UNIVERSITY OF OSLO
Department of
Informatics

Automated volume measurements in echocardiography by utilizing expert knowledge

Jøger Hansegård

July 4, 2008

Abstract

Left ventricular (LV) volumes and ejection fraction (EF) are important parameters for diagnosis, prognosis, and treatment planning in patients with heart disease. These parameters are commonly measured by manual tracing in echocardiographic images, a procedure that is time consuming, prone to inter- and intra-observer variability, and require highly trained operators. This is particularly the case in three-dimensional (3D) echocardiography, where the increased amount of data makes manual tracing impractical. Automated methods for measuring LV volumes and EF can therefore improve efficiency and accuracy of echocardiographic examinations, giving better diagnosis at a lower cost.

The main goal of this thesis was to improve the efficiency and quality of cardiac measurements. More specifically, the goal was to develop rapid and accurate methods that utilize expert knowledge for automated evaluation of cardiac function in echocardiography.

The thesis presents several methods for automated volume and EF measurements in echocardiographic data. For two-dimensional (2D) echocardiography, an atlas based segmentation algorithm is presented in *paper A*. This method utilizes manually traced endocardial contours in a validated case database to control a snake optimized by dynamic programming. The challenge with this approach is to find the most optimal case in the database. More promising results are achieved in triplane echocardiography using a multiview and multi-frame extension to the active appearance model (AAM) framework, as demonstrated in *paper B*. The AAM generalizes better to new patient data and is based on more robust optimization schemes than the atlas-based method. In triplane images, the results of the AAM algorithm may be improved further by integrating a snake algorithm into the AAM framework and by constraining the AAM to manually defined landmarks, and this is shown in *paper C*. For 3D echocardiograms, a clinical semi-automated volume measurement tool with expert selected points is validated in *paper D*. This tool compares favorably to a reference measurement tool, with good agreement in measured volumes, and with a significantly lower analysis time. Finally, in *paper E*, fully automated real-time segmentation in 3D echocardiography is demonstrated using a 3D active shape model (ASM) of the left ventricle in a Kalman filter framework. The main advantage of this approach is its processing performance, allowing for real-time volume and EF estimates.

Statistical models such as AAMs and ASMs provide elegant frameworks for incorporating expert knowledge into segmentation algorithms. Expert knowledge can also be utilized directly through manual input to semi-automated methods, allowing for manual initialization and correction of automatically determined volumes. The latter technique is particularly suitable for clinical routine examinations, while the fully automated 3D ASM method can extend the use of echocardiography to new clinical areas such as automated patient monitoring.

In this thesis, different methods for utilizing expert knowledge in automated segmentation algorithms for echocardiography have been developed and evaluated. Particularly in 3D echocardiography, these contributions are expected to improve efficiency and quality of cardiac measurements.

Preface

This thesis has been submitted to the Faculty of Mathematics and Natural Sciences at the University of Oslo in partial fulfillment of the requirements for the degree Philosophiæ Doctor. The work was carried out at GE Vingmed Ultrasound and at the Digital Signal Processing and Image Analysis group at the Department of Informatics, during a three-year scholarship supported by the Norwegian Research Council and GE Vingmed Ultrasound. My supervisors have been Stein Inge Rabben, Anne Schistad Solberg, and Erik Steen.

Acknowledgments

Jøger Hansegård
Oslo, July 4, 2008

List of included papers

This thesis is based on the following five papers, referred to in the text by their letters (A-E). The review process is indicated for each paper.

A   **Knowledge based extraction of the left ventricular endocardial boundary from 2D echocardiograms**
J. Hansegård, E. Steen, S. I. Rabben, A. H. Torp, H. Torp, S. Frigstad, and B. Olstad,
IEEE Ultrasonics Symposium, vol. 3, pp. 2121-2124, IEEE August 2004.
Conference paper. (Peer reviewed abstract).

B   **Detection of the myocardial boundary in the left ventricle from simultaneously acquired triplane ultrasound images using multi view active appearance motion models**
J. Hansegård, S. Urheim, E. Steen, H. Torp, B. Olstad, S. Malm, and S. I. Rabben,
IEEE Ultrasonics Symposium, vol. 4, pp. 2267-2270, IEEE September 2005.
Conference paper. (Peer reviewed abstract).

C   **Constrained active appearance models for segmentation of triplane echocardiograms**
J. Hansegård, S. Urheim, K. Lunde, and S. I. Rabben,
IEEE Transactions on Medical Imaging, vol. 26 (10), pp. 1391-1400, IEEE October 2007.
Peer reviewed journal paper.

D   **Semi-automated quantification of left ventricular volumes and ejection fraction by real-time three-dimensional echocardiography**
J. Hansegård, S. Urheim, K. Lunde, S. Malm and S. I. Rabben,
Submitted for publication.

E   **Real-time active shape models for segmentation of 3D cardiac ultrasound**
J. Hansegård, F. Orderud, and S. I. Rabben,
in *12th International Conference on Computer Analysis of Images and Patterns (CAIP 2007)*, ser. Lecture Notes in Computer Science, Walter G. Kropatsch, Martin Kampel and Allan Hanbury, Eds. vol. 4673, pp. 157-164, Springer 2007.
Peer reviewed conference paper.

Related publications not included in this thesis

1. **Real-time tracking of the left ventricle in 3D echocardiography using a state estimation approach**
   F. Orderud, J. Hansegård, and S. I. Rabben,
   in *10th International Conference on Medical Imaging and Computer-Assisted Intervention (MICCAI 2007), ser.* Lecture Notes in Computer Science, N. Ayache, S. Ourselin, and A. Maeder, Eds. vol. 4791, pp. 858-865, Springer 2007.

Contents

Paper A
Knowledge Based Extraction of the Left Ventricular Endocardial Boundary from 2D Echocardiograms
J. Hansegård, E. Steen, S. I. Rabben, A. H. Torp, H. Torp, S. Frigstad, and B. Olstad
IEEE Ultrasonics Symposium, vol. 3, pp. 2121-2124, IEEE August 2004.

Paper B
Detection of the myocardial boundary in the left ventricle from simultaneously acquired triplane ultrasound images using multi view active appearance motion models
J. Hansegård, S. Urheim, E. Steen, H. Torp, B. Olstad, S. Malm, and S. I. Rabben
IEEE Ultrasonics Symposium, vol. 4, pp. 2267-2270, IEEE September 2005.

Paper C
Constrained Active Appearance Models for Segmentation of Triplane Echocardiograms
J. Hansegård, S. Urheim, K. Lunde, and S. I. Rabben
IEEE Transactions on Medical Imaging, vol. 26 (10), pp. 1391-1400, IEEE October 2007.

Paper D
Semi-automated quantification of left ventricular volumes and ejection fraction by real-time three-dimensional echocardiography
J. Hansegård, S. Urheim, K. Lunde, S. Malm and S. I. Rabben
Submitted for publication.

Paper E
Real-Time Active Shape Models for Segmentation of 3D Cardiac Ultrasound
J. Hansegård, F. Orderud, and S. I. Rabben
in 12th International Conference on Computer Analysis of Images and Patterns (CAIP 2007), ser. Lecture Notes in Computer Science, Walter G. Kropatsch, Martin Kampel and Allan Hanbury, Eds. vol. 4673, pp. 157-164, Springer 2007.

Chapter 1

Introduction and rationale

The left ventricle of the heart is responsible for pumping blood through the body, and a well functioning ventricle is essential for patient health and wellbeing. Heart diseases, such as coronary artery disease, myocardial infarction, heart failure, and ischemia, are often manifested through reduced contractility of the heart muscle, leading to lowered pumping capacity, or ejection fraction (EF), and changes in ventricular volume [114, 115]. Left ventricular (LV) volumes and EF are therefore important measurements for diagnosis, prognosis, and treatment of patients with heart disease [100, 102, 107], and reliable determination of these parameters is of high clinical interest.

Echocardiography is an inexpensive and safe method for investigating the morphology and function of the heart [114], and is therefore the primary tool to non-invasively assess heart function in the clinic [45, 100]. The recommended method for measuring volumes and EF using echocardiography is by manually tracing the boundary of the ventricular cavity at end diastole and end systole in two orthogonal view-planes, and then computing the volume by the biplane method of disks [56, 85, 86].

Three-dimensional (3D) echocardiography has recently been introduced as a routine clinical tool [46, 72, 84]. This allows for acquiring multiple (typically three) view-planes simultaneously [6], and acquisition of dense volumetric data. In triplane imaging, improved accuracy of manually measured volumes and EF has been reported [63]. Dense volumetric data provide much more information about the ventricle than the two-dimensional (2D) modality, and the quality of extracted 2D images is improved by reducing foreshortening [46], providing better visualization of wall function, and improving the accuracy of measured volumes and EF [49, 60, 99].

The quality of echocardiographic images depends on many factors, such as speckle, thermal noise, low contrast, shadows, dropouts, reverberations, anisotropic resolution, and foreshortening. In addition, delineation of the endocardial boundary should by convention treat papillary muscles as part of the chamber volume [56]. Since accurate assessment of cardiac volumes and EF is a challenging task, this is usually done by highly trained and experienced physicians. Manual assessment of cardiac volumes is, however, time consuming and expensive, and can be prone to inter- and intra-observer variability [49]. In 3D echocardiography, manual tracing becomes impractical compared to 2D images, because of the increased amount of data.

It is therefore a clinical need for automated methods that allows for rapid, reproducible, and accurate measurements of cardiac volumes and EF in 2D, multiplane, and 3D echocardiographic data. Such methods should be robust with respect to the highly varying local image evidence in echocardiograms, and they should adhere to clinical conventions.

1.1   Aim of study

The main goal of this thesis was to improve efficiency and quality of cardiac measurements. More specifically, the goal was to develop rapid and accurate methods that utilize expert knowledge for automated evaluation of cardiac function in echocardiography.

1.2   Summary of contributions

The focus of this thesis has been to develop automated tools for measuring left ventricular volumes and EF in 2D, triplane, and 3D echocardiographic data. It has therefore been important to establish methods that are able to automatically segment the endocardial boundary in echocardiograms. Since an expert is able to do good segmentation even in moderately difficult cases, particular focus has been on how to incorporate some of the expert's knowledge into the segmentation algorithms.

The main contributions in this thesis include:

- An atlas-based method for utilizing information from a case database of expert annotated echocardiograms in a snake segmentation algorithm (Paper A).

- An active appearance model-based method for semi-automated multi-frame segmentation of triplane echocardiograms, giving temporally and spatially consistent segmentation between frames and views (Paper B).

- A framework for combining active appearance models with other segmentation algorithms such as dynamic programming-based snakes, which also permits incorporation of manually defined landmarks into the segmentation algorithm (Paper C).

- A clinical tool for semi-automated segmentation of 3D echocardiographic data, and for extraction of training data for 3D active shape models (Paper D).

- A method for fully automated real-time segmentation in 3D echocardiography using 3D active shape models (Paper E).

1.3   Outline of the thesis

This thesis is organized as follows. Chapter 2 gives a brief introduction to heart anatomy and function, echocardiography, historical perspectives, and an introduction to different types of segmentation algorithms. Chapter 3 covers related work, and summarizes concepts that are assumed to be known in the papers. Chapter 4 summarizes the individual papers, and a discussion of the contributions is given in chapter 5. Chapter 6 concludes this thesis and discusses possible directions for further work.

Chapter 2

Background

Automated interpretation of echocardiographic data is an interdisciplinary subject, requiring clinical understanding of heart anatomy and function, knowledge of image segmentation algorithms, as well as knowledge of the ultrasound image formation, processing, and display. This chapter gives a brief introduction to these fields.

2.1 Cardiology

The heart is a muscular organ, located in the chest and protected by the rib cage. The main purpose of the heart is to pump blood through the body. It consists of four main chambers, the left atrium, the left ventricle, the right atrium, and the right ventricle, as shown in Fig. 2.1. The left side of the heart is responsible for pumping oxygenated blood from the lungs throughout the body, i.e. the systemic circuit. The right side pumps de-oxygenated blood from the body into the lungs. This is called the pulmonary circuit. Since the systemic circuit has a higher blood resistance than the pulmonary circuit, the left part of the heart is larger and stronger than the right part. For both systems, the atria handle inflow to the heart, and pump blood into the corresponding ventricles. In turn, the ventricles pump blood out from the heart, into the body or into the lungs.

The atria are separated from the ventricles by directional valves allowing blood to flow from the atria into the ventricles, but not in the opposite direction. The mitral valve is located between the left atrium and the left ventricle, while the tricuspid valve is located between the right atrium and the right ventricle. These valves can hold the high ventricular blood pressure because they are anchored by thin strings, or chordae, to papillary muscles, which are attached to the inside of the ventricular wall. Reflux of blood back into the ventricles is prevented by a separate set of directional valves. The pulmonary valve is located at the outflow tract of the right ventricle, while the aortic valve is located at the outflow tract of the left ventricle.

The heart wall consists mainly of muscular tissue called the myocardium, which is composed of specialized cardiac muscle cells bundled into muscle fibers. These fibers are organized in multiple layers, each having different orientation. The interior of the heart chambers has a folded structure called trabeculae. The endocardium is the innermost layer of tissue that lines the chambers of the heart, and separates the myocardium from blood. This layer mainly consists of endothelial cells and connective tissue. The outer lining of the heart is called the epicardium and consists mainly of connective tissue. The epicardium forms the inner part of the heart sac called the pericardium, which contains the heart and the roots of the great vessels.
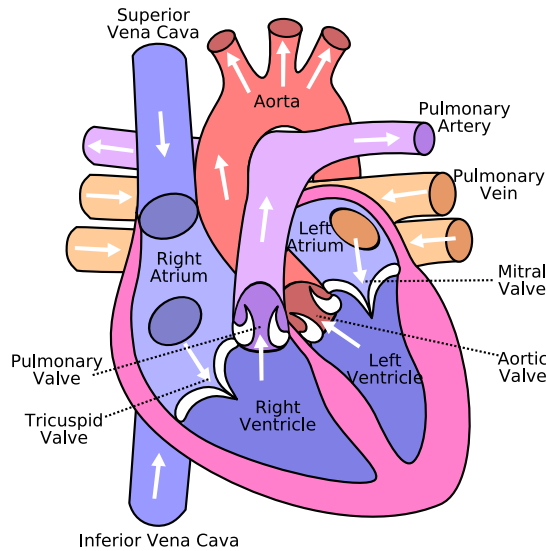
Figure 2.1: *A schematic cut through the heart, showing the main chambers, the valves, and the connected blood vessels. From [117].*

The pumping action of the heart causes the left ventricular blood volume to change during a heartbeat, as shown in Fig. 2.2. During the systolic contraction, blood is ejected from the ventricle through the aortic valve, and the ventricular volume is reduced. The mitral valve is closed, preventing blood-flow into the atrium. At beginning of diastole, the ventricle relaxes, and the reduced ventricular pressure allows blood to flow from the atrium into the ventricle through the mitral valve. This is known as rapid filling. Diastasis occurs when the pressure gradient over the mitral valve has been equalized, and volume is relatively constant. The duration of the heartbeat varies from stroke to stroke, mainly caused by different duration of this stage. In late diastole, the atrium contracts, causing the ventricle to become completely filled.

2.2   Measuring cardiac function

Cardiac function can be measured in several ways, but left ventricular volumes, stroke volume, and EF are among the most commonly used parameters [60, 109]. Stroke volume (SV) is the difference between the end diastolic volume (EDV) and the end systolic volume (ESV), and is a measure of how much blood the heart pumps in each stroke. Ejection fraction (EF) is the ratio of stroke volume to end diastolic volume, measuring the percentage of the end diastolic volume being emptied in each stroke. EF is a good indicator of general heart state and a useful predictor of clinical outcome [60, 102, 107], as cardiac diseases often leads to reduced heart contractility [53]. SV and EF are often called global parameters, since they only take the total volume of the left ventricle into account. Global parameters can therefore be used to express the general state of the heart, but they cannot be used alone to predict which part of the cardiac muscle is affected.

In a population of healthy adults, the EDV measured by 2D echocardiography is on average 127 ml in men and 98 ml in women, and the ESV is on average 50 ml in men and 42 ml in women, with an EF of approximately 60% for both sexes [116].
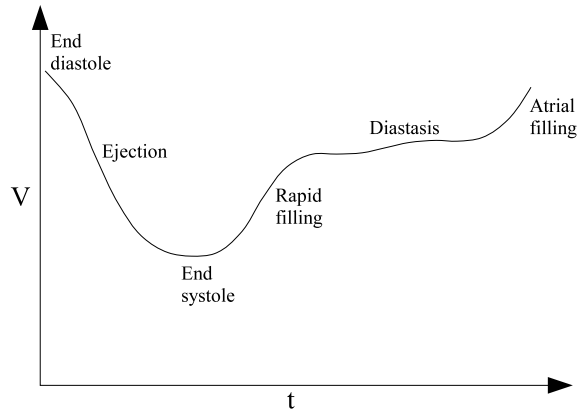
Figure 2.2: *This figure shows a left ventricular time-volume curve for one cardiac cycle measured in 3D ultrasound using deformable model segmentation.*

Ventricular volumes and EF are assessed by measuring the volume enclosed by the endocardium and the mitral annulus. Clinical recommendations state that papillary muscles should be regarded as part of the blood volume [56]. In clinical routine, volumes are measured either by eyeballing, or by manual tracing in one or multiple views. Different protocols for volume computations from manual tracing exist, but the most common way is to use the bi-plane method of disks (MOD) [85, 86], where the endocardial boundary is traced manually in two nearly orthogonal images, whereby a stack of ellipses are fitted to the traced contours. Volumes and EF have also been measured in a single view using the area-length formula, $V = 8A^2/(3\pi L)$ [86] where A is the area enclosed by the traced contour, while L is the length from the base to the ventricular apex.

2.3   Echocardiography

2.3.1   Ultrasound and interaction with tissue

Ultrasound is mechanical pressure waves with frequencies inaudible for the human ear. In medical applications, the typical frequency range is 2-10 MHz, similar to electromagnetic radio frequency (RF) waves. In soft tissue, the speed of sound is approximately as in water; $c =$1540 m/s, giving wavelengths λ in the range of 0.15-0.75 mm. Ultrasound wave pulses are typically generated using an ultrasound transducer consisting of piezo-electric crystals that vibrate when exposed to a high frequency electric potential. As a transmitted ultrasound pulse traverses through the body, it interacts with the tissue, causing scattering, reflection, and absorption of the wave energy. Absorption is caused by the conversion of kinematic wave energy to heat, and is frequency dependent. Scattering occurs when the ultrasound wave interacts with objects of size less than the wavelength, whereas reflections occur at tissue interfaces. Scattering and reflection causes a fraction of the transmitted energy to be echoed back to the transducer. These echoes can be measured as electric signals generated by the same piezo-electric crystals used to transmit the pulse. The time from the pulse is transmitted, to the echo is received is proportional to the depth of the scatterers or tissue interfaces. From Huygens' principle, the wave front generated by a piezo-electric crystal can be treated as
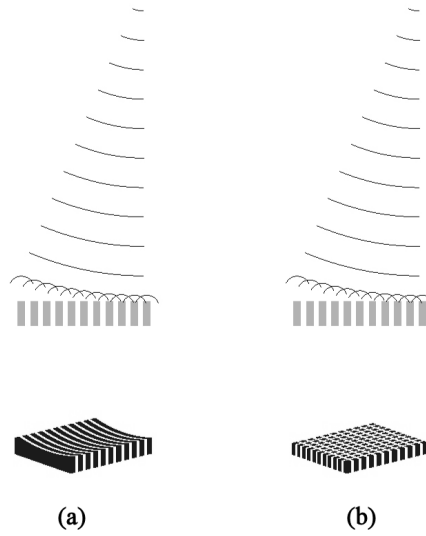
Figure 2.3: *Steering and focusing of the ultrasound beam is done by applying small individual delays to the signal passed to each element, thereby shaping and directing the wavefront propagating from the transducer. Similarly, delays are applied to the received signal from each element to obtain a receive focus and steering. Figure (a), shows a transducer consisting of a linear array of elements. This type of transducer can produce 2D images by steering and focusing the beam in a single plane. Figure (b) shows a 2D matrix transducer. By controlling each of the elements of the 2D matrix transducer individually, the ultrasound beam can be steered in both the azimuth and the elevation plane to produce 3D images. Adapted from [15].*

spherical waves transmitted from an infinite number of infinitesimal point sources that are superimposed. Since every crystal has a certain size, or aperture, the resulting wave front is directive and can be thought of as a beam. In its focus point, the width D_F of a focused beam can be computed from:

$$D_F = \frac{\lambda}{D}F = F_\# \lambda \,,$$

(assuming -3dB beam width), where λ is the wavelength, F is the distance to the focus point, D is the aperture diameter, and $F_\#$ is the F-number of the imaging system. A larger aperture gives a more focused beam. The attenuation compensated amplitude of the received signal at a time τ after the wave pulse was transmitted, is primarily a function of impedance gradients of the tissue in a limited spatial volume at depth $r = c\tau/2$, with a radial resolution determined by the length, or bandwidth, of the transmitted pulse, and lateral resolution determined by the aperture and frequency. The combination of radial and lateral resolution determines the system's acoustic point spread function.

2.3.2   Transducer types

Ultrasound enables us to do spatially localized measurements of tissue properties, which is the fundamental condition for ultrasound imaging. In medical imaging, an ultrasound image is a meaningful visualization of localized measurements of tissue properties, where an important criteria is to display as detailed and precise information as possible, with high
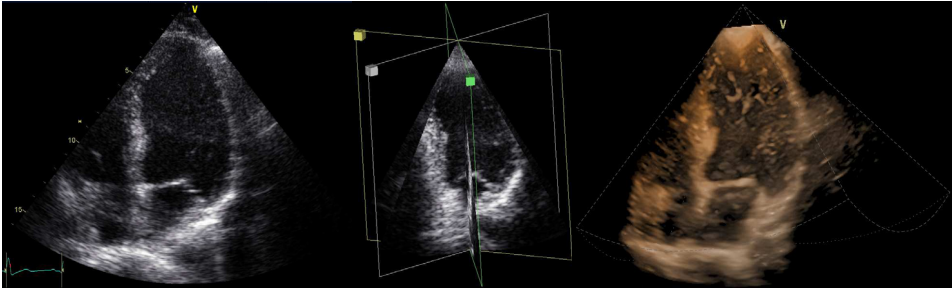
Figure 2.4: *A conventional 2D echocardiogram (left), triplane echocardiogram (middle) and 3D echocardiogram (right).*

spatial resolution and signal to noise ratio (SNR). The design of the transducer is therefore important. Improved lateral resolution can be achieved by an increased aperture, or by focusing the ultrasound beam, either by curving the transducer surface, or by using a phased array of multiple piezo-electric crystals.

Delaying the transmit signal to the central elements gives coherent oscillations in the focus. An electronically controlled phased array can also permit dynamic receive focus depth to match the depth of the origin of the echo, by rapidly decreasing the processing delay of the central elements.

Depending on the clinical application, different types of transducers can be used, including single element transducers, annular arrays, phased arrays, linear arrays, or matrix transducers, each using different focusing techniques [5]. For trans-thoracic echocardiography, the most commonly used transducer types are the phased array transducers and matrix transducers. Phased array transducers have a fixed focus in the elevation plane, but in the azimuth plane, the ultrasound beams are steered by gradually increasing the delay of the signal to each transducer element individually, allowing the formation of 2D images, as shown in Fig. 2.3. Multiplane images can be produced using 2D phased array transducers by mechanically rotating the transducer [104], whereas 3D data can be acquired by mechanically sweeping the transducer, while recording its position and orientation using a position sensor [36, 73]. Matrix transducers allow for steering in both the azimuth and elevation plane, and are thus capable of producing 2D, multiplane, and 3D data directly. Examples of 2D, triplane, and 3D images are given in Fig. 2.4.

2.3.3   Ultrasound processing chain

In a modern ultrasound scanner, the analog electric signal from each of the transducer's elements are converted to digital representations by a receiver. To produce images, these signals have to be processed through a series of steps as shown in Fig. 2.5.

Focusing and steering of the received signals are done by the beam former by applying small delays on each signal, before summation. Due to attenuation of ultrasound energy by the tissue, the reflected signal from deep structures will have smaller amplitudes than reflections from shallow structures. When displaying an ultrasound image, it is desirable to have comparable intensities from different depths, and attenuation compensation is therefore applied, usually assuming a simplified attenuation model with exponential decay, even if different tissue types contributes with varying attenuation.
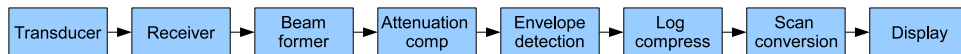
Figure 2.5: *Building blocks of the ultrasound processing chain.*

Information about tissue properties in the received RF signal is mainly encoded in the signal's amplitude. The signal amplitude is therefore extracted by envelope detection. In a digital system, the detected amplitude has a high dynamic range, typically encoded as 16 bit integer values. If this signal was mapped directly to pixel intensities, usually in the range 0-255, signal from weak scatterers would be suppressed, and only specular reflections would be visible on the screen. Log compression is therefore used to achieve a non-linear compression of the dynamic range. Log compression is typically on the form

$$y = log(ax + b) \quad,$$

where the output signal y is a compressed version of the input signal x, with compress gain a and compress offset b.

In general, the beam layout used during image acquisition does not match the bitmap shown by the display unit. The detected and log-compressed signal must therefore be transformed geometrically to match the coordinate system of the display, through a process known as scan conversion. For linear array transducers, this is merely re-sampling and interpolation of the beams, whereas for fan geometries, the coordinates are mapped through a polar coordinate transform.

2.3.4   Challenges for segmentation

Based upon the physical principles of ultrasound imaging, the images produced by echocardiography have several limitations that will have impact on segmentation algorithms:

1. The relationship between image intensity and physical properties of the tissue is highly complex. Whereas other imaging modalities such as X-ray have a relatively simple Lambert-Beer relationship between signal strength and physical tissue properties, the pixel intensity in an ultrasound image is a result of both reflection and interference patterns. Different types of tissue are therefore often distinguished by subtle variation in speckle patterns rather than intensity changes. This has also impact on the intensity distribution in ultrasound images, giving a highly non-Gaussian density [16, 106].

2. When imaging through the thorax, the transducer array must be small enough to fit between the ribs. Compared to abdominal ultrasound images, the small aperture of echocardiographic transducers gives broader ultrasound beams, resulting in relatively low resolution. Shadowing caused by ribs can further reduce the effective aperture, and in turn, the image resolution. Fat absorbs acoustic energy, causing a low signal to noise ratio in obese persons, in turn, giving low quality images with poor contrast.

3. The ultrasound image is highly anisotropic, particularly because the radial resolution is usually higher than the lateral resolution. Also, transducers have a fixed transmit focus within a single frame, making the lateral resolution depth dependent. Additionally, the angle dependency of the effective aperture causes lower image resolution at the extremes of the imaged sector.
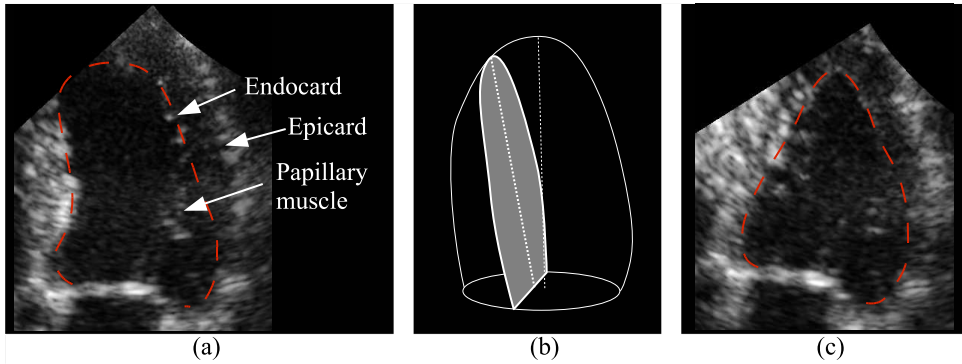
Figure 2.6: *The fibrous structure of the epicardium can in many cases show up clearer than the endocardium, making accurate tracing of the endocardial border (red dashed line) difficult (a). Also, papillary muscles are distinct structures that by convention should be included in the volume of the chamber. Foreshortening occurs when the image plane is oblique to the ventricular main axis (b). The result is a smaller apparent chamber volume (c).*

4. Shadows occur when ultrasound energy is completely absorbed or reflected, causing parts of the imaged object to disappear, a phenomenon known as dropouts. This effect is typically caused by the lungs and ribs or insufficient transducer-skin contact, causing parts of the heart to be invisible in the images.

5. Artifacts occur when reflections from shallow structures are reflected again by the transducer surface and back into the tissue, giving rise to reverberations that can appear as artificial structures in the image. Similar effects can also arise from grating lobes that are caused by the limited aperture of the phased array transducer used in echocardiography.

6. The intensity of the back-scattered signal depends on transducer orientation relative to the myocardial muscle fiber direction. The back-scattered signal intensity is low when the ultrasound beam is parallel to the fiber direction. Also, the definition of the endocardium can vary between regions in the image, as the blood/tissue interface is often almost tangential to the ultrasound beam direction. The folded structure of the trabeculae also gives lower reflection of acoustic energy. As illustrated by Fig. 2.6(a), the epicardium is often depicted more clearly than the endocardium in ultrasound images because of its fibrous structure giving more reflected ultrasound energy.

7. The temporal resolution of ultrasound imagery is limited by the speed of sound. Especially for 3D echocardiography, this can have impact on the effective frame rate. A too low frame rate causes under-sampling of the LV contraction pattern. In addition, de-correlation of speckle can impair algorithms that rely on matching image data from different frames.

Additionally, several challenges are imposed by clinical and physiological factors. The most important factors being:

1. Foreshortening occurs when the imaging plane slices through the heart in an oblique angle. In echocardiography, this is a particularly challenging problem compared to other imaging modalities, because of the dependency upon a few thoracic windows.

Foreshortening causes reduction in the apparent ventricular length, and the area of the imaged cavity is reduced, as illustrated in Fig. 2.6(b-c)

2. Papillary muscles are often prominent features in the echocardiograms, with acoustic properties similar to myocardium. For volume measurements, papillary muscles are by convention usually treated as part of the LV cavity [56, 86] as illustrated by Fig. 2.6(a). Consequently, a contour or surface used for volume measurements should not necessarily be co-located with blood-tissue transitions. The intensity transition from blood to myocardium is often lower than the transition from myocardium to epicardium, and the endocardial boundary is therefore not even co-located with the strongest image gradients.

All of these factors have impact on measurements of cardiac parameters. Low contrast, speckle, dropouts, and reverberation make it difficult, even for an expert, to accurately trace the endocardial boundary. Low image resolution and foreshortening have direct impact on the measured volumes.

2.4   A historical perspective

The development of medical ultrasound for diagnostic purposes began in the late forties, with the Austrian Dussik brothers [28, 29] and the British John J. Wild [108]. Using an ultrasound radar simulator operating at 15 MHz, Wild was able to measure the thickness of dog intestines in 1949 by plotting the amplitude of the received signal as a function of time, a technique known as A-mode imaging. The first M-mode images of the heart, where reflected amplitude is plotted as a function of depth and time, were published in 1954 [30], and this modality was quickly adopted to clinical use.

Computerized estimation of LV volumes and EF was first reported twenty years later by Ledley and Wilson [57]. They used regional thresholding along with a simple contour finding algorithm to extract the blood/tissue interface from digitized M-mode images, whereby the LV volumes could be computed using an ellipsoid approximation.

2D images of the heart were first reconstructed from M-mode images, and the first real-time scanner that gained popularity was developed by Bom in 1971 [10]. Ten years later, in 1980-1981, Skorton et al. [91, 92] presented one of the earliest automated segmentation algorithms for this modality, based upon a combination of histogram thresholding and Sobel edge detection. During the early 80s, several approaches towards segmentation algorithms for 2D echocardiography were presented based upon similar techniques, using combinations of image smoothing, thresholding, and edge operators along with simple heuristic boundary search algorithms [1, 18, 111, 112]. Later in the 80s, advances were made towards the utilization of temporal information in sequences of 2D echocardiograms, by introducing optical flow [61]. The first step towards "modern" segmentation techniques in echocardiography was reported in 1988 by Klinger et al. [54]. They used simulated annealing and a Markov random field to adapt what we today recognize as an active contour, or deformable model, to the image data. Image features, shape smoothness, contour area, and temporal behavior were modeled as separate elements that each contributed to the deformation of the contour. This technique showed many resemblances to the "Snake" active contour presented by Kass et al. [52] the same year, where the shape of an evolving contour was modeled through its internal and external energies.

During the 90s and 00s, the contributions to automated segmentation of 2D echocardiography were too numerous for a thorough review in this brief introduction, but some contributions

where particularly noticeable. Chalana et al. [17] used an extension to the snake algorithm to simultaneously track the endocardial and epicardial borders. In addition to using classical edge detection algorithms, Kucera and Martin [55] used a finite element model with region based external forces, which deformed the contour by maximizing the difference between image data inside and outside the contour, according to a statistical model of image intensities. Jacob et al. [47, 48] formulated contour detection and tracking in a Kalman filter framework, where a kinematic model describes the temporal behavior of the contour, which was modeled by a shape space model, trained on contours extracted manually from previously seen images. The problem was formulated in a Bayesian framework, and the Kalman filter was used to track the contour through the cardiac cycle. Bosch et al. [12, 97] proposed to use active appearance motion models (AAMMs) for time-continuous segmentation of 2D echocardiograms. Spatial and temporal shape properties of the endocardial wall, as well as texture information in the echocardiograms, were modeled using an active appearance model trained on a database of segmented training images.

The earliest approaches to 3D echocardiography were based upon reconstruction from 2D images, and were reported already in 1974 [26]. Development in matrix transducer technology later allowed for simultaneous acquisition of two orthogonal imaging planes [93], and later real-time volumetric imaging [88]. Automated segmentation algorithms for 3D echocardiograms were presented in the late 80s by Tamura [101], by segmentation in multiple 2D images using a radial search algorithm. Song et al. [94] formulated the surface detection problem in a Bayesian framework, to combine image evidence with prior knowledge of LV shape. They used a convex combination of manually delineated surfaces from a database of shapes to represent the LV surface. Synthetic images where generated from the surface representation using the distance transform, and matched to the data. The Nelder-Mead simplex algorithm was used for the optimization. Angelini et al. [4] used snakes driven by an expanding balloon force in multiple short axis views extracted from the 3D data set. A finite difference approximation scheme was used for optimization. Pre-processing of image data was done using multidimensional space-frequency analysis with brushlet functions and nonlinear denoising enhancement to reduce the impact of speckle, and a gradient operator was used to extract image energies. Montagnat et al. [70] formulated the surface representation using simplex meshes, extending the basic 3D formulation to four dimensions (4D) by introducing temporal constraints. 4D anisotropic diffusion was applied to the image data to reduce speckle, and external forces were generated as a combination of gradient forces and regional edge evidence. Corsi et al. [25] used a 3D level set method on each frame in the 4D image sequence. Based upon a traditional inflation speed function, they proposed a new term to the level set equation tailored to echocardiography, reducing leakage of the surface through "holes" caused by missing data. A Laplacian of Gaussian filter was used for pre-processing, and image force extraction. Initialization was done manually to get an initial surface close to the true endocardial boundary. Van Stralen et al. [98] introduced a semi-automatic endocardial border detection method for 4D echocardiography. Data were acquired using a fast rotating phased array transducer, and segmentation was done in the rotated 2D images. Edge templates were extracted from manually traced contours in four views. The interpolated edge templates were used for tracking in both time and space. Dynamic programming was used for contour detection, and internal surface properties were modeled through continuity constraints during optimization. Orderud [77] used a Kalman filter based approach, where the LV surface was modeled using a rigid ellipsoid model. The Kalman filter was used to estimate the ellipsoid pose, including position, orientation, and scale, and the transition criterion [83] was used to extract contour position estimates. This is the first report on real-time segmentation in 4D
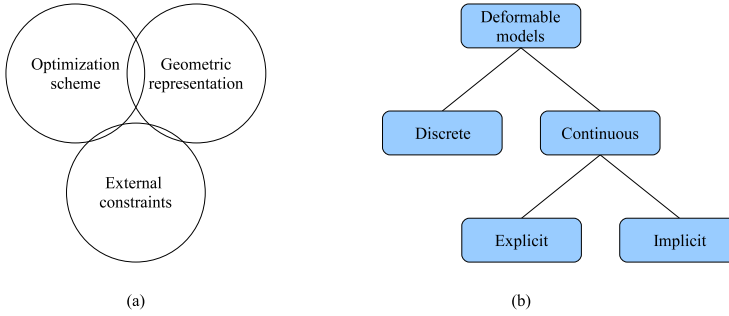
Figure 2.7: *Segmentation algorithms based upon deformable models are characterized by three hallmark criteria (a), including the geometric representation, external constraints, and optimization scheme. They are further classified by their geometric representation (b) (adapted from [71]).*

ultrasound. The method was later improved [78] by allowing for local deformations through a quadratic spline-based surface representation coupled with a global pose transform, giving higher regional accuracy than the rigid ellipsoid model. Following this approach, Hansegård et al. applied this Kalman filter framework to active shape models (ASMs) [42]. The ASM was trained on LV surfaces traced using a semi-automated clinical segmentation tool (4DLVQ), resulting in a shape space of physiologically realistic LV shapes, while retaining real-time performance.

The advances within automated analysis of echocardiograms has been numerous, and several reviews exist [34, 39, 64, 74, 90]. For historical reviews of the development of echocardiography, see [31, 32, 51, 82, 89, 105].

2.5   Deformable model segmentation

As illustrated by the previous sections, a wide variety of segmentation algorithms for echocardiography has been studied. Many of these methods are based upon, or have strong resemblance with, a particular class of segmentation algorithms known as deformable models. Deformable models are mathematically defined contours or surfaces that are fitted to the target image under the influence of internal and external constraints. A brief overview of different types of deformable models, and their application to echocardiography, is given in the following sections. The references found in these sections are not meant to constitute an exhaustive list of techniques, but they are included to illustrate the similarities and differences between different deformable models.

Segmentation algorithms based upon deformable models can be characterized by three hallmark criteria, shown in Fig. 2.7(a). *First*, different types of deformable models are classified by their geometrical representation. The geometrical representation usually dictates a set of shape constraints that describe its ability to represent different structures. *Second*, different deformable models use different types of optimization strategies to fit the model to the derived external constraints, while complying with the internal constraints. *Third*, the segmentation algorithms are associated with a set of external constraints, determining how image features are translated into shape changes. These external constraints can often be interchanged between different types of deformable models, but restrictions may be dictated by the surface representation and optimization scheme.

2.5.1   Geometric representations

Already when it comes to choice of surface representation, a great number of formulations have been proposed. A simple categorization of the different types is given in Fig. 2.7(b). The most striking difference between different approaches is whether the model is defined continuously over the contour or surface, or discretely in a finite number of points [71]. Typical examples of discrete models include triangulated meshes [98], simplex meshes [38, 71], discrete contours or snakes [52], and point distribution models [23, 24, 48]. When it comes to continuous models, it is natural to separate between implicit and explicit representation. Implicit models, such as level sets [25, 79, 110], describe the contour or surface embedded in a higher dimensional function. Explicit models have a direct geometric formulation, and include for instance spline models [33], deformable templates [66], or superquadrics [14, 103].

All deformable models have some kind of internal constraints, allowing apriori knowledge of the target object's shape to be taken into account. These constraints enforce some degree of smoothness or continuity to the model, for instance by counteracting stretching or bending. The term shape space is often used to describe the space of all possible shapes that can be represented given the geometric representation. Internal constraints enforce the model to lie within a subset of all possible shapes, hopefully giving a subset of shapes that are better suited for the segmentation problem at hand. In the snake formulation, the internal constraints restrict the shape space to a sub-space of smooth models, and have curvature preserving properties (see section 3.1). Variations over this theme are widely used, for instance in level-set based segmentation [25] and simplex meshes [70]. Point distribution models are a special type of geometric representation, by representing the shape by an average shape with deformation modes superimposed, giving a linear shape space (see section 3.2). In this formulation, internal constraints can be formulated by restricting the maximum allowed amount that is added of each deformation mode. Internal constraints can also include some sort of regularization towards a template shape [48].

2.5.2   Optimization techniques

It is common to formulate the deformation process as the minimization of an energy functional, often denoted by the symbol E. This energy represents the total quality of the geometry, its distance to a reference shape or state, and its quality of fit to the external constraints. Another way of formulating the deformation process is to introduce a law of motion for the deformable model, and internal and external constraints are therefore formulated as forces acting upon the model. The optimization technique is highly dependent on the choice of geometric representation, but a few basic approaches are more common than others. For snakes, Kass et al. [52] used an iterative gradient descent technique solved using finite differences, and the contour could be trapped in local minima. Amini et al. [2] used dynamic programming to obtain a globally optimal solution. This approach has later been used in several works [83, 98]. Level-set methods involve various optimization schemes, most commonly finite differences [25] or fast marching [110]. Active shape models can be solved directly or iteratively using a least squares fit. If a motion model is incorporated, the Kalman filter can give a similar closed form solution [9, 48]. For simplex mesh fitting, a law of motion is commonly applied and solved iteratively using finite differences [38, 71]. Active appearance model fitting is done using a gradient descent technique, but with precomputed "gradient matrices" trained on a training set to reduce the computational complexity.

2.5.3   Enforcing temporal continuity

In segmentation of echocardiography, temporal continuity is an important aspect. Some features are not visible in still images, but emerge in image sequences. There are several ways to enforce temporally consistent segmentation. The simplest method is by initializing the segmentation by the result in the previous frame [68]. In off-line processing, where all image frames are known beforehand, temporal constraints can be applied through the external constraints, by enforcing for instance first- or second order temporal continuity [35], by regularization the displacement of mesh nodes against a model of the ventricular contraction pattern [38], or by introducing regularization trajectories [70]. Active appearance models can also be extended by incorporating the ventricle's contraction pattern in a phase-normalized heartbeat into the model, giving temporal 2D segmentation [12]. Here, temporal continuity is modeled through an average contraction pattern with different contraction variations superimposed, giving a smooth temporal behavior. In 2D echocardiography, temporal continuity can be achieved using a 3D segmentation approach, where time is treated similarly to spatial dimensions, making phase normalization unnecessary [69]. Temporal continuity can also be achieved using a motion model [9, 48] giving stochastic control of the temporal behavior of a parametric model such as the point distribution model.

2.5.4   External constraints

Deformable models are influenced by external constraints that drive the model deformation towards image features, for instance strong edges. Usually, the choice of external constraints is highly application specific, but for echocardiographic data most approaches can be categorized into four main groups.

Force fields. A common method is to pre-compute a 2D or 3D force field that is valid for the full image domain. Since this force field is defined everywhere in the image, the forces acting on a point on the deformable model can be looked up directly from this force field. Kass et al. [52] used a simple gradient operator to produce an edge magnitude image, also known as a potential surface [67]. The forces acting on the contour were computed by taking the gradient of this potential surface. Another way of computing the force field is by taking the gradient of the smoothed edge map produced by an edge detector such as the Canny detector [35]. A similar approach is to use optical flow [61, 67], where the force field represents estimates of motion present in image sequences. In noisy and anisotropic ultrasound data, multi-scale implementations are often used [17, 52] to avoid being trapped in local minima. The computation of the force field can be computationally expensive, but in an iterative optimization scheme such as for snakes or level sets [25], the cost can still be justified, since the force computations are done only once for the entire image domain.

Pixel profiles. A different approach is to search for edges or other image features along pixel profiles sampled across the deformable surface or contour. These profiles can either be sampled along the contour or surface normal [66], radially relative to a defined center of gravity [87], or using other geometries [41]. Several different approaches towards feature extraction along such profiles have been proposed. Setarehdan et al. used fuzzy multi-scale edge detection using wavelets [87]. Mignotte et al. [65] searched for points that maximized the differences between the average pixel intensity between the outer and inner part of the profile. Rabben et al. [83] used a similar approach, but

also minimized the standard deviation of pixel values inside and outside the selected point. Blake et al. [8] used normal displacements as a method to map extracted pixel profiles onto a parametric contour, using assimilated wavelet-based features. For active shape model segmentation, the Mahalanobis distance between extracted contours and a statistical edge model is often used [22], but because of the high intensity variations and speckle noise found in echocardiographic images, this approach has had limited use in this modality.

Regional constraints. The third approach of computing external constraints, suitable for closed contours or surfaces, is to evaluate properties of image regions within and outside of the deformable model. The external constraint is formulated to produce boundaries that give a best possibly separation of homogeneous regions with respect to for instance intensity distribution [55, 66]. Angelini et al. formulated this in a Mumford-Shah level set framework [3].

Template matching. In the active appearance model, external constraints are formulated as template matching. Changes to the model parameters are computed from the pixel difference between the image and a template generated by the model [19].

2.5.5   Hybrid models

Each type of deformable model has different strengths and weaknesses. Several authors have therefore proposed hybrid models, combining aspects of different approaches, in an effort to overcome typical shortcomings of the "standard" algorithms in a particular segmentation problem. Often, hybrid methods involve initializing a method b with outputs from a method a, such as in Oost et al. [76], where segmentation by active appearance models was combined with a dynamic programming-based snake to improve the final segmentation accuracy of X-ray angiography. Another approach is to switch back and forth between two methods such as in Mitchell et al. [69], where two approaches, active shape models and active appearance models, were combined to help avoid local minima and to improve segmentation results of cMRI data. Hansegård et al. [42] proposed to combine active appearance models with the snake algorithm in a fully integrated manner, such that for each iteration, the new output of the snake algorithm was used to correct the active appearance model (paper C). The output from the active appearance model was in turn used as an initial contour for the snake.

Chapter 3

Background for papers

The characteristics of echocardiograms put special design requirements on an automated segmentation algorithm. Additional requirements arise since the ultimate goal of the method is to be used in a clinical setting. An automated volume measurement algorithm should be:

Accurate. The method should produce accurate contours, and be able to follow weak edges. The accuracy of volumes derived from the detected contour or surface must agree well with the true LV volume.

Interactive and compliant with clinical conventions. The algorithm must produce physiologically realistic contours or surfaces that the user will agree upon. Given different individual and institutional conventions for interpretation of echocardiograms, this can be challenging. Therefore, the method should allow for manual interaction to adjust the results according to the clinician's judgment.

Efficient. A standard echocardiographic examination is usually performed in 10 to 30 minutes, and because of the increasing cost of modern healthcare, this time should ideally be shortened. A clinical volume measurement method should therefore be fast and easy to use, taking no longer than 2-3 minutes for analysis of one patient, including manual initialization and interaction. Otherwise, the clinician will stick to manual tracing or even eyeballing.

Repeatable. The method must be repeatable. This means that the method should give comparable results each time the analysis is repeated, even if this is done by a different clinician. In addition, it should give reproducible results in different images of the same patient. This restricts the amount of manual interaction that can be permitted.

Robust. The method should handle the various challenges related to echocardiography. It should be insensitive towards spatial intensity variations and reverberations. It should be able to handle non-Gaussian pixel distributions and low contrast between blood and tissue. Further, the method must handle varying image resolution, and missing wall evidence in parts of the image. For time sequences, the method should be invariant to frame-rate, and it must not be affected by valve motions and intensity changes over time, as well as changes in speckle pattern.

In summary, the algorithm should provide accurate, rapid, robust, repeatable, and clinically acceptable results in data with relatively poor image quality. From chapter 2, it is obvious that the number of different algorithms available for segmentation of echocardiograms is
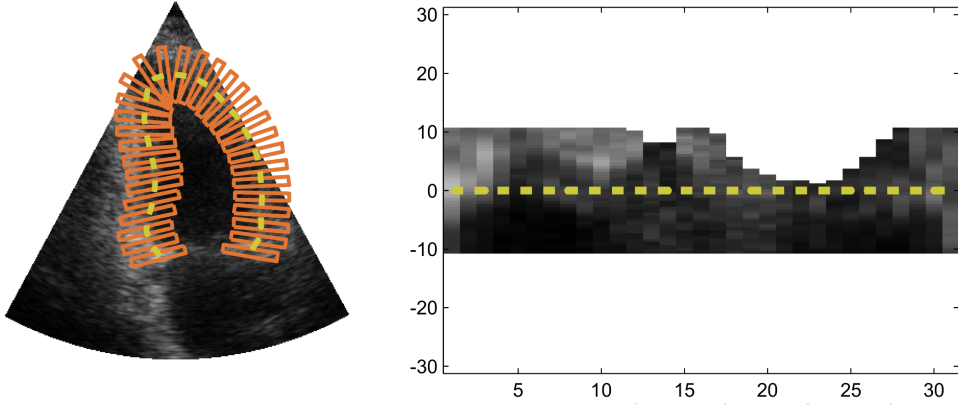
Figure 3.1: *Extraction of edge profiles for segmentation by dynamic programming.*

huge. This thesis focuses on three main techniques, which are believed to adhere to the above requirements.

Snakes solved by dynamic programming are well known within the field of image analysis. Their main advantage is that the dynamic programming algorithm efficiently provides optimal solutions given the chosen constraints, and they can easily be controlled by manual interaction. Active shape models and active appearance models were chosen mainly because they produce physiologically realistic shapes, and they generalize well to new data. Active appearance models were of particular interest, since they do not rely on finding the strongest edges in the image, but use a texture matching technique that permits incorporation of clinical conventions. The subsequent sections give brief introductions to the different segmentation algorithms used throughout this thesis.

3.1 Dynamic programming

Active contour models, also called snakes, introduced by Kass et al. [52], are among the most popular segmentation algorithms for medical imagery, and this work is cited in most works involving deformable models. Snakes are based upon an energy-minimizing contour, guided by external energies E_{ext} derived from image features. Smoothness of the contour is controlled by internal energies E_{int}, making the contour able to resist bending. Parametrically, the geometry of the snake can be described by a 2D curve $\mathbf{v}(s) = (x(s), y(s))^T$, and the total energy of the snake can be written as

$$E_{snake} = \int_0^1 E_{int}(\mathbf{v}(s)) + E_{ext}(\mathbf{v}(s)) \, ds \quad . \tag{3.1}$$

This equation is usually discretized and solved in an iterative gradient descent fashion. An alternative, introduced by Amini et al. [2], is to use dynamic programming [7] as a rapid way of finding a globally optimal solution to (3.1). Dynamic programming is based upon the principle of optimality, stating that in an optimal sequence of choices, each sub-sequence must also be optimal.

In boundary detection, this is done by first sampling edge features from candidate points along search profiles across a pre-defined initial contour, as shown in Fig. 3.1. The edge evidence value at each candidate point along each search profile is represented by nodes with

cost x_k^m in a graph, where $k = 1 \ldots n$ is the candidate point index in a single candidate profile, and $m = 1 \ldots M$ is the search profile index. The number of layers M in the graph equals the number of search profiles, and the number of states n in each layer represents the number of candidate points extracted along each profile.

Each graph node is assigned a cost value depending on its edge evidence, such that a candidate point with a strong edge has a low cost. In addition, the partial paths between nodes in neighboring layers are assigned partial path costs $g^m(i,k)$ representing the cost of a transition from node x_i^m to node x_k^{m+1} depending on some smoothness criterion. Typically, the cost of a large state jump is high, whereas a low cost is associated with staying in the same state when moving between two layers. The optimal path through this graph can be found using the following algorithm [96]:

1. Specify initial costs $C(x_i^1)$ of all nodes in the first graph layer, $i = 1, \ldots, n$ and partial path costs $g^m(i,k)$, $m = 1, \ldots, M-1$

2. Repeat step 3 for all graph layers $m = 1, \ldots, M-1$.

3. Repeat step 4 for all states $k = 1, \ldots, n$ in graph layer m.

4. Let the cost of choosing node k in the next graph layer $m+1$ be expressed as

$$ C\left(x_k^{m+1}\right) = \min_{i=-w/2\ldots w/2} \left[C\left(x_{k+i}^m + g^m(i,k)\right) \right] \quad, $$

where w is a search window width. Set pointer from node x_k^{m+1} back to node x_i^{m*}; where * denotes the optimal predecessor.

5. Find an optimal node x_k^{M*} in the last graph layer M and obtain an optimal path by backtracking through the pointers from x_k^{M*} to x_i^{1*}.

Using this algorithm, $E_{\text{snake}} = C\left(x_k^{M*}\right)$ represents the minimum total snake energy. External energies are encoded in the node costs, whereas internal constraints are encoded in the partial path costs.

The dynamic programming based snake was used in combination with other segmentation techniques by Hansegård et al. [41, 42] (papers A and C).

3.2   Active shape models

Statistical models include a family of models that are trained on a set of observations of the object of interest. The model's ability to represent new shapes or appearances depends on the variability of the examples in the training sets.

Point distribution models (PDMs) introduced by Cootes and Taylor [23] have become very popular for modeling the shape of objects where some degree of similarity is found between objects of the same class. PDMs capture the average shape and shape variations found in a training set, and are parameterized such that each model parameter represents orthogonal deformation modes.

Training is usually done by manual or automated annotation in images of the object class of interest. The resulting annotations are normalized using Procrustes alignment [37] to remove trivial pose variations, including scale, rotation and position, giving the normalized shapes $\mathbf{x}_i = (\mathbf{v}_1^T, \mathbf{v}_2^T, \cdots, \mathbf{v}_m^T)^T$. Here, $\mathbf{v}_i = (x_i, y_i)^T$ in 2D images, and $\mathbf{v}_i = (x_i, y_i, z_i)^T$ in 3D images. After estimating the average shape $\bar{\mathbf{x}}$, PCA [80] is applied to the shape vectors to
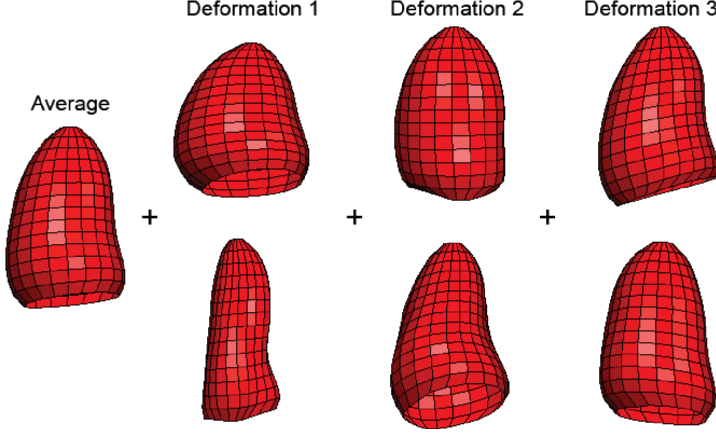
Figure 3.2: *The figure shows both the average shape, and the three first deformation modes (mean ± 3 standard deviations) of an LV point distribution model trained on echocardiograms from 21 patients.*

obtain the matrix $\mathbf{P}_s$ containing the corresponding eigenvectors. Any shape $\mathbf{x}$ in the training set can be reconstructed through the average shape and a linear combination of deformation modes as:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s \mathbf{b}_s \ . \tag{3.2}$$

The eigenvectors $\mathbf{P}_s$ now represent orthogonal deformation modes around the average shape, and the relative contribution of each eigenvector in each training set is determined by the corresponding eigenvalue λ_i. A compact shape description is obtained by discarding eigenvectors with little contribution to the shape variation in the training set. New shapes can be generated from this model by selecting a parameter vector $\mathbf{b}_s$, and computing the resulting shape using (3.2). Typically, the elements of $\mathbf{b}_s$ are restricted to lie within some proportion of the corresponding eigenvalue to ensure statistically plausible shapes. An example showing the variation of the three first deformation modes of a 3D PDM is given in Fig. 3.2.

Active shape models extend point distribution models with an update scheme to optimize the model parameters together with the parameters of a global pose transform to achieve best possible fit to the target image. In the 2D case, the segmentation algorithm uses the following iterative scheme [23]:

1. Create an instance $\mathbf{X}(\mathbf{b}_s, \mathbf{p})$ of the shape model with parameters $\mathbf{b}_s$ and pose parameters $\mathbf{p}$ in the image frame by applying a pose transform $T_{\mathbf{p}}$.

2. Perform edge detection in pixel profiles sampled from the image along the contour normals at each of the model points to suggest new node positions $\mathbf{X} + d\mathbf{X}$.

3. Compute the incremental changes $d\mathbf{x}$ in the local model coordinate system by applying the inverse pose transform.

4. Compute the update to the model parameters from $d\mathbf{b} = \mathbf{P}_s^T d\mathbf{x}$.

This update scheme is equivalent to a weighted least squares algorithm, and can be used to obtain good fit in a few iterations [23].

Note that the terms active shape model and point distribution model are often interchanged in the literature, as it is in the papers of this thesis.

3.2.1 Kalman filter-based active shape models

ASMs can be fitted to measured image features using the least squares algorithm presented above. However, in image sequences of a moving object, independent segmentation in subsequent frames can lead to inconsistent temporal behavior and poor segmentation accuracy. One solution to this problem is to introduce a motion model [9] describing the temporal dynamics of the pose and shape parameters, or states $\mathbf{t}_k = \left(\mathbf{b}_s^T, \mathbf{p}^T\right)$, of the statistical model. This motion model takes the form:

$$\mathbf{t}_{k+1} = \mathbf{A}_1\mathbf{t}_k + \mathbf{A}_2\mathbf{t}_{k-1} + \mathbf{B}_0\mathbf{w}_k \ \ .$$

Here, the predicted model state $\mathbf{t}$ in the next frame $k+1$ is represented by a weighted sum of the current and previous states at times k and $k-1$ respectively, and a noise term $\mathbf{w}_k$ that allows prediction error to be incorporated in the model. The matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ can be adjusted to control properties such as damping and inertia, while the matrix $\mathbf{B}_0$ controls the uncertainty added by the prediction.

Edge detection is performed in extracted pixel profiles perpendicular to the predicted contour, similarly as for the basic ASM search, and suggested updates to the model's state parameters are computed to give best possible fit to the edge measurements. In addition, each edge measurement is associated with a measurement error that represents the uncertainty of the estimated edge position.

The states of the ASM can now be propagated through time as a multivariate Gaussian distribution using a Kalman filter [48, 50] that takes both the uncertainties of the motion model's prediction and the edge detection uncertainties into account. As for the basic active shape model, this gives a weighted least squares solution, but the motion model is used to regularize the temporal behavior of the model. The main advantage of this technique is that temporal consistency is obtained in a closed form solution, eliminating the need for iterative refinement. This approach was used by Hansegård et al. [40] for real-time and fully automatic volume measurements in 3D echocardiography using a 3D ASM (paper E).

Besides tracking of ASMs, this technique can also be used with other types of deformable models [8]. Orderud presented studies where the extended Kalman filter was used to segment and track the left ventricle in volumetric ultrasound data in real time using a rigid ellipsoid model [77], and using a spline based model [78].

3.3 Active appearance models

Active appearance models (AAMs) [19] extend the ASM by including a statistical image texture model in addition to a PDM. The texture model can generate new unseen images of the target object as the mean texture with a linear combination of texture variation modes superimposed.

In the same way as the shape of the object is represented by a coordinate vector $\mathbf{x}$, the object's texture is represented by a normalized texture vector $\mathbf{g}$, sampled from the training image. PCA is then applied to the entire cohort of texture vectors, to obtain a texture model with variation modes $\mathbf{P}_g$ formed by the eigenvectors of the covariance matrix of the texture vectors, the average texture $\bar{\mathbf{g}}$, and texture parameters $\mathbf{b}_g$:

$$\mathbf{g} \;=\; \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g \;. \tag{3.3}$$

This texture model can be combined with the PDM in (3.2) by a third PCA on the concatenated shape and texture parameters, leading to the equations for generation of new shapes $\mathbf{x}$ and textures $\mathbf{g}$:

$$\mathbf{x} \;=\; \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{c}$$
$$\mathbf{g} \;=\; \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{c} \;. \tag{3.4}$$

Here $\mathbf{Q}_g$ and $\mathbf{Q}_s$ denote the basis vectors for texture and shape of the combined model, and $\mathbf{c}$ are the model parameters controlling both shape and texture at the same time.

Application of AAMs to segmentation of image data is done using the following optimization strategy. The difference between the generated model texture $\mathbf{g}_m$ given by (3.4) and the normalized image texture $\mathbf{g}_s$ is expressed by the residual vector $\mathbf{r}(\mathbf{p}) = \mathbf{g}_s - \mathbf{g}_m$. Assuming linearity, the quality of the model fit is improved by modifying $\mathbf{p}$ according to

$$\delta\mathbf{p} = -\left(\frac{\delta\mathbf{r}}{\delta\mathbf{p}}^T \frac{\delta\mathbf{r}}{\delta\mathbf{p}}\right)^{-1} \frac{\delta\mathbf{r}}{\delta\mathbf{p}}^T \mathbf{r}(\mathbf{p}) \;, \tag{3.5}$$

where $\frac{\delta\mathbf{r}}{\delta\mathbf{p}}$ is precomputed during model training by applying perturbations to the model parameters and measuring the effect on the residuals. The classical fitting scheme for the basic AAM is an iterative procedure where model parameters are updated for each iteration using (3.5) according to the following algorithm [19]:

1. Sample the texture under the current shape $\mathbf{X}$ using the image normalization function $T_{\mathbf{u}}$, $\mathbf{g}_s = T_{\mathbf{u}}^{-1}(\mathbf{g}_{im})$.

2. Evaluate the residual vector $\mathbf{r}(\mathbf{p}) = \mathbf{g}_s - \mathbf{g}_m$, and the current error, $E = |\mathbf{r}|^2$.

3. Compute the predicted displacements according to (3.5).

4. Update the model parameters $\mathbf{p} \to \mathbf{p} + k\delta\mathbf{p}$, where initially $k = 1$.

5. Calculate the new points, $\mathbf{X}'$ and model frame texture $\mathbf{g}_m'$.

6. Sample the image at the new points to obtain $\mathbf{g}_{im}'$.

7. Calculate the new error vector, $\mathbf{r}' = T_{\mathbf{u}'}^{-1}(\mathbf{g}_{im}') - \mathbf{g}_m'$.

8. If $|\mathbf{r}'|^2 < E$ then accept the new estimate, otherwise, try at $k = 0.5$, $k = 0.25$, etc.

This procedure is repeated until no improvement is made to the error, and convergence is assumed. This scheme has been proven to work well in many applications, although different extensions have been proposed [20, 27]

Major extensions to the basic AAM include the multiview AAM [58, 75, 76, 113], where multiple view planes of the heart are modeled simultaneously. Bosch et al. [12, 13] introduced the active appearance motion model, AAMM, where the temporal behavior of the left ventricle is incorporated in a single model. Hansegård et al [42, 43] combined these models into a multiview and multi-frame model for segmentation of triplane echocardiograms (papers B and C).

Chapter 4

Summary of papers

4.1 Paper A

Knowledge based extraction of the left ventricular endocardial boundary from 2D echocardiograms

J. Hansegård, E. Steen, S. I. Rabben, A. H. Torp, H. Torp, S. Frigstad, and Bjørn Olstad
Published in *Proc. IEEE Ultrasonics Symposium*
Volume 3, 2004

The goal of the work resulting in this paper was to automatically detect the endocardial boundary of the left ventricle in 2D echocardiograms by incorporating knowledge of LV shape and intensity signature into the detection algorithm. The main contribution of this paper was a method where expert knowledge from a manually segmented case database could be integrated into a snake algorithm [52]. In traditional snake algorithms, detection is initialized by a single initial shape, and edge features are extracted directly from the patient image using some kind of edge model, such as the maximum gradient. To incorporate expert knowledge into this algorithm, multiple segmentations were performed taking initial shape and edge templates from each database case, or atlas. Candidate pixel profiles were extracted from the patient image and matched against corresponding edge templates computed from each atlas. The internal energy term of the snake was designed to preserve the curvature of the atlas shape, while the external energy favored high correlation coefficient between the candidate profiles and the edge templates. All database cases were mapped to the patient image using this technique. After segmentation, the endocardial contours were extracted, and the snake energies represented the goodness of fit for each atlas. The atlas with the lowest snake energy was assumed to provide the best shape, giving the best segmentation result.

This framework was evaluated on 19 patients by comparing automatically detected contours against manually determined contours using the derived EDV, ESV, EF, and the pairwise point distances at ED and ES respectively. The strength of this method was its ability to utilize expert knowledge directly in the snake algorithm. Good results were achieved in some cases, but the lowest total snake energy did not provide a sufficiently robust selection criterion.

4.2 Paper B

Detection of the myocardial boundary in the left ventricle from simultaneously acquired triplane ultrasound images using multi view active appearance motion models

J. Hansegård, S. Urheim, E. Steen, H. Torp, B. Olstad, S. Malm and S. I. Rabben

In this paper, we presented an algorithm for detecting the LV myocardial boundary from simultaneously acquired triplane echocardiograms. In triplane echocardiograms, LV shapes, textures, and contraction patterns are highly correlated between the three planes. By utilizing this correlation in a detection algorithm, we expected improved results compared to independent contour detection in the three image planes. The main contribution of this paper was a multiview extension of the active appearance motion models (AAMMs) [12, 97], capable of utilizing the correlation between multiple planes and frames inherent in echocardiographic recordings, to do spatially and temporally consistent segmentation. The AAM framework provided an elegant way of combining expert information about the LV shape and contraction pattern from multiple views into a single model.

The multiview AAMM was built using manually traced triplane echocardiograms from 20 participants. A spatial band-pass filter was used for speckle removal and normalization of the ultrasound images. The shapes and image textures extracted from the training sets from three planes and three frames, at ED, ES, and mid-systole, were combined using PCA, to form a compact model that included the major modes of variation of LV shapes, textures, and contraction patterns. This model was evaluated using a leave one out strategy, by comparing automatically determined ED and ES volumes and EF against volumes determined by manual tracing.

The results from the experiments were promising, with strong correlation and relatively small bias in volumes compared to manual tracing. The EF measurements were, however, not satisfactory, mainly caused by a "too normal" contraction pattern due to a relatively small training database.

4.3   Paper C

Constrained active appearance models for segmentation of triplane echocardiograms
J. Hansegård, S. Urheim, K. Lunde and S. I. Rabben
Published in *IEEE Trans. Med. Imaging*
Volume 26, Issue 10, 2007

This paper presented multiview and multi-frame active appearance models for segmentation of the left ventricle in triplane echocardiograms, and is a continuation of the work in **paper B**[1]. The paper's main contribution was a general way of integrating local edge detector based segmentation algorithms and manual interaction into the AAM framework. The method was based upon a constrained AAM algorithm, previously presented by Cootes [21]. The potential of this method was demonstrated by coupling the multiview AAMM with a snake [52] optimized using dynamic programming [2], with the gradient criterion [83] as local edge detector. For each iteration of the AAM, each point on the AAM shape was constrained towards the output of the snake, with a weight dependent on the local edge evidence estimated using the gradient criterion. Manual constraints were incorporated similarly, but with a higher weight, since manually defined landmarks were assumed highly accurate. Normalization of the ultrasound images was performed by mapping the skewed ultrasound histogram to a Gaussian distribution according to the method of Bosch et al. [12].

---

[1]Paper B was included as part of this thesis since the experience acquired during this work had several implications for the choices made in paper C.

We showed that combining the AAM with a dynamic programming based snake algorithm, gave more accurate and reproducible volume and EF measurements than the two methods alone. Constraining the model to manually defined landmarks further improved the results. We also demonstrated that volumes and EF estimates improved by using triplane echocardiograms compared to single plane data. The effect of "under-training" was still present, albeit in a lesser extent than for the unconstrained multiview AAMM.

4.4   Paper D

Semi-automated quantification of left ventricular volumes and ejection fraction by real-time three-dimensional echocardiography

J. Hansegård, S. Urheim, K. Lunde, S. Malm and S. I. Rabben

Submitted for publication

In this paper, we presented an evaluation of a clinical semi-automated volume quantification tool (4DLVQ) for 3D echocardiography. The main contribution of the work leading to this paper was a deformable model based segmentation algorithm[2] that provided accurate and repeatable estimates of LV volumes, while providing simple but powerful editing capabilities for manual correction of the detected surface. To avoid foreshortening during initialization of the model, manual image alignment was applied to extract the clinically correct apical views. Initialization of the deformable model was done by selecting three landmarks in three apical views at ED and ES. By disabling temporal forces, pure 3D segmentation was run in these two frames to provide immediate feedback to the user. Editing the detected surfaces was done manually by adding landmarks that pulled the deformable model towards the corresponding position. Temporal forces were then enabled to allow full 4D segmentation, giving volume estimates for each frame within the cardiac cycle. The 4DLVQ tool was integrated as a part of the EchoPAC workstation (GE Vingmed Ultrasound, Horten, Norway). Evaluation was done in volumetric echocardiograms from 35 patients by comparing the LV volumes and EF obtained using 4DLVQ against a commercially available volume measurement tool (TomTec LV-Analysis, TomTec Imaging Systems, Unterschleissheim, Germany). We demonstrated that 4DLVQ produced repeatable estimates of LV volumes and EF, with good agreement to the TomTec tool. However, since 4DLVQ required a less accurate initialization, higher repeatability was demonstrated, with a significantly shorter analysis time. We therefore concluded that 4DLVQ could serve well as a clinical analysis tool.

4.5   Paper E

Real-time active shape models for segmentation of 3D cardiac ultrasound

J. Hansegård, F. Orderud and S. I. Rabben

Published in: 12th International Conference on Computer Analysis of Images and Patterns, (CAIP 2007), Lecture notes in Computer Science

Volume 4673, 2007

The goal of the work leading to this paper was to develop a fully automated method for robust and accurate LV surface detection in 3D echocardiograms. The main contribution of the paper was a 3D extension to the active shape model (ASM) [23, 24] that was coupled with

---

[2]An overview of the segmentation algorithm is given in the paper, but due to the proprietary nature of the algorithm, details have been omitted.

a second order motion model [9] to ensure better temporal behavior than with the standard ASM. Optimization was done using an extended Kalman filter [48] for efficient segmentation in volumetric data. Training of the 3D ASM was done using LV shapes extracted by the semi-automated segmentation tool developed in **Paper D**. These shapes also served as manual reference for evaluation of the proposed method. A transition criterion [83] was used for edge detection. Evaluation in 21 patients demonstrated high robustness, with relatively low bias in EDV, ESV, and EF, and narrow limits of agreements compared to the reference. The method permitted fully automated segmentation, without the need for user interaction, potentially improving the reproducibility compared to semi-automated methods. We did, however, observe that the model had a tendency to get confused by papillary muscles. A nice property of this approach was its efficiency, allowing real-time segmentation, without problems related to foreshortening. In addition, the robustness was excellent, giving successful segmentation, with reasonably good accuracy, in all test data sets.

Chapter 5

Discussion

Paper A was written before 3D ultrasound was released as a commercial product, and it is therefore focused on 2D segmentation. Atlas based techniques had become popular for many segmentation tasks within medical imaging [81, 62], providing a simple way of incorporating expert knowledge into segmentation algorithms. However, during the work on paper A, it became clear that our hybrid multiple atlas and snake algorithm lacked ability to generalize between cases, and a good match in edge signatures did not correlate well with shape similarity, and the minimum total snake energy was not suitable for selecting the best atlas. In addition, matching the patient image against a huge set of atlases would become highly computationally expensive. Since this was purely a 2D method, foreshortening would inevitably prohibit accurate volume measurements, and temporal information was not utilized to improve the segmentation. Further work should therefore be based upon a technique that generalized better between different database cases, and which provided a more robust optimization scheme.

When beginning the work of **paper B**, 3D echocardiography had recently been commercialized, making real-time triplane imaging available, where foreshortening was expected to constitute less problems. This required an algorithm for coupled view segmentation. Based on previous experiences using multiple atlases, I realized that shape and image information had to be condensed into one model, which allowed for generalization between example cases. Active appearance models [19] had been shown to give robust segmentation in sequences of 2D ultrasound images [12], and multiple view segmentation was demonstrated in CT cardiac imagery [58, 75, 76]. The AAM provided an elegant way combining information from multiple cases, thereby generalizing better than individual atlases. Paper B presented a multi view and multi frame active appearance model that modeled multiple views and frames of a cardiac cycle simultaneously. This approach demonstrated good segmentation results, but it did have some problems. Unfortunately, foreshortening was a problem also in triplane imagery, and expert physicians used tracing conventions that compensated for this issue. Compared to previous work [12], adding multiple views caused "under-training" of the model, and it tended to fail due to imaging artifacts and the lack of clear edge evidence in many of the ultrasound recordings. Additionally, the local accuracy in good data was in some cases lower than what could be obtained with traditional deformable model techniques.

It was clear that manual interaction would be needed for a clinically usable method. Cootes had previously presented an algorithm for constraining the AAM to known landmarks [21]. This algorithm was well suited to manually control the AAM, but it also gave an opportunity to incorporate other local segmentation algorithms into the AAM framework. **Paper C**, presented a multi-view and multi-frame AAM for triplane echocardiograms, where

the AAM could be constrained to manually defined landmarks, and also landmarks determined by a dynamic programming based snake algorithm. As a result, segmentation accuracy and robustness was significantly improved compared to the method presented in paper B. One of the challenges with our triplane segmentation algorithm was that the pose in each view was treated independently, with only a weak coupling between the views. We speculated if a 3D shape model would improve the results in cases with severe foreshortening.

By the time paper C was finished, real-time 3D echocardiography had developed into a clinical imaging modality, but clinical tools for volume measurements in 3D ultrasound data were lacking. It was therefore decided to extend my work on statistical models to handle volumetric data. In order to construct such models, a relatively large number of pre-segmented volumetric data sets were needed. Doing this manually would be a highly time-consuming task, and a semi-automated segmentation tool was therefore required. At the same time, GE Vingmed Ultrasound decided to develop their own tool for volume quantification of the left ventricle and I got the offer to participate in development of the segmentation algorithm. This way, I got first hand access to a clinical segmentation tool for building statistical models for volumetric data. Previous experience had shown that manual interaction to initialize and correct the segmentation was critical in a clinical tool. We therefore developed an interactive 4D LV quantification tool (4DLVQ) capable of doing segmentation in sequences of 3D volumes. A validation study of this tool was performed in **paper D**.

The next step was now to investigate the potential of statistical models for segmentation of volumetric data. Active appearance models did indeed provide big advantages compared to our initial atlas-based approach. Unfortunately, the texture models did not always give sufficient degrees of freedom to match the high intensity variations found between ultrasound images. Also, the optimization scheme was not robust enough for fully automated segmentation in all cases. The computational performance of AAMs was reasonable for triplane segmentation, but 3D AAMs were expected to be too computationally expensive for an efficient clinical tool. Orderud had successfully demonstrated that segmentation of 3D echocardiography could be performed in real-time using a deformable spline model in a Kalman filter-based framework [78]. This spline model did, however, not restrict deformations to physiologically realistic shapes. Looking into Orderud's work, I realized that the spline model could easily be exchanged by a different type of surface representation, namely the active shape model (ASM) [23]. Using training shapes extracted using the 4DLVQ tool, a 3D ASM of the left ventricle was constructed and fed into the Kalman filter framework. The results of this work were presented in **paper E**, where robust and fully automated segmentation was achieved in volumetric data. Compared to the work of Orderud et al. [78], the 3D ASM improved LV volume and EF estimates, and produced more physiologically realistic surfaces.

5.1   Relations to algorithmic design requirements

In the introduction of chapter 3, five design requirements for a clinical segmentation algorithm were defined. In summary, the algorithm should: 1) produce accurate contours, follow weak edges, and give accurate volume measurements, 2) comply with clinical conventions, produce physiologically realistic contours, and allow for interaction. Further it should 3) be efficient and fast and easy to use, as well as 4) repeatable and 5) robust in poor image quality data. In this section, the proposed methods are discussed in relation to these requirements.

Paper A. The main strengths of the hybrid atlas and snake algorithm presented in paper A were twofold. First, it produced shapes that were highly constrained to actual heart shapes, making the results *physiologically realistic*. Second, since the database edge signatures were derived from manually traced contours, the method was not restricted to detecting the strongest edges in the images directly, but was rather matching database signatures against the target image. This technique was developed in an effort to make the model *follow weak edges* along the endocardial wall instead of the often more prominent epicardium. *Clinical conventions* and information about papillary muscles were also handled by this technique. Potentially, these factors could provide accurate and realistic shapes that the user would agree upon. *User interaction* was permitted during initialization of the model, and due to the dynamic programming based snake framework, it would be possible to incorporate manual correction of the detected contour into the algorithm. However, a database that included a sufficiently large number of pathologies and imaging situations would make execution time too long for *efficient* clinical use. Even if the *repeatability* of this technique was not studied, small changes in initialization between repetitions can cause different selection of database cases, potentially introducing relatively large differences in repeated measurements. The method did also have shortcomings related to *robustness* towards challenges in ultrasound images, such as reverberations and regional intensity variations. To ensure that all possible imaging situations were covered, the database would become impractically large.

Paper B. The multiview AAM introduced in paper B produced *physiologically realistic* shapes since the contours were computed from a shape space trained on manually traced ventricles. The texture matching technique also provided a framework that could *follow the weak edges* of the endocardial border, and for incorporating *clinical and institutional conventions*, since for instance papillary muscles were integrated into the texture model. The local *accuracy* of the detected contours was unfortunately low even in images with strong edge evidence. This might have been related to "under-training" of the model, but our impression was that the standard AAM convergence scheme was not sufficiently *robust* to handle the different image artifacts, poor contrast, and low signal to noise ratio sufficiently well. Regionally varying signal from different locations in the image tended to affect the algorithm more than the texture difference between the cavity and the myocardium. The pose transform estimates were particularly impaired by poor convergence, and the model had therefore a tendency to end up in local minima. Various alternative convergence schemes were evaluated [20, 27], but none were found to be noticeable better than the standard AAM optimization algorithm. The method was relatively *robust* towards dropouts, since multiple views were used to guide the model. Dropouts in one image plane could therefore be corrected by utilizing image information from other views, both because a weak coupling of the pose transform in each view was employed, and because of the strong shape prior inherent with such models. We did not study the *repeatability* of this technique, but poor convergence can potentially cause relatively large variations in detected volumes and EF between repeated measurements, as small changes to the initialization of the model can result in contours from different local minima. We speculated if a multi scale technique would reduce the dependency on manual initialization to avoid local minima, thereby improving *repeatability*, convergence radius, and *robustness* of this algorithm.

Paper C. Constraining the AAM to manually defined landmarks, allowing for *interactive* segmentation, solved many of the challenges with the standard AAM presented in Paper B. By manually restricting some points of the model, the problems related to the convergence of the pose transform could almost completely be eliminated, producing more *accurate* and *repeatable* results. The effects of "under-training" were less prominent since a larger range of

the model parameters could effectively be utilized. Further development of this technique would be to allow for manually editing the detected contour by adding more manually defined landmarks, but care must then be taken to avoid over-constraining the model. Snakes based upon dynamic programming have several advantages such as high performance while obtaining globally optimal solutions given a set of shape constraints. This makes them relatively *robust* and potentially also *repeatable*. Modeling of local shape constraints such as smoothness is relatively simple, but it can be difficult to restrict the global shape of the snake to a subspace of physiologically realistic shapes. Several limitations of the snake and the AAM algorithms were overcome by combining these techniques as presented in paper C. This hybrid approach preserved the AAM's ability to produce *physiologically realistic* shapes, and the snake algorithm improved the *local accuracy* of the detected contour by driving the AAM towards strong edges. Of particular interest was the framework's ability to dynamically tune the relative contribution of the edges detected by the snake in regions with poor edge evidence. As a result, the proposed method produced *accurate* contours, with derived volumes and EF that agreed well with manual tracing. Both the AAM and the snake have favorable computational complexity, and this technique is expected to provide sufficient performance for *efficient* clinical use.

Paper D. The 4DLVQ tool presented in paper D proved to be well suited as a clinical tool. By proper tuning of the internal shape constraints of the model, we obtained a model that produced *physiologically realistic* shapes, even if no explicit shape information was encoded in the model, such as with the AAM. By analyzing the statistics of the pixel intensity in regions around the deformable model, we were able to create an edge detector that preferred the endocardial border to the often more prominent epicardial border. The model therefore *followed weak edges* well, while retaining a nice shape in locations with dropout. Utilization of 3D data was a key factor to improve segmentation results, since missing information in parts of the volume could be compensated for by utilizing edge evidence in nearby locations, and foreshortening could be avoided completely by manually aligning the views before initialization of the model. Volumes and EF derived from the detected surfaces agreed well with a different commercially available segmentation tool, and 4DLVQ proved an *efficient* clinical tool. Evaluation of *repeatability* demonstrated that the technique produced consistent results for repeated measurements even with manual initialization and editing. *Robustness* towards image artifacts was high, but manual correction was often needed to compensate for papillary muscles.

Paper E. The 3D ASM segmentation algorithm presented in paper E did, similarly to the AAM, utilize a trained PCA-model to restrict its shapes to a space of *physiologically realistic* surfaces. Importantly, volumes and EF estimated by the 3D ASM agreed well with 4DLVQ, and the *regional accuracy* of the method was good, but due to the strong shape prior, the *local accuracy* seemed lower than for the 4DLVQ tool. The presented technique did not incorporate statistical knowledge of image texture, but was using a simple transition edge detector. It was therefore occasionally confused by papillary muscles and large dropouts. This can turn out as a challenge for its use as a routine volume measurement tool in the clinic, since there is currently no way of correcting the result in cases where the model fails. The *robustness* of the 3D ASM was however excellent, and since it is a fully automated tool the *reproducibility* and *ease of use* is expected to be high compared to manual or semi-automated techniques. The main advantage of this approach was its unsurpassed *efficiency*, providing real-time measurements of LV volumes and EF.

5.2   Achievement of goals

The main goal of this thesis was to develop automated methods that utilize expert knowledge for improving the quality of cardiac measurements in echocardiography. The goal did not put any restrictions on the type of measurements, but I decided to focus on automated assessment of LV volumes and EF, since they are among the most important parameters for prognosis and diagnosis of patients with cardiac disease. Experience had shown that automated volume measurement methods often failed in difficult ultrasound data, while an expert could still perform reasonably good measurements. It was therefore natural to try to incorporate some of the expert's knowledge into the automated methods.

One might argue that virtually all segmentation algorithms do to some extent incorporate expert knowledge. In the basic snake algorithm, or in level-set based methods, this knowledge is formulated as smoothness constraints of the evolving deformable model, and knowledge of edge features are modeled in the edge detector. Thresholding techniques or region-based segmentation techniques involve incorporation of some kind of knowledge about intensity distributions. However, to incorporate *explicit* clinical knowledge into an automated segmentation algorithm has not received too much research interest. This is perhaps due to a desire to find a general-purpose segmentation algorithm that can be used across modalities, and for any type of images. Atlas based techniques, on the other hand, use a strategy where clinical knowledge is used directly, by matching validated images onto new patient images, assuming that the clinical decisions made in the matched atlas are still valid for the new patient. Due to the large amount of variation in both shape and intensity distribution between patients, multiple atlases were used, assuming that at least one of the atlases should match the patient image well. Active appearance models and active shape models can be thought of as generalized atlases. They are trained on validated cases, but information from multiple cases is condensed into a single model. In my experience, this is a far better method for incorporating expert knowledge into a segmentation algorithm than pure atlas-based methods, because AAMs and ASMs have the ability to generalize between cases. I believe that statistical models can improve the accuracy and reproducibility of cardiac measurements, and the results of paper C are highly promising. Still, further evaluation of the techniques is required.

During the work on this thesis, it was clear that encoding expert knowledge into the segmentation algorithms could improve volume and EF estimates. However, in images with poor quality the automated algorithms could still fail. In these cases, it should be possible to utilize the physician's knowledge directly by allowing for manual input to correct the automatically detected contour or surface. Constraining three or more nodes on a LV model vastly reduces the degrees of freedom of the segmentation problem, giving significantly improved results. In accurate and efficient clinical volume measurement tools, manual interaction is therefore a necessary part of the workflow.

3D echocardiography is gaining popularity as a clinical tool for cardiac diagnosis, and has the potential of significantly improving clinical decision-making [44]. However, manual tracing in volumetric data is highly impractical in clinical practice. Few commercially available volume measurement tools exist, and there is a clinical need for simpler and more effective tools to fully utilize this potential. It was therefore important to develop automated volume measurement algorithms for volumetric ultrasound data. This thesis presents contributions to the field of active appearance models that potentially can make this technique suitable for semi-automated analysis of triplane ultrasound data in the clinic, improving repeatability and efficiency of echocardiographic quantification. The 4DLVQ tool proved to be a rapid and accurate method for volume quantification in 3D echocardiography.

Its efficient workflow for both initialization and editing of the detected surface resulted in short analysis times, potentially making 4DLVQ suitable for on-line use on the ultrasound scanner during cardiac examinations. 3D ASMs have the potential of instantly displaying LV volume and EF measurements on the screen, even before the physician has had the time to do eyeballing. This fully automated method still require improvements to replace semi-automated methods for clinical routine examinations, but with improvements to for instance the edge detector, this approach will have many interesting clinical applications. The technique can be used for simplifying acquisition by providing automatically aligned views in 3D or triplane echocardiography, automated scanner setup, segmentation optimized rendering, automated ROI extraction for colorflow processing, automated extraction of specific cardiac views, or renderings. The method can also be used for initialization of other segmentation algorithms.

In conclusion, I expect that the results in this thesis will contribute to improved efficiency and quality of routine cardiac examinations, giving better evaluation of cardiac function at a lower cost.

Chapter 6

Conclusion

This thesis has focused on how expert knowledge and manual interaction can be incorporated into automated methods to improve evaluation of cardiac function in echocardiography.

Several extensions to existing segmentation algorithms for automated assessment of LV volumes and EF have been presented, including atlas-based methods, statistical models, and classical deformable models. A hybrid snake and atlas based segmentation algorithm was used for segmentation of 2D echocardiograms. In this algorithm, expert knowledge was incorporated by controlling the snake using shapes and textures from similar examples in a validated case database. More promising results were achieved in triplane echocardiography using a multiview and multi-frame extension to the active appearance model framework. This approach utilized a statistical model that was trained on manually traced contours. Further improvements were achieved by constraining the model to manually defined landmarks, and by integrating a snake algorithm into the AAM framework. For 3D echocardiograms, a clinical semi-automated volume measurement tool has been validated. In this tool, expert knowledge was incorporated by allowing the user to manually edit the detected surface. Finally, fully automated real-time segmentation in 3D echocardiography was presented using a 3D active shape model of the left ventricle in a Kalman filter framework.

The most promising results were obtained by automated segmentation in 3D data, as the increased amount of information in these images, compared to conventional 2D echocardiography, can be used to compensate for several of the typical challenges of echocardiographic data. This modality is believed to dominate development of new segmentation strategies because 3D echocardiography provides more accurate volume measurements than 2D techniques, and because manual quantification in this modality becomes impractical.

Automated segmentation algorithms for volume measurements in echocardiography must produce contours that agree well with the true endocardial border, but perhaps more important is that the method produces results that the physician trust. As a rule of thumb, a clinical tool for reliable volume measurements has to allow for manual interaction to correct the segmentation. In some situations, such as intra-operative monitoring, this is not feasible, and robust fully automated methods should therefore be pursued.

Different methods for utilizing expert knowledge in automated segmentation algorithms for echocardiography have been developed and evaluated in this thesis. Particularly in 3D echocardiography, these contributions are expected to improve efficiency and quality of cardiac measurements.

6.1   Future work

The two most promising contributions of this thesis are the 4DLVQ tool and the 3D active shape model.   Both techniques were used to measure LV volumes and EF in 3D echocardiograms.   We have performed preliminary studies using the 3D ASM for segmentation in triplane data with encouraging results. By using a true 3D shape model for segmentation of triplane echocardiograms, challenges related to foreshortening can be reduced.

Our main focus has been on automated assessment of left ventricular volume, but automated quantification of the right ventricle and the atria are currently becoming active research areas. Further topics will also involve simultaneous multi chamber segmentation, improved anatomical models including for example the aortic outlet tract, and more accurate models that include valve motion.   Our framework for controlling the active appearance models with external constraints could provide a useful tool for combining traditional image segmentation with other external constraints such as velocity fields extracted either by Doppler techniques or speckle tracking.

Further development in scanner front end processing capabilities and probe design will improve image quality, frame-rate, and field of view of volumetric echocardiography. These factors will in turn improve the accuracy of automated segmentation algorithms, and allow for new clinical applications. Simultaneous segmentation of the endocardial and epicardial border for LV mass computation is an old research area that may play a more important role as image quality improves. Stress echocardiography is an area that will benefit from improved frame rate, where real-time segmentation techniques, for instance using statistical models, may play an important role. Improved image quality will also make automated quantification methods for 2D imagery more feasible.

Fully automated segmentation algorithms open up for several new interesting clinical applications. Such techniques allow for intra-operative monitoring, automated acquisition setup, real-time feedback during cardiac interventions, and new visualization techniques. Fully automated algorithms can also be used for training of clinical personnel by providing automated annotation, or by advising on scanning techniques.

Another interesting field is automated diagnosis from automatically interpreted cardiac data. Several attempts have been made to classify pathologies based upon ventricular shape parameters [11, 59, 95]. Further work should also focus on automated diagnostic support to simplify the clinical decision process, by presenting examples of similar pathologies fetched from an annotated case database to the user during scanning.

The field of medical imaging is moving forward rapidly. Echocardiography has had the advantage of safe and rapid acquisition of high resolution 3D images, and since it is a cheap and portable technique, it has been able to compete with other modalities such as magnetic resonance imaging and computed tomography. However, these modalities are continuously developing, and we must endeavor to improve 3D echocardiography as a clinical tool, for instance by providing new and better automated quantification tools. 3D imaging is currently reserved for high-end ultrasound scanners, but the trend goes towards smaller and more compact scanners for widespread use also in primary care. Soon, portable, or even hand-held scanners may come with 3D imaging capabilities, and more, less experienced operators will use ultrasound as a diagnostic tool.   Simplifying 3D echocardiography is therefore an important challenge. The work in this thesis is a step in this direction, making 3D echocardiography a more useful clinical tool.

References

[1] A. J. Buda, E. J. Delp, C. R. Meyer, J. M. Jenkins, D. N. Smith, F. L. Bookstein, and B. Pitt, "Automatic computer processing of digital 2-dimensional echocardiograms," *Am J Cardiol*, vol. 52, no. 3, pp. 384–389, Aug. 1983.

[2] A. A. Amini, T. E. Weymouth, and R. C. Jain, "Using dynamic programming for solving variational problems in vision," *IEEE Trans Pattern Anal Machine Intell*, vol. 12, no. 9, pp. 855–867, Sep. 1990.

[3] E. D. Angelini, S. Homma, G. Pearson, J. W. Holmes, and A. F. Laine, "Segmentation of real-time three-dimensional ultrasound for quantification of ventricular function: A clinical study on right and left ventricles," *Ultrasound Med Biol*, vol. 31, no. 9, pp. 1143–1158, Sep. 2005.

[4] E. D. Angelini, A. F. Laine, S. Takuma, J. W. Holmes, and S. Homma, "LV volume quantification via spatiotemporal analysis of real-time 3-D echocardiography," *IEEE Trans Med Imaging*, vol. 20, pp. 457–469, 2001.

[5] B. A. J. Angelsen, *Waves, signals and signal processing in medical ultrasonics*. Dept. of Physiology and Biomedical Engineering, Norwegian University of Science and Technology, Trondheim, Norway, 1995, vol. I.

[6] L. P. Badano, E. Dall'Armellina, M. J. Monaghan, M. Pepi, M. Baldassi, M. Cinello, and P. M. Fioretti, "Real-time three-dimensional echocardiography: Technological gadget or clinical tool?" *J Cardiovasc Med*, vol. 8, no. 3, pp. 144–162, Mar 2007.

[7] R. Bellman, *Dynamic Programming*, N. J. Princeton, Ed. Princeton University Press, 1957.

[8] A. Blake and M. Isard, *Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1998.

[9] A. Blake, M. Isard, and D. Reynard, "Learning to track the visual motion of contours," *Artif Intell*, vol. 78, no. 1-2, pp. 179–212, 1995.

[10] N. Bom, C. T. Lancee, J. Honkoop, and P. G. Hugenholtz, "Ultrasonic viewer for cross-sectional analyses of moving cardiac structures," *Biomed Eng*, vol. 6, no. 11, pp. 500–503, 1971.

[11] J. Bosch, F. Nijland, S. Mitchell, B. Lelieveldt, O. Kamp, J. Reiber, and M. Sonka, "Computer-aided diagnosis via model-based shape analysis: Automated classification of wall motion abnormalities in echocardiograms," *Acad Radiol*, vol. 12, no. 3, pp. 358–367, 2005.

[12] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance motion models," *IEEE Trans Med Imaging*, vol. 21, no. 11, pp. 1374–1383, Nov. 2002.

[13] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Fully automated endocardial contour detection in time sequences of echocardiograms by active appearance motion models," in *Comp. Cardiol.*, vol. 28, 2001, pp. 93–96.

[14] A. Bosnjak, A. Bosnjak, V. Torrealba, M. Acuna, M. Bosnjak, B. Solaiman, G. Montilla, and C. Roux, "Segmentation and VRML visualization of left ventricle in echocardiographic images using 3D deformable models and superquadrics," in *Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE*, V. Torrealba, Ed., vol. 3, 2000, pp. 1724–1727 vol.3.

[15] S. Brekke, "Techniques for enhancement of temporal resolution in three-dimensional echocardiography," Ph.D. dissertation, Norwegian University of Science and Technology, 2007.

[16] C. Burckhardt, "Speckle in ultrasound B-mode scans," *IEEE Trans Sonics Ultrason*, vol. 25, no. 1, pp. 1–6, 1978.

[17] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim, "Multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Trans Med Imaging*, vol. 15, pp. 290–298, 1996.

[18] S. M. Collins, D. J. Skorton, E. A. Geiser, J. A. Nichols, D. A. Conetta, N. G. Pandian, and R. E. Kerber, "Computer-assisted edge detection in two-dimensional echocardiography: Comparison with anatomic data," *Am J Cardiol*, vol. 53, no. 9, pp. 1380–1387, May 1984.

[19] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans Pattern Anal Machine Intell*, vol. 23, no. 6, pp. 681–685, June 2001.

[20] T. F. Cootes and C. J. Taylor, "An algorithm for tuning an active appearance model to new data," in *In British Machine Vision Conference*, 2006, pp. 919–928.

[21] T. F. Cootes and C. J. Taylor, "Constrained active appearance models," in *Proc. IEEE Int. Conf. Computer Vision*, vol. 1, July 2001, pp. 748–754.

[22] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Proc. SPIE Med. Imag.*, vol. 4322, 2001, pp. 236–248.

[23] T. F. Cootes and C. J. Taylor, "Active shape models - 'Smart snakes'," in *Proc. British Machine Vision Conf.*   Springer-Verlag, 1992, pp. 266–275.

[24] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - Their training and application," *Comput Vis Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[25] C. Corsi, G. Saracino, A. Sarti, and C. Lamberti, "Left ventricular volume estimation for real-time three-dimensional echocardiography," *IEEE Trans Med Imaging*, vol. 21, pp. 1202–1208, 2002.

[26] D. L. Dekker, R. L. Piziali, and E. Dong, "A system for ultrasonically imaging the human heart in three dimensions," *Comput Biomed Res*, vol. 7, no. 6, pp. 544–553, Dec. 1974.

[27] R. Donner, M. Reiter, G. Langs, P. Peloschek, and H. Bischof, "Fast active appearance model search using canonical correlation analysis," *IEEE Trans Pattern Anal Machine Intell*, vol. 28, no. 10, pp. 1690–1694, 2006.

[28] K. T. Dussik, "Über die Möglichkeit hochfrequente mechanische Schwingungen als diagnostisches hilfsmittel zu verwerten," *Z Neurol Psychiat*, pp. 174–153, 1942.

[29] K. T. Dussik, F. Dussik, and I. Wyt, "Auf dem wege zur Hyperphonographie des Gehirnes," *Wien Med Wochenschr*, vol. 97, pp. 425–529, 1947.

[30] I. Edler and C. H. Hertz, "The use of ultrasonic reflectoscope for the continuous recording of the movements of heart walls." *Kung Fysiogr Sällsk Lund F*, vol. 24, no. 3, pp. 1–19, May 1954.

[31] I. Edler and K. Lindström, "The history of echocardiography," *Ultrasound Med Biol*, vol. 30, no. 12, pp. 1565–1644, Dec. 2004.

[32] H. Feigenbaum, "Evolution of echocardiography," *Circulation*, vol. 93, no. 7, pp. 1321–1327, Apr. 1996.

[33] M. Figueiredo, M. Figueiredo, J. Leitao, and A. Jain, "Unsupervised contour representation and estimation using B-splines and a minimum description length criterion," *IEEE Trans Image Process*, vol. 9, no. 6, pp. 1075–1087, 2000.

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

[35] N. Friedland and D. Adam, "Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing," *IEEE Trans Med Imaging*, vol. 8, pp. 344–353, 1989.

[36] E. A. Geiser, M. Ariet, D. A. Conetta, S. M. Lupkiewicz, L. G. Christie, and C. R. Conti, "Dynamic three-dimensional echocardiographic reconstruction of the intact human left ventricle: Technique and initial observations in patients," *Am Heart J*, vol. 103, no. 6, pp. 1056–1065, Jun. 1982.

[37] C. Goodall, "Procrustes methods in the statistical analysis of shape," *J R Stat Soc*, vol. 53, no. 2, pp. 285–339, 1991.

[38] O. Gérard, A. C. Billon, J. Rouet, M. Jacob, M. Fradkin, and C. Allouche, "Efficient model-based quantification of left ventricular function in 3-D echocardiography," *IEEE Trans Med Imaging*, vol. 21, no. 21, pp. 1059–1068, Sep. 2002.

[39] A. Hammoude, "Endocardial border identification in two-dimensional echocardiographic images: Review of methods," *Comput Med Imaging Graph*, vol. 22, pp. 181–193, 1998.

[40] J. Hansegård, F. Orderud, and S. Rabben, "Real-time active shape models for segmentation of 3D cardiac ultrasound," in *12th International Conference on Computer Analysis of Images and Patterns, CAIP 2007*, ser. Lect Notes Comput Sci, vol. 4673 LNCS, Vienna, 2007, pp. 157–164.

[41] J. Hansegård, E. Steen, S. Rabben, A. H. Torp, H. Torp, S. Frigstad, and B. Olstad, "Knowledge based extraction of the left ventricular endocardial boundary from 2D echocardiograms," in *2004 IEEE Ultrasonics Symposium*, Y. M.P., Ed., vol. 3, Montreal, Que., 2004, pp. 2121–2124.

[42] J. Hansegård, S. Urheim, K. Lunde, and S. Rabben, "Constrained active appearance models for segmentation of triplane echocardiograms," *IEEE Trans Med Imaging*, vol. 26, no. 10, pp. 1391–1400, 2007.

[43] J. Hansegård, S. Urheim, E. Steen, H. Torp, B. Olstad, S. Malm, and S. Rabben, "Detection of the myocardial boundary in the left ventricle from simultaneously acquired triplane ultrasound images using multi view active appearance motion models," in *2005 IEEE Ultrasonics Symposium*, vol. 4, Rotterdam, 2005, pp. 2267–2270.

[44] J. L. Hare, C. Jenkins, S. Nakatani, A. Ogawa, C. Yu, and T. H. Marwick, "Feasibility and clinical decision-making with 3D echocardiography in routine practice," *Heart*, vol. 94, no. 4, pp. 440–445, 2008.

[45] G. S. Hillis and P. Bloomfield, "Basic transthoracic echocardiography," *Br Med J*, vol. 330, no. 7505, pp. 1432–1436, Jun. 2005.

[46] J. Hung, R. Lang, F. Flachskampf, S. K. Shernan, M. L. McCulloch, D. B. Adams, J. Thomas, M. Vannan, and T. Ryan, "3D echocardiography: A review of the current status and future directions," *J Am Soc Echocardiogr*, vol. 20, no. 3, pp. 213–233, Mar. 2007.

[47] G. Jacob, J. A. Noble, C. Behrenbruch, A. D. Kelion, and A. P. Banning, "A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography," *IEEE Trans Med Imaging*, vol. 21, no. 3, pp. 226–238, 2002.

[48] G. Jacob, J. A. Noble, M. Mulet-Parada, and A. Blake, "Evaluating a robust contour tracker on echocardiographic sequences," *Med Image Anal*, vol. 3, no. 1, pp. 63–75, Mar. 1999.

[49] L. D. Jacobs, I. S. Salgo, S. Goonewardena, L. Weinert, P. Coon, D. Bardo, O. Gerard, P. Allain, J. L. Zamorano, L. P. de Isla, V. Mor-Avi, and R. M. Lang, "Rapid online quantification of left ventricular volume from real-time three-dimensional echocardiographic data," *Eur Heart J*, vol. 27, pp. 460–468, Nov. 2006.

[50] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[51] O. Kamp and M. Cramer, "History of echocardiography in the netherlands: 30 years of education and clinical applications," *Neth Heart J*, vol. 16, no. 1, pp. 16–20, 2008.

[52] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int J Comput Vis*, vol. 1, no. 1, pp. 321–331, Jan. 1988.

[53] A. M. Katz, *Physiology of the heart*, 2nd ed.   Raven Press, Ltd., 1992.

[54] J. W. Klingler, C. L. Vaughan, T. D. Fraker, and L. T. Andrews, "Segmentation of echocardiographic images using mathematical morphology," *IEEE Trans Biomed Eng*, vol. 35, no. 11, pp. 925–934, 1988.

[55] D. Kucera and R. W. Martin, "Segmentation of sequences of echocardiographic images using a simplified 3D active contour model with region-based external forces," *Comput Med Imaging Graph*, vol. 21, no. 1, pp. 1–21, 1997.

[56] R. M. Lang, M. Bierig, R. B. Devereux, F. A. Flachskampf, E. Foster, P. A. Pellikka, M. H. Picard, M. J. Roman, J. Seward, J. S. Shanewise, S. D. Solomon, K. T. Spencer, M. S. J. Sutton, and W. J. Stewart, "Recommendations for chamber quantification." *J Am Soc Echocardiogr*, vol. 18, no. 12, pp. 1440–1463, 2005.

[57] F. D. Ledley and J. B. Wilson, "Computer analysis of ultrasoundcardiograms," *Comput Biol Med*, vol. 4, no. 1, pp. 27–28, Jun. 1974.

[58] B. P. F. Lelieveldt, M. Üzümcü, R. J. van der Geest, J. H. C. Reiber, and M. Sonka, "Multi-view active appearance models for consistent segmentation of multiple standard views: Application to long- and short-axis cardiac MR images," vol. 1256, pp. 1141–1146, 2003.

[59] K. Leung and J. Bosch, "Localized shape variations for classifying wall motion in echocardiograms," *Lect Notes Comput Science*, vol. 4791, no. PART 1, pp. 52–59, 2007.

[60] X. Lu, M. Xie, D. Tomberlin, B. Klas, V. Nadvoretskiy, N. Ayres, J. Towbin, and S. Ge, "How accurately, reproducibly, and efficiently can we measure left ventricular indices using M-mode, 2-dimensional, and 3-dimensional echocardiography in children?" *Am Heart J*, vol. 155, no. 5, pp. 946–953, May 2008.

[61] G. E. Mailloux, G. E. Mailloux, A. Bleau, M. Bertrand, and R. Petitclerc, "Computer analysis of heart motion from two-dimensional echocardiograms," *IEEE Trans Biomed Eng*, vol. BME-34, no. 5, pp. 356–364, 1987.

[62] J. B. Maintz and M. A. Viergever, "A survey of medical image registration." *Med Image Anal*, vol. 2, no. 1, pp. 1–36, Mar 1998.

[63] S. Malm, S. Frigstad, E. Sagberg, P. A. Steen, and T. Skjarpe, "Real-time simultaneous triplane contrast echocardiography gives rapid, accurate and reproducible assessment of left ventricular volumes and ejection fraction: A comparison with magnetic resonance imaging," *J Am Soc Echocardiogr*, vol. 19, no. 12, pp. 1494–1501, Dec. 2006.

[64] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Med Image Anal*, vol. 1, no. 2, pp. 91–108, Jun. 1996.

[65] M. Mignotte and J. Meunier, "A multiscale optimization approach for the dynamic contour-based boundary detection issue." *Comput Med Imaging Graph*, vol. 25, no. 3, pp. 265–275, 2001.

[66] M. Mignotte, J. Meunier, and J. Tardif, "Endocardial boundary estimation and tracking in echocardiographic images using deformable templates and Markov random fields," *Pattern Anal Appl*, vol. 4, pp. 256–271, 2001.

[67] I. Mikic, S. Krucinski, and J. D. Thomas, "Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates," *IEEE Trans Med Imaging*, vol. 17, pp. 274–284, 1998.

[68] A. Mishra, P. K. Dutta, and M. K. Ghosh, "A GA based approach for boundary detection of left ventricle with echocardiographic image sequences," *Image and Vis Comput*, vol. 21, pp. 967–976, 2003.

[69] S. C. Mitchell, J. G. Bosch, B. P. F. Lelieveldt, R. J. van der Geest, J. H. C. Reiber, and M. Sonka, "3-D active appearance models: Segmentation of cardiac MR and ultrasound images," *IEEE Trans Med Imaging*, vol. 21, no. 9, pp. 1167–1178, Sep. 2002.

[70] J. Montagnat and H. Delingette, "4D deformable models with temporal constraints: Application to 4D cardiac image segmentation," *Med Image Anal*, vol. 9, pp. 87–100, 2005.

[71] J. Montagnat, H. Delingette, and N. Ayache, "A review of deformable surfaces: Topology, geometry and deformation," *Image Vis Comput*, vol. 19, no. 14, pp. 1023–1040, 2001.

[72] V. Mor-Avi, L. Sugeng, L. Weinert, P. MacEneaney, E. G. Caiani, R. Koch, I. S. Salgo, and R. M. Lang, "Fast measurement of left ventricular mass with real-time three-dimensional echocardiography: Comparison with magnetic resonance imaging," *Circulation*, vol. 110, no. 13, pp. 1814–1818, Sep. 2004.

[73] W. E. Moritz, A. S. Pearlman, D. H. Mc Cabe, D. K. Medema, M. E. Ainsworth, and M. S. Boles, "An ultrasonic technique for imaging the ventricle in three dimensions and calculating its volume," *IEEE Trans Biomed Eng*, vol. BME-30, no. 8, pp. 482–492, 1983.

[74] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans Med Imaging*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[75] C. R. Oost, B. P. F. Lelieveldt, M. Üzümcü, H. J. Lamb, J. H. C. Reiber, and M. Sonka, "Multi-view active appearance models: Application to X-ray LV angiography and cardiac MRI," in *Proc. Conf. Information Processing in Medical Imaging*, July 2003, pp. 234–245.

[76] E. Oost, G. Koning, M. Sonka, P. V. Oemrawsingh, J. H. C. Reiber, and B. P. F. Lelieveldt, "Automated contour detection in X-ray left ventricular angiograms using multiview active appearance models and dynamic programming," *IEEE Trans Med Imaging*, vol. 25, no. 9, pp. 1158–1171, Sep. 2006.

[77] F. Orderud, "A framework for real-time left ventricular tracking in 3D+T echocardiography, using nonlinear deformable contours and Kalman filter based tracking," in *Comput Cardiol*, 2006.

[78] F. Orderud, J. Hansgård, and S. Rabben, "Real-time tracking of the left ventricle in 3D echocardiography using a state estimation approach," in *10th International Conference on Medical Imaging and Computer-Assisted Intervention, MICCAI 2007*, ser. Lect Notes Comput Sci, N. Ayache, S. Ourselin, and A. Maeder, Eds., vol. 4791, no. Part 1. Brisbane: Springer-Verlag, 2007, pp. 858–865.

[79] N. Paragios, "A level set approach for shape-driven segmentation and tracking of the left ventricle," *IEEE Trans Med Imaging*, vol. 22, no. 6, pp. 773–776, Jun. 2003.

[80] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philos Mag*, vol. 2, no. 2, pp. 559–572, 1901.

[81] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation." *Annu Rev Biomed Eng*, vol. 2, pp. 315–337, 2000.

[82] M. H. Picard, R. L. Popp, and A. E. Weyman, "Assessment of left ventricular function by echocardiography: A technique in evolution," *J Am Soc Echocardiogr*, vol. 21, no. 1, pp. 14–21, Jan. 2008.

[83] S. I. Rabben, A. H. Torp, A. Støylen, S. Slørdahl, K. Bjørnstad, B. O. Haugen, and B. Angelsen, "Semiautomatic contour detection in ultrasound M-mode images," *Ultrasound Med Biol*, vol. 26, no. 2, pp. 287–296, 2000.

[84] I. S. Salgo, "Three-dimensional echocardiographic technology," *Cardiol Clin*, vol. 25, no. 2, pp. 231–239, May 2007.

[85] N. Schiller, H. Acquatella, T. Ports, D. Drew, J. Goerke, H. Ringertz, N. Silverman, B. Brundage, E. Botvinick, R. Boswell, E. Carlsson, and W. Parmley, "Left ventricular volume from paired biplane two-dimensional echocardiography," *Circulation*, vol. 60, no. 3, pp. 547–555, Sep. 1979.

[86] N. B. Schiller, P. M. Shah, M. Crawford, A. DeMaria, R. Devereux, H. Feigenbaum, H. Gutgesell, N. Reichek, D. Sahn, and I. Schnittger, "Recommendations for quantitation of the left ventricle by two-dimensional echocardiography." *J Am Soc Echocardiogr*, vol. 2, no. 5, pp. 358–367, Sep. 1989.

[87] S. Setarehdan, S. Setarehdan, and J. Soraghan, "Automatic cardiac LV boundary detection and tracking using hybrid fuzzy temporal and fuzzy multiscale edge detection," *IEEE Trans Biomed Eng*, vol. 46, no. 11, pp. 1364–1378, 1999.

[88] T. Shiota, M. Jones, M. Chikada, C. E. Fleishman, J. B. Castellucci, B. Cotter, A. N. DeMaria, O. T. von Ramm, J. Kisslo, T. Ryan, and D. J. Sahn, "Real-time three-dimensional echocardiography for determining right ventricular stroke volume in an animal model of chronic right ventricular volume overload," *Circulation*, vol. 97, no. 19, pp. 1897–1900, May 1998.

[89] S. Singh and A. Goyal, "The origin of echocardiography: A tribute to Inge Edler." *Tex Heart Inst J*, vol. 34, no. 4, pp. 431–438, 2007.

[90] D. Skorton and S. Collins, "Quantitation in echocardiography," *Cardiovasc Intervent Radiol*, vol. 10, no. 6, pp. 316–331, Nov. 1987.

[91] D. J. Skorton, C. A. McNary, J. S. Child, F. C. Newton, and P. M. Shah, "Digital image processing of two dimensional echocardiograms: Identification of the endocardium," *Am J Cardiol*, vol. 48, no. 3, pp. 479–486, Sep. 1981.

[92] D. J. Skorton, C. A. McNary, J. S. Child, and P. M. Shah, "Computerized image processing in cross-sectional echocardiography," *Am J Cardiol*, vol. 45, no. 2, pp. 403–403, Feb. 1980.

[93] J. E. Snyder, J. Kisslo, and O. von Ramm, "Real-time orthogonal mode scanning of the heart. I. System design." *J Am Coll Cardiol*, vol. 7, no. 6, pp. 1279–1285, Jun 1986.

[94] M. Song, R. M. Haralick, F. H. Sheehan, and R. K. Johnson, "Integrated surface model optimization for freehand three-dimensional echocardiography," *IEEE Trans Med Imaging*, vol. 21, pp. 1077–1090, 2002.

[95] M. Sonka, J. G. Bosch, B. P. F. Lelieveldt, S. C. Mitchell, and J. H. C. Reiber, "Computer-aided diagnosis via model-based shape analysis: Cardiac MR and echo," *International Congress Series*, vol. 1256, pp. 1013–1018, Jun. 2003.

[96] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, Analysis and Machine Vision*, second edition ed., M. Sonka, Ed.   PWS Publishing, 1998.

[97] M. Sonka, B. P. F. Lelieveldt, S. C. Mitchell, J. G. Bosch, R. J. van der Geest, and J. H. C. Reiber, "Active appearance motion model segmentation," *Proc Digit Comput Video*, pp. 64–68, 2001.

[98] M. van Stralen, J. G. Bosch, M. M. Voormolen, G. van Burken, B. J. Krenning, C. T. Lancée, N. de Jong, and J. H. C. Reiber, "A semi-automatic endocardial border detection method for 4D ultrasound data," in *Proc Med Image Comput Computer Assist Intervent (MICCAI 2004)*, ser. Lect Notes Comput Sci, vol. 3216.   Springer Berlin / Heidelberg, 2004, pp. 43–50.

[99] L. Sugeng, V. Mor-Avi, L. Weinert, J. Niel, C. Ebner, R. Steringer-Mascherbauer, F. Schmidt, C. Galuschky, G. Schummers, R. M. Lang, and H. J. Nesser, "Quantitative assessment of left ventricular size and function. Side-by-side comparison of real-time three-dimensional echocardiography and computed tomography width magnetic resonance reference," *Circulation*, vol. 114, pp. 654–661, Aug. 2006.

[100] K. Swedberg, J. Cleland, H. Dargie, H. Drexler, F. Follath, M. Komajda, L. Tavazzi, O. A. Smiseth, A. Gavazzi, A. Haverich, A. Hoes, T. Jaarsma, J. Korewicki, S. Levy, C. Linde, J.-L. Lopez-Sendon, M. S. Nieminen, L. Pierard, and W. J. Remme, "Guidelines for the diagnosis and treatment of chronic heart failure: Executive summary (update 2005): The task force for the diagnosis and treatment of chronic heart failure of the european society of cardiology," *Eur Heart J*, vol. 26, no. 11, pp. 1115–1140, Jun. 2005.

[101] S. Tamura, K. Yata, M. Matsumoto, T. Matsuyama, T. Shimazu, and M. Inoue, "Plan-based boundary extraction and 3-D reconstruction for orthogonal 2-D echocardiography," *Pattern Recogn*, vol. 20, no. 2, pp. 155–162, 1987.

[102] G. Taylor, J. Humphries, E. Mellits, B. Pitt, R. Schulze, L. Griffith, and S. Achuff, "Predictors of clinical course, coronary anatomy and left ventricular function after recovery from acute myocardial infarction," *Circulation*, vol. 62, no. 5, pp. 960–970, Nov. 1980.

[103] D. Terzopoulos, D. Terzopoulos, and D. Metaxas, "Dynamic 3D models with local and global deformations: Deformable superquadrics," *IEEE Trans Pattern Anal Machine Intell*, vol. 13, no. 7, pp. 703–714, 1991.

[104] M. Voormolen, B. Krenning, C. Lancee, F. ten Cate, J. Roelandt, A. van der Steen, and N. de Jong, "Harmonic 3-D echocardiography with a fast-rotating ultrasound transducer," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 53, no. 10, pp. 1739–1748, 2006.

[105] R. C. Waag and R. Gramiak, "Methods for ultrasonic imaging of the heart," *Ultrasound Med Biol*, vol. 2, no. 3, pp. 163–170, Jun. 1976.

[106] R. Wagner, R. Wagner, S. Smith, J. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," *IEEE Trans Sonics Ultrason*, vol. 30, no. 3, pp. 156–163, 1983.

[107] H. White, R. Norris, M. Brown, P. Brandt, R. Whitlock, and C. Wild, "Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction," *Circulation*, vol. 76, no. 1, pp. 44–51, Jul. 1987.

[108] J. J. Wild, *Inventive Minds: Creativity in Technology*. Oxford University Press, 1992, ch. The Origin of Soft-Tissue Ultrasonic Echoing and Early Instrumental Application to Clinical Medicine, pp. 115–141.

[109] M. J. Wood and M. H. Picard, "Utility of echocardiography in the evaluation of individuals with cardiomyopathy," *Heart*, vol. 90, no. 6, pp. 707–712, Jun. 2004.

[110] J. Y. Yan and T. G. Zhuang, "Applying improved fast marching method to endocardial boundary detection in echocardiographic images," *Pattern Recognit Lett*, vol. 24, pp. 2777–2784, 2003.

[111] L. Zhang and E. Geiser, "An effective algorithm for extracting serial endocardial borders from 2-dimensional echocardiograms," *IEEE Trans Biomed Eng*, vol. 31, no. 6, pp. 441–447, 1984.

[112] W. Zwehl, R. Levy, E. Garcia, R. Haendchen, W. Childs, S. Corday, S. Meerbaum, and E. Corday, "Validation of a computerized edge detection algorithm for quantitative two-dimensional echocardiography," *Circulation*, vol. 68, no. 5, pp. 1127–1135, Nov. 1983.

[113] M. Üzümcü, R. J. van der Geest, M. Sonka, H. J. Lamb, J. H. C. Reiber, and B. P. F. Lelieveldt, "Multiview active appearance models for simultaneous segmentation of cardiac 2- and 4-chamber long-axis magnetic resonance images," *Invest Radiol*, vol. 40, pp. 195–203, Apr. 2005.

[114] E. Braunwald, Ed., *Heart disease*, 4th ed. W. B. Saunders company, 1992, vol. 1.

[115] E. Braunwald, Ed., *Heart disease*, 4th ed. W. B. Saunders company, 1992, vol. 2.

[116] C. Lentner and C. Lentner, Eds., *Heart and Circulation*, ser. Geigy Scientific Tables. Basel: CIBA-GEIGY Ltd., 1990, vol. 5.

[117] "Wikipedia." [Online; accessed 02-Jul-2008] http://en.wikipedia.org/wiki/Left_ventricle

Paper A

Knowledge Based Extraction of the Left Ventricular Endocardial Boundary from 2D Echocardiograms

J. Hansegård, E. Steen, S. I. Rabben, A. H. Torp, H. Torp, S. Frigstad, and B. Olstad

Knowledge Based Extraction of the Left Ventricular Endocardial Boundary from 2D Echocardiograms

Jøger Hansegård*, Erik Steen†, Stein Inge Rabben†, Anders H. Torp†, Hans Torp‡, Sigmund Frigstad†, and Bjørn Olstad†*Department of Informatics, University of Oslo, Oslo, Norway†GE Vingmed Ultrasound, Horten, Norway ‡Department of circulation and imaging, Norwegian University of Science and Technology, Trondheim, Norway

Abstract

Extraction of the endocardial boundary of the left ventricle is a key challenge in cardiac ultrasound imaging. The cardiac anatomy may be difficult to determine automatically without incorporating knowledge of both wall shape and intensity signature into the detection algorithm. The aim of this study is to establish a framework for knowledge based extraction of the left ventricular endocardial boundary. The method is based upon the Snake algorithm, where internal and external energy terms are combined into a Snake energy. Instead of using the patient image directly for calculation of the external energy, we propose to use the correlation between geometrically normalized images from the patient and from the database. The ventricular shapes from the database cases are used to compute the internal energy term.

One boundary is detected for each case, hence a selection criterion is required. The total Snake energy is evaluated for this purpose and compared to manual selection of the best case.

As a preliminary verification of the framework, the ventricular end diastolic and end systolic areas and the ventricular ejection fraction where calculated from the detected boundaries for a set of patient cases, using both manual and automatic database case selection. Using manual case selection, the results are encouraging, but the total Snake energy did not provide a sufficiently robust selection criterion.

The strength of the proposed method is its ability to utilize expert knowledge directly for extraction of the endocardial boundary from ultrasound data. Using manual selection of the best case, the calculated parameters from detected boundaries were in good agreement with manual delineation. Further work is required to find a robust selection criterion.

I. Introduction

Automatic delineation of the endocardial boundary is a key challenge in cardiac ultrasound imaging. Several clinically important parameters, such as ventricular ejection fraction (EF), stroke volume and cardiac output can be computed based on this boundary.

For extraction of the endocardial boundary, three common approaches include pixel segmentation techniques, Snakes, and case based methods such as Active Appearance Models.

Image segmentation techniques using for example thresholding are attractive due to their simplicity [1], [2]. These methods usually utilizes a local pixel neighbourhood to classify pixels into the classes 'blood' or 'tissue.' However, due to cardiac ultrasound characterista such as noise, reverberations, and drop outs, methods based entirely on pixel information within local neighbourhoods tend to fail.

Snakes [3], [4] are popular as edge detectors in echocardiography since apriori knowledge about boundary shape and cross boundary intensity signatures is easily incorporated as energy functions. These energy functions include external forces caused by pixel properties, while internal forces constraint the contour shape. Usually the same internal and external
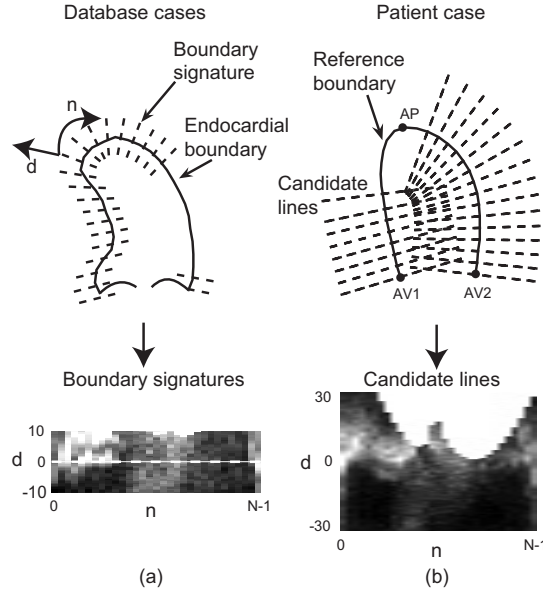
Figure 1.   (a) Database boundary signatures are generated by sampling the pixel values along lines across the manually outlined endocardial boundary. Each line *n* reflects the intensity profile in a neighborhood around the boundary. These signatures are the basis for a geometrically normalized boundary signature image, which is used as template for boundary detection. (b) Since the actual endocardial boundary is not known in the patient case, a reference boundary is established as a spline through three manually outlined landmarks, including the two atrioventricular points (AV1 and AV2) and the apex (AP). The pixel values along candidate lines are sampled normal to the reference boundary, and a candidate image is established. Note that pixels outside the image sector appear in white. These pixels are not used for boundary detection. The distance *d* from the boundary is measured in millimeters.

energy functions are used for the whole image and also for all patient cases. This is a limitation in cardiac boundary detection since the cross boundary signature may vary at different locations in the image, and the ventricular shape may vary significantly from case to case due to individual and pathological variations.

Case based methods include, amongst others, Active Shape Models and Active Appearance Models [5]–[7]. Such models are trained on a set of known cases, which allows the model to adapt to the characteristics of each new case, as long as the model was trained on a similar example.

The aim of this study is to establish an improved Snake algorithm which combines some aspects of the traditional Snake approach and case based methods. This is done by modifying the Snake algorithm to utilize database cases as templates for ventricular cross-boundary pixel signatures and ventricular shape. Since one boundary will be detected for each database case, a selection criterion for the case giving the best boundary detection is required. A sub goal is therefore to evaluate if the total Snake energy can be used as a selection criterion.

II. METHOD

We propose a method where the examples in a case database are used as templates for detection of the left ventricular endocardial boundary.

Instead of using traditional gradient or transition criterions to evaluate the boundary strength in a patient image, we use the pixel signatures in a small neighborhood across the database case's boundary as templates. For each database boundary signature, we have corresponding candidate lines spanning from the cavity into the myocardium of the patient
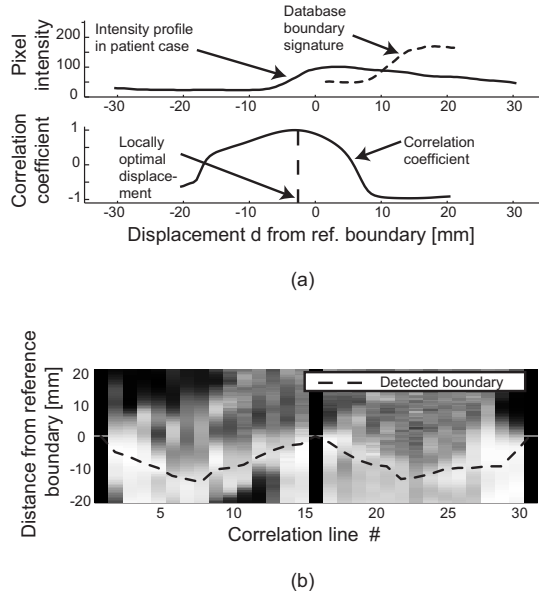
(a)

(b)

Figure 2.   (a) An external energy image is computed by the line wise correlation coefficient when shifting the database boundary signature incrementally along the patient case candidate line. The correlation at each displacement of the boundary signature is recorded for all geometrically corresponding pairs of candidate lines and boundary signatures. (b) The resulting correlation sequences can be interpreted as a correlation image. The optimum horizontal traverse with respect to the combined energy function is included.

case. We search for probable boundary locations along these candidate lines by evaluating the correlation between the database boundary signatures and the candidate line pixel values in the patient cases.

Shape and boundary energy terms are combined using appropriate weight functions, and dynamic programming is used to find the globally optimal boundary location.

Each database case results in one detected boundary. A selection criterion is finally used to pick the database case that provides best boundary detection. The corresponding boundary is used as the detected ventricular boundary.

A. Preprocessing

Spatiotemporal filtering of the tissue data is applied to reduce the amount of speckle and noise in the images. This is implemented as an edge preserving median filter which operates on time sequences of tissue data.

B. Geometrical normalization of database and patient cases

The endocardial boundary is manually outlined in all the database cases. Cross boundary signatures are generated by first establishing sampling lines as illustrated in figure 1a. The cross boundary intensity profiles form geometrically normalized boundary images which are stored in the case database. The actual boundary is located along the horizontal center axis of the boundary images.

Geometric normalization of the patient case is performed on-line, based on the manually identified atrioventricular (AV) plane and apex point (AP). A spline through the landmarks serves as reference boundary for anatomical normalization of the patient tissue data, as

Table I
ERROR IN ESTIMATED PARAMETERS

Parameter	Computer to observer distance	
	Manual selection	Automatic selection
ES area(cm^2)	-0.70 ±1.4	-0.069± 3.1*
ED area (cm^2)	-0.074±2.6	0.56 ± 4.0*
EF(%)	-3.9 ±8.0	-3.9 ±16.0
ES Point error (mm)	3.4 ±0.93	4.1 ± 1.5
ED Point error (mm)	3.2 ±0.76	3.7 ± 1.3

* ($p < 0.05$) vs. manual SD. The means were not significantly different ($p < 0.05$)

illustrated in figure 1b. Pixel values are sampled along candidate lines normal to the reference boundary, resulting in a geometrically normalized candidate image.

C. Boundary detection

Boundary detection in the geometrically normalized patient case is formulated as an energy minimization problem, where the total energy is composed of external and internal energies of the boundary. The total energy of the problem is expressed as:

$$E_{total} = \sum E_{ext} + \sum E_{int} \tag{1}$$

The external energies E_{ext} are calculated as the correlation between corresponding columns in the candidate image and boundary signature image, as illustrated in figure 2. The value of the correlation image at a certain line n and displacement d gives us the correlation between the corresponding patient case candidate line and database boundary signature when the boundary signature is shifted incrementally along the candidate line. Therefore, the position of a correlation maximum indicates a high probability of a boundary at the specified distance from the reference boundary. Hence, boundary detection is now reduced to find a horizontal traverse of the correlation image that gives high probability of a boundary and at the same time a physiologically probable boundary shape. The correlation image can be written as a function $\mathbf{R}(d,n)$, where d corresponds to the displacement, and n represents one of the N pairs of candidate lines and boundary signatures. The total energy of a horizontal traverse of the correlation image is formulated as:

$$\sum E_{ext} = \sum_{n=0}^{N-1} \mathbf{R}(d_n, n) \tag{2}$$

where $\{d_n\}$ represents a horizontal traverse.

We see that the sliding correlation coefficient estimate may result in several maxima for one pair of candidate lines and boundary signatures. Therefore, we need a shape criteria that ensures a physiologically probable boundary. Parallelism between the detected boundary and the manually outlined database boundary is used for this purpose. The database boundary can be expressed by its distance from the reference boundary $\{d_n^{db}\}$, and the internal energy term can be expressed as:

$$\sum E_{int} = \sum_{n=1}^{N-1} \left| (d_n - d_{n-1}) - (d_n^{db} - d_{n-1}^{db}) \right| \tag{3}$$

The traverse $\{d_n^{opt}\}$ of the correlation image which is optimal with respect to our energy function in (1) is found efficiently by dynamic programming [4], [8], and results in a minimal
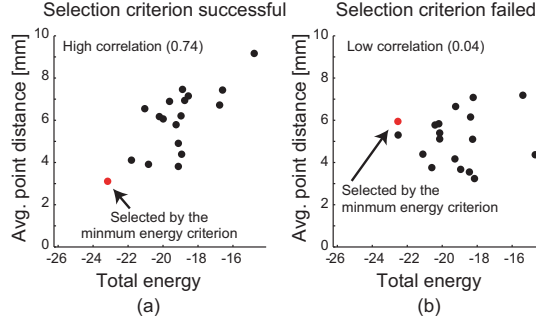
Figure 3.   These plots illustrates the average point distance between the detected boundary and the manually outlined boundary as a function of Snake energy. In pane a, the correlation is high, and the minimum snake energy is a good selection criterion. However, the correlation between point distance and Snake energy is typically as illustrated in pane b, and the proposed selection criterion will fail.

total energy E_{opt}. The detected boundary is calculated from the reference boundary and the optimal traverse $\left\{ d_n^{opt} \right\}$.

Each database case results in one detected boundary for the patient case, and a selection criterion is finally applied to select one. We have evaluated the minimum total boundary energy E_{opt} against manual selection of the best database boundary.

III. RESULTS

A database containing two-chamber apical views was established using a Vivid 7 ultrasound scanner (GE Vingmed Ultrasound, Horten, Norway). In total 19 cases was recorded and evaluated leaving one out between each experiment. For each case, the boundary was evaluated for both end systole (ES) and end diastole (ED) defined by maximum long axis shortening and mitral valve closing respectively. Figure 4a shows the patient case and the detected boundary at ED, along with the corresponding database case. Figure 4b shows the result of unfavourable case selection.

Each detected boundary was compared to manual delineation, calculating the average boundary point to point distance and ventricular area at ES and ED. For EF calculation, the long axis length was defined as the distance from the center of the AV plane to its distal point of the endocardial boundary. The differences between parameters estimated from the detected boundaries and manually outlined boundaries resulted in a computer to observer distance (COD) for each parameter. The COD mean and standard deviations are included in table I for both automatic case selection and for manual selection of the best case. The variance in the estimated areas is larger for automatic selection than for manual selection ($p < 0.05$, f-test). The means were not significantly different ($p < 0.05$, t-test).

To evaluate the potential of the total Snake energy as a selection criterion, the average point to point distance between the manually outlined and the detected boundaries were computed as a function of total Snake energy. The results for two patient cases are shown in figure 3. In pane a, the correlation between point to point distance and Snake energy is high. For this patient case, the Snake energy is a good selection criterion. However, the correlation is usually lower, and a typical example is shown in pane b, where the proposed selection criterion is suboptimal.

IV. DISCUSSION

We have developed a framework for knowledge based detection of the left ventricular endocardial boundary, which enables use of template cases within the Snake algorithm.
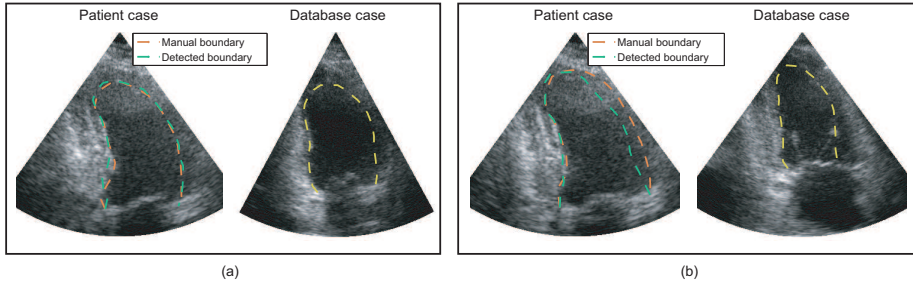
Figure 4.   Pane a shows an example of successful boundary detection in a patient case. The detected boundary is shown in green. A manually delineated boundary for the patient case is included in orange for comparison. The corresponding database case with the manually delineated boundary is also included. Pane b shows a case where boundary detection failed due to incorrect case selection.

One of the major potential of this approach is the ability to incorporate expert knowledge into the boundary detector. Due to noise, reverberations, and drop outs in cardiac ultrasound images, the strongest gradient in an image does not necessarily correspond directly to a tissue boundary. Using database cases as examples, our method can overcome such problems.

The Snake energy based selection criterion has been shown to fail in many cases even if some of the database cases resulted in good boundary detection. This is illustrated in figure 3 and confirmed by the higher variance in estimated areas when using automatic case selection compared to manual case selection (table I). In many cardiac ultrasound images, the signal from the endocardium is very low compared to the signal from the epicardium. If such a case is used as a database template for a patient case with strong signal from the endocardium, the epicardial boundary signature from the database case might correlate well with the endocardial boundary signature of the patient case. Consequently, the epicardium will be detected instead of the endocardium. As long as the shape criterion is not violated, the total Snake energy can be very low for such cases due to high correlation, and the selection criterion fails.

Another limitation of the proposed method is its dependency on a diverse image database. An artifact or pathology that is not present in the database will not necessarily be handled satisfactory. Performance issues will inevitably arise for large databases since all database cases must be evaluated for each patient case.

V. CONCLUSION

We have established a framework that enables the Snake algorithm to utilize example cases from a database for endocardial boundary detection. Visual selection of database case gave good boundary detection, while the total Snake energy is not sufficiently robust as a case selection criterion.

ACKNOWLEDGEMENTS

References

[1] J. Klingler, C. Vaughan, J. Fraker, T.D., and L. Andrews, "Segmentation of echocardiographic images using mathematical morphology," *IEEE Transactions on Biomedical Engineering*, vol. 35, no. 11, pp. 925–934, 1988.

[2] W. Ohyama, T. Wakabayashi, F. Kimura, S. Tsuruoka, and K. Sekioka, "Automatic left ventricular endocardium detection in echocardiograms based on ternary thresholding method," in *15th International Conference on Pattern Recognition*, vol. 4, 2000, pp. 320–323.

[3] A. Amini, T. Weymouth, and R. Jain, "Using dynamic programming for solving variational problems in vision," *IEEE Trans PAMI*, vol. 12, no. 9, pp. 855–867, 1990.

[4] S. I. Rabben, A. H. Torp, A. Støylen, S. Slørdahl, K. Bjørnstad, B. O. Haugen, and B. Angelsen, "Semiautomatic contour detection in ultrasound m-mode images," *Ultrasound in Med. & Biol.*, vol. 26, no. 2, pp. 287–296, 2000.

[5] J. Bosch, S. Mitcheell, B. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. Reiber, "Fully automated endocardia1 contour detection in time sequences of echocardiograms by active appearance motion models," *Computers in Cardiology*, vol. 28, pp. 93–96, 2001.

[6] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance motion models," *IEEE Trans Medical Imaging*, vol. 21, no. 11, pp. 1373–1383, 2002.

[7] M. Sonka, B. P. Lelieveldt, S. C. Mitchell, J. G. Bosch, R. J. van der Geest, and J. H. Reiber, "Active appearance motion model segmentation," *Second International Workshop on Digital and Computational Video (DCV'01)*, pp. 64–68, 2001.

[8] R. Bellman, *Dynamic Programming*, N. Princeton, Ed. Princeton University Press, 1957.

Paper B

Detection of the myocardial boundary in the left ventricle from simultaneously acquired triplane ultrasound images using multi view active appearance motion models

J. Hansegård, S. Urheim, E. Steen, H. Torp, B. Olstad, S. Malm, and S. I. Rabben

Detection of the myocardial boundary in the left ventricle from simultaneously acquired triplane ultrasound images using multi view active appearance motion models

Jøger Hansegård[*], Stig Urheim[§], Erik Steen[†], Hans Torp[‡], Bjørn Olstad[†], Siri Malm[‡] and Stein Inge Rabben[†] [*]Department of Informatics, University of Oslo, Oslo, Norway [†]GE Vingmed Ultrasound, Horten, Norway [‡]Department of Circulation and Imaging, Norwegian University of Science and Technology, Trondheim, Norway [§]Rikshospitalet University Hospital, Oslo, Norway

Abstract

We report a new algorithm for detecting the LV myocardial boundary from simultaneously acquired triplane US image sequences using Multi View Active Appearance Motion Models. Coupled boundary detection in three planes can potentially increase the accuracy of LV volume measurements, and also increase the robustness of the boundary detection over traditional methods. A database of triplane image sequences from full cardiac cycles, including the standard A4CH, A2CH, and ALAX views were established from 20 volunteers, including 12 healthy persons and 8 persons suffering from heart disease. For each dataset the LV myocardial boundary was manually outlined, and the ED and ES frames were determined visually for phase normalization of the cycles. The evaluation of the MVAAMM was performed using a leave one out approach. The mean point distance between manually and automatically determined contours were 4.1±1.9 mm, the volume error was 7.0±14 ml, and fractional volume error was 8.5±16 %. Volume detection using the automatic method showed excellent correlation to the manual method (R^2=0.87). Common ultrasound artefacts such as dropouts were handled well by the MVAAMM since the detection in the three image planes were coupled. The views with the largest point distance had one or more foreshortened views. A larger training database may improve the performance in such cases.

I. Introduction

Automatic detection of the myocardial boundary of the left ventricle (LV) is a key challenge in cardiac ultrasound (US) imaging, since several clinically important parameters such as ejection fraction (EF), stroke volume and cardiac output can be computed from this boundary.

A. Motivation

The latest generation of three dimensional (3D) cardiac US scanners are capable of recording multiple imaging planes simultaneously. Thus, the three apical views, including the apical four chamber (A4CH), two chamber (A2CH), and long axis (ALAX) views can be acquired at the same time. This is called triplane acquisition, and is illustrated in figure 1. Using triplane images, the accuracy of the estimated volume is expected improve over traditional single plane methods. It is therefore a need for automatic methods for detection of the myocardial boundary in triplane recordings.
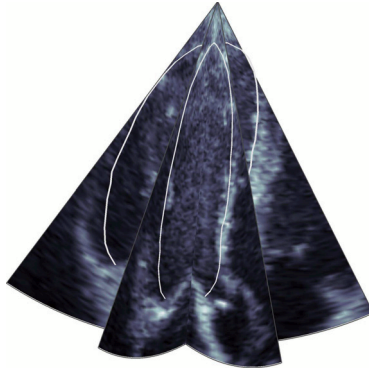
Figure 1.   The 3D acquisition configuration for triplane recordings results in three different imaging planes which can be tuned to each patient case, resulting in the standard apical four chamber, two chamber, and long axis views. Examples of detected contours using the Multi View Active Appearance Motion Model are shown as white lines.

B. Background

The Active Appearance Model (AAM) was introduced by Cootes and Taylor [1] as a generalization to Active Shape Models introduced earlier [2]. This is a generative deformable model that performs well for segmentation of noisy data. The AAMs combine knowledge of object shape and texture from a segmented image database, using principal component analysis (PCA) to extract the major components of variation in the training data. This method was extended by Bosch *et al* to include motion patterns of the heartbeat into the model [3], resulting in a method called Active Appearance Motion Models (AAMM). This model performed better than the classic AAM for cardiac US data. Multi View Active Appearance Models (MVAAMs) [4] are another extension to the AAM framework, which can do coupled boundary detection in multiple views, but only in single frames.

C. Contribution

We propose a new algorithm for time continuous detection of the LV myocardial boundary in simultaneously acquired triplane US image sequences. This algorithm is a combination of the AAMM and the MVAAM, resulting in a Multi View Active Appearance Motion Model (MVAAMM). The MVAAMM incorporates an expert's knowledge of the spatial correlation between imaging planes and temporal correlation between image frames from different parts of the cardiac cycle. Since the algorithm has knowledge of this correlation, we obtain probable ventricular shapes even if data is missing in parts of the dataset, for example caused by drop outs. In the rest of this report we will describe this model, and give an idea of its performance based on experiments on a small dataset of triplane echocardiograms.

II. METHOD

The MVAAMM is a deformable model trained on data segmented manually by an expert. The LV myocardial boundary of a new patient is extracted by fitting the model to the triplane echocardiograms. The steps involved in training and fitting the model are explained in the following sections.

A. Preprocessing

The myocardial boundaries were outlined manually in the triplane data sets by an expert, resulting in contours enclosing the US signal from the ventricular cavity. The MVAAMM

depends upon texture patches which include signal from the myocardium, and a contour inflation method was developed to grow the contours outwards a specified distance. First, the contour was converted to an implicit contour, resulting in a bitmap of the curve. The inflated contours were then found from the distance maps generated using a morphological distance transform. The same approach was used for deflation of detected contours.

For removal of speckle noise, the ultrasound images were filtered using a difference of Gaussians bandpass filter, where the lower 3 dB pass frequency was on average 0.047 mm^{-1}, while the upper 3 dB pass frequency was on average 0.11 mm^{-1}. These frequencies corresponds to structures of 21 mm and 9.1 mm respectively. The 30 dB band stop frequencies were 0.0084 mm^{-1} and 0.21 mm^{-1}, corresponding to 12 cm and 4.8 mm.

B. MVAAMM formulation

Building the MVAAMM consists of three steps. First we build a model for the ventricular shapes. Then we build a similar model for the textures within the shapes. The third step is to combine these two models into a model of appearance.

1) Modeling the ventricular shapes: Given a manually segmented US dataset of V views and F frames, the myocardial boundary in each view and frame can be defined by N landmarks. The first and last landmark are true anatomical landmarks, corresponding to the atrioventricular (AV) plane, while the remaining landmarks are distributed uniformly along the boundary. Each landmark can be identified uniquely by their spatial x, y, and z coordinates along with time t, but for convenience we adopt an image centric coordinate system, where each landmark is defined by two spatial coordinates x and y. For the case with a single frame f and a single view v, we concatenate the coordinates of the landmarks in this frame and view into a shape vector:

$$\mathbf{x}_{vf} = (x_1, y_1, x_2, y_2, \ldots, x_N, y_N)^T$$

We describe the ventricular shapes for all frames and views by a shape vector:

$$\mathbf{x} = (\mathbf{x}_{11}^T, \ldots, \mathbf{x}_{V1}^T, \mathbf{x}_{12}^T, \ldots, \mathbf{x}_{V2}^T, \ldots, \mathbf{x}_{1F}^T, \ldots, \mathbf{x}_{VF}^T)^T$$

The frame number refer to the normalized cardiac cycle, where $f = f_{ED}$ corresponds to the end diastolic (ED) frame, while $f = f_{ES}$ corresponds to the end systolic (ES) frame. The remaining $F - 2$ frames are uniformly distributed between these time events. From a cardiac US database of S manually segmented cases, we obtain S ventricular shapes. Normalization of the shapes to unit size and zero mean combined by alignment using Procrustes' method gives us the shape vectors $\{\mathbf{x}_i\}, i = 1, \ldots S$. Now, the mean shape $\bar{\mathbf{x}}$ and shape covariance matrix $\mathbf{C}$ can be estimated,

$$\bar{\mathbf{x}} = \frac{1}{S}\sum_{i=1}^{S}\mathbf{x}_i \quad \text{and} \quad \mathbf{C} = \frac{1}{S-1}\sum_{i=1}^{S}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$$

and PCA gives us the shape eigenvector matrix $\mathbf{P}_s$ along with the eigenvalues, or shape coefficients, $\mathbf{b}_s$ of the training set.

The aligned and normalized database cases can now be approximated by a sum of the mean shape $\bar{\mathbf{x}}$, and a linear combination of the eigenvectors $\mathbf{P}_s$ using the shape coefficients as weights $\mathbf{x} \approx \bar{\mathbf{x}} + \mathbf{P}_s\mathbf{b}_s$.
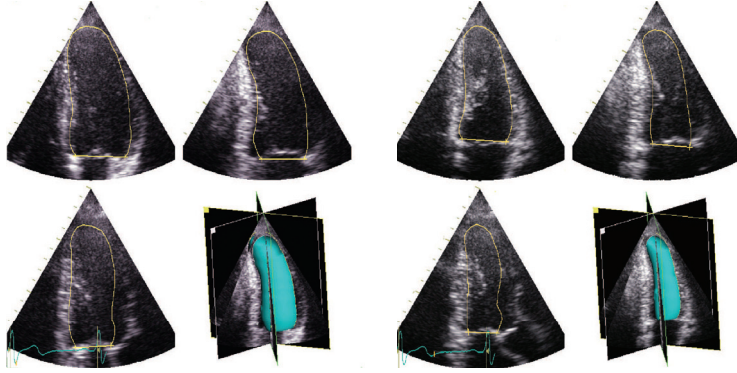
Figure 2.   The detected contours at ED (left) and ES (right) is shown along with a volume model generated from the contours.

2) Modeling the ventricular textures: A model of the textures in the training set is created using the same principles as for the shapes, by rearranging the pixel values within the shapes into a texture vector. This is done by first defining a common reference texture geometry, in our case we chose the mean ventricular shape. The texture patches within the expanded manually defined contours were warped onto this reference frame using thin plate splines [5]. This results in one texture vector $\mathbf{g}_i$ for each case i in the database. If the number of pixels in the reference coordinate system for view v and frame f is $n_{v,f}$, and the pixel value number j in this view and frame has a intensity $g_{i,j,v,f}$, the texture vectors can be expressed as:

$$\mathbf{g}_i = (\{g_{i,j,v,f}\}_{j=1,v=1,f=1}^{n_{v,f},V,F})^T, \quad i=1,\ldots,S$$

The texture vectors are normalized using the non linear normalization method described by Bosch *et al* [3], followed by the standard AAM normalization procedure [1].

As for the shapes, we use PCA to obtain the texture eigenvectors $\mathbf{P}_g$ and the associated coefficients $\mathbf{b}_s$ of the training set. The textures in the training set can now be approximated by a sum of the mean texture and a linear combination of the eigenvectors $\mathbf{g} \approx \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g$. Then we form a combined shape and texture vector from the coefficients:

$$\mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix}$$

where $\mathbf{W}_s$ is a weight matrix. A third PCA is applied to these vectors, obtaining $\mathbf{b} = \mathbf{P}_c \mathbf{c}$ where $\mathbf{P}_c$ is the eigenvectors, and $\mathbf{c}$ is the coefficients of appearance. $\mathbf{P}_c$ can be split in two parts, one representing shapes, and one representing textures:

$$\mathbf{P}_c = \begin{pmatrix} \mathbf{P}_{cs} \\ \mathbf{P}_{cg} \end{pmatrix}$$

New textures and shapes can be generated by choosing the elements in the appearance coefficient vector:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s \mathbf{W}_s^{-1} \mathbf{P}_{cs} \mathbf{c}, \quad \mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{P}_{cg} \mathbf{c}$$

C. MVAAMM matching

Boundary detection in a target patient dataset is done by finding the parameters $\mathbf{c}$ of the MVAAMM and pose transformation $\mathbf{p}$ that minimizes the squared sum of pixel difference $\mathbf{r}(\mathbf{c}, \mathbf{p})$ between the model texture and the patient texture sampled under the model shape.

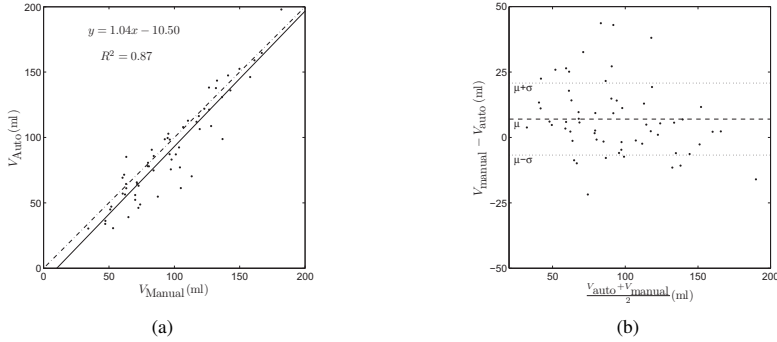(a)   (b)

Figure 3.   The automatically detected boundaries are compared to the manually determined boundaries in this scatter plot (a). The dotted line shows the line of identity, while the solid line shows the linear regression line. The Bland-Altmann plot (b) indicates a small bias in the data, and a t-test reveals that this bias is statistically significant ($P < 0.05$).

We use the standard AAM matching algorithm, where the model is fitted iteratively to the target triplane data set, starting with the average model and manually defined initial pose. $\mathbf{r}(\mathbf{c}, \mathbf{p})$ is calculated for each iteration, and the required change in parameters is found from $\delta\mathbf{c} = -\mathbf{R}_g\mathbf{r}(\mathbf{c}, \mathbf{p})$ where $\mathbf{R}_g$ is precalculated on the training set as described in [1]. This matching is performed on all views and frames simultaneously. For pose estimation, the MVAAMM matching algorithm differs slightly from the AAM search. In the AAM formulation, pose and model parameters are coupled, but for triplane recordings this is not suitable, due to operator dependent variations in probe orientation relative to the cardiac long axis. Therefore, the pose is estimated separately for each view v using the equation $\delta\mathbf{p}_v = -\mathbf{R}_{s,v}\mathbf{r}(\mathbf{c}, \mathbf{p}_v)$, where we have separate $R_{s,v}$ for each view.

When the model has converged, the detected contours are deflated using the technique described previously, and a dynamic 3D surface model is generated using spline interpolation. The ventricular volumes are computed from this model.

III. EXPERIMENTS AND RESULTS

The automatic boundary detector was evaluated on a database of triplane echocardiograms from 20 volunteers, including 12 healthy persons, and 8 persons suffering from heart disease. Patients with asynchronies were excluded from the study. The ventricular wall was outlined manually in the three planes at ED, ES and mid systole (MS) resulting in 9 contours and texture patches from each dataset. Variations within the training set is of critical importance when training the models, and we used a leave-one-out technique for training and evaluation. We trained 20 models from the database, and one dataset was left out from each model. The model was then evaluated on the dataset that was left out.

Detection was initialized in the first frame by manually identifying three landmarks, including the mitral valve plane and the LV apex in the first frame of each image plane. The mean ventricular shape was rotated, scaled, and placed in the image frame to match the landmarks. As an initial guess of the texture, we used the mean texture.

The detected myocardial boundaries were compared to the manually outlined boundaries using average point distance, calculated as the mean of the euclidean distances between pairs of corresponding landmarks. We also computed the error of the automatically determined volumes (V_{Auto}) compared to volumes from manual segmentation (V_{Manual}) as $V_{err} = V_{Manual} - V_{Auto}$. The fractional error ($V_{\%err} = 1 - V_{Auto}/V_{Manual}$) describes the error relative to the magnitude of the measurement.

Table I
MVAAMM-RESULTS

Results (mean ±stddev)	
Mean Point Distance (mm)	4.1±1.9
Error (V_{Manual}-V_{Auto})	7.0±14*
Volume fractional error (1-V_{Auto}/V_{Manual}) (%)	8.5 ± 16
Volume regression (y=V_{Auto}, x=V_{Manual}) (ml)	$y = 1.0x - 11$
Volume correlation coefficient (R^2)	0.87

* Significantly different from 0 ($p < 0.05$)

The mean landmark distance between manually and automatically determined contours were 4.1 mm, with a standard deviation of 1.9 mm. The volumes from automatically detected boundaries are compared to the volumes from the manually determined boundaries in figure 3(a), which includes data from ED, ES, and MS from the 20 leave-one-out experiments. As we see of the scatter plot, there is a high correlation between the manually and automatically determined volumes ($R^2 = 0.87$). The automatic method underestimates the volume by 7 ml with a standard deviation of 14 ml as shown in the Bland-Altmann plot in figure 3(b). The error relative to manually determined volumes is 8.5%±16%. An example of detected contours at ED and ES is shown in figure 2, and the results are summarized in table I.

We observed that dropouts were handled remarkably well compared to other edge detectors, according to our experience. We also observed that the data sets with the highest point distance between manually and automatically determined contours corresponded to recordings with foreshortening in one or more of the views. The results are slightly less accurate than, but comparable to, the ones obtained by Bosch *et al* [3]. This is expected, since we have trained our models on a smaller number of datasets, giving fewer variational modes in the models. The correlation between EFs derived from manually and automatically determined contours is poor ($R^2 = 0.29$). A relatively small variation in the training set is probably one reason, as is the MVAAMM's tendency to favor a too normal contraction pattern [3]. These challenges will be topics for further studies of the algorithm.

IV. DISCUSSION

We have developed a multi view active appearance motion model for detection of the LV myocardial boundary in simultaneously acquired triplane ultrasound echocardiograms. The model performs detection in all planes and multiple frames simultaneously. We have tested the model on a small database of triplane echocardiograms with encouraging results, comparable to AAMMs trained on a larger data set.

Since the three planes are coupled during detection, drop outs are handled remarkably well. Foreshortening of the views is suspected to contribute to the bias of the estimated volumes. A larger training database with a larger variation of both anatomy and artefacts will probably improve the performance of the algorithm.

ACKNOWLEDGEMENTS

REFERENCES

[1] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *ECCV '98: Proc 5th European Conf. on Computer Vision*, vol. 2. Springer-Verlag, 1998, pp. 484–498.

[2] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," in *IPMI*, 1993, pp. 33–47.

[3] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance motion models," *IEEE Trans Medical Imaging*, vol. 21, no. 11, pp. 1374–1383, 2002.

[4] C. R. Oost, B. P. F. Lelieveldt, M. Üzümcü, H. Lamb, J. H. C. Reiber, and M. Sonka, "Multi-view active appearance models: Application to x-ray lv angiography and cardiac mri," in *IPMI*, 2003, p. 234.245.

[5] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, 1989.

Paper C

Constrained Active Appearance Models for Segmentation of Triplane Echocardiograms

J. Hansegård, S. Urheim, K. Lunde, and S. I. Rabben

Constrained active appearance models for segmentation of triplane echocardiograms

Jøger Hansegård, *Member, IEEE,* Stig Urheim, Ketil Lunde,
and Stein Inge Rabben, *Member, IEEE*

Abstract

This paper presents multiview and multiframe active appearance models (AAMs) for left ventricular segmentation in triplane echocardiograms. We describe a general way of integrating local edge detector based segmentation algorithms into the AAM framework. The feasibility of this approach is evaluated by comparing an AAM constrained by a dynamic programming (DP) based snake with an unconstrained AAM, and an AAM constrained by manually defined landmarks.

A leave-one-out validation scheme was used for training and testing of the methods. Evaluation was done in 36 patients suffering from various heart diseases, using manually determined volumes and ejection fractions (EF) as reference. The segmentation was initialized by manual selection of the mitral annulus and apex in three imaging planes. The differences, in volume, between manual segmentation and the best automatic method (DP-constrained AAM) were -3.1 ± 20 ml (mean±SD) at end-diastole and 0.61 ± 13 ml at end-systole. The difference in EF was -1.3 ± 6.3 %, comparable to the inter-observer variability.

We show that (1) constraining the model to manually defined landmarks improves volume and EF estimates compared to unconstrained AAMs, (2) further improvement is achieved using a DP-constrained AAM, and (3) segmentation in triplane echocardiograms gives higher accuracy than single plane data.

I. Introduction

LEFT ventricular (LV) volumes and ejection fraction (EF) are important clinical measures in diagnosis and prognosis of patients with heart diseases. Conventionally, LV volumes are measured manually in two-dimensional (2D) echocardiography, using the disk summation method on one or two apical views [1], [2].

Triplane imaging is a new ultrasound (US) modality allowing simultaneous acquisition of three apical planes rotated about the long axis of the left ventricle as shown in Fig. 1. Potentially, the triplane modality will give more accurate volume measurements since all planes are recorded simultaneously and the geometric relations between the image planes are known. As for 2D echocardiography, volumes in triplane data are measured manually, a time-consuming task prone to inter- and intra-observer variability [3]. Hence, there is a need for automated extraction of the myocardial boundary from triplane echocardiograms.

Automatic segmentation of echocardiograms is a challenging task for several reasons. US image quality is affected by speckle noise, shadows from the lungs and ribs, and energy absorption in subcutaneous fat [4]. Further, foreshortening occurs when the imaging plane is not slicing through the true ventricular apex, causing underestimated volumes. Additionally, papillary muscles are distinct structures within the cavity that the echocardiographer may want to exclude [5], and the mitral valve, when open during diastole, may be misinterpreted as a part of the wall by a naive segmentation approach. An experienced echocardiographer is able to do good volume measurements despite these challenges because he or she has developed strategies for compensating for the different sources of errors.

Traditionally, automatic segmentation of US images has been based upon active contour models, also called "Snakes", introduced by Kass *et al.* [6], where the contour is fitted to an object by minimizing its internal and external energy. Amini *et al.* [7] showed that dynamic
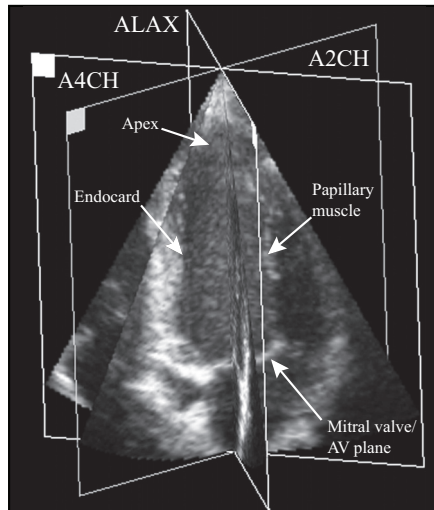
C-1

Figure 1.   Triplane echocardiography allows simultaneous acquisition of three apical planes, including the standard apical four chamber (A4CH), two-chamber (A2CH), and the long axis (ALAX) views.

programming is a well-suited optimization strategy for active contours. A disadvantage of these methods is that expert knowledge cannot be modeled directly.

The active shape model (ASM) introduced by Cootes and Taylor [8], [9] incorporates expert knowledge in a global shape model. Later the same authors introduced the active appearance model (AAM) where knowledge of the object's texture is also included [10], [11]. Due to the robustness of AAMs to noisy data, these models are particularly suitable for segmentation in US data. To model the temporal behavior of the left ventricle in 2D echocardiography, Bosch *et al.* [12] introduced the active appearance motion model (AAMM), where data from the entire cardiac cycle is combined into a single model. Segmentation of multiview data from cardiac magnetic resonance imaging (MRI) was presented in [13] and [14], where a multiview AAM utilized spatial correlation between multiple imaging planes. Combined multiview and multiframe AAMs were presented in [15], [16], and [17]. A third extension to the basic AAM is the constrained active appearance model (CAAM) [18], where the model is constrained to known contour points.

Initialization of the AAM can have significant impact on the final result, especially if the model is initialized far from the object of interest [19]. The texture matching technique used for AAMs is well suited for producing contours with nice global properties, but our experience is that in echocardiograms with poor image quality, the local accuracy is lower than what can be obtained by active contours optimized by dynamic programming (DP). A related problem specific for the AAMM is its tendency to prefer a "too-normal" contraction pattern [20] causing poor EF measurements.

These limitations were addressed in [19] by using manual initialization, and by using DP to improve the local quality of the AAM segmentation. However, the AAM was not constrained by the manually defined landmarks, and DP was applied as a post-processing step, not as an integrated part of the AAM segmentation algorithm.

Contributions

In this paper, we describe a general method for integrating segmentation algorithms based on local image evidence with the AAM framework. Further, we demonstrate the use of

multiview AAMMs for time continuous segmentation of triplane echocardiograms, and we present a method for constraining the multiview AAMM to manually defined landmarks by using the CAAM algorithm.

II. Background

A. Basic active appearance models

The AAM is described in detail in [11], [21], and [20], but to make this text self-contained, we describe the basic steps here.

In a 2D image, an object's shape can be modeled as a coordinate vector

$$\mathbf{x} = (x_1, y_1, x_2, y_2, \ldots x_n, y_n)^T$$

of $2n$ elements, while its texture vector $\mathbf{g}$ is found by warping the object's intensities onto a fixed reference frame. The shapes from all training sets are aligned using Procrustes alignment [22] to remove pose differences, while the texture vectors are normalized to zero mean and unit variance [10]. A statistical model for shape and texture can be written in the form

$$
\begin{aligned}
\mathbf{x} &= \bar{\mathbf{x}} + \mathbf{P}_s \mathbf{b}_s \\
\mathbf{g} &= \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g
\end{aligned}
\tag{1}
$$

where $\bar{\mathbf{x}}$ and $\bar{\mathbf{g}}$ are the mean shape and texture, $\mathbf{P}_s$ and $\mathbf{P}_g$ are basis vectors for shape and texture, and $\mathbf{b}_s$ and $\mathbf{b}_g$ are the shape and texture model parameters. The basis vectors are obtained by applying principal component analysis (PCA) to the shape and texture vectors of a training set. These models are combined by a third PCA on the concatenated shape and texture parameters

$$
\mathbf{b} = \begin{pmatrix} W\mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \begin{pmatrix} \mathbf{W}\mathbf{P}_s^T(\mathbf{x} - \bar{\mathbf{x}}) \\ \mathbf{P}_g^T(\mathbf{g} - \bar{\mathbf{g}}) \end{pmatrix}
$$

where $\mathbf{W}$ is a suitable weight matrix for combining vectors of different quantities. The resulting AAM can be written as

$$
\begin{aligned}
\mathbf{x} &= \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{c} \\
\mathbf{g} &= \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{c}
\end{aligned}
\tag{2}
$$

where $\mathbf{Q}_g$ and $\mathbf{Q}_s$ denote the basis vectors for texture and shape of the combined model, and $\mathbf{c}$ are the model parameters controlling both shape and texture at the same time. This model will produce points in the normalized model frame, but a suitable transformation $T_{\mathbf{t}}(\mathbf{x})$ with parameters $\mathbf{t}$ transforms these points into the image frame $\mathbf{X}$. Commonly, $T_{\mathbf{t}}$ is a similarity transform with four parameters describing rotation, scale, and translation.

The model parameters $\mathbf{c}$ and pose parameters $\mathbf{t}$ define the position of the model points in the image frame $\mathbf{X}$, and can be combined to a single parameter vector $\mathbf{p}^T = (\mathbf{c}^T | \mathbf{t}^T)$ using concatenation. During matching, pixels in this region of the image $\mathbf{g}_{im}$, are sampled and projected into the texture model frame $\mathbf{g}_s = T_{\mathbf{u}}(\mathbf{g}_{im})$. In the standard AAM formulation, the transform $T_{\mathbf{u}}$ shifts and scales the texture intensities. The difference between the model texture $\mathbf{g}_m$ given by (2) and the normalized image texture $\mathbf{g}_s$ is expressed by the residual vector $\mathbf{r}(\mathbf{p}) = \mathbf{g}_s - \mathbf{g}_m$, and its sum of squares $E(\mathbf{p}) = \mathbf{r}^T \mathbf{r}$ is used as a measure of the model fit. Assuming linearity, E is minimized by modifying $\mathbf{p}$ according to

$$\delta \mathbf{p} = -\mathbf{R}\mathbf{r}(\mathbf{p}) \tag{3}$$

where $\mathbf{R}$ is the pseudo inverse of the Jacobian of $\mathbf{r}(\mathbf{p})$ with respect to the model parameters $\mathbf{p}$ estimated from the training data by numerical differentiation:

$$\mathbf{R} = \left(\frac{\delta \mathbf{r}^T}{\delta \mathbf{p}} \frac{\delta \mathbf{r}}{\delta \mathbf{p}} \right)^{-1} \frac{\delta \mathbf{r}^T}{\delta \mathbf{p}} . \tag{4}$$

Assuming a fixed Jacobian throughout the search, the Jacobian can be precomputed during model training by applying perturbations to the model parameters and measuring the effect on the residuals.

The classical fitting scheme for the basic AAM is an iterative procedure [10] where model parameters are updated for each iteration using (3).

B. Independent active appearance models

Independent AAMs [23] are slightly different from the basic AAM in that the shape and texture are modeled separately using (1) rather than using a combined model. Therefore, training of independent AAMs is done by applying perturbations to the pose parameters $\mathbf{t}$, shape parameters $\mathbf{b}_s$, and texture parameters $\mathbf{b}_g$ separately, and we obtain three pseudo inverses, $\mathbf{R}_t$, $\mathbf{R}_s$, and $\mathbf{R}_g$. They are used during matching to compute updates to the shape, texture, and pose parameters separately using (3).

C. Constrained active appearance models (CAAMs)

The CAAM, originally described in [18] allows contour points to be constrained to known positions with given variances.

We assume that the texture residuals $\mathbf{r}$ are uniform Gaussian with variance σ_r^2 and that the model parameters are Gaussian with diagonal covariance $\mathbf{S}_p^2$. Since we know the positions of some contour points, we can construct a landmark vector $\mathbf{X}_0$ of the same length as the shape vector $\mathbf{x}$, where unknown points are set to zero. To this landmark vector, we associate a diagonal covariance matrix $\mathbf{S_X}^2$ where unknown contour points have infinite variance, and known points are assigned a variance representing the confidence of the constraint. We can then compute the distances $\mathbf{d}$ from $\mathbf{X}_0$ to the corresponding model points $\mathbf{X}$ as a function of the model configuration from $\mathbf{d}(\mathbf{p}) = \mathbf{X} - \mathbf{X}_0$. Using a maximum a-posteriori (MAP) formulation, it can be shown [18] that a measure of the quality of fit of the model is

$$E = \sigma_r^{-2} \mathbf{r}^T \mathbf{r} + \mathbf{p}^T \mathbf{S}_p^{-1} \mathbf{p} + \mathbf{d}^T \mathbf{S_X}^{-1} \mathbf{d}$$

and we obtain the update step of the CAAM by solving

$$\mathbf{A} \delta \mathbf{p} = -\mathbf{a} \tag{5}$$

where

$$\begin{aligned} \mathbf{A} &= \sigma_r^{-2} \frac{\delta \mathbf{r}^T}{\delta \mathbf{p}} \frac{\delta \mathbf{r}}{\delta \mathbf{p}} + \mathbf{S}_\mathbf{p}^{-1} + \frac{\delta \mathbf{d}^T}{\delta \mathbf{p}} \mathbf{S_X}^{-1} \frac{\delta \mathbf{d}}{\delta \mathbf{p}} \\ \mathbf{a} &= \sigma_r^{-2} \frac{\delta \mathbf{r}^T}{\delta \mathbf{p}} \mathbf{r}(\mathbf{p}) + \mathbf{S}_\mathbf{p}^{-1} \mathbf{p} + \frac{\delta \mathbf{d}^T}{\delta \mathbf{p}} \mathbf{S_X}^{-1} \mathbf{d} . \end{aligned} \tag{6}$$

Here both pose transformation and shape changes are folded into the vector $\mathbf{p} = (\mathbf{c}^T | \mathbf{t}^T)^T$, and the Jacobian of the distances $\frac{\delta \mathbf{d}}{\delta \mathbf{p}}$ can be decomposed as $\frac{\delta \mathbf{d}}{\delta \mathbf{p}} = \left(\frac{\delta \mathbf{d}}{\delta \mathbf{c}} | \frac{\delta \mathbf{d}}{\delta \mathbf{t}} \right)$. The Jacobians $\frac{\delta \mathbf{r}}{\delta \mathbf{p}}$ of the residual and $\frac{\delta \mathbf{d}}{\delta \mathbf{c}}$ can be assumed constant and are therefore precomputed during training. The change in distance as a function of perturbations to model pose $\frac{\delta \mathbf{d}}{\delta \mathbf{t}}$ is computed for each iteration of the model update sequence.

The matrix $\mathbf{S_X}^2$ can be regarded as a weight matrix where low values indicate high confidence. Infinite values mean that no restrictions are put on the corresponding points. Too low values can potentially over-constrain the model resulting in poor segmentation; also, the number of constraints will have impact on the resulting segmentation.

III. Segmentation method

In this section, we present our multiview AAMM for time continuous segmentation of triplane echocardiograms. We describe methods for constraining this model to manually defined landmarks, and we show how segmentation algorithms based on local image evidence can be integrated with the AAM algorithm.

A. Multiview AAMMs

First we build a shape motion vector per view by concatenating shapes from all F frames $\mathbf{x}_m = (\mathbf{x}_1^T | \mathbf{x}_2^T | \ldots | \mathbf{x}_F^T)^T$. After pose alignment, the shape motion vectors from all N views are concatenated to give one multiview shape motion vector for each training set

$$\mathbf{x}_{vm} = (\mathbf{x}_{m,1}^T | \mathbf{x}_{m,2}^T | \ldots | \mathbf{x}_{m,N}^T)^T \ .$$

Applying PCA on the combined shapes results in a basis for the shape model according to (1). This shape model differs from the basic AAM in that all frames and views are modeled simultaneously, controlled by a single parameter vector.

In triplane echocardiography, all views are acquired simultaneously from a single point of view. The LV poses in each of the three image planes are therefore strongly related, but since the LV is often shifted and rotated relative to the probe axis, this relationship is non-trivial. We are therefore using separate similarity transforms $T_{\mathbf{t}}(\mathbf{x})$ for each view, each governed by four parameters, giving a total of 12 pose parameters for a multiview AAMM with three views. A weak coupling of the pose between each view is enforced by weighting the rotation, scale, and vertical position towards the mean over the views after each iteration of the AAM fitting scheme. Pose variations through the cardiac cycle are included in the model, and the same pose transforms are used for each frame.

Construction of the texture model is similar to the shape model. First, the image patches are warped onto the average shape of the corresponding frame and view. The texture vectors are then concatenated for each view separately to obtain a texture motion vector $\mathbf{g}_m$, and the resulting texture motion vectors from each view are combined into a single multiview texture motion vector

$$\mathbf{g}_{vm} = (\mathbf{g}_{m,1}^T | \mathbf{g}_{m,2}^T | \ldots | \mathbf{g}_{m,N}^T)^T \ .$$

Ultrasound imagery has highly non-Gaussian histograms giving the texture vectors roughly inversely exponential or chi-square pixel distributions [20]. Therefore, the standard AAM normalization procedure results in highly skewed intensity histograms. PCA assumes that each element of the texture vectors are approximately Gaussian over the training set. Therefore, Bosch *et al.* [20] introduced a non-linear normalization procedure for US data. This was done by establishing an empirical pixel distribution from all of the range normalized texture vectors. Each texture vector was then mapped to the standardized normal distribution separately. In our experiments, this procedure is employed prior to the standard AAM normalization, and PCA is finally used to create a texture model, which combines information from all frames and views.

For algorithms based upon the CAAM framework, a third PCA was applied resulting in a single multiframe and multiview model of texture and shape. For unconstrained AAMMs, we used the independent AAM formulation.

Optimum perturbations for estimating the Jacobians were found experimentally, with two perturbations for each parameter. For independent AAMs, perturbations of both shape and texture parameters were ±0.5 standard deviations (SDs) of the total variation of the model parameters in the training set. Pose parameter perturbations were $\pm10\%$ in scale, ±6 degrees in rotation, and ±2 pixels in horizontal and vertical translation for each view separately.

In our experiments, the multiview AAMM included three views and six frames. The shape within a single frame and view consisted of 40 points, resulting in combined shape vectors consisting of 720 points. The corresponding texture vectors had approximately 43000 elements. For single plane models, the combined shape vectors consisted of 240 points, and the texture vectors had approximately 15000 elements.

B. Landmark constrained AAMMs

The extension of CAAMs to multiple views and frames is straightforward. Assume that we have constructed landmark vectors $\mathbf{X}_{0,i}$ with associated diagonal covariance matrices for all frames and views in the data set. If some frames or views have no corresponding landmarks, all elements of the landmark vector are zero, while the diagonal elements of the covariance matrix are infinity. For each view separately, we can now construct new landmark motion vectors by concatenating vectors from all F frames: $\mathbf{X}_{0,m}^T = (\mathbf{X}_{0,1}^T|\mathbf{X}_{0,2}^T|\ldots,|\mathbf{X}_{0,F}^T)$. These vectors are in turn concatenated into a multiview landmark motion vector $\mathbf{X}_{0,mv}^T = (\mathbf{X}_{0,m,1}^T|\mathbf{X}_{0,m,2}^T|\ldots|\mathbf{X}_{0,m,N}^T)$. Construction of the diagonal covariance $\mathbf{S}_{\mathbf{X},mv}^2$ follows the same pattern. The multiview landmark motion vectors, shape vectors, and texture vectors can be plugged directly into (6), and optimum parameter updates are found by solving (5) as for the standard CAAM.

Training of the CAAM is similar to the standard AAM approach, but in addition to computing the Jacobian of the texture residuals, we estimate $\frac{\delta \mathbf{d}}{\delta \mathbf{c}}$ by perturbing each parameter ±0.5 SDs of the variation within the training set.

Assuming that manually defined landmarks are accurate, the values of the associated weight matrix $\mathbf{S}_{\mathbf{X},mv}^2$ should be kept as low as possible without over-constraining the model. In a pilot study, we found experimentally that a SD of 0.01 pixels gave contours in good agreement with the manually defined landmarks without over-constraining the model. This parameter depends on the image resolution, which in our data was 1.4 ± 0.18 pixels/mm (mean $\pm$SD). In a more advanced implementation, prior knowledge of inter- and intra-observer variability can easily be incorporated in the algorithm.

C. Integrating local segmentation algorithms with AAMs

The CAAM framework is not restricted only to manually defined landmarks, since segmentation algorithms based on local image evidence can be used to extract edge information with associated edge confidence. In its simplest form, a local edge detector finds the strongest edges $\mathbf{X}_0$ independently along normals of the AAM shape using for example a gradient criterion, and the confidences of all edges are assumed equal. The solution to the CAAM is found iteratively from (5), applying the edge detector after each iteration to obtain new edge positions. In echocardiography, local edge evidence may vary highly at different locations, and the contribution of the local edge detector to the segmentation should therefore reflect the confidence of the measured edge positions. This is accomplished through the covariance matrix $\mathbf{S}_{\mathbf{X}}^2$, causing the edge detector to dominate locally if edge evidence is strong, otherwise segmentation is controlled by the texture residuals. Due to speckle noise, a simple edge detector based upon independent point measurements will give highly noisy edge estimates in US images. We are therefore using DP to obtain more globally correct edge measurements.

DP constrained multiview AAMMs can be implemented using the following iterative scheme:

1) Use DP to find an optimum contour in each frame and view separately given an initial shape, resulting in a landmark vector $\mathbf{X}_0$ and edge energy $\mathbf{E}_0$ associated with each point along the contour.
2) Based on the energies, we express the confidence $\mathbf{S}_{\mathbf{X}}^{-1} = w_{DP}\mathbf{I}M(\mathbf{E}_0)$ at each point of $\mathbf{X}_0$ using a function $M()$ which maps energies to variances. Here, w_{DP} is an experimentally determined weight factor. The mapping function can for example be a constant, giving equal confidence in all points obtained from the DP algorithm, or a linear mapping to the range $[0,1]$.
3) The landmark vectors and diagonal covariance matrices from all frames and views are combined, and optimal parameter updates are obtained by solving (5).
4) Repeat from 1 until convergence.

In our experiments, we chose a linear mapping function. The weight $w_{DP} = 100$ was determined experimentally to give a good trade-off between global and local accuracy, by matching the energies from the DP algorithm to the contribution from the texture residuals.

D. Dynamic programming optimized active contours

After each iteration of the CAAM update algorithm, the detected contour $T_\mathbf{t}(\mathbf{x})$ is used as an initial contour for a dynamic programming based active contour. This initial contour can be written as a sequence of points $(\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_n)$ where each vector $\mathbf{i}_i$ represents the point's x and y coordinates. For each initial point, m candidate points $\{\mathbf{p}_i^1, \mathbf{p}_i^2, \ldots, \mathbf{p}_i^m\}$ are sampled with spacing d along the associated contour normal. Using dynamic programming, we determine the active contour $(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n)$, where each point $\mathbf{p}_i$ on the contour is selected from the corresponding candidate points to minimize the total contour energy expressed as

$$E_{tot}(\mathbf{p}_n) = w_{curv} \sum_{i=2}^{n} E_{curv}(\mathbf{p}_{i-1}, \mathbf{p}_i)$$
$$+ w_{dist} \sum_{i=1}^{n} E_{dist}(\mathbf{p}_i) + w_{grad} \sum_{i=1}^{n} E_{grad}(\mathbf{p}_i, l) \quad .$$

The first term is a smoothness constraint where

$$E_{curv}(\mathbf{p}_{i-1}, \mathbf{p}_i) = (D(\mathbf{p}_i, \mathbf{i}_i) - D(\mathbf{p}_{i-1}, \mathbf{i}_{i-1}))^2$$

is the change in distance from the initial contour, and $D()$ represents the Euclidean distance between two points. The second term

$$E_{dist}(\mathbf{p}_i) = e^{\left(\frac{-(i-n/2)}{s}\right)^2} |D(\mathbf{p}_i, \mathbf{i}_i)|$$

is a weighted distance from the initial contour, preventing the active contour from diverging too much from the AAM contour. The exponential weight ensures that the contour is close to the initial contour in a region of size s around the apex. $E_{grad}(\mathbf{p}_i, l)$ is the gradient at a candidate point along the contour normal averaged over a window of size l.

IV. CLINICAL EVALUATION

A. Data material

Apical triplane echocardiograms from 39 adult patients selected from the Department of Cardiology, Rikshospitalet-Radiumhospitalet University Hospital (old myocardial infarction, valve disease, dilated cardiomyopathy, pulmonary hypertension, heart transplant, and structurally normal hearts), were recorded using a Vivid 7 scanner (GE Vingmed Ultrasound,

Norway) equipped with a 2D phased array transducer (3V). The three imaging planes were aligned to the apical four-chamber (A4CH), apical two-chamber (A2CH), and the apical long axis (ALAX) views respectively (Fig. 1). Patients with arrhythmia were not included. Three data sets were excluded due to poor image quality and considerable out-of-plane motion during the cardiac cycle, making manual tracing impossible.

For the remaining 36 data sets, observer 1 (SU) traced the myocardial boundary manually in all three planes at end-diastole (ED) and end-systole (ES) using an online triplane volume tool (GE Vingmed Ultrasound, Norway). Volumes and EF were computed using a spline-based geometry model [3] and used as a manual reference. The analysis was repeated by a second independent observer (KL) for evaluation of inter-observer variability.

B. Model training

The clinical tool used by observer 1 and observer 2 did not permit storing of traced contours. For model training, the myocardial boundary was therefore traced manually in each view by a third observer (JH) and validated by observer 1. If observer 1 did not agree, the contour was re-traced accordingly. This was done for six frames, including ED, ES, two systolic frames, and two diastolic frames in each view, giving a total of 18 contours for triplane echocardiograms.

The timing of the ED and ES frames were determined manually, primarily by finding the aortic valve opening (AVO) and aortic valve closing (AVC), secondarily by visual identification of the largest and smallest LV volumes. The intermediate frames were determined by linear interpolation. Spatial coherence between corresponding points from different shapes was obtained by equidistant resampling to 40 points.

Before texture sampling, the width and height of all shapes were increased by 35% and 10% respectively, to ensure inclusion of signal from the myocardial tissue without producing shapes that were too large to fit within the imaged sector. The endocardial boundary was recovered from the segmented shapes by a corresponding shrink factor. For performance reasons, the images were resampled to 100×100 pixels. This was a limitation in our prototype AAM framework, and should not be regarded as a general limitation of the AAM formulation.

We used a leave-one-out approach to train 36 different models, where different data sets were left out from each model for evaluation. All models were generated retaining 96% shape variance and 90% texture variance. The same approach was used to train single plane AAMMs, but only the A4CH view was included.

C. Experiments

To compare the different segmentation algorithms, we performed segmentation using the following approaches:
1) Unconstrained multiview AAMM.
2) Multiview AAMM constrained by 18 manually defined landmarks (LM-CAAMM).
3) Multiview AAMM constrained by both manually defined landmarks and landmarks obtained by DP (DP-CAAMM).
4) Same as 3, but only including the A4CH view.

The manually defined landmarks were extracted from the contours traced by observer 3 to avoid contribution of initialization inaccuracies when comparing the methods. The landmarks included two points identifying the atrio-ventricular (AV) plane and one point identifying the LV apex in each view at ED and ES, giving six landmarks for each view. This procedure resulted in a total of 18 landmarks for triplane experiments and six landmarks for single plane experiments. All models were initialized to mean texture and shape, while the model

Table I
SUCCESS RATES AND DIFFERENCES BETWEEN SEMI-AUTOMATIC SEGMENTATION AND MANUAL REFERENCE (MEAN±SD)

Method	EDV [ml]	ESV [ml]	EF [%]	Landmark distance [mm]	Success rate [%]
Unconstrained AAMM	−7.3±20	−2.5±22†	−1.5±11†	4.3±3.5	83
LM-CAAMM	−3.3±24	−4.9±18	0.66±8.9†	2.7±1.9	100
DP-CAAMM	−3.1±20	0.61±13	−1.3±6.3	3.4±2.3	100
Single plane DP-CAAMM	−8.4±35†	1.6±23†	−4.2*±11†	2.8±2.0	100
Only DP	−18*±20	−9.6*±18	−1.3±7.4	0.80±2.3	100
Inter-observer	13*±19	9.9*±15	−1.7±6.3	-	100

* Bias significantly different from 0 ($p < 0.05$).
† SD significantly different from inter-observer variability ($p < 0.05$).

pose was initialized by first fitting a custom-made 2D spline model to the three landmarks of each ED view. This model was evaluated at 40 locations, and the initial AAMM pose was found by aligning the AAMM to the spline model using Procrustes analysis. To ensure comparable initialization for all methods, unconstrained multiview AAMMs were initialized using the same approach.

For each of the above experiments, leave-one-out models were evaluated on the left-out data sets. For triplane data sets, end-diastolic volume (EDV) and end-systolic volume (ESV) were computed using a spline-based geometry model [3] and compared to the manually defined standard established by observer 1. EF was computed as $EF = (EDV - ESV)/EDV \cdot 100\%$. For single plane data sets, EDV and ESV were computed directly from the detected contours using the area length formula $V = 8A^2/(3\pi L)$ where A is the area of the closed contour, and L is the distance from the AV plane to the apex. The accuracy of the single plane model versus the triplane model was assessed using the pairwise distance from the points on the detected A4CH contours to the training shapes averaged over six frames.

D. Statistical analysis

The bias (mean difference) of the algorithms was evaluated by comparing automatically determined EDV, ESV, and EF to the manual reference (two-tailed t-test assuming equal means), while the SD of each method was compared to the inter-observer variability (two-tailed F-test assuming equal SD).

To find the best method, the bias of each algorithm were compared in a pair-wise manner (two-tailed paired t-test on the mean assuming equal means and possibly unequal variances, and two-tailed paired F-test on the SD assuming equal SD). A *p*-value less than 0.05 was considered significant for all tests.

As a criterion for successful segmentation, we adopted the criterion of Bosch *et al.* [20], where an average point distance < 8mm (inter-observer variability + 3 SD) was regarded successful. This criterion was used to remove obvious convergence failure from further statistical analysis.

V. RESULTS

Experiments were carried out on triplane echocardiograms using unconstrained multiview AAMMs, landmark constrained multiview AAMMs (LM-CAAMM), and for multiview AAMMs constrained by both landmarks and DP (DP-CAAMM). We also evaluated the DP-CAAMM trained on single plane data showing the A4CH view for comparing triplane and single plane segmentation. The DP-CAAMM algorithm was chosen since it gave the best segmentation in both single plane and triplane data. Separate experiments were carried out using only DP for comparison with the DP-CAAMM.
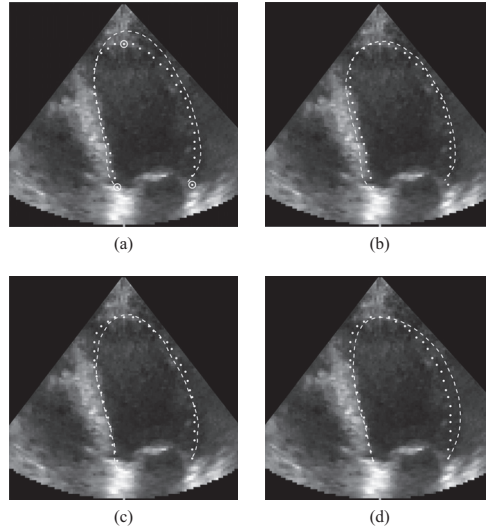
Figure 2.   Comparison of the different triplane algorithms (dashed lines) with manually traced contours (dotted lines.) Segmentation has been performed on six frames and three views simultaneously, and all images show the ED frame of the A4CH view. (a) Unconstrained multiview AAMM. White circles indicate manually defined landmarks. (b) Manually constrained multiview LM-CAAMM. (c) DP constrained multiview DP-CAAMM. (d) DP algorithm alone.

A. Constrained versus unconstrained AAMs

The unconstrained multiview AAMM had a success rate of 83% (30 of 36 patients) (Table I row 1). As illustrated in Fig. 2(a), the agreement between the manually defined landmarks and the unconstrained multiview AAMM were generally low. After exclusion of convergence failures, this distance was on average 4.3 ± 3.5 mm (Table I row 1). Compared to observer 1, the unconstrained multiview AAMM had no significant bias in EDV, ESV, or EF estimates (Table I row 1). However, the SD of both ESV and EF estimates were significantly higher than the inter-observer variability ($p = 0.03$ and $p < 0.01$ respectively) (Table I row 6). Fig. 4(f) indicates a relatively low correlation coefficient of automatically determined EFs compared to the manual reference, even if correlation coefficients of EDV and ESV were high (Fig. 4(a)).

The multiview LM-CAAMM offered a success rate of 100% (36 of 36 patients) (Table I row 2). As illustrated in Fig. 2(b), constraining the model improves the agreement with the manually traced contours, and the average distance between automatically and manually defined landmarks was reduced to 2.7 ± 1.9 mm (Table I row 2). There is no significant bias in EDV, ESV, or EF, while the SD in EF is significantly higher than the inter-observer variability ($p = 0.04$) (Table I row 6). Compared to the unconstrained AAMM, the LM-CAAMM has higher correlation coefficients for EF (Fig. 4(f) and 4(g)).

B. AAMs constrained by dynamic programming

The DP-CAAMM had a success rate of 100% (Table I row 3), and as illustrated by Fig. 2(b) and 2(c), the accuracy of the contour is generally higher than for the LM-CAAMM. There is no significant bias in EDV, ESV, or EF (Table I, row 3), and compared to the inter-observer variability (Table I row 6), the SD of the DP-CAAMM is not significantly different for EDV, ESV, or EF.

Compared to the LM-CAAMMs, the SD of EF measurements is significantly lower ($p = 0.04$). Further, the correlation coefficients for EDV, ESV, and EF are higher than for any
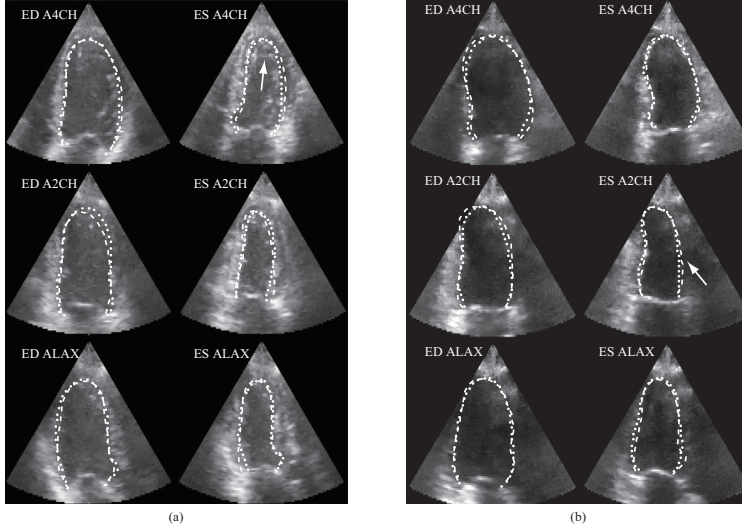
Figure 3. Examples showing typical results from automatic segmentation in triplane data sets using a DP-CAAMM (dashed line) compared to manual tracing (dotted line). The figure only shows the ED and ES frames, but segmentation was done in six frames and three views simultaneously. The algorithm handles typical clinical challenges such as foreshortening (indicated by white arrow in (a)) and dropout (indicated by white arrow in (b)) well.

of the other methods (Fig. 4(c) and 4(h)). The average distance between the automatically and manually defined landmarks was 3.4 ± 2.3 mm (Table I row 3), which is slightly higher than for the LM-CAAMM.

For evaluation of the DP-CAAMM, it was necessary to evaluate the performance of DP alone. The DP algorithm alone performed well in cases with strong edges, but the temporal behavior was generally poor, and dropouts were not handled well. The bias was significant for both EDV ($p < 0.01$) and ESV ($p < 0.01$) (Table I, row 5.) The bias and SD of the EF estimates were similar to DP-CAAMM, but DP-CAAMM has a slightly better correlation coefficient to the manual reference (0.93/0.94 for EDV and ESV respectively for DP versus 0.95/0.97 for DP-CAAMM.) Volumes estimated by the DP-CAAMM have significantly less bias for both ESV ($p < 0.01$) and EDV ($p < 0.01$) than DP alone (Table I, row 3 and 5).

Fig. 3(a) shows an example of a typical multiview and multiframe segmentation using the DP-CAAMM technique, where the model has been constrained by manually defined landmarks. The images suffer from foreshortening, especially in ES, but the model still performed well. Fig. 3(b) illustrates how the model behaves in cases with dropouts.

C. Single plane versus triplane

The EDV and ESV based on the single plane model have no significant bias compared to the manual reference (Table I, row 4), but there is a significant bias in EF ($p = 0.03$). In addition, the SD of the EDV, ESV, and EF estimates are significantly higher than for both the triplane algorithm and the inter-observer variability with $p < 0.01$ in all cases (Table I, row 4 and 6). The average point distance for single plane AAMMs was 3.7 ± 2.7 mm, while for the A4CH contour of the triplane AAMMs the distance was slightly lower, 3.6 ± 2.4 mm.

Figures 4(d) and 4(i) show low correlation coefficients between the single plane estimates and the manual reference for EDV, ESV, and EF.
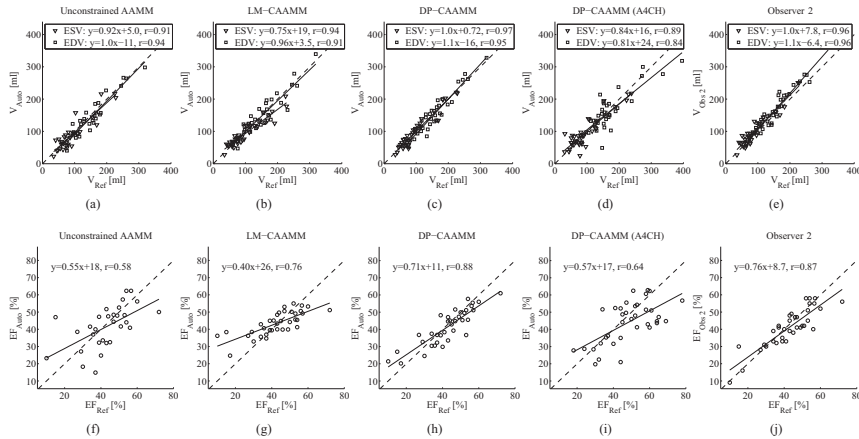
Figure 4. Comparison of automatically and manually determined volumes and EF. Solid lines represent the regression equations, and dotted lines indicate the unit line $y = x$. Results for the unconstrained multiview AAMM versus manual reference are shown in (a) and (f), the multiview LM-CAAMM is shown in (b) and (g), while the multiview DP-CAAMM is shown in (c) and (h). The results for the single plane DP-CAAMM are shown in (d) and (i). Figures (e) and (j) show observer 2 versus the reference.

VI. DISCUSSION

In this paper we introduced: (1) multiview AAMMs for improved volume and EF estimates in triplane echocardiographic data, (2) the use of CAAMs for further improved accuracy and robustness by constraining the model to manually defined landmarks, and (3) a method for integration of local edge detector based segmentation algorithms with the AAM framework, effectively combining the local segmentation power of DP based segmentation algorithms with the global properties of AAMs.

To our knowledge, this is the first report on semi-automatic segmentation of simultaneously acquired triplane echocardiograms. In addition, we have not seen any earlier reports where the CAAM-framework is used for cardiac segmentation. We believe that combining local segmentation methods with the AAM is a promising solution for segmentation of data with highly varying local image evidence.

A. Clinical applicability

We observed that the multiview LM-CAAMM was handling many of the challenges of cardiac US. Papillary muscles were successfully excluded according to recommendations[5], and the model was able to compensate for foreshortening. Information from multiple frames was successfully exploited to give improved temporal behavior.

In clinical practice, efficiency and minimum user interaction is of high priority. Previous work [20] has focused on efficiency by implementing fully automatic segmentation algorithms requiring minimum user interaction. Our impression is that more weight must be put on robustness and manual interaction by allowing the model to be guided by manually defined landmarks.

B. Triplane versus single plane segmentation

Triplane echocardiograms give higher accuracy of volume and EF estimates than single plane echocardiograms with similar image quality using the DP-CAAMM algorithm. Volumes estimated from three planes require fewer geometric assumptions about the LV

anatomy than the area-length formula. Further, the quadratic relationship between area and volume causes small errors in the single plane contour to give large errors in the volume estimates.

The average point distance was only slightly better for the A4CH contours extracted from the triplane segmentation than for the single plane segmentations. We speculate if this is a result of "under-training" due to a relatively small training set. By adding additional spatial deformation modes, a larger training set is required to span the higher dimensional shape and texture space. The "under-training" is however not necessarily a disadvantage, since a strong prior can give a more robust model in poor data. We also believe that the multiview AAMM would have given better results if the model's 2D pose in each view were derived from a common three-dimensional (3D) pose. This leads to fewer independent pose variables, allowing only physically possible combinations over the views. However, the LV is generally not aligned with the probe. Therefore, the asymmetry of the ventricular shape makes extraction of 3D pose from 2D views a non-trivial task. This issue requires further attention.

C. Constraining the active appearance model

Constraining the multiview AAMM improved segmentation robustness and accuracy, increasing the success rate from 83% to 100% compared to the unconstrained AAMM. Manually constrained multiview AAMMs also produced contours agreeing better with the manually defined landmarks. This was especially the case for the LV apex, where image noise and foreshortening caused poor segmentation using the unconstrained model. The model's ability to follow user defined landmarks is of high clinical importance. First, since the landmarks are selected by an expert they are usually good approximations of the true landmarks, and should be used as guidance for the model. Secondly, a segmentation method producing contours disagreeing with manually defined landmarks will gain little user confidence. Initializing the model manually is a time consuming task, but analysis time is not necessarily increased since the number of failures are reduced and accuracy is improved. Our framework is not limited only to manually defined landmarks, and fully automated segmentation can potentially be achieved using an automated landmark detection algorithm [24]. Also, the definition of ED and ES should be automated to reduce analysis time.

In our experiments, all models were constrained by landmarks extracted from the contours traced by observer 3. In a clinical setting, the accuracy of these landmarks is subject to inter- and intra-observer variability. It remains to be shown if knowledge of inter- and intra-observer variability can be utilized to improve segmentation in cases where manually defined landmarks are inaccurate.

As shown in Fig. 4(g), the manually constrained multiview AAMM has a tendency to estimate a "too-normal" EF despite the constraints. This is related to the "under-training" issue discussed in section VI-B.

D. Active appearance models constrained by DP

We have provided a general way of integrating segmentation algorithms based on local image evidence into the AAM framework. This framework accepts integration with a wide range of segmentation algorithms, including for example direct measurements of strongest edges, Kalman filter based algorithms [25], or even ASMs. We used the DP algorithm, which has been widely used for segmentation of echocardiograms.

Our study shows that the DP constrained multiview AAMM gives the best estimates of LV volumes and EF, and the method produces contours in good agreement with manually
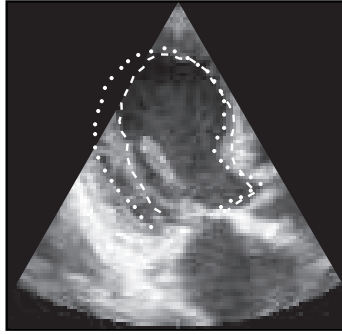
Figure 5.   An example of segmentation (dashed line) using the DP-CAAMM where the DP-algorithm is weighted too high, forcing the detected contour to follow a papillary muscle instead of the myocardial boundary (dotted line).

traced contours with respect to both global shape and local accuracy. Both volume and EF estimates are dramatically improved compared to the multiview AAMM only constrained by manually defined landmarks.

Our approach, with a DP algorithm fully integrated with a constrained AAM, combines the local segmentation accuracy of the DP algorithm with the global properties of the AAM. By weighting the contribution of the DP algorithm depending on local edge evidence, the DP algorithm dominates when wall evidence is strong; otherwise, deformation is controlled by the AAM. When integrating the DP algorithm with the AAM framework, a confidence measure must be associated with the DP generated contour. Improper confidence measures can give unwanted consequences as shown in Fig. 5, where the DP algorithm has too high weight causing the segmented contour to follow the papillary muscle instead of the myocardial boundary.

A pertinent question is if the combination of AAM and DP can be reduced to an ASM, which does not have an inherent texture model. In a pilot study, we found that the texture model was necessary for tracking the AV-plane and for the shape of the segmentation.

The idea of combining the AAM framework with other segmentation algorithms can be extended to include strain measurements from speckle tracking or doppler imaging. This can potentially give a more physiologically realistic LV model, which also describes torsion and stretching of the myocardial tissue. The framework can also be used in coupled AAMs for simultaneous segmentation of the epi- and endocardial boundaries. Further, the framework can easily be extended to higher dimensions for use in 3D echocardiography.

E. Comparison with related work

Oost *et al.* [19] describe hybrid AAM/DP segmentation of X-ray LV angiograms where a DP based active contour is initialized by the output from the AAM algorithm. They do not feed the output from the DP algorithm back into the AAM, and the resulting contours do not have to lie within the model shape space. They reported results from semi-automatic segmentation where the model was initialized by three user defined landmarks, but in contrast to our approach, the model was free to deform without being constrained by the landmarks.

In [20], Bosch *et al.* report a low error between manually and automatically determined area EF ($0.66 \pm 5.5\%$) using fully automatic AAMMs. We are not able to reproduce these results using unconstrained AAMMs. We speculate if the inferior image quality in triplane compared to standard 2D echocardiograms is the reason. In the future, improved image

quality in triplane echocardiography can make fully automatic segmentation techniques for this modality more feasible.

F. Limitations

We used a leave-one-out approach for evaluation of the segmentation methods due to a limited number of data sets. This technique can lead to over-fitting of parameters to the available data set, but since test data were never included in the training set, our results should give a good impression of how the algorithms perform.

The AAMMs were constrained by landmarks extracted from the training shapes traced by observer 3. These shapes were validated by observer 1, and the segmented volumes are expected to agree better with observer 1 than observer 2. Comparison with intra-observer variability would therefore have been a stricter criterion than the inter-observer evaluation used in our study.

Patients with arrhythmia were not included in our study, because this would require a much larger population of patients to model the pathological LV contraction patterns correctly. A similar argument applies to the exclusion of recordings with severe out-of-plane motion. This is an inherent challenge with AAMs, because care must be taken when assembling the training sets to ensure a balanced composition of a wide range of pathologies and imaging situations.

Triplane echocardiography has been shown to give accurate LV volume measurements [3], but MRI is still accepted as the gold standard, and would have provided a better ground truth for our evaluation. These limitations should be addressed in a larger scale clinical evaluation.

VII. CONCLUSION

We have developed a landmark and DP constrained multiview and multiframe active appearance model for segmentation of the left ventricle in simultaneously acquired triplane echocardiograms. Evaluation in 36 patients shows that (1) constraining the model to manually defined landmarks improves volume and EF estimates, (2) further improvement is achieved using the DP-algorithm to constrain the multiview AAMM, and (3) segmentation in triplane echocardiograms give superior results compared to single plane data.

REFERENCES

[1] N. B. Schiller, P. M. Shah, M. Crawford, A. DeMaria, R. Devereux, H. Feigenbaum, H. Gutgesell, N. Reichek, D. Sahn, and I. Schnittger, "Recommendations for quantitation of the left ventricle by two-dimensional echocardiography." *J. Amer. Soc. Echocardiogr.*, vol. 2, no. 5, pp. 358–367, Sep. 1989.

[2] J. S. Gottdiener, J. Bednarz, R. Devereux, J. Gardin, A. Klein, W. J. Manning, A. Morehead, D. Kitzman, J. Oh, M. Quinones, N. B. Schiller, J. H. Stein, and N. J. Weissman, "American society of echocardiography recommendations for use of echocardiography in clinical trials," *J. Amer. Soc. Echocardiogr.*, vol. 17, no. 10, pp. 1086–1119, Oct. 2004.

[3] S. Malm, S. Frigstad, E. Sagberg, P. A. Steen, and T. Skjarpe, "Real-time simultaneous triplane contrast echocardiography gives rapid, accurate and reproducible assessment of left ventricular volumes and ejection fraction: A comparison with magnetic resonance imaging," *J. Amer. Soc. Echocardiogr.*, vol. 19, no. 12, pp. 1494–1501, Dec. 2006.

[4] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[5] R. M. Lang, M. Bierig, R. B. Devereux, F. A. Flachskampf, E. Foster, P. A. Pellikka, M. H. Picard, M. J. Roman, J. Seward, J. S. Shanewise, S. D. Solomon, K. T. Spencer, M. S. J. Sutton, and W. J. Stewart, "Recommendations for chamber quantification." *J. Amer. Soc. Echocardiogr.*, vol. 18, no. 12, pp. 1440–1463, 2005.

[6] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 1, pp. 321–331, Jan. 1988.

[7] A. A. Amini, T. E. Weymouth, and R. C. Jain, "Using dynamic programming for solving variational problems in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 9, pp. 855–867, Sep. 1990.

[8] T. F. Cootes and C. J. Taylor, "Active shape models - 'smart snakes'," in *Proc. British Machine Vision Conf.* Springer-Verlag, 1992, pp. 266–275.

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[10] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, June 2001.

[11] ——, "Active appearance models," in *Proc. Eur. Conf. Computer Vision*, H. Burkhardt and B. Neumann, Eds., vol. 2. Springer-Verlag, 1998, pp. 484–498.

[12] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Fully automated endocardial contour detection in time sequences of echocardiograms by active appearance motion models," in *Proc. Computers in Cardiology*, vol. 28, Sep. 2001, pp. 93–96.

[13] B. P. F. Lelieveldt, M. Üzümcü, R. J. van der Geest, J. H. C. Reiber, and M. Sonka, "Multi-view active appearance models for consistent segmentation of multiple standard views: application to long- and short-axis cardiac MR images," *International Congress Series*, vol. 1256, pp. 1141–1146, 2003.

[14] C. R. Oost, B. P. F. Lelieveldt, M. Üzümcü, H. Lamb, J. H. C. Reiber, and M. Sonka, "Multi-view active appearance models: Application to X-ray LV angiography and cardiac MRI," in *Proc. Conf. Information Processing in Medical Imaging*, July 2003, pp. 234–245.

[15] M. Üzümcü, R. J. van der Geest, M. Sonka, H. J. Lamb, J. H. C. Reiber, and B. P. F. Lelieveldt, "Multiview active appearance models for simultaneous segmentation of cardiac 2- and 4-chamber long-axis magnetic resonance images," *Invest. Radiol.*, vol. 40, pp. 195–203, Apr. 2005.

[16] J. Hansegård, S. Urheim, E. Steen, H. Torp, B. Olstad, S. Malm, and S. I. Rabben, "Detection of the myocardial boundary in the left ventricle from simultaneously acquired triplane ultrasound images using multi view active appearance motion models," in *Proc. IEEE Ultrasonics Symposium*, vol. 4, 2005, pp. 2267– 2270.

[17] J. G. Bosch, "Automated contour detection in echocardiographic images," Ph.D. dissertation, Leiden University, 2006. [Online]. Available: http://hdl.handle.net/1887/4423

[18] T. F. Cootes and C. J. Taylor, "Constrained active appearance models," in *Proc. IEEE Int. Conf. Computer Vision*, vol. 1, July 2001, pp. 748–754.

[19] E. Oost, G. Koning, M. Sonka, P. V. Oemrawsingh, J. H. C. Reiber, and B. P. F. Lelieveldt, "Automated contour detection in X-ray left ventricular angiograms using multiview active appearance models and dynamic programming," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1158–1171, Sep. 2006.

[20] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance motion models," *IEEE Trans. Med. Imag.*, vol. 21, no. 11, pp. 1374–1383, Nov. 2002.

[21] T. F. Cootes, "Statistical models of appearance for computer vision," The University of Manchester, Tech. Rep., 2001. [Online]. Available: http://www.isbe.man.ac.uk/~bim/Models/app_models.pdf

[22] C. Goodall, "Procrustes methods in the statistical analysis of shape," *J. Royal Stat. Soc. B*, vol. 53, no. 2, pp. 285–339, 1991.

[23] I. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135–164, Nov. 2004.

[24] A. Torp, S. Rabben, A. Stoylen, H. Ihlen, K. Andersen, L.-A. Brodin, and J. Olstod, "Automatic detection and tracking of left ventricular landmarks in echocardiography," in *IEEE Ultrasonics Symposium*, vol. 1, 2004, pp. 474–477.

[25] D. Comaniciu, X. S. Zhou, and S. Krishnan, "Robust real-time myocardial border tracking for echocardiography: An information fusion approach," *IEEE Trans. Med. Imag.*, vol. 23, no. 7, pp. 849–860, 2004.

Paper D

Semi-automated quantification of left ventricular volumes and ejection fraction by real-time three-dimensional echocardiography

J. Hansegård, S. Urheim, K. Lunde, S. Malm and S. I. Rabben

Semi-automated quantification of left ventricular volumes and ejection fraction by real-time three-dimensional echocardiography

Jøger Hansegård[1], Stig Urheim[2], Ketil Lunde[2], Siri Malm[3], and Stein Inge Rabben[4]

[1] Dept. of Informatics, University of Oslo, Norway
[2] GE Vingmed Ultrasound
[3] Dept. of Cardiology, Rikshospitalet University Hospital , Oslo, Norway
[4] Harstad University Hospital, Harstad, Norway

Abstract. Recent studies have shown that real-time three-dimensional (3D) echocardiography (RT3DE) gives more accurate and reproducible left ventricular (LV) volume and ejection fraction (EF) measurements than traditional two-dimensional methods. We have developed a new semi-automated tool (4DLVQ) for volume measurements in RT3DE, and we sought to evaluate the accuracy and repeatability of this method.
LV end-diastolic volumes (EDV), end-systolic volumes (ESV), and EF measured using 4DLVQ were compared with a commercially available semi-automated analysis tool (TomTec 4D LV-Analysis ver. 2.2) in 35 patients. Repeated measurements were performed to investigate inter- and intra-observer variability.
Average analysis time of the new tool was 141s, significantly shorter than 261s using TomTec ($p < 0.001$). Bland Altman analysis revealed high agreement of measured EDV, ESV, and EF compared to TomTec ($p = NS$), with bias and 95% limits of agreement of 2.1 ± 21 ml, -0.88 ± 17 ml, and $1.6 \pm 11\%$ for EDV, ESV, and EF respectively. Intra-observer variability of 4DLVQ vs. TomTec was 7.5 ± 6.2 ml vs. 7.7 ± 7.3 ml for EDV, 5.5 ± 5.6 ml vs. 5.0 ± 5.9 ml for ESV, and $3.0 \pm 2.7\%$ vs. $2.1 \pm 2.0\%$ for EF ($p = NS$). The inter-observer variability of 4DLVQ vs. TomTec was 9.0 ± 5.9 ml vs. 17 ± 6.3 for EDV ($p < 0.05$), 5.0 ± 3.6 ml vs. 12 ± 7.7 for ESV ($p < 0.05$), and $2.7 \pm 2.8\%$ vs. 3.0 ± 2.1 for EF ($p = NS$).
In conclusion, the new analysis tool gives rapid and reproducible measurements of LV volumes and EF, with good agreement compared to another RT3DE volume quantification tool.

Introduction

Left ventricular (LV) volumes and ejection fraction (EF) are important parameters for diagnosis and prognosis of patients with heart disease (Mandinov et al, 2000; Taylor et al, 1980; White et al, 1987). Traditionally, LV volumes are measured by manual tracing in two sequentially acquired two-dimensional (2D) echocardiograms, using the biplane method of disks (MOD) (Gottdiener et al, 2004; Schiller et al, 1989). The spatial under-sampling of the ventricle, inherent with such 2D techniques, requires geometric assumptions about the LV shape. Foreshortening, occurring when the image plane is oblique to the ventricular main axis, also introduces errors in MOD measurements in 2D echocardiography (Gutiérrez-Chico et al, 2005; Jenkins et al, 2004).

Real-time three-dimensional (3D) echocardiography (RT3DE) (also known as four-dimensional (4D) echocardiography) is gaining popularity as a routine clinical tool (Monaghan, 2006), and has a significant potential of improving clinical decision-making (Hare et al, 2008). Of particular interest is the improved accuracy and repeatability of volume and EF measurements, compared to conventional 2D techniques (Gopal et al, 1995; Gutiérrez-Chico et al, 2005; Sugeng et al, 2006; Jacobs et al, 2006). However, manual analysis of 3D data is time-consuming and impractical. Thus, clinical use of volume measurements from RT3DE requires simple and efficient automated analysis tools.

Currently, two commercially available volume measurement tools for RT3DE exist on the market; QLAB (Philips, Andover, MA, USA), and TomTec's 4D LV-Analysis software (TomTec Imaging Systems, Unterschleissheim, Germany). Different versions of the TomTec

tool have been verified against cardiac magnetic resonance imaging (cMRI) in several studies (Chan et al, 2006; Jenkins et al, 2004; Kühl et al, 2004; Soliman et al, 2007), showing excellent agreement of measured LV volumes and EF.

One of the challenges with the TomTec analysis tool is that it requires manual tracing of the endocardial border in three apical planes for initialization and manual correction of the detected surface (Chukwu et al, 2007). The tool is therefore relatively difficult and time-consuming to use, especially for non-expert users, since the exact position of the endocardial border can be difficult to determine in many situations.

We present a new semi-automated tool for 4D LV volume quantification (4DLVQ) in RT3DE. This tool provides a simple user interface and an efficient workflow by eliminating the need for manual tracing, making the tool simpler to use for non-expert users.

The objective of this study was to evaluate the agreement of LV volumes and EF measured by 4DLVQ compared to TomTec, to evaluate the repeatability of these parameters, and to determine the potential of 4DLVQ as a clinical tool.

Materials and methods

Volume quantification tool

A volume quantification tool was developed for rapid semi-automated detection of the LV endocardial border in RT3DE, based upon a 3D energy minimizing deformable model (McInerney and Terzopoulos, 1996). The deformable model is evolved using an iterative deformation scheme, under the influence of internal and image derived forces, temporal forces, and user defined landmark forces.

The model's internal forces ensure second order shape continuity of the detected object by counteracting stretching and bending of the surface. In clinical data, a large variation of LV shapes is expected due to different pathologies and individual variations between patients. We have therefore avoided encoding explicit a priori information about the LV shape into the shape constraints.

Image forces are derived from the volumetric data using a local edge detector, utilizing a combination of gradient and transition information (Rabben et al, 2000). These image forces pull the surface towards image edges within a region around the deformable model. To ensure consistent surface detection from frame to frame, and to give a smooth time-volume curve, temporal forces constrain the model to second-order temporal continuity. By disabling temporal forces, the deformable model can be used for surface detection in single frames. This mode was used for initial surface detection at end-diastole (ED) and end-systole (ES).

User input is facilitated by generating spring-like forces that pull the deformable model towards the spatial location of user-defined landmarks (Kass et al, 1988).

The relative weights of the image and model forces were tuned manually on more than 100 training data sets to obtain a good trade-off between accuracy and robustness in different imaging situations and for various pathologies.

After completed surface detection, LV volumes were derived from the triangulated surfaces by summation of all triangular patches using the divergence theorem (Goldman, 2004).

Patient selection

Evaluation was done in two data sets from previous studies with a total of 56 patients (77% men, age 23-76) submitted to echocardiography at St. Olavs hospital, Trondheim, Norway, ((Malm et al, 2006), $n = 20$, exclusion criteria: atrial fibrillation, severe extra-cardiac disease),

and Rikshospitalet-Radiumhospitalet University Hospital, Oslo, Norway, ((Hansegård et al, 2007b), *n* = 36, exclusion criteria: arrhythmia). All patients gave informed written consent to participation, and the studies conformed to the declaration of Helsinki, with approval from the regional committees of medical ethics.

Apical volumetric imaging was performed by experienced operators using a Vivid 7 scanner (GE Vingmed Ultrasound, Horten, Norway) and a 3D transducer (3V). Electrocardiogram (ECG) gated sub-volumes were acquired from four consecutive cardiac cycles during breathhold, with the participants in the left lateral recumbent position. The sub-volumes were automatically stitched to a sequence of full 3D volumes covering the entire LV, and stored digitally for analysis.

In 20 (36%) of the 56 patients, less than 70% of the myocardium was visualized, caused by a too narrow imaging sector, shadows, or dropouts. Stitching artefacts were found in one additional patient (2%). These factors would prohibit reliable volume measurements (Tighe et al, 2007), and the total of 21 patients (38%), were therefore excluded from further analysis. Of note in this study, was that patients were not pre-selected from 2D or 3D image quality.

Analysis

Analysis was performed by an expert operator using a customized EchoPAC workstation (GE Vingmed Ultrasound, Horten, Norway) with the 4DLVQ software integrated. When entering the tool, the user was presented with a quad-screen, showing cine loops of three apical views with $60°$ inter-plane spacing, and one short axis (SAX) view. If required, the apical views were manually corrected to show the standard apical four-chamber (A4CH), apical two-chamber (A2CH), and apical long axis (ALAX) views, thereby eliminating foreshortening. When this anatomical alignment step was complete, the ED frame was automatically detected from the ECG and, if necessary, manually corrected by visually determining the maximum cavity area.

While displaying the ED frame, surface detection was initialized by manually selecting two points identifying the mitral annulus and one point identifying the LV apex in each of the three apical views shown in Fig. 1(a). After the total of nine landmarks had been defined at ED, non-temporal 3D surface detection was immediately performed to extract the endocardial border and to compute the EDV. The time required for a full 3D surface detection was less than one second in all cases. Cross-sections of the detected 3D surface were displayed in three apical views and three SAX views distributed between the LV apex and base, as shown in Fig. 1(a), to allow visual verification of the detected surface. A fourth user controlled SAX plane was used to further inspect the surface detection result. If necessary, the displayed LV surface could, at this time, be edited by manually adding new landmark points using the trackball. An interactive workflow was achieved by automatically repeating 3D surface detection after each new landmark until a satisfactory result was obtained. An overview of the tool workflow is given in Fig. 2(a).

The above procedure was repeated for the ES frame. End systole was detected automatically from the ECG and corrected manually, primarily to agree with aortic valve closing, secondarily by visually determining the minimal cavity area.

When surface detection was complete for both ED and ES, preliminary measurements of EDV, ESV, and EF were presented to the user. Our experience was, however, that ED and ES could be more robustly determined from a full time-volume curve. Full 4D surface detection was therefore performed to detect LV surfaces for each frame in the entire cardiac cycle. This typically required less than 10 seconds, depending on the frame- and heart rate. This allowed a full time-volume curve to be computed, and the maximum and minimum volumes were presented on the screen as EDV and ESV respectively along with the derived EF, as shown in

(a)

(b)
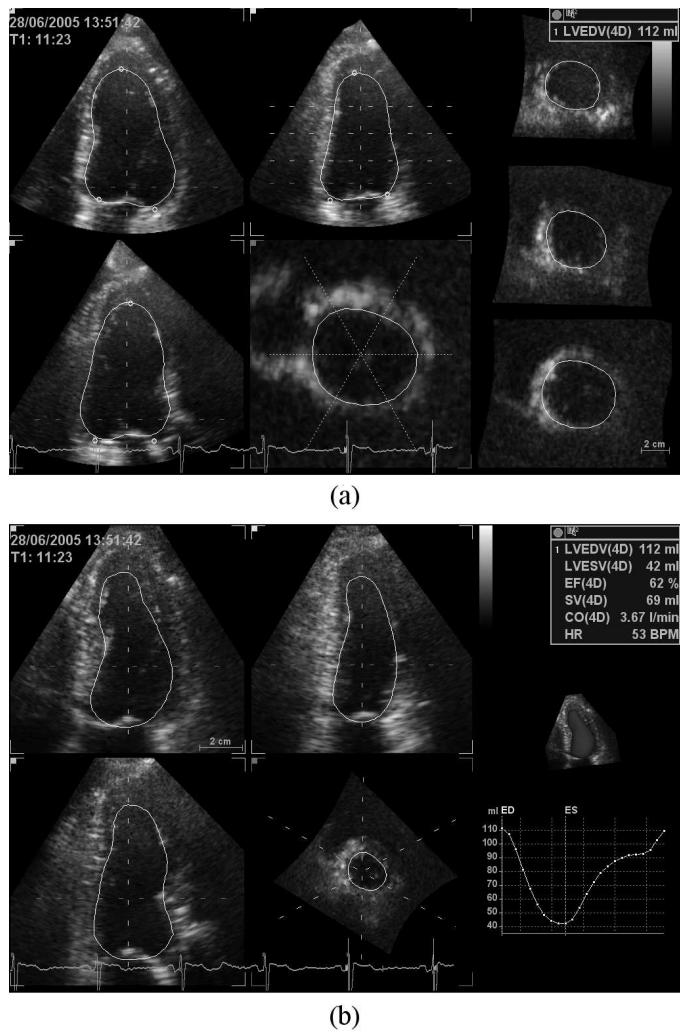
Fig. 1. An example of LV surface detection at ED using 4DLVQ is shown in (a). The standard apical views (top left, top middle, bottom left) were obtained by manual alignment. Each small circle in the apical views indicates a manually defined landmark used for initialization of surface detection. One SAX view (middle, bottom) was dynamically updated to reflect the trackball position, in order to facilitate precise landmark positioning and verification of the 3D surface detection. If necessary, additional landmarks could be added to leave papillary muscles within measured cavity volume. Three extra SAX views distributed between apex and base (right) were used to further verify the detected surface. The complete 4D surface detection at ES with time-volume curve is shown in (b).
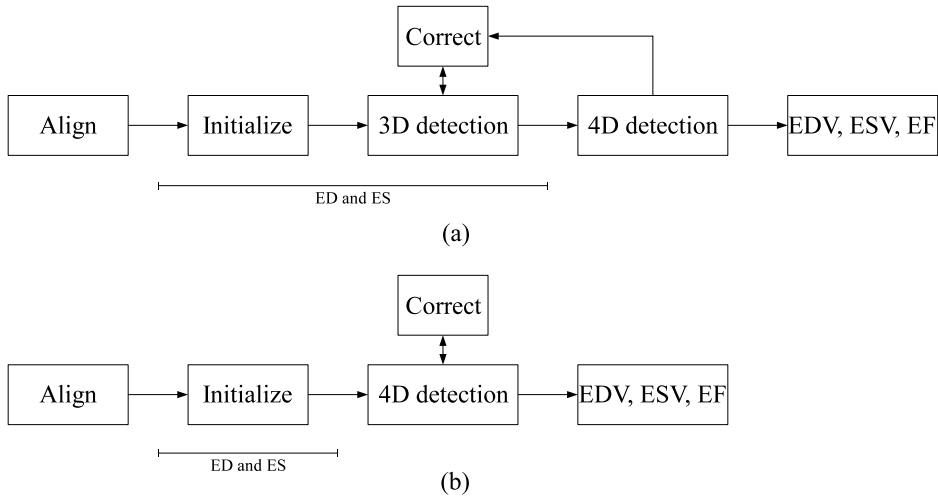
Fig. 2. Comparison of workflow for 4DLVQ (a) and TomTec (b). With 4DLVQ, each view is aligned to the standard apical view. Initialization is done at ED using 3 clicks in each apical view. 3D surface detection is automatically triggered, and if necessary the detected 3D surface is edited manually by adding landmarks. This is an interactive procedure where the 3D surface detection is repeated after each new landmark. Once completed with ED, the same procedure is repeated for ES. When both ED and ES are finished, a full 4D surface detection is performed to obtain EDV, ESV, and EF along with a time-volume curve. At this stage it is possible to edit the surfaces if improvements are required. With TomTec, initialization is done by accurate tracing at both ED and ES in each view after alignment to the standard apical views. 4D surface detection is done directly after initialization, and correction is done to the resulting surfaces if required.

Fig. 1(b). If required, the LV surfaces computed during this step could also be edited further. In these cases, non-temporal 3D surface detection was performed after each additional manually defined landmark as for ED and ES. A full 4D surface detection was then manually triggered once again, to update the time-volume curve. When a satisfactory result was obtained, EDV, ESV, and EF was recorded for analysis.

Separate analysis was done using an EchoPAC workstation with a TomTec 4D LV-Analysis plug-in ver. 2.2. An overview of the TomTec workflow is given in Fig. 2(b). The apical views were aligned to the standard A4CH, A2CH, and ALAX views, similarly as for the 4DLVQ tool, to eliminate foreshortening. Initialization of the TomTec surface detection algorithm was done by manually tracing the endocardial border using a spline-based annotation tool in the three apical views. Contrary to 4DLVQ, manual tracing was completed at both ED and ES in each apical view before continuing to the next view. Manual editing of the LV surface after surface detection could be avoided by tracing the endocardial border as accurately as possible during initialization (Chukwu et al, 2007). While tracing, a cine-loop of the corresponding view displayed the traced endocardial border to ensure temporal consistency between ED and ES. After tracing in the A4CH, A2CH, and the ALAX views, 4D LV surface detection was performed. Two manually adjustable apical planes were used to validate the automatically detected surfaces. The tool allowed for manually editing the LV surfaces, but due to the accurate initialization procedure, this was not used. EDV, ESV, and EF measurements were derived from the automatically detected LV surfaces, and recorded for analysis. These values were used as reference values for evaluation of 4DLVQ.

For both methods, the analysis time was measured from the start of analysis of volumetric data until the volume and EF measurements were displayed on the screen. The analysis time was reported as average time $\pm$ standard deviation (SD).

Table 1. Population mean and range of EDV, ESV, and EF measured by 4DLVQ and TomTec.

	EDV [ml]	ESV [ml]	EF [%]
4DLVQ	137 (88-243)	76 (39-180)	47 (13-65)
TomTec	135 (86-273)	77 (41-182)	45 (12-63)

Repeatability

Intra-observer variability for 4DLVQ and TomTec was assessed in all 35 patients by the primary operator after at least 14 days. A second expert operator assessed inter-observer variability in 10 randomly selected patients. Both operators were blinded to previous measurements.

To compare repeatability with other studies, we adopted the method of Jacobs et al (2006) and Sugeng et al (2006). First, the absolute difference between two repeated measurements was computed for each patient as shown in Eq. (1). The absolute differences were reported as mean $\pm$ SD over all patients. Secondly, % variability was defined as the absolute difference between two single measurements, normalized by the average of the two measurements in the same patient as shown in Eq. (2). The % variability was reported as mean $\pm$ SD over all patients.

$$\text{Abs. variability} = |\text{meas. 2} - \text{meas. 1}| \tag{1}$$

$$\text{\% variability} = \frac{|\text{meas. 2} - \text{meas. 1}|}{\text{avg of meas. 1 and 2}} \cdot 100\% \tag{2}$$

Statistical analysis

The relationship between 4DLVQ and the TomTec reference values was analyzed by linear regression and Bland Altman analysis (Bland and Altman, 1986). The latter was used to evaluate the agreement between the two methods (two-tailed t-test on the differences with a null hypothesis of zero difference and $p < 0.05$ regarded as significant). The agreement between the two methods was reported as the mean difference (bias) and the corresponding 95% limits of agreement. For comparison with other studies, Pearson's correlation coefficient between the two methods was also reported.

Agreement between inter- and intra-observer variability of the two methods was computed from the absolute differences between repeated measurements (Mann-Whitney U test with $p < 0.05$ regarded as a significant).

Analysis time was compared using a Mann-Whitney U test with $p < 0.05$ regarded as significant.

Results

Semi-automated analysis was feasible in all 35 data sets. A full 4D analysis using 4DLVQ required 141 ± 37 s including image alignment, initialization, and manual correction of the detected surfaces. This was significantly quicker than for TomTec ($p < 0.001$), which required 261 ± 63 s. The maximum time required for the two tools was 266 and 392 seconds respectively. The 4DLVQ tool required on average three additional manually defined landmarks to correct the initially detected surfaces.

TomTec and 4DLVQ yielded similar results for population mean and range for EDV, ESV, and EF (Table 1). Fig. 3(a-c) shows the agreement between 4DLVQ and TomTec. The mean
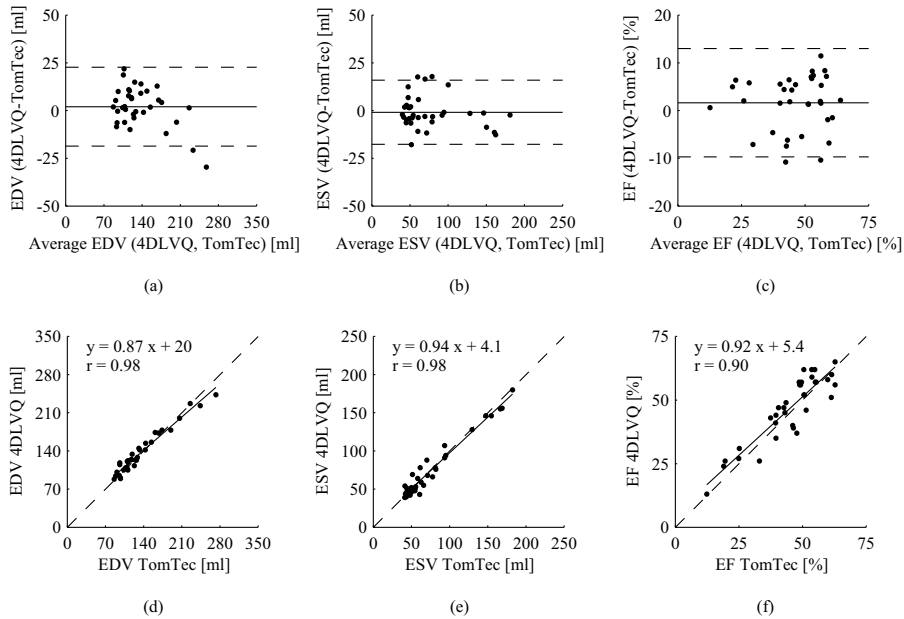
Fig. 3. Bland Altman analysis of EDV, ESV, and EF measured by 4DLVQ compared to TomTec is shown in (a), (b), and (c) respectively. Average difference (solid) is shown along with 95% limits of agreement (dashed). EDV, ESV, and EF measured by 4DLVQ is plotted against TomTec in (d), (e), and (f) along with unit line (dashed), regression line (solid), and Pearson's correlation coefficient r.

differences and 95% limits of agreement for EDV were 2.1 ± 21 ml, -0.88 ± 17 ml for ESV, and $1.6 \pm 11\%$ for EF. The differences between the two tools were not statistically significant for EDV, ESV, or EF. The small bias and narrow 95% limits of agreements are strong indications that 4DLVQ gives results that are highly comparable with TomTec.

Fig. 3(d-f) shows the relationship between EDV, ESV, and EF measured by 4DLVQ and TomTec. As expected when comparing measurements which both are subject to some measurement error, linear regression gave slopes that were less than unity for all parameters (0.87, 0.94, and 0.92 for EDV, ESV, and EF respectively). The zero intercepts were 20 ml for EDV, 4.1 ml for ESV, and 5.4 % for EF. The correlation coefficients of parameters measured by 4DLVQ and TomTec were 0.98, 0.98, and 0.90 for EDV, ESV, and EF respectively.

The results from the inter- and intra-observer analysis of 4DLVQ and TomTec are shown in Table 2. This analysis reveals low intra-observer variability for both methods. An important observation is that the inter-observer variability was significantly better with 4DLVQ than with TomTec for both EDV and ESV.

Discussion

We have evaluated a new semi-automated method for rapid quantification of LV volumes and EF in volumetric echocardiograms, and compared this to a reference volume quantification tool (TomTec) in 35 patients.

Conventional 2D methods for assessment of LV volumes and EF, such as the biplane MOD have several well-known shortcomings. First, the sparse sampling of the LV requires geometric modeling with inherent assumptions of the ventricular shape. Due to between-subject and

Table 2. Inter- and intra-observer variability for TomTec and 4DLVQ.

	Inter-observer		Intra-observer	
	Abs.	%	Abs.	%
EDV				
4DLVQ [ml]	9.0 ± 5.9	5.9 ± 3.7	7.5 ± 6.2	5.8 ± 4.5
TomTec [ml]	$17^* \pm 6.3$	11 ± 4.5	7.7 ± 7.3	5.5 ± 5.2
ESV				
4DLVQ [ml]	5.0 ± 3.6	5.6 ± 4.0	5.5 ± 5.6	8.0 ± 7.6
TomTec [ml]	$12^* \pm 7.7$	15 ± 8.8	5.0 ± 5.9	6.0 ± 7.3
EF				
4DLVQ [%]	2.7 ± 2.8	5.9 ± 5.8	3.0 ± 2.7	7.1 ± 6.5
TomTec [%]	3.0 ± 2.1	8.5 ± 6.5	2.1 ± 2.0	4.9 ± 4.4

*Significantly different from 4DLVQ ($p < 0.05$). EDV, end-diastolic volume; ESV, end-systolic volume; EF, ejection fraction. Abs. values are population mean $\pm$ SD of absolute differences between repeated measurements. % values are population mean $\pm$ SD of absolute differences of repeated measurements normalized by the average of the two repeated measurements.

pathological variations, these models can only serve as approximations of the true ventricular shape. Second, the 2D methods can be highly affected by foreshortening because it can be difficult to acquire 2D echocardiograms that are properly aligned with the LV main axis.

In contrast to 2D methods, our 3D deformable model-based volume quantification method eliminated the need for geometric assumptions of the LV shape. Since the displayed views were manually aligned to the true anatomical LV main axis, foreshortening was also completely avoided. These factors make automated 4D quantification better suited for accurate and reproducible measurements of ventricular volumes and EF than manual 2D methods. Also, an automated method can provide time-volume curves from a full cardiac cycle, giving more accurate EF estimates. The time-volume curve can potentially also be used to improve echocardiographic diagnosis, by providing information about timing of cardiac events, and filling rates in diastolic function analysis.

Analysis time was significantly shorter with 4DLVQ than with TomTec. This difference was mainly due to the simplified 4DLVQ initialization procedure, requiring only nine easily located landmarks at ED and ES, instead of triplane tracing used for the TomTec analysis. Because of this simplified initialization, a few additional manually defined landmarks were often required to include papillary muscles in the LV volume, but since this was done interactively with immediate visual feedback, the overhead was minimal. One might argue that the TomTec analysis time could be reduced by a less accurate initialization, but it has been shown that accurate initialization is required to avoid time-consuming manual editing of the detected LV surface (Chukwu et al, 2007). Soliman et al (2007) reported a TomTec analysis time of 360 ± 120 seconds, confirming that our results are representable with respect to TomTec analysis time.

We have shown that EDV, ESV, and EF assessed by 4DLVQ compares well to measurements performed by TomTec, with small bias, narrow 95% limits of agreement, and high repeatability. Since the agreement between the two methods was high for EDV, ESV, and EF, we conclude that that 4DLVQ performs at least as well as TomTec in clinically realistic data, even with a lower analysis time.

Several studies have presented repeatability assessment for various versions of TomTec. Soliman et al. (2007) reported inter-observer variability of 6.4 ± 7.8 ml, 7.8 ± 9.7 ml, and $7.1 \pm 6.9\%$, and intra-observer variability of 4.7 ± 3.2 ml, 6.1 ± 5.8 ml, and $6.6 \pm 7.4\%$ for EDV, ESV, and EF respectively. Our results demonstrated similar intra-observer variability,

but higher inter-observer variability with TomTec. We speculate if the discrepancy in inter-observer variability can be explained by the manual initialization procedure provided by TomTec. Slight differences in tracing conventions caused a bias between the observers of approximately 9 ml at ED and ES. This bias corresponds to a systematic tracing error of less than 1 mm for typical chamber sizes, but the impact on the variability parameters used in this study is evident. These differences in tracing conventions are not evident with 4DLVQ, which require less user input during initialization.

Differences between the workflows provided by TomTec and 4DLVQ are illustrated in Fig. 2. With 4DLVQ, initialization of all views is completed at ED before continuing to ES. TomTec uses a different strategy, where manual initialization is completed at both ED and ES in each view before proceeding with the next view. The traced contours are shown in a cine-loop preview display during tracing, providing additional information to ensure consistent contours between ED and ES. It has been claimed that manual editing in TomTec only has local impact on the detected surfaces (Soliman et al, 2007). We experienced in several cases that manually editing the surface caused it to "slip" from the endocardial border outside of the edited area, even in cases with strong edge evidence. Also, TomTec does not provide immediately updated surface detection during editing, whereas 4DLVQ provided immediate feedback, giving better control over the detected surface.

The clinical feasibility of RT3DE relies on simple and efficient analysis tools, ideally integrated as a part of the scanner software to facilitate on-line analysis during examinations. This puts strong constraints on the performance and ease of use of the tool, also with respect to manual correction of automatically detected surfaces. The presented volume quantification tool was implemented as an off-line analysis tool for use on the EchoPAC workstation. We have shown that a full 4D analysis can be done in less than 3 minutes, also in patients where manual correction was needed. This indicates that the tool is well suited for on-line analysis.

All in all, 4DLVQ seems to be a reliable clinical tool, which provides rapid and repro-ducible measurements of LV volumes and EF with good agreement compared to TomTec. It has a simple workflow that makes it easy to use for non-expert users. Recent development within the field of automated landmark detection in RT3DE (van Stralen et al, 2008) may be utilized to completely eliminate the need for manual initialization. Also, promising results have been presented using fully automated real-time 4D surface detection methods (Hansegård et al, 2007a; Orderud et al, 2007). In patients with poor acoustic properties, semi-automated methods are still preferred, since they allow for manual correction of the auto-matically determined LV volumes. But in the future, fully automated methods will improve efficiency and repeatability of echocardiography examinations.

It has been shown that the accuracy of LV volumes and EF is highly correlated to the amount of myocardium that is visualized in the RT3DE recording (Tighe et al, 2007). We de-fined an image quality threshold of 70% myocardium visibility, causing exclusion of 36% of the patients. Of note for this study was that patients were not pre-selected for image quality. The need for combining sub-volumes from four cardiac cycles caused exclusion of one pa-tient due to inability to hold breath throughout the acquisition. Future improvements to probe design and front-end processing capabilities are expected to give increased field of view, and less need for ECG gating, while improving image contrast and signal to noise ratio. These factors will improve feasibility and accuracy of automated assessment of LV function.

Accurate EF measurements are of high clinical importance. cMRI is currently accepted as the gold standard for LV quantification, and several studies have shown that TomTec compares well with cMRI, giving good agreement in measured EF, but with slightly under estimated volumes (Chan et al, 2006; Jenkins et al, 2004; Kühl et al, 2004; Soliman et al, 2007). This

bias is explained by differences in how the two modalities visualize trabeculae and valves, and also partial volume effects in cMRI (Barbier et al, 2007). Since 4DLVQ provides volume and EF measurements that agree well with TomTec, it is reasonable to believe that 4DLVQ will give similar results in a comparison with cMRI. The next natural step is therefore to compare 4DLVQ against cMRI.

Conclusion

We have presented a new volume quantification tool for automated EDV, ESV, and EF measurements in volumetric echocardiograms. The tool compared well to a commercially available analysis tool (TomTec 4D LV-Analysis), with high repeatability, and a significantly shorter analysis time. This is an important step towards wide spread use of RT3DE in clinical routine.

References

Barbier C, Johansson L, Lind L, Ahlström H, Bjerner T. The exactness of left ventricular segmentation in cine magnetic resonance imaging and its impact on systolic function values. Acta Radiol. 2007;48(3):285–291.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307–310.

Chan J, Jenkins C, Khafagi F, Du L, Marwick TH. What is the optimal clinical technique for measurement of left ventricular volume after myocardial infarction: A comparative study of 3-dimensional echocardiography, single photon emission computed tomography, and cardiac magnetic resonance imaging. J Am Soc Echocardiogr. 2006;19(2):192–201.

Chukwu EO, Katz AS, Toole RS, Schapiro W, Reichek N, Gopal AS. Prerequisites for avoiding manual boundary editing using semi-automated left ventricular quantitation by realtime three-dimensional echocardiography. In: J Am Coll Cardiol, 56th annual scientific session. 2007.

Goldman RN. Area of planar polygons and volume of polyhedra. In: Arvo J (Ed.), Graphics Gems II. Academic Press. 2004.

Gopal AS, Shen Z, Sapin PM, Keller AM, Schnellbaecher MJ, Leibowitz DW, Akinboboye OO, Rodney RA, Blood DK, King DL. Assessment of cardiac function by three-dimensional echocardiography compared with conventional noninvasive methods. Circulation. 1995;92(4):842–853.

Gottdiener JS, Bednarz J, Devereux R, Gardin J, Klein A, Manning WJ, Morehead A, Kitzman D, Oh J, Quinones M, Schiller NB, Stein JH, Weissman NJ. American society of echocardiography recommendations for use of echocardiography in clinical trials. J Am Soc Echocardiogr. 2004;17(10):1086–1119.

Gutiérrez-Chico J, Zamorano JL, Isla LPD, Orejas M, Almeíra C, Rodrigo JL, Ferreirós J, Serra V, Macaya C. Comparison of left ventricular volumes and ejection fractions measured by three-dimensional echocardiography versus two-dimensional echocardiography and cardiac magnetic resonance in patients with various cardiomyopathies. Am J Cardiol. 2005;95(6):809–813.

Hansegård J, Orderud F, Rabben SI. Real-time active shape models for segmentation of 3D cardiac ultrasound. Lect Notes Comput Sci. 2007a;4673:157–164.

Hansegård J, Urheim S, Lunde K, Rabben SI. Constrained active appearance models for segmentation of triplane echocardiograms. IEEE Trans Med Imaging. 2007b;26(10):1391–1400.

Hare JL, Jenkins C, Nakatani S, Ogawa A, Yu C, Marwick TH. Feasibility and clinical decision-making with 3D echocardiography in routine practice. Heart. 2008;94(4):440–445.

Jacobs LD, Salgo IS, Goonewardena S, Weinert L, Coon P, Bardo D, Gérard O, Allain P, Zamorano JL, de Isla LP, Mor-Avi V, Lang RM. Rapid online quantification of left ventricular volume from real-time three-dimensional echocardiographic data. Eur Heart J. 2006;27(4):460–468.

Jenkins C, Bricknell K, Hanekom L, Marwick TH. Reproducibility and accuracy of echocardiographic measurements of left ventricular parameters using real-time three-dimensional echocardiography. J Am Coll Cardiol. 2004;44(4):878–886.

Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. Int J Comput Vis. 1988;1(4):321–331.

Kühl HP, Schreckenberg M, Rulands D, Katoh M, Schäfer W, Schummers G, Bücker A, Hanrath P, Franke A. High-resolution transthoracic real-time three-dimensional echocardiography: Quantitation of cardiac volumes and function using semi-automatic border detection and comparison with cardiac magnetic resonance imaging. J Am Coll Cardiol. 2004;43(11):2083–2090.

Malm S, Frigstad S, Sagberg E, Steen PA, Skjarpe T. Real-time simultaneous triplane contrast echocardiography gives rapid, accurate and reproducible assessment of left ventricular volumes and ejection fraction: A comparison with magnetic resonance imaging. J Am Soc Echocardiogr. 2006;19(12):1494–1501.

Mandinov L, Eberli FR, Seiler C, Hess OM. Diastolic heart failure. Cardiovasc Res. 2000;45(4):813–825.

McInerney T, Terzopoulos D. Deformable models in medical image analysis: A survey. Med Image Anal. 1996;1(2):91–108.

Monaghan MJ. Role of real time 3D echocardiography in evaluating the left ventricle. Heart. 2006;92:131–136.

Orderud F, Hansegård J, Rabben SI. Real-time tracking of the left ventricle in 3D echocardiography using a state estimation approach. Lect Notes Comput Sci. 2007;4791:858–865.

Rabben SI, Torp AH, Støylen A, Slørdahl S, Bjørnstad K, Haugen BO, Angelsen B. Semiautomatic contour detection in ultrasound M-mode images. Ultrasound Med Biol. 2000;26(2):287–296.

Schiller NB, Shah PM, Crawford M, DeMaria A, Devereux R, Feigenbaum H, Gutgesell H, Reichek N, Sahn D, Schnittger I. Recommendations for quantitation of the left ventricle by two-dimensional echocardiography. J Am Soc Echocardiogr. 1989;2(5):358–367.

Soliman OII, Krenning BJ, Geleijnse ML, Nemes A, van Geuns R, Baks T, Anwar AM, Galema TW, Vletter WB, Cate FJT. A comparison between QLAB and TomTec full volume reconstruction for real time three-dimensional echocardiographic quantification of left ventricular volumes. Echocardiography. 2007;24(9):967–974.

Sugeng L, Mor-Avi V, Weinert L, Niel J, Ebner C, Steringer-Mascherbauer R, Schmidt F, Galuschky C, Schummers G, Lang RM, Nesser HJ. Quantitative assessment of left ventricular size and function. Side-by-side comparison of real-time three-dimensional echocardiography and computed tomography width magnetic resonance reference. Circulation. 2006;114(7):654–661.

Taylor G, Humphries J, Mellits E, Pitt B, Schulze R, Griffith L, Achuff S. Predictors of clinical course, coronary anatomy and left ventricular function after recovery from acute myocardial infarction. Circulation. 1980;62(5):960–970.

Tighe DA, Rosetti M, Vinch CS, Chandok D, Muldoon D, Wiggin B, Dahlberg ST, Aurigemma GP. Influence of image quality on the accuracy of real time three-dimensional echocardiography to measure left ventricular volumes in unselected patients: A comparison with gated-SPECT imaging. Echocardiography. 2007;24(10):1073–1080.

van Stralen M, Leung KYE, Voormolen MM, de Jong N, van der Steen AFW, Reiber JHC, Bosch JG. Time continuous detection of the left ventricular long axis and the mitral valve plane in 3-D echocardiography. Ultrasound Med Biol. 2008;34(2):196–207.

White H, Norris R, Brown M, Brandt P, Whitlock R, Wild C. Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. Circulation. 1987;76(1):44–51.

Paper E

Real-Time Active Shape Models for Segmentation of 3D Cardiac Ultrasound

J. Hansegård, F. Orderud, and S. I. Rabben

Real-Time Active Shape Models for Segmentation of 3D Cardiac Ultrasound

Jøger Hansgård[1], Fredrik Orderud[2], and Stein I. Rabben[3]

[1] University of Oslo, Norway, `jogerh@ifi.uio.no`
[2] Norwegian University of Science and Technology, Norway, `fredrik.orderud@idi.ntnu.no`
[3] GE Vingmed Ultrasound, Norway, `stein.rabben@med.ge.com`

Abstract. We present a fully automatic real-time algorithm for robust and accurate left ventricular segmentation in three-dimensional (3D) cardiac ultrasound. Segmentation is performed in a sequential state estimation fashion using an extended Kalman filter to recursively predict and update the parameters of a 3D Active Shape Model (ASM) in real-time. The ASM was trained by tracing the left ventricle in 31 patients, and provided a compact and physiological realistic shape space. The feasibility of the proposed algorithm was evaluated in 21 patients, and compared to manually verified segmentations from a custom-made semi-automatic segmentation algorithm. Successful segmentation was achieved in all cases. The limits of agreement (mean±1.96SD) for the point-to-surface distance were 2.2±1.1 mm. For volumes, the correlation coefficient was 0.95 and the limits of agreement were 3.4±20 ml. Real-time segmentation of 25 frames per second was achieved with a CPU load of 22%.

1 Introduction

Left ventricular (LV) volumes and ejection fraction (EF) are among the most important parameters in diagnosis and prognosis of heart diseases. Recently, real-time three-dimensional (3D) echocardiography was introduced. Segmentation of the LV in 3D echocardiographic data has become feasible, but due to poor image quality, commercially available tools are based upon a semi-automatic approach [1,2]. Furthermore, most reported methods are using iterative and computationally expensive fitting schemes. These factors make real-time segmentation in 3D cardiac ultrasound challenging.

Prior work by Blake *et al.* [3,4] and Jacob *et al.* [5,6], have shown that a state estimation approach is well suited for real-time segmentation in 2D imagery. They used a Kalman filter, which requires only a single iteration, to track the parameters of a trained deformable model based on principal component analysis (PCA), also known as Active Shape Models (ASMs) [7]. ASMs can be trained on manually traced LV contours, resulting in a sub-space of physiologically probable shapes, effectively exploiting expert knowledge of the LV anatomy and function. For segmentation of 3D cardiac data, Van Assen *et al.* [8] introduced the 3D ASM. However, there are to our knowledge no reports of real-time implementations of 3D ASMs.

Based on the work in [3,4,5,6], real-time LV segmentation of 3D cardiac ultrasound was recently introduced by Orderud [9]. He used an extended Kalman filter for robust tracking of a rigid ellipsoid LV model. Later this framework has been extended to use a flexible spline-based LV model coupled with a global pose transform to improve local segmentation accuracy [10]. However, expert knowledge of LV anatomy could not be modeled directly.

To utilize expert knowledge of LV anatomy during segmentation, we propose to use a 3D ASM for real-time segmentation of 3D echocardiograms, by extending the framework described in Orderud [9]. The 3D ASM, trained on LV shapes traced by an expert, gives a compact deformable model which is restricted to physiologically realistic shapes. This model is fitted to the target data in real-time using a Kalman filter. The feasibility of the algorithm is demonstrated in 21 patients, where we achieve real-time segmentation of the LV shape, and instantaneous measurements of LV volumes and EF.

2   Shape Model

A set of 496 triangulated LV training meshes were obtained from 31 patients using a custom-made segmentation tool (GE Vingmed Ultrasound, Norway). The training tool provides manual editing capabilities. When necessary, the user hence did manual editing of the segmentation to make it equivalent to manual tracing.

Building the ASM requires pair-wise point correspondence between shapes from different patients [7,8]. We developed a reparametrization algorithm for converting triangulated LV training shapes into quadrilateral meshes. This algorithm produced meshes with 15 longitudinal and 20 circumferential segments, with vertices approximately identifying unique anatomical positions. The meshes were aligned separately to remove trivial pose variations, such as scaling, translation and rotation.

From the aligned training set, the mean vertex position $\bar{\mathbf{q}}_i$ was computed, and PCA was applied on the vertex distribution to obtain the N_x most dominant eigenvectors. In normalized coordinates, the ASM can be written on the form

$$\mathbf{q}_i(\mathbf{x}_l) = \bar{\mathbf{q}}_i + \mathbf{A}_i \mathbf{x}_l \ , \tag{1}$$

where the position of a vertex $\mathbf{q}_i$ is expressed as a linear combination of the associated subspace of the N_x most dominating eigenvectors combined into the $3 \times N_x$ deformation matrix $\mathbf{A}_i$. Here, $\mathbf{x}_l$ is the local state vector of the ASM. The expression for the ASM can be optimized assuming that the deformation at vertex $\bar{\mathbf{q}}_i$ is primarily directed along the corresponding surface normal $\bar{\mathbf{n}}_i$ of the average mesh. This is done by projecting the deformation matrix $\mathbf{A}_i$ onto the surface normal, giving an N_x-dimensional vector of projected deformation modes $\mathbf{A}_i^{\perp} = \bar{\mathbf{n}}_i^T \mathbf{A}_i$. The optimized expression for the ASM can now be written on the form

$$\mathbf{q}_i(\mathbf{x}_l) = \bar{\mathbf{q}}_i + \bar{\mathbf{n}}_i (\mathbf{A}_i^{\perp} \mathbf{x}_l) \ , \tag{2}$$

reducing the number of multiplications by a factor of three.

Due to the quadrilateral mesh structure of the ASM, a continuous surface is obtained using a linear tensor product spline interpolant. An arbitrary point on the ASM in normalized coordinates can be expressed as $\mathbf{p}_l(\mathbf{x}_l) = \mathbf{T}_l|_{(u,v)}$ where (u,v) represents the parametric position on the surface, and the local transformation $\mathbf{T}_l$ includes the deformation and interpolation applied to the mean mesh. By coupling this model with a global pose transformation $\mathbf{T}_g$ with parameters $\mathbf{x}_g$ including translation, rotation, and scaling, we obtain a surface

$$\mathbf{p}(\mathbf{x}) = \mathbf{T}_g(\mathbf{p}_l(\mathbf{x}_l), \mathbf{x}_g) \tag{3}$$

in real-world coordinates, with a composite state vector $\mathbf{x}^T \equiv [\mathbf{x}_g^T, \mathbf{x}_l^T]$. An illustration showing the steps required to generate the ASM is shown in Fig. 1. In our experiments we used 20 eigenvectors, describing 98% of the total variation within the training set.
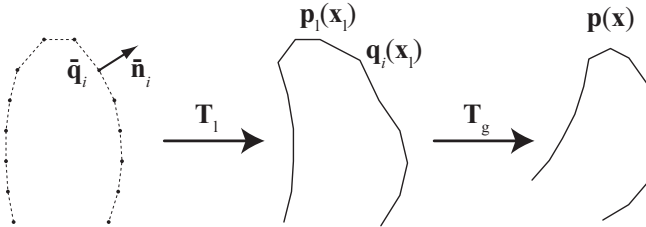
Fig. 1. A point $\mathbf{p}(\mathbf{x})$ on the ASM is generated by first applying a local transformation $\mathbf{T}_l$ described by the ASM state vector $\mathbf{x}_l$ on the mean shape, followed by a global pose transformation $\mathbf{T}_g$ to obtain a shape in real-world coordinates.

3   Tracking Algorithm

The tracking algorithm extends prior work by Orderud [9,10], to enable usage of 3D ASMs in the Kalman filter for real-time segmentation. This is accomplished by using the ASM shape parameters $\mathbf{x}_l$ directly, in addition to the global pose parameters $\mathbf{x}_g$, in the Kalman filter state vector.

3.1   Motion Model

Modeling of motion in addition to position can be accomplished in the prediction stage of the Kalman filter by augmenting the state vector to contain the last two successive state estimates. A *motion model* which predicts the state $\bar{\mathbf{x}}$ at timestep $k+1$, is then expressed as

$$\bar{\mathbf{x}}_{k+1} - \mathbf{x}_0 = \mathbf{A}_1(\hat{\mathbf{x}}_k - \mathbf{x}_0) + \mathbf{A}_2(\hat{\mathbf{x}}_{k-1} - \mathbf{x}_0) \ , \tag{4}$$

where $\hat{\mathbf{x}}_k$ is the estimated state from timestep k. Tuning of properties, like damping and regularization towards the mean state $\mathbf{x}_0$ for all deformation parameters, can then be accomplished by adjusting the coefficients in matrices $\mathbf{A}_1$ and $\mathbf{A}_2$. Prediction uncertainty can similarly be adjusted by manipulating the process noise covariance matrix $\mathbf{B}_0$ used in the associated covariance update equation. The latter will then restrict the change rate of parameter values.

3.2   Measurement Processing

Edge-detection is based on *normal displacement* measurements v_i [4], which are calculated by measuring the radial distance between detected edge-points $\mathbf{p}_{\text{obs},i}$ and the contour surface $\mathbf{p}_i$ along selected search normals $\mathbf{n}_i$. These displacements are coupled with associated measurement noise r_i to weight the importance of each edge, based on a measure of edge confidence. Measurement vectors are calculated by taking the normal projection of the composite state-space Jacobian for the contour points

$$\mathbf{h}_i^T = \mathbf{n}_i^T \left[\frac{\partial \mathbf{T}_g(\mathbf{p}_l, \mathbf{x}_g)_i}{\partial \mathbf{x}_g} \quad \frac{\partial \mathbf{T}_g(\mathbf{p}_l, \mathbf{x}_g)_i}{\partial \mathbf{x}_l} \right] \ , \tag{5}$$

which is the concatenation of a global and a local state-space Jacobi matrix. The global Jacobian is trivially the state-space derivative of the global pose transformation, while the local Jacobian has to be derived, using the chain-rule for multivariate calculus, to propagate surface points on the spline through mesh vertices, and finally to the ASM shape parameters:

$$\frac{\partial \mathbf{T}_g(\mathbf{p}_l, \mathbf{x}_g)}{\partial \mathbf{x}_l} = \sum_{n \in x,y,z} \frac{\partial \mathbf{T}_g(\mathbf{p}_l, \mathbf{x}_g)}{\partial \mathbf{p}_{l,n}} \sum_{j \in 1..N_q} \left(\frac{\partial \mathbf{T}_l(\mathbf{x}_l)_n}{\partial \mathbf{q}_j} \cdot \bar{\mathbf{n}}_k \right) \mathbf{A}_j^\perp \ . \tag{6}$$

Here, $\partial \mathbf{T}_g(\mathbf{p}_l, \mathbf{x}_g)/\partial \mathbf{p}_l$ is the spatial derivative of the global transformation, and $\partial \mathbf{T}_l(\mathbf{x}_l)/\partial \mathbf{q}_j$ is the spatial mesh vertex derivative of the spline interpolant.

3.3   Measurement Assimilation and State Update

All measurements are assimilated in *information space* prior to the state update step. Assumption of independent measurements leads to very efficient processing, allowing summation of all measurement information into an information vector and matrix of dimensions invariant to the number of measurements:

$$\mathbf{H}^T \mathbf{R}^{-1} \mathbf{v} = \sum_i \mathbf{h}_i r_i^{-1} v_i \tag{7}$$

$$\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} = \sum_i \mathbf{h}_i r_i^{-1} \mathbf{h}_i^T \ . \tag{8}$$

The updated state estimate $\hat{\mathbf{x}}$ at timestep k can then by computed by using the information filter formula for measurement update [11], and the updated error covariance matrix $\hat{\mathbf{P}}$ is calculated directly in information space:

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + \hat{\mathbf{P}}_k \mathbf{H}^T \mathbf{R}^{-1} \mathbf{v} \tag{9}$$

$$\hat{\mathbf{P}}_k^{-1} = \bar{\mathbf{P}}_k^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \ . \tag{10}$$

Using this form, we avoid inversion of matrices with dimensions larger than the state dimension.

4   Evaluation

4.1   Data Material

For evaluation of the proposed algorithm, apical 3D echocardiograms of one cardiac cycle from 21 adult patients (11 diagnosed with heart disease) were recorded using a Vivid 7 scanner (GE Vingmed Ultrasound, Norway) with a 3D transducer (3V). In all patients, meshes corresponding to the endocardial boundary were determined using a custom-made semi-automatic segmentation tool (GE Vingmed Ultrasound, Norway). The segmentations were, if needed, manually adjusted by an expert to serve as independent references equivalent to manual tracing.

4.2   Experimental Setup and Analysis

Edge measurements were done perpendicular to the mesh surface within a distance of ± 1.5 cm to the surface at approximately 450 locations, using a simple edge model based on the transition criterion [12]. The ASM was initialized to the mean shape, and positioned in the middle of the volume in the first frame. Segmentation was performed on the evaluation set by running the algorithm for a couple of heartbeats, to give the ASM enough time to lock on to the LV.

The accuracy of the ASM was assessed using the mean of absolute point-to-mesh distances between the ASM and the reference, averaged over one cardiac cycle. Volume differences (bias) between the ASM and the reference were calculated for each frame. End-diastolic volume (EDV), end-systolic volume (ESV), and EF $((\mathrm{EDV} - \mathrm{ESV})/\mathrm{EDV} \cdot 100\%)$ were compared to the manually verified reference (two-tailed t-test assuming zero difference), with 95% limits of agreement (1.96 standard deviations (SD)). EDV and ESV were computed as the maximum and minimum volume within the cardiac cycle respectively.

5   Results

We observed that common challenges with 3D cardiac ultrasound, such as drop-outs, shadows, and speckle noise were handled remarkably well, and segmentation was successful in all of the 21 patients. Some examples are shown in Fig. 2(b-d).

Table 1. Segmentation results showing ASM versus reference.

	Volume [ml]	EDV [ml]	ESV [ml]	EF [%]
Difference (mean$\pm$1.96SD)	3.4*$\pm$20	-5.9*$\pm$21	6.2*$\pm$19	-7.7*$\pm$12
Correlation coeff. (r)	0.95	0.91	0.91	0.74

* Significantly different from 0, p<0.05.

(a)

(b)

(c)

(d)
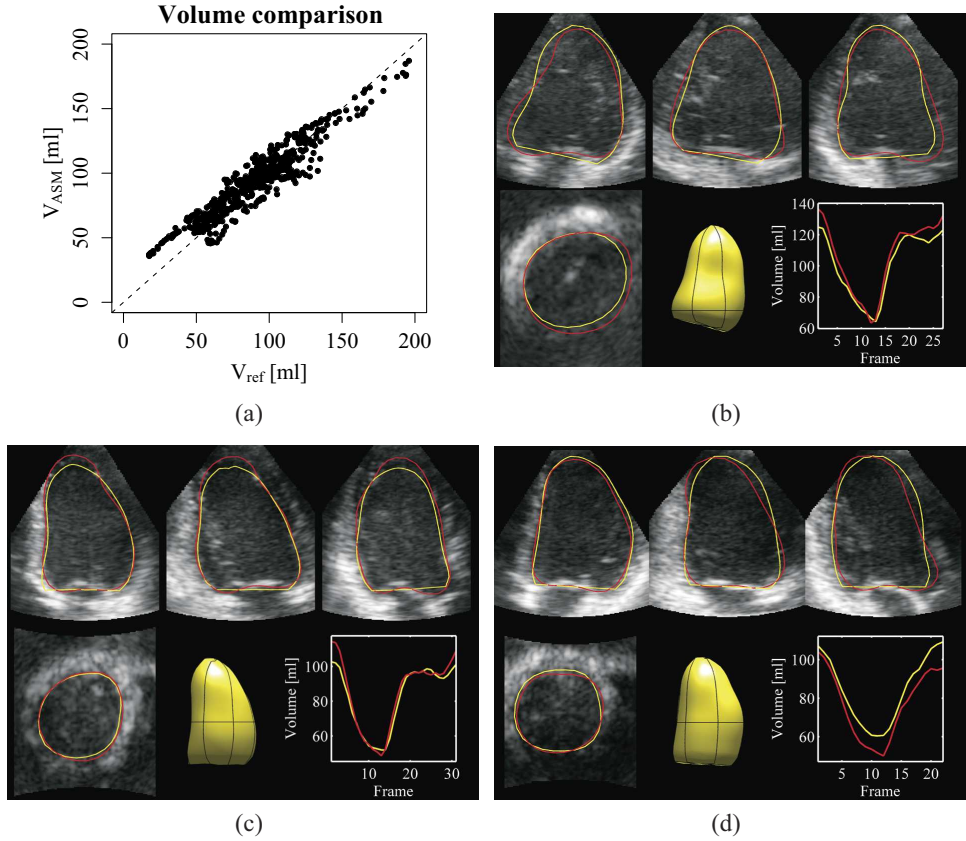
Fig. 2. LV volumes obtained by the ASM (V_{ASM}) in all 21 patients is compared to the reference (V_{ref}) and shown with the identity line (dashed) in (a). In (b-d), the end-diastolic segmentation (yellow) in three patients is compared to the reference (red) and shown along with the volume curve for one cardiac cycle.

The limits of agreement (mean$\pm$1.96SD) for the point-to-surface distance were 2.2$\pm$1.1 mm, indicating good overall agreement between the ASM and the reference. From Tab. 1, column 2, we see that the limits of agreement for volumes were 3.4$\pm$20 ml, with a strong correlation (r=0.95). The volume correspondence between the ASM and the reference is shown in Fig. 2(a).

We also found a strong EDV correlation (r=0.91), with a bias and 95% limits of agreement of -5.9$\pm$21 ml (Tab. 1, column 3). The correlation in ESV was 0.91, with limits of agreement of 6.2$\pm$19 ml, (Tab. 1, column 4), while the correlation in EF was 0.74, with limits of agreement of -7.7$\pm$12% (Tab. 1, column 5).

The CPU load required to maintain real-time segmentation at 25 frames per second (fps) was approximately 22% on a 2.16 GHz Intel Core 2 Duo processor.

6   Discussion

We have presented a fully automatic real-time algorithm for robust and accurate LV segmentation in 3D cardiac ultrasound. This was achieved by combining a 3D ASM with a Kalman

filter based tracking algorithm. The feasibility of the algorithm was demonstrated in 21 patients.

Computational performance was excellent with a CPU load of 22% at 25 fps. Compared to traditional ASM update schemes, the Kalman filter gives good segmentation in a single iteration, allowing real-time implementations.

Contours detected by the ASM showed good overall agreement with the reference shapes, both with respect to point-to-mesh distances and volumes. There was a significant bias in estimated EDV, ESV, and EF, but with relatively narrow 95% limits of agreement. We speculate if the bias is primarily caused by the simple edge detector used, and better results are expected using a more advanced edge detector. Robustness was high in the evaluation set, with successful segmentation in all patients.

Since no user-interaction is required, the algorithm provides rapid analysis of LV function, and it can potentially provide higher reproducibility than semi-automatic methods. Adding means of manual corrections when segmentation fails will be subject for further studies.

The algorithm was evaluated on a population with varying image quality, but for evaluation of clinical applicability, the algorithm must be tested on a larger population. Also, an inherent challenge when using ASMs for clinical applications, is that care must be taken when assembling the training set to ensure inclusion of a sufficiently wide range of pathologies.

Traditional applications where our algorithm fits well includes rapid analysis of LV volumes, EF, and regional function. With real-time segmentation, we expect new applications to emerge, such as patient monitoring, and automated operator guidance.

7   Conclusion

We have developed a fully automatic algorithm for real-time segmentation of the left ventricle in 3D cardiac ultrasound. Initial evaluation in 21 patients is very promising, suggesting that this method is applicable in a clinical setting.

References

1. Jacobs, L.D., Salgo, I.S., Goonewardena, S., Weinert, L., Coon, P., Bardo, D., Gerard, O., Allain, P., Zamorano, J.L., de Isla, L.P., Mor-Avi, V., Lang, R.M.: Rapid online quantification of left ventricular volume from real-time three-dimensional echocardiographic data. European Heart Journal **27** (November 2006) 460–468
2. Sugeng, L., Mor-Avi, V., Weinert, L., Niel, J., Ebner, C., Steringer-Mascherbauer, R., Schmidt, F., Galuschky, C., Schummers, G., Lang, R.M., Nesser, H.J.: Quantitative assessment of left ventricular size and function. Side-by-side comparison of real-time three-dimensional echocardiography and computed tomography width magnetic resonance reference. Circulation **114** (August 2006) 654–661
3. Blake, A., Curwen, R., Zisserman, A.: A framework for spatiotemporal control in the tracking of visual contours. International Journal of Computer Vision **11**(2) (October 1993) 127–145
4. Blake, A., Isard, M.: Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion. Springer-Verlag New York, Inc., Secaucus, NJ, USA (1998)
5. Jacob, G., Noble, J.A., Kelion, A.D., Banning, A.P.: Quantitative regional analysis of myocardial wall motion. Ultrasound in Medicine & Biology **27**(6) (June 2001) 773–784
6. Jacob, G., Alison Noble, J., Mulet-Parada, M., Blake, A.: Evaluating a robust contour tracker on echocardiographic sequences. Medical Image Analysis **3**(1) (March 1999) 63–75
7. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - Their training and application. Computer Vision and Image Understanding **61**(1) (January 1995) 38–59
8. van Assen, H.C., Danilouchkine, M.G., Behloul, F., Lamb, H.J., van der Geest, R., Reiber, J.H.C., Lelieveldt, B.P.F.: Cardiac LV segmentation using a 3D active shape model driven by fuzzy inference. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI. Volume 2878 of Lecture Notes in Computer Science., Springer Berlin / Heidelberg (2003) 533–540

9. Orderud, F.: A framework for real-time left ventricular tracking in 3D+T echocardiography, using nonlinear deformable contours and kalman filter based tracking. In: Computers in Cardiology. (2006)
10. Orderud, F., Hansegård, J., Rabben, S.I.: Real-time tracking of the left ventricle in 3D echocardiography using a state estimation approach. Submitted to: Medical Image Computing and Computer-Assisted Intervention - MICCAI (2007)
11. Comaniciu, D., Zhou, X.S., Krishnan, S.: Robust real-time myocardial border tracking for echocardiography: An information fusion approach. Medical Imaging, IEEE Transactions on **23**(7) (2004) 849–860
12. Rabben, S.I., Torp, A.H., Støylen, A., Slørdahl, S., Bjørnstad, K., Haugen, B.O., Angelsen, B.: Semiautomatic contour detection in ultrasound M-mode images. Ultrasound in Med. & Biol. **26**(2) (Feb. 2000) 287–296