2007

# Segmentation And Computer-aided Diagnosis Of Cardiac MR Images Using 4-D Active Appearance Models

Honghai Zhang
*University of Iowa*

# SEGMENTATION AND COMPUTER-AIDED DIAGNOSIS OF CARDIAC MR IMAGES USING 4-D ACTIVE APPEARANCE MODELS

by

Honghai Zhang

An Abstract

Of a thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

July 2007

Thesis Supervisor: Professor Milan Sonka

## ABSTRACT

The four-dimensional (4-D) cardiac MR images contain rich information about the static and dynamic properties of the heart, which were not fully utilized in clinical practice for quantitative analysis – a difficult task for humans, which can be achieved by computer-aided image analysis and diagnosis. In this thesis, the 4-D Active Appearance Model (AAM) was used to achieve highly automated computer segmentation of the left and right ventricles (LV and RV) and the diagnosis of normal and tetralogy of Fallot (TOF) patients. The whole process was implemented in four stages: data construction, model construction, computer segmentation, and computer-aided diagnosis.

The data construction stage overcame most inherent limitations of cardiac MR imaging and produced high-quality 4-D ventricular image with isotropic voxels, complete coverage and no respiratory motion artifacts. A manual tracing application was developed to trace the ventricular surfaces in a true 4-D context and produced accurate independent standard for model construction and segmentation validation.

In the model construction stage, the 4-D AAMs were constructed using a custom designed automatic landmarking and texture mapping procedure with high efficiency.

In the computer segmentation stage, the 4-D AAMs were applied to segment the left and right ventricles of 25 normal and 25 TOF patient scans. The segmentation achieved accurate results measured by signed surface positioning errors. On normal hearts, the average signed errors were 0.3±2.3 mm for LV and 0.1±3.4 mm for RV. On TOF hearts with large shape variability, the errors were -1.5±3.2 mm for LV and

-0.9±4.3 mm for RV. Other error metrics such as relative overlapping also indicated good segmentation accuracies.

In the computer-aided diagnosis stage, 100% normal/TOF classification was achieved using the novel 4-D ventricular function indices – the shape modal indices. The longitudinal analysis performed on subjects with multiple annual scans showed that the normal subjects exhibited smaller variances of these 4-D indices than TOF patients, which demonstrated the potential of using them as disease status determinants. In addition, the quantitative 4-D indices provided more information about the dynamic properties of the heart and identified patient-specific features that were not sensed by human expert observers.

Abstract Approved: _____

Thesis Supervisor

_____

Title and Department

_____

Date

# SEGMENTATION AND COMPUTER-AIDED DIAGNOSIS OF CARDIAC MR IMAGES USING 4-D ACTIVE APPEARANCE MODELS

by

Honghai Zhang

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

July 2007

Thesis Supervisor: Professor Milan Sonka

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

---

PH.D. THESIS

---

This is to certify that the Ph.D. thesis of

Honghai Zhang

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Electrical and Computer Engineering at the July 2007 graduation.

Thesis Committee: _____
Milan Sonka, Thesis Supervisor

_____
Gary E. Christensen

_____
Andreas Wahle

_____
Joseph M. Reinhardt

_____
Thomas D. Scholz

## TABLE OF CONTENTS

## LIST OF TABLES

Table

## LIST OF FIGURES

Figure

## LIST OF ALGORITHMS

Algorithm

# CHAPTER 1
# INTRODUCTION

The state-of-the-art medical imaging techniques are capable of producing four-dimensional images with high spatial and temporal resolutions. They provide the physicians with more direct knowledge about the static and dynamic properties of the target organ and enable more accurate and reliable diagnosis. However, the rich information contained in these images is not fully utilized for quantitative analysis – a difficult task for humans, which can be achieved by computer-aided image analysis and diagnosis.

The medical image analysis performed by a human expert accomplishes two tasks. First, the target organ in the image is identified. Second, a set of features of the target organ are extracted and used to form the diagnoses: normal or diseased, the disease severity, and the associated potential risks. The success of this human approach relies on the knowledge formed in a 'average-and-difference' manner. In cardiac image analysis, the 'average' knowledge about the appearance of the heart is used to locate the cardiac borders and roughly identify whether the subject is normal or diseased. The 'difference' part of the knowledge describes how much a certain heart can differ from the average heart and is used to achieve accurate cardiac border locations and disease stage identifications. The human nature determines that the forming of such 'average-and-difference' knowledge is a closed-loop process, the 'average' and 'difference' are constantly learned and refined with the accumulation of experience.

The computer-aided medical image analysis performs the same tasks as humans

by using quantitative features to form decisions, but translating subjective human knowledge to mathematical expressions often suffers from over-simplification and the ability of learning from experience is often lost. An important observation of the nature of the human's 'average-and-difference' knowledge is that it can be described by a statistical model in which the 'average' is the mean and 'difference' is the second and higher orders of statistics. The computer-based methods that form, learn, and utilize such statistical information embedded in the human knowledge and closely emulate the closed-loop human process are the *model-based* analyses, which also have a bonus advantage of being able to identify subtle but important features that are often invisible to human eyes.

In this study, the model-based approach was applied to the segmentation and diagnosis of a specific congenital heart disease – the post-operative tetralogy of Fallot. The implementation and capability of the model-based approach in four-dimensional were studied using magnetic resonance images.

## 1.1 Tetralogy of Fallot

Congenital heart disease is a broad term refers to abnormalities of the cardiac structure and function caused by abnormal or disordered heart development before birth. Described in 1672 by Niels Stensen and in 1888 by Etienne Fallot, the *tetralogy of Fallot* (TOF) is a combination of four heart defects that cause insufficiently oxygenated blood to be pumped to the body. It occurs in approximately 5 out of 10,000 infants and accounts for 10–15% of all congenital heart diseases. Figure 1.1 shows the classic form of TOF that includes the following four defects [1–3, 28].

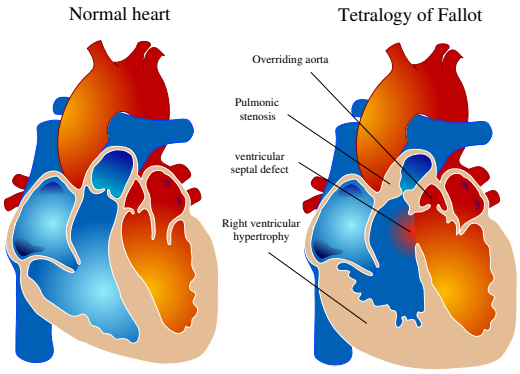



Figure 1.1: The structure of the normal heart and four defects of tetralogy of Fallot. (Public domain image made by Mariana Ruiz Villarreal.)

1. *Ventricular septal defect (VSD)*: A hole between the two ventricles of the heart. It is centered around the most superior aspect of the septum and is often single and large.
2. *Pulmonic stenosis*: Right ventricular outflow tract obstruction, a narrowing at or just below the pulmonary valve. The stenosis is mostly the result of hypertrophy of the septoparietal trabeculae, however the deviated outlet septum is believed to play a role.
3. *Overriding aorta*: The aortic valve is not restricted to the left ventricle, thus having bi-ventricular connections. The aortic root can be moved anteriorly or override the septal defect, but it is still on the right side of the pulmonary artery root.
4. *Right ventricular hypertrophy*: The right ventricle is more muscular than normal. Due to the mis-arrangement of the external ventricular septum, the right

ventricular wall increases in size to deal with the increased obstruction to the right outflow tract. This feature is now generally agreed to be a secondary anomaly, as the level of hypertrophy generally increases with age.

The common treatment for TOF is surgery. A temporary operation may be done at first if the infant is small and the complete repair comes later. The typical temporary operation is a shunt operation performed to provide adequate blood flow to the lung and to allow the infant to grow big enough for a full repair. The shunt is built between the aorta and the pulmonary artery and will be removed in the full repair. In the complete repair, the ventricular septal defect is closed with a patch, the right ventricular outflow tract is opened by removing some thickened muscle below the pulmonary valve, the pulmonary valve is repaired or removed, and the peripheral pulmonary arteries that go to both lungs are enlarged. Sometimes a tube is also placed between the right ventricle and the pulmonary artery and this procedure is called a Rastelli repair.

After the successful complete repair, the TOF patients still need regular follow-ups since several long-term complications and risks may develop and therefore need more surgeries or medications. Some common complications are listed below [49].

- *Right ventricular failure:* It may happen especially if the surgery created severe pulmonary valve insufficiency, the regurgitation of blood backwards from the pulmonary artery into the right ventricle.
- *Electrical conduction abnormalities:* Every TOF patient has right bundle branch block secondary to the congenital VSD. Sewing the patch into the ventricular

septum can create heart block or failure of the upper atria to conduct or communicate with the lower ventricles. A permanent pacemaker is occasionally needed.

- *Arrhythmias:* The surgery on the ventricles can introduce an infrequent but life-threatening arrhythmia – postoperative ventricular tachycardia.
- *Residual hole in the ventricular septum:* Another possible outcome, with oxygenated blood passing from the left side of the heart to the right (shunting).

It is essential to monitor the changes of cardiac function for TOF patients before and after surgery with lifetime follow-up care. Since it is not preferred to operate on infants whose hearts are still growing and can still function 'normally' by providing enough oxygenated blood to the body, the cardiac function is an important determinant of the necessity and optimal timing of the surgery. After surgery, the cardiac function assessment helps the physicians to identify any potential risks and complications. The cardiac function indices that are important for TOF patients follow-up care are the indicators of the ventricular function.

## 1.2 Ventricular Function Analysis

### 1.2.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is widely used in clinical ventricular functional analysis due to its advantages such as high spatial and temporal resolutions, favorable signal-to-noise ratio (SNR), and flexible imaging locations. Cardiovascular Magnetic Resonance (CMR) refers to the techniques and protocols focused on imaging of the heart.

Each CMR image is a 2-D image acquired from a slice of the heart. In order to reduce the scan time and achieve desired spatial resolution and SNR, the slice thickness is often chosen to be larger than the in-plane resolution. A 3-D image representing the whole heart is composed by stacking of CMR images and the resulting 3-D image therefore has *anisotropic* voxels. A typical voxel size of a ventricular CMR image is $1.5 \times 1.5 \times 8$mm, where 8mm is the slice thickness and 1.5mm is the in-plane resolution.

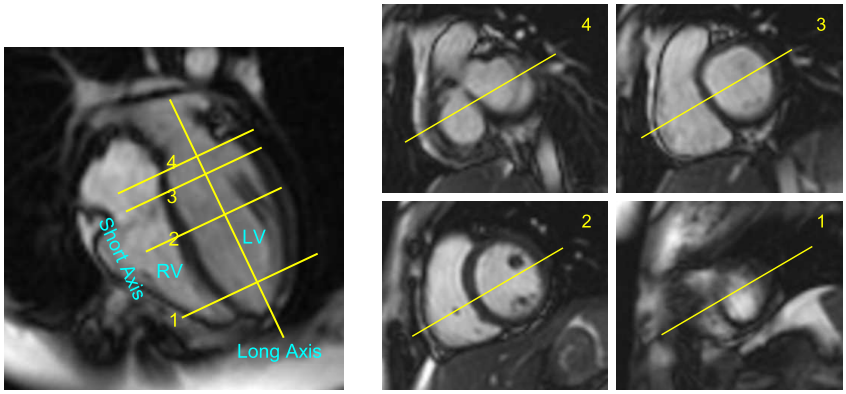
A 4-D (3-D+time) image representation of the heart is acquired from multiple cardiac cycles. To acquire a 4-D image of a complete cardiac cycle with $N$ phases, ECG gating is used to trigger the data acquisition. A cardiac cycle is defined as the time between two peaks of the ECG R-wave. Each R-peak triggers the acquisition of $N$ images from one fixed slice location at all phases of a cardiac cycle until images from all prescribed slices are acquired. The resulting 4-D image is a composite image constructed from multiple cardiac cycles and therefore may have the following artifacts.

- The anisotropic imaging voxel may introduce partial volume effect – a loss of resolution caused by multiple features present in the same imaging voxel. For example, a voxel may contains both water and fat and the resulting image intensity is then neither of fat nor of water.
- The motion artifact caused by motion of the entire or part of the object during acquisition. It typically results in blurring of images and/or shifting of target object in images acquired from different breathing phases.

- Flow artifact caused by flowing blood or fluids in the body.

### 1.2.2 Clinical Ventricular MR Imaging

Two types of CMR image sequences are commonly used in ventricle analysis: the long-axis view and short-axis view. Figure 1.2 shows some sample of such images. At the beginning of a CMR scan, a set of *localizer* images are acquired from three standard orthogonal orientations (transverse, sagittal and coronal views) with large field of view. From them, the location and orientation of the heart is approximated and then a *four-chamber slice* is selected as shown in Figure 1.2a where all four heart chambers, the left and right ventricles and atria, are clearly visible.




(a) (b)

Figure 1.2: The long- and short-axis views of the heart. (a) A long-axis (four-chamber) image. (b) The short-axis images from different locations.

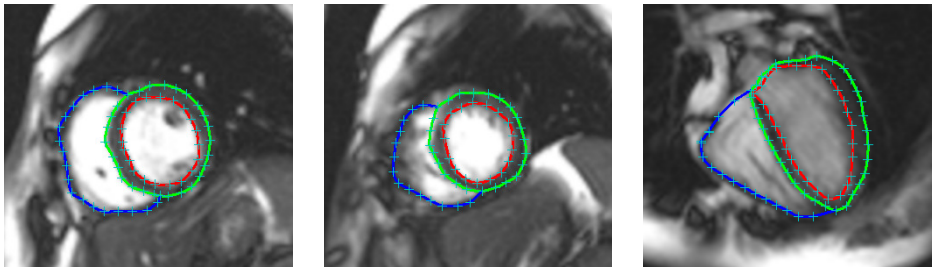
Based on the four-chamber slice, the long-axis images are acquired from planes parallel to the four-chamber slice with constant slice thickness. Note that not all long-axis images contain clear view of all four heart chambers. On the four-chamber

slice, the long-axis of the left ventricle (LV) is estimated by connecting the LV apex and the center of the LV base as shown in Figure 1.2a. The short-axis images are then acquired from planes perpendicular to both the LV long-axis and the long-axis imaging planes. Four sample short-axis view images from the ventricular apexes, mid-sections, bases, and atria are shown in Figure 1.2b and their approximate imaging plane locations are marked in Figure 1.2a.

### 1.2.3 Ventricular Image Analysis

The conventional ventricular MR image analysis is performed only on the short-axis images using expert manual tracing. Two LV cardiac borders, the endocardial and epicardial borders, can be traced. Only one right ventricle (RV) border can be identified due to the insufficient contrast of the RV myocardium on the short-axis images. Figure 1.3 shows several examples of the long- and short-axis images with manual tracings. On the short-axis images, the LV endocardial border is not always defined at locations with maximum intensity gradients. The papillary muscle is included in the blood pool so the LV endocardial border has a circular shape. Only half of the LV epicardial border can be easily identified while the other half has to be estimated. The manual tracing of the RV involves some estimation in the context of the neighboring anatomic structures. Finding ventricular bases is difficult on the short-axis images but relatively easy on the long-axis images. Although the ventricular valves are almost invisible, their plane locations can be estimated by humans when the atria are visible.

The manual tracings are often performed only on two cardiac phases, the end-



(a) (b) (c)

Figure 1.3: Examples of ventricular images with manual tracings. (a,b) The short-axis image tracings. (c) The long-axis image tracing.

diastole and the end-systole, identified from visual inspection of changes of the LV in size. Several ventricular function indices can be derived. The ventricular volume is often calculated as the summation of ventricular areas enclosed by the cardiac borders of all slices multiplied by the slice thickness. The LV mass can be derived from the LV myocardium volume. The ejection fraction (EF) is one of the important ventricular function indices derived from end-diastolic volume (EDV) and end-systolic volume (ESV) as

$$\mathrm{EF} = \frac{\mathrm{EDV} - \mathrm{ESV}}{\mathrm{EDV}} \times 100\% \tag{1.1}$$

It has long been noticed that the major source of error in these volumetric measurements derived from the short-axis images is the inclusion or exclusion of slices at the level of the ventricular bases [20, 38]. The MR images with anisotropic voxels provide incomplete information and the valves that separate ventricles and atria are almost invisible on short-axis images. The partial volume effects at the ventricular bases make the identification of the most basal slice more difficult [32]. In clinical

practice, the ventricle coverage is often sacrificed for consistency and a common criterion for defining the ventricular base is to include any short-axis image that contains at least 50% circumferential extent of clearly identifiable myocardium. It is also common to include the papillary muscle in the LV blood pool [46]. This criterion and the partial volume effects can been seen in Figure 1.3, especially in base and apex images.

Even with the defined rules for short-axis image tracing, the inter- and intra-observer variabilities still cannot be eliminated. On the other hand, the apex, the mitral and the tricuspid valve planes can be easily identified in the long-axis view images. Several methods that utilize long-axis images were proposed and shown to have improved coverage of the LV. The method proposed by Bloomer et al. [6] uses the radial long-axis images in LV volume measurement. Swingen et al. [46] proposed a feedback-assisted method that puts guide points on both the long- and short-axis images and uses guide points to reconstruct the LV in 3-D. Note that these methods only apply to the LV whose shape is more regular and cylindrical, the shape information embedded in the long-axis images is only partially utilized.

### 1.2.4 Limitations

As a composite image, the 4-D MR image suffers from anisotropic voxels and inherent MR artifacts. In addition, some imperfections of the MR imaging protocol also affect the resulting image quality. Such imperfections include inaccurate estimation of the long-axis plane orientation and LV long-axis direction, inaccurate short-axis plane orientation, and insufficient number of slices to cover the heart completely.

The manual tracing of cardiac borders is a tedious task with inevitable inter- and intra-observer variabilities. The ventricular volumes calculated are only rough approximations. The resulting volumetric indices are therefore inaccurate and contain no shape information. In addition, these indices are only derived from one or two cardiac phases and cannot fully describe the 4-D nature the heart.

## 1.3 Motivations

In 4-D ventricular MR images, more static and dynamic information of the heart is visible in high resolution and quality and more features of congenital heart diseases are found and intensively studied. But achieving automated quantitative measurements of these features and using them in patient care as reliable disease indicators are often difficult mainly due to the lack of computer segmentation and analysis methods.

The model-based segmentation methods, specifically, active shape model (ASM) [14, 16] and active appearance model (AAM) [13] are well suited for the ventricular segmentation task. During the past decades, these methods have been successfully applied to ventricular segmentation of MR images in 2-D [37], 3-D [36, 48, 50], bi-temporal 3-D [45], and recently 4-D [39]. But the common pitfall of the current 3-D and 4-D applications in MR images is that only the short-axis images are used and the conventional manual tracings on them are treated as the independent standard and used to construct the model. The problems of the incomplete and inconsistent ventricle coverage are not fully addressed.

One of the MR artifacts, the respiratory motion artifact, can be modeled as simple shifting or mis-alignment of heart in 2-D images [21] and therefore can be solved by

rigid registration. A 3-D version of the motion correction is proposed by Stegmann et al. [44] but only utilizes short-axis images and thus relies on the assumption that the LV long-axis is always perpendicular to the short-axis imaging plane, which is often not true in the clinical image data.

Most importantly, the long-axis images, which are always acquired in standard ventricular MR protocol, are not fully utilized or not considered at all in most current ventricular data construction and model-based segmentation approaches.

## 1.4 Goals

The final goal of this research is to provide quantitative disease indices for TOF that can be used to achieve computer-aided diagnosis (CAD). To achieve this, a set of novel 4-D ventricular function indices that fully or partially describe the 4-D nature of the ventricles are proposed. As a proof of concept, the feasibility of using these indices to achieve normal and TOF classification was fully evaluated. Then the potential of using these indices as disease stage determinants and early indicators of future TOF induced cardiac risks was explored.

In order to derive 4-D function indices, the accurate 4-D segmentation of the left and right ventricles is required. The 4-D AAM segmentation was used due to its advantages of incorporated human knowledge and inherent capability of producing 4-D function indices. In this research, an efficient 4-D AAM method was implemented and used for ventricle segmentation.

The success of AAM relies on high quality image and accurate manual tracings. In this reserach, the 4-D data construction procedure was developed for ventricular

MR images, it include image data interpolation, motion correction, image fusion, true 4-D manual tracing and shape interpolation. Most importantly, the long-axis images and the embedded shape information were fully utilized.

In summary, the whole cardiac MR analysis process is divided into four pipeline stages, each with its own aim. Then the four goals to be achieved in this study are listed as follows and their achievements are reviewed in 7.1 on page 108,

1. *4-D data construction:* Build 4-D data sets (images and labeled volumes) that describe the true 4-D nature of the ventricles with isotropic voxels, complete ventricle coverage, and free of motion artifacts.
2. *4-D model construction:* Develop efficient methods to build and train the 4-D AAMs of the ventricles.
3. *4-D AAM segmentation:* Use the created 4-D AAMs to achieve accurate 4-D ventricular segmentations.
4. *Computer-aided diagnosis:* Develop novel 4-D ventricular function indices that can describe the difference between normal and TOF subjects and test the potentials of using them to achieve CAD of TOF prognosis.

# CHAPTER 2
# STATISTICAL MODELS IN MEDICAL IMAGE ANALYSIS

## 2.1 Introduction

Most image segmentation methods can be categorized as either image-based or model-based. The image-based methods extract image features such as pixel intensity and edge strength and combine them with some prior knowledge about the target object to achieve segmentation. It is always difficult to translate complex and subjective prior knowledge into a simple mathematical form. Such limitation prevents the image-based approach from achieving satisfactory results in ventricular segmentation.

In the model-based segmentation approach, a statistical model is first created and trained from a population of manually traced samples so the human knowledge embedded in the manual tracing is incorporated into the model. Then the automatic segmentation becomes an iterative process, the model is fit onto the target image and iteratively refined based on the knowledge embedded in the model and image features. In this chapter the mathematical background and general procedure of two major model-based methods, active shape model and active appearance model, are described.

The general approach of CAD is to use some quantitative metrics to achieve diagnosis. The design of a specific CAD method has two tasks: 1) identify metrics to use and how to measure them; 2) choose appropriate classification method to achieve diagnosis. The choice of metrics is often more important. The conventional way of choosing metrics is the heuristic approach: identify metrics from experience

and then test whether diagnosis is possible. In this case, finding the optimal set of metrics is always difficult, it is also a process of mapping complex human knowledge to quantitative measurements. There is no guarantee that the selected metrics are the unique disease characteristics or they can utilize all the information provided by the segmentation.

The statical model incorporates and closely emulates the 'average-and-difference' human knowledge and therefore becomes the appropriate choice of metrics. More specifically, the statistical model implicitly translates the 'average-and-difference' knowledge into 'mean-and-variance' quantitative metrics which also contain most information provided the segmentation. The potential of using model-based approach in CAD is fully explored in Chapter 6.

## 2.2 Active Shape Model

### 2.2.1 Point Distribution Model

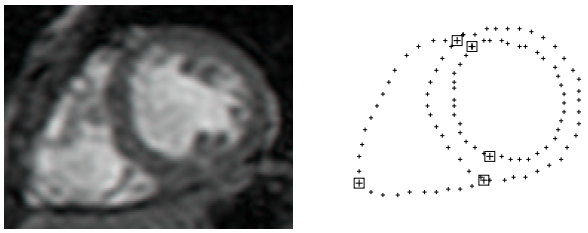
In 1977, Harshman et al. [30] proposed a method to describe the tongue shape, and the variations of tongue shape when pronouncing English vowels were analyzed. This work suggested the feasibilities of using evenly distributed points as shape representation, and using statistical analysis to model shape variations. In the early 1990's, a model-based approach in shape analysis was proposed by Cootes and Taylor [10, 11, 15, 18] and named as *Point Distribution Model* (PDM).

In PDM, the shape of an object is represented by a set of $n$ corresponding points on its boundary called *landmarks*. The most important requirement of choosing landmarks is the point correspondence, i.e., all the $i$th landmarks in the population

must come from the 'same' location of the object. An example of landmarks selected on a short-axis MR image is shown in Figure 2.1. With $n$ defined landmarks, the shape of an object is then written as a *shape vector* $\boldsymbol{s}$

$$\boldsymbol{s} = [x_1, y_1, z_1, \cdots, x_n, y_n, z_n]^T, \text{ or } \boldsymbol{s} = [x_1, y_1, \cdots, x_n, y_n]^T, \tag{2.1}$$

where $(x_i, y_i, z_i)$ or $(x_i, y_i)$ is the coordinate of the $i$th landmark. The dimensionality of such shape vector is $3n$ for 3-D surface, $2n$ for 2-D contour.



(a) (b)

Figure 2.1: Landmarks defined on a short-axis MR image. (a) The original image. (b) The landmarks (marked by +) representing the ventricular shape.

In order to analyze the shape variations, the information contained in the shape vector must remain the same regardless of any scaling, rotation, and translation applied to the object. The effects of these Euclidean transforms must be removed first to align all shapes to a common reference frame. The procedure of such shape alignment is known as the *Procrustes Analysis* [29] listed in Algorithm 2.1 and illustrated using an example of aligning ventricular shapes in Figure 2.2.

The mean shape produced by the Procrustes Analysis is a first order statistical

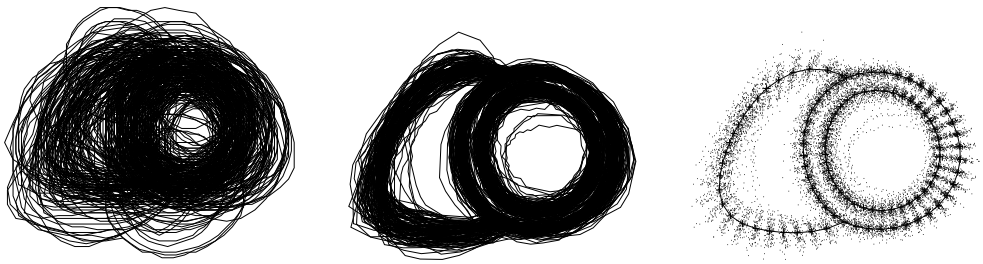
**Input**: Sample shape population, $\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_M$.
**Output**: Aligned shapes, $\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_M$, and a *mean shape* $\overline{\boldsymbol{s}}$.

Align the centroids all shapes to remove the effects of translation.
Arbitrarily select a sample shape as the initial $\overline{\boldsymbol{s}}$.
**while** $\overline{\boldsymbol{s}}$ *has not converged* **do**
    **for** *each* $\boldsymbol{s}_i$ **do**
        Transform $\boldsymbol{s}_i$ so it is aligned to $\overline{\boldsymbol{s}}$.
    **end**
    Calculate a new $\overline{\boldsymbol{s}}$ from the aligned $\boldsymbol{s}_i$ using

$$\overline{\boldsymbol{s}} = \sum_{i=1}^{M} \boldsymbol{s}_i / M. \tag{2.2}$$

**end**

Algorithm 2.1: The Procrustes Analysis



(a) (b) (c)

Figure 2.2: The Procrustes Analysis. The sample shapes before and after (a,b) alignment. (c) The aligned landmarks have approximately Gaussian distributions.

description of the shape population. Figure 2.2c shows that the distribution of the resulting landmarks is approximately Gaussian. The second order description, the variance, can be found by *Principal Component Analysis* (PCA) [33]. The $M$ $N$-dimensional sample shape vectors, are treated as $M$ points in a $N$-dimensional space and they are often correlated. The PCA generates $N$ new orthogonal bases (axes) originated at the mean shape such that under these new bases, the correlation of

samples is minimized.

Each such new basis, also called *principal component* or *PCA mode*, is a direction vector and the projection of a sample onto it is a measurement of variance. The PCA modes are often sorted such that the first mode represents the greatest variance of the population and the following modes have decreasing variances. The decreasing of variances makes it possible to approximate a sample with fewer modes and consequently represent a sample vector in a more compact form. The effects and benefits of applying PCA to a 2-D point cloud are shown in Figure 2.3. The procedure of PCA is as follows.

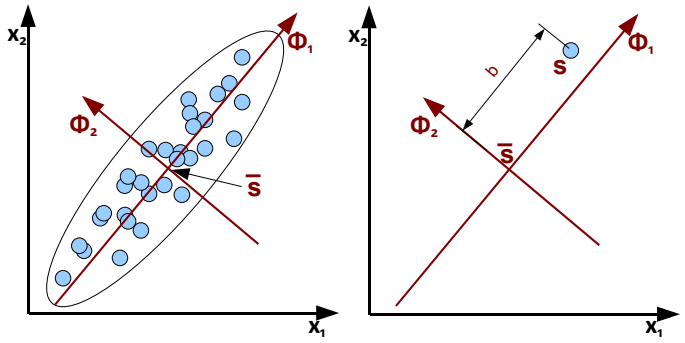



Figure 2.3: Applying PCA to 2-D points in $(x_1, x_2)$ space produces components $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$. A point $\boldsymbol{s}$ is approximated as $\boldsymbol{s} \approx \overline{\boldsymbol{s}} + \boldsymbol{\phi}_1 b$.

1. Compute the mean shape of the aligned shape population using equation (2.2).
2. Compute the correlation of the aligned shapes as a covariance matrix $\mathbf{C}$.

$$\mathbf{C} = \frac{1}{M}\mathbf{B}\mathbf{B}^T, \quad \mathbf{B} = \left[(\boldsymbol{s}_1 - \overline{\boldsymbol{s}})|(\boldsymbol{s}_2 - \overline{\boldsymbol{s}})|\cdots|(\boldsymbol{s}_M - \overline{\boldsymbol{s}})\right], \tag{2.3}$$

3. Transform $\mathbf{C}$ to a diagonal matrix $\mathbf{\Lambda}$ using eigen decomposition of $\mathbf{C}$ such that

$$\mathbf{C\Phi} = \mathbf{\Lambda\Phi}, \quad \text{and} \quad \mathbf{\Phi}^T\mathbf{C\Phi} = \mathbf{\Lambda}. \tag{2.4}$$

The diagonal elements of $\mathbf{\Lambda}$ are the eigenvalues $\lambda_i$ sorted as $\lambda_i \geq \lambda_{i+1}$. The columns of matrix $\mathbf{\Phi}$ are the eigenvectors $\boldsymbol{\phi}_i$ associated with $\lambda_i$.

4. Any shape vector $\boldsymbol{s}$ is then defined by $\overline{\boldsymbol{s}}$ and the new bases specified by $\mathbf{\Phi}$ as

$$\boldsymbol{s} = \overline{\boldsymbol{s}} + \mathbf{\Phi}\boldsymbol{b}, \quad \text{i.e.} \quad \boldsymbol{b} = \mathbf{\Phi}^{-1}(\boldsymbol{s} - \overline{\boldsymbol{s}}) = \mathbf{\Phi}^T(\boldsymbol{s} - \overline{\boldsymbol{s}}). \tag{2.5}$$

The vector $\boldsymbol{b}$ is unique for each individual sample while $\overline{\boldsymbol{s}}$ and $\mathbf{\Phi}$ are common for all samples. Then $\boldsymbol{b}$ becomes the new representation of $\boldsymbol{s}$ and is called *modal indices*. The $i$th element of $\boldsymbol{b}$, $b_i$, is a quantitative measure of the derivation between $\boldsymbol{s}$ and $\overline{\boldsymbol{s}}$ along the $i$th principal component and is often referred as the $i$th *mode* or the $i$th *modal index* of $\boldsymbol{s}$.

5. Since the smaller the eigenvalue, the less its influence on the shape variance, a cutoff $z$ is selected using a pre-selected cutoff ratio $\alpha$ as

$$\sum\nolimits_{i=1}^{z} \lambda_i \geq \alpha \sum\nolimits_{i=1}^{N} \lambda_i, \quad 0 < \alpha < 1. \tag{2.6}$$

A typical value of $\alpha$ is 97%. The compact representation of the model and the

sample is then achieved by

$$\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \cdots, \boldsymbol{\phi}_z], \quad \boldsymbol{b} = [b_1, b_2, \cdots, b_z]^T \tag{2.7}$$

such that any shape vector $\boldsymbol{s}$ is approximated by

$$\boldsymbol{s} \approx \overline{\boldsymbol{s}} + \boldsymbol{\Phi}\boldsymbol{b}. \tag{2.8}$$

The standard deviation $\sigma_i$ of the $i$th PCA mode is $\sqrt{\lambda_i}$. If the vector $\boldsymbol{b}$ is constrained such that $b_i \in [-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]$, 99.7% of the obsevations from the population is covered. For each $b_i$, varying its value introduces complex shape variations otherwise cannot be described or quantitatively measured. An example of such shape variations is shown in Figure 2.4.

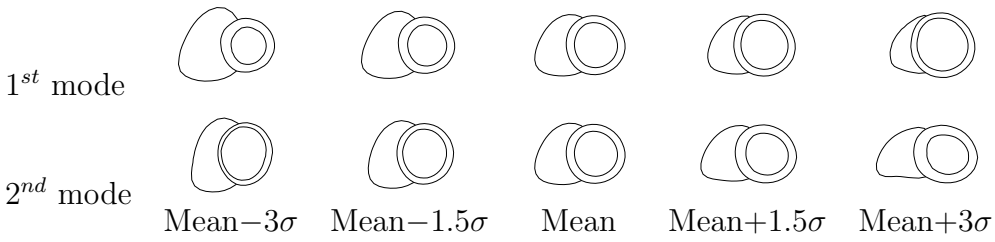



Figure 2.4: The shape variations associated with the two strongest shape modes.

In equation (2.3), the size of the covariance matrix $\mathbf{C}$ is $N \times N$, and the complexity of eigen decomposition is $O(N^3)$. However, if $M \ll N$, a smaller $M \times M$ covariance

matrix $\mathbf{D}$ is only needed as

$$\mathbf{D} = \frac{1}{M}\mathbf{B}^T\mathbf{B}. \tag{2.9}$$

The eigen decomposition of $\mathbf{D}$ finds eigenvectors $\mathbf{\Phi}_D$ and eigenvalues $\mathbf{\Lambda}_D$ and the non-zeros eigenvalues solved for $\mathbf{D}$ are equal to the eigenvalues solved for $\mathbf{C}$. The eigenvectors $\mathbf{\Phi}_C$ solved for $\mathbf{C}$ are

$$\mathbf{\Phi}_C = \mathbf{B}\mathbf{\Phi}_D. \tag{2.10}$$

The columns of $\mathbf{\Phi}_C$ must be normalized to unit magnitude to maintain orthonormality requirement for the calculation of modal index vector $\boldsymbol{b}$.

### 2.2.2 Active Shape Model Segmentation

The knowledge about the object shape embedded in the PDM can be used in segmentation and this model-based segmentation method is called *Active Shape Model* (ASM) [14,16]. The resulting segmentation in the target image space is defined as the target shape $\boldsymbol{s}'$, as the model shape $\boldsymbol{s}$ in the model space. The relationship between the $\boldsymbol{s}'$ and $\boldsymbol{s}$ is specified by a transform $\mathcal{T}$ such that $\boldsymbol{s}' = \mathcal{T}(\boldsymbol{s})$. The general procedure of the ASM is listed in Algorithm 2.2.

The searching for more suitable landmark locations is the *model refinement* step. The simplest way to do that is to search a new location for each landmark along the normal direction of the current target shape. How to measure the suitability of the candidate landmarks is learned from an extra *training* step.

For the $i$th landmark, from all the images in the population, $k$ pixels are chosen

**Input**: The target image. $\overline{\boldsymbol{s}}$, $\boldsymbol{\Lambda}$, and $\boldsymbol{\Phi}$ of the PDM.
**Output**: The target shape $\boldsymbol{s}'$, the model shape $\boldsymbol{s}$, and modal index vector $\boldsymbol{b}$.
Initialize with $\boldsymbol{s} = \overline{\boldsymbol{s}}$ and $\boldsymbol{b} = \mathbf{0}$.
Estimate $\mathcal{T}$ to approximately fit $\boldsymbol{s}'$ onto the target image.
**while** $\boldsymbol{b}$ *has not converged* **do**
    Update $\boldsymbol{s}'$ by moving landmarks to more suitable locations.
    Align updated $\boldsymbol{s}'$ to $\overline{\boldsymbol{s}}$ with a new $\mathcal{T}$.
    $\boldsymbol{s} = \mathcal{T}^{-1}(\boldsymbol{s}')$.
    $\boldsymbol{b} = \boldsymbol{\Phi}^T(\boldsymbol{s} - \overline{\boldsymbol{s}})$ and apply constraints of $b_i \in [-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]$.
    $\boldsymbol{s} = \overline{\boldsymbol{s}} + \boldsymbol{\Phi}\boldsymbol{b}$.
    $\boldsymbol{s}' = \mathcal{T}(\boldsymbol{s})$.
**end**

Algorithm 2.2: Active Shape Model segmentation

on each side along the normal direction of the object border. From these $2k + 1$ pixels, the intensity derivatives are normalized and put into a profile vector $\boldsymbol{g}_i$. For all such profile vectors, a profile model is created by calculating a mean profile $\overline{\boldsymbol{g}}_i$ and a covariance matrix $\mathbf{G}_i$. During the model refinement, a profile vector $\boldsymbol{g}$ is sampled from the target image and its fitness is calculated by a Mahalanobis distance

$$f(\boldsymbol{g}) = (\boldsymbol{g} - \overline{\boldsymbol{g}}_i)^T \mathbf{G}_i^{-1} (\boldsymbol{g} - \overline{\boldsymbol{g}}_i). \tag{2.11}$$

The more suitable landmarks are chosen as the points with minimal Mahalanobis distances.

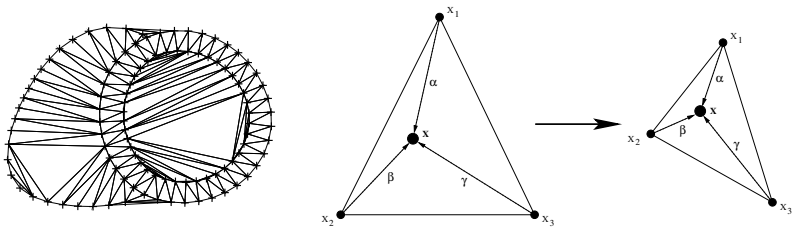
This ASM method has inherent limitations, the intensity information such as the texture pattern of the target object is not included in the model, and only part of such information is utilized in the model refinement step.

## 2.3 Active Appearance Model

### 2.3.1 Statistical Texture Model

The *texture*, the pixel intensity pattern, is another important feature of the object. In order to model the texture, similar to the shape landmarking, the first step is to form a *texture vector* $\boldsymbol{t}$ of pixel intensities. To satisfy the fixed length and point correspondence requirements, the image has to be mapped to a *shape-free* reference frame where the pixel correspondence can be defined. The straightforward but computational expensive approach for such shape-free mapping is to use elastic registration of images.

A more efficient mapping method utilizes the existing shape vectors. A 2-D version of such method is illustrated in Figure 2.5. On the mean shape of the PDM, a pixel grid is defined first and the triangular mesh is built upon the landmarks as shown in Figure 2.5a. Any point in the pixel grid is then uniquely associated with a triangle and has its barycentric coordinate defined. Then the shape-free mapping is performed in the barycentric coordinate system as illustrated in Figure 2.5b.




(a) (b)

Figure 2.5: The 2-D shape-free texture mapping. (a) The mean shape is divided into a set of triangles. (b) Specified by barycentric coordinates $(\alpha, \beta, \gamma)$, $\boldsymbol{x}$ is an unique point within the triangle with various shapes.

Analogous to the shape alignment step, the global intensity variations of brightness and contrast need to be normalized. This iterative texture normalization procedure is listed in Algorithm 2.3 and it produces a mean texture vector $\bar{\boldsymbol{t}}$ of zero-mean and unit-variance.

**Input**: Texture vectors, $\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_M$.
**Output**: Normalized texture vectors, $\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_M$, and a mean texture $\bar{\boldsymbol{t}}$.

Arbitrarily select one texture vector as $\bar{\boldsymbol{t}}$.
**while** $\bar{\boldsymbol{t}}$ *has not converged* **do**
  Normalize $\bar{\boldsymbol{t}}$ as

$$\bar{\boldsymbol{t}} \Leftarrow (\bar{\boldsymbol{t}} - \mu(\bar{\boldsymbol{t}})\mathbf{1})/\sigma(\bar{\boldsymbol{t}}). \tag{2.12}$$

  **for** *each* $\boldsymbol{t}_i$ **do**
    Compute $\alpha$ and $\beta$ by

$$\beta = \mu(\boldsymbol{t}_i), \quad \alpha = \frac{(\boldsymbol{t}_i - \beta\mathbf{1}) \cdot (\boldsymbol{t}_i - \beta\mathbf{1})}{(\boldsymbol{t}_i - \beta\mathbf{1}) \cdot \bar{\boldsymbol{t}}}. \tag{2.13}$$

    Align $\boldsymbol{t}_i$ to $\bar{\boldsymbol{t}}$ by

$$\boldsymbol{t}_i \Leftarrow (\boldsymbol{t}_i - \beta\mathbf{1})/\alpha. \tag{2.14}$$

  **end**
  Compute a new mean texture $\bar{\boldsymbol{t}}$ as

$$\bar{\boldsymbol{t}} = \sum_{i=1}^{M} \boldsymbol{t}_i / M. \tag{2.15}$$

**end**

Algorithm 2.3: Texture normalization

The PCA is applied to the normalized texture vectors and generates a statistical texture model such that any sample texture vector can be approximated by

$$\boldsymbol{t} \approx \bar{\boldsymbol{t}} + \boldsymbol{\Phi}\boldsymbol{b}. \tag{2.16}$$

Figure 2.6 shows the texture variations associated with the first two modes using the short-axis MR image example.

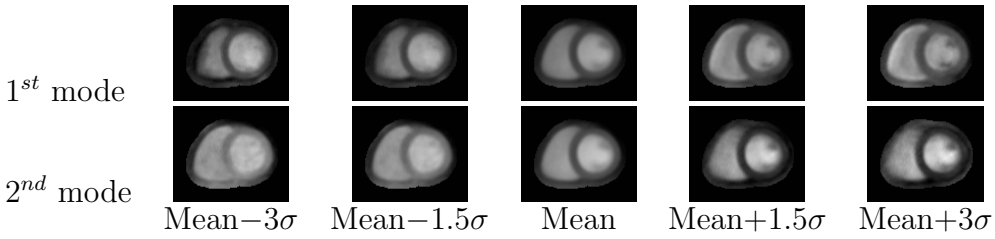



Figure 2.6: The texture variations associated with two strongest PCA modes. The shape is confined to the mean shape.

### 2.3.2 Statistical Appearance Model

The *appearance*, the combination of shape and texture, of the object is modeled by analyzing the appearance vector $\boldsymbol{a}$, the concatenation of shape modal index vector $\boldsymbol{b}_s$ and the texture modal index vector $\boldsymbol{b}_t$ with appropriate weighting [12, 17].

$$\boldsymbol{a} = \begin{bmatrix} \mathbf{W}_s \boldsymbol{b}_s \\ \boldsymbol{b}_t \end{bmatrix} = \begin{bmatrix} \mathbf{W}_s \boldsymbol{\Phi}_s^T (\boldsymbol{s} - \overline{\boldsymbol{s}}) \\ \boldsymbol{\Phi}_t^T (\boldsymbol{t} - \overline{\boldsymbol{t}}) \end{bmatrix}, \quad \mathbf{W}_s = r\mathbf{I}, \quad r^2 = \sum\nolimits_i \lambda_{t_i} \Big/ \sum\nolimits_j \lambda_{s_j}. \tag{2.17}$$

The subscripts $s$ and $t$ indicate the parameters are from the shape and texture models, respectively. The weighting factor $\mathbf{W}_s$ is a diagonal matrix used to compensate the difference in units of the shape and texture models since the the shape model is based on distance and the the texture model is based on intensity.

Applying another PCA to the appearance vector $\boldsymbol{a}$ yields the *Statistical Appear-*

*ance Model* which approximates the appearance of an object by

$$\boldsymbol{a} \approx \boldsymbol{\Phi}_a \boldsymbol{b}_a, \tag{2.18}$$

where the subscript $a$ indicates the parameters of the appearance model. Since the weighting factor $\mathbf{W}_s$ is chosen such that the sum of all $\mathbf{W}_s \boldsymbol{b}_s$ and $\boldsymbol{b}_t$ in the population is zero, the appearance model has no mean appearance. The modal index vector $\boldsymbol{b}_a$ is the compact form of the appearance of the object and it can synthesize the shape and texture of an object by

$$\boldsymbol{\Phi}_a = \begin{bmatrix} \boldsymbol{\Phi}_{a_s} \\ \boldsymbol{\Phi}_{a_t} \end{bmatrix}, \quad \boldsymbol{s} = \overline{\boldsymbol{s}} + \boldsymbol{\Phi}_s \mathbf{W}_s^{-1} \boldsymbol{\Phi}_{a_s} \boldsymbol{b}_a, \quad \boldsymbol{t} = \overline{\boldsymbol{t}} + \boldsymbol{\Phi}_t \boldsymbol{\Phi}_{a_t} \boldsymbol{b}_a \tag{2.19}$$

Figure 2.7 shows the appearance variations associated with two strongest PCA modes using the short-axis MR image example.

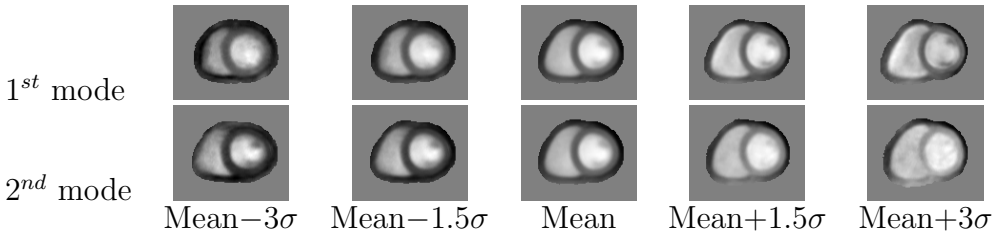



Figure 2.7: The appearance variations associated with two strongest PCA modes. The 'mean' is the combination of the mean shape and the mean texture.

### 2.3.3 Active Appearance Model Segmentation

The statistical appearance model incorporates both shape and texture information learned from manual segmentation and therefore can be used in segmentation, such model-based segmentation method is *Active Appearance Model* (AAM) [12, 17]. The general procedure of AAM segmentation is similar to the ASM segmentation. The mean shape and texture are first fit to the target image, then the appearance model is iteratively refined and PCA modes restricted.

In the model space, there exists the PCA appearance and model synthesized shape and texture of the object as model vectors $\boldsymbol{b}_a$, $\boldsymbol{s}$ and $\boldsymbol{t}$, where $\boldsymbol{s}$ and $\boldsymbol{t}$ are calculated from $\boldsymbol{b}_a$. A transform $\mathcal{T}_s$ such that $\boldsymbol{s}' = \mathcal{T}_s(\boldsymbol{s})$ maps the model shape to target shape $\boldsymbol{s}'$ in the target image space. The target texture $\boldsymbol{t}'$ is sampled based on $\boldsymbol{s}'$ and mapped back to model space by an image to model space mapping $\mathcal{T}_t(\boldsymbol{t}')$. The goal of AAM is to find the optimal $\boldsymbol{b}_a$, $\mathcal{T}_s$ and $\mathcal{T}_t$ that minimize the error metrics $E$ of a texture difference vector $\boldsymbol{r}$ as

$$E = \boldsymbol{r}^T\boldsymbol{r}, \quad \boldsymbol{r} = \boldsymbol{t} - \mathcal{T}_t(\boldsymbol{t}'). \tag{2.20}$$

For each sample in the population, its associated $\mathcal{T}_s$, $\mathcal{T}_t$ and $\boldsymbol{b}_a$ are already optimal after the AAM is constructed, a new $\boldsymbol{r}$ can be computed by displacing these parameters away from the optimal $\boldsymbol{p}$ by $\delta\boldsymbol{p}$ to learn the relation between $\delta\boldsymbol{p}$ and $\boldsymbol{r}$. This learning process is the *AAM training*. The conventional training method is the fixed Jacobian training, where the desired relation is specified by a Jacobian matrix

$\mathbf{R}$ as

$$\delta \boldsymbol{p} = -\mathbf{R}\boldsymbol{r}(\boldsymbol{p}), \quad \mathbf{R} = \left[\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{p}}^T \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{p}}\right]^{-1} \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{p}}^T. \tag{2.21}$$

Although $\mathcal{T}_s$, $\mathcal{T}_t$ and $\boldsymbol{b}_a$ are typically correlated, they are often optimized separately in practice to reduce the computational cost of the training. However, the estimation of $\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{p}}$ is still expensive and is roughly approximated by

$$\frac{\mathrm{d}r_i}{\mathrm{d}p_j} = \sum\nolimits_k w(\delta p_{jk})(r_i(\boldsymbol{p} + \delta p_{jk}) - r_i(\boldsymbol{p})), \tag{2.22}$$

where $w(x)$ is a normalized Gaussian weighting function. The typical choices for maximum parameter displacements are: $0.5\sigma_a$ of $\boldsymbol{b}_a$; 10% of scaling, 3–5 pixels of translation, 10 degree of rotation for $\mathcal{T}_s$; 10% for $\mathcal{T}_t$. The number of values sampled within the displacement range is typically 10 to 20.

The AAM training produces three Jocobians $\mathbf{R}_s$, $\mathbf{R}_t$, and $\mathbf{R}_a$ for model parameters $\mathcal{T}_s$, $\mathcal{T}_t$ and $\boldsymbol{b}_a$, respectively. The AAM segmentation is then a model matching process as listed in Algorithm 2.4. Two examples of applying the AAM segmentation method on the short-axis MR images are shown in Figure 2.8, where the results of five matching iterations are shown.

## 2.4 Hybrid Model-based Segmentation

The AAM is efficient in locating the target object and producing good results, but the local shape accuracy is limited by the lack of shape details contained in the model and the existing multiple local minimums in the model matching. On the other hand, ASM is more flexible than AAM for it does not force texture constraints on the result.

**Input**: The target image, $\mathbf{R}_s$, $\mathbf{R}_t$, and $\mathbf{R}_a$, and the appearance model.
**Output**: The optimal model parameters $\mathcal{T}_s$, $\mathcal{T}_t$ and $\boldsymbol{b}_a$ of the target object and the target shape $\boldsymbol{s}'$.

Set $\boldsymbol{b}_a = 0$. Initialize $\mathcal{T}_s$ and $\mathcal{T}_t$ to place $\boldsymbol{s}'$ approximately on the target image.
Compute $\boldsymbol{r}$ and $E$ using equation (2.20).
**while** *$E$ has not converged* **do**
    Compute $\delta\mathcal{T}_s = -\mathbf{R}_s\boldsymbol{r}$, $\delta\mathcal{T}_t = -\mathbf{R}_t\boldsymbol{r}$, and $\delta\boldsymbol{b}_a = -\mathbf{R}_a\boldsymbol{r}$.
    Set $\boldsymbol{k} = [1, 1.5, 0.5, 0.25, 0.125, \cdots]$, $E' = E$, and $i = 1$
    **while** $E' \geq E$ **do**
        $\mathcal{T}_s' = \mathcal{T}_s + k_i\delta\mathcal{T}_s$, $\mathcal{T}_t' = \mathcal{T}_t + k_i\delta\mathcal{T}_t$, and $\boldsymbol{b}_a' = \boldsymbol{b}_a + k_i\delta\boldsymbol{b}_a$.
        Compute $\boldsymbol{r}'$ and $E'$ from $\mathcal{T}_s'$, $\mathcal{T}_t'$ and $\boldsymbol{b}_a'$.
        **if** $E' < E$ **then**
            $\mathcal{T}_s = \mathcal{T}_s'$, $\mathcal{T}_t = \mathcal{T}_t'$, $\boldsymbol{b}_a = \boldsymbol{b}_a'$, $E = E'$, $\boldsymbol{r} = \boldsymbol{r}'$
        **else**
            Set $i = i + 1$ to test the next value of $\boldsymbol{k}$.
        **end**
    **end**
**end**

Algorithm 2.4: Active Appearance Model segmentation

Its extra freedom in shape deformation often provides the extra momentum needed to bring the AAM model matching out of a local minimum and therefore improves local shape accuracy. A hybrid segmentation approach was proposed by Mitchell et al. [37] and was applied to 2-D segmentation of short-axis MR images. This approach combines the strengths of AAM and ASM and overcomes their limitations. A sample result is shown in Figure 2.9.

In the hybrid approach, the AAM is first performed to produce $\boldsymbol{b}_s^{(A)}$, $\boldsymbol{b}_t^{(A)}$, and $\mathcal{T}_s^{(A)}$. The ASM is then performed using the AAM result as the initialization, it produces $\boldsymbol{b}_s^{(S)}$ and $\mathcal{T}_s^{(S)}$. A hybrid appearance $\boldsymbol{b}_a^{(H)}$ is then computed as

$$\boldsymbol{b}_a^{(H)} = [\boldsymbol{\Phi}_a^{(A)}]^T \begin{bmatrix} \mathbf{W}_s^{(A)}(\rho\boldsymbol{b}_s^{(S)} + (1-\rho)\boldsymbol{b}_s^{(A)}) \\ \boldsymbol{b}_t^{(A)} \end{bmatrix}, \tag{2.23}$$

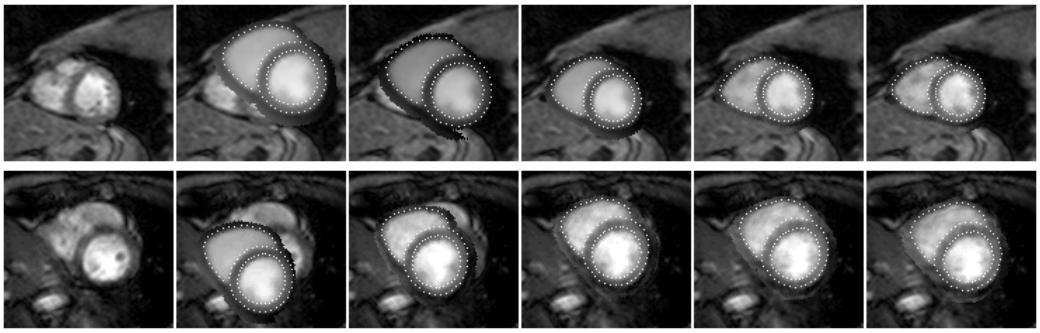
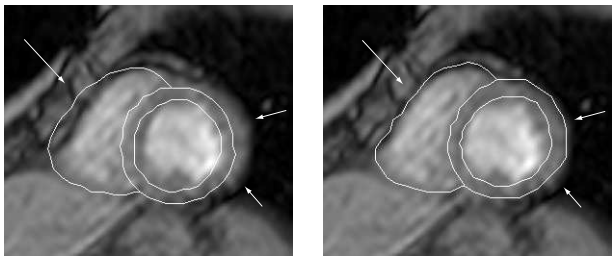


Figure 2.8: Two iteration sequences of 2-D AAM segmentations. The first column is the target images.

The superscript $(A)$ indicates AAM parameters and the superscript $(S)$ indicates ASM parameters. A hybrid shape transform $\mathcal{T}_s^{(H)}$ is

$$\mathcal{T}_s^{(H)} = \rho\mathcal{T}_s^{(S)} + (1-\rho)\mathcal{T}_s^{(A)}. \tag{2.24}$$



(a) (b)

Figure 2.9: Comparison of segmentation accuracy. (a) Conventional AAM. (b)Hybrid AAM/ASM segmentation.

The hybrid method can be easily fit into the AAM procedure listed in Algorithm 2.4, where each iteration of model matching is followed by the extra ASM and hybrid

stage as described in equations (2.23, 2.24). A final AAM iteration is often preferred so the resulting shape is 'valid'.

## 2.5 Challenges of 4-D AAM

The most important requirement of any model-based segmentation is the model must be constructed and trained on a population with high quality ground truths in image as well as manual tracings. As to ventricular MR images, such ground truths are difficult to acquire due to the inherent limitations of the MR images and the manual tracing methods.

The methods and algorithms listed in this chapter can be directly extended to 4-D case once the ventricular shapes of a $P$-phase cardiac cycle is written as a 4-D shape vector $\boldsymbol{s}$ as

$$\boldsymbol{s} = [\boldsymbol{s}_{(1)}^T, \boldsymbol{s}_{(2)}^T, \cdots, \boldsymbol{s}_{(P)}^T]^T, \tag{2.25}$$

where each $\boldsymbol{s}_{(i)}$ is a 3-D shape vector as defined in equation (2.1). But creating such 4-D shape vector manually is almost impossible, a good automated landmarking algorithm is the key of 4-D AAM construction.

Increasing the image space from 2-D or 3-D to 4-D dramatically increases the dimensionality of shape and texture vectors and the ASM and AAM will suffer from the 'curse of dimensionality' in various aspects.

- To model the statistical properties of shape and texture vectors in a high-dimensional space, large population is always preferred. But large population of 4-D data, especially accurate manual tracings, are very difficult to acquire.

- Due to the increased computational cost of error metrics $E$, including all possible values of model parameters in AAM training is impossible. Further simplification, assuming the elements of model parameter are uncorrelated, is unavoidable. Such simplification can degrade the model training quality.
- Also due to the high cost of computing $E$, the 4-D AAM often cannot be performed with as many as matching iterations as 2-D and 3-D AAMs.
- In 4-D AAM, more model parameters are to be optimized therefore there are more local minimums in the parameter search space. The possibility of the model matching stuck at a local minimum is increased.

# CHAPTER 3
# 4-D DATA CONSTRUCTION

## 3.1 4-D Image Construction

### 3.1.1 Image Data Management

The MRI scanners output acquired images as DICOM files that contain the image data and various information associated with the image. A typical DICOM file contains:

- The 2-D gray scale image data with 16-bit intensity resolution.
- Patient information such as name, hospital ID, sex, date of birth (anonymized in the DICOM files used for this study ).
- MR sequence description added by the operator such as 'true FISP short-axis'. The exact information in this description is generally not well defined and varies.
- Geometric and timing information indicating where and when a 2-D slice image is acquired. A slice is defined in the absolute MR scanner coordinate system and is identified by: the coordinates of the upper-left corner, two unit vectors indicating the directions of image axes, image width and height measured by pixels, physical size of the pixels, the physical slice thickness, and a time offset indicating the acquisition time with respect to the starting of a cardiac cycle – the peak of ECG R-wave.

The number of DICOM files included in a complete 4-D MR study is very large. Each imaging sequence contains 15–30 phases per cardiac cycle, with 5–15 slices per phase. Besides the long- and short-axis sequences, other sequences such as 3-D

localizers are also included. Although each DICOM file contains all the necessary information for data reconstruction, searching through all the files for small piece of information is expensive. To efficiently manage the 4-D MR data, the following data management steps were performed.

1. Extract non-image information from all DICOM files. Sort and organize them in a hierarchy fashion such that a complete study contains a four-level tree from root to leaf as *study* $\hookrightarrow$ *sequence* $\hookrightarrow$ *phase* $\hookrightarrow$ *slice*. Record the tree structure into a XML file.
2. Manually check and modify the XML files if necessary. The manual modification often involves assigning unique and consistent sequence descriptions. Some operations such as merging or splitting some imperfectly described or acquired data sequences are automated.
3. Based on the hierarchy data description, extract 2-D images from DICOM files and combine them into series of 4-D ANALYZE images.
4. Write geometrics and timing information to customized ASCII data information files, one for each 4-D image.

The DICOM standard is complicated and the exact information stored is scanner dependent. The data management steps standardize the data format and file naming scheme, remove dependency on DICOM standard and correct any error or inconsistency in data description. Although the manual inspection and modification is a little tedious, it only needs to be performed once. Most importantly, the number of files used for a 4-D sequence is reduced from hundreds to three (ANALYZE header and

image and a ASCII information file), which is more manageable.

### 3.1.2 Image Interpolation and Rotation

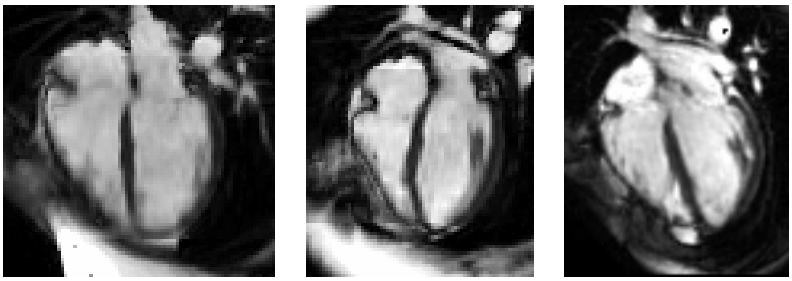
The next step is to construct 4-D images of isotropic voxels and fixed number of phases per cardiac cycle by interpolation. The nearest-neighbor method is used in temporal interpolation for two reasons. First, the MR protocol only guarantees that the acquisition of the first phase is triggered by the ECG R-peak and approximately 90–100% of the cardiac cycle is covered by the sequence. Second, any higher order temporal interpolation will introduce blurring of images.

After temporal interpolation, spatial interpolation is performed to produce isotropic voxels in each 3-D image. The spatial interpolation of ventricular MR images is mainly preformed along the $Z$ axis that is perpendicular to the $XY$ imaging plane. The *shifted linear interpolation* [7] is used due to its two advantages: 1) the computational cost is close to that of the standard linear interpolation; 2) the SNR is close to or higher than that of the cubic interpolation and the image edges are better preserved. It also agrees to the fact that any $Z$ slice that is not adjacent to the sampling point should have no contribution to the interpolation.

Image rotation is performed together with spatial interpolation. Based on the imaging plane orientations, the MR images are rotated such that in the resulting 3-D images built from short(long)-axis images, the $XY$ plane is parallel to the short(long)-axis imaging plane and the $XZ$ plane is parallel to the long(short)-axis imaging plane.

### 3.1.3 Imperfections of Ventricular MR Images

Figures 3.1 and 3.2 show several example 2-D slices of the interpolated and rotated 3-D images. The image rotation helps us to see the imperfections of ventricular MR imaging protocol. An accurate estimated LV long-axis direction corresponds to the image shown in Figure 3.1a, where the LV long-axis is approximately vertical. Inaccurate LV long-axis direction estimations produce images shown in Figures 3.1b and 3.1c.

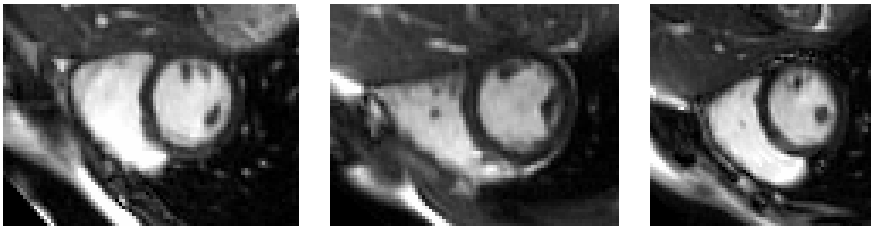


(a) (b) (c)

Figure 3.1: Various LV long-axis direction estimations.

If the orientation of the long-axis imaging plane is correctly estimated, the resulting short-axis should look like Figure 3.2a, where the RV is exactly on the left side of the LV. Inaccurate long-axis imaging plane estimations make their associated short-axis images look like Figures 3.2b and 3.2c, where the RVs are rotated.

The orientation of the heart plays a vital role in identifying the ventricular bases and apexes in a 3-D context. The MR protocol imperfections make the manual tracing in 3-D more difficult.



(a) (b) (c)

Figure 3.2: Effects of various long-axis imaging plane orientation estimations.

### 3.1.4 Motion Correction and Image Fusion

After image interpolation and rotation, a 4-D short-axis (SA) image is reconstructed from the original short-axis MR images and a 4-D long-axis (LA) image from the original long-axis MR images. Two typical orthogonal slices of such reconstructed 4-D images at end-diastole are shown in Figures 3.3a and 3.3b. The existing respiratory motion artifacts are obvious in the bottom image of Figure 3.3a, where the motion artifacts are reflected as mis-alignments of ventricles in the short-axis slices, which agrees to the result of a study by Dornier et al. [21]. If the short-axis MR images are acquired with accurate LV long-axis direction estimations, the motion artifacts can be corrected by aligning the centroids of LV in the short-axis slices so they form a line parallel to the $Z$ axis of the SA image [44]. When the MR protocol imperfections exist, applying this simple motion correction approach can potentially change the ventricular shapes.

The long-axis view provides complementary information of the ventricles and was utilized in this study to remove the dependency of accurate imaging plane orientations. The motion correction was achieved by registering (translation only) corre-

sponding slices of SA and LA image. The top image of Figure 3.3a was aligned to the top image of Figure 3.3b, then the bottom image of Figure 3.3b was aligned to the bottom image of Figure 3.3a. The resulting images after motion correction are shown in Figures 3.3c and 3.3d.

Due to the MR protocol variability and scan time constraints, the short-axis view images may not cover the ventricles completely. The image quality of $XZ$ slices (bottom image of Figure 3.3c) is not as good as the corresponding slices from the long-axis view (bottom image of Figure 3.3d) due to interpolation. As the final step, the motion corrected SA and LA images are fused into a single 4-D image as shown in Figure 3.3e, which has a better ventricle coverage as well as an improved image quality.

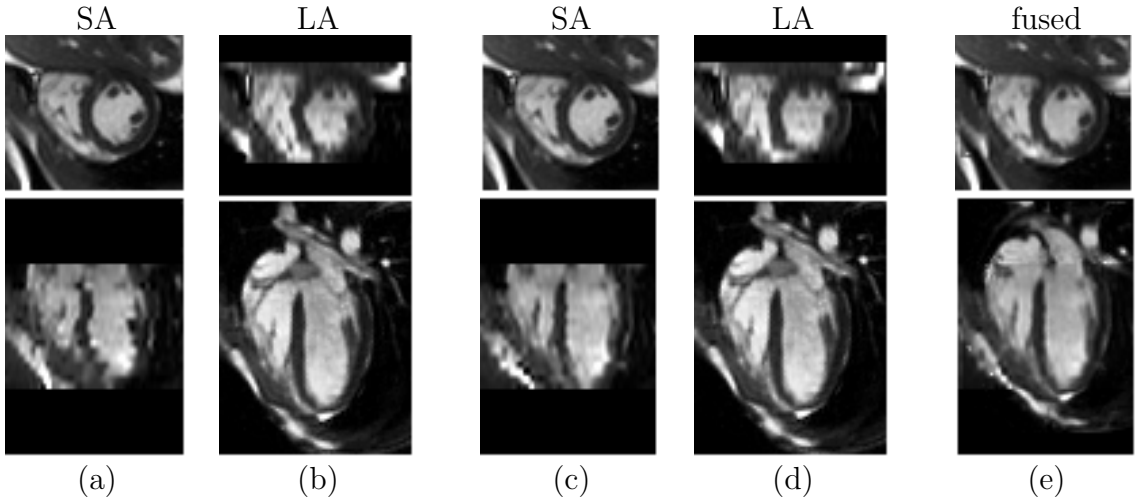
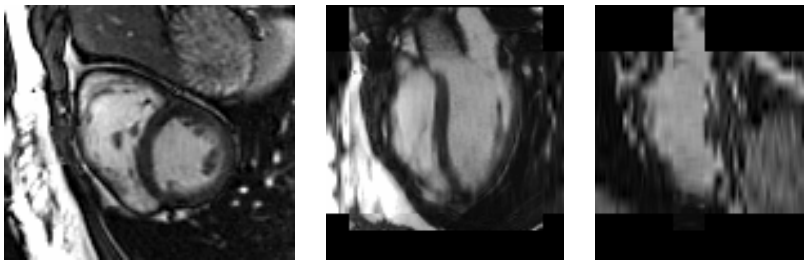



Figure 3.3: Motion correction and image fusion. Orthogonal slices of (a) reconstructed SA image, (b) reconstructed LA image, (c) SA image after motion correction, (d) LA image after motion correction, and (e) the final fused image.

However, the motion correction relies on the coverage of the long-axis view. If only a small number (two or three) of long-axis images are available, only a small portion of the ventricles is included on the short-axis slices of the reconstructed LA images. In this case, the motion correction was still able to produce a smooth ventricular shape on the $XZ$ slice similar to the bottom image of Figure 3.3d. But the long-axis images cannot provide enough information to correct any possible motion along the $Y$ axis of the SA image. Without two-chamber images acquired from planes perpendicular to both short- and long-axis planes, whether the motion correction is successful cannot be fully assessed. Such a 'motion corrected' and fused image is shown in Figure 3.4.



(a) $XY$ slice (b) $XZ$ slice (c) $YZ$ slice

Figure 3.4: The 'motion corrected' and fused image when only tow or three long-axis MR images are available.

The heart only occupies a small portion of the standard ventricular MR image and is the only moving organ. Before the motion correction, a rectangle region of interest (ROI) for the heart must be specified to exclude the static part of image such as chest wall and body fat. Although the ROI often has to be specified manually, it only need to be drawn on a single slice of the first phase. This ROI reduces the size of

the fused image and computational cost associated with all the data processing steps after it.

Before the image fusion, the overall intensity patterns of the SA and LA images are adjusted by histogram matching, some discontinuity of intensity pattern in the result is inevitable but are often subtle and has little influence on the modeling of texture and AAM segmentation.

### 3.1.5 Subjective Quality Assessment

A total of 127 ventricular MR scans were used, each containing long- and short-axis images. Visual inspection was used to identify the existence of MR protocol imperfections. It was found that 33 scans have inaccurate LV long-axis directions (see Figure 3.1), and 58 scans have inaccurate long-axis imaging planes (see Figure 3.2).

The existence of motion artifacts and the performance of motion correction were also assessed by visual inspection. Although quantitative evaluation of motion correction was reported in [44], such approach is operator-dependent and time-consuming.

Among the 127 MR scans tested, visual assessment found 26 scans with visible motion artifacts similar to Figure 3.3a, and 5 scans have even larger motions. Of the 31 scans with motion artifacts, 21 scans have enough long-axis view images so the motion correction and image fusion performed very well and produced fused 4-D images with similar quality as shown in Figure 3.3e. On the 96 scans without visible motion artifacts, the motion correction was also performed by default and did not introduce any unnecessary motion artifacts.

However, 10 scans with motion artifacts reflected as shifting of ventricles along the $X$ axis only have two or three long-axis images available. The motion correction successfully produced smooth ventricular borders on the $XZ$ slices. Due to the limited coverage of long-axis images, the performance of motion correction on the $YZ$ slices cannot be assessed. An example of such resulting image is shown in Figure 3.4.

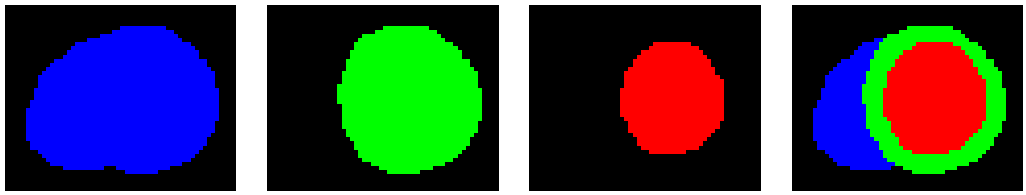
## 3.2 4-D Volume Construction

### 3.2.1 Conventional Manual Tracing

In this study, the *conventional manual tracing* is especially used to refer to the clinical approach of ventricular MR image segmentation. The ventricular borders are manually traced by experts on original 2-D MR images as contours. Although both long- and short-axis images can be traced, the correlation between them as point-to-point correspondence is often not available or very difficult to utilize for human experts. They can only rely on limited spatial and temporal context information when switching images from different slices and phases back and forth. The limitations of this conventional approach was discussed in detail in Chapter 1.

In this study, three ventricular borders were traced: the LV epi- and endocardial borders and the RV epicardial border. The two LV borders were always defined as closed contours. The manual tracing utilizes Catmull-Rom splines [8] so the operator only needs to define and modify the control points of splines without drawing the border pixel by pixel. The 2-D ventricular shape is represented by boundary pixels or labeled regions. The translation between them is achieved by boundary extraction or flood-fill and overlapping. Figure 3.5 shows the process of region overlapping. The

3-D ventricular labeled volume is created by stacking tracings from multiple slices together, but it still has *anisotropic* voxels.



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

Figure 3.5: Overlapping the labeled images of (a) the union of the LV and RV, (b) the whole LV, and (c) the LV blood pool creates (d) the multi-object labeled image whose labels are RV, LV myocardium, and LV blood pool.

### 3.2.2 Shape Interpolation

Similar to image interpolation, the anisotropic labeled volume needs to be interpolated so the true 3-D ventricular shape can be created. Unlike image intensity, the labels cannot be directly interpolated. Instead, the shape interpolation is achieved in the domain of distance maps as proposed by Raya and Udupa [41]. A distance transform $\mathcal{DT}$ is defined as

$$\mathbf{D}^{(i)} = \mathcal{DT}(\mathbf{L}^{(i)}), \tag{3.1}$$

where $\mathbf{L}$ is the input labeled image, $\mathbf{D}$ is the output distance map. The superscript $(i)$ indicates that only the volume labeled by $i$ is considered. In the resulting distance map $\mathbf{D}$, each location is assigned to a value indicates its distance to the surface of the volume of interest. Positive values are arbitrarily assigned to exterior locations and

negative values to interior locations. The brute-force distance transform is expensive with complexity of $O(N^2)$. A linear time algorithm proposed by Maurer [35] that utilizes Voronoi diagram was used in this study.

After performing interpolation on the distance map, a new distance map with isotropic voxels is created. Such distance map can be translated back to labeled volume by an inverse distance transform $\mathcal{DT}^{-1}$ as

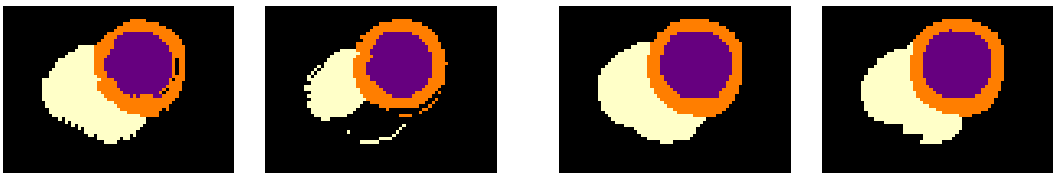
$$\mathbf{L}^{(i)} = \mathcal{DT}^{-1}(\mathbf{D}^{(i)}), \tag{3.2}$$

where all voxels with negative distances are labeled with $i$ and others as background. The inverse transform is actually achieved by thresholding, where a zero threshold is used in equation (3.2). If non-zero threshold is used, it becomes surface growing or shrinking.

The typical slice distance of ventricular MR is 4 to 5 times that of the image pixel size. The shape interpolation in 3-D can introduce another artifact as shown in Figure 3.6a where the 2-D ventricular shape on a interpolated slice is not a single connected region or a 'valid' ventricle. In order to create 'valid' 3-D ventricular shapes, an additional 3-D surface smoothing step was performed. The resulting 2-D slices of the smoothed ventricles are shown in Figure 3.6b.

### 3.2.3 Volume Fusion by Shape Estimation

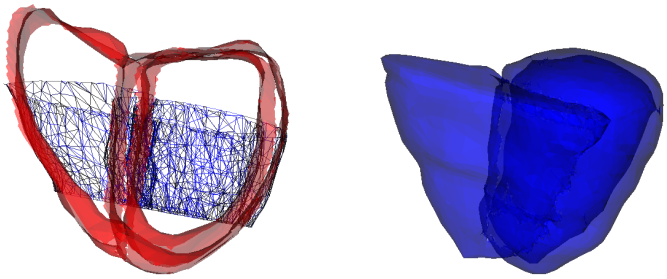
After shape interpolation, two sets of ventricular labeled volumes are created from the long- and short-axis views, respectively. An example of such volumes is shown



(a) (b)

Figure 3.6: Creating 'valid' ventricles. (a) Artifacts due to shape interpolation. (b) After 3-D surface smoothing, 'valid' ventricular shapes are created.

in Figure 3.7a, where neither covers the ventricles completely due to the limitations of conventional manual tracing. An additional volume fusion step was performed to estimate the shapes of the ventricular bases and apexes and create a set of labeled volumes that cover the ventricles completely as shown in Figure 3.7b. Two types of shape estimation approaches were tested: *scaling-and-shifting*, and *cookie-cutter*.



(a) (b)

Figure 3.7: An example of volume fusion. (a) The surfaces of ventricles defined by the long-axis and short-axis view labeled volumes. (b) The surfaces of the complete ventricles after shape estimation.

The scaling-and-shifting approach assumes that in the base or apex region not covered by the short-axis tracing, the short-axis slice ventricular shape is a scaled

and shifted version of the last traced short-axis base or apex slice. The scaling and shifting parameters are estimated from the long-axis view labeled volume. In the cookie-cutter approach, the last traced base and apex short-axis slices are first extended to infinity without scaling and shifting. Then the long-axis slice with the largest ventricular areas is used as a 'cookie cutter' to remove the extra 'dough'. In both approaches, the base of the RV is arbitrarily capped at the location where the LV and RV begins to separate in the long-axis view.

In the example shown in Figure 3.7, both shape estimation approaches created a fused volume similar to Figure 3.7b. When tested on larger population, visual inspection showed different resulting shapes and both approaches have their pros and cons. Some extra quantitative analyses of the shape estimation approaches showed a common weakness of both approaches: they are not robust enough to guarantee the overall correctness and consistency of the fused ventricular volumes. The main problem is only a few (two or three) long-axis images can be accurately traced by the experts, thus their contribution to the shape estimation is limited.

### 3.2.4 True 4-D Manual Tracing

Although it creates approximate ventricular volumes that are better than conventional manual tracing in coverage, the volume fusion can not provide high quality independent standards for model construction and segmentation validation. A new manual tracing application was developed to perform the manual tracing in a true 4-D context. A screenshot of this application is shown in Figure 3.8.

The input of this tracing application is the 4-D image constructed by motion

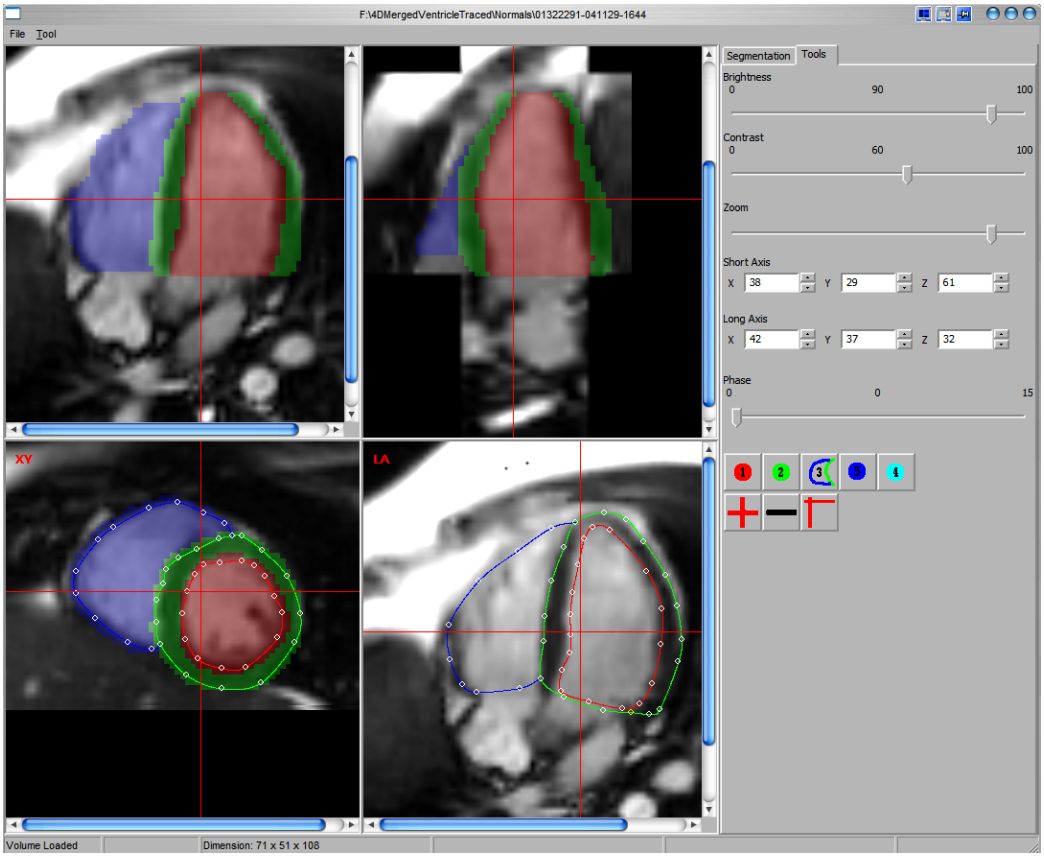



Figure 3.8: A screenshot of 4-D manual tracing application.

correction and image fusion. For convenience, the $XY$ slices of this 4-D image are parallel to the short-axis imaging plane, $XZ$ slices are parallel to the long-axis imaging plane. The three orthogonal views of the image are displayed in the top-row panels and the bottom-left panel where the cross hair indicates the locations of a 3-D point in all three orthogonal slices. As a reference, the 4-D image is also defined in another orientation whose $XY$ slices are parallel to the long-axis imaging plane as shown in the bottom right panel also with the point location indicated the by the cross hair. This four-panel and cross hair display put the user in a true 4-D context of the heart

where 3-D and 4-D navigation is unambiguous.

In this 4-D environment, any slice in the bottom-row panels can be traced – not limited to the ones that correspond to original MR images – so large shape changes can be accurately captured by densely traced slices. Since it is more difficult to trace large numbers of long-axis images, the tracings on them were only used as references to assist accurate short-axis tracing and were not used in shape interpolation. After the 4-D manual tracing, the shape interpolation was performed using only short-axis tracings. The resulting ventricular volumes can be loaded by the tracing application for visual inspection as colored regions in Figure 3.8. In addition, this tracing application can be used as a generic 4-D image viewer with volume overlay capability, a very convenient feature for visual inspection of 4-D segmentation.

### 3.2.5 Quantitative Quality Assessment

Due to its inherent incomplete and inconsistent ventricle coverage, the conventional tracing cannot be used as ground truth to assess the quality of 4-D manual tracing. Without measurements from other imaging modalities, the quality of 4-D tracing cannot be directly assessed. Fortunately, it has been shown that for normal subjects the left ventricle stroke volume (LVSV) is about 5–6% larger than the right ventricle stroke volume (RVSV) [4, 26], the agreement of LVSV and RVSV of normal subjects was thus used as indirect tracing quality metrics. The agreement was assessed by the stroke volume ratio as $LVSV/RVSV$ and zero-intercept linear regression of RVSV versus LVSV. A total of 25 MR scans of normal subjects were analyzed and three sets of ventricular volumes were measured as follows.

- *Conventional*: The conventional manual tracings were performed by an expert and verified by another senior expert and consensus of the experts was reached. Motion correction was performed and traced ventricular borders were translated accordingly. The ventricular volumes were then calculated from the results of shape interpolation of the short-axis tracings.
- *Fused*: Starting from the conventional tracings, the long-axis tracings were used for shape estimation and volume fusion. The cookie-cutter approach was used since visual inspections showed it preserves the consistency of ventricle coverage better than the scaling-and-shifting approach.
- *True 4-D*: The conventional tracings were loaded into the 4-D tracing program and improved in the 4-D context. The resulting short-axis tracings were used in shape interpolation.

Table 3.1 lists the analysis results of stroke volume agreement. The conventional method did not produce good stroke volume agreement. The volume fusion achieved good agreement in average but still has a large standard deviation and a poor correlation. The stroke volume ratios of the true 4-D results agree to those reported in literatures [4, 26] with small standard deviation. The correlation of the 4-D results is much improved and suggests a very good stroke volume agreement.

| | LVSV/RVSV | | | linear regression | |
|---|---|---|---|---|---|
| | mean±SD | min | max | slope | R |
| Conventional | 1.20±0.29 | 0.67 | 2.07 | 0.87 | 0.67 |
| Fused | 1.05±0.18 | 0.62 | 1.47 | 0.97 | 0.63 |
| True 4-D | 1.08±0.08 | 0.94 | 1.26 | 0.93 | 0.94 |

Table 3.1: Comparison of stroke volume agreements.

To assess the improvement in ventricle coverages, the end-diastolic volume (EDV) and end-systolic volume (ESV) of LV and RV were measured from the true 4-D tracings and were compared with those from the conventional method and calculated as percentage of increase. The analysis results are listed in Table 3.2, where the results of 25 TOF patient scans are also included. Since more ventricular base slices were identified and traced by the 4-D manual tracings, large part of ventricular volumes that were missed or sacrificed for consistency by the conventional tracing method was recovered, especially for TOF patients and at end-systoles of LV and RV.

| | Normals | | | TOFs | | |
|---|---|---|---|---|---|---|
| | mean±SD | min | max | mean±SD | min | max |
| LVEDV | 31%±14% | 8% | 72% | 52%±23% | 13% | 113% |
| LVESV | 52%±25% | 24% | 116% | 90%±42% | 22% | 194% |
| RVEDV | 31%±19% | 1% | 82% | 65%±51% | 15% | 246% |
| RVESV | 34%±24% | 2% | 95% | 94%±73% | 16% | 366% |

Table 3.2: Percentage of ventricular volume increases of 4-D method compared with conventional tracing.

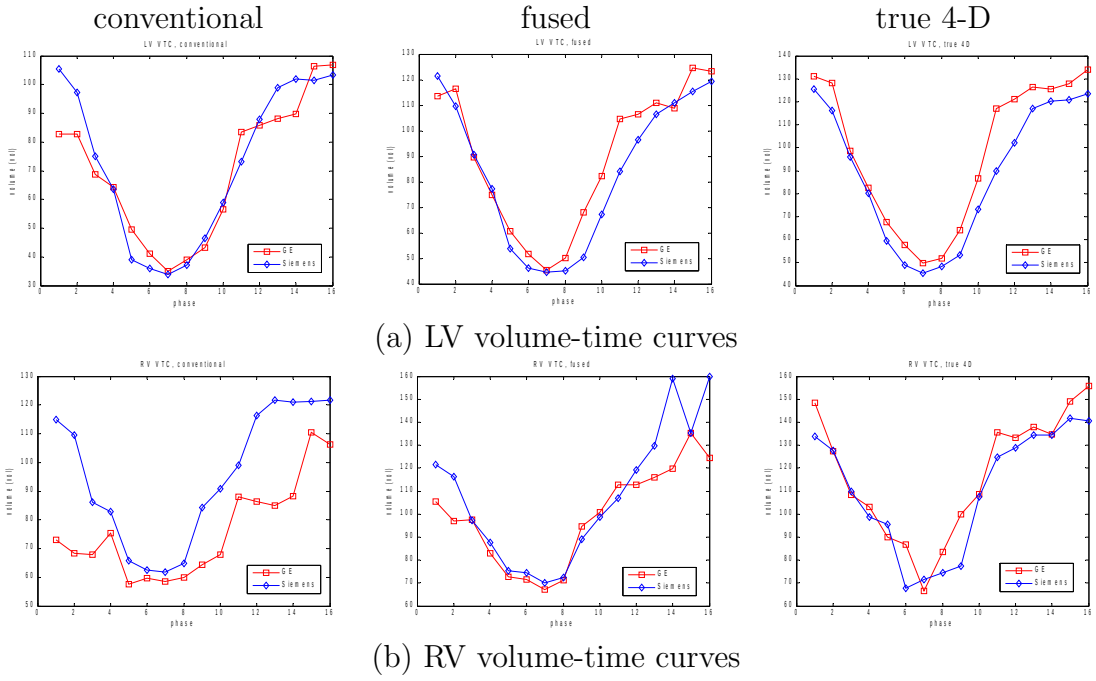
One normal subject has two same-day scans first by a GE Signa scanner with voxel size of 1.72×1.72×6 mm, then by a Siemens Avanto scanner with voxel size of 2.08×2.08×6 mm. The ventricular volumes measured from the true 4-D tracings are listed in Table 3.3. The true 4-D method produced consistent volumetric measurements in spite of difference in scanner and imaging resolution.

To further assess the performance of different volume construction methods, the volume-time curves (VTC) of LV and RV of the two same-day normal subject scans are plotted for all three methods in Figure 3.9. The true 4-D method achieved not

| True 4-D | LV | | | | RV | | | |
|---|---|---|---|---|---|---|---|---|
| | EDV | ESV | SV | EF | EDV | ESV | SV | EF |
| GE scan | 134 | 50 | 84 | 63% | 156 | 67 | 89 | 57% |
| Siemens scan | 126 | 45 | 80 | 64% | 142 | 68 | 74 | 52% |

Table 3.3: Comparison of ventricular volumes (ml) measured by true 4-D method from two same-day scans of a normal subject.

only the best volume-time consistency but also the smoothest curves, which agree to the continuity nature of ventricular contraction and relaxation.




(a) LV volume-time curves

(b) RV volume-time curves

Figure 3.9: Comparison of volume-time curves of same-day scans of a normal subject.

## 3.3 Summary

The 4-D ground truths of ventricular images and volumes were constructed and thus provide reliable training data for the next stage, building the 4-D AAM. The

available long- and short-axis images were fully utilized to achieve automated motion correction, image fusion, accurate 4-D manual tracing, and shape interpolation. Both subjective and objective assessments showed that most of the inherent limitations of MR imaging protocols and conventional analysis were overcome by these techniques.

The conventional manual tracing is time-consuming and the true 4-D tracing is more tedious even with the help of the developed application. Our failed attempts of volume fusion proved that there is no shortcut for accurate volume construction. The quantitative assessments of true 4-D tracing quality was only applicable to scans of normal subjects. For TOF patients with unknown disease stages and RV regurgitation, the relationship between stroke volumes of LV and RV is unknown and therefore cannot be used for quantitative assessment. Fortunately, the same set of rules were used for normal and TOF tracings and the visual inspection of the smoothness of the volume-time curves can also provide some assistance to the manual tracing. By visual inspection, the true 4-D tracings of TOF showed similar quality as those of normal subjects.

# CHAPTER 4
# 4-D MODEL CONSTRUCTION

## 4.1 Automated Landmark Generation

### 4.1.1 Landmarking Strategies

To represent the 3-D shape of the LV surfaces, a brute-force approach can be used [36]. A fixed number of short-axis slices are sampled first. On each slice, the ventricular borders are represented by points defined in the polar coordinate system with equal angle increment. Then a 3-D mesh surface is constructed such that the mesh triangles are built from points from two neighboring short-axis slices. The number and distribution of short-axis slices have to be carefully chosen to make sure that any large shape change from slice to slice is captured. This approach works very well when the borders of the object on the arbitrarily chosen slices are roughly circular, but applying this approach to the RV surface is rather difficult.

A previously proposed general-purpose approach is to use minimum description length (MDL) [19]. The point correspondence is built by optimal encoding: an MDL criterion to evaluate the quality of correspondence in terms of the ability to encode the whole training set for a given landmark distribution. The MDL encoded models optimize model properties such as compactness and specificity. The major drawback of this approach is its high computational cost. As reported in [19], creating landmarks for less than 50 2-D shapes took several hours to finish. It makes applying such approach to 4-D shapes almost impossible.

Another way to formulate the landmarking problem is to use registration together

with a *template*. Such a template-based approach was proposed by Frangi et al. [25] and a slightly modified (for efficiency) version is shown in Figure 4.1 using 2-D shapes for illustration. The whole process contains the following three steps.

1. *Template creation*: The sample shapes of the population are aligned to a common reference space using a set of affine transforms $\mathcal{T}_{affine}$ that are restricted to rotation, translation, and scaling. The aligned shapes are then blended (averaged) to a template shape.
2. *Landmarking the template*: Create landmarks on the template shape. Then apply individual inverse transform $\mathcal{T}_{affine}^{-1}$ to map the landmarks back to the sample space (shown as transformed landmarks in Figure 4.1) so they are very close to the original sample shapes.
3. *Landmark propagation*: Apply a set of elastic transforms $\mathcal{T}_{elastic}$ to propagate the transformed landmarks onto the the borders (surfaces) of the original sample shapes.

In this study, the actual implementations of the three steps of the landmarking procedure are slightly different from those proposed by Frangi et al. [25]. They were custom designed for the 4-D ventricle case where the computational cost is the major concern.

Because the goal of shape alignment is to make the transformed landmarks as close to the sample shape as possible, using *anisotropic scaling* in the affine transform is a more appropriate choice. Note that another shape aligning step – or more precisely the Procrustes Analysis – will be performed independently again later to build the

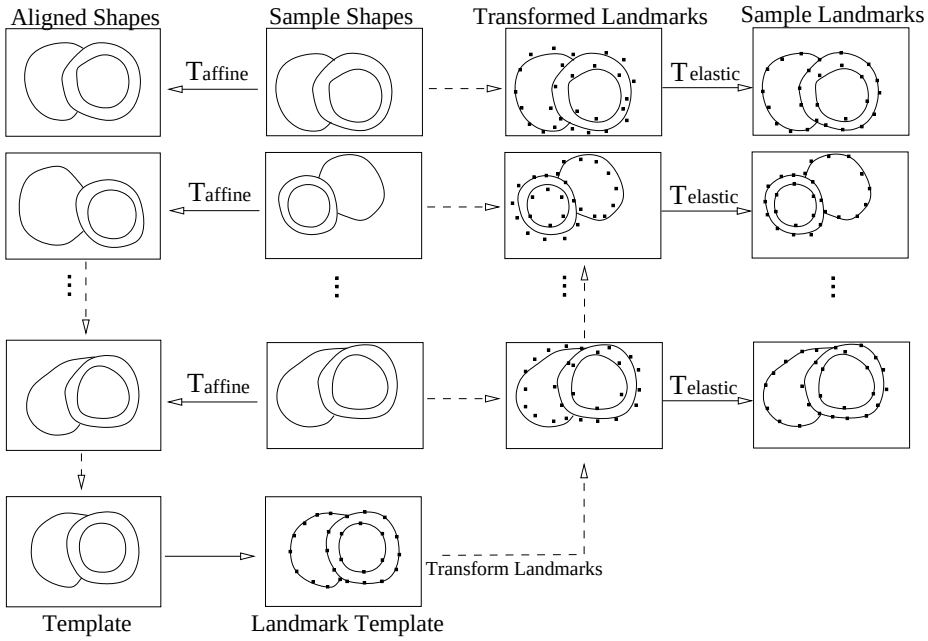



Figure 4.1: Flowchart of automated landmark generation.

model.

By intuition, if the sample population contains $M$ 4-D samples (each has 16 3-D shapes), Figure 4.1 should have $M$ sample shapes and the resulting template is 4-D. But when the 4-D shapes have embedded motion patterns, the 4-D alignment cannot remove the effects of the motion completely and the resulting 4-D template could be 'invalid' – possible concaved ventricular base and apex. In this study, the automated landmarking was actually performed in 3-D by treating the sample population as $16M$ 3-D sample shapes so a single 'valid' 3-D template was created. An extra benefit of this approach is the same number of landmarks with the same surface topology are used for all cardiac phases.

### 4.1.2 Template Creation

Starting from the sample shapes represented by labeled volumes, $\mathbf{L}_1, \mathbf{L}_2, \cdots, \mathbf{L}_M$, the iterative procedure that creates the template labeled volume $\overline{\mathbf{L}}$ is very similar to Algorithm 2.1. It produces aligned labeled volumes, $\mathbf{L}'_1, \mathbf{L}'_2, \cdots, \mathbf{L}'_M$, they can be blend into a template labeled volume $\overline{\mathbf{L}}$ using distance transform as

$$\overline{\mathbf{L}} = \mathcal{DT}^{-1}\left[\frac{1}{M}\sum\nolimits_{i=1}^{M}\mathcal{DT}(\mathbf{L}'_i)\right]. \tag{4.1}$$

The transform $\mathcal{T}^{(i)}_{affine}$ aligns sample labeled volume $\mathbf{L}_i$ to the template $\overline{\mathbf{L}}$ such that the difference between $\mathbf{L}'_i = \mathcal{T}^{(i)}_{affine}(\mathbf{L}_i)$ and $\overline{\mathbf{L}}$ is minimized. In this study, this transform was found by iterative closest point (ICP) registration [5, 9]. Since the distance transform is computed by the Voronoi diagram based algorithm [35], the closest points to the surface of $\mathbf{L}'_i$ can be easily found on the Voronoi diagram.

The ICP registration requires a good initialization of the transform. When the LV and RV are treated as a single object and landmarked together, such initialization is trivial. Due to its unique shape, when only the LV is landmarked and modeled, its rotation must be initialized in the context of the RV.

### 4.1.3 Landmarking the Template

The created template labeled volume is translated into another shape representation – triangular mesh surface. The most commonly used algorithm to extract the object mesh surface is the *marching cubes* [34], it takes 3-D distance map as input and produces iso-surface of the object consists of vertices and triangles. In this study,

a variant of the marching cubes – the *marching tetrahedrons* was implemented to directly use labeled volume with multiple labels as input without distance transform calculation. The basic idea of marching tetrahedrons is illustrated in Figure 4.2, where the situation of four tetrahedron vertices belong to four different objects is treated as a 'invalid' case.

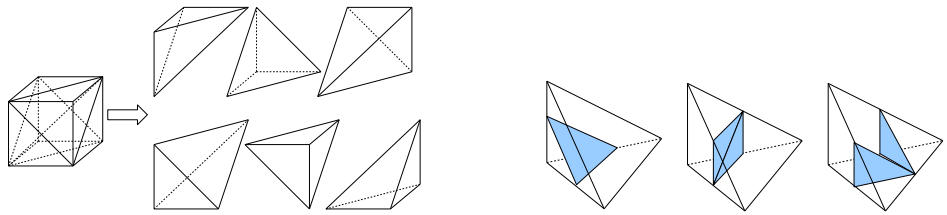
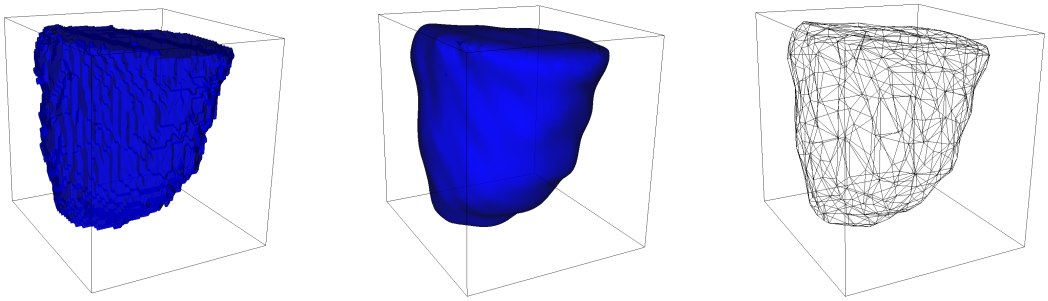



Figure 4.2: Marching tetrahedrons. (a) Dividing a cube into six tetrahedrons. (b) Creating surface triangles (shadowed triangles) based on different combinations of vertex labels. Left and middle: the four vertices belong to two objects. Right: the four vertices belong to three objects.

The customized marching tetrahedrons algorithm does not produce conventional iso-surface, so the shapes of the voxels are kept as shown in Figure 4.3a, where the staircase-like surface is not a true representation of the ventricular shape. The surface smoothing filter implemented in VTK [43], *vtkWindowedSincPolyDataFilter*, was then applied. It adjusts the vertex coordinates using a windowed sinc function interpolation kernel to relax the mesh [47] and can be thought as a special designed low-pass filter that eliminates the high frequency components – the staircases. Although the locations of vertices are modified during the smoothing, their movements are restricted

within a very small range. Figure 4.3b shows the resulting smoothed surface. Note this surface smoothing algorithm is also used to create 'valid' ventricular shapes after shape interpolation.

The smoothed surface still consists of very large number of vertices and triangles. To form the template landmarks, the number of vertices must be reduced and such a algorithm is the *surface decimation*. In this study, a VTK filter *vtkQuadricDecimation* was used. It uses the quadric error measurement as the error metric in eliminating vertices and triangles. The resulting decimated surface often have 'evenly' distributed vertices and triangles of similar sizes, which meet the requirements of a good set of landmarks. Further details of this algorithm can be found in [27, 31]. Figure 4.3c shows the resulting surface that consists of 302 vertices and 600 triangles.



(a) (b) (c)

Figure 4.3: Examples of LV epicardial template surfaces: (a) after marching tetrahedrons, (b) after surface smoothing, and (c) after surface decimation.

Before the details of landmark propagation are discussed, the different representations of the 3-D ventricular shape are summarized here. In the sample image space,

the 3-D ventricular shape can be defined by: a set of 2-D contours, labeled volume, distance map, a dense triangular mesh, and the landmarks that form the shape vector and also a sparse mesh when surface triangles topology is defined. In the statistical model space, the ventricular shape is defined by landmarks or the associated PCA modes. In general, those shape representations are interchangeable and the translation from one representation to another is always feasible without losing important shape features although losing some shape details is inevitable. Figure 4.4 shows all the 3-D shape representations and their relationships. Note that the extension to 4-D case is very simple – a 4-D shape is a set of 3-D shapes from all cardiac phases.

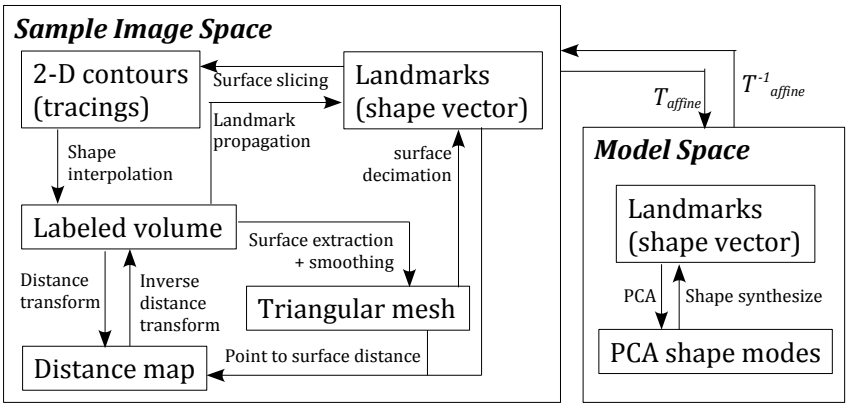



Figure 4.4: 3-D shape representations

### 4.1.4 Landmark Propagation

The landmark propagation is formulated as an elastic registration problem, which can be solved by many different ways all with various choices of cost functions to

minimize. Most of the time, the ground truth of such elastic registration is unavailable and the algorithm can be very expensive. In the 4-D ventricular case, reducing the computational cost is a major concern and by utilizing anisotropic scaling in template creation the transformed landmarks are reasonably close to the target sample surface, so an efficient method was designed. The $N$ landmark points are written as $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_N$, where $\boldsymbol{p}_i = (x_i, y_i, z_i)$. An elastic transform $h(\cdot)$ is applied to each landmark point such that $\boldsymbol{p}_i$ is modified by $\boldsymbol{p}_i \Leftarrow h(\boldsymbol{p}_i)$. The transform $h(\cdot)$ is chosen to be an implicit function that is actually applied by iterative shape deformation. The cost function to be iteratively minimized is

$$C(\boldsymbol{p}_1, \cdots, \boldsymbol{p}_N) = w_s C_s + w_r C_r, \tag{4.2}$$

which is the combination of a similarity cost $C_s$ and a regularization cost $C_r$ with appropriate weightings $w_s$ and $w_r$. If the target labeled volume is $\mathbf{L}$, $C_s$ is computed by

$$C_s = \sum\nolimits_i |\mathbf{D}(\boldsymbol{p}_i)|^2, \quad \mathbf{D} = \mathcal{DT}(\mathbf{L}), \tag{4.3}$$

where $\mathbf{D}$ is the distance map of $\mathbf{L}$ computed by distance transform $\mathcal{DT}$. It can be derived that iteratively minimizing $C_s$ is equivalent to iteratively moving $\boldsymbol{p}_i$ along the gradient direction of $\mathbf{D}(\boldsymbol{p}_i)$ such that

$$\boldsymbol{p}_i \Leftarrow \boldsymbol{p}_i + w_s \mathbf{D}(\boldsymbol{p}_i) \frac{\partial \mathbf{D}(\boldsymbol{p}_i)}{\partial \boldsymbol{p}_i}. \tag{4.4}$$

The minimization of $C_s$ propagates the landmarks onto the sample surface, is independent of the neighborhood structure of any mesh vertex, and cannot preserve the surface smoothness and the relative sizes of triangles associated with any mesh vertex. The regularization constraint $C_r$ is used to preserve the spatial smoothness of the transformation. The constraint used is the Dirichlet integral that can be calculated in the surface mesh case by

$$C_r = \frac{1}{2}\int_{\Omega} |\nabla h|^2 = \frac{1}{4}\sum\nolimits_{edge\ j} (\cot\alpha_j + \cot\beta_j)|a_j|^2, \tag{4.5}$$

where $|a_j|$ is the length of the $j$th edge, $\alpha_j$ and $\beta_j$ are two angles associated with the edge. Their definitions are shown in Figure 4.5 where the $j$th edge is the edge between points $\boldsymbol{p}_i$ and $\boldsymbol{q}_j$. It can be shown that when $C_r$ is minimized, vertex $\boldsymbol{p}_i$ must satisfy [40]

$$\boldsymbol{p}_i = \frac{\sum_j (\cot\alpha_j + \cot\beta_j)\boldsymbol{q}_j}{\sum_j (\cot\alpha_j + \cot\beta_j)}. \tag{4.6}$$

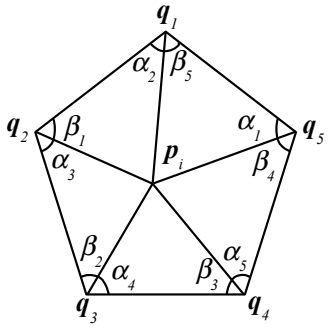



Figure 4.5: Vertex $\boldsymbol{p}_i$ of the surface mesh and its associated neighboring vertices $\boldsymbol{q}_j$.

Combining the optimization of $C_s$ and $C_r$ together, the landmark $\boldsymbol{p}_i$ is iteratively updated by

$$\boldsymbol{p}_i \Leftarrow \boldsymbol{p}_i + w_s \mathbf{D}(\boldsymbol{p}_i) \frac{\partial \mathbf{D}(\boldsymbol{p}_i)}{\partial \boldsymbol{p}_i} + w_r \left( \frac{\sum_j (\cot \alpha_j + \cot \beta_j) \boldsymbol{q}_j}{\sum_j (\cot \alpha_j + \cot \beta_j)} - \boldsymbol{p}_i \right). \tag{4.7}$$

The proper weightings, $w_s$ and $w_r$ are empirical values related to number of iterations. It was found using more iterations and smaller weightings generate better result by visual inspection. This implementation of landmark propagation is very fast, it takes approximately 2 seconds for a set of less than 1000 landmarks to be propagated using 200 iterations ($w_s$=0.03, $w_r$=0.01). Measured by landmark to target surface distance, the propagation achieved sub-voxel ($\approx 0.1 \pm 0.2, |\max| < 1$ voxel) accuracy.

## 4.2 Statistical Model Construction

### 4.2.1 Building Shape Models

In this study, after the landmarks were created, the statistical shape model was constructed using the same procedure described in Chapter 2. The shape alignment method used for Procrustes Analysis was the ICP registration of surface vertices as used for template creation. Depending on the purposes of the shape model, the scaling of the ICP registration used was isotropic or anisotropic.

- *Isotropic scaling* was used to create shape model for CAD – identifying the shape difference between normal and TOF subjects. In this case, the potential disease related shape variabilities in the ventricular height along the LV long-axis, the short-axis shape of the ventricles (circular or ellipsoidal LV, crescent

or half-moon-like RV) were kept intact.

- *Anisotropic scaling* was used to create shape model for segmentation. In this case, some global shape features (preserved for CAD purpose as above) were removed from the model. To start the segmentation, the mean shape was first scaled with anisotropic scaling and fit onto the target image, it brought the model shape much closer to the target than isotropic scaling and therefore can speedup the segmentation convergence and reduce the chance of converging to local minimums.

### 4.2.2 Building Appearance Models

The texture vector was constructed by the shape-free texture mapping in 3-D. The 3-D Delaunay tetrahedralization was first performed on the template landmarks. Since it always produced a convex surface from the input 3-D points without considering the actual input surface topology, some of the resulting tetrahedrons, whose centroids are outside the actual template surface, were removed. Then the point correspondence of all interior points were built upon the barycentric coordinate system.

Given a tetrahedron defined by its four vertices $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$, and $\boldsymbol{x}_4$, any point $\mathbf{x}$ inside it can be written as

$$\mathbf{x} = \alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2 + \gamma\boldsymbol{x}_3 + \delta\boldsymbol{x}_4, \tag{4.8}$$

where $\alpha + \beta + \gamma + \delta = 1$ and $0 \leq \alpha, \beta, \gamma, \delta \leq 1$ must be true.

With the coordinates of vertices written as $\boldsymbol{x}_i = [x_i, y_i, z_i]$ and the coordinates of

the point within the tetrahedron written as $\mathbf{x} = [x, y, z]$, five $4 \times 4$ matrices can be defined as

$$\mathbf{D}_0 = \begin{bmatrix} \boldsymbol{x}_1 & 1 \\ \boldsymbol{x}_2 & 1 \\ \boldsymbol{x}_3 & 1 \\ \boldsymbol{x}_4 & 1 \end{bmatrix}, \mathbf{D}_1 = \begin{bmatrix} \mathbf{x} & 1 \\ \boldsymbol{x}_2 & 1 \\ \boldsymbol{x}_3 & 1 \\ \boldsymbol{x}_4 & 1 \end{bmatrix}, \mathbf{D}_2 = \begin{bmatrix} \boldsymbol{x}_1 & 1 \\ \mathbf{x} & 1 \\ \boldsymbol{x}_3 & 1 \\ \boldsymbol{x}_4 & 1 \end{bmatrix}, \mathbf{D}_3 = \begin{bmatrix} \boldsymbol{x}_1 & 1 \\ \boldsymbol{x}_2 & 1 \\ \mathbf{x} & 1 \\ \boldsymbol{x}_4 & 1 \end{bmatrix}, \mathbf{D}_4 = \begin{bmatrix} \boldsymbol{x}_1 & 1 \\ \boldsymbol{x}_2 & 1 \\ \boldsymbol{x}_3 & 1 \\ \mathbf{x} & 1 \end{bmatrix}, \tag{4.9}$$

from which the barycentric coordinates $(\alpha, \beta, \gamma, \delta)$ of $\mathbf{x}$ is calculated by

$$\alpha = \frac{|\mathbf{D}_1|}{|\mathbf{D}_0|}, \quad \beta = \frac{|\mathbf{D}_2|}{|\mathbf{D}_0|}, \quad \gamma = \frac{|\mathbf{D}_3|}{|\mathbf{D}_0|}, \quad \delta = \frac{|\mathbf{D}_4|}{|\mathbf{D}_0|}, \tag{4.10}$$

where $|\mathbf{D}|$ represents the determinant of matrix $\mathbf{D}$. With $(\alpha, \beta, \gamma, \delta)$ solved for a point $\mathbf{x}$, mapping $\mathbf{x}$ to its corresponding point in a new tetrahedron is done by applying equation (4.8) with new vertices and same $(\alpha, \beta, \gamma, \delta)$ as illustrated in Figure 4.6.

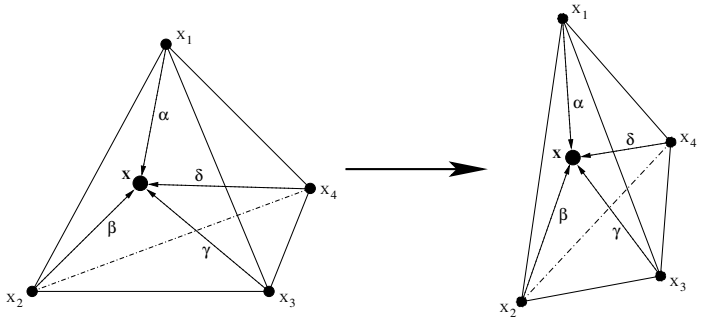



Figure 4.6: Shape-free texture mapping in 3-D using barycentric coordinates.

Unlike infarcted hearts, TOF does not change the texture of the ventricles – at least there are no visible changes in MR images. The shape feature is the dominant

difference between normal and TOF. Therefore the appearance model was only used for segmentation and not for CAD. The AAM segmentation can converge faster with better result if the texture of the surrounding region of the object is included since it gives the AAM matching more 3-D context to work with.

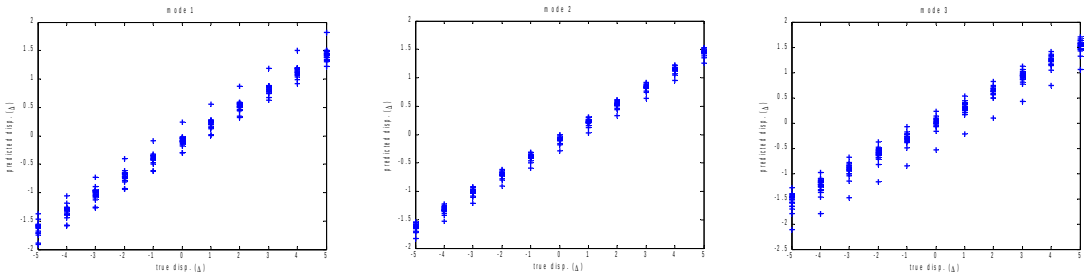
To find the texture of the surrounding region, a *virtual surface* was created by enlarging the landmark surface (the most exterior one if multiple surfaces exist) by a small scaling ratio such as 0.2. Then the same 3-D Delaunay and barycentric coordinate approach was used to build the texture vector. However, the new set of vertices of the virtual surface was not explicitly modeled, thus the name *virtual surface*. During the segmentation, whenever the shape changes, the virtual surface is updated from the current shape and the texture vector is re-sampled from the target image. As to the ventricular segmentation, if the LV and RV are modeled and segmented separately, no virtual surface was defined for LV since the LV epicardial surface, most part of which often has to be estimated in manual tracing, provides the needed 3-D context information, only the RV had its associated virtual surface defined.

## 4.3 Model Training and Quality Assessment

As described in Chapter 2, a sample object is expressed by three sets of parameters, appearance PCA modes $\boldsymbol{b}_a$, an affine transform $\mathcal{T}_s$, and a texture transform $\mathcal{T}_t$. Finding the Jacobian associated with $\mathcal{T}_t$ by model training is not necessary since it is more efficient to find $\mathcal{T}_t$ during AAM segmentation simply by texture normalization. The actual goal of model training is to compute Jacobian matrices for $\boldsymbol{b}_a$ and $\mathcal{T}_s$.

The affine transform is defined by 9 parameters (rotation, translation, and scaling). If the size of the training population is 20, the dimensionality of $\boldsymbol{b}_a$ is 20 or less. If 10 displacement values within a certain range are used in training, to cover the entire training space, the texture difference has to be evaluated $10^{29}$ times for each sample in the population, an impossible task for the 4-D AAM training.

To simplify the model training, we have to assume, without knowing whether it is true or not, that these parameters are not related, so the number of texture differences to be evaluated for each sample is reduced to 10×29=290. The training of a population contains 20 normal LVs using this simplified scheme only took approximately 15 minutes on a PC with 2.4GHz Pentium4 CPU. If the training space is exhausted, the estimated running time is $9.8\times10^{21}$ years.

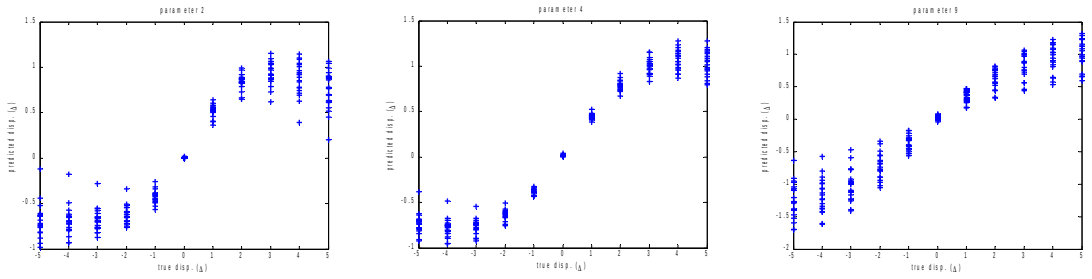


(a) 1st mode (b) 2nd mode (c) 3rd mode

Figure 4.7: 4-D AAM training quality tested on three strongest modes. $x$-axis: true displacement, $y$-axis: displacement predicted by model training ($\Delta = 0.1\sigma$).

Since the training procedure is much simplified, we have to test (at least once) whether the resulting Jacobians can make correct parameter displacement predictions in segmentation. However, the quality test has to use the same scheme as the training

using assume-to-be-unrelated values. Figures 4.7 and 4.8 show the results of quality tests on appearance modes and affine transform using a population of 20 normal LVs.



$\Delta = 2°$ (a) rotation (roll) | $\Delta = 1$ (voxel) (b) translation along $X$ | $\Delta = 2\%$ (c) scaling along $Z$

Figure 4.8: 4-D AAM training quality tested on three affine transform parameters. $x$-axis: true displacement, $y$-axis: displacement predicted by model training.

In the ideal case, the data points in those predicted-vs-true plots should approximately form a line of unit-slope and zero-intercept. In Figure 4.7, a very good linear relation is observed with an approximate slope of 0.5 and near-zero intercept. Although the slope indicates underestimated displacements, the training of appearance modes can still be considered very good since the linearity is more important than slope – we can always apply a constant scaling to the predicted displacements in segmentation.

On the other hand, the quality of affine parameter training shown in Figure 4.8 is worse, the preferred linear relationship only exists in a very small range of displacements. The overall patterns in the reasonably small training space – $\pm 10°$ rotation, $\pm 5$ voxels of translation, $\pm 10\%$ scaling – are more sinusoidal than linear.

From the above training quality assessments several important rules of segmentation strategy design are derived.

- We can rely on the capability of trained Jacobians in optimizing appearance modes but only when the model is fitted closely to the target.
- Because stronger modes are often associated with larger shape variations, they should be optimized first.
- The trained Jacobians of the affine transform are not reliable. They might contribute to the fine tuning of the segmentation but cannot be trusted for model initialization.
- Brute-force optimization of affine transform is necessary for model initialization and bring the model matching out of local minimums.

## 4.4 Summary

In this chapter, an automatic landmarking method is specially designed for 4-D ventricular shapes. Based on the landmarks created, the 4-D AAM is constructed and the quality of model training is tested. The results of model training quality tests are considered in choosing the customized segmentation strategies in the next chapter.

# CHAPTER 5
# 4-D AAM SEGMENTATION

## 5.1 Segmentation Strategies

Building the independent standard for model construction and segmentation validation is a very time-consuming process even using the developed true 4-D tracing application, therefore the size of the sample population is limited. In order to achieve good segmentation results on the TOF subjects with large shape variability, the 4-D AAM ventricular segmentations were performed using several customized strategies: model and segment normal and TOF subjects separately, model and segment LV and RV separately, use *anisotropic* scaling in AAM, and manually initialize the mean shape onto the target image.

Modeling and segmenting normal and TOF subjects using a combined model has some drawbacks. The mean shape of the combined model is neither a typical normal nor TOF heart, the segmentation therefore requires more iterations to converge. The normal hearts have less shape variability than the TOF hearts, if trained with a small population, the combined model could easily capture most of the shape variabilities of the TOF group and the common difference between normal and TOF groups while some of the shape variabilities of the normal group may be lost. In addition, training a combined model is more expensive than training two separated normal and TOF models, each only contains half the samples as the combined model. However, using separate normal and TOF models requires the segmentation of an unknown data to be performed twice and then choose the best result, which can be done by visual in-

spection or automatically comparing the model matching errors. In practice, whether a subject is normal and TOF patient is often obvious, so only one segmentation is actually needed. Another advantage of the separated models is when applying the AAM segmentation method to other types of cardiac diseases, the normal model does not need to be reconstructed and re-trained.

The left and right ventricles can either be treated together as a single entity using a combined LV+RV model or separately. The most visible feature of the TOF on MR image is the RV dilation that can happen at various locations of the RV as illustrated in Figure 5.1. The RV may wrap around the LV (Figure 5.1a) or expand away from the LV (Figure 5.1b) without changing the LV's shape. The RV may also expand toward the LV (Figure 5.1c) so that the LV's shape on the short-axis view is not circular anymore. On the long-axis view (Figure 5.1(d,e)), the RV may elevate toward the right atrium so that the LV and RV are not side-by-side anymore. A combined LV+RV model in this case adds some extra and unnecessary global constraints on the relationship (relative size and position) of the LV and RV and therefore restricts the capability of AAM segmentation. The separated LV and RV models give the AAM segmentation more freedom in shape deformation.

Even the normal hearts have various ventricular heights along the long-axis of the LV. During the manual tracing, it was found that the correct identification of the ventricular bases is still difficult and can only be achieved when the atria's shapes are determined. Although the model contains some knowledge about the surrounding intensity patterns of the ventricles, the atria are not entirely included. Some prelim-

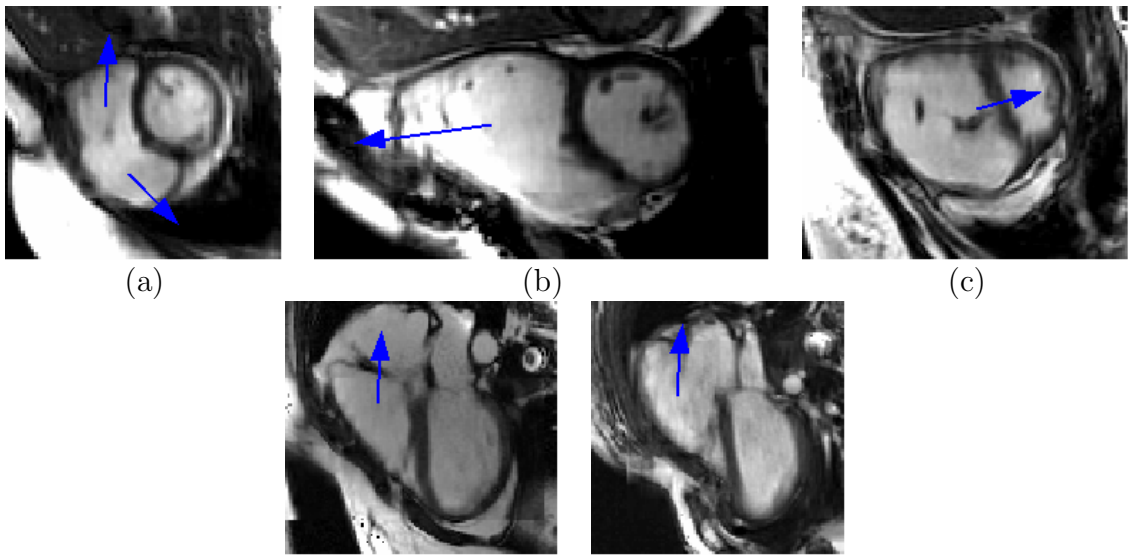



Figure 5.1: Different types of RV dilation showing in short-axis view (a,b,c) and long-axis view (d,e).

inary tests of the 4-D AAM segmentations and the model training quality tests (See 4.3) showed that the segmentation is very sensitive to how the mean shape is initially fitted onto the target image. Using manual initialization and *anisotropic scaling* in the model construction and segmentation were chosen due to their advantages and flexibilities. The anisotropic scaling removes the variability in the ventricular height from the model. It also leaves the task of identifying the ventricular bases and apexes to the user. The manual initialization of the mean shape is more robust than automatic affine transform optimization, which has be shown to have poor quality in 4.3. Although it requires some manual operations, it only needs to be performed using the first cardiac phase, and anisotropic scaling also helps the user to achieve better manual initialization without knowing the shape variations associated with the PCA

modes.

## 5.2 Experimental Methods

The study population consists of normal and TOF groups. The normal group contains 25 MR scans of 24 subjects, one subject has two MR scans in the same day on different scanners. The TOF group contains 25 MR scans of 24 patients with repaired TOF and resultant pulmonic regurgitation, one patient has two MR scans of six months apart on different scanners. The normal group was mostly scanned by a Siemens Avanto scanner with 2-D resolution of 2.05±0.15 mm and slice thickness of 6–10 mm. The TOF group was mostly scanned by a GE Signa scanner with 2-D resolution of 1.48±0.19 mm and slice thickness of 6–10 mm. The MR imaging protocol used in this study is the steady-state free precession. Short-axis images completely cover the ventricles were acquired. The number of long-axis images acquired varied within 2 to 10, with at least one image clearly showing the four heart chambers. The original MR images went through the 4-D data construction steps described in Chapter 3. The 4-D manual tracings were performed by a cardiologist expert and then verified by another senior expert.

The 4-D AAM segmentations were performed using the strategies described in 5.1 and the hybrid AAM/ASM algorithm described in 2.4. In addition to the manual affine transform initializations, the transform parameters were also fine tuned by a random-restart hill climbing optimization [42] within a small range at the beginning of each model matching iteration.

The performance of the segmentation was assessed using *holdout* validation. For

each normal or TOF group, the 25 scans were repeatedly divided into a training group of 20 scans and a testing group of 5 scans. The 4-D AAMs were created and trained from the training group and the segmentations were then performed on the testing group. The whole procedure was repeated five times until all the 25 scans were segmented. The two same-day repeated normal scans were put into the same testing group and the two six-month-apart TOF scans were put into different testing groups.

In addition to the AAM and manual results, another set of *synthesized* results were created for performance assessment. The samples in the testing group were aligned to the model built from the training group and their corresponding appearance PCA modes were calculated and restricted within $\pm 5\sigma$, from which the shape vectors of the testing samples were calculated and translated to surface and labeled volume representations to form the *synthesized* results. The synthesized results represent the *optimal* results can be achieved by AAM segmentation using a given training group. Then three sets of testing and reference result combinations were compared as: *AAM vs. manual*, *AAM vs. synthesized*, and *synthesized vs. manual*.

The manual initialization of the mean shape was performed by setting the affine transform parameters (rotation, scaling, and translation) to align the mean shape to the manual results without knowing the 'optimal' values found by the synthesized results. One to ten sets of parameters were tested and the best segmentation result – identified by surface positioning error – was chosen to be the final AAM result.

The segmentation performance was assessed by three types of error metrics:

- *Surface positioning error* is defined as the 3-D distance between two surfaces. If signed error is used, negative error means the testing point is inside the reference surface. Three ventricular surfaces were tested: the LV endocardial (LV Endo) surface, the LV epicardial (LV Epi) surface, and the RV epicardial surface.
- *Relative label overlap* is defined as the overlapping ratio of two binary label sets $A$ and $B$ as $A \cap B/(A \cup B)$. The three set of labels tested are: the LV blood pool (LV Endo), the whole LV (blood pool and myocardium, LV Epi), and the RV.
- *Ventricular capacity* is measured from the labeled volumes in milliliters. The conventional end-diastolic and end-systolic volumes of LV and RV were measures as LVEDV, LVESV, RVEDV, and RVESV after the end-diastolic and end-systolic phases were automatically identified from the results being tested.

The first two types of error metrics were measured for each cardiac phase and for all phases combined; the testing result was also roughly divided into three ventricular sections – apex, mid-ventricle, and base – for error assessment.

## 5.3 Results of Normal Hearts

### 5.3.1 Overall Segmentation Errors

The overall surface positioning and relative overlap errors measured form the segmentations of the normal hearts are listed in Table 5.1. The signed errors show small biases, sub-voxel average accuracies and standard deviations around 1.5 voxels. The absolute errors show close to 1.0 voxels of average accuracies, less than 2.0 voxels of

standard deviations. The segmentations of LV epicardial surface are slightly worse than but very close to those of LV endocardial surface, showing that the AAM successfully incorporated expert knowledge about the estimated LV epicardial surface and utilized it in the segmentation. Comparing the errors of LV and RV segmentations, the performance on RV is worse than on LV in the standard deviation of positioning errors and average overlap ratios, it indicates that the segmentation accuracy is affected by the nature of the shape. In this case the irregularity of the RV, especially at the junction of the LV and RV, cannot be fully captured by the model into PCA modes as global features.

| Normal | Surface positioning errors | | | | Overlap |
|---|---|---|---|---|---|
| LV Endo | sgn voxel | sgn mm | abs voxel | abs mm | |
| AAM vs manual | 0.13±1.38 | 0.27±2.83 | 1.11±1.71 | 2.27±3.50 | 0.77±0.03 |
| AAM vs syn. | -0.32±1.13 | -0.66±2.31 | 0.95±1.71 | 1.94±3.51 | 0.81±0.04 |
| Syn. vs manual | 0.48±1.16 | 0.98±2.37 | 1.00±1.27 | 2.04±2.59 | 0.80±0.03 |
| LV Epi | sgn voxel | sgn mm | abs voxel | abs mm | |
| AAM vs manual | -0.12±1.45 | -0.26±2.98 | 1.16±1.96 | 2.37±4.01 | 0.80±0.03 |
| AAM vs syn. | -0.40±1.22 | -0.81±2.51 | 1.06±1.92 | 2.17±3.95 | 0.82±0.04 |
| Syn. vs manual | 0.30±1.13 | 0.61±2.32 | 0.92±1.30 | 1.88±2.66 | 0.84±0.02 |
| RV | sgn voxel | sgn mm | abs voxel | abs mm | |
| AAM vs manual | 0.04±1.64 | 0.09±3.37 | 1.36±2.14 | 2.79±4.38 | 0.66±0.04 |
| AAM vs syn. | -0.21±1.30 | -0.44±2.66 | 1.10±1.88 | 2.25±3.86 | 0.72±0.05 |
| Syn. vs manual | 0.31±1.33 | 0.64±2.71 | 1.09±1.54 | 1.22±3.15 | 0.73±0.04 |

Table 5.1: The overall surface positioning and relative overlap errors of normal hearts expressed in mean±SD.

The limitation of PCA modes as global descriptors is also reflected in the synthesized vs. manual comparison. It indicates that some local shape features are not included in the PCA modes, therefore the standard deviations of positioning errors are larger than 1.0 voxels and the relative overlap ratios are less than 90% or even as

low as 73% for RV.

Due to the limitation of the PCA modes, the errors of AAM vs. manual comparison cannot outperform those of synthesized vs. manual comparison, which is clear shown in Table 5.1. However, the AAM vs. synthesized comparison show no close to perfect accuracy due to two reasons. First, the 'optimal' affine transform that aligns the target and mean shapes is always hard to be manually initialized and automatically optimized in 4-D case. Second, the high computational cost limits the number of model matching iterations can be performed in practice and the real-world AAM result is just a close-to-optimal one, possibly also a local minimum. In this study, starting from a given initial affine transform, a whole iteration of optimizing all PCA modes took 5 to 10 minutes on a PC with Pentium4 2.4GHz CPU when the C++ implementation was custom designed for performance.

### 5.3.2 Segmentation Errors of Repeated Scans

The overall segmentation errors of the repeated normal scans are listed in Table 5.2. Although having different 2-D imaging resolutions (1.72 mm for GE scanner, 2.08 mm for Siemens scanner), the positioning errors measured by millimeters showed no clear difference as well as the relative overlap ratios. The segmentation accuracies achieved are comparable to those listed in Table 5.1 without decrease, sometimes even better.

| Normal | Surface positioning errors | | | | Overlap |
|---|---|---|---|---|---|
| LV Endo | sgn voxel | sgn mm | abs voxel | abs mm | |
| GE | 0.05±1.26 | 0.09±2.17 | 1.01±1.58 | 1.74±2.72 | 0.81 |
| Siemens | -0.44±1.11 | 0.91±2.32 | 0.93±1.21 | 1.94±2.54 | 0.80 |
| RV | sgn voxel | sgn mm | abs voxel | abs mm | |
| GE | -0.29±1.75 | -0.51±3.01 | 1.46±2.48 | 2.51±4.26 | 0.66 |
| Siemens | -0.01±1.67 | -0.01±3.47 | 1.37±2.16 | 2.85±4.50 | 0.64 |

Table 5.2: Segmentation errors (comparing with the manual result) of the repeated scans of a normal subjects expressed in mean±SD.

### 5.3.3 Segmentation Errors of Ventricular Sections

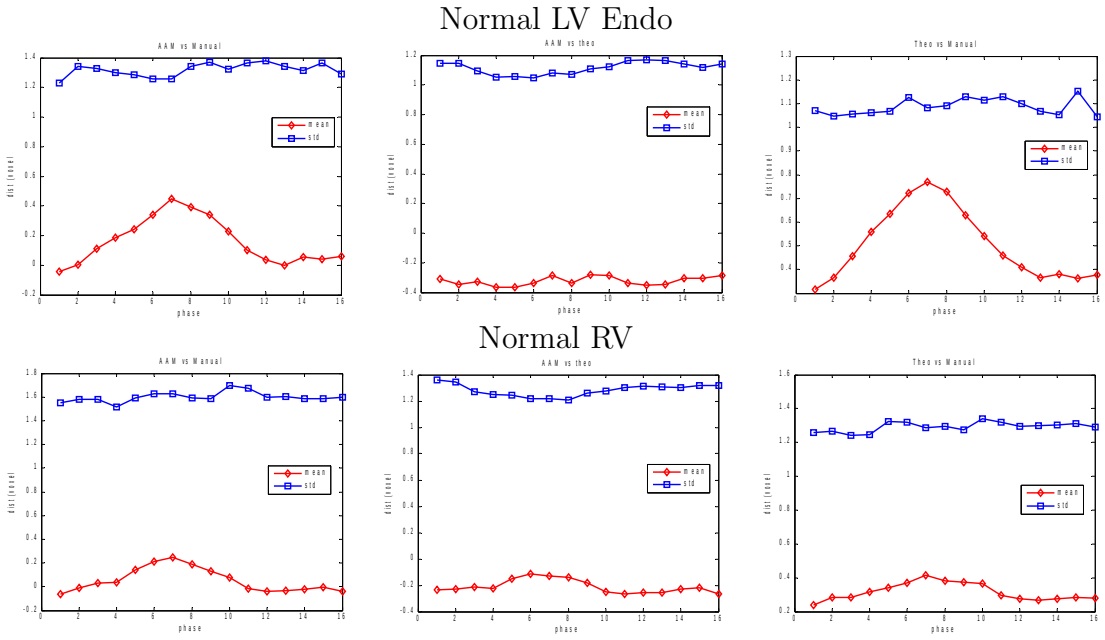
The surface positioning and relative overlap errors for the three ventricular sections of normal hearts are listed in Table 5.3. It shows that the segmentation is more sensitive to the orientation of the heart in ventricular apexes and bases. Since the mid-ventricle sections are easy to trace and provide reliable information for the affine transform initialization, the segmentation of them outperformed those of apexes and bases and of the overall ventricles, sometimes even close to those of the overall synthesized vs. manual comparison. Although the approximate locations of ventricular apexes are easier to find on the MR images than those of the bases, achieving accurate segmentation of these small sections may be more difficult, which explains the smaller overlap ratios and larger positioning deviations of the apexes.

### 5.3.4 Temporal Curves of Segmentation Errors

In order to further analyze the segmentation performances at different cardiac phases, the temporal curves of the surface positioning errors and relative overlap ratios are plotted in Figures 5.2 and 5.3, each showing the curves of the mean and standard deviation of errors.

| | Surface positioning errors | | | | Overlap |
|---|---|---|---|---|---|
| | sgn voxel | sgn mm | abs voxel | abs mm | |
| Normal LV Endo, AAM vs manual | | | | | |
| apex | 0.76±1.16 | 1.55±2.37 | 1.22±1.30 | 2.50±2.65 | 0.68±0.09 |
| mid. | -0.06±1.19 | -0.11±2.43 | 0.98±1.60 | 2.01±3.27 | 0.83±0.04 |
| base | -0.02±1.38 | -0.05±2.81 | 1.16±1.86 | 2.37±3.81 | 0.75±0.04 |
| Normal RV, AAM vs manual | | | | | |
| apex | 0.85±1.77 | 1.74±3.64 | 1.71±2.00 | 3.51±4.11 | 0.46±0.08 |
| mid. | 0.01±1.42 | 0.02±2.92 | 1.24±1.92 | 2.54±3.94 | 0.72±0.05 |
| base | -0.26±1.54 | -0.54±3.16 | 1.30±2.24 | 2.67±4.59 | 0.67±0.05 |

Table 5.3: Segmentation errors for three ventricular sections of normal hearts expressed in mean±SD.




(a) AAM vs manual (b) AAM vs syn. (c) Syn. vs manual

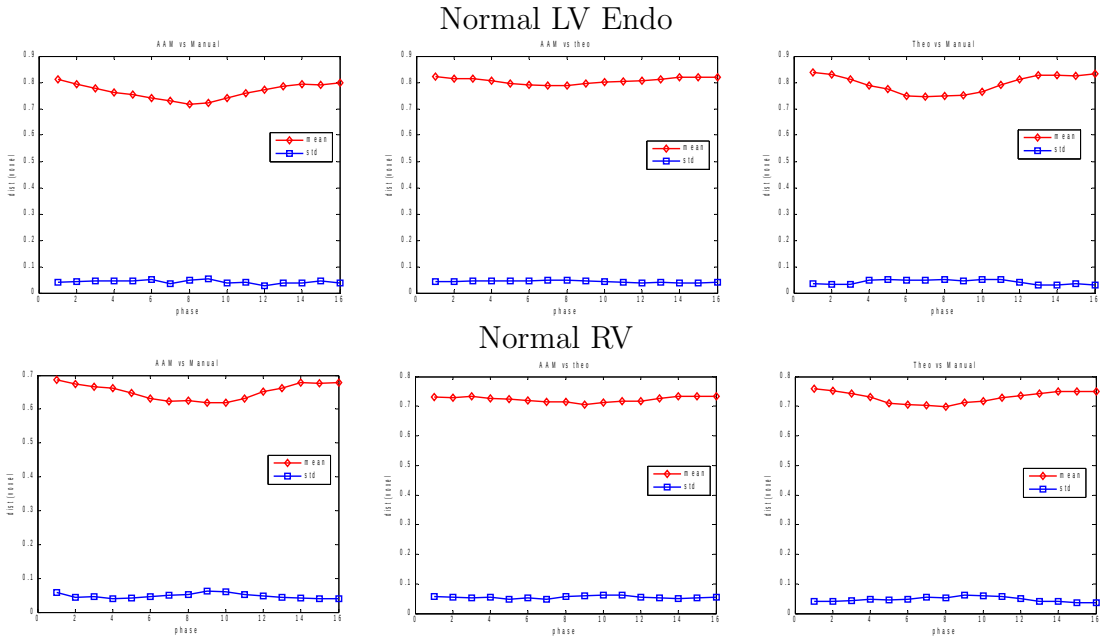
Figure 5.2: The surface positioning error curves of normal hearts. Bottom curves (in red): mean error, top curves (in blue): error standard deviation.

The MR acquisition is triggered by ECG gating, so the end-diastole appears at the beginning or the end of the cardiac cycle and the the end-systole appears in the middle of the cardiac cycle. The plotted curves of synthesized vs. manual comparison show that the model cannot exactly capture the contraction and relaxation pattern of the heart, especially the LV, therefore the mean LV positioning error curve has a peak and the mean overlap ratio curves have a valley in the middle of the cardiac cycle. The same kinds of peaks and valleys also exist in the curves of AAM vs. manual comparison. However, the curves of AAM vs. synthesized comparison shows that the segmentation follows the contraction and relaxation patterns contained in the model, hence no fluctuation of the mean curves (except the RV positioning errors). All the standard deviation curves are not affected by the contraction and expansion of the ventricles.

### 5.3.5 Ventricular Capacity Agreements

The end-diastolic and end-systolic volumes of each ventricle were analyzed together by linear regression and Bland-Altman analysis. The resulting plots are shown in Figures 5.4 and 5.5. The plotted lines in the linear regression plots represent the linear relationship found by the regressions. The Y-axis of the Bland-Altman plots is the percentage difference between method $M_1$ and $M_2$ as

$$100\% \times \frac{M_2 - M_1}{0.5(M_2 + M_1)}. \tag{5.1}$$

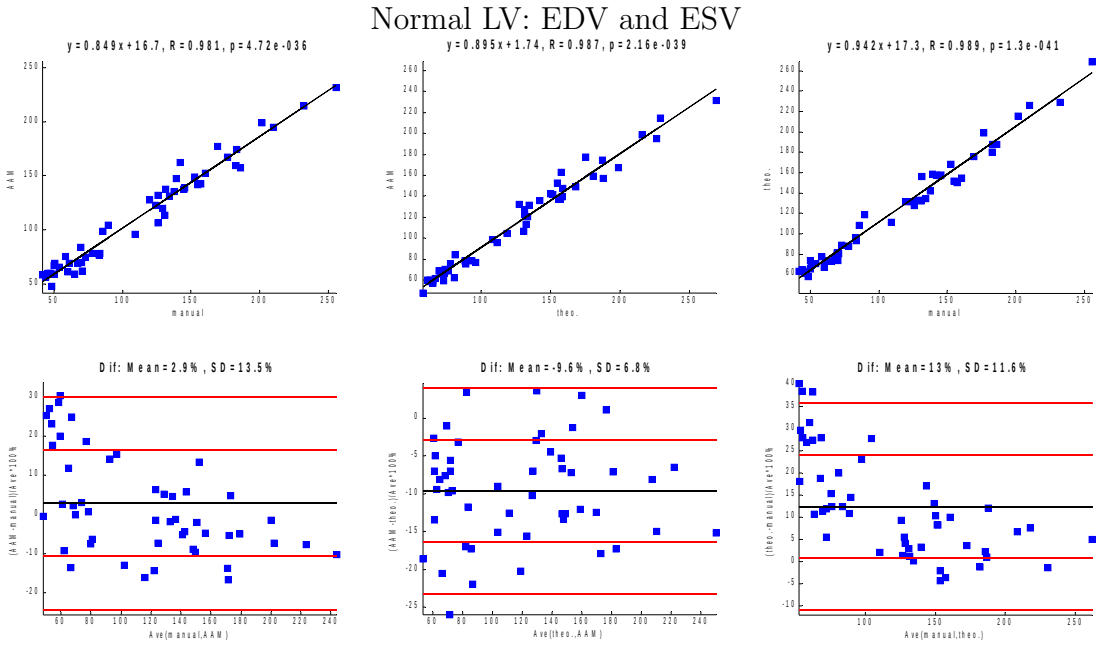



(a) AAM vs manual (b) AAM vs syn. (c) Syn. vs manual

Figure 5.3: The relative overlap curves of normal hearts. Top curves (in red): mean error, bottom curves (in blue): error standard deviation.

The five plotted lines in the Bland-Altman plots from top to bottom are: mean-2SD, mean-SD, mean, mean+SD, and mean+2SD, respectively. The numerical results of the analyses are listed in Table 5.4.

All the linear regression results show very good correlations of results ($R > 0.92, p \ll 0.05$). The distributions of errors in the Bland-Altman plots indicate the same properties of the segmentation performance as in the analyses of the other two error metrics: limitation of PCA modes as global descriptor, slightly worse results for RV. How the PCA modes cannot fully capture the contraction and relaxation patterns of the ventricles is more obvious here, where the ESVs are mostly overestimated and the EDVs slightly underestimated in synthesized vs. manual comparison. It is




(a) AAM vs manual (b) AAM vs syn. (c) Syn. vs manual

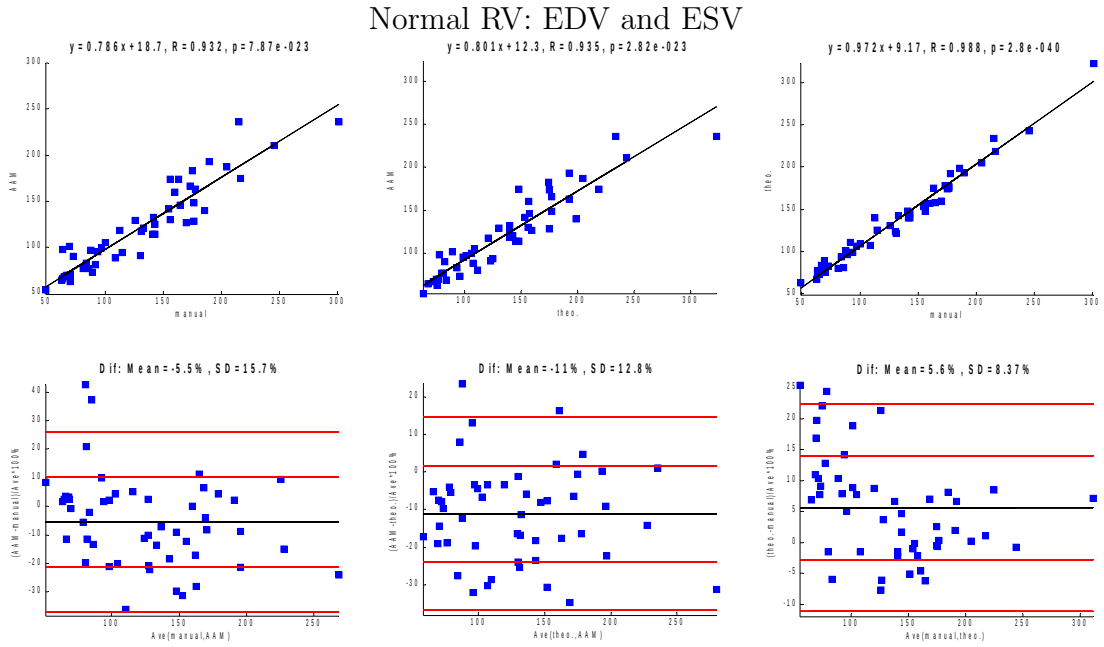
Figure 5.4: Linear regression (top row) and Bland-Altman plots (bottom row) of the normal LVEDVs and LVESVs.

also clear that the segmentation results are not the 'optimal' ones suggested by the synthesized results but local minimums with underestimated ventricular capacities.

## 5.4 Results of TOF Hearts

### 5.4.1 Overall Segmentation Errors

The overall surface positioning and relative overlap errors measured from the segmentations of TOF hearts are listed in Table 5.5. They are worse than the results of the normal hearts. The signed errors show small biases of -1.0 to -2.0 voxels, and standard deviations of 2.0 to 3.0 voxels. The absolute errors show 1.5 to 2.5 voxels of average errors, less than 5.0 voxels of standard deviations. Compared with the results of normal subjects measured by millimeters, the increases in positioning




(d) AAM vs manual (e) AAM vs syn. (f) Syn. vs manual

Figure 5.5: Linear regression (top row) and Bland-Altman plots (bottom row) of the normal RVEDVs and RVESVs.

| Normal | Linear regression | | Bland-Altman | |
|---|---|---|---|---|
| LV | Equation | R | Mean | SD |
| AAM vs. manual | $y = 0.85x + 16.7$ | 0.98 | 2.9% | 13.5% |
| AAM vs. syn. | $y = 0.90x + 1.7$ | 0.99 | -9.6% | 6.8% |
| Syn. vs. manual | $y = 0.94x + 17.3$ | 0.99 | 13.0% | 11.6% |
| RV | Equation | R | Mean | SD |
| AAM vs. manual | $y = 0.79x + 18.7$ | 0.93 | -5.5% | 15.7% |
| AAM vs. syn. | $y = 0.80x + 12.3$ | 0.93 | -11.0% | 12.8% |
| Syn. vs. manual | $y = 0.97x + 9.2$ | 0.99 | 5.6% | 8.4% |

Table 5.4: Linear regressions and Bland-Altman analyses of the normal hearts.

standard deviations are 1.5 to 2.0 mm, and the overlap ratios show obvious decreases in means and increases in standard deviations.

| TOF | Surface positioning errors | | | | Overlap |
|---|---|---|---|---|---|
| LV Endo | sgn voxel | sgn mm | abs voxel | abs mm | |
| AAM vs manual | -1.00±2.18 | -1.45±3.18 | 1.94±3.68 | 2.84±5.37 | 0.66±0.06 |
| AAM vs syn. | -1.59±1.91 | -2.31±2.79 | 2.09±4.19 | 3.05±6.10 | 0.67±0.07 |
| Syn. vs manual | 0.58±1.81 | 0.85±2.66 | 1.54±2.06 | 2.26±3.02 | 0.76±0.05 |
| LV Epi | sgn voxel | sgn mm | abs voxel | abs mm | |
| AAM vs manual | -1.70±2.25 | -2.47±3.29 | 2.35±4.66 | 3.43±6.80 | 0.68±0.06 |
| AAM vs syn. | -2.10±2.02 | -3.06±2.95 | 2.50±5.07 | 3.65±7.38 | 0.68±0.08 |
| Syn. vs manual | 0.35±1.77 | 0.52±2.61 | 1.46±2.11 | 2.15±3.09 | 0.81±0.04 |
| RV | sgn voxel | sgn mm | abs voxel | abs mm | |
| AAM vs manual | -0.63±2.93 | -0.85±4.28 | 2.45±4.30 | 3.56±6.22 | 0.61±0.06 |
| AAM vs syn. | -0.70±2.69 | -0.92±3.92 | 2.32±4.10 | 3.37±5.90 | 0.66±0.07 |
| Syn. vs manual | 0.11±2.12 | 0.14±3.10 | 1.71±2.67 | 2.50±3.91 | 0.74±0.04 |

Table 5.5: The overall surface positioning and relative overlap errors of TOF hearts expressed in mean±SD.

Although other similar observations such as the limitation of PCA modes as global descriptor can be made as those from the results of the normal hearts, a clear difference in the accuracies of LV endocadial and epicardial surfaces is observed in TOF patients, indicating the existence of more variability of LV myocardium in TOF and the AAM's inability of capturing such changes.

### 5.4.2 Segmentation Errors of 'Repeated' Scans

The overall segmentation errors of the two six-month-apart TOF scans are listed in Table 5.6. They were scanned with different 2-D resolutions, 1.56 mm for GE scanner, and 2.03 mm for Siemens scanner. Unlike the repeated normal scans, the results on Siemens scan show clear decrease in segmentation accuracy not only due to the lower resolution but also its inaccurate LV long-axis direction in the MR image,

which makes manual tracing and AAM segmentation more difficult.

| TOF | Surface positioning errors | | | | Overlap |
|---|---|---|---|---|---|
| LV Endo | sgn voxel | sgn mm | abs voxel | abs mm | |
| 1st (GE) | -0.10±1.62 | -0.15±2.54 | 1.29±2.14 | 2.01±3.34 | 0.78 |
| 2nd (Siemens) | -0.97±1.62 | -1.97±3.29 | 1.57±3.01 | 3.18±6.17 | 0.66 |
| RV | sgn voxel | sgn mm | abs voxel | abs mm | |
| 1st (GE) | -0.41±2.49 | -0.65±3.89 | 1.97±3.44 | 3.07±5.38 | 0.64 |
| 2nd (Siemens) | 0.54±2.09 | 1.09±4.24 | 1.75±2.42 | 3.55±4.91 | 0.63 |

Table 5.6: Segmentation errors (comparing with the manual result) of the six-month-apart scans of a TOF patient expressed in mean±SD.

### 5.4.3 Segmentation Errors of Ventricular Sections

The surface positioning and relative overlap errors for the three ventricular sections of TOF hearts are listed in Table 5.7. It still shows that the segmentation is more sensitive the orientation of the heart in ventricular apexes and bases. The TOF patients often have larger RV (especially RV bases) than normal subjects and various types and their combinations of RV elevation (aneurysm toward the right atrium), which is more difficult to trace as well as for computer segmentation to find, so larger error standard deviations are observed in the RV bases.

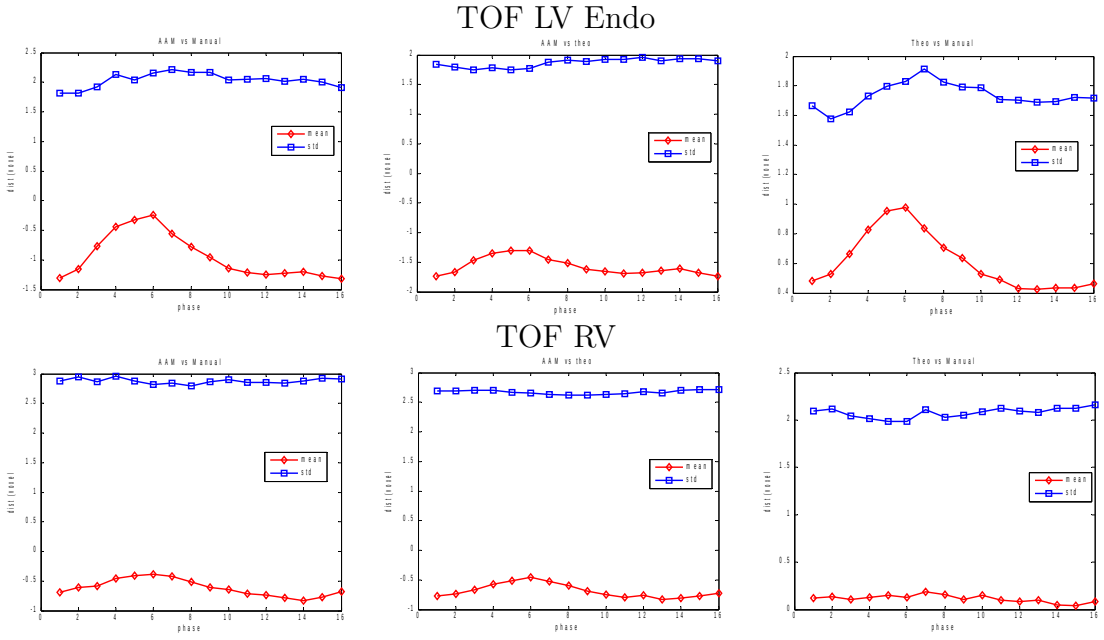
### 5.4.4 Temporal Curves of Segmentation Errors

The temporal curves of the surface positioning errors and relative overlap ratios are plotted in Figures 5.6 and 5.7, each showing the curves of the mean and standard deviation of errors.

Most of the observations made on the temporal curves of normal hearts still hold

| | Surface positioning errors | | | | Overlap |
|---|---|---|---|---|---|
| | sgn voxel | sgn mm | abs voxel | abs mm | |
| TOF LV Endo, AAM vs manual | | | | | |
| apex | -0.57±1.95 | -0.86±2.86 | 1.80±3.16 | 2.63±4.64 | 0.60±0.08 |
| mid. | -1.15±1.98 | -1.69±2.89 | 1.88±3.65 | 2.75±5.34 | 0.74±0.05 |
| base | -1.12±2.21 | -1.60±3.22 | 2.08±3.94 | 3.03±5.72 | 0.63±0.07 |
| TOF RV, AAM vs manual | | | | | |
| apex | 0.44±2.51 | 0.69±3.65 | 2.25±3.19 | 3.27±4.62 | 0.57±0.10 |
| mid. | -0.74±2.60 | -1.04±3.80 | 2.33±4.11 | 3.39±5.97 | 0.70±0.07 |
| base | -1.17±2.99 | -1.62±4.36 | 2.67±4.94 | 3.87±7.12 | 0.56±0.08 |

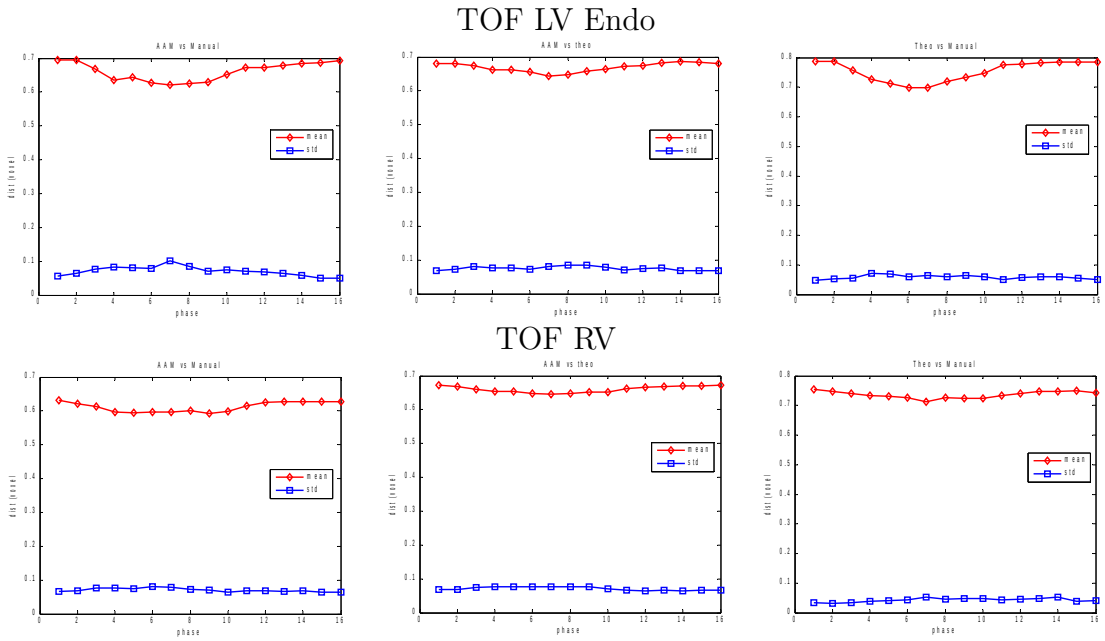
Table 5.7: Segmentation errors of ventricular sections of TOF hearts expressed in mean±SD.




(a) AAM vs manual (b) AAM vs syn. (c) Syn. vs manual

Figure 5.6: The surface positioning error curves of TOF patients. Bottom curves (in red): mean error, top curves (in blue): error standard deviation.

here, but the positioning standard deviation curve of synthesized vs. manual comparison of LV (top of Figure 5.6c) shows a similar pattern as the mean error, indicating the TOF LVs may contain more shape variability related to the contraction and relaxation of the heart that is not captured by the model.



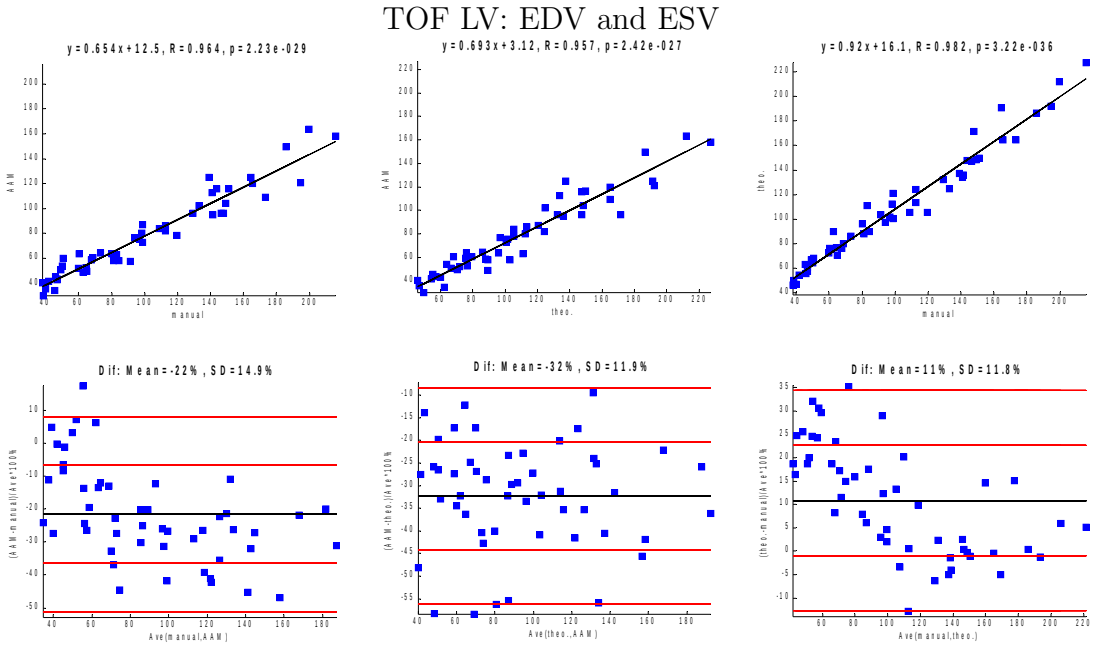

(a) AAM vs manual (b) AAM vs syn. (c) Syn. vs manual

Figure 5.7: The relative overlap curves of TOF patients. Top curves (in red): mean error, bottom curves (in blue): error standard deviation.

### 5.4.5 Ventricular Capacity Agreements

The end-diastolic and end-systolic volumes of each ventricle were analyzed together by linear regression and Bland-Altman analysis. The resulting plots are shown in Figures 5.8 and 5.9. The numerical results of the analyses are listed in Table 5.8.

All the linear regression results still show very good correlations of results ($R >$



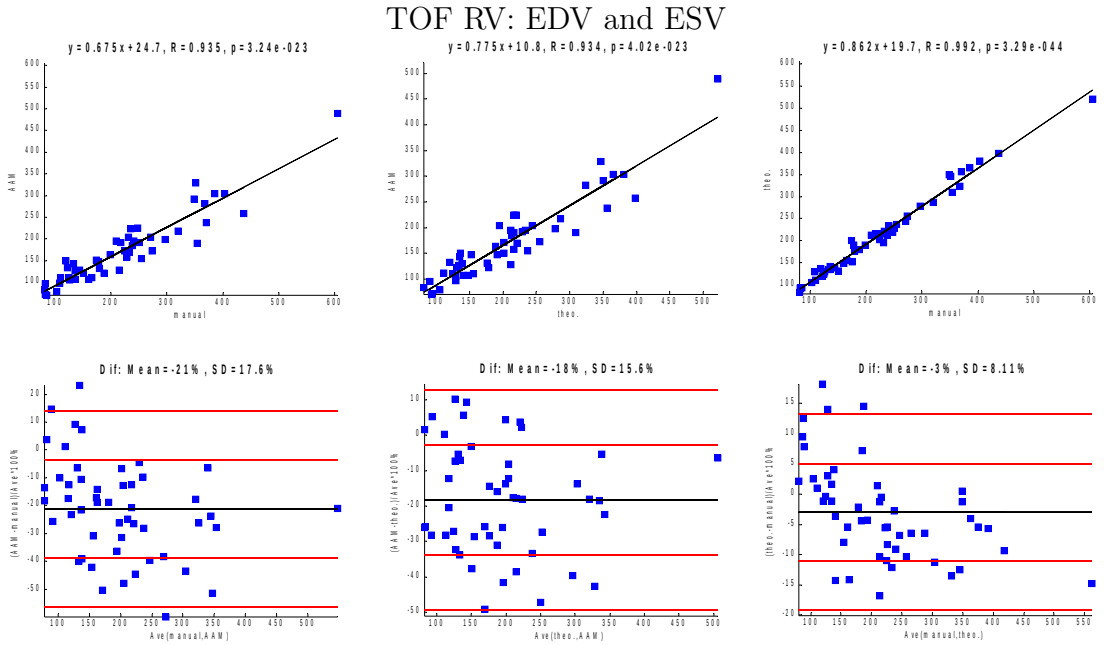

(a) AAM vs manual (b) AAM vs syn. (c) Syn. vs manual

Figure 5.8: Linear regression (top row) and Bland-Altman plots (bottom row) of TOF LVEDVs and LVESVs.

$0.93, p \ll 0.05$). Similar observations can be made on the error distributions from the Bland-Altman plots as those from the results of normal hearts. The results of TOF patients also show larger error standard deviations and more capacity underestimations in the Bland-Altman plots than the normal subjects.

## 5.5 Summary and Discussion

The main difficulty in achieving accurate 4-D AAM ventricular segmentations is the small training size. Without substantially large sample population, the PCA modes as global descriptors cannot capture all the shape variabilities existing in the population. Although the ventricular shapes of neighboring cardiac phases are similar, the overall contraction and relaxation pattern of the ventricles are still not fully




(d) AAM vs manual (e) AAM vs syn. (f) Syn. vs manual

Figure 5.9: Linear regression (top row) and Bland-Altman plots (bottom row) of TOF RVEDVs and RVESVs.

captured by the AAM using a limited population.

By using the custom designed segmentation strategies, the 4-D AAM method successfully produced very good segmentation results on normal hearts. An example of the segmentation result at the first and eighth cardiac phases is shown in Figure 5.10 in three orthogonal views. On the specific short-axis slice shown, the manual result does not represent the true RV shape in the image due to shape interpolation, while the AAM segmentation of RV is slightly more agreeable to the the true shape. In the long-axis view, the segmented RV is lower than the manual result, especially on the eighth phase, showing that the AAM had some difficulties in identifying the location of RV base and following the contraction of RV. Also in the long-axis view,

| TOF | Linear regression | | Bland-Altman | |
|---|---|---|---|---|
| LV | Equation | R | Mean | SD |
| AAM vs. manual | $y = 0.65x + 12.5$ | 0.96 | -22.0% | 14.9% |
| AAM vs. syn. | $y = 0.69x + 3.1$ | 0.96 | -32.0% | 11.9% |
| Syn. vs. manual | $y = 0.92x + 16.1$ | 0.98 | 11.0% | 11.8% |
| RV | Equation | R | Mean | SD |
| AAM vs. manual | $y = 0.68x + 24.7$ | 0.94 | -21.0% | 17.6% |
| AAM vs. syn. | $y = 0.78x + 10.8$ | 0.93 | -18.0% | 15.6% |
| Syn. vs. manual | $y = 0.86x + 19.7$ | 0.99 | -3.0% | 8.1% |

Table 5.8: Linear regressions and Bland-Altman analyses of TOF hearts.

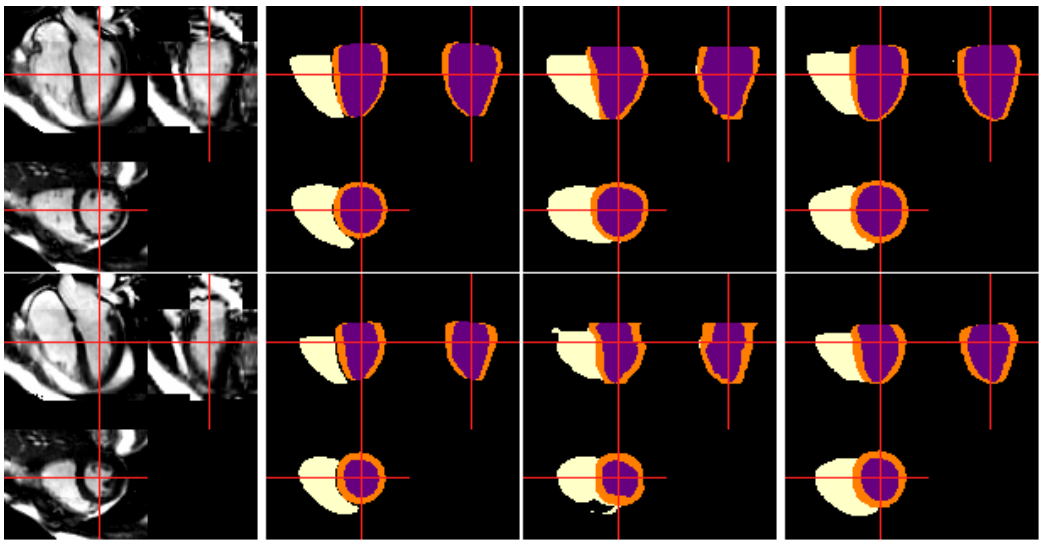



Figure 5.10: An example of a normal heart. Top row: first phase, bottom row: eighth phase.

although the image indicates slight tilting of the LV, it was not exactly captured by the manual tracing and synthesized result on the first phase, therefore not reflected in the AAM segmentation that is initialized based on the first phase manual result. In spite of these localized inaccuracies, the AAM segmentation still produced a very satisfactory result in the overall sense.

When AAM is used to capture larger and more complex shape and texture variabilities of the TOF hearts. The same kinds of localized inaccuracies are also observed but worse than those of the normal hearts. An example of the segmentation result at the first and eighth cardiac phases is shown in Figure 5.11 in three orthogonal views. It is more obvious that not all the RV dilation patterns are captured by the model and hence reflected in the different RV shapes between synthesized and manual results. The RV dilation changes the LV's shape in a complex way such that at the shown short-axis slice, the LV is ellipsoidal on the first phase but still close to circular on the eighth phase, and this shape deformation is only partially reflected in the segmentation result. Again, in spite of these localized inaccuracies, some important RV shape variability such as the dilation in the apex region is successfully captured by the model and found by the AAM segmentation with still satisfactory result.

In both the results shown in Figures 5.10 and 5.11, the PCA modes as global descriptors do show their strengths in smoothing out some artifacts such as bumpy surfaces in the manual results, which are introduced by the shape interpolation and small inaccuracies in manual tracings, and therefore produce more plausible ventricular surfaces.

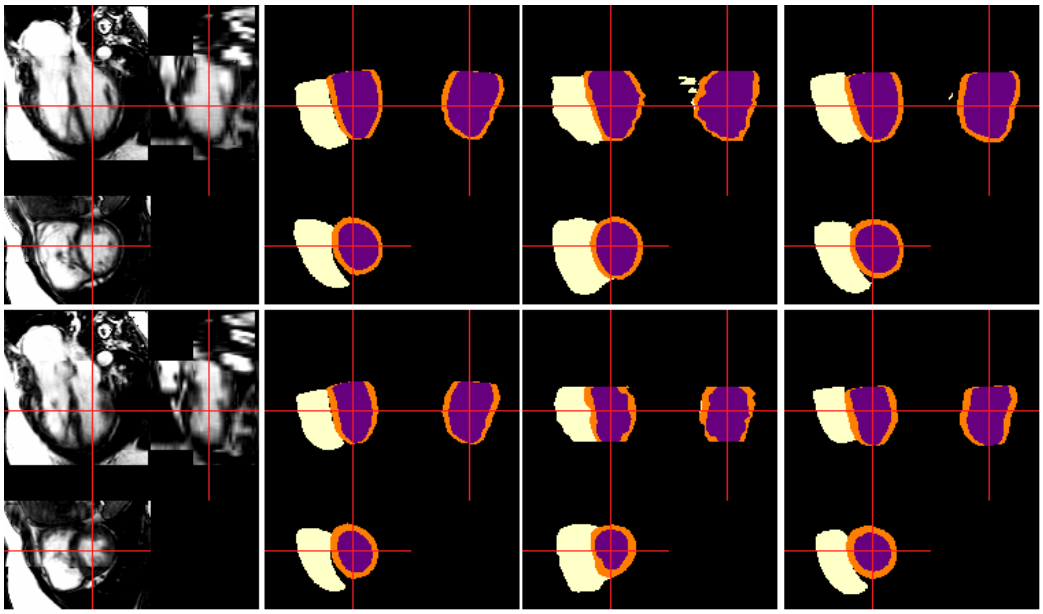



Figure 5.11: An example of a TOF heart. Top row: first phase, bottom row: eighth phase.

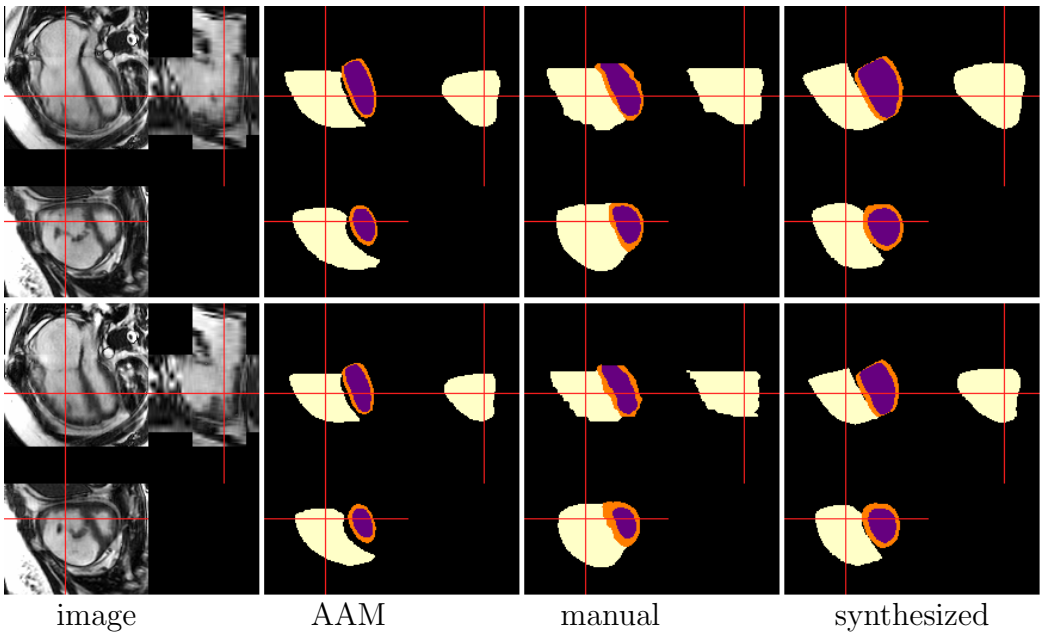



Figure 5.12: An example of a TOF heart with severe RV dilation. Top row: first phase, bottom row: eighth phase.

The image and results of a TOF patient suffers from severe RV dilation is shown in Figure 5.12. This is the only manually traced patient who have most types of RV dilation except aneurysm toward the right atrium. No matter which currently available training group is used, this TOF is always a outliner and the model cannot capture the severe shape changes, indicated by the larger difference between synthesized and manual results. However, by setting some extreme *anisotropic* scaling values in the initialization of the affine transform, the resulting segmentation is closer to the manual result than the associated synthesized result. If the independent standard of this subject is unknown, the segmentation although inaccurate provides a very good starting point for any further manual modification on it.

In summary, although mainly affected by small training population, the 4-D AAM still produced very good segmentation results in a global sense. However, the 4-D AAM did not fully capture the contraction and relaxation pattern of the heart and therefore may preventing the ventricular capacity based function indices such as ejection fraction and volume-time curve from representing the true nature of the cardiac motion.

# CHAPTER 6
# COMPUTER-AIDED DIAGNOSIS

The final goal of the cardiac MR image analysis is the computer-aided diagnosis derived from the computer segmentation results. Using the 4-D AAM segmentation results, two novel ventricular function indices were derived as PCA shape modes and ventricular volume-time curves. The capabilities of using these indices to achieve normal/TOF classification and potential TOF disease progression identification are explored in this chapter. For comparison, the capability of the conventional index, ventricular ejection fraction (EF), is also tested.

## 6.1 Experimental Methods

In addition to the 25 normal and 25 TOF scans used for the AAM segmentation validations, 10 second-year scans of normal subjects and 9 second-year scans of TOF patients were added to the study population. This 69-scan population contains 24 normal subjects and 24 TOF patients. Among them, 9 normal subjects have first- and second-year scans, one normal subject has two repeated first-year scans and one second-year scan; one TOF patient has two first-year scans of six months apart and one second-year scan. The second-year scans were segmented by the 4-D AAM methods using models created from the manual tracing results of all the 25 first-years normal/TOF scans.

From the segmentation results of manual tracings and 4-D AAM, the ejection fractions of the RV as RVEF, and the volume-time curves (VTC) of the LV, RV, and LV and RV concatenated were analyzed. The magnitudes of the VTCs were

normalized within the range of $[0, 1]$ to remove the effects of the global scaling. Note that in the concatenated LV+RV VTCs, the relative size differences in LV and RV were kept. The PCA was applied to the VTCs that were treated as vectors of length 16 or 32, and the PCA modes of the VTCs were classified.

Two statistical shape models of the combined LV and RV, one from the manual tracing results of 50 scans, one from the AAM results of the 69 scans, were created. The *isotropic* scaling was used in the *Procrustes Analysis* to keep the information about the relative size difference of LV and RV and the variability in ventricular height along the long-axis in the resulting models. The shape variation associated with the strongest PCA modes were visualized and inspected, and then the PCA modes were classified.

The normal/TOF classifications using the PCA modes (of shape or VTC) were performed using *leave-one-out* validation, and the classification methods were $k$-nearest neighbor ($k$-NN) [22,24] and linear discriminant analysis (LDA) [22,23] using various number of PCA modes. The performance of the classification was measured by classification rate – the percentage of scans correctly identified as normal or TOF patients.

The differences between the first- and second-year scans were analyzed using various number of shape PCA modes. The differences were first measured as the *absolute distance* between two scans in the shape model space. In this model space defined by PCA modes, an $M$-dimensional plane ($M$ is the number of strongest PCA modes used) that separates the normal and TOF was then found by the LDA method. The

identified plane serves as the decision boundary for normal/TOF classification, its normal direction is approximately pointing from the centroid of normal scans to that of the TOF patient scans in the model space, such that any scan with negative distance to the decision boundary is classified as normal and positive distance means TOF patient. The difference between the first- and second-year scans ($2^{nd} - 1^{st}$) is defined as the change of their distances to the identified plane, this change is referred as the *relative distance*. The definitions of absolute and relative scan distances are illustrated in Figure 6.1 in the model space of two PCA modes.

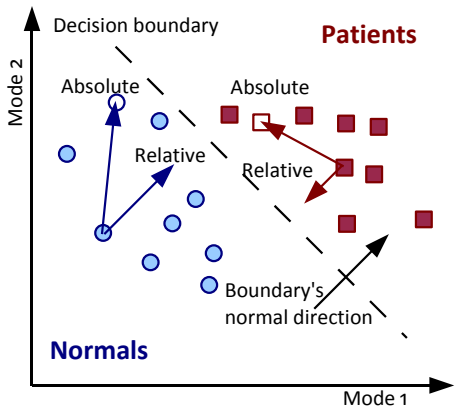
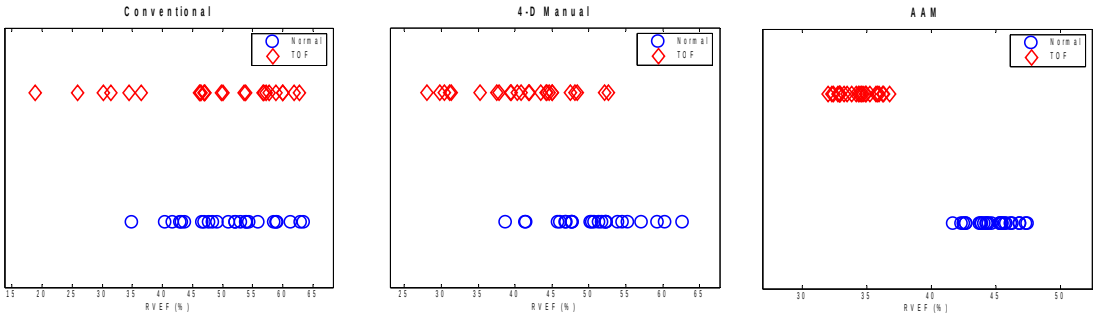



Figure 6.1: The definitions of absolute and relative scan distances.

## 6.2 RVEF Analysis

The RVEFs of 50 MR scans calculated from the conventional manual tracings, true 4-D manual tracings, and AAM segmentations are plotted in Figure 6.2 with their means and standard deviations listed. The results in Figure 6.2(a,b) agree to

the clinical fact that TOF patients generally have smaller RVEFs due to the dilated RV but the RVEF itself cannot fully distinguish normal and TOF patients. Also note that since the true 4-D tracings reduced some uncertainties in identifying RV bases, the resulting RVEFs have smaller standard deviation than that of the conventional method.




normal: 51%±7%
TOF: 48%±12%
(a) Conventional

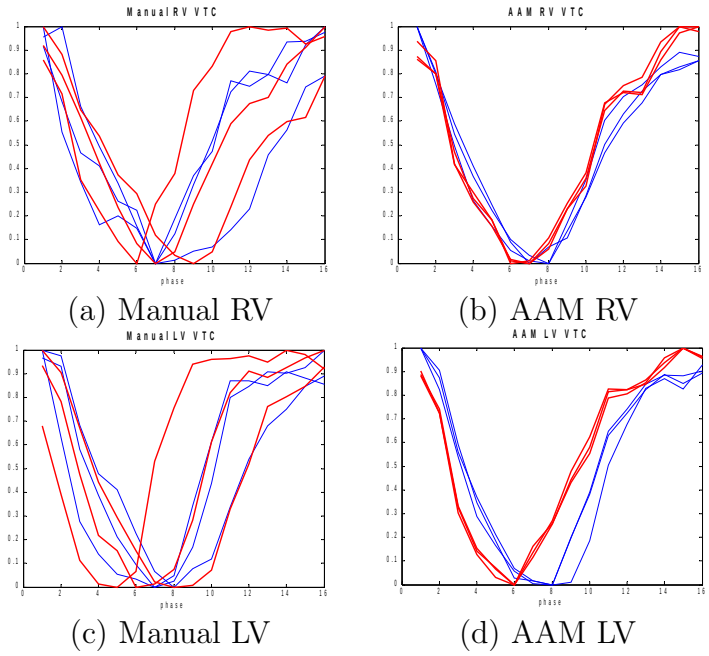
normal: 51%±6%
TOF: 41%±7%
(b) 4-D manual

normal: 45%±2%
TOF: 34%±1%
(c) AAM

Figure 6.2: Distributions of RVEFs

Although Figure 6.2c indicates that there is a clear distinction between normal and TOF patients in their RVEFs, it does not represent the real difference of normal and TOF in RVEF. As seen in the segmentation performance assessments in Chapter 5, the 4-D AAM cannot exactly follow the cardiac motion pattern. Several randomly selected VTCs of RV in Figure 6.3(a,b) further illustrates such limitation of 4-D AAM segmentations, where the large variability in the end-systolic time of the TOF RV is not fully captured by the model and consequently not reflected in the segmentations. Such limitation of the model also exists in LV as shown in Figure 6.3(c,d). In the

AAM results for both LV and RV, the end-systole of TOF hearts is always earlier than the normal hearts.




(a) Manual RV (b) AAM RV

(c) Manual LV (d) AAM LV

Figure 6.3: Several VTCs created from manual and AAM segmentation results. Red thick lines are TOF VTCs and blue fine lines are normal VTCs.

The classification performance achieved by using RVEFs is listed in Table 6.1, where the number of scans correctly classified and the classification rates are both shown. It indicates that the true 4-D manual segmentation produced better RVEF measures than conventional manual tracing such that the decreasing of RVEF in TOF patients is mostly captured. Although the AAM results achieved 100% normal/TOF classification, the RVEF measured from AAM segmentation amplified the difference between normal and TOF hearts incorrectly as described earlier.

| Method | Normal | TOF | Method | Normal | TOF |
|---|---|---|---|---|---|
| Conventional, 25 Normals, 25 TOFs | | | | | |
| 1-NN | 12(48%) | 18(72%) | LDA | 14(56%) | 12(48%) |
| Manual, 25 Normals, 25 TOFs | | | | | |
| 1-NN | 23(92%) | 22(88%) | LDA | 21(84%) | 20(80%) |
| AAM, 25 Normals, 25 TOFs | | | | | |
| 1-NN | 25(100%) | 25(100%) | LDA | 25(100%) | 25(100%) |

Table 6.1: The classification performance achieved using RVEFs.

### 6.3 VTC Analysis

The classification performance achieved using the VTCs derived from the true 4-D manual tracings is listed in Table 6.2. It suggests that the TOF changes the cardiac motions of both LV and RV. It is interesting that although most of the time the LV shows no clear deformation, its cardiac motion pattern might be changed more than that of the RV. When the VTCs of LV and RV are combined, the RV dilation is included in the combined VTCs so classification rates of 100% for normal subjects and 96% for TOF patients are achieved.

| Method | Normal | TOF | Method | Normal | TOF |
|---|---|---|---|---|---|
| Manual LV VTC, 25 Normals, 25 TOFs | | | | | |
| 1-mode 3-NN | 22(88%) | 20(80%) | 1-mode LDA | 19(76%) | 17(72%) |
| 2-mode 5-NN | 20(80%) | 22(88%) | 2-mode LDA | 19(76%) | 15(60%) |
| Manual RV VTC, 25 Normals, 25 TOFs | | | | | |
| 1-mode 1-NN | 16(64%) | 15(60%) | 4-mode LDA | 17(72%) | 15(60%) |
| 3-mode 1-NN | 17(72%) | 17(72%) | 5-mode LDA | 16(64%) | 15(60%) |
| Manual LV+RV VTC, 25 Normals, 25 TOFs | | | | | |
| 1-mode 1-NN | 22(88%) | 23(92%) | 1-mode LDA | 24(96%) | 23(92%) |
| 1-mode 5-NN | 24(96%) | 24(96%) | 2-mode LDA | 25(100%) | 24(96%) |

Table 6.2: The classification performance achieved using PCA modes of VTCs of the manual volumes.

Using the AAM segmentation results of the same 50 MR scans, the achieved

classification performance is listed in Table 6.3, where the best classification rate achieved for normal and TOF is 100%. Similarly, 100% normal/TOF classificaiton was achieved when the AAM results of all 69 MR scans were used. But again, these classification results do not represent the true difference between normal and TOF in VTCs due to the limitation of AAM segmentation in capturing cardiac motion pattern.

| Method | Normal | TOF | Method | Normal | TOF |
|---|---|---|---|---|---|
| AAM LV VTC, 25 Normals, 25 TOFs | | | | | |
| 1-mode 1-NN | 23(92%) | 22(88%) | 1-mode LDA | 24(96%) | 23(92%) |
| 2-mode 1-NN | 25(100%) | 25(100%) | 2-mode LDA | 25(100%) | 25(100%) |
| AAM RV VTC, 25 Normals, 25 TOFs | | | | | |
| 1-mode 1-NN | 25(100%) | 25(100%) | 1-mode LDA | 25(100%) | 25(100%) |
| AAM LV+RV VTC, 25 Normals, 25 TOFs | | | | | |
| 1-mode 1-NN | 24(96%) | 24(96%) | 1-mode LDA | 24(96%) | 25(100%) |
| 3-mode 1-NN | 25(100%) | 25(100%) | 5-mode LDA | 25(100%) | 25(100%) |

Table 6.3: The classification performance achieved using PCA modes of VTC of the AAM segmented volumes.

The volume-time curves only partially describe the complex cardiac motion by the changes of a global measure – the ventricular capacity. The shape of the VTCs are mainly defined by the end-diastolic and end-systolic phases but finding these phases correctly in manual tracings can be difficult. There are two ways to find them, one is to observe the changes of ventricular areas on a fixed short-axis slice and define the phases with maximal and minimal ventricular areas as end-diastole and end-systole, respectively. The other way is to calculate the ventricular volumes of all phases and define the end-diastole and end-systole as phases with maximal and

minimal volumes. In practice, the phases identified by both methods often do not exactly agree to each other. The short-axis slice method ignores the possible cardiac motion of the heart along the long-axis so the images of different phases may actually come from slightly different parts of the ventricles. The volume calculation is very sensitive to the identification of the ventricular base, which is always difficult and can dramatically change the resulting ventricular volume measures.

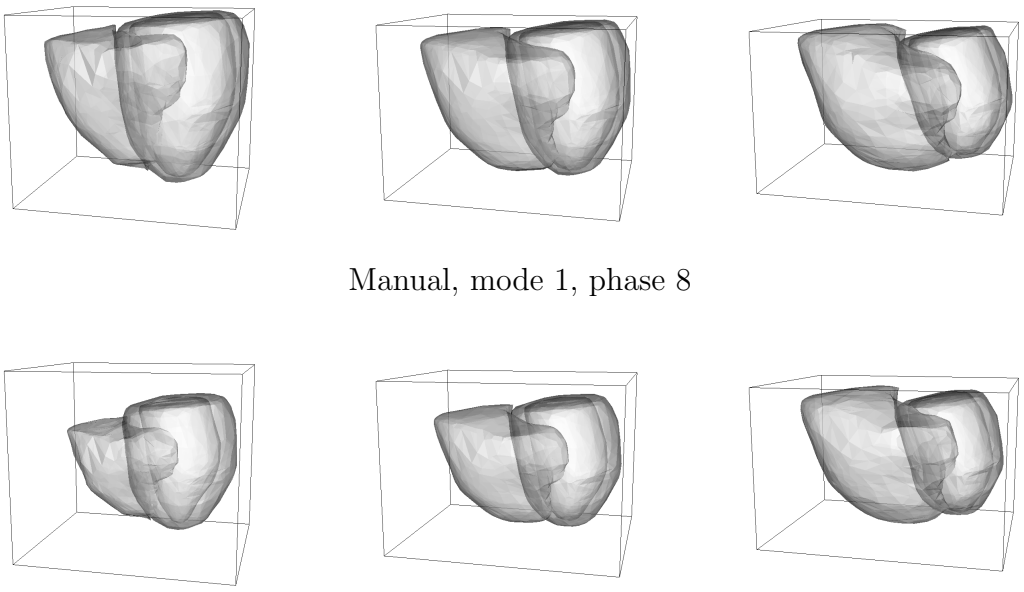
The VTCs of the manual tracings may not represent the real cardiac motion pattern and the AAM segmentation has its limitation in capturing all possible motion pattern using a small training population. Although the classification results listed above indicate the existence of a difference between the volume-time curves of the normal and TOF patients, finding the true form of this difference will require larger population, more accurate identifications of end-diastolic and end-systolic phases in both manual and computer segmentations. And such requirements also apply to any further study of the ejection fraction.

## 6.4 Shape Model Analysis

Figure 6.4 shows the shape variations associated with the first PCA mode of the shape model created from the manual tracings of 50 MR scans. It can be seen that the first PCA mode is mostly related to the variability of the size of the RV with respect to the LV so the main difference between normal and TOF is captured in this mode. Figure 6.5 shows the shape variations associated with the second mode, this mode mostly represents the variability of the ventricular height along the long-axis.

Figures 6.6 and 6.7 show the shape variations associated with the first and second

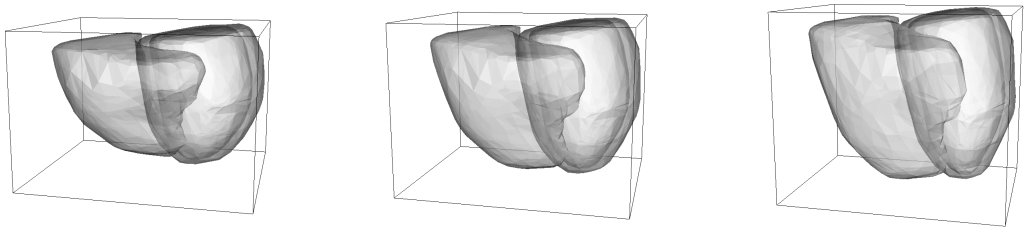
Manual, mode 1, phase 1



(a) mean-2$\sigma$ (b) mean (c) mean+2$\sigma$

Figure 6.4: Shape variations associated with the first PCA mode of the shape model created from manual tracings.

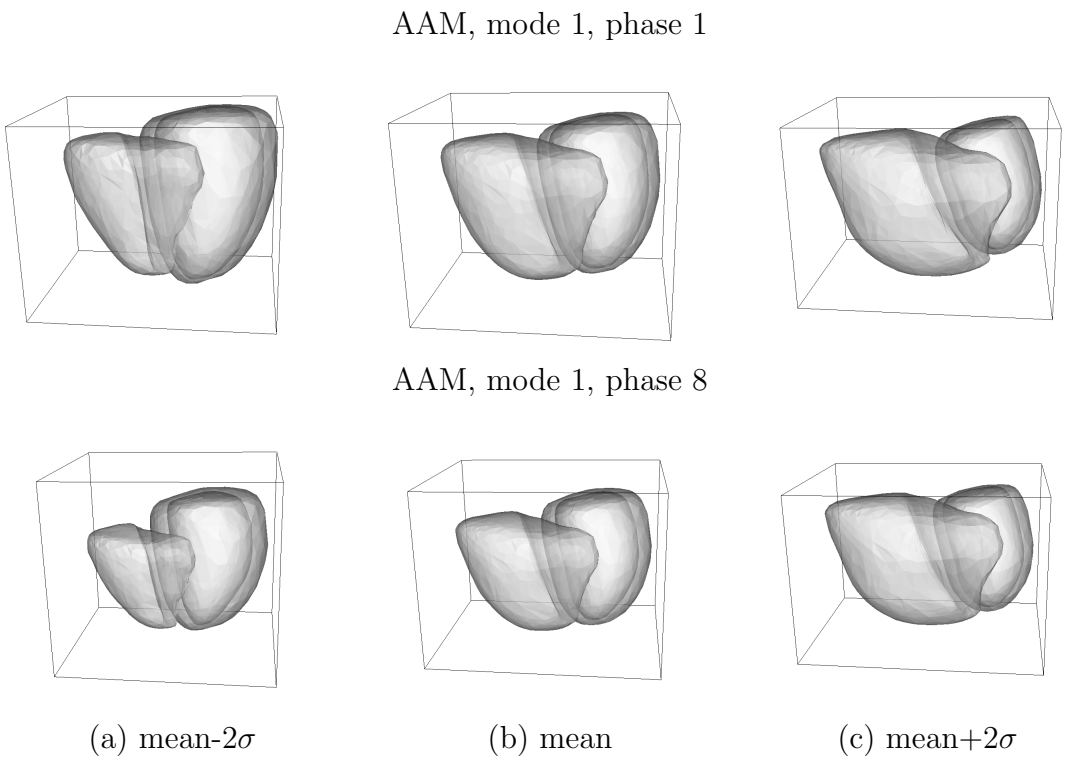
Manual, mode 2, phase 1



(a) mean-2$\sigma$ (b) mean (c) mean+2$\sigma$

Figure 6.5: Shape variations associated with the second PCA mode of the shape model created from manual tracings.

PCA modes of the shape model created from AAM segmentations of the same 50 MR scans. In spite of the various segmentation inaccuracies, the first PCA mode still represents the main difference between normal and TOF patients in a very similar way as in Figure 6.4. The shape variations associated with the second PCA mode mainly represent the changes of the relative locations of LV and RV, which are different from those in Figure 6.5.

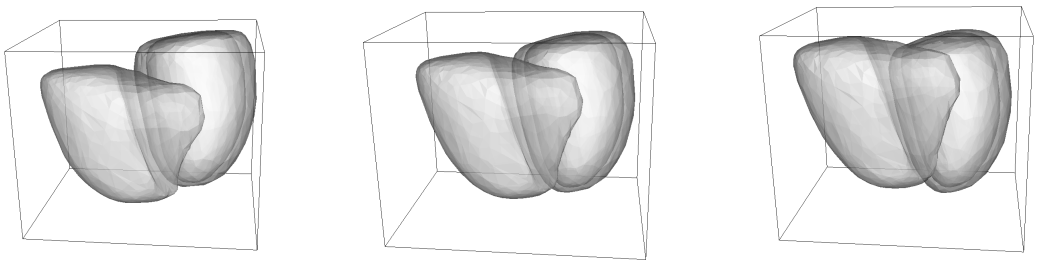



(a) mean-$2\sigma$ (b) mean (c) mean+$2\sigma$

Figure 6.6: Shape variations associated with the first PCA mode of the shape model created from AAM segmentations.

From the same set of 50 MR scans but two different shape models created from manual tracings and AAM segmentation results, the distributions of the first and

AAM, mode 2, phase 1

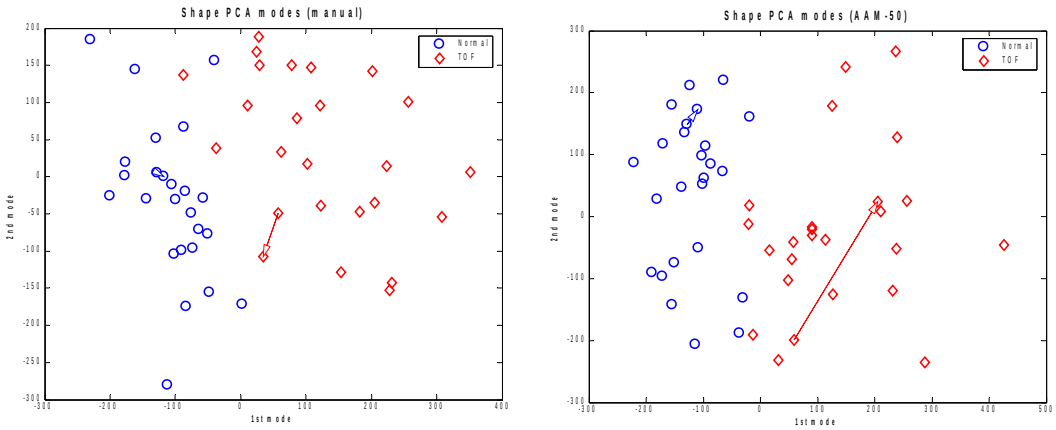


(a) mean-2$\sigma$ (b) mean (c) mean+2$\sigma$

Figure 6.7: Shape variations associated with the second PCA mode of the shape model created from AAM segmentations.

second PCA modes are shown in Figure 6.8, where the lines with arrows indicating the changes of PCA modes of the repeated scans. Although the difference of repeated TOF scans in PCA modes of AAM results is larger than that of the manual results, the shape variations associated with the second mode are different between the two models as shown in Figures 6.5 and 6.7, and the specific AAM segmentation is not guaranteed to be the optimal one. However, the distributions of shape modes show that the shape model created from the AAM segmentation results successfully captures the main difference between normal and TOF in the first shape mode. Compared with the model created from manual tracings, the gap between normal and TOF is not increased in Figure 6.8b, which means the shape PCA modes created from the AAM segmentations result do not incorrectly amplify the difference between normal and TOF hearts.

Table 6.4 lists the classification performance achieved using shape PCA modes. Unlike the classification results listed in Tables 6.2 and 6.3, there is no clear distinction




(a) Manual (b) AAM

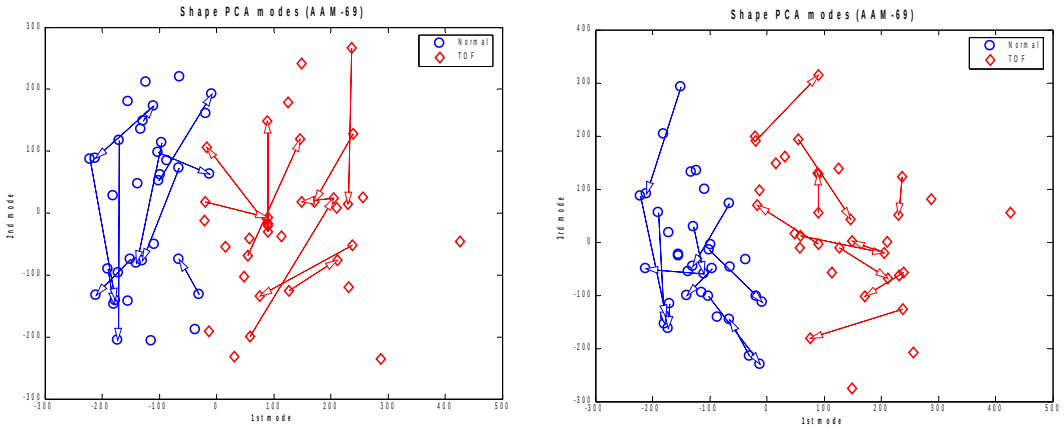
Figure 6.8: Two strongest shape PCA modes of manual and AAM results.

between the classification performances of models created from manual tracings and AAM segmentations. It shows that the shape model is a more robust ventricular feature representation than the VTCs when various segmentation inaccuracies exist. Using the manual segmentation results, all normal subjects and 92% of the TOF patients can be correctly identified. Using the AAM segmentation results, all normal and TOF subjects can be correctly classified only using five PCA shape modes.

| Method | Normal | TOF | Method | Normal | TOF |
|---|---|---|---|---|---|
| Manual, 25 Normals, 25 TOFs | | | | | |
| 1-mode 3-NN | 24(96%) | 23(92%) | 1-mode LDA | 25(100%) | 22(88%) |
| 2-mode 1-NN | 23(92%) | 23(92%) | 3-mode LDA | 25(100%) | 21(84%) |
| 3-mode 1-NN | 25(100%) | 23(92%) | 5-mode LDA | 24(96%) | 21(84%) |
| AAM, 25 Normals, 25 TOFs | | | | | |
| 1-mode 1-NN | 24(96%) | 24(96%) | 1-mode LDA | 25(100%) | 22(88%) |
| 2-mode 1-NN | 24(96%) | 24(96%) | 3-mode LDA | 25(100%) | 24(96%) |
| 3-mode 1-NN | 25(100%) | 25(100%) | 5-mode LDA | 25(100%) | 25(100%) |
| AAM, 35 Normals, 34 TOFs | | | | | |
| 1-mode 1-NN | 33(94%) | 32(94%) | 1-mode LDA | 35(100%) | 30(88%) |
| 3-mode 5-NN | 34(97%) | 32(94%) | 3-mode LDA | 35(100%) | 32(94%) |
| 5-mode 5-NN | 35(100%) | 34(100%) | 5-mode LDA | 35(100%) | 34(100%) |

Table 6.4: The classification performance achieved using shape PCA modes.

Figure 6.8 also shows that the PCA modes of TOF patients covers a larger range than those of normal subjects so the PCA modes can be potentially used as disease status indicators. Figure 6.9 shows the changes of several strongest PCA modes of the subjects with first-year and second-year scans, where the shown vectors indicating the changes from first-year to second-year scans. Although the length and direction of these vectors are approximately random. Their projections on the first mode – the most important mode for normal/TOF separation – show that the changes of normal subjects are smaller than those of the TOF patients.




(a) 1st and 2nd modes (b) 1st and 3rd modes

Figure 6.9: Distributions and changes of the three strongest shape PCA modes of AAM results.

The changes of PCA modes in the model space were measured by absolute and relative distances, whose definitions are illustrated in Figure 6.1. The distribution, reported as mean and standard deviation, of changes of PCA modes are listed in Table 6.5.

| | Absolute | | Relative | |
|---|---|---|---|---|
| Modes | Normal | TOF | Normal | TOF |
| 1 | 51±34 | 83±53 | -4±64 | 4±103 |
| 2 | 141±87 | 165±67 | 6±60 | 0±95 |
| 3 | 197±77 | 183±68 | -24±53 | -4±91 |
| 4 | 208±71 | 206±74 | -21±53 | -6±89 |
| 5 | 226±65 | 215±75 | 1±47 | -6±72 |

Table 6.5: The changes of PCA modes of subjects with multiple scans.

If only one PCA mode is used, the absolute changes of normal subjects show smaller mean and standard deviation than those of the TOF, which agree to the observations made on Figure 6.9. When more than one PCA modes are used, the difference in absolute changes of normal and TOF is almost invisible due to the fact that the AAM segmentation results are just local minimums in most of the time and the vectors of PCA mode changes are not directly related to those seen by the classifiers.

On the other hand, the relative change measures the changes from the perspective of the specific classifier and therefore better describes any potential status changes of the subjects. It can be seen from Table 6.5 that the normal subjects consistently show smaller variances of relative changes than TOF patients. Eight of the ten mean relative changes show no tendency of normal changing to TOF and vice versa. The other two larger mean relative changes are from the normal subjects with the negative values indicating a tendency of normal subjects move away from the decision surface and thus much less likely to be classified as TOF subjects.

## 6.5 Summary and Discussion

The ventricular function indices of RVEF and volume-time curves have clear and easy to understand physiological meaning. The normal/TOF classification performed using these indices indicated some difference between normal and TOF patient. However, these indices contain limited amount of information and are sensitive to the manual and computer segmentation errors. Further study of them is only possible when computer segmentations with improved accuracy, especially in capturing the cardiac motion, are available.

The shape PCA modes on the other hand turned out to be less sensitive to the segmentation errors and achieved complete normal/TOF classification. The longitudinal analysis showed that the normal subjects exhibit less variability in the shape PCA modes than TOF subjects. Therefore the PCA modes can be potentially used as TOF disease status determinants.

The capability of all these indices in computer-aided diagnosis is still limited by the current small population and segmentation inaccuracies. But the current analysis results already showed their strengths in identifying subtle difference between normal and TOF patients. If larger population is analyzed, the accuracy of AAM segmentation can be improved and consequently the performance of the CAD.

# CHAPTER 7
# CONCLUSION

## 7.1 Achieved Goals

The four goals of this work are listed in 1.4 on page 12 as: 4-D data construction, 4-D model construction, 4-D AAM segmentation, and computer-aided diagnosis. The reviews of what have been achieved are as follows.

### 7.1.1 4-D Data Construction

The image fusion procedure overcame most of the inherent limitations of cardiac MR imaging and produced high-quality 4-D ventricular images with isotropic voxels, complete ventricle coverage, and no respiratory motion artifacts. The whole procedure only requires a few user interactions of specifying the correct MR sequence description and an approximate ROI on a single phase. If only a few (two or three) long-axis images are available, the motion correction algorithm cannot completely remove all possible respiratory motion artifacts, but the algorithm is very robust and did not introduce any additional new artifacts.

The developed 4-D manual tracing application provided the user with a true 4-D context of the heart based on the high-quality 4-D image constructed by the developed image fusion technique. In this 4-D context, finding the correct locations of the ventricular bases and apexes became much easier. The direct tests on ventricle coverage improvement and indirect tests on the stroke volume agreement of normal hearts showed that the 4-D manual tracing application produced more accurate and reliable results than conventional tracings performed in a 2-D or pseudo 3-D context.

The manual tracing of ventricular images is still a time-consuming process. Tracing a complete 16-phase cardiac cycle takes approximately four to five hours depending on the image quality. Although the approximate locations of ventricular bases and apexes are easier to find in the 4-D context, determining the ventricular borders accurately in these regions is still difficult due to the variability in the ventricle orientation in the image.

The volume fusion by distance interpolation is very efficient. However, in order to generate 'valid' short-axis cross sections, an expensive 3-D smoothing step has to be performed that takes 5 to 10 minutes on a 4-D data set.

### 7.1.2 4-D Model Construction

The most difficult but essential part of the model construction is the automatic landmarking. Its most important requirement is the point correspondence. Designed within a template-based framework, the landmarking problem is translated into a series of registration problems that can be solved by many methods but the ground truth is often inaccurate or hard to define. In this 4-D case, theoretical correctness had to be sacrificed for efficiency in choosing the appropriate registration algorithm. The chosen algorithm using a cost function of simple distance measure combined with regularization constraint proved to be highly efficient and the resulting models correctly captured the important shape and texture features of the ventricles. Using these models, the segmentation and disease classification performed in later stages all achieved very good results.

The training quality of the constructed 4-D AAMs were fully tested and it was

found that the conventional model matching error metrics – RMS of texture difference – can guide the PCA appearance modes correctly toward their optimums but cannot lead the model-to-image space transform correctly due to the complex surrounding anatomic structures existed in the cardiac MR images. It caused the AAM segmentation to be very sensitive to the model initialization.

### 7.1.3 4-D AAM Segmentation

The inherent property of the 4-D AAM as a model-based segmentation method is that it only includes global descriptions of shape and appearance variabilities thus cannot achieve good local segmentation accuracy. In addition, the size of the training population directly affects the segmentation performance. The customized segmentation strategies and hybrid of AAM and ASM approach were used to overcome some of the limitations of the 4-D AAM method. The segmentation results tested on 25 normal and 25 TOF patient scans showed good segmentation accuracies measured by signed surface positioning errors. On normal subjects, the average signed errors are 0.3±2.3 mm for LV endocardial surface and 0.1±3.4 mm for RV surface. On TOF patients with large shape variability, the errors are -1.5±3.2 mm for LV endocardial surface and -0.9±4.3 mm for RV surface. Other error metrics such as relative overlapping also indicated good segmentation accuracies.

Further detailed analyses of the measured error metrics showed that the AAM segmentation cannot exactly identify the cardiac motion pattern such that the segmentation errors on the end-systolic phase were slightly larger than those on the end-diastolic phase. The variability in the ventricle orientation also affected the seg-

mentation accuracy such that the errors on the ventricular bases and apexes were slightly larger than those on the mid-ventricle sections.

Overall, the segmentation results produced by the 4-D AAM were very good and the subsequent computer-aided diagnosis achieved similar performance as that using manual tracing results when PCA shape modes were analyzed. The segmentation only needs to be initialized on the first cardiac phase with user interaction. The segmentation of a 4-D image using several (less than five) different initializations can be finished within 30 to 45 minutes on a PC with 2.4 GHz Pentium4 CPU depending on the size of the image. The required manual modifications of the computer segmentation results are mostly global fine-tunings of the rotation, scaling and translation and some local fine-tunings of surfaces at small portions of the ventricles.

### 7.1.4 Computer-Aided Diagnosis

Using the ventricular volume-time curves derived from the computer segmentation results, 100% normal/TOF classification was achieved with simple classifiers of $k$-NN and LDA. Compared with the classification results using VTCs derived from the manual tracings, it was found that the inability of following cardiac motion in the computer segmentation amplified the subtle difference between normal and TOF patients. However, the VTCs derived from the manual tracings showed that TOF changed the caridac motion patterns of both left and right ventricles and the changes were captured by VTCs.

The analyses performed on the shape PCA modes showed that they were less sensitive to the various segmentation errors. The main difference of normal and TOF

patients was correctly captured by the strongest shape modal index of both models created from manual and computer segmentation results and was not amplified by the computer segmentation. The distribution of shape modal indices showed a clear distinction between normal and TOF patients and 100% normal/TOF classification was achieved.

Further longitudinal analyses performed on subjects with multiple annual MR scans showed that the TOF patients exhibited larger variances in the changes of the shape modal indices than normal subjects, which demonstrated the potential of using the modal indices as disease status determinants.

## 7.2 Summary

In conclusion, all the four proposed goals of this thesis were achieved. A complete 4-D cardiac MR image analysis pipeline was implemented. The pipeline input is a set of (usually hundreds of) 2-D MR images scanned and stored as DICOM files. It not only produced accurate 4-D segmentation of both left and right ventricles, but also produced the computer-aided results in the form of novel 4-D ventricular function indices that can be used to achieve 100% normal/TOF classification, identify subtle but important disease characteristics, and potentially early detection of disease progression.

## REFERENCES

[1] R. H. Anderson and M. Tynan, "Tetralogy of Fallot – a centennial review," *International Journal of Cardiology*, vol. 21, pp. 219–232, 1988.

[2] R. H. Anderson and P. M. Weinberg, "The clinical anatomy of tetralogy of Fallot," *Cardiol. Young*, vol. 15, pp. 38–47, 2005.

[3] M. M. Bartelings and A. C. G. de Groot, "Morphogenic considerations on congenital malformations of the outflow tract part 1: common arterial trunk and tetralogy of Fallot," *International Journal of Cardiology*, vol. 32, pp. 213–230, 1991.

[4] P. Berthout, J. C. Cardot, M. Baud, R. Faivre, J. Verdenet, A. C. Bidet, J. P. Bassand, R. Bidet, and J. P. Maurat, "Factors influencing the quantitication of valvular regurgitation by gated equilibrium radionuclide angiography," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 9, pp. 112–114, 1984.

[5] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, pp. 239–256, 1992.

[6] T. N. Bloomer, S. Plein, A. Radjenovic, D. M. Higgins, T. R. Jones, J. P. Ridgway, and M. U. Sivananthan, "Cine MRI using steady state free precession in the radial long axis orientation is a fast accurate method for obtaining volumetric data of the left ventricle," *Journal of Magnetic Resonance Imaging*, vol. 14, pp. 685–692, 2001.

[7] T. Blu, P. Thévenaz, and M. Unser, "Linear interpolation revitalized," *IEEE Transactions on Image Processing*, vol. 13, no. 5, pp. 710–719, May 2004.

[8] E. Catmull and R. Rom, "A class of local interpolating splines," in *Computer Aided Geometric Design*, 1974.

[9] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image Vision Comput.*, vol. 10, no. 3, pp. 145–155, 1992.

[10] T. F. Cootes, D. Cooper, C. J. Taylor, and J. Graham, "A trainable method of parametric shape description," in *The British Machine Vision Conference*, 1991, pp. 54–61.

[11] T. F. Cootes, D. H. Cooper, C. J. Taylor, and J. Graham, "Trainable method of parametric shape description," *Image Vision Computing*, vol. 10, no. 5, pp. 289–294, 1992.

[12] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *European Conference on Computer Vision*, 1998, pp. 484–498.

[13] ——, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, 2001.

[14] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," *Image and Vision Computing*, vol. 12, no. 6, pp. 355–366, 1994.

[15] T. F. Cootes and C. J. Taylor, "Active shape models – "smart snakes"," in *The British Machine Vision Conference*, 1992, pp. 266–275.

[16] ——, "Active shape model search using local grey-level models: A quantitative evaluation," in *The British Machine Vision Conference*, 1993, pp. 639–648.

[17] ——, "Modelling object appearance using the grey-level surface," in *The British Machine Vision Conference*, 1994, pp. 479–488.

[18] T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham, "Training models of shape from sets of examples," in *The British Machine Vision Conference*, 1992, pp. 9–18.

[19] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Trans. Med. Imag.*, no. 5, pp. 525–537, 2002.

[20] J. F. Debatin, S. S. Nadel, H. D. Sostman, C. E. Spritzer, A. J. Evans, and T. M. Grist, "Magnetic resonance imaging – cardiac ejection fraction measurements: Phantom study comparing four different methods," *Investigative Radiology*, vol. 27, no. 3, pp. 198–204, 1992.

[21] C. Dornier, M. K. Ivancevic, P. Thevenaz, and J. P. Vallee, "Improvement in the quantification of myocardial perfusion using an automatic spline-based registration algorithm," *Journal of Magnetic Resonance Imaging*, vol. 18, no. 2, pp. 160–168, 2003.

[22] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification.* John Wiley & Sons, Inc., 2001.

[23] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[24] E. Fix and J. L. Hodges, "Discriminatory analysis, nonparametric discrimination: Consistency properties," USAF School of Aviation Medicine, Tech. Rep., 1951.

[25] A. Frangi, D. Rueckert, J. Schnabel, and W. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *IEEE Trans. Med. Imag.*, vol. 21, pp. 1151–1166, 2002.

[26] E. J. Gandsman, D. L. North, R. S. Shulman, and E. W. Bough, "Measurement of the ventricular stroke volume ratio by gated radionuclide angiography," *Radiology*, vol. 138, pp. 161–165, 1981.

[27] M. Garland and P. S. Heckbert, "Surface simplification using quadric error metrics," in *SIGGRAPH 97 Proceedings*, 1997, pp. 209–218.

[28] M. A. Gatzoulis, G. D. Webb, and P. E. F. Daubeney, *Diagnosis and Management of Adult Congenital Heart Disease*. Churchill Livingstone, 2003.

[29] G. Goodall, "Procrustes methods in the statistical analysis of shape," *Journal of the Royal Statistical Society B*, vol. 53, no. 2, pp. 285–339, 1991.

[30] R. Harshman, P. Ladefoged, and L. Goldstein, "Factor analysis of tongue shapes," *The Journal of the Acoustical Society of America*, vol. 62, no. 3, pp. 693–707, 1977.

[31] H. Hoppe, "New quadric metric for simplifying meshes with appearance attributes," in *IEEE Visualization 1999*, 1999, pp. 59–66.

[32] T. Ibrahim, C. Weniger, and M. Schwaiger, "Effects of papillary muscles and trabecule on left ventricular parameters in cine-magnetic resonance images(MRI)," *Journal of Cardiovascular Magnetic Resonance*, vol. 1, pp. 1–6, 1999.

[33] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer Verlag, 1986.

[34] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *Computer Graphics*, vol. 21, pp. 163–169, 1987.

[35] C. R. Maurer, R. Qi, and V. Raghavan, "A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, pp. 265–270, 2003.

[36] 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. Imag.*, vol. 21, pp. 1167–1178, 2002.

[37] S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. C. Reiber, and M. Sonka, "Multistage hybrid active appearance model matching: Segmentation of left and right ventricles in cardiac MR images," *IEEE Trans. Med. Imag.*, vol. 20, no. 5, pp. 415–423, 2001.

[38] P. M. Pattynama, J. Doornbos, J. Hermans, E. E. van der Wall, and A. de Roos, "Magnetic resonance evaluation of regional left ventricular function: Effects of through plane motion," *Investigative Radiology*, vol. 27, no. 9, pp. 681–685, 1992.

[39] D. Perperidis, R. H. Mohiaddin, P. J. Edwards, and D. Rueckert, "Segmentation of cardiac MR and CT image sequences using model-based registration of a 4D statistical model," in *Medical Imaging 2007: Proceedings of the SPIE*, 2007.

[40] U. Pinkall and K. Polthier, "Computing discrete minimal surfaces and their conjugates," *Experimental Mathematics*, vol. 2, no. 1, pp. 15–36, 1993.

[41] S. P. Raya and J. K. Udupa, "Shape-based interpolation of multidimensional objects," *IEEE Trans. Med. Imag.*, vol. 9, no. 1, pp. 32–42, 1990.

[42] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Prentice-Hall, 1995, ch. 4.

[43] W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-Oriented Approach To 3D Graphics*. Kitware, Inc., 2004.

[44] M. B. Stegmann, H. Olafsdottir, and H. B. W. Larsson, "Unsupervised motion-compensationof multi-slice cardiac perfusion MRI," *Medical Image Analysis*, vol. 9, no. 4, pp. 394–410, 2005.

[45] M. B. Stegmann and D. Pedersen, "Bi-temporal 3D active appearance models with applications to unsupervised ejection fraction estimation," in *Medical Imaging 2005: Proceedings of the SPIE*, vol. 5747, Apr 2005, pp. 336–350.

[46] C. M. Swingen, R. T. Seethamraju, and M. Jerosch-Herold, "Feedback-assisted three-dimensional reconstruction of the left ventricle with MRI," *Journal of Magnetic Resonance Imaging*, vol. 17, pp. 528–537, 2003.

[47] G. Taubin, T. Zhang, and G. Golub, "Optimal surface smoothing as filter design," IBM, Tech. Rep. RC-20404, 1996.

[48] H. C. van Assen, M. G. Danilouchkine, F. Behloul, H. J. Lamb, R. J. van der Geest, J. H. C. Reiber, and B. P. F. Lelieveldt, "Cardiac LV segmentation using a 3D active shape model driven by fuzzy inference," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, 2003, pp. 533–540.

[49] A. Zeballos and M. Manka, "Tetralogy of Fallot." [Online]. Available: http://www.emedicinehealth.com/articles/11205-1.asp

[50] Z. Zhao, "Novel 3D statistical shape models for segmentation of medical images," Ph.D. dissertation, Nanyang Technological University, Singapore, 2005.