dataset and then a locally affine mode is applied in which each chamber would move independently to account for varying thicknesses of septal and valve plane widths as well as different cardiac phases.

6.5.1.2 Adaptive resolution of snake (explicit and implicit)

The snakes in this thesis are affected by the partial volume effect of papillary muscles and speckle. One approach to avoid the need for good prefiltering is to use a multiresolution approach to the segmentation. A positive side effect to this methodology is the speedup in the segmentation algorithms. Possible drawbacks with this method may be increased partial volume effect or dropping out of thin structures that were visible in the full resolution image. If the image was to remain at the original resolution and only the snake changes scale it may be more adaptive to these artefacts. Note the presence of a shape prior may avoid the need to use this adaptation.

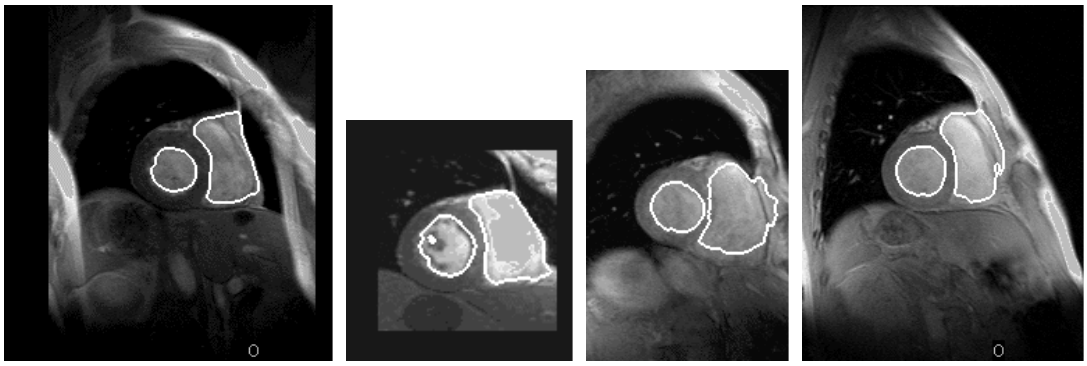
6.5.1.3 Optimisation approaches

Currently the algorithm implementation is highly unoptimised speed wise since the program writes out intermediate images to disk every few iterations as well as logs of various variable values for debugging purposes. Although the C++ program is running in release mode it could be sped up even further by using direct memory access instead of access functions to the image, explicit mesh and level set arrays within the classes. If these optimisations as well as the adaptive resolution of the

snakes are implemented the speedup may be up to twice or three times the current evolution time (~10 minutes for 3D foetal heart datasets on Pentium 4 1.9GHz 1GB RAM Windows XP Professional).

6.5.2 Application of snake to MRI modality

This thesis tested automatic segmentation of foetal echocardiograms. It is possible to apply MSSCD to cardiac MRI datasets with little or no change. Whilst investigating possible use of MSSCD outside ultrasound I applied this algorithm to 2D MRI images of the adult heart to show the performance and limitations outside foetal echocardiography. It appears that the LV is segmented without any difficulty (Figure 75). The RV snake however, can leak into the space beyond the epicardium since the wall is very thin and there is little contrast of the space outside the heart compared to the blood pool chamber (refer to Figure 75c,d). However, this is less of a problem when there is an air space between the epicardium and the chest wall (Figure 75a,b). If the images were segmented in 3D instead of 2D the additional dimension may provide some constraint via curvature to reduce the amount of leakage. Use of atlas based approaches for whole heart in MRI would also be useful constraining the segmentation in cardiac MRI e.g [113].



a b c d

Figure 75 Example segmentation of adult cardiac MR images (2D) using the MSSCD algorithm. The LV appears on the *left hand side* of each image.

The example presented in Figure 75 is not the only potential segmentation use of the MSSCD algorithm. It could be applied to adult echocardiography, ovary echocardiography, inner ear auditory imaging and other types of images where organs with multipart regions are required to be segmented. The +SP term can be incorporated into the algorithms TIMS and MSSCD for more reliability in the segmentation in the presence of spurious edge artefacts and wall dropouts. If a large number of datasets are acquired then in addition to the mean the natural variation in the heart can be modelled using level set statistical shape modelling such as in Leventon's work [65].

REFERENCES

[1] Mitchell, S. C., Korones, S. B., and Berendes, H. W., "Congenital heart disease in 56,109 births. Incidence and natural history," *Circulation*, vol. 43, no. 3, pp. 323-332, Mar.1971.

[2] Esh-Broder, E., Ushakov, F. B., Imbar, T., and Yagel, S., "Application of free-hand three-dimensional echocardiography in the evaluation of fetal cardiac ejection fraction: a preliminary study," *Ultrasound in Obstetrics & Gynecology*, vol. 23, no. 6, pp. 546-551, 2004.

[3] Deng, J., "Terminology of three-dimensional and four-dimensional ultrasound imaging of the fetal heart and other moving parts," *Ultrasound in Obstetrics & Gynecology*, vol. 22, no. 4, pp. 336-334, 2003.

[4] Frangi, A. F., Niessen, W. J., and Viergever, M. A., "Three-dimensional modeling for functional analysis of cardiac images: A review," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 2-25, 2001.

[5] Farlex. The Free Dictonary. *http://medical-dictionary.thefreedictionary.com*, vol. 2009.

[6] Wikimedia. Wikipedia - The Free Encyclopedia. *http://en.wikipedia.org*, vol. 2009.

[7] Montagnat, J., Sermesant, M., Delingette, H., Malandain, G., and Ayache, N., "Anisotropic filtering for model-based segmentation of 4D cylindrical echocardiographic images," *Pattern Recognition Letters*, vol. 24, no. 4-5, pp. 815-828, Feb.2003.

[8] Chang, F.-M., Hsu, K.-F., Ko, H.-C., Yao, B.-L., Chang, C.-H., Yu, C.-H., Liang, R.-I., and Chen, H.-Y., "Fetal heart volume assessment by three-dimensional ultrasound," *Ultrasound in Obstetrics & Gynecology*, vol. 9, no. 1, pp. 42-48, 1997.

[9] Davis, L. E. and Hohimer, A. R., "HEMODYNAMICS AND ORGAN BLOOD-FLOW IN FETAL SHEEP SUBJECTED TO CHRONIC ANEMIA," *American Journal of Physiology*, vol. 261, no. 6, pp. R1542-R1548, 1991.

[10] Ashmead, G. A., "Ultrasound diagnosis of anomalies," in Ashmead, G. A. (ed.) *Essentials of maternal fetal medicine* USA: Chapman and Hall International Thomson Publishing, 1997, pp. 43-63.

[11] Kucera D and Martin RW, "Segmentation of sequences of echocardiographic images using a simplified 3D active contour model with region based external forces," *Computerized Medical Imaging and Graphics*, vol. 21, no. 1, pp. 1-21, 1997.

[12] Lassige TA, Benkeser PJ, Fyfe D, and Sharma S, "Comparison of septal defects in 2D and 3D echocardiography using active contour models,"

Computerized Medical Imaging and Graphics, vol. 24, no. 6, pp. 377-388, 2000.

[13] Richens D, Rougon N, and Mousseaux E, "Segmentation by deformable contours of MRI sequences of the left ventricle for quantitative analysis," *IEE 354 Proceedings - International Conference on Image Processing and Applications*, pp. 393-396, 1992.

[14] Sklansky, M. S., Nelson, T., Strachan, M., and Pretorius, D., "Real-time three-dimensional fetal echocardiography: initial feasibility study," *Journal of Ultrasound in Medicine*, vol. 18, no. 11, pp. 745-752, Nov.1999.

[15] Nelson, T. R., Pretorius, D. H., Sklansky, M., and HagenAnsert, S., "Three-dimensional echocardiographic evaluation of fetal heart anatomy and function: Acquisition, analysis, and display," *Journal of Ultrasound in Medicine*, vol. 15, no. 1, pp. 1-9, 1996.

[16] Deng, J., Yates, R., Birkett, A. G., Ruff, C. F., Linney, A. D., Lees, W. R., Hanson, M. A., and Rodeck, C. H., "Online motion-gated dynamic three-dimensional echocardiography in the fetus - Preliminary results," *Ultrasound in Medicine and Biology*, vol. 27, no. 1, pp. 43-50, 2001.

[17] Zosmer, N., Jurkovic, D., Jauniaux, E., Gruboeck, K., Lees, C., and Campbell, S., "Selection and identification of standard cardiac views from three-dimensional volume scans of the fetal thorax," *Journal of Ultrasound in Medicine*, vol. 15, no. 1, pp. 25-32, Jan.1996.

[18] Lewin, P. A. Diagnostic ultrasound - a glimpse into the next decade. *World Medical Association Business Briefing - Global Healthcare 2000*, vol. World Medical Association. 2000.

[19] Deng, J., Gardener, J. E., Rodeck, C. H., and Lees, W. R., "Fetal echocardiography in three and four dimensions," *Ultrasound in Medicine and Biology*, vol. 22, no. 8, pp. 979-986, 1996.

[20] Deng, J., Ruff, C. F., Linney, A. D., Lees, W. R., Hanson, M. A., and Rodeck, C. H., "Simultaneous use of two ultrasound scanners for motion-gated three-dimensional fetal echocardiography," *Ultrasound in Medicine and Biology*, vol. 26, no. 6, pp. 1021-1032, 2000.

[21] Sklansky, M. D., Nelson, P., and Pretorius, M. D., "Usefulness of Gated Three-Dimensional Fetal Echocardiography to Reconstruct and Display Structures Not Visualized With Two-Dimensional Imaging," *The American Journal of Cardiology*, vol. 80, no. 5, pp. 665-668, Sept.1997.

[22] Fish, P., *Physics and instrumentation of diagnositic medical ultrasound* Wiltshire: John Wiley and Sons Ltd., 1996.

[23] MCDICKEN, W. N., *Diagnositc Ultrasounics Principles and use of instruments*, 3rd Edition ed. 1991.

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

[25] Anderson, M. E. and Trahey, G. E. A seminar on k-space applied to medical ultrasound. Duke University. 2000.

[26] Cosgrove, D. Harmonic imaging: tissues and microbubbles. *http://www.smirg.org/lectures/lecture_id2.php?sl=1*, vol. 2002.

[27] Wojcik, G., Mould, J., Ayter, S., and Carcione, L. A study of second harmonic generation by focused medical transducer pulses. 2, pp. 1583-1588. IEEE Ultrasonics Symposium, 1998. Proceedings. 1998.

[28] Gonçalves, L. F., Lee, W., Chaiworapongsa, T., Espinoza, J., Schoen, M. L., Falkensammer, P., Treadwell, M., and Romero, R., "Four-dimensional ultrasonography of the fetal heart with spatiotemporal image correlation," *American Journal of Obstetrics and Gynecology*, vol. 189, no. 6, pp. 1792-1802, 2003.

[29] Deng, J., Sullivan, I. D., Yates, R., Vogel, M., Mcdonald, D., Linney, A. D., Rodeck, C. H., and Anderson, R. H., "Real-time three-dimensional fetal echocardiography - optimal imaging windows," *Ultrasound in Medicine & Biology*, vol. 28, no. 9, pp. 1099-1105, Sept.2002.

[30] Powis, R. L. and Powis, W. J., *A thinker's guide to ultrasonic imaging* Baltimore: Urban & Schwarzenberg, 1984.

[31] GE. ultrasound definitions. *http://www.amershamhealth.com/medcyclopaedia/medical/index.asp*,vol. 2004.

[32] Bucsko, J. K. Imaging the eye with very-high frequency ultrasound. *Radiology Today*, vol. 5, no. 19, pp. 10-16. 2004.

[33] Ouellette, J. Seeing with sound - acoustic microscopy advances beyond failure analysis. *The Industrial Physicist*, vol. pp. 14-17. 2004.

[34] Thijssen, J. M., "Ultrasonic speckle formation, analysis and processing applied to tissue characterization," *Pattern Recognition Letters*, vol. 24 pp. 659-675, 2003.

[35] Anderson, M. E. Spatial and temporal coherence. *http://dukemil.egr.duke.edu/Ultrasound/k-space/node8.htm*, vol. 2000.

[36] Goodman, J. W., "Some fundamental properties of speckle," *Journal of the Optical Society of America*, vol. 66, no. 11, pp. 1145-1150, 1976.

[37] Marr D and Hildreth E, "Theory of edge detection," *Proceeding Royal Society London B*, vol. 207 pp. 187-217, 1980.

[38] Perona, P. and Malik, J., "Scale-space and edge detection using anisotropic diffusion," *PAMI*, vol. 12, no. 7, pp. 629-639, 1990.

[39] Kass M, Witkin A, and Terzopoulos D, "Snakes: Active Contour Models," *International Journal of Computer Vision*, vol. 1 pp. 321-331, 1988.

[40] Chalana, V., Linker, D. T., Haynor, D. R., and Kim, Y. M., "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Transactions on Medical Imaging*, vol. 15, no. 3, pp. 290-298, 1996.

[41] Ranganath S, "Contour extraction from cardiac MRI studies using snakes," *IEEE Transactions on Medical Imaging* , vol. 14, no. 2, pp. 328-338, 1995.

[42] Santarelli, M. F., Positano, V., and Benassi, A. A new algorithm for 3D automatic detection and tracking of cardiac wall motion. 26, pp. 133-136. Hannover, Germany. Computers in Cardiology. 1999.

[43] Cohen, L. D., "On Active Contour Models and Balloons," *CVGIP: Image Understanding*, vol. 53, no. 2, pp. 211-218, 1991.

[44] Kang DJ, "A fast and stable snake algorithm for medical images," *Pattern Recognition Letters*, vol. 20 pp. 507-512, 1999.

[45] McInerney T and Terzopoulos D, "A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis," *Computerized Medical Imaging and Graphics*, vol. 19, no. 1, pp. 69-83, 1994.

[46] McInerney, T. and Terzopoulos D, "Topology adaptive deformable surfaces for medical image volume segmentation," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 840-850, 1999.

[47] Xu, C. and Prince, J. L. Gradient Vector Flow: a new external force for snakes. pp. 66-71. IEEE Computer Vision and Pattern Recognition. 1997.

[48] Sanchez, P. J., Zapata, J., and Ruiz, R., "An Active Contour Model Algorithm for Tracking Endocardiac Boundaries in Echocardiographic Sequences," *Critical Reviews in Biomedical Engineering*, vol. 28, no. 3&4, pp. 487-492, 2000.

[49] Tao, Z. and Tagare, D. Tunneling descent level set segmentation of ultrasound images. Christensen, G. E. and Sonka, M. 3565, pp. 750-761. Berlin, Heidelberg, Springer-Verlag. Information Processing in Medical Imaging. 2005.

[50] Dindoyal, I., Lambrou, T., Deng, J., Ruff, C. F., Linney, A. D., and Todd-Pokropek, A. An active contour model to segment foetal cardiac ultrasound data. pp. 77-80. University of Sheffield, UK. Proceedings of Medical Image Understanding and Analysis. 2003.

[51] Bosnjak, A., Torrealbar, V., Montilla, G., Villegas, H., Burdin, V., Solaiman, B., and Roux, C. Segmentation, reconstruction, modeling and 3D visualization of the ventricles in echocardiographics images. pp. 260-264. Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on. 2001.

[52] Corsi, C., Saracino, G., Sarti, A., and Lamberti, C., "Left ventricular volume estimation for real-time three-dimensional echocardiography," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1202-1208, 2002.

[53] Caselles, V., Kimmel, R., and Guillermo, S., "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61-79, 1997.

[54] Wang, J. and Li, X., "Controlled accurate searches with balloons," *Pattern Recognition*, vol. 36, no. 3, pp. 827-843, Mar.2003.

[55] Delingette, H. and Montagnat, J., "New algorithms for controlling active contours shape and topology," *Lecture Notes in Computer Science*, vol. 1843 pp. 381-395, 2000.

[56] Montagnat, J. and Delingette, H. Space and time shape constrained deformable surfaces for 4D medical image segmentation. *Medical Image Computing and Computer-Assisted Intervention - Miccai 2000*, vol. 1935, pp. 196-205. MICCAI 2000. 2000.

[57] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. Training models of shape from sets of examples. pp. 9-18. Proceedings of British Machine Vision Conference. 1992.

[58] McInerney, T. and Terzopoulos, D. Topologically adaptable snakes. pp. 840-845. Computer Vision, 1995. Proceedings., Fifth International Conference on. 1995.

[59] Corsi, C., Borsari, M., Conseganati, F., Sarti, A., Lambert, C., Travaglini, A., Shiota, T., and Thomas, J. D., "Left ventricular endocardial surface detection based on real-time 3D echocardiographic data," *European Journal of Ultrasound*, vol. 13, no. 1, pp. 41-51, 2001.

[60] Chan, T. F. and Vese, L. A., "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277, 2001.

[61] Gibou, F. and Fedkiw, R. A fast hybrid k-means level set algorithm for segmentation. Stanford Technical Report. 2002.

[62] Sarti, A., Corsi, C., Mazzini, E., and Lamberti, C. Maximum likelihood segmentation of ultrasound images with Rayleigh distribution. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 52, no. 6, pp. 947-960. 2005.

[63] Sonka, M., Hlavac, V., and Boyle, R., *Image Processing, Analysis and Machine Vision*, 2nd Edition ed. Brooks/Cole Publishing Company, USA, 1999.

[64] Sethian, J. A., *Level set methods evolving interfaces in geometry, fluid mechanics, computer vision and materials science*, 1st edition ed. University Of Cambridge, 1996.

[65] Leventon, M. E., Grimson, W. E. L., and Faugeras, O. Statistical shape influence in geodesic active contours. 1, pp. 316-321. Computer Vision and Pattern Recognition. 2000.

[66] Paragios, N., Rousson, M., and Ramesh, V. Knowledge-based registration & segmentation of the left ventricle: a level set approach. *Applications of Computer Vision*, vol. pp. 37-42. 2002.

[67] Cremers, D., Sochen, S., and Schnorr, C. Towards recognition-based variational segmentation using shape priors and dynamic labelling. Isle of Skye, Scotland. Scale space methods in computer vision. 2003.

[68] Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. E., and Willsky, A., "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 137-154, 2003.

[69] Storn, R. and Price, K. Differential Evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces. 1995.

[70] Dindoyal, I., Lambrou, T., Deng, J., and Todd-Pokropek, A. Fully automated shape prior applied to the foetal heart - preliminary results. Manchester. Proceedings of Medical Imaging Understanding and Analysis. 2006.

[71] Dindoyal, I., Lambrou, T., Deng, J., and Todd-Pokropek, A. Level set snake algorithms on the fetal heart. pp. 864-867. Arlington, VA, USA. IEEE International Symposium on Biomedical Imaging. 2007.

[72] Sarti, A., "Subjective surfaces: a geometric model for boundary completion," *International Journal of Computer Vision*, vol. 46, no. 3, pp. 201-221, 2002.

[73] Dindoyal, I., Lambrou, T., Deng, J., Ruff, C. F., Linney, A. D., Rodeck, C. H., and Todd-Pokropek, A. Level set segmentation of the fetal heart. Frangi, A. F. pp. 123-132. Barcelona, Spain, Springer-Verlag. Functional Imaging and Modeling of the Heart, Lecture Notes in Computer Science. 2005.

[74] Sethian, J. A., "Adaptive fast marching and level set methods for propagating interfaces," *Acta.Math.Univ.Comenianae*, vol. LXVII, no. 1, pp. 3-15, 1998.

[75] Chan, T. F. and Vese, L. A. A level set algorithm for minimizing the Mumford-Shah functional in image processing. pp. 161. Washington DC, USA. VLSM 2001. 2001.

[76] Vese, L. A. and Chan, T. F., "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271-293, 2002.

[77] Azhari, H., Weiss, J. L., Rogers, W. J., Siu, C. O., and Shapiro, E. P., "A noninvasive comparative study of myocardial strains in ischemic canine hearts using tagged MRI in 3-D," *AJP - Heart and Circulatory Physiology*, vol. 268, no. 5, pp. H1918-H1926, May1995.

[78] Bradski, G. and Thrun, S. Optical Flow, Feature Tracking, Normal Flow. *www.robots.stanford.edu/cs223b04/CS%20223-B%20L9%20Optical%20Flow.ppt*, vol. 1.

[79] Crum, W. R., Camara, O., and Hill, D. L. G., "Generalized overlap measures for evaluation and validation in medical image analysis," *IEEE Transactions on Medical Imaging*, vol. 25, no. 11, pp. 1451-1461, 2006.

[80] Kuo, H.-C., Chang, F.-M., Wu, C.-H., Yao, B.-L., and Liu, C.-H., "The primary application of three-dimensional ultrasonography in obstetrics," *American Journal of Obstetrics and Gynecology*, vol. 166, no. 3, pp. 880-886, 1992.

[81] Levental, M., Pretorius, D. H., Sklansky, M. S., Budorick, N. E., Nelson, T. R., and Lou, K., "Three-dimensional ultrasonography of normal fetal heart: comparison with two-dimensional imaging," *Journal of Ultrasound in Medicine*, vol. 17, no. 6, pp. 341-348, June1998.

[82] Guerra, F. A., Isla, A. I., Aguilar, R. C., and Fritz, E. G., "Use of free-hand three-dimensional ultrasound software in the study of the fetal heart," *Ultrasound in Obstetrics & Gynecology* , vol. 16, no. 4, pp. 329-334, 2000.

[83] Deng, J., Birkett, A. G., Kalache, K. D., Hanson, M. A., Peebles, D. M., Linney, A. D., Lees, W. R., and Rodeck, C. H., "Conversion of umbilical arterial Doppler waveforms to cardiac cycle triggering signals: A preparatory study for online motion-gated three-dimensional fetal echocardiography," *Ultrasound in Medicine and Biology*, vol. 27, no. 1, pp. 51-59, 2001.

[84] Tomasi, C. and Manduchi, R. Bilateral filtering for gray and color images. pp. 839-847. Washington DC, USA, IEEE. IEEE International Conference on Computer Vision. 1998.

[85] Ye, X., Noble, J. A., and Atkinson, D., "3-D freehand echocardiography for automatic left ventricle reconstruction and analysis based on multiple acoustic windows," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1051-1058, 2002.

[86] Mulet-Parada, M. and Noble, J. A., "2D+T acoustic boundary detection in echocardiography," *Medical Image Analysis*, vol. 4 pp. 21-30, 2000.

[87] Yu, Y. and Acton, S. T., "Speckle Reducing Anisotropic Diffusion," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1260-1270, 2002.

[88] Xiao, G. F., Brady, M., Noble, J. A., and Zhang, Y. Y., "Segmentation of ultrasound B-mode images with intensity inhomogeneity correction," *IEEE Transactions on Medical Imaging*, vol. 21, no. 1, pp. 48-57, 2002.

[89] Hughes, D. I. and Duck, F. A., "Automatic attenuation compensation for ultrasonic imaging," *Ultrasound in Medicine and Biology*, vol. 23, no. 5, pp. 651-664, 1997.

[90] Han, C. Y., Lin, K. N., Wee, W. G., Mintz, R. M., and Porembka, D. T., "Knowledge-based image analysis for automated boundary extraction of transesophageal echocardiographic left-ventricular images," *IEEE Transactions On Medical Imaging*, vol. 10, no. 4, pp. 602-610, 1991.

[91] Furuie, S. S., Rebelo, M. S., Gutierrez, M. A., and Melo, C. P. Segmentation of left ventricle myocardium in MRI using endocardium layer expansion model. pp. 141-144. IEEE Computers in Cardiology. 1997.

[92] Friedland, N. and Adam, D., "Automatic Ventricular Cavity Boundary Detection from Sequential Ultrasound Images Using Simulated Annealing," *IEEE Transactions on Medical Imaging*, vol. 8, no. 4, pp. 344-353, 1999.

[93] Sanchez-Ortiz, G. I., Wright, G. J. T., Clarke, N., Declerck, J., Banning, A. P., and Noble, J. A., "Automated 3-D echocardiography analysis compared with manual delineations and SPECT MUGA," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1069-1076, 2002.

[94] Lynch, M., Ghita, O., and Whelan, P. F., "Automatic segmentation of the left ventricle cavity and myocardium in MRI data," *Computers in Biology and Medicine*, vol. 36, no. 4, pp. 389-407, 2006.

[95] Paragios, N., "A Variational Approach for the Segmentation of the Left Ventricle in Cardiac Image Analysis," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 345-362, Dec.2002.

[96] Bernard, O., Touil, B., Gelas, A., Prost, R., and Friboulet, D. Segmentation of Myocardial Regions in Echocardiography Using the Statistics of the Radio-Frequency Signal. 4466, no. 2007, pp. 433-442. Salt Lake City, UT, USA, Lecture Notes in Computer Science. FIMH 2007. 2007.

[97] El Berbari, R., Bloch, I., Redheuil, A., Angelini, E. D., Mousseaux, E., Frouin, F., and Herment, A. Automated Segmentation of the Left Ventricle Including Papillary Muscles in Cardiac MagneticResonance Images. 4466, no. 2007, pp. 453-462. Salt Lake City, UT, USA, Lecture Notes in Computer Science. FIMH 2007. 2007.

[98] Hill, A. and Taylor, C. J., "Model-Based Image Interpretation Using Genetic Algorithms," *Image Vision Computing*, vol. 10, no. 5, pp. 295-300, 1992.

[99] Parker, A. D., Hill, A., Taylor, C. J., Cootes, T. F., Jin, X. Y., and Graham, J. Application of point distribution models to the automated analysis of echocardiograms. pp. 25-28. Computers in Cardiology. 1994.

[100] Hamarneh, G. and Gustavsson, T. Combining snakes and active shape models for segmenting the human left ventricle in echocardiographic images. 27, pp. 115-118. 2000.

[101] Hamarneh, G. and Gustavsson, T. Statistically constrained snake deformations. 3, pp. 1610-1615. 2000.

[102] Song, M. Z., Haralick, R. M., Sheehan, F. H., and Johnson, R. K., "Integrated surface model optimization for freehand three- dimensional echocardiography," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1077-1090, 2002.

[103] Mitchell, S. C., Bosch, J. G., Leieveldt, B. P. F., Van der Geest, R. J., Reiber, J. H. C., and Sonka, M., "3-D Active Appearance Models: segmentation of cardiac MR and ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1167-1178, 2002.

[104] van Assen, H. C., Danilouchkine, M. G., Behloul, F., Lamb, H. J., Van der Geest, R. J., Reiber, J. H. C., and Lelieveldt, B. P. F. Cardiac LV segmentation using a 3D active shape model driven by fuzzy inference. *Medical Image Computing and Computer-Assisted Intervention - Miccai 2003, Pt 1*, vol. 2878, pp. 533-540. Montreal, Canada. 2003.

[105] Besl, P. J., "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 0162-8828, pp. 239-256, 1992.

[106] Fritscher, K. D., Pilgram, R., and Schubert, R. Automatic cardiac 4D segmentation using level sets. Frangi, A. F., Petia, R. I., Santos, A., and Hernandez, M. 3504, pp. 113-122. Spain, Springer. Functional Imaging and Modeling of the Heart. 2005.

[107] Chan, T. F. and Vese, L. A., "Active Contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277, 2001.

[108] Lynch, M., Ghita, O., and Whelan, P. F., "Left-ventricle myocardium segmentation using a coupled level-set with a priori knowledge," *Computerized Medical Imaging and Graphics*, vol. 30 pp. 255-262, 2006.

[109] Wang, X., Schaerer, J., Suejung, H., Qian, Z., Metaxas, D., Chen, T., and Axel, L. Reconstruction of Detailed Left Ventricle Motion from tMRI Using Deformable Models. pp. 60-69. Salt Lake City, UT, USA, Lecture Notes in Computer Science. FIMH 2007. 2007.

[110] Szilágyi, S. M., Szilágyi, L., and Benyó, Z. Volumetric analysis of the heart using echocardiography. 4466, no. 2007, pp. 81-90. Salt Lake City, UT, USA, Lecture Notes in Computer Science. FIMH 2007. 2007.

[111] Peters, J., Ecabert, O., Carsten, M., Kneser, R., Groth, A., and Weese, J. Automatic Whole Heart Segmentation in Static Magnetic Resonance Image Volumes. pp. 402-410. Springer Berlin. MICCAI. 2007.

[112] Zheng, Y., Georgescu, B., Barbu, A., Scheuering, M., and Comaniciu, D. Four-Chamber Heart Modeling and Automatic Segmentation for 3D Cardiac CT Volumes. Rio de Janeiro, Brazil. ICCV 2007. 2007.

[113] Zhuang, X., Rhode, K., Arridge, S., Razavi, R., Hill, D., Hawkes, D., and Ourselin, S. An Atlas-Based Segmentation Propagation Framework Using Locally Affine Registration - Application to Automatic Whole Heart Segmentation. 5242, no. 2008, pp. 425-433. New York, Lecture Notes in Computer Science. MICCAI 2008. 2008.

[114] Jacob, G., Noble, A., and Blake, A. Robust contour tracking in echocardiographic sequences. pp. 408-413. Bombay, India, IEEE. Computer Vision, 1998. Sixth International Conference on. 1998.

[115] Declerck, J., Feldmar, J., and Ayache, N., "Definition of a four-dimensional continuous polar transformation for the tracking and the analysis of left-ventricle motion," *Medical Image Analysis*, vol. 2, no. 2, pp. 197-213, 1998.

[116] Malassiotis, S. and Strintzis, M. G., "Tracking the left ventricle in echocardiographic images by learning heart dynamics," *IEEE Transactions on Medical Imaging*, vol. 18, no. 3, pp. 282-290, 1999.

[117] Gerard, O., Billon, A. C., Rouet, J. M., Jacob, M., Fradkin, M., and Allouche, C., "Efficient model-based quantification of left ventricular function in 3-D echocardiography," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1059-1068, Sept.2002.

[118] Boukerroui, D., Noble, J. A., and Brady, J. M. Velocity estimation in ultrasound images: a block matching approach. Proceedings of Information Processing in Medical Imaging (IPMI). 2003.

[119] Cho, J. and Benkeser, P. J., "Cardiac segmentation by a velocity-aided active contour model," *Computerized Medical Imaging and Graphics*, vol. 30, no. 1, pp. 31-41, 2005.

[120] Lynch, M., Ghita, O., and Whelan, P. F., "Segmentation of the Left Ventricle of the Heart in 3-D+t MRI Data Using an Optimized Nonrigid Temporal Model," *IEEE Transactions on Medical Imaging*, vol. 27, no. 2, pp. 105-203, 2008.

[121] Piccoli, L., Dahmer, A., Scharcanski, J., and Navaux, P. O. A. Fetal echocardiographic image segmentation using neural networks. 2, no. 465, pp. 507-511. Image Processing and its Applications 1999, Seventh International Conference. 1999.

[122] Fernandes, D., Lemos de Siqueira, M., and Navaux, P. O. A. Segmentation of fetal echocardiographic images using self-organizing maps and oscillatory neural networks. pp. 55-60. Florianópolis. VI Simposio Ibero-Americanoem Reconecimento de Padroes (SIARP'2001). 2001.

[123] Wang, D. L. Relaxation oscillators and networks. *Wiley Encyclopedia of Electrical and Electronics Engineering*, vol. 18, pp. 396-405. John Wiley & Sons. 1999.

[124] Siqueira, M. L., Scharcanski, J., and Navaux, P. O. A., "Echocardiographic image sequence segmentation and analysis using self-organizing maps," *Journal of VLSI Signal Processing*, vol. 32 pp. 135-145, 2002.

[125] Wu, Y.-T., Chen, H.-Y., Chen, L.-F., Lee, P.-L., Yeh, T.-C., and Hsieh, J.-C., "Fetal heart tracking in ultrasound image sequences using coarse-to-fine wavelet," *Journal of Medical and Biological Engineering*, vol. 22, no. 2, pp. 67-74, 2002.

[126] Dindoyal, I., Lambrou, T., Deng, J., and Todd-Pokropek, A. Automatic segmentation of low resolution fetal cardiac data using snakes with shape priors. pp. 538-543. Istanbul, Turkey. IEEE Image and Signal Processing and Analysis. 2007.

[127] de Siqueira, M. L., Muller, D. N., and Navaux, P. O. A. Localization of cardiac cavities using hierarchically search in 2D grayscale ultrasound images. pp. 225-228. Slovakia. Proceedings of IWSSIP 2008: International conference on systems, signals and image processing. 2008.

[128] Tutschek, B. and Sahn, D. J., "Semi-automatic segmentation of fetal cardiac cavities: progress towards an automated fetal echocardiogram," *Ultrasound in Obstetrics & Gynecology*, vol. 32 pp. 176-180, 2008.

[129] Dindoyal, I., Lambrou, T., Deng, J., and Todd-Pokropek, A. Automatic segmentation of fetal heart chambers from echocardiography. 4, no. Suppl 1, pp. 44-45. Berlin, Germany. International Journal of CARS. 2009.

[130] Lobregt, S. and Viergever, M. A., "A discrete dynamic contour model," *Medical Imaging, IEEE Transactions on*, vol. 14, no. 1, pp. 12-24, 1995.

[131] itk-SNAP. no. 1.6.0.1. http://www.itksnap.org/. 2008.

[132] Bhat, A. H., Corbett, V., Carpenter, N., Liu, N., Liu, R., Wu, A., Hopkins, G., Sohaey, R., Winkler, C., Sahn, C. S., Sovinsky, V., Li, X., and Sahn, D. J., "Fetal ventricular mass determination on three-dimensional echocardiography studies in normal fetuses and validation experiments," *Circulation*, vol. 110 pp. 1054-1060, 2004.