# Left ventricle functional analysis in 2D+t contrast echocardiography within an atlas-based deformable template model framework


**Ramón Casero Cañas**

Wolfson Medical Vision Laboratory,

Department of Engineering Science,

University of Oxford

A thesis submitted for the

degree of Doctor of Philosophy

at the University of Oxford

Trinity Term 2008

# CONTENTS

# Left ventricle functional analysis in 2D+t contrast echocardiography within an atlas-based deformable template model framework

Ramón Casero Cañas, Oriel College

For the degree of Doctor of Philosophy, Trinity Term 2008

## Abstract

This biomedical engineering thesis explores the opportunities and challenges of 2D+t contrast echocardiography for left ventricle functional analysis, both clinically and within a computer vision atlas-based deformable template model framework.

A database was created for the experiments in this thesis, with 21 studies of contrast Dobutamine Stress Echo, in all 4 principal planes. The database includes clinical variables, human expert hand-traced myocardial contours and visual scoring.

First the problem is studied from a clinical perspective. Quantification of endocardial global and local function using standard measures shows expected values and agreement with human expert visual scoring, but the results are less reliable for myocardial thickening.

Next, the problem of segmenting the endocardium with a computer is posed in a standard landmark and atlas-based deformable template model framework. The underlying assumption is that these models can emulate human experts in terms of integrating previous knowledge about the anatomy and physiology with three sources of information from the image: texture, geometry and kinetics.

Probabilistic atlases of contrast echocardiography are computed, while noting from histograms at selected anatomical locations that modelling texture with just mean intensity values may be too naive. Intensity analysis together with the clinical results above suggest that lack of external boundary definition may preclude this imaging technique for appropriate measuring of myocardial thickening, while endocardial boundary definition is appropriate for evaluation of wall motion.

Geometry is presented in a Principal Component Analysis (PCA) context, highlighting issues about Gaussianity, the correlation and covariance matrices with respect to physiology, and analysing different measures of dimensionality.

A popular extension of deformable models —Active Appearance Models (AAMs)— is then studied in depth. Contrary to common wisdom, it is contended that using a PCA texture space instead of a fixed atlas is detrimental to segmentation, and that PCA models are not convenient for texture modelling.

To integrate kinetics, a novel spatio-temporal model of cardiac contours is proposed. The new explicit model does not require frame interpolation, and it is compared to previous implicit models in terms of approximation error when the shape vector changes from frame to frame or remains constant throughout the cardiac cycle.

Finally, the 2D+t atlas-based deformable model segmentation problem is formulated and solved with a gradient descent approach. Experiments using the similarity transformation suggest that segmentation of the whole cardiac volume outperforms segmentation of individual frames. A relatively new approach —the inverse compositional algorithm— is shown to decrease running times of the classic Lucas-Kanade algorithm by a factor of 20 to 25, to values that are within real-time processing reach.

In memoriam, grandpa Ramón Casero Olmedo (1916-2002).

And to granny, Rosario Díaz Santos.

# Acknowledgements

First I would like to thank my grandparents Ramón and Rosario, to whom this thesis is dedicated, for their unconditional love and for being such extraordinary role models.

My gratitude to my supervisor Prof. Alison Noble, for giving me the opportunity to work in this project, her guidance, comments and corrections. To Prof. Andrew Zisserman, whose unwavering enthusiasm for scientific discussion was a marvellous source of inspiration, that showed me the way. To my examiners, Prof. Sir Mike Brady and Dr. Boudewijn Lelieveldt, for a very enjoyable viva, their corrections and suggestions. To the sponsors who contributed financially to my research: The Engineering and Physical Sciences Research Council (EPSRC), Mirada Solutions Ltd., Oriel College, and the Department of Engineering Science, University of Oxford. To Dr. Javier Bermejo, Prof. Harald Becher and Dr. Jon Timperley, for their clinical mentoring and data collection. To Jean Lofts, who took care of everything IT in the lab and was always ready to give a hand. To the esteemed colleagues who contributed with their comments: Dr. Simon Baker, Dr. Alejandro Frangi, Dr. Tom Minka, Dr. Miguel Mulet, and Dr. Peter Kohl.

To OSS Watch, where I happily worked for the last 2 years of my DPhil. Especially to my bosses, Randy Metcalfe and Ross Gardler. If Randy gave me a job when I was in trouble, and got me back on my feet, Ross saw that I did not stop walking, more even so

towards the end of this thesis. And to the rest of the team, Elena Blanco, Rowan Wilson, Stuart Yeates, James Wilson, Gabriel Hanganu, Sebastian Rahtz, Michael Fraser, and the Oxford University Computing Services (OUCS) staff at large, who were so supportive and great to work with.

To Dr. Carlos Bousoño, Dr. Sancho Salcedo, Prof. Aníbal Figueiras and Dr. Manel Martínez, for awakening my interest in research back in Madrid. To Prof. Jörg Mühlbacher, for his guidance and mentoring in Linz.

To my friends in Oxford, who were with me in joy and sorrow. Yves, Olivia and little Eva Martelli, my surrogate family here all these years; for so many good memories, everything from Calvados to nappy changing was a blast. Kate Hibbs, for being wonderful, for all your support, looking forward to more hot dates at G&D's. Joel Leong, brilliant in his many manifestations, but none better than his friendship. Rohan Loveland, whose charming company made up for the short length of his stays. The Gluttony Society, those gourmands in Oriel College who every week gathered secretly and tried breathtaking, sophisticated, exotic, international dishes in a dinner followed by sensual excess like chocolate cake and philosophy: Sasha Svyryd, Sumalee Khosla, Amber Riaz, Adél Bothma, András Szigeti, Tanya Balinska, Peter Rauschenberger, Yuki Kawaguchi and the Suvan family. To Thabit Al-Murani, for his kindness, pilates training, and his sportsmanship in defeat at *Puerto Rico*. James Ambrose, the best of housemates despite the whale songs. Ellen Helsper, for her help with the national survey, and the good fun.

To my uncles Pedro Joaquín Casero and Jesús Vicente Cañas, and my friend May Jimeno, who besides their constant love and support, really helped me to focus when everything went tits up with this thesis.

To my mates from the Medical Vision Lab (MVL), Argonauts or otherwise, for making research more fun: Bobby Ali, Nate Cahill, Vicente Grau, Nick Hughes, Niranjan Joshi, Mike Kadour, Cat and Matt Kelly, Sarina Mansor, Andrew McLennan, Olivier Noterdaeme, Ingmar Posner, Dominique van de Sompel, Grace Vesom, Etienne von Lavante and Carolina Wessel. To Adrian Nish, Mike Hunt, Ed Tarleton, Mike Clements, Farrah

Bhatti, Susan Valladares, Nazar Fedorchuk, and the rest of the Oxford bunch. To the *An ideal husband* crew, for a fantastic experience: Paul Tosio, Naomi Weiss, Samuel Evans, Edward Tyerman, Luba Mandzy, Susan Harris-Huemmert, David McGrath, Olly Woodford, Kathleen Mountjoy and Barbara Whitfield. To Silvia González, Oihana Cabodevilla, Mariona Lleal, Isabel Lasala, Lidia de Tienda, Jesús Vega, Sulzhan Bali and the rest of the NOOC people.

To my parents Ramón and Rosario, for their love, for bringing me up and encouraging me to work hard. To Ana and Elena, for being such cool, talented and lovely sisters. To the rest of my uncles and aunts, Miguel Ángel, María José, Ana Mari, Edmundo, José Ángel, Eli, Matilde, Satur, Luis, Rosa and Pilar, for taking care of me all these years. To the rest of my large extended family, cousins, uncles and aunts in several degrees, and Pepe Pérez and family from Chiclana.

To my best friend, Jaime Farinos, for all these years of being always there, for our trips, the banter, the support. To the rest of my seguidmechicos family, for a life together: Javier Díez, Manuel Freire, César Bustos, Samuel Llano, Martín Fernández, Mar Gallego, Macarena Jiménez and little Beatriz Bustos. To Eva González, because nobody has talked more about so many things for so long as we have, for all the things we have shared, and the things we have to share. To Guillermo López Gallego, a constant stream of ingenuity, and a very dear friend. To María Muñoz, my lighthouse, for all the love. To Serena Celletti, Nadia Mirante, Goran Sarajlic and the rest of the Erasmus mates. To Irantzu Bilbao, Sonia Cabistany and Aitor Olea. To the Carlos III gang, Daniel García, Pablo Barrera, Guillermo López Acón, Raúl Ferrero, Gabriel Cuervo, Sara García, Ruth Ronda, Eduardo Morgado, Nacho Ballesteros, Emilio Lastra, Miguel Ángel Cano, Jorge Jiménez, Carlos García, Blanca Ferrero, Sonia Madrid, Gorka Guardiola, Raquel Manso, Sara Pérez, Carlos Sánchez, Álvaro Pruneda, Iván Moreno and Sergio Horcajo. To Laura Mabberley, a friend beautiful inside and outside. To Lora Appel and family, for an unforgettable time past.

# CHAPTER 1

# Introduction

## 1.1 Overview

This thesis explores the opportunities and challenges of 2D+t echocardiography for Left Ventricle (LV) functional analysis, both clinically and as a computer vision problem.

The problem I tackle in particular is automatically finding (or *segmenting*) smooth approximations to the LV myocardial boundaries in 2D+t Power Modulation contrast echocardiography cine loops. I then use those smooth boundaries to compute quantitative measures of cardiac function such as Ejection Fraction, endocardial motion and myocardial thickening. I discuss said quantitative measures and compare them to visual functional scoring of each muscular segment, a qualitative measure commonly used by clinicians.

In this thesis, the problem of myocardial boundary segmentation is formulated in an atlas-based deformable template model framework. This framework was chosen because 1) it is widely used and well understood in computer vision; 2) it allows to integrate prior knowledge about image pixel intensities, cardiac geometry variability and cardiac kinetics; and 3) it is fast compared to other segmentation methods (e.g. level sets).

Within the standard formulation, prior knowledge about pixel intensities is encoded into a fixed atlas. Prior knowledge about cardiac geometry variability is encoded into a template, the parameters of a similarity transformation, and a Principal Component Analysis (PCA) shape space.

Furthermore, I explore two extensions to the standard formulation. First, the Active Appearance Model (AAM), very popular in the field, gives flexibility to the atlas intensity values creating a PCA texture space. It has been assumed in the literature that such space provides better segmentation than a fixed atlas, but I contend that this is not necessarily the case. I also provide an exhaustive discussion of other key aspects of the AAM that is missing in the literature.

And second, I propose a novel temporal extension to the standard formulation that allows to integrate prior information about cardiac kinetics. I compare this extension, that models time as an explicit variable, to an existing implicit formulation. I also evaluate the improved performance when segmenting whole cardiac cycles with the temporal extension compared to segmenting individual frames.

In terms of implementation, I discuss the speed up provided by the inverse compositional algorithm approach over the classic Lucas-Kanade algorithm. All automatic segmentation results are validated against a database of hand segmentations traced by a human expert.

To sum up, the main contributions of this thesis are a database of contrast echocardiography cine loops, with expert hand segmentation of the myocardium and visual functional scoring of each muscular segment; a novel temporal extension of PCA shape models; an in-depth critique of AAMs; and the implementation of the extended shape model using the inverse compositional algorithm.

## 1.2 Motivation

The motivation for this thesis arises from the following scenario. The UK is one of the regions in Europe with highest cardiovascular mortality rates [129]. Ischaemic Heart Disease (IHD) —caused by reduced blood supply to the myocardium or cardiac muscle— alone accounted for 17% of all deaths in 2005 [66]. A widely used tool for non-invasive diagnosis of cardiac disease is echocardiography —an imaging modality that uses ultrasound to display real-time images of the heart— due to its low cost, safety, availability and ease of use.

Current guidelines for the clinical application of echocardiography state that the most common recommendation is evaluation of Left Ventricle (LV) function [40], one of the reasons being that LV diastolic disorders are an early indicator of coronary heart disease [173]. Function is usually assessed visually or quantitatively from endocardial motion and myocardial thickening. However, one of the main drawbacks of echocardiography is image quality, as spurious echos render the endocardial border as a 'blurred haze' (e.g. [12, sec. 1.1.1]) and in up to 20% of studies do not allow good enough visualisation (e.g. [186]). So-called contrast agents can be injected into the vascular system, to safely improve image quality. But even so, human expert evaluation requires extensive training [144], is resource intensive, and is subject to error.

Therefore, availability of computerised tools that could help human experts in the evaluation of LV function from contrast echocardiography, like the tools studied in this thesis, would be of great value. In this respect, automated segmentation and functional quantification would make evaluation faster, cheaper, more widely available, more reproducible and more robust. That in turn would enable earlier detection and better assessment of IHD, information that would ultimately benefit a large number of patients.

## 1.3 Breakdown by chapter

### Ch. 2: Data

This chapter introduces the necessary background about two-dimensional Power Modulation contrast echocardiography cine loops, and presents the database created for the experiments in this thesis.

Echocardiography imaging basically consists in sweeping a region of interest in the heart with a beam of ultrasonic waves, and processing and displaying the resulting echos on a screen. To improve the image quality, a contrast agent can be injected into the patient. Contrast agents are gas-filled microbubble suspensions that amplify echos and substantially improve endocardial definition. Power Modulation is a technique that cancels echos from tissue, further improving border definition. Moreover, it enables myocardial perfusion visualisation. This feature gave rise to a clinical hypothesis tested in this thesis, whether perfusion highlights the myocardium well enough to enable segmentation of the LV external boundary.

The database contains 21 studies of Power Modulation contrast Dobutamine Stress Echo B-mode cine loops of 1 cardiac cycle each, in all 4 principal planes —2-chamber (2C), 3-chamber (3C), 4-chamber (4C) and short axis (SAX) at the papillary muscles level. The database includes clinical variables (age, sex, history of myocardial infarction, ...), human expert hand-traced smooth approximations to the myocardial contours, anatomical landmarks and visual functional scoring (normokinetic, hypokinetic, akinetic, dyskinetic).

The cine loop acquisition protocol is explained in its clinical context. Visual scoring and hand tracing of myocardial boundaries were performed off-line by a Level 3 expert, and used as the criterion standard (gold standard). Hand tracing took for two months to complete, being by far the main limiting factor in the database size.

Contours were traced with a software application that allows to draw approximating splines on ultrasound frames. Approximating splines were later converted to interpolating splines, and interpolating splines were converted to configurations of anatomical

landmarks and pseudo-landmarks, a convenient representation for the geometry models used in the rest of this thesis.

## Ch. 3: Clinical evaluation

Current scientific consensus highlights the importance of LV functional evaluation using fundamental imaging echocardiography, whether to assess the health of the whole LV (global function), or of individual myocardial segments (local function).

This chapter reviews clinical measures of LV function, and discusses how they can be computed from smooth contours like those in the database. This sets the scene for the rest of this thesis, that tackles the problem of automatically segmenting myocardial contours.

Clinical best practice for functional analysis is compiled from guidelines and standards from the American College of Cardiology, American Heart Association and American Society of Echocardiography.

Standard measures are computed for global (Ejection Fraction) and local (Fractional Area Change) endocardial function, the latter on a standard 16-segment model. Local function is evaluated from wall thickening too. Experiments compare these quantitative measures to expert visual scoring.

Four issues are inspected: 1) Whether local functional evaluation is always better than global evaluation; 2) whether Power Modulation contrast echo enables both endocardial motion and wall thickening assessment; 3) merits and demerits of endocardial motion and wall thickening assessment; and 4) limitations and caveats to translate functional measures into diagnostic information about cardiac disease.

Finally, six patients with outlying EF measures are discussed in depth, illustrating the limitations of contrast 2D echocardiography functional evaluation.

## Ch. 4: Cardiac segmentation and deformable models

Myocardial segmentation in contrast echocardiography cine loops is a hard problem. Drawing inspiration from human experts, this thesis aims to integrate three sources of information into the segmentation algorithm: texture, geometry and kinetics. This chapter focuses on texture and geometry, while kinetics will be dealt with in Ch. 5.

Ultrasound images are strongly textured with speckle, a deterministic interference pattern. Training set images are aligned and warped to compute intensity histograms for selected anatomical locations in the SAX plane, and discussed with regard to Gaussianity, distribution tails, and the effect of the variance in the segmentation algorithm residuals.

As advanced in Ch. 2, myocardial geometry is represented with (continuous) smooth interpolating splines and anatomical landmarks traced by an expert. The spline control points provide the landmarks and pseudo-landmarks for a (discrete) point configuration representation, without loss of information. Contours are aligned and mapped using Procrustes analysis to a mean shape, so that location, size and rotational effects are removed.

Principal Component Analysis (PCA) is introduced, as a popular method in the literature to compute a space of shape variability called the Point Distribution Model (PDM).

Dimensionality of the shape model is assessed under two criteria: the variance criterion from PDM literature and the Generalised Information Criterion. The former uses an arbitrary threshold, while the latter estimates the threshold from the noise level in the data. I propose an anatomical/functional criterion that computes the threshold such that the maximum approximation error for the model of the training contours is 1.5 pixels, as myocardial thickness can typically be as small as 15 pixels in end diastole.

PCA assumes a multi-dimensional distribution of the data. This assumption is found to be invalid for the shape data. I propose a method that gaussianises each shape variable to improve the overall Gaussianity of the data. The validity of this method is checked in terms of the compactness of the 2D shape model, and skewness and kurtosis measures on

the PCA projections of the data.

The segmentation framework of deformable template models is presented, and different templates are compared, for instance the Active Shape Model (ASM) and Border Active Appearance Model (AAM) templates. In order to integrate previous information about myocardial ultrasound texture, atlas-based deformable models are introduced. A probabilistic atlas is a measure of a statistic for each sampling point in the template, typically the mean. The statistic is computed with the same alignment and warp process mentioned above for ultrasound histograms. Atlases for contrast 2D echocardiography are computed and discussed with regard to the feasibility of endocardial and external wall segmentation in Power Modulation contrast echocardiography.

AAMs are actually an extension of atlas-based models, where the atlas is allowed to change within a PCA texture space. A review of AAM literature shows their popularity in the field. The rest of sections in this chapter present an in-depth analysis of specific components of deformable models, with special emphasis on AAMs.

First, intervolume intensity normalisation is reviewed. I propose to compute its parameters from the whole cine loop, instead of individual frames, to avoid the bias caused by the larger number of bright pixels in end diastole than end systole.

Combining different types of variables in PCA models is examined next, and it is contended that a simple PCA on combined variables is equivalent to PCA on the coefficients of separate PCA models for each variable. These results will be useful to model kinetics in Ch. 5.

PCA is usually applied to the covariance matrix of the data, but the correlation matrix can be used too. The study of both cases illustrates physiological relationships between landmarks and the role of variable scaling. The latter is used to propose a scaling factor for combination of variables that will be used in Ch. 5 to model kinetics too.

A comparison between AAM and atlas-based segmentation is presented. While AAMs have been used in the literature for over a decade as segmentation algorithms, I could not find published experimental evidence that they perform better than atlas-based de-

formable models. Based on the discussion of the projected-out formulation for the AAM, and segmentation experiments, I contend that atlas-based segmentation actually may outperform AAM segmentation.

A further problem with AAMs is the feasibility of computing texture models in practice. The dimensionality of shape data is low enough for reasonably sized training data sets, but texture vectors are at least one or two orders of magnitude larger. Practical requirements to build PCA texture models are examined, in terms of time, human resources and computing power.

Based on the results from this chapter, segmentation will be formulated as an atlas-based deformable model algorithm in Ch. 6.

### Ch. 5: Temporal extension of cardiac PCA shape models

This chapter presents the integration of prior information on cardiac kinetics by extending the PCA 2D shape model to 2D+t.

Previous 2D+t implicit models tackled this problem by interpolating all data volumes to the same number of frames, and concatenating the shape vectors from each frame in a cine loop together. In addition, concatenation reduces the size of the training data set while increasing the number of variables to model.

To overcome those issues, I propose an explicit 2D+t cyclic shape model, that generates shapes as a function of shape coefficients and time. Linear time is mapped to two temporal variables in quadrature to obtain a cyclic model. This is also explained from a Kernel PCA (KPCA) perspective as data linearisation in feature space. Shape and temporal variables are combined with the scaling factor from Ch. 4.

I propose an extension of 2D Procrustes alignment to 2D+t, so that pose and scale are still removed, but kinetic effects are preserved. I also propose a reparameterisation of time to reflect asymmetries in the cardiac contraction cycle.

The model is assessed visually, in terms of its physiological plausibility. It is also

evaluated quantitatively and compared to the implicit model under two different scenarios: 1) all frames have different shape vectors and 2) all frames from a volume share the same shape vector.

Finally, two observed deviations from the expected behaviour that the 2D+t model does not accommodate are studied: out of plane data, and the effect of functional abnormalities on the temporal variable.

### Ch. 6: Atlas-based deformable model segmentation

This last chapter poses the segmentation problem in an atlas-based deformable model framework. A review of implementation schemes from the original forward additive Lucas-Kanade to the most recent inverse compositional algorithm is outlined. The inverse compositional algorithm allows to precompute expensive steps in the Gauss-Newton iterative optimisation algorithm, and thus, it is much faster.

Hierarchical schemes are introduced to explain the separate formulation of a similarity transformation and a warp for the template deformation.

The 2D segmentation algorithm is extended with the explicit 2D+t cyclic shape model from Ch. 5. The 2D+t model imposes a smoothness constraint, and thus, segmentation of the whole volume is expected to be more robust than frame by frame. To test this hypothesis, segmentation results from three methods are compared: 1) individual frames with the inverse compositional algorithm; 2) whole volumes with the forward additive algorithm; and 3) whole volumes with the inverse compositional algorithm. The three approaches are evaluated in terms of convergence, accuracy and speed.

# CHAPTER 2

# Data

## 2.1 Echocardiography

Echocardiography is a non-invasive diagnostic tool widely used for the assessment of heart disease. It basically consists of sweeping a line, plane or, lately, volume, with a linear beam of ultrasonic waves, in the order of MHz, that intersects the heart in a region of interest.

When the beam crosses an interface between two media (tissue, vessels, blood plasma, red blood cells) with different acoustic impedance, part of the ultrasonic wave is transmitted, and part reflected. The reflection is specular (the angle of incidence equals the angle of reflection) for interfaces larger than the wave length, or scattered (in all directions) for smaller interfaces. *Backscatter* is the signal reflected back on the transducer. Backscatter from the specular echos is processed by the ultrasound machine to display the real-time *fundamental image* of the heart on a screen. A typical ultrasound system image formation diagram is displayed in Fig. 2.1.

Some advantages of echocardiography over other cardiac imaging procedures are its low cost (3 to 20 times cheaper than the others) [139], rapid evaluation, safety, that it is

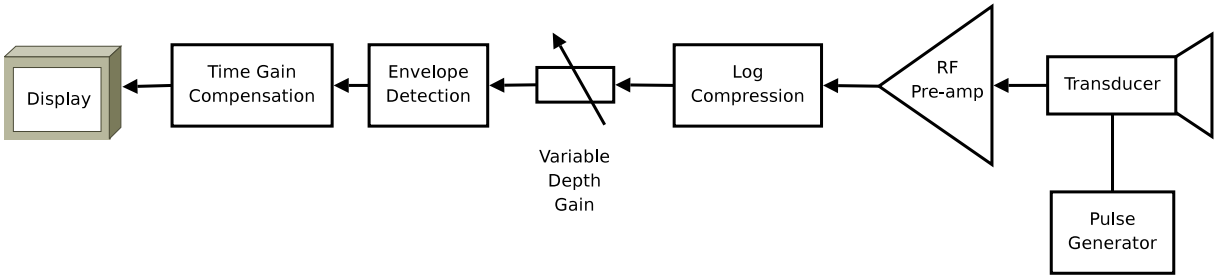



Figure 2.1: Typical ultrasound system diagram. (Adapted from Mulet Parada [125, Ch. 3].)

non-invasive, it does not require a special room with magnetic or radiation isolation, the echocardiography machine can be taken to the patient's bedside, or into the operation theatre, and its power consumption is like that of a computer system, while e.g. an MRI system needs around 100KW. The main problem with 2D echocardiography is poor image quality in patients with a bad *acoustic window*, but image quality can be improved with contrast agents, as discussed below. The second problem is misalignment. In apical planes, misalignment occurs when the interrogation plane does not lie on the long axis, and this causes apical foreshortening, which in turn introduces errors in global and local measures (e.g. [96, 134]). In SAX planes, misalignment occurs when the interrogation plane is not orthogonal to the long axis. Misalignment is no longer an issue with 3D echocardiography, because the whole volume is captured instead of just a plane, but 3D technology is out of the scope of this thesis.

## 2.2 Contrast echocardiography

In normal fundamental imaging interrogation, echos produced by red blood cells can be 1,000-10,000 times weaker than those produced by tissue. Thus, blood image areas are sensitive to interference from spurious echos, effectively rendering the endocardial border as a 'blurred haze' [12, sec 1.1.1]. To overcome this problem, contrast agents have been developed over the last 40 years [128].

Echocardiographic contrast agents are gas-filled microbubble suspensions that are in-

troduced into the vascular system. Microbubbles produce strong echos, and hence improve the signal to noise ratio in the image.

The clinical indication of first generation contrast agents (Albunex, Echovist, Levovist) was to linearly amplify the blood backscatter, but the development of the second generation of contrast agents (SonoVue, Optison, Luminity/Definity), with smaller and more persistent bubbles, enabled new imaging techniques such as harmonic imaging [128].

Harmonic imaging exploits non-linearities in the propagation media that produce the emission of echos at multiples of the fundamental frequency (*harmonics*), in addition to the fundamental frequency echos. Contrast harmonic imaging makes use of the strong non-linear response of microbubbles in an acoustic field with approx. peak pressure between 100 kPa and 1 MPa [12, sec 1.3.1]. This non-linear response is much weaker in tissue, so band-pass filtering the second harmonic produces an image where regions with contrast agent appear much brighter. Especially the blood pool is highly contrasted from the myocardium (*opacified*), and thus the endocardial border definition is significantly enhanced [40, sec. V.A] [128, 133, 186]. This is known as Left Ventricle Opacification (LVO), and is the principal indication second generation contrast agents have been approved for, in Europe, the USA and other countries [20, 128] [12, sec. 1.2.2]. The safety and tolerance of contrast agents are extremely favourable [75, 178]. Currently, contrast agents are recommended by the American Society of Echocardiography (ASE) for studies where the endocardial definition is suboptimal —when at least 2 segments[1] of the LV cannot be visualised [186]. This is typically the case with obese patients and those with lung disease, who comprise 10% to 20% of all echocardiography examinations [128]. But contrast agents can be useful to improve border definition and image alignment (it helps to avoid off-axis scanning) even in studies with good image quality [133].

LVO can be further improved using new low power techniques such as Power Modulation, that prevent bubble destruction and artifacts produced by second harmonic imaging [133, 177]. Power Modulation consists in sending a sequence of pairs of pulses, the first

[1]Left Ventricle segment models will be presented in section 3.3.1.

with a small amplitude $A$, and the second with an amplitude $A/2$. Both pulses have the same shape when leaving the transducer, and are reflected by tissue and microbubbles. At the receptor, the second reflection is multiplied by 2, and substracted from the first reflection. As tissue is assumed to behave linearly for a small $A$, the substraction is almost zero. Microbubbles, on the other hand, behave non-linearly, change the pulse's shape and thus, the substraction returns a harmonic signal (see e.g. [192, Ch. 2]).

Power Modulation, in effect, removes tissue backscatter and so enhances the endocardial border definition by increasing the contrast between the blood pool and the muscle. In addition, it enables Myocardial Contrast Echocardiography (MCE) for perfusion evaluation, because the small pulse amplitude minimises microbubble destruction. A downside to Power Modulation is that the frame rate is lower (16-25 Hz) than for second harmonic imaging, but it is still fast enough for functional evaluation. In a very recent review, myocardial perfusion has been found to have the same diagnostic accuracy as radionuclide imaging for the detection of CAD, while being cheaper, more portable and safer [20]. Therefore, it can be expected that perfusion evaluation will receive increasing interest in the near future. Our contrast DSE database studies were obtained using Power Modulation, but perfusion evaluation is beyond the scope of this thesis.

Instead, the reason for using Power Modulation in this thesis was the clinical hypothesis that myocardial perfusion would help segmentation by highlighting the muscle area thanks to intramural and epicardial vessels. The merit of this hypothesis will be criticised in section 3.4.2.

## 2.3 Contrast DSE database

A contrast Dobutamine Stress Echocardiography (DSE) database was built for this thesis. DSE is a clinical test where the drug dobutamine is slowly injected into the patient. Dobutamine *stresses* the heart, making it beat faster and allowing the clinicial to assess Ischaemic Heart Disease (IHD) in a controlled environment.

Cine loops were recorded at baseline (rest or no stress), intermediate and peak stress levels. Only baseline data was used in my experiments, as stress analysis is beyond the scope of this thesis. The reason to obtain the baseline data from DSE studies is that as stress echocardiography always requires optimal image quality and image quality decreases with stress [133], contrast is used more frequently. For example, contrast is used in 60% of all stress examinations in the John Radcliffe Hospital, Oxford [13].

The database contains 2D Power Modulation contrast echocardiography loops from 21 patients, acquired at the John Radcliffe Hospital by Prof. Harald Becher and Dr. Jon Timperley (both Level 3 echocardiographers[2]). The reason for the limited size of the database has more to do with the time and resources constraints imposed by hand segmentation of the myocardium than with the capacity of the hospital to acquire cine loops. This will be explained in more detail in section 2.4.

Each cine loop is gated at end diastole and contains 1 cardiac cycle. Images were acquired with a Philips Sonos 5500 (also known as HP or Agilent Sonos 5500) echocardiography system, and saved to computer files in the proprietary format Agilent DSR TIFF. Each file contains a stack of JPEG-compressed YBR-coded greyscale 2D images with 256 levels of grey in the ultrasound window, but for the electrocardiogram (ECG, see Fig. 2.3b), that is green.

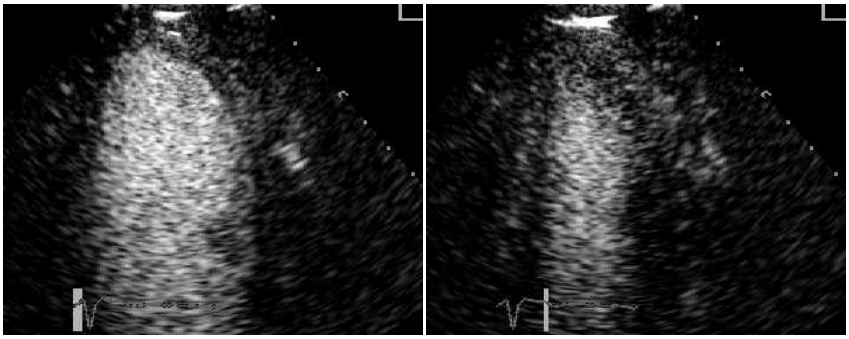
Each patient was imaged in the 4 principal planes, as illustrated by Fig. 3.2a: 2-chamber (2C), 3-chamber (3C), 4-chamber (4C) and short axis (SAX) at the papillary muscles level. The standard protocol used at the John Radcliffe Hospital was followed [178]. Examples of the frames in the database are shown in Fig. 2.2. To help with the interpretation of the images, end diastole frames from each plane are labelled in Fig. 2.3. The contrast and stress agents used were Bracco SonoVue (0.8 ml/min) and dobutamine, respectively. The contrast and stress agents were injected through a intravenous (IV) cannula connected to two pumps —one for each agent— through a three way tap. The contrast pump is specific for SonoVue; it provides constant agitation and can be prepared

[2]The competence description of Level 3 echocardiographers is presented in section 4.2.

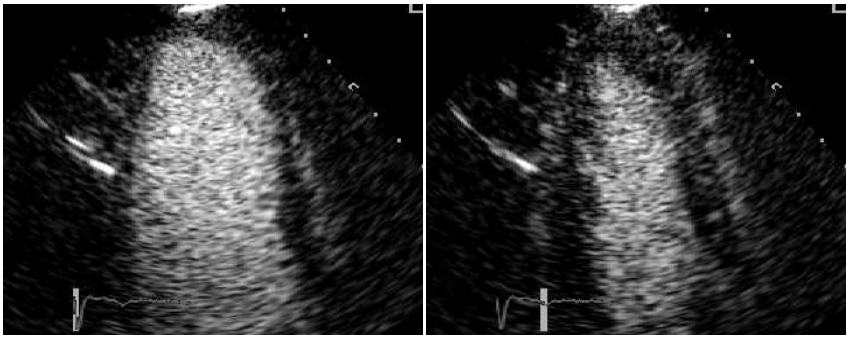
2 min prior to the study, and started by the sonographer with a remote control.

The clinical history of all patients was tabulated, and visual local function scoring for each segment was performed by Prof. Becher. Local function and visual scoring will be explored in detail in the next chapter in section 3.3. Myocardial contours were hand traced for all frames, as explained below in section 2.4.

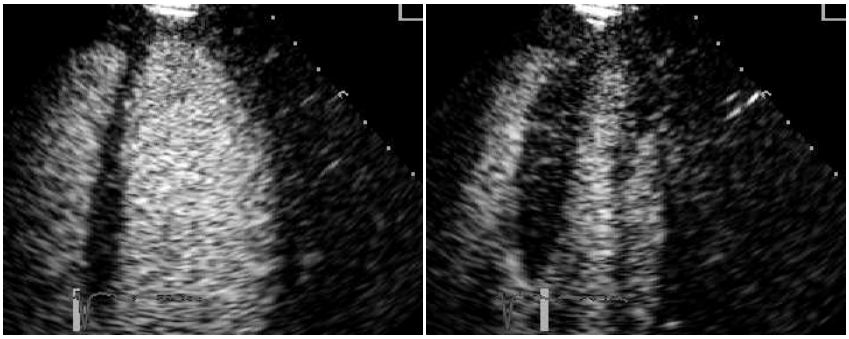
A Java wrapper was coded around proprietary legacy C/C++ code owned by Mirada Solutions Ltd. (now Siemens Molecular Imaging) to read the DSR TIFF image data, anonymise it and save it to standard-compliant DICOM files using the US Multi-Frame Image Module [130, A.7.4], so that it can be easily loaded from Matlab or other software applications. In Matlab, all frames were converted to greyscale. The last frame of each cardiac cycle only contains half the number of rows of previous frames. Those missing rows were interpolated so that all frames have the same size.



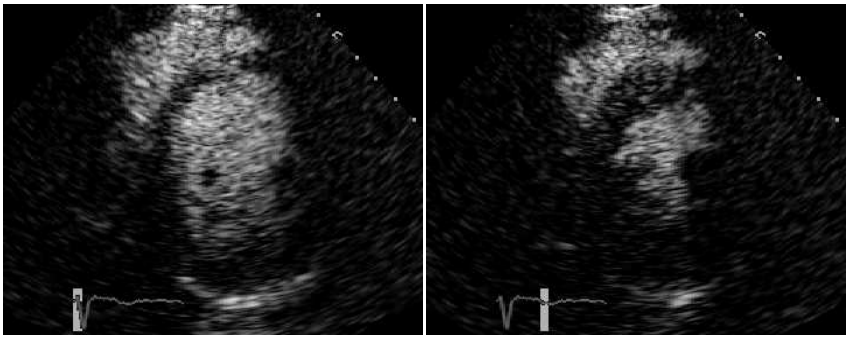
(a) 2C end diastole. (b) 2C end systole.



(c) 3C end diastole. (d) 3C end systole.

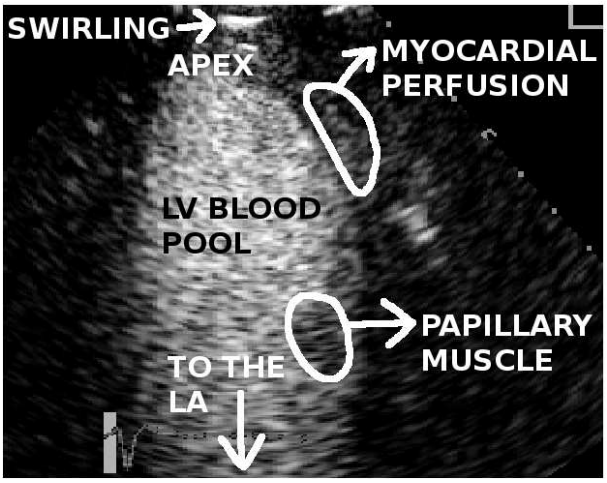


(e) 4C end diastole. (f) 4C end systole.

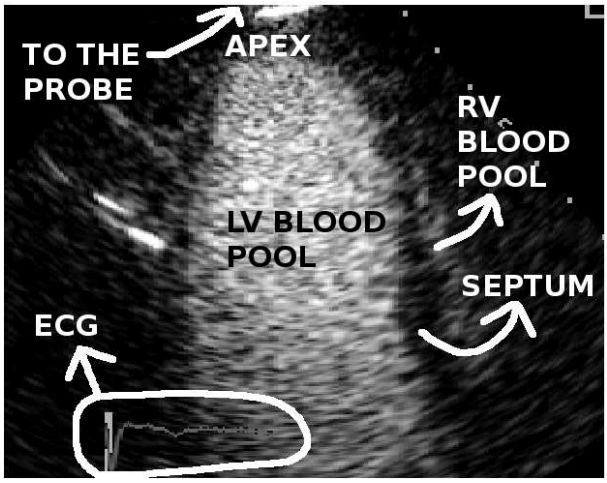


(g) SAX end diastole. (h) SAX end systole.

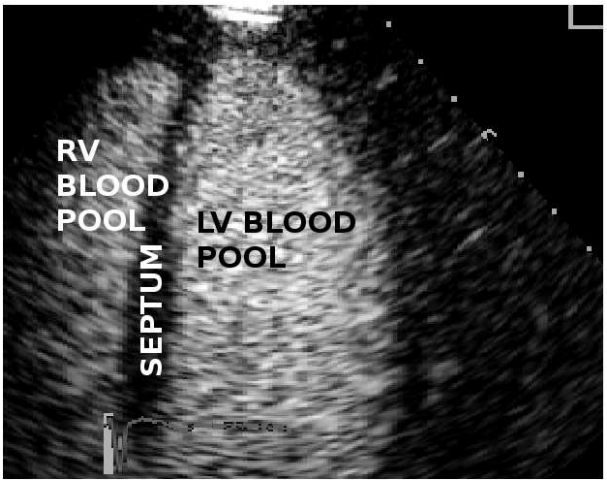
Figure 2.2: Examples of Power Modulation contrast frames, in end diastole and end systole, for the four principal planes: 2C, 3C, 4C and SAX.



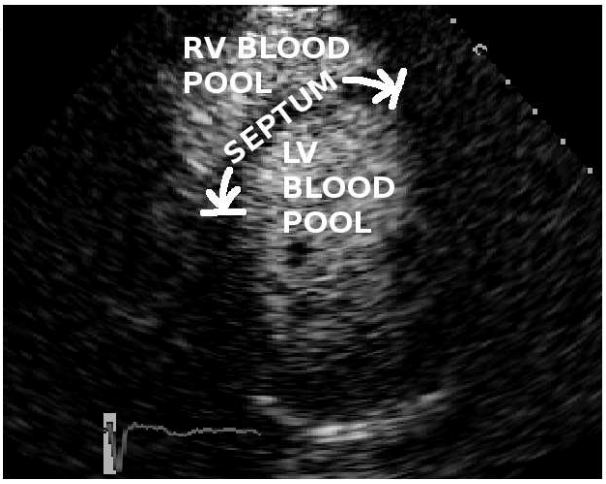

(a) 2C end diastole.




(b) 3C end diastole.




(c) 4C end diastole.




(d) SAX end diastole.

Figure 2.3: Commented Power Modulation contrast frames, in end diastole, for the four principal planes: 2C, 3C, 4C and SAX. LV: Left Ventricle. RV: Right Ventricle. LA: Left Atrium. ECG: Electrocardiogram.

## 2.4 Contours and landmark configurations

Endocardial and epicardial contours were initialised by Dr. Timperley using a custom built software application with a Graphical User Interface (GUI), called Quamus (Mirada Solutions Ltd.). The GUI displays a frame of an echocardiography loop and allows the expert to draw continuous contours adding control points with the mouse, and then save them to a text file. These contours will be called Quamus contours. In brief, they are closed quadratic approximating splines; control points do not lie on the contour and there is no correspondence between them from contour to contour (see Fig. 2.4a for an example, and Appendix A for details).

Quamus contours were corrected by Prof. Becher, by dragging control points with the mouse on a computer. This took for 1+ hour almost everyday for 2 months. Hand segmentation is very slow for three reasons: 1) Human experts are extremely busy and have a great amount of responsibility at the hospital, so it is hard for them to allocate time for hand tracing; 2) image quality is poor and makes the expert browse back and forth through the cine loop to extrapolate contour points; and 3) clicking and dragging points with the mouse on a computer screen, frame by frame, is a painstaking process. Due to the slowness of the process, it was not feasible to include more studies in the database.

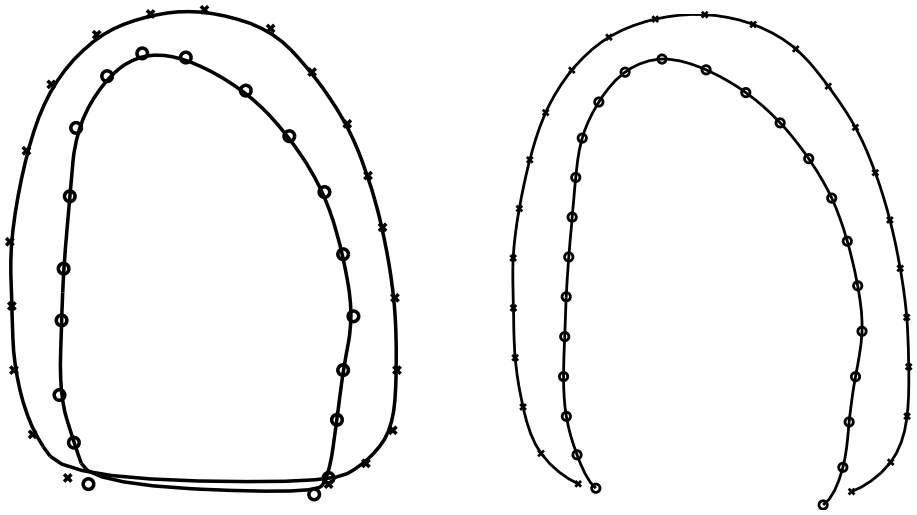
Quamus contours were sampled with a constant number of points, and a natural interpolating cubic spline[3] was fitted to them. The reason to use interpolating, rather than approximating, splines is that for interpolating splines, control points belong to the contour, while for approximating splines, control points are external to it. I will propose in section 4.4 that a convenient representation of the geometry for cardiac contours is a configuration of landmarks and pseudo-landmarks that belong to the contour and define a sparse correspondence between curves. This configuration of landmarks and pseudo-landmarks can be easily obtained from the control points of interpolating splines, but not

[3]Natural cubic spline fitting is implemented in Matlab by function `cscvn()`, using centripetal parametrisation for the knot vector [102].

from the control points of approximating splines.

While it is true, in general, that interpolating splines could be susceptible to *ringing* (large oscillations of the interpolated contour around the true myocardial border), in practice this is not an issue for our data. The myocardial representation used in this thesis is very smooth, and distance between control points is large enough to preclude ringing.

In apical planes, the base segment was then removed, as it does not correspond to any anatomical feature. Sampling points were placed on anatomical landmarks and at



(a) Quamus approximating quadratic splines, 19 and 20 control points endocardium and epicardium, respectively.

(b) Natural interpolating cubic splines, 25 control points for both endocardium and epicardium.

Figure 2.4: Difference between approximating and interpolating splines. 2C endocardial ($\circ$) and epicardial ($\times$) contours and control points.

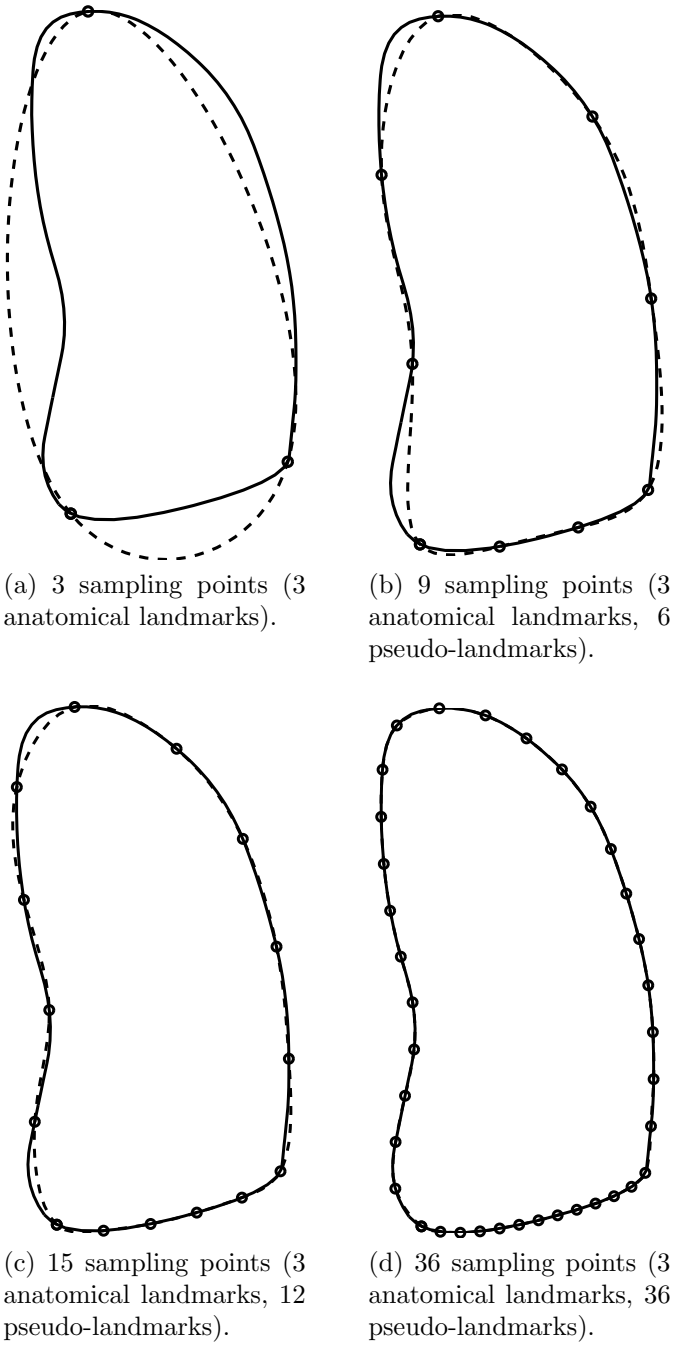
equidistant arc lengths between them: in apical planes, there is 1 anatomical landmark at the apex, and 2 at both sides of the mitral valve; in SAX planes, 1 at the infero-septal beginning of the RV. Thus, sampling points are a collection of landmarks and pseudo-landmarks. The difference between Quamus approximating splines and the resampled natural cubic interpolating splines is illustrated in Fig. 2.4. To assure that no relevant information was lost with the sampling, the number of control points was increased until the distance between the Quamus and the cubic contour was less than 1 pixel, as illustrated

by Fig. 2.5. It was found that apical and SAX contours can be conveniently represented by 25 and 15 pseudo-landmarks, respectively.

To sum up, cubic interpolating splines fulfil 3 conditions: 1) they approximate expert hand-traced contours to an arbitrarily small error with an increasing number of sampling points, that in practice has been found to be small; 2) the sampling points become control points of the spline, so landmarks and pseudo-landmarks become the minimum set of points that contains all the information about the spline; and 3) the sampling points establish a sparse correspondence between contours. Table 2.1 summarises the comparison between Quamus contours and cubic interpolating splines.

| Quamus quadratic approximating spline | Natural cubic interpolating spline |
|---|---|
| Hand traced by expert | Approximation to Quamus contour to arbitrarily small error |
| Uniform knot vector | Centripetal parametrisation knot vector |
| Control points $\notin$ contour | Control points $\in$ contour |
| 3 or 1 landmarks | 25 or 15 (pseudo) landmarks |
| Control points $\neq$ landmarks | Control points = (pseudo) landmarks |
| No sparse correspondence between landmarks | Sparse correspondence between (pseudo) landmarks |

Table 2.1: Comparison of Quamus contours to natural cubic interpolating splines.



(a) 3 sampling points (3 anatomical landmarks).

(b) 9 sampling points (3 anatomical landmarks, 6 pseudo-landmarks).

(c) 15 sampling points (3 anatomical landmarks, 12 pseudo-landmarks).

(d) 36 sampling points (3 anatomical landmarks, 36 pseudo-landmarks).

Figure 2.5: Approximating the natural interpolating cubic spline (dashed) to the Quamus contour (solid). The natural cubic spline is interpolated from different number of points sampled from the Quamus contour. Sampling points are placed on anatomical landmarks and at equidistant arc lengths between them: in apical planes, there is 1 anatomical landmark at the apex, and 2 at both sides of the mitral valve; in SAX planes, 1 at the infero-septal beginning of the RV.

# CHAPTER 3

# Clinical evaluation

## 3.1 Background

Current guidelines for the clinical application of echocardiography were published in 2003 by Cheitlin et al. [40], and state that the most common recommendation is evaluation of Left Ventricle (LV)[1] systolic function [40, sec. V]. The LV is important because oxygen-rich blood is pumped out from it at high pressure into the systemic circulation, LV diastolic disorders are an early indicator of coronary heart disease [173], and it is also most conveniently positioned near the ultrasound probe. This evaluation can be performed using linear measurements (M-mode), but the principal non-invasive method is 2D echocardiography, as it enables assessment of global and regional systolic function [40, sec. V], both visual and quantitative.

Standard recommendations by the American Society of Echocardiography (ASE) for quantification of LV systolic function in 2D echo have been available for almost 20 years [152, 153] and they are continually updated. These recommendations were updated and

[1] Right ventricle evaluation is more problematic due to 'its heavy trabecular pattern and to the difficulty in obtaining standardised imaging planes' [40, sec. V.B.1.g].

extended to other chambers by Lang et al. [96] in 2005. In the following sections, methods for the assessment of global and local function are presented, followed by special considerations about contrast echocardiography. Finally, the methods are discussed using experimental results from our database of 21 patients.

## 3.2 Global function

For global function, the recommended and best validated quantitative measure is Ejection Fraction (EF) [96], defined as

$$\mathrm{EF} = \frac{V_{\mathrm{ED}} - V_{\mathrm{ES}}}{V_{\mathrm{ED}}} \tag{3.1}$$

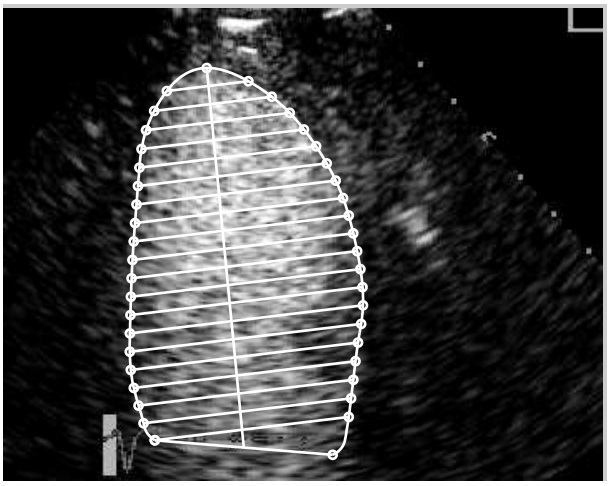
where $V_{\mathrm{ED}}$, $V_{\mathrm{ES}}$ are the end-diastolic volume and end-systolic volume, respectively. EF can be assessed by experts using visual estimation, but quantitative methods may be more reproducible [40, sec. V.A]. The most common way of quantifying LV volumes is applying the biplane method of disks (modified Simpson's rule) [96].

The biplane method of disks assumes that the endocardial boundaries for the 2C and 4C planes are known. The method comprises the following steps [152, 153]: First, the mitral annulus mid-point is computed. The apical long axis is defined as the segment that links the mitral annulus mid-point and the apex. Sections are computed perpendicular to the long axis, under the assumption that corresponding sections in 2C and 4C form an elliptical cylinder or disk. It is common to use $N = 20$ sections. The extra section at the base has triangular form. Note that some details on how to compute volumes were not specified by [152, 153]. Specifically, I approximated the volume of the extra section at the base with an equivalent rectangle with height the distance between the mitral annulus mid-point and the first disk. I also used the numerical integration midpoint rule to compute the diameters of each disk, $d_{2\mathrm{C}}$, $d_{4\mathrm{C}}$. For the extra section at the base, I used as diameter the length of the first disk boundary. The method is illustrated in Fig. 3.1. In order to include the base section, I had to tweak the definition from [152, 153] to express

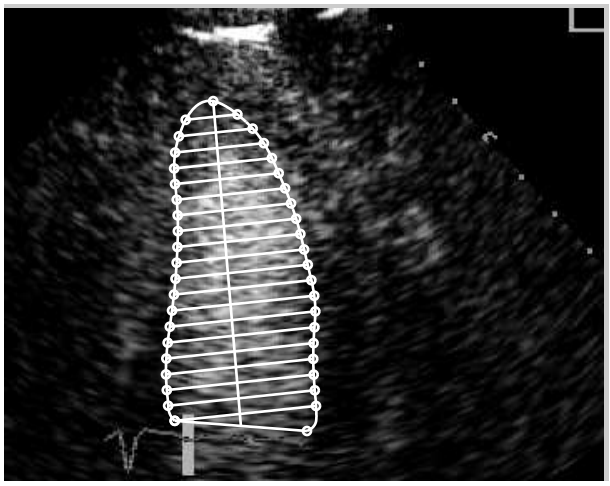
the total volume $V$ as

$$V = \frac{\pi L_{\text{mid}}}{4} d_{2\text{C},0} d_{4\text{C},0} + \frac{\pi (L_{\text{apical}} - L_{\text{mid}})}{4} \sum_{i=1}^{N} d_{2\text{C},i} d_{4\text{C},i} \tag{3.2}$$

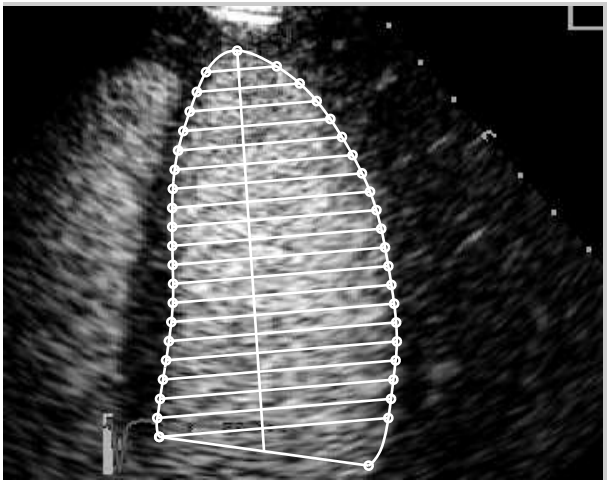
where $L_{\text{mid}}$ is the distance from the mitral annulus mid-point to the first section boundary, $L_{\text{apical}}$ is the length of the apical long axis, $d_{2\text{C},0} d_{4\text{C},0}$ are the equivalent diameters for the base section, and $d_{2\text{C},i} d_{4\text{C},i}$ are the diameters of the $i$-th disk, $i = 1, \ldots, N$.



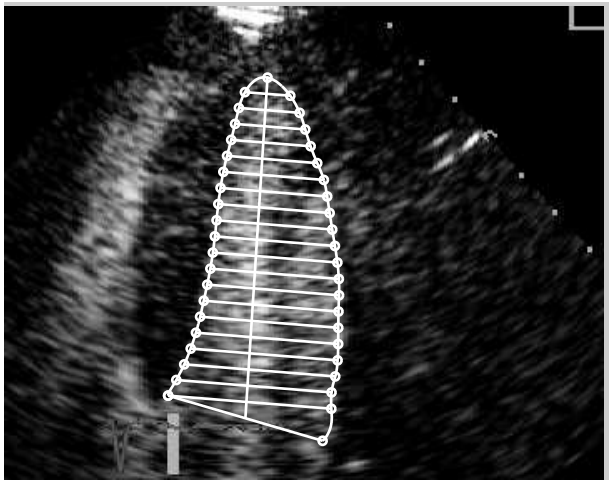
(a) 2C end diastole (ED).



(b) 2C end systole (ES).



(c) 4C end diastole (ED).



(d) 4C end systole (ES).

Figure 3.1: Biplane method of disks (modified Simpson's rule) for Patient 001.

## 3.3 Local function

Global function evaluation is limited, as usually functional abnormalities are local. The severity of local abnormalities can be underestimated in a global measure, or even completely masked by a functional increment in the rest of the myocardium. Thus, it is of great interest to develop an evaluation methodology for local function. Nonetheless, the 2003 Guidelines admitted that despite the availability of many useful methods, it is still controversial which one is optimal [40, sec. V.B].

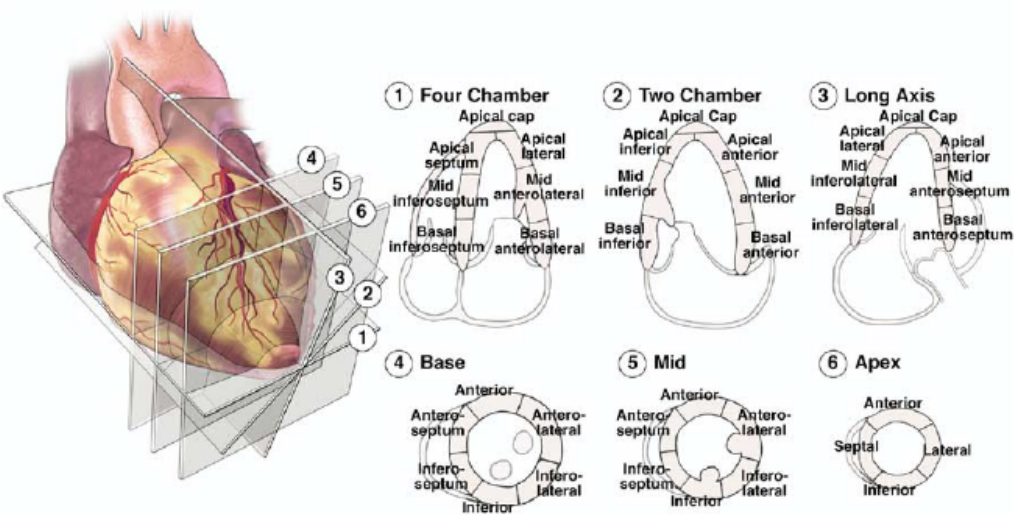
### 3.3.1 Segment model

To evaluate LV local function, the myocardium is partitioned into imaginary segments. There are many partitions used in different modalities, the more standard being the 16-segment model proposed in 1989 by the ASE [153], and updated by the American Heart Association in 2002 to the 17-segment model in Fig. 3.2a [37] by adding a segment for the apical cap. But in fact, the 16-segment model is more appropriate for functional assessment, as the apical cap does not move in the normal apex [96].

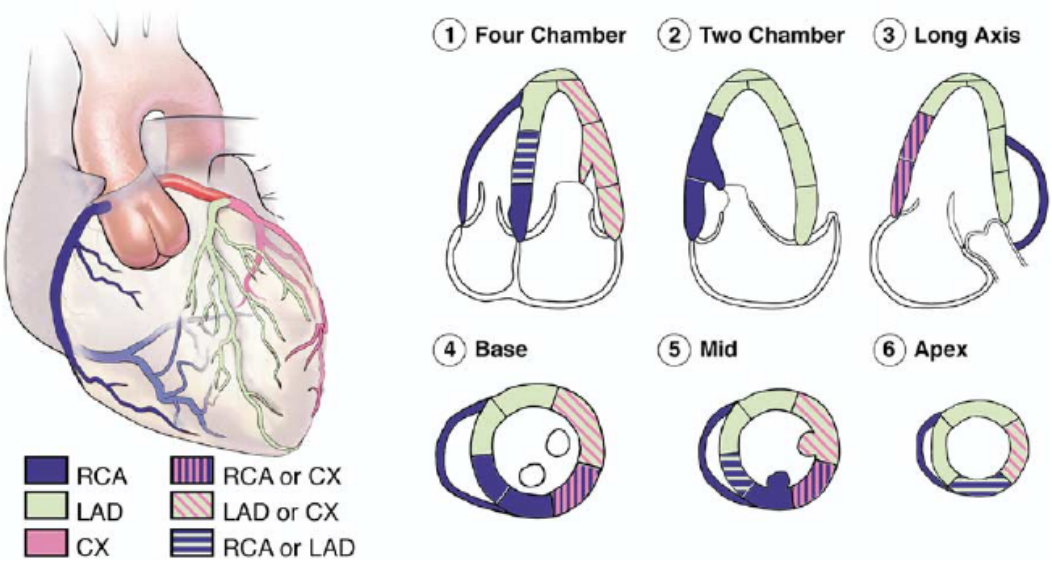
The reasons behind the 16-segment model were [153]: '1) Anatomic logic. 2) Easy identification of the segments using internal anatomical landmarks. 3) Relationship of the segments to known coronary arterial supply. 4) A uniform scoring system for grading the severity of segmental wall motion abnormalities'. Points 3) and 4) are discussed in sections 3.3.2 and 3.3.3, respectively.

### 3.3.2 Coronary artery supply

Despite being a reason for the 16-segment model, unfortunately the relationship between segments and coronary arteries is not straightforward. It is generally accepted that there is a connection between Coronary Artery Disease (CAD) and regional wall-motion abnormalities (RWMAs), but this connection has noteworthy limitations that are not always acknowledged in the literature. First, partial/total occlusion of a major coronary artery




(a) Segmental analysis of LV walls.




(b) distribution of coronary arteries.

Figure 3.2: 17-segment model of LV walls and distribution of coronary arteries: Right coronary artery (RCA), left anterior descending (LAD) and Circumflex (CX). (Reprinted from Lang et al. [96] with permission).

can result in ischemia/infarction and thus RWMA in the segments it feeds, but if the myocardium is fed by secondary arteries, the problem in the major artery may remain inconspicuous. Second, RWMA may arise from a variety of causes (e.g. left bundle branch block), and not only CAD (e.g. [160]). Finally, Lang et al. [96] proposed Fig. 3.2b as a good first approximation for the typical distribution of coronary arteries, but noted that arterial distribution varies between subjects. In fact, different publications do not agree in the coronary artery distribution for certain segments, as shown in Table 3.1.

| Segment | Plane | [96] | [37] | [151] | [16] |
|---|---|---|---|---|---|
| Basal inferoseptum | 4C | RCA | RCA | RCA | RCA |
| Mid inferoseptum | 4C | RCA/LAD | RCA | LAD | RCA |
| Apical septum | 4C | LAD | LAD | LAD | LAD |
| Apical lateral | 4C | LAD/CX | CX | LAD | LAD |
| | 3C | LAD | CX | N/A | LAD |
| Mid anterolateral | 4C | LAD/CX | CX | CX | CX |
| Basal anterolateral | 4C | LAD/CX | CX | CX | CX |
| Basal inferior | 2C | RCA | RCA | RCA | RCA/CX |
| Mid inferior | 2C | RCA | RCA | RCA | RCA/CX |
| Apical inferior | 2C | LAD | RCA | RCA | LAD |
| Apical anterior | 2C | LAD | LAD | LAD | LAD |
| | 3C | LAD | LAD | N/A | LAD |
| Mid anterior | 2C | LAD | LAD | LAD | LAD |
| Basal anterior | 2C | LAD | LAD | LAD | LAD |
| Basal inferolateral | 3C | RCA/CX | CX | CX | RCA/CX |
| Mid inferolateral | 3C | RCA/CX | CX | CX | RCA/CX |
| Mid anteroseptum | 3C | LAD | LAD | LAD | LAD |
| Basal anteroseptum | 3C | LAD | LAD | LAD | LAD |

Table 3.1: Assignment of segments to coronary arterial territories from different authors: Lang et al. [96] (the most recent and authoritative), Cerqueira et al. [37], Sawada et al. [151], Beleslin et al. [16]. RCA = Right Coronary Artery, LAD = Left Anterior Descending, CX = Left Circumflex Coronary Artery, N/A = Not available.

### 3.3.3 Scoring system

Local function evaluation in the clinical setting usually employs a qualitative scoring system. The scoring system assigns one of the following values to each segment, after visual inspection [96, 153]: 1 = normokinesis or hyperkinesis, 2 = hypokinesis, 3 = akinesis,

4 = dyskinesis (paradoxical systolic motion), 5 = aneurysmal (diastolic deformation). Dyskinetic segments are asynchronous with the rest, i.e. they increase the size of the cavity in systole, while the rest of segments decrease it. When building the contrast DSE database, scoring of dyskinetic segments was split into 2 sub-scores: A = asynchronous normokinesis, B = asynchronous hypokinesis.

While the scoring system provides a standardised framework for expressing a qualitative assessment of regional LV function, research has also focused on finding quantitative measures, mostly derived from endocardial motion, myocardial thickening and myocardial perfusion.

### 3.3.4 Endocardial motion

In 2D echocardiography, LV endocardial motion is easier to assess than thickening, because the blood/myocardial boundary is better imaged than the external boundary (formed by the epicardium and the Right Ventricle's endocardium). There are several reasons: the epicardial interface is less echogenic, the RV endocardium is heavily trabeculated, and sometimes part of the myocardium falls outside the ultrasound window or is occluded by a rib shadow[2].

Segmental endocardial motion can be quantified using line or area measurements. Using ventriculograms from 34 patients, Gelberg et al. [69] found that area measures are more sensitive and specific to detect abnormalities than radial (from the wall to the centroid) or chord (from wall to wall) line measures. Carstensen et al. [35] compared several Fractional Area Change (FAC) measures. Centroid methods partition the cavity area as illustrated by Fig. 3.3a. FACext is computed using floating external reference points to create the partition in Fig. 3.3b. Centroid methods can be fixed (FACfix, using the ED centroid for all frames) or floating (FACfloat, computing a different centroid for

[2]H. Becher and J. Timperley (John Radcliffe Hospital) noted that rib shadows were a major obstacle for epicardial segmentation when we were creating the contrast DSE database. This was later confirmed by M. Mulet-Parada in a personal communication, from his own experience while writing his DPhil Thesis [125] on LV segmentation and tracking.

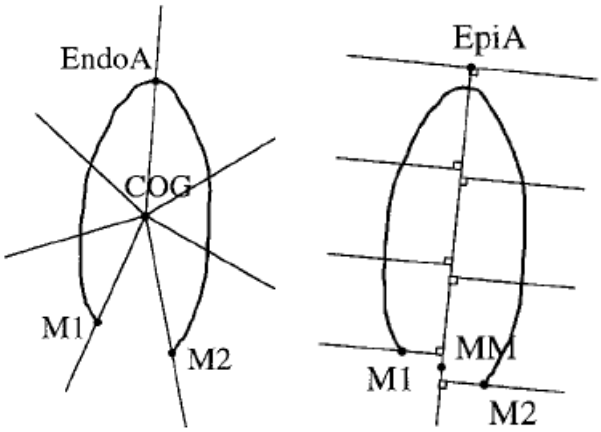
each frame). Carstensen et al. did not find significant differences between centroid and external reference methods. Floating reference systems were found to introduce error in the measure, but are better suited when there is substantial intrathoracic motion of the heart. Jacob et al. [83, 84] proposed measuring endocardial excursion as the distance from a landmark at ES and ED in a Principal Component Analysis (PCA) shape model[3], and computed the excursion for each segment as the average excursion of 4 landmarks. The measure was normalised by the largest excursion of all segments. Experiments were run on 4 patients and 1 patient, respectively, so the results are inconclusive. Caiani et al. [32] used a FACfloat method in SAX contours obtained from MRI images. Bermejo et al. [18] used contrast echocardiograms to validate semi-automatic segmentation of the endocardium. Endocardial motion was measured on 27 patients using radial shortening, i.e. change in the distance from myocardial points to their Nearest Neighbours (NNs) on the long axis (a similar idea to FACext, but without integrating the area and using internal reference points). Radial shortening values showed noticeable overlapping between normal and abnormal, although when thresholds to identify RWMA were adjusted to the typical endocardial excursion of each segment, discrimination accuracy was very high (area under the ROC = 0.87).

### 3.3.5 Myocardial thickening

Myocardial thickening is recommended by the ASE as a complement to improve functional evaluation [96]. Endocardial motion evaluation alone can underestimate RWMA, if ischemic or infarcted myocardium is dragged by healthy tissue; or it can overestimate it, if healthier adjacent regions are affected by 'tethering, disturbance of regional loading conditions, and stunning' [96].

Evaluating myocardial thickness in 2D echocardiography with linear measures can be seen as a sparse or dense correspondence problem between the endocardial and external boundaries. Different approaches to measuring wall thickness are reviewed and discussed

[3]Principal Component Analysis (PCA) shape models will be explained in detail in Ch. 4.




(a) Using a centroid. (b) Using external reference points.

Figure 3.3: Fractional Area Change (FAC) measure methods. (Reprinted from Carstensen et al. [35] with permission).

in the rest of this section, taking into account that there are no standard clinical recommendations.

Radial shortening (e.g. as used for wall motion by [18]) could be extended to wall thickening, but it has been shown that wall deformation follows directions towards different centres, and that assuming a single centroid introduces evaluation errors [185]. Dumesnil et al. [59] traced endocardial and epicardial contours on ventriculograms from 32 patients. Noting that the shortest distance between both contours is not necessarily the true wall thickness, they preferred the term 'wall dynamics' rather than 'wall thickness', but the latter is commonly used in the literature. The authors computed thickness as the distance between endocardial points and their Nearest Neighbours (NNs) on the epicardium, and averaged values within the same segment. Similarly to EF, they defined Wall Thickening (WT) as

$$\mathrm{WT} = \frac{T_{\mathrm{ES}} - T_{\mathrm{ED}}}{T_{\mathrm{ED}}} \tag{3.3}$$

where $T$ is thickness, and found that even though mean values for normal and abnormal segments were statistically significantly different ($p < 0.2$), the overlap between both distribution was large. NNs provide a reasonable approximation to wall thickness when both

myocardial boundaries are parallel to the long axis, but not in basal and apical segments, as illustrated by Fig. 3.4a. Mizushige et al. [123] used M-mode echocardiography orthogonal to the wall to measure WT on 28 patients, and found a significant correlation with coronary stenosis. Jacob et al. [83, 84] extended their wall motion measure (see above) to compute WT, and applied it to a few case studies. This assumes a correspondence between the endocardium and external boundary through the shape space that has no known anatomical interpretation. A popular method in medical literature is the *centreline method* proposed by Sheehan et al. [162], illustrated in Fig. 3.4c. Von Land et al. [185] noted that the centreline method suffered from chords crossing each other, and that in order to compute the centreline, a correspondence must already be known. They proposed an iterative algorithm that at each step uses the current correspondence to recompute the centreline, and vice versa; the algorithm also avoids crossings and smooths sudden curve changes. A similar solution, the iterative average curve method, was later and independently proposed by Chalana and Kim [38], but it does not prevent chord crossings. The evolution of the algorithm is illustrated by Fig. 3.4c and 3.4d. Papademetris et al. [135, 136], independently proposed the *Symmetric Nearest Neighbour (SNN) correspondence algorithm*. This algorithm avoids the need for iterations. The SNN idea is, in fact, similar to that used by von Land et al. [185] to find reliable correspondences. Examples for the SNN correspondence algorithm are shown in Fig. 3.4 for apical and SAX planes.

### 3.3.6 Myocardial perfusion

Finally, myocardial perfusion is defined as tissue blood flow at the capillary level (e.g. [20]). Perfusion is a relatively new measure of regional function, thanks to the introduction of new generation contrast agents and imaging techniques. Perfusion evaluation was mentioned briefly in section 2.2, but it is beyond the scope of this thesis.

## 3.4 Experimental results

### 3.4.1 Global function experiments

Baseline studies from the Power Modulation contrast DSE database were used to evaluate global function in 21 patients (10 normal, 11 abnormal). Endocardial contours from 2C and 4C planes were used. The frames with the largest and smallest areas were labelled as ED and ES, respectively. LV volumes (ED and ES) were computed applying the biplane method of disks to the 2C and 4C planes contours, as described in section 3.2. EF was computed using Eq. (3.1).

Fig. 3.5a shows box-and-whisker plots for EF stratified as normal and abnormal patients. Overlapping notches indicate lack of evidence to reject the null hypothesis that the medians of the two groups are equal at the 5% significance level. That is, there is no evidence that median EF is statistically different in normal and abnormal studies and thus, diagnosis would not be possible with this measure and criterion. But lack of statistical significance does not mean lack of clinical interpretation. The failure of the null hypothesis test can be attributed to the large EF variance in the abnormal group. To overcome this problem, another experiment was run where EF was corrected for mean scoring (computed as the average scoring for all 12 segments in 2C and 4C). The results suggest a linear decrease of EF with increasing scoring values, as shown in Fig. 3.5b. Comparing the corrected EF to the severity of abnormality intervals recommended by the ASE [96, Table 6] suggest that this experiment's results slightly overestimate EF. The corrected EF graph also allows easier identification of 6 outliers. Information about the outliers has been summarised in Fig. 3.6. Each case is discussed in the rest of this section, as it can help to understand the limitations of global function evaluation from contrast echocardiography.

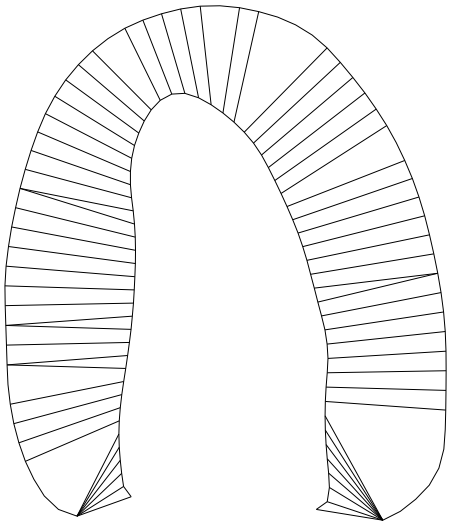
**Patient 006:** ($\nabla$, Fig. 3.6a) This patient is severely abnormal. Scoring reflects an akinetic apex, dyskinesis in the inferior wall and hypokinesis in the lateral and anterolateral walls. The EF=44% looks abnormal enough, but when corrected with the mean

scoring, it is overestimated by approx. 10%-15%. The myocardial contours indicate that the reason is a large displacement in the anterolateral endocardial wall in 4C (expected to be hypokinetic to akinetic). Given that the apical endocardium's lack of movement in 2C agrees with the scoring, and that the patient has a history of myocardial infarction, the large displacement in 4C could be attributed to failed hand tracing or to a shift in the interrogation plane halfway through the cycle. The latter argument is supported by jump in epicardial motion visible in the cine loop.

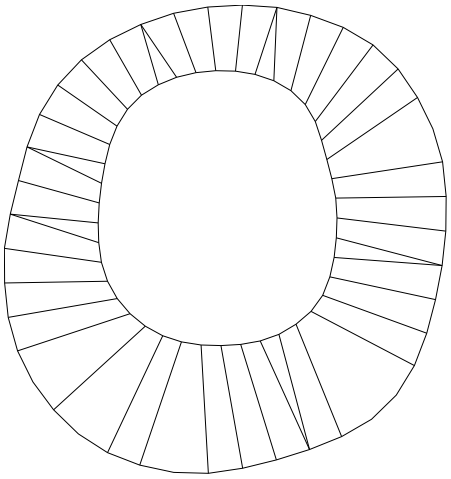
**Patient 012:** (◯, Fig. 3.6b) This patient is hypokinetic in 3 segments, but his EF is high (EF=69%). One of the hypokinetic segments is basal inferoseptal in 4C. The normal excursion in this segment is very small, so even if it is hypokinetic, its impact in the EF is negligible. The other 2 segments are part of the anterior wall in 2C. The apical lateral segment has reduced excursion, but the mid lateral segment seems to be dragged by the healthy basal lateral segment.

**Patients 022, 038 and 039:** (□, ♢ and +; Figs. 3.6c, 3.6d and 3.6e). These patients have large EFs, between 70% and 90%. Inspection of their corresponding contours suggest that the interrogation plane is off-axis, mostly in the 2C plane. This causes an apparent large movement of the external wall, apical foreshortening and overestimation of the EF.

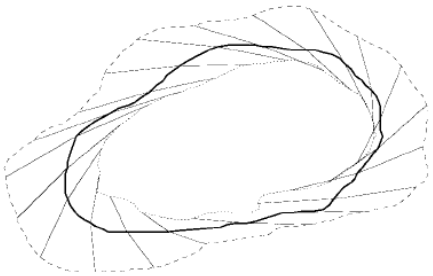
**Patient 040:** (△, Fig. 3.6f) Visual inspection of the 2C images suggest insufficient LV opacification, that is worse towards the base. Hand tracings of the endocardium and inferior wall at the base actually intersect (this is anatomically unfeasible). Thus, endocardial tracings in 2C cannot be trusted in this case. This patient is particularly outlying because both her EF and mean scoring are simultaneously too large. The large mean scoring can be explained by the two dyskinetic septal segments. They contribute as hypokinetic segments in terms of EF, but with values of 4 to the mean scoring value.



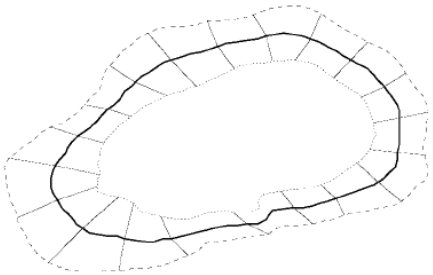
(a) External NNs to endocardial points (2C).



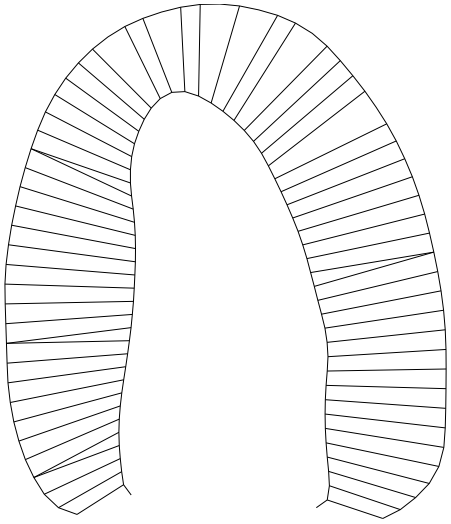
(b) External NNs to endocardial points (SAX).



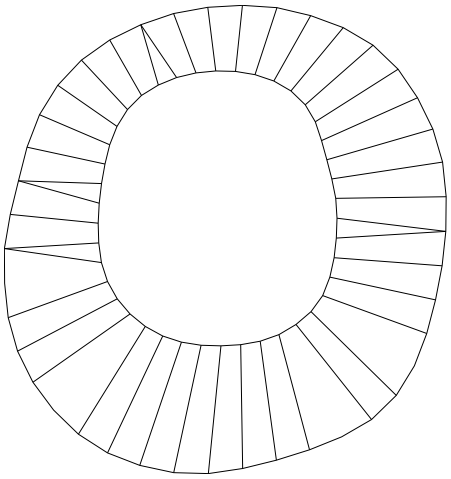
(c) AC (initialisation).



(d) AC (after 5 iterations).

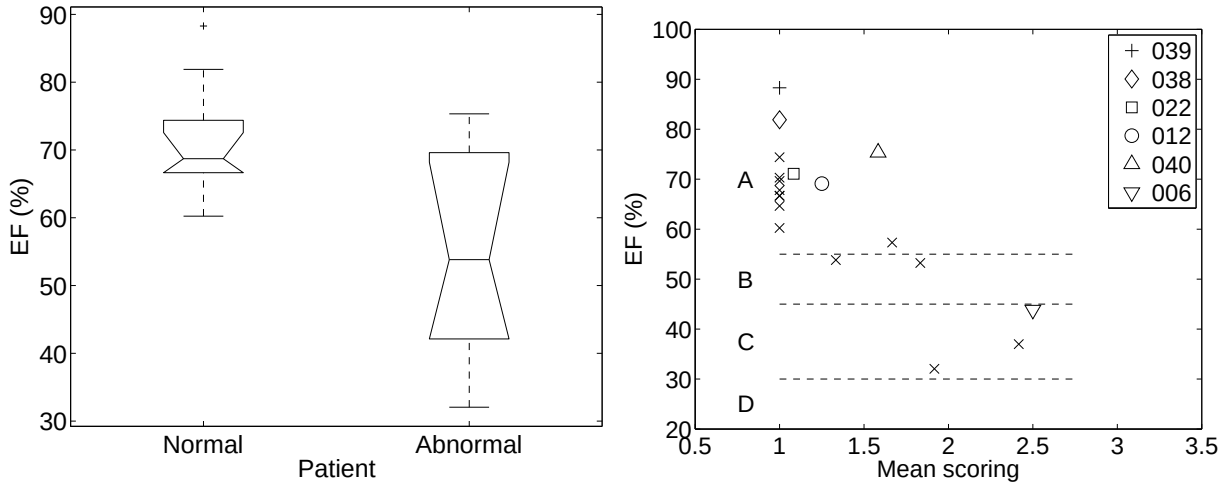


(e) SNNs (2C).



(f) SNNs (SAX).

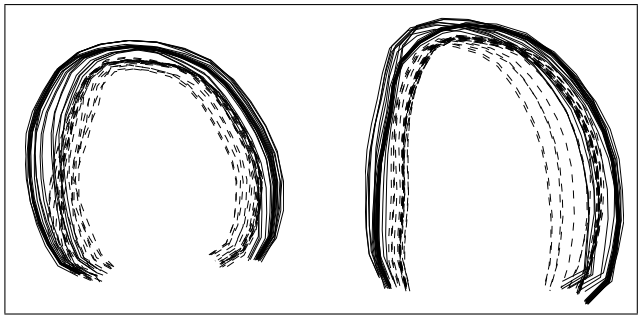
Figure 3.4: Line measures of wall thickness. NN: Nearest Neighbour, computed as closest external boundary point to each endocardial point. SNN: Symmetric Nearest Neighbour correspondence by Papademetris et al. [135, 136]. AC: Average curve method by Chalana and Kim [38], an improvement on the centreline method by Sheehan et al. [162], but without the chord crossing correction of von Land et al. [185]. (c) and (d): Reprinted from Chalana and Kim [38] ©1997 IEEE, with permission.



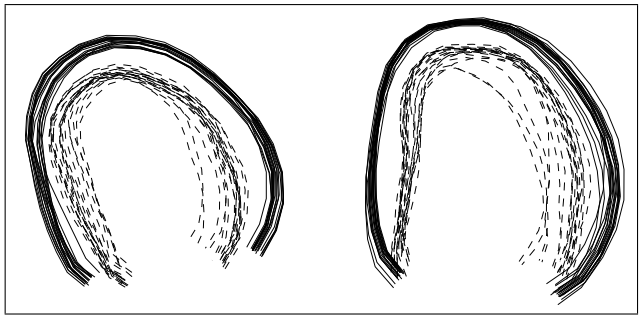

(a) Normal vs. abnormal patients.

(b) EF vs. mean scoring.

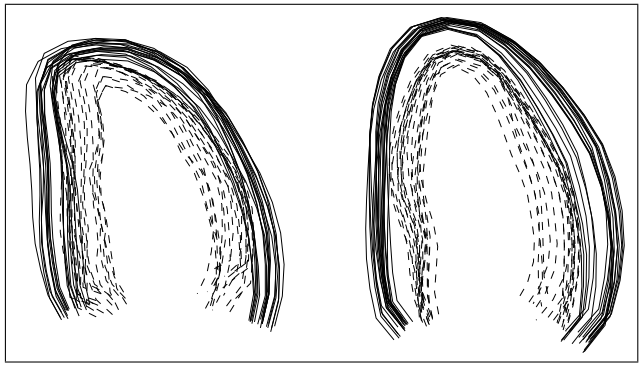
Figure 3.5: Global function assessment with Ejection Fraction (EF) and mean scoring. (a) Stratification as abnormal/normal shows significant overlap and makes it harder to spot outliers. Overlapping notches indicate lack of evidence to reject the null hypothesis that the medians of the two groups are equal at the 5% significance level. (b) Plotting EF vs. mean scoring in 2C and 4C shows EF within a reasonable interval for normals and a roughly linear relationship for abnormal. Inlying data points are marked with symbol '×'. Outlying data points are marked with the symbols in the legend. Mean scoring was computed as the average scoring for all 12 segments in 2C and 4C, using the following scale: 1 = normokinesis or hyperkinesis, 2 = hypokinesis, 3 = akinesis, 4 = asynchronous normokinesis, 5 = asynchronous hypokinesis. Horizontal dashed lines show the severity of abnormality intervals recommended by the ASE [96, Table 6]: Reference range (A), Mildly abnormal (B), Moderately abnormal (C), Severely abnormal (D). Some outlying patients are discussed in the main text: 039 (+), 038 (♢), 022 (□), 012 (◯), 040 (△), 006 (∇).



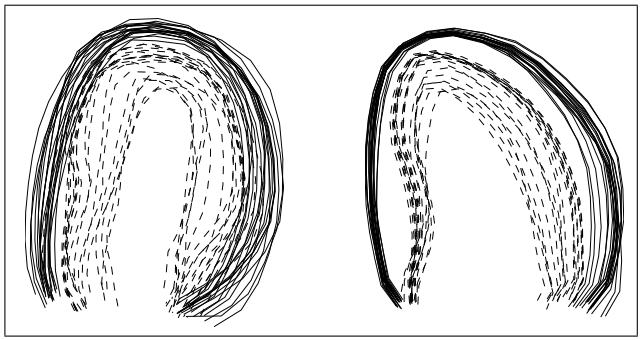
(a) Patient 006 (∇), EF=44%, SC=2.5.
$SC_{2C}$ = 2, 2, 3, 3, 1, 1. $SC_{4C}$ = B, B, 3, 3, 2, 2.



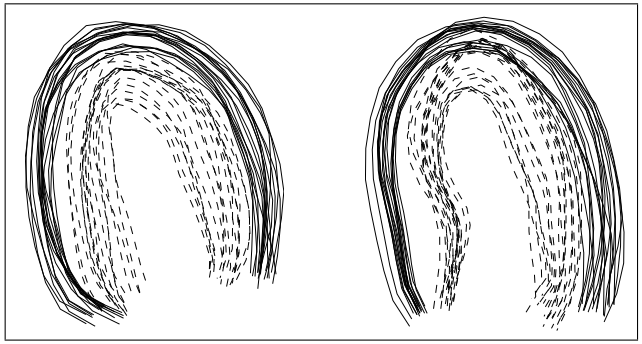
(b) Patient 012 (◯), EF=69%, SC=1.2.
$SC_{2C}$ = 1, 1, 1, 2, 2, 1. $SC_{4C}$ = 2, 1, 1, 1, 1, 1.



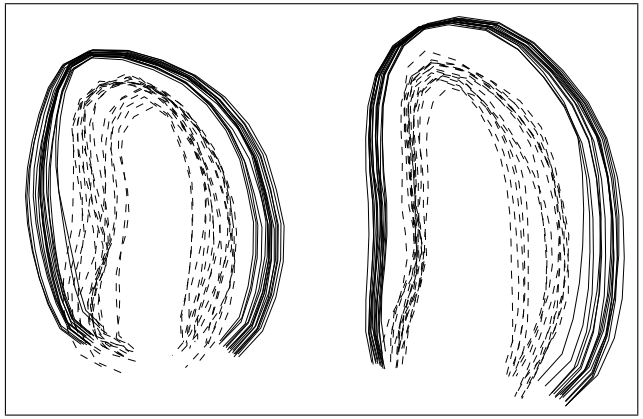
(c) Patient 022 (□), EF=71%, SC=1.1.
$SC_{2C}$ = 1, 1, 2, 1, 1, 1. $SC_{4C}$ = 1, 1, 1, 1, 1, 1.



(d) Patient 038 (♢), EF=82%, SC=1.0.
$SC_{2C}$ = 1, 1, 1, 1, 1, 1. $SC_{4C}$ = 1, 1, 1, 1, 1, 1.



(e) Patient 039 (+), EF=88%, SC=1.0.
$SC_{2C}$ = 1, 1, 1, 1, 1, 1. $SC_{4C}$ = 1, 1, 1, 1, 1, 1.



(f) Patient 040 (△), EF=75%, SC=1.6.
$SC_{2C}$ = 1, 1, 1, 2, 1, 1. $SC_{4C}$ = 1, A, A, 1, 1, 1.

Figure 3.6: Suspected Ejection Fraction outliers. Expert hand tracing of LV myocardial contours. Solid: external boundary. Dashed: endocardium. Each box displays the 2C (left) and 4C (right) views. $SC_{2C}$, $SC_{4C}$: Scoring for each of 6 segments in 2C and 4C (clockwise). SC: Mean scoring value for both planes. Scores: 1 = normokinesis or hyperkinesis, 2 = hypokinesis, 3 = akinesis, A = asynchronous normokinesis, B = asynchronous hypokinesis.

### 3.4.2 Local function experiments

Evaluation of local function from the contrast DSE database is more limited than global function. Because of functional heterogeneity and the lack of a general contractility model, each segment needs to be treated separately. The lack of a general contractility model will be addressed in Ch. 5 —with some preliminary visual results about functional heterogeneity presented in Fig. 5.5. While incidence of global abnormality is high in the database (11 of 21 patients), incidence of local abnormalities for a given segment is much lower, as abnormal patients usually have several normokinetic segments too. In addition, 2 of the abnormal patients have dyskinetic segments. Thus, there is not enough data to perform a similar analysis to the previous section.

This notwithstanding, local functional analysis in the database is still interesting and illustrates the methods discussed in section 3.3. In particular, the following local function measures were implemented and tested: Fractional Area Change for endocardial wall motion, with fixed and floating centroids (FACfix and FACfloat, respectively), as explained in section 3.3.4; and Wall Thickening (WT) for myocardial thickening, using the SNN correspondence method to measure distances, as explained in sec. 3.3.5.

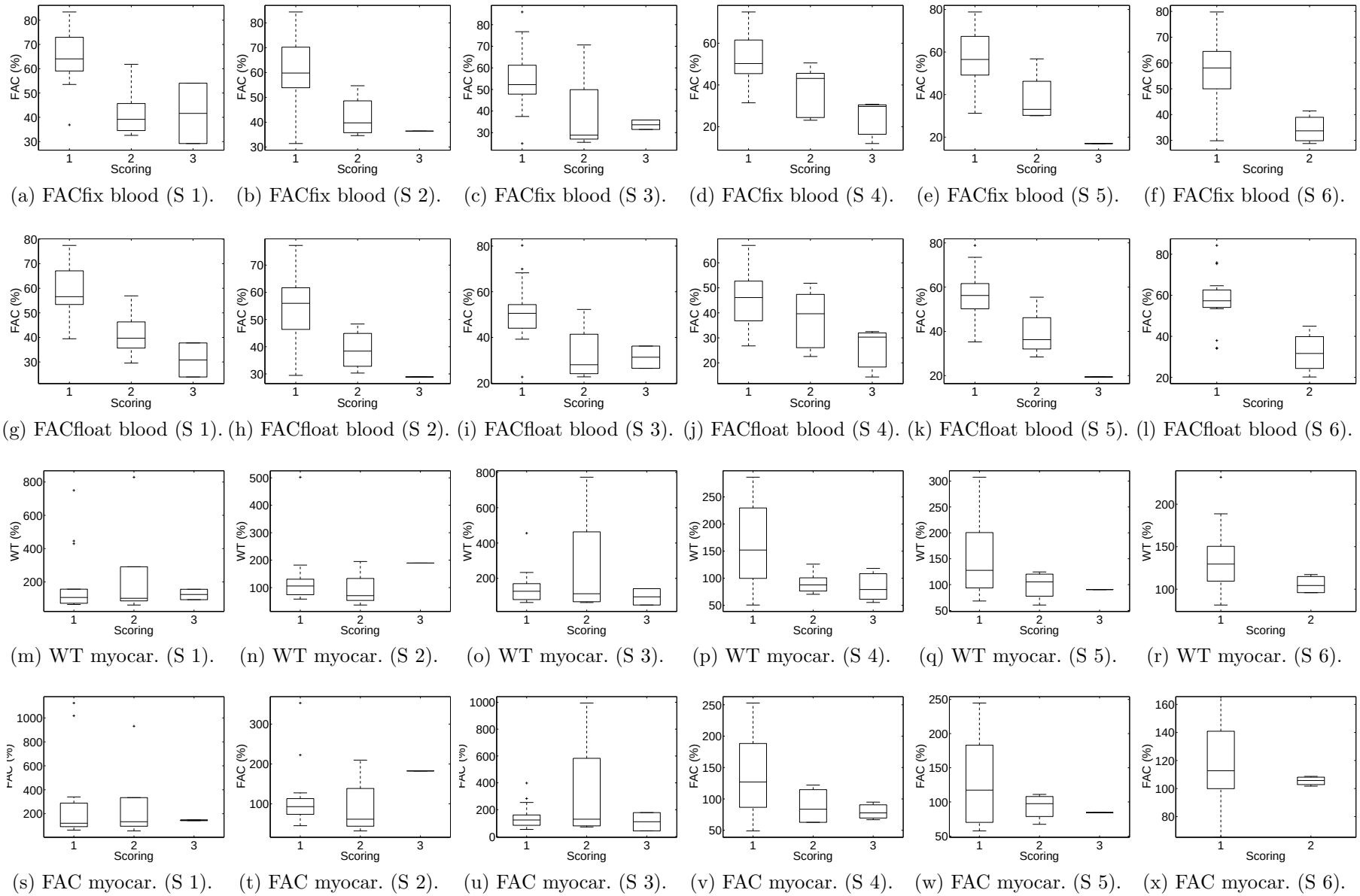
The results are presented with box-and-whisker plots in Figs. 3.7 to 3.10. Each page corresponds to one plane (2C, 3C, 4C, SAX), each row to an evaluation method, and each column to a segment. The 16-model explained in sec. 3.3.1 is used. Segments are numbered from 1 to 6 clockwise; in SAX planes, the first segment is the inferoseptal.

In broad terms, endocardial motion measures (FACfix and FACfloat) are similar to each other. Likewise, myocardial thickening measures (WT and FAC myocardium) are similar to each other too.

Endocardial motion measures appear to discriminate better between normal and abnormal segments than myocardial thickening. The dataset is too small to show any statistically significant differences, but this observation is not merely speculative, as it is consistent throughout the 4 planes. This result is not expected after the discussion in

section 3.3.5. This could be explained because when FACfix and FACfloat are normalised by the area value at ED, this area is maximum (expanded cavity). On the contrary, when WT and FAC myocardium are normalised by the distance or area value at ED, those magnitudes are minimal (myocardium is thinnest). Hence, myocardial thickening measures are in principle less robust against noise and errors.

But a more interesting explanation can be developed from the imaging technique limitations. As discussed in sec. 2.2, Power Modulation was chosen to further improve endocardial delineation through better LV opacification, and under the hypothesis that perfusion would highlight the myocardium. However, visual inspection of the studies shows that perfusion backscatter is not strong enough, and that the external boundary is virtually invisible in large areas of the image. Manual tracing of the external wall required a lot of interpolation from the experts, for instance using part of the apex and the base to extrapolate the middle section of the external wall, or using signals from epicardial vessels to approximate its location. It is possible that errors in the hand tracing of the external wall are large enough to preclude Power Modulation contrast echocardiography as an appropriate technique to measure myocardial thickening. Further research is needed to substantiate this claim.



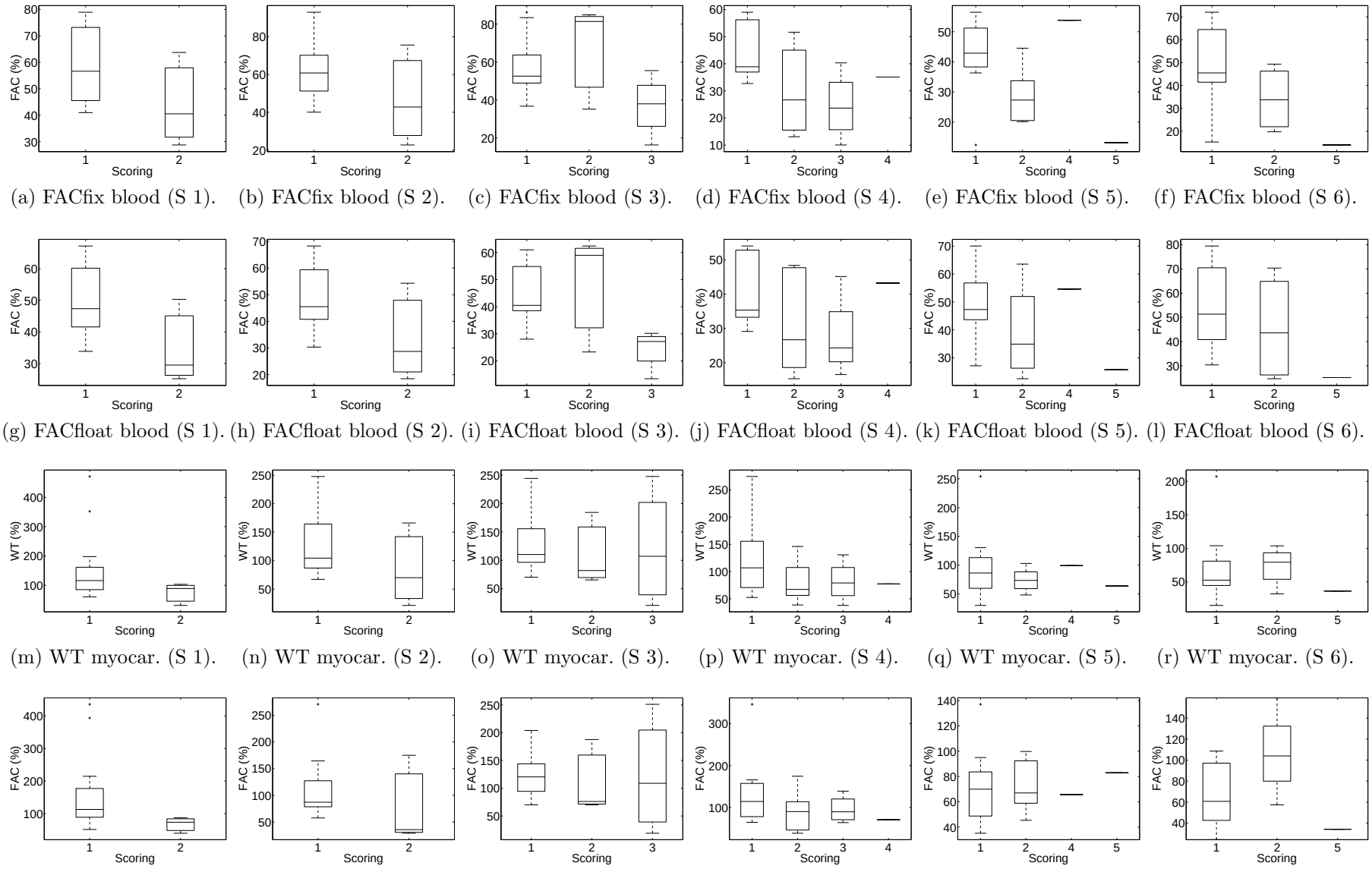

(a) FACfix blood (S 1). (b) FACfix blood (S 2). (c) FACfix blood (S 3). (d) FACfix blood (S 4). (e) FACfix blood (S 5). (f) FACfix blood (S 6).

(g) FACfloat blood (S 1). (h) FACfloat blood (S 2). (i) FACfloat blood (S 3). (j) FACfloat blood (S 4). (k) FACfloat blood (S 5). (l) FACfloat blood (S 6).

(m) WT myocar. (S 1). (n) WT myocar. (S 2). (o) WT myocar. (S 3). (p) WT myocar. (S 4). (q) WT myocar. (S 5). (r) WT myocar. (S 6).

(s) FAC myocar. (S 1). (t) FAC myocar. (S 2). (u) FAC myocar. (S 3). (v) FAC myocar. (S 4). (w) FAC myocar. (S 5). (x) FAC myocar. (S 6).

Figure 3.7: Local function measures for 2C. 21 patients. Scoring: 1 = normokinesis or hyperkinesis, 2 = hypokinesis, 3 = akinesis, 4 = asynchronous normokinesis, 5 = asynchronous hypokinesis.



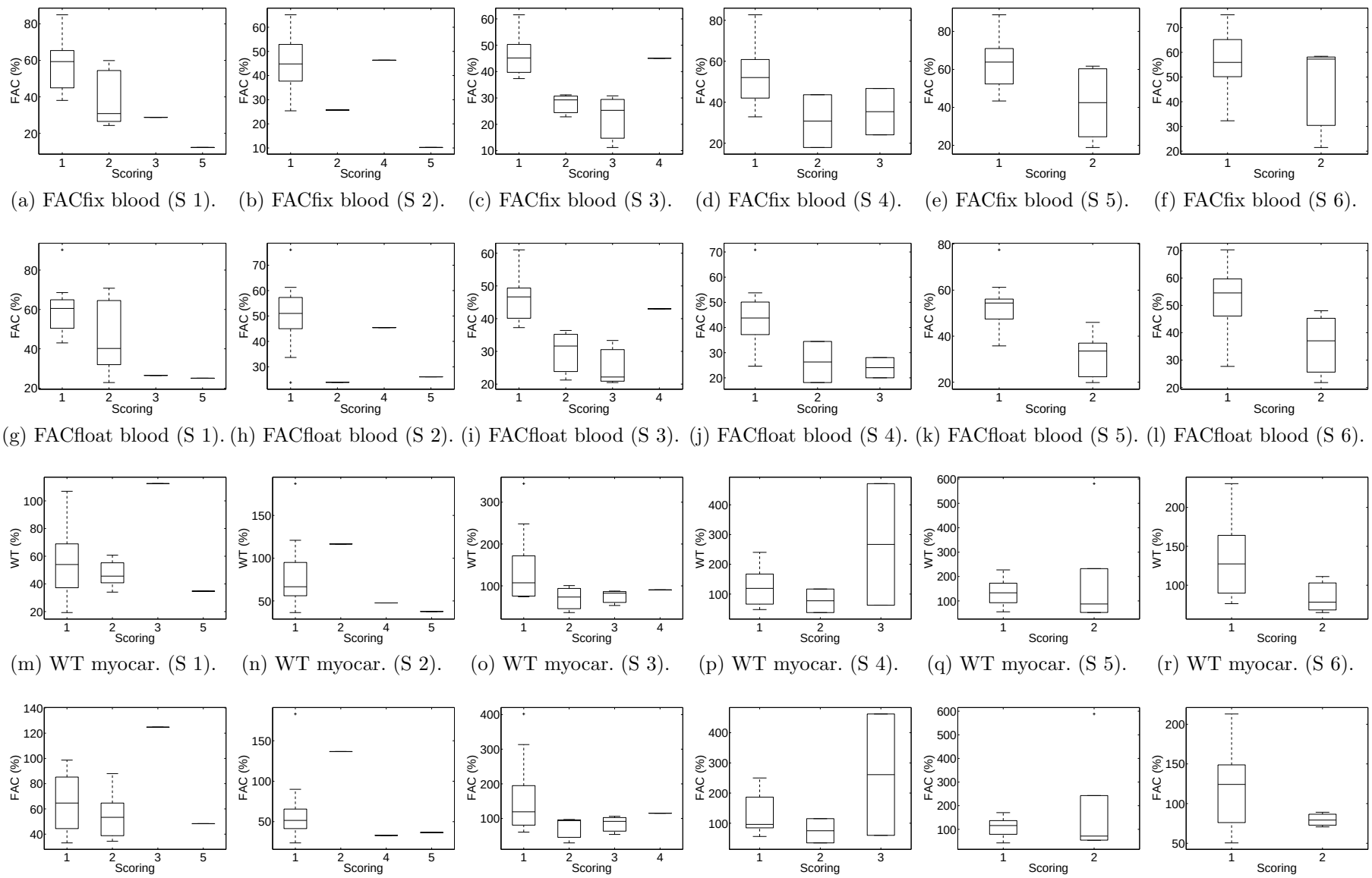

(a) FACfix blood (S 1). (b) FACfix blood (S 2). (c) FACfix blood (S 3). (d) FACfix blood (S 4). (e) FACfix blood (S 5). (f) FACfix blood (S 6).

(g) FACfloat blood (S 1). (h) FACfloat blood (S 2). (i) FACfloat blood (S 3). (j) FACfloat blood (S 4). (k) FACfloat blood (S 5). (l) FACfloat blood (S 6).

(m) WT myocar. (S 1). (n) WT myocar. (S 2). (o) WT myocar. (S 3). (p) WT myocar. (S 4). (q) WT myocar. (S 5). (r) WT myocar. (S 6).

(s) FAC myocar. (S 1). (t) FAC myocar. (S 2). (u) FAC myocar. (S 3). (v) FAC myocar. (S 4). (w) FAC myocar. (S 5). (x) FAC myocar. (S 6).

Figure 3.8: Local function measures for 3C. 21 patients. Scoring: 1 = normokinesis or hyperkinesis, 2 = hypokinesis, 3 = akinesis, 4 = asynchronous normokinesis, 5 = asynchronous hypokinesis.



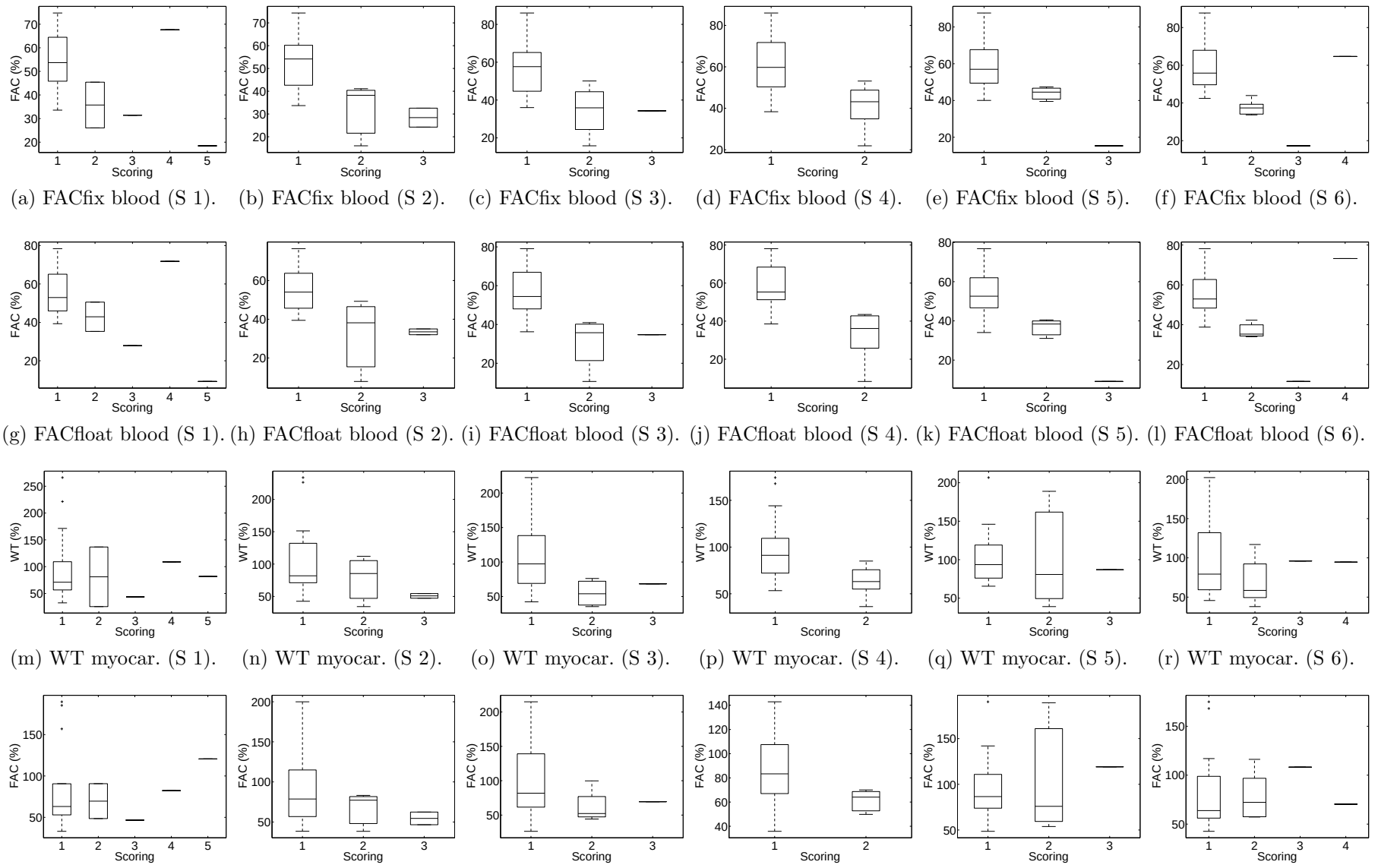

(a) FACfix blood (S 1). (b) FACfix blood (S 2). (c) FACfix blood (S 3). (d) FACfix blood (S 4). (e) FACfix blood (S 5). (f) FACfix blood (S 6).

(g) FACfloat blood (S 1). (h) FACfloat blood (S 2). (i) FACfloat blood (S 3). (j) FACfloat blood (S 4). (k) FACfloat blood (S 5). (l) FACfloat blood (S 6).

(m) WT myocar. (S 1). (n) WT myocar. (S 2). (o) WT myocar. (S 3). (p) WT myocar. (S 4). (q) WT myocar. (S 5). (r) WT myocar. (S 6).

(s) FAC myocar. (S 1). (t) FAC myocar. (S 2). (u) FAC myocar. (S 3). (v) FAC myocar. (S 4). (w) FAC myocar. (S 5). (x) FAC myocar. (S 6).

Figure 3.9: Local function measures for 4C. 21 patients. Scoring: 1 = normokinesis or hyperkinesis, 2 = hypokinesis, 3 = akinesis, 4 = asynchronous normokinesis, 5 = asynchronous hypokinesis.




(a) FACfix blood (S 1). (b) FACfix blood (S 2). (c) FACfix blood (S 3). (d) FACfix blood (S 4). (e) FACfix blood (S 5). (f) FACfix blood (S 6).

(g) FACfloat blood (S 1). (h) FACfloat blood (S 2). (i) FACfloat blood (S 3). (j) FACfloat blood (S 4). (k) FACfloat blood (S 5). (l) FACfloat blood (S 6).

(m) WT myocar. (S 1). (n) WT myocar. (S 2). (o) WT myocar. (S 3). (p) WT myocar. (S 4). (q) WT myocar. (S 5). (r) WT myocar. (S 6).

(s) FAC myocar. (S 1). (t) FAC myocar. (S 2). (u) FAC myocar. (S 3). (v) FAC myocar. (S 4). (w) FAC myocar. (S 5). (x) FAC myocar. (S 6).

Figure 3.10: Local function measures for SAX. 21 patients. Scoring: 1 = normokinesis or hyperkinesis, 2 = hypokinesis, 3 = akinesis, 4 = asynchronous normokinesis, 5 = asynchronous hypokinesis.

## 3.5 Summary and conclusions

Current scientific consensus highlights the importance of Left Ventricle (LV) functional evaluation using fundamental imaging echocardiography, whether to assess the health of the whole LV (global function), or of individual myocardial segments (local function).

This chapter discussed functional measures obtained from smooth approximations to the LV myocardial boundaries in 2D+t Power Modulation contrast echocardiography cine loops. This discussion, as well as the conclusions presented in this section, set the scene for the rest of this thesis, that tackles the problem of automatically segmenting the aforementioned LV smooth contours.

The recommended and best validated quantitative measure for global function is Ejection Fraction (EF), usually computed using the biplane method of disks (modified Simpson's rule). Information provided by EF is limited, though, as local function abnormalities can be underestimated or masked in a global measure.

To tackle this problem, the myocardium can be conveniently partitioned into 16 imaginary segments, and each segment evaluated separately. Local measures are more sensitive than global measures, but they have some limitations too. For instance, there is a connection between Coronary Artery Disease (CAD) and Regional Wall-Motion Abnormalities (RWMAs), but CAD $\not\Rightarrow$ RWMAs, RWMAs $\not\Rightarrow$ CAD, and not all coronary arteries map to the same myocardial segments in all subjects. Those limitations notwithstanding, local function evaluation is of great clinical value.

Local function evaluation in the clinical setting usually employs a standard *qualitative* scoring system, with categories such as normokinetic, hypokinetic, etc. There is no scientific consensus on which *quantitative* measures are optimal, though. For endocardial motion, I selected Fractional Area Change (FAC). For myocardial thickening I selected FAC and Wall Thickening (WT), computed from the Symmetric Nearest Neighbour (SNN) correspondence.

For endocardial motion, FAC was computed with respect to fixed (FACfix) and floating

(FACfloat) centroids, but both methods showed similar results. Likewise, results for myocardial thickening using WT or FAC were similar.

Global evaluation results suggest that it is hard to build a normal/abnormal classifier based on EF alone, due to the wide range of EF values for abnormal subjects. Correcting global EF values with mean local scoring values, though, showed better separability and, more importantly, allowed for the detection of outliers. EF outliers had several expected causes:

- Shifts in the interrogation plane halfway through the cycle, or the interrogation plane being off-axis. These are problems intrinsic to 2D imaging, and will not be an issue with 3D imaging.

- Hypokinetic segments being dragged by healthy adjacent muscle. This is related to the physiology, and illustrates the limitations of global and local endocardial motion evaluation.

- Errors in hand tracing due to very poor image quality produced by lack of contrast, as not even contrast agents can currently make all studies valid for functional evaluation.

As mentioned above, it has been assumed that local function evaluation supersedes global function evaluation. But experiments in this chapter found some drawbacks of local evaluation, as functional heterogeneity and the lack of a general contractility model force each segment to be treated separately. Thus, many more studies are required to characterise local function than global function. Our database of 21 patients does not contain enough cases with the different levels of local function abnormalities in each segment for this type of analysis. Moreover, given its cost and slowness, it is clear that enlarging the database to an appropriate size by expert hand segmentation would be unfeasible in practice.

This problem highlights another important application of automatic segmentation of smooth approximations to the LV myocardial boundaries sought by this thesis. Not only

will fast automatic segmentation help with quantitative clinical evaluation, but it will also help to compile databases large enough to build statistical models of normal and abnormal local function measures.

Finally, this chapter discussed the ASE recommendation of using myocardial thickening as a complement to improve functional evaluation. Results from my experiments measuring WT and FAC suggest that myocardial thickening values are not reliable. On the one hand, this could be due to myocardial thickening measures being less robust against noise and errors. On the other hand, this could mean that our clinical hypothesis was incorrect. Power Modulation was chosen assuming that perfusion would highlight the myocardium. However, visual inspection of the studies shows that perfusion backscatter is not strong enough, and that the external boundary is virtually invisible in large areas of the image. Therefore, it is possible that errors in the hand tracing of the external wall are large enough to preclude Power Modulation contrast echocardiography as an appropriate technique to measure myocardial thickening. Further research is needed to substantiate this claim.

# CHAPTER 4

# Cardiac segmentation and deformable models

## 4.1 Introduction

The previous chapter showed the interest of finding or *segmenting* smooth myocardial boundaries for functional evaluation of the Left Ventricle (LV), but hand tracing is a tedious and time consuming task unfit for the clinical setting. Availability of computer algorithms that can segment myocardial boundaries in the echocardiography image would help human experts make faster, more reproducible and accurate diagnoses.

This chapter analyses a segmentation framework to find myocardial boundaries on the Power Modulation contrast echocardiographic 2D+t cine loops. First, section 4.2 presents the segmentation problem, the difficulties it meets, and the pieces of information available to human experts and computers: texture, geometry and kinetics. To analyse texture, section 4.3 presents histograms with the distribution of intensity values in the contrast echo image. Next, section 4.4 reviews ways to represent the geometry of cardiac contours as configurations of landmarks, and how they can be mapped using Procrustes Alignment to obtain so-called *shapes* (geometries where location, size and rotational effects

are removed). Section 4.5 explains Principal Component Analysis (PCA) as a method to compute shape spaces, or spaces of shape variation. The next two sections study specific aspects of the PCA shape space: section 4.6, its dimensionality; and section 4.7, the importance of the Gaussianity assumption in the 2D shape model. Section 4.8 explored atlas-based deformable template models. In particular, this section studied how these models integrate a probabilistic atlas (texture) and a template (geometry) in a segmentation algorithm with a border search region. Integration of kinetics was left for Ch. 5. An extended version of atlas-based deformable models are Active Appearance Models (AAMs). A historical review and a description of AAMs are presented in section 4.9. The next sections analyse specific aspects of AAMs: intervolume intensity normalisation in section 4.10, how to combine shape and texture variables in section 4.11, the meaning of the correlation and covariance matrices in section 4.12, and how to compute a scaling factor for combination of variables in section 4.13. Section 4.14 notes that AAMs and atlas-based segmentation have not been compared in the literature, even though the former is an extension of the latter, and proceeds to evaluate their performance. Section 4.15 discusses the feasibility of computing PCA shape and texture models, considering realistic lengths of the corresponding vectors, and of training data sets. Finally, section 4.16 summarises the findings and conclusions from this chapter.

## 4.2 The segmentation problem

Contrast echocardiography data (see examples in Fig. 2.2) is inherently hard for myocardial segmentation, both for edge detection and texture based approaches.

Edge detection methods find that there are not clearly defined edges because the image is heavily textured, the echo depends on the local orientation of the contour to the beam, the endocardium is trabeculated, and the epicardial interface returns at best a weak echo. Besides, there are some well-known artifacts: the papillary muscles come in and out of the image plane and create edges as strong as for the endocardium; ribs and the lung obstruct

the beam and mask out the actual cardiac border while at the same time creating a new strong edge at the boundary of the shadow; and intramural arteries produce strong edge signals within the myocardium.

Texture based approaches, on the other hand, find that intensity distribution is non-Gaussian, an assumption in many methods, e.g. PCA and other least-squares model fitting approaches; it does not reflect the physical properties of the imaged material; papillary muscles have similar texture to the myocardium; and rib and lung shadows can remove texture almost completely from a large region of the image [114].

An artifact common to both approaches is the *attenuation field*. Ultrasound images are created from a linear array of ultrasound beams, that are attenuated differently depending on their propagation path. Thus, the response at a certain depth is a function of the reflector's echogenicity as well as the previous history of the beam (for attenuation field estimation methods see e.g. [80, 189]). This problem is exacerbated in contrast ultrasound, due to the high acoustic impedance of microbubbles [114].

Another artifact characteristic of contrast echo is *swirling*, caused by bubble destruction near the probe, so that contrast is lost in the near field [114]. The opposite problem is having too much contrast, so intensity levels are saturated, and detail is lost near the apex.

There are two reasons why human experts can dismiss artifacts and fill in missing information. First, they combine many sources of information, e.g. patient data and medical history, several cardiac views, and temporal evolution of edges and texture on the image. And second, they use their experience and knowledge to interpret all the data. A Level 3 echocardiographer (the highest level of training), at the very minimum, has trained for 12 months, performed 300 transthoracic two-dimensional and Doppler echocardiograms, and interpreted 750, plus 500 annually for maintenance of competence, besides completing specific training for contrast and stress echo [144]. That is, human experts have learned temporal models of shape and texture, so they can recognise and ignore an intramural signal, or interpolate a segment of a contour in a blackout area using

information outside the shadow.

The scope of this thesis is limited to contrast echo functional analysis of the LV (as opposed to, for example, fundamental imaging or Doppler echo). There are 3 main pieces of information that human experts extract from the images:

**Texture:** The texture in the myocardium is not only darker than in the blood pool, and lighter than outside the heart, but it also has a visibly characteristic textural pattern.

**Geometry:** Analysis is simplified assuming that endocardium and epicardium are smooth continuous boundaries.

**Kinetics:** Images are visualised as a movie, rather than as individual frames; this not only shows whether the myocardium is contracting correctly, but also helps to identify artifacts.

The rest of this chapter analyses texture and geometry, and the integration of both sources of information, whereas the integration of kinetics is left for Ch. 5.

## 4.3 Texture

Ultrasound images are strongly textured with so-called *speckle*. Speckle is a deterministic interference pattern that originates in the image because the transducer acts as a coherent source and as a non-coherent detector which is much bigger than the wavelength. It is the same pattern that appears in Synthetic Aperture Radar (SAR) or lasers. The ultrasound beam is reflected by many scatterers with uncorrelated phases per resolution cell, so that the received signal is the superposition of many small incoherent contributions. This can be modelled as a complex Gaussian random variable with independent and identically distributed (i.i.d.) real and imaginary components. An envelope detector removes the phase information and the image is formed as pixel intensity from the squared magnitude of each resolution cell, usually modelled with a Rayleigh distribution [57].

The image formation system has several non-linear processing stages too, typically Time Gain Compensation (TGC), log compression, several gain controls that can be manually adjusted by the operator, and bilinear interpolation of the scan-lines. The diagram of a typical ultrasound system was shown if Fig. 2.1, and more details about the image formation process can be found, for example, in Mulet Parada [125, Ch. 3, Appendix A].

Unlike other image modalities, e.g. X-rays, speckle does not present a simple relationship between physical properties of reflectors (tissue, blood and boundaries) and intensity levels. In the mid-1980s it was generally agreed that speckle was an annoyance that reduced resolution, and thus diagnostic accuracy, and much effort has been devoted to filter it out (e.g. [9, 31]). More recently, Speckle Tracking Echocardiography has taken advantage of the fact that the size and distribution of scatterers produce unique speckle patterns on the myocardium, that can be tracked throughout the cardiac cycle to evaluate strain (see e.g. [142, 166] for recent publications).

Availability of ultrasound texture models is limited. Tao et al. [176] made a brief review of available theoretical non-contrast speckle models, but noted that they are referred to the transducer, and do not take into account the full image formation process in the ultrasound machine, which is often unknown because manufactures do not make details available (Fig. 2.1 is just a basic diagram). Kaplan and Ma [87] proposed a log-compressed Rayleigh pdf to model intensity values in B-mode ultrasound images, taking into account logarithmic compression performed by the machine. This model has not been tested with actual echocardiography data, neither includes Time Gain Compensation (TGC), harmonic imaging or other factors that could arguably affect the intensity pdf. Tao et al. [176] studied models empirically, testing normal, lognormal, Weibull and gamma distributions for Goodness of Fit (GoF) in regions of interest (tissue and blood) of B-mode SAX frames, and found that none of them fitted the data with statistical significance. Histograms of regional intensity visually suggested left-skewed distributions, nonetheless [176].

To my best knowledge, there are no models for pixel intensity distribution in contrast

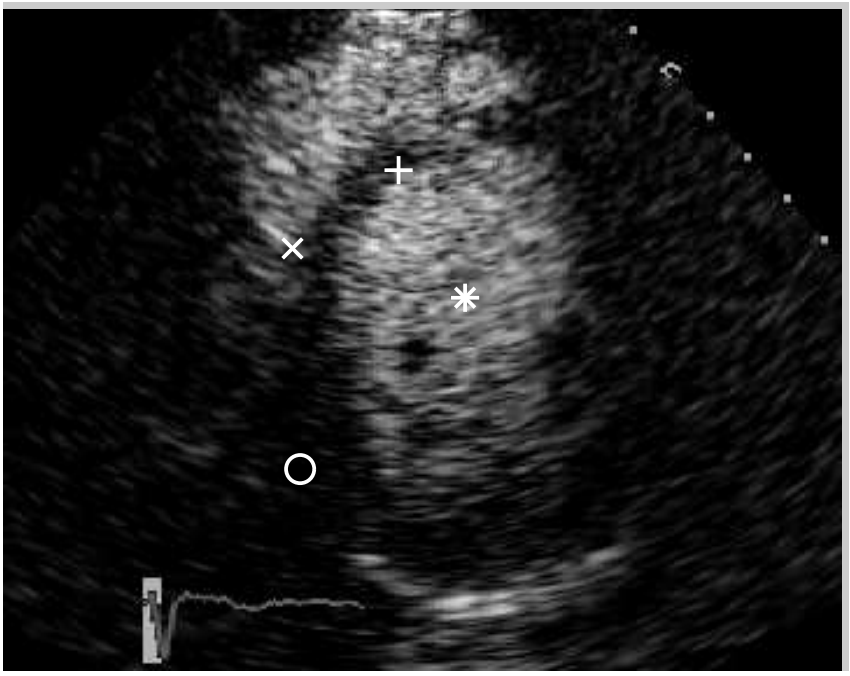


Figure 4.1: Four sampling points in a sample SAX frame: dark region left of LV (O), RV blood pool (X), tissue between RV and LV (+), LV blood pool ($\star$).

echocardiography, so the following experiment was performed to illustrate the texture information. A mean SAX contour was computed from the hand traced contours of 316 SAX frames (all frames from 21 patients). Frames were aligned with a similarity transformation (see section 6.3), and warped with a thin-plate spline (TPS), using landmarks from the myocardial contours and the shapel model's mean contour to define the mappings. Intensity values from aligned images were sampled with a *triple-mask*.

Details on the similarity transformation can be found in section 6.3, whereas details on the TPS warp can be found in Appendix C. For the moment, just assume that we have a way to sample each ultrasound frame so that each given sampling point represents the same anatomical location in all frames. That is, we can study the intensity distribution of each anatomical location over a population of size 316 frames, as opposed to the intensity distribution of a block of pixels in a single frame (e.g. as in [176]). Fig. 4.1 displays four sampling points that were selected to illustrate the differences in the distributions of different regions of the image. Fig. 4.2 displays the histograms for each pixel. To evaluate whether the intensity distributions depend on the image scale, the original 256x360 frames

were scaled down by a factor of 0.2. The resulting histograms are plotted in Fig. 4.3.

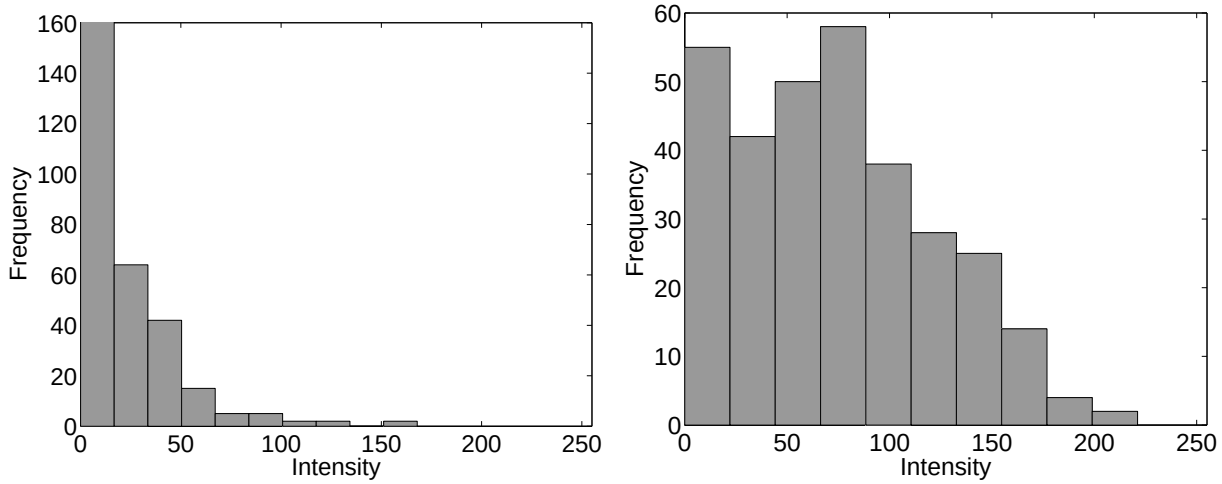
The histograms illustrate that the distribution of a pixel does not change very much over scales, and that tissue pixels in one region are similar to tissue pixels in another region. Similarly, pixels in blood regions are similar. In addition, the variance of the distribution increases with the mean value. This is consistent with the linear relationship observed by Tao et al. [176] in non-contrast regions. Fig. 4.2d presents a distribution that looks quite symmetric. Kadour[1] noted that this could be due to saturation in the image, i.e. the distribution is not really symmetric, but the right tail (very bright values) has been clipped off as the maximum intensity value is 255. Likewise, while Fig. 4.2a suggests an exponential distribution, it could also be that the left side of the distribution has been clipped off due to underexposure. Kadour's observation is consistent with the findings of [176], who had to remove from their study all frames with underexposed or saturated values in order to compute parametric texture models.

The discussion above has outlined some of the issues to be tackled for texture modelling in Power Modulation images. In-depth texture modelling is beyond the scope of this thesis, but still two observations can be made.

First, pixel intensities do not have a Gaussian distribution. The AAM texture approach that will be discussed in section 4.11 is based on a PCA space computed from pixel intensity values. But sections 4.14 and 4.15 will argue against the convenience of using PCA texture models altogether, so intensity gaussianisation has not been tackled in the main body of this thesis. Nonetheless, some preliminary results are presented in Appendix E.

And second, histograms in Fig. 4.2 strongly suggest that residuals variances will not be similar for different regions of the image in the segmentation algorithm's least-squares optimisation presented in Ch. 6. That is, the least-squares solution will be sub-optimal in the Gauss-Markov theorem sense, and the solution will depend mostly on the brighter areas (larger intensity values). But as these areas contain most of the signal in the image, it

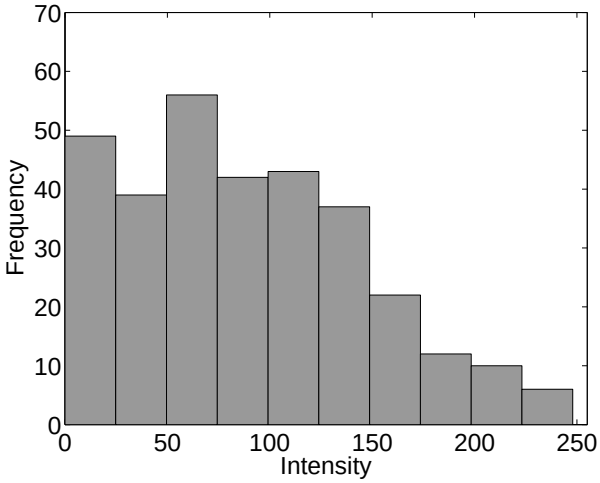
[1] M. Kadour. Medical Vision Laboratory, University of Oxford, UK. Personal communication.



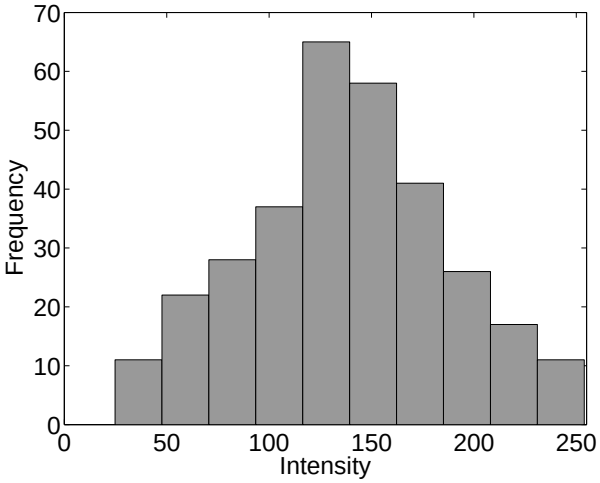

(a) Dark tissue region (O).




(b) RV blood pool (X).

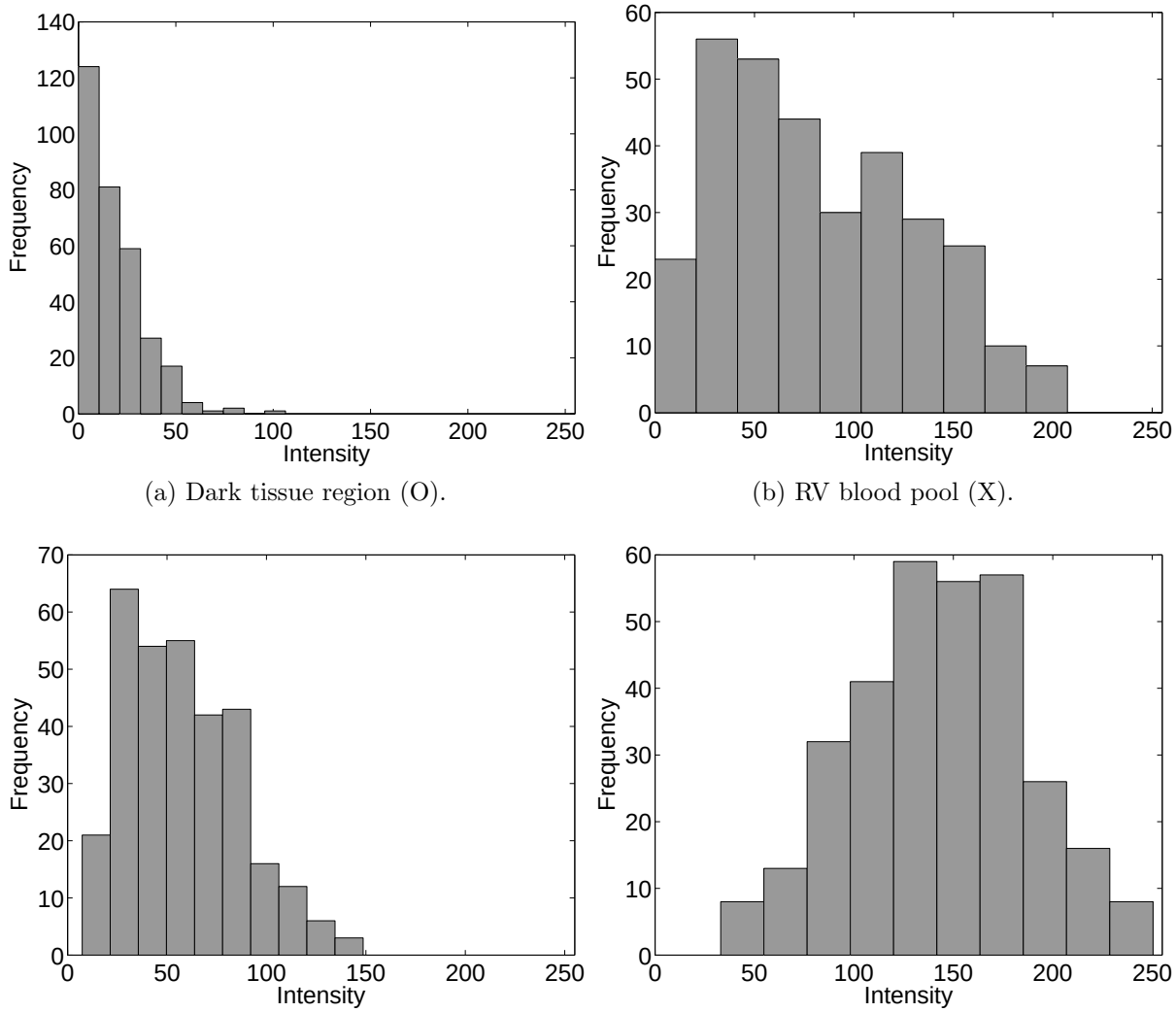



(c) Tissue between RV and LV (+)




(d) LV blood pool ($\star$)

Figure 4.2: Histograms for contrast ultrasound intensity distributions of selected locations in SAX view. Note that the locations are illustrated on a specific frame in Fig. 4.1, but the histograms are computed from the intensity values of that location sampled in every frame of the database.




(a) Dark tissue region (O).

(b) RV blood pool (X).

(c) Tissue between RV and LV (+).

(d) LV blood pool ($\star$).

Figure 4.3: As in Fig. 4.2, but frame size scaled by a factor 0.2.

will be assumed that the segmentation algorithm will still be guided in the right direction.

## 4.4 Geometry, Procrustes alignment and shape

The idea of finding geometric models for the Left Ventricle (LV) is well-established. Yettram and Vinson [191] reviewed papers from 1952 to 1977 that approximated the LV as an ellipsoid, and proposed a non-ellipsoidal 3D finite element model painting with a marker on a plaster figurine sculptured using two plywood templates as a reference. In a review of computer vision in LV segmentation during the second half of the 20th century, Suri [174] noted that there had been 'an explosive growth of model-based LV segmentation and its modelling' from the late 1970s. For a review of 3D models for cardiac functional analysis see Frangi et al. [67].

Bookstein [24] introduced the description of shape based on configurations of landmarks (points with a biological correspondence) from morphometrics to computer vision in the mid-1980s, to explain geometry and deformations. A configuration of landmarks is a vector $s$

$$s = [X_{11},\ \ldots,\ X_{P1},\ X_{12},\ \ldots,\ X_{P2}]^\top \tag{4.1}$$

where $[X_{i1}, X_{i2}]$ are the Cartesian coordinates of the $i$-th landmark (there are $P$ landmarks in total). Kendall [90] noted that this was an approximation to his own work, and worded his original definition of shape [89] as 'Shape is what remains when location, size, and rotational effects are filtered out'. Statistical shape analysis was a mature field by the late-1990s (e.g. Dryden and Mardia [58]). The location, size and rotational effects are the components of a similarity transformation, and are usually filtered out using *Procrustes analysis*. Gower [71] proposed the Generalised Procrustes analysis method to compute iteratively the translation, scaling and rotation transformations that minimises the distance between a set of *configurations* or discrete sets of 2D landmarks and maps them

onto a reference coordinate frame. The method also computes the mean configuration

$$\bar{s} = \frac{1}{M} \sum_{i=1}^{M} s_i \tag{4.2}$$

Rohlf and Slice [146] modified the scaling part of the algorithm and called it Least-Squares Fit Generalised Orthogonal Procrustes Analysis (LSFGOPA). The method in [146] contains an error that causes the iterative alignment to diverge sometimes. I propose a fix and also reformulate some operations to speed them up in Appendix B.

Landmarks provide an anatomical intra- and intersubject correspondence that is exploited in this thesis to model the relationship between geometry, shape variations and kinetics. To explain endocardial geometry, a combination of anatomical and pseudo-landmarks is necessary. Our experts could identify 3 anatomical landmarks in apical planes (1 at the apex, and 2 at both sides of the mitral valve) and 1 in the SAX plane (at the infero-septal beginning of the RV). For functional assessment, endocardial trabeculations are smoothed out as in Fig. 3.2 and the papillary muscles left inside the cavity. Hence, myocardial boundaries can be modelled with simple continuous curves. Our clinical experts proposed sampling at equidistant arc length between anatomical landmarks.

Continuous curves based on B-splines have been used in computer vision from early on [19], and have several interesting properties: easy interactive design; local support; easy and accurate evaluation; they are continuous and can be smooth, but can also contain corners by duplication of interior knots. In addition, they allow the representation of continuous curves as sets of landmarks, thus bridging continuous and discrete shape models. For instance, Baumberg and Hogg [10] used spline control points as landmarks. However, curves in [10] were approximating splines (i.e. the control points are not curve points) and in practice finding the control points is cumbersome and does not provides an intuitive set of landmarks for the correspondence between contours. This is the same problem that was found for Quamus contours in section 2.4. Alternative approaches have used as landmarks the coefficients of other basis decompositions, instead of B-splines. For

example, the Discrete Cosine Transform (DCT) [72] or a Wavelet Transform (WT) [55]. Two points from those publications are worth mentioning. Hamarneh and Gustavsson [72] claimed that using DCT coefficients 'eliminate[s] the need for point correspondence'. This is not the case, as in order to compute the DCT, curves need to be parameterised, and this in fact implies solving a dense correspondence problem. Davatzikos et al. [55] obtained better segmentation results using WT coefficients, but only for data sets with many fewer training shapes than landmarks.

I propose using the control points of periodic interpolating cubic splines, as described in section 2.4. This way no information of the continuous curve is lost, and control points are landmarks or pseudo-landmarks that define a sensible anatomical correspondence.

## 4.5 PCA space of shape variations or shape model

Staib and Duncan [167] noted that segmentation of natural objects found in biomedical images is 'doomed' if local information is not constrained by a space of shape variations, or shape model, that should be as 'specific as possible' and be incorporated 'explicitly, specifically, and early in the analysis'. That is, the LV geometry discussed in section 4.4 requires to be embedded into a shape model.

Principal Component Analysis (PCA) [77, 138], also known as the Karhunen-Loève transform, is one of the most popular methods in Statistics for modelling, dimensionality reduction and denoising. It was introduced by Sirovich and Kirby [164] in computer vision literature as a dimensionality reduction method for face images in the late-1980s.

PCA finds a basis of orthonormal vectors that span the data set. The first vector has the direction of maximum variance of the data. The next component has the direction of maximum variance amongst those orthogonal to the first, and so on. Cootes et al. [47] proposed computing a shape space applying PCA to landmark configurations, and called

it the Point Distribution Model (PDM)

$$s = \bar{s} + V\,b \tag{4.3}$$

where $\bar{s}$ is the mean shape, $V$ is the shape space matrix, and $b$ is the *coefficient or Principal Components (PC)* vector. In this model, $s$ is the configuration of $P = n/2$ landmarks vector in Eq. (4.1). In terms of statistical analysis, $s$ is a vector with $n$ random variables. The model is learned from a training set of $M$ examples $S = [s_1,\ \ldots,\ s_M]$. The mean shape $\bar{s}$ is given by (4.2). The eigenvectors or *loading vectors* $V = [v_1,\ \ldots,\ v_M]$ are computed using PCA; that is, as solutions to the eigenproblem

$$Cv = \lambda v \tag{4.4}$$

where $\lambda$ is an eigenvalue, and $C$ is the covariance matrix

$$C = \frac{1}{M}\sum_{i=1}^{M}(s_i - \bar{s})(s_i - \bar{s})^\top = \frac{1}{M}\tilde{S}\tilde{S}^\top \tag{4.5}$$

of the centred training set $\tilde{S}$

$$\tilde{s}_i = s_i - \bar{s} \tag{4.6}$$

If there are more variables than training vectors, i.e. $n > M$, then (4.4) is solved faster using Multidimensional Scaling (MDS) (e.g. [150], writing $v_i$ as a linear combination of training vectors[2]

$$v_i = \begin{cases} \tilde{S}a_i'/\sqrt{\lambda_i'}, & \lambda_i' \neq 0 \\ 0, & \lambda_i' = 0 \end{cases} \tag{4.7a}$$

$$\lambda_i = \lambda_i'/M \tag{4.7b}$$

[2]With this definition, both $\|v_i\|^2 = 1$, $\|a_i\|^2 = 1$.

where $a_i$ are called the *coefficient eigenvectors*. In matrix form

$$V = \tilde{S} A' \Lambda'^{1/2} \tag{4.8}$$

where $\Lambda'$ is a diagonal matrix with $\lambda'_i$ as the $i$-th element in the main diagonal. It can be shown (e.g. [158]) that left multiplying by $\tilde{S}$, and substituting $C$ by (4.5) and $v$ by (4.7), the eigenproblem (4.4) is equivalent to

$$\tilde{K} a'_i = \lambda'_i a'_i \tag{4.9}$$

where

$$\tilde{K} = \tilde{S}^\top \tilde{S} \tag{4.10}$$

## 4.6 Dimensionality of shape models

Dimensionality is a measure of the flexibility or degrees of freedom of a model. Selecting the dimensionality of a shape model can be seen as a compromise between *Goodness of Fit* (GoF) (how well the model fits the data) and *parsimony* (how simple the model is). In the PDM, this compromise is achieved keeping only the $k \leq n$ eigenvectors of $V$ with the largest eigenvalues. When the shape model is used in segmentation, other considerations are relevant too, e.g. whether more flexible shape models increase the number of local minima in the target function, and so impair the convergence of the optimisation algorithm. This will be discussed in more detail in section 6.2.

In this section, three approaches to compute the model dimensionality based are presented. 1) The ratio of explained variance is an empirical method that is widely used, but without a straightforward connection to clinical measures. 2) The Generalised Information Criterion is a family of theoretical methods that attempt to identify the cut-off point where eigenvectors model noise instead of data. 3) I propose, as an application specific heuristic, an anatomical criterion based on the approximation error of landmarks.

PDM literature regarding dimensionality selection is very limited. Cootes et al. [47] originally proposed the *variance criterion*, that selects $k$ so that a certain proportion $f_v$ of the total variance $V_T = \sum_i \lambda_i$ of the data is explained by the model

$$\sum_{j=1}^{k} \lambda_i \geq f_v V_T \tag{4.11}$$

Cootes and Taylor [46] mentioned, but did not try, using an estimate of noise variance or the approximation error in a leave-one-out scheme too. Recent publications (e.g. [103]) use the ratio of explained variance to fix the dimensionality of the model. Values selected for $f_V$ have been e.g. 90%, 92%, 97% [15], 95% [50, 97, 98, 168], 96% [47], 98% [42, 48, 62], 90%-95% [1], without justification, or relating them to any useful anatomical criterion. To my best knowledge, Stegmann [170, sec. 7.5] analysed alternatives for computing $k$ for the first time in PDM literature: cross-validation, bootstrapping and a data permutation version of parallel analysis, deciding to use the latter. Stegmann warned, however, that Monte Carlo simulation was necessary, and that experiments were 'done with replacement due to the massive effort involved in keeping track of the permutations'. Very recently, Mei et al. [116] used resampling techniques and measured the stability of PCA eigenvector directions[3] to compute the intrinsic dimensionality of the model, i.e. the number of modes that explain anatomical variation as opposed to noise. Their results strongly suggest that the $95\%V_T$ rule is unreliable, but it is worth noting two issues. First, that they replaced the arbitrary threshold $f_V = 95\%$ by another arbitrary threshold —a significance level of 5%— neither of which can be easily related to a measure related to anatomy or physiology. And second, that their $t$-test is heterodox, as they set as the null-hypothesis $H_0 : \xi(\alpha_1) > \xi(a_i)$, where $\xi$ is a *mode instability* measure, and $\alpha_1,\ a_i$ are noise and data terms. Canonically, the null-hypothesis should be $H_0 : \xi(\alpha_1) = \xi(a_i)$, and the alternative hypothesis, $H_1 : \xi(\alpha_1) > \xi(a_i)$. However, Mei et al.'s paper was published too close to

[3] Mei et al. [116] based their measure of eigenvector stability on the correspondence among PCA modes computed from different bootstrap replicates.

submission of this thesis to discuss their methodology in depth.

Model selection techniques exist in Statistics literature, but a systematic review is beyond the scope of this thesis. Nonetheless, I will focus on a family of widely used criteria that can be obtained from an alternative formulation of PCA, the Information Criteria, and that are very fast to compute.

Tipping and Bishop [179] showed that finding the intrinsic dimensionality of a PCA model can be formulated as a Factor Analysis problem where the error term is isotropic. This is known as Probabilistic PCA (PPCA) [180]. In more detail, the training data is modeled as a function of a loading matrix $A$, latent variables $y$ and isotropic noise $e \sim \mathcal{N}(0,\ \sigma^2\mathbb{I})$

$$s = Ay + e \tag{4.12}$$

The model parameters $A,\ \sigma^2$ are estimated so that they maximise the log likelihood function

$$L(\hat{A},\ \hat{\sigma}^2) = -\frac{M}{2}\ln|\hat{A}\hat{A}^\top + \hat{\sigma}^2\mathbb{I}| - \frac{M}{2}\mathrm{tr}\left(C(\hat{A}\hat{A}^\top + \hat{\sigma}^2\mathbb{I})^{-1}\right) \tag{4.13}$$

where $M$ is the sample size, $C$ is the covariance matrix given by (4.5), and

$$\hat{\sigma}^2 = \frac{1}{n-k}\sum_{i=k+1}^{n}\lambda_i \tag{4.14a}$$

$$\hat{A} = V_k\left(\Lambda_k - \hat{\sigma}^2\mathbb{I}\right)^{1/2} \tag{4.14b}$$

where $V_k,\ \Lambda_k$ are the first $k$ columns of the $V,\ \Lambda$ matrices, respectively. While non-singular covariance matrices are positive definite and, thus, have all eigenvalues $\lambda_i > 0$, it should be noted that $C$ is an estimate subject to sampling error and small violations of the underlying hypotheses, e.g. Gaussianity or independence. In practice, this can produce one or more small negative eigenvalues, a problem not acknowledged in the PDM literature, but studied for the equivalent Factor Analysis [184] and MDS [101] formulations. In case the negative eigenvalues have a small modulus, they can be attributed to noise and disregarded. Within the PPCA framework, finding the optimal $k$ is formulated as

minimising the Generalised Information Criterion (GIC) [78, 159]

$$k_{opt} = \min_k J(\hat{A},\ \hat{\sigma}^2,\ k) \tag{4.15}$$

The GIC can be expressed as

$$J(\hat{A},\ \hat{\sigma}^2,\ k) = -2L(\hat{A},\ \hat{\sigma}^2) + C(M)D(k) \tag{4.16}$$

where $D(k)$ is the number of free parameters or dimensionality of the model

$$D(k) = nk + 1 - \frac{k(k-1)}{2} \tag{4.17}$$

and $C(M)$ is a weighting factor. That is, a compromise is found between how well the model represents the data ($L$) and the model complexity ($D$). Depending on $C(M)$, the GIC is known by different names: Akaike's Information Criterion (AIC), Consistent AIC (CAIC) or Bayesian Information Criterion (BIC) (for pointers see e.g. [78])

$$C(M) = \begin{cases} 2, & \text{AIC} \\ \ln(M) + 1, & \text{CAIC} \\ \ln(M), & \text{BIC} \end{cases} \tag{4.18}$$

To improve the performance of AIC, the unbiased version of AIC, the finite sample corrected AIC (AICc), was defined as [81]

$$\text{AICc} = \text{AIC} + \frac{2(D(k)+1)(D(k)+2)}{n - D(k) - 2} \tag{4.19}$$

However, Hu and Xu [78] found that AIC tends to overfit the model, so it is to be expected that AICc actually performs worse. Hu and Xu [78] also found that CAIC tends to underfit the model, while BIC is more accurate. Minka [120] used the PPCA framework

to compute the posterior probability density of the training dataset $S$ given the model and proposed Laplace's Information Criterion (LIC) as

$$p(S|k) \approx p(U) \left( \prod_{j=1}^{k} \lambda_j \right)^{-M/2} \hat{v}^{-M(n-k)/2} (2\pi)^{(m+k)/2} |A_Z|^{-1/2} M^{-k/2} \tag{4.20}$$

where

$$m = nk - \frac{k(k+1)}{2} \tag{4.21}$$

$$p(U) = 2^{-k} \prod_{i=1}^{k} \Gamma \left( \frac{n-i+1}{2} \right) \pi^{-(n-i+1)/2} \tag{4.22}$$

$$\hat{v} = \frac{\sum_{j=k+1}^{n} \lambda_j}{n-k} \tag{4.23}$$

$$|A_Z| = \prod_{i=1}^{k} \prod_{j=i+1}^{n} \left( \hat{\lambda}_j^{-1} - \hat{\lambda}_i^{-1} \right) (\lambda_i - \lambda_j) M \tag{4.24}$$

$$\hat{\lambda}_i = \begin{cases} \lambda_i, & i \leq k \\ \hat{v}, & \text{otherwise} \end{cases} \tag{4.25}$$

Values in (4.20) can be larger than floats, so to avoid overflowing the $\ln(p(S|k))$ is computed instead[4]. The intrinsic dimensionality of the model $k_{opt}$ is given by

$$k_{opt} = \max_k \ln(p(S|k)) \tag{4.26}$$

Minka [120] also proved that BIC is an approximation of LIC when the terms that grow with $M$ are dropped. Synthetic experiments suggested that LIC outperforms BIC.

To illustrate the differences among the above approaches, the optimal dimensionality for each criterion was computed on a training dataset of 20 baseline studies in 2C. The results are displayed in Fig. 4.4 and Table 4.1. Computing times are very small, so time differences are not significant in practice. To place dimensionality values in anatomical

[4]T. Minka. Media Laboratory, MIT, USA. Personal communication. T. Minka's implementation of LIC can be downloaded from `http://research.microsoft.com/~minka/papers/pca/`.

context, the following mean landmark distance error was used

$$d_{\text{mean}} = \frac{1}{P}\sum_{i=1}^{P} \|X(i) - X'_k(i)\| \tag{4.27}$$

where $P$ is the number of landmarks in a contour, $X(i)$ is the $i$-th landmark, and $X'_k(i)$ is the approximation to $X(i)$. A good choice of distance units is pixels, as all images have the same size but different spatial resolution. In brief, to compute the approximation $X'_k$, landmarks are first projected onto the first $k$ eigenvectors of the shape space and then back to input space. This process is explained step by step in the rest of this section.

Rewriting the definition of the configuration vector $s$ from Eq. (4.1), so that it matches the notation above

$$s = [X(1)_1,\ \ldots,\ X(P)_1,\ X(1)_2,\ \ldots,\ X(P)_2]^\top \tag{4.28}$$

where $X(i) = [X(i)_1,\ X(i)_2]$. From the definition of the PDM in Eq. (4.3), that I reproduce here for convenience

$$s = \bar{s} + Vb$$

the shape space $V$ is truncated so that only the first $k$ eigenvectors are kept. The truncated space will be called $V_k$. The projection $b_k$ of $s$ onto $V_k$ is

$$b_k = V_k^\top (s - \bar{s})$$

The back projection onto input space is

$$s'_k = \bar{s} + V_k b_k \tag{4.29a}$$

$$= \bar{s} + V_k V_k^\top (s - \bar{s}) \tag{4.29b}$$

If and only if all eigenvectors have been kept and $V_k$ is full rank, then $V_k V_k^\top = \mathbb{I}$, $s'_k = s$.

| | 95% $V_T$ | 98% $V_T$ | Anat. | CAIC | BIC | LIC | AIC | AICc |
|---|---|---|---|---|---|---|---|---|
| $k_{opt}$ | 8 | 10 | 10 | 48 | 51 | 51 | 51 | 51 |
| time (ms) | 4.5 | 4.5 | N/A | 83.1 | 84.5 | 39.5 | 85.0 | 82.4 |
| median $d_{\text{mean}}$ (pix) | 2.2 | 1.3 | 1.3 | 1.1e-2 | 4.3e-15 | 4.3e-15 | 4.3e-15 | 4.3e-15 |

Table 4.1: Quantitative measures from Fig. 4.4. Time values correspond to computations on a PC with a Xeon Dual CPU 2.66GHz and 2GB of RAM.

From Table 4.1 and Fig. 4.4, information criteria (CAIC, BIC, LIC, AIC, AICc) results suggest that almost all 52 variables are significant to explain the data. But while the dimensionality chosen by information criteria may be optimal in a noise or redundancy sense, it is too high from a practical point of view. Fig. 4.4 shows that subpixel accuracy is possible with as few as 15 eigenvectors. Lowering the shape model's dimensionality as much as possible is important, because a too flexible shape model could impair the convergence of the segmentation algorithm, as it was pointed out at the beginning of this section.

The dimensionality obtained by the variance criterion is even smaller than 15, but as discussed above, there is no reliable method to fix the threshold. For 95% $V_T$, the optimal dimensionality is 8 eigenvectors, and the median $d_{\text{mean}}$ is 2.2 pixels. For 98% $V_T$, the optimal dimensionality is 10 eigenvectors, and the median $d_{\text{mean}}$ is 1.3 pixels, a 41% decrease.

To overcome the limitations of both the information and variance criteria, I propose as an anatomical/functional criterion the dimensionality $k$ that makes the median $d_{\text{mean}} \leq$ 1.5 pixels, as myocardial thickness can typically be as small as 15 pixels in end diastole. The dimensionality for the anatomical criterion is 10 eigenvectors in the example above. This value matches the value of the 98% $V_T$ criterion in this instance, but we have no reasons to assume that it will do so in all cases.

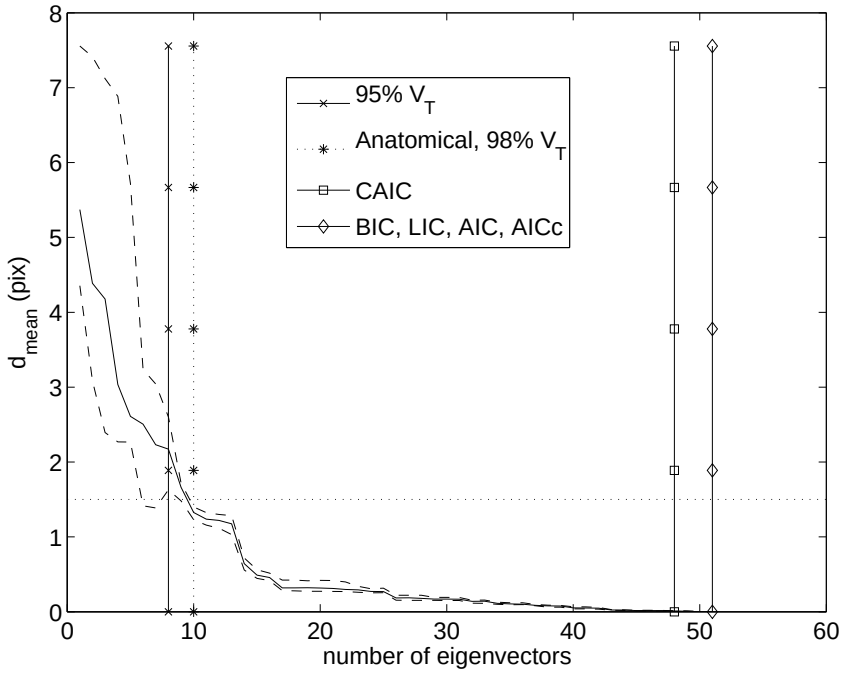



Figure 4.4: Endocardial 2D+t model dimensionality according to several criteria discussed in the main text. The 2D+t model is the explicit cyclic model proposed in section 5.2. The curvy graph corresponds to $d_{\text{mean}}$ (Dashed: 95% CI. Solid: Median). Vertical lines correspond to $k_{opt}$ values. The horizontal dotted line corresponds to the anatomical criterion threshold of 1.5 pixels. Data set from 20 patients, 334 frames, 50 spatial variables (25 points) and 2 temporal variables. Quantitative measures are available in Table 4.1.

## 4.7 Gaussianising shape vectors

PCA assumes that the data has a multidimensional normal distribution. When this is not the case, the resulting model may not be able to describe the data correctly. Bosch et al. [26] pointed out this problem for ultrasound intensity data, but the same ideas can be applied to shape vectors. They proposed aggregating all intensity variables as a single variable $g$ and applying the gaussianising transformation $T : \mathbb{R} \mapsto \mathbb{R}$

$$\hat{g} = T(g) = F_Z^{-1}(F_g(g)) \tag{4.30}$$

where $F_Z$ is the theoretical Cumulative Distribution Function (CDF) of a standardised Gaussian variable, $Z \sim \mathcal{N}(0,\ 1)$ and $F_g$ is the estimated CDF of the variable to be gaussianised. Although not mentioned by Bosch et al. [26], it can be shown that $T$ is optimal in the sense that the Kullback-Leibler divergence between the normal distribution and the distribution of $\hat{g}$ is minimal (e.g. [41, 149]). Gaussianisation has been explored as a multidimensional problem too. For example, Chen and Gopinath [41] proposed an iterative method where in each iteration Independent Component Analysis (ICA) is used to separate the input variables into least dependent components. Then a mixture of Gaussians is used to approximate $F_g$ and each component is gaussianised. The aggregated approach of [26] is too naive, as it assumes that all variables have the same distribution. The method in [41] is computationally expensive and does not provide a simple way to invert the gaussianisation transformation.

I propose a method sharing the best features of [26] and [41]. To avoid the iterative procedure of [41] and have an easy way to degaussianise the data, it can be hypothesised that the individual gaussianisation of input variables would gaussianise the PCA components. This hypothesis has two theoretical limitations. First, PCA uncorrelates the input variables, a necessary but not sufficient condition of independence, so not even Gaussianity of the components guarantees Gaussianity of the input variables. And second, input variables are a linear transformation of the PCA components. If we had Gaussian and

independent components, then the input variables would be normally distributed as well (see e.g. [147]), but not necessarily vice versa.

To gaussianise each input variable, $F_g$ can be estimated using parametric models, e.g. Gaussian mixture models [41]; normal, log normal, Weibull, gamma [176]. But the clipping of the distribution tails observed in section 4.3 can be a problem for the estimation of the parameters. In addition, for Gaussian mixture models it is necessary to determine the number of Gaussians with model selection techniques. Thus, in practice it is more convenient to use a non-parametric estimator like the empirical CDF (ECDF), e.g. [26, 149].

In addition, the mean intensity level and variance characterise the underlying material in textures, and the mean coordinate value and its variance contains important information about the shape, so $F_Z$ in (4.30) will not be standardised, i.e. $F_{Z,i}$ is the distribution of a normal variable $Z_i \sim \mathcal{N}(\mu_i,\ \sigma_i)$ where $\mu_i,\ \sigma_i$ are estimated from $g_i$. The non standardised normal CDF can be formulated as

$$F_Z(z) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{z-\mu}{\sigma\sqrt{2}}\right)\right) \tag{4.31}$$

where erf is the error function

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z \exp(-t^2)dt \tag{4.32}$$

Using (4.31) in (4.30) the gaussianised texture is

$$\hat{g} = \sigma\sqrt{2}\,\operatorname{erf}^{-1}(2F_g(g) - 1) + \mu \tag{4.33}$$

However, it should be noted that $F_Z$ has infinite tails and that the tangent quickly tends to 0, while the ECDF of $F_g$ has finite tails. This means that in practice, only values within the domain and range of the ECDF should be gaussianised, and outside values should be scaled linearly to avoid small perturbations in the gaussianised data producing

large errors when the gaussianisation is inverted.

To assess the proposed gaussianisation method, it is necessary to define a Goodness of Fit (GoF) measure. Originally, Sirovich and Kirby [164] proposed a measure in the spirit of the Signal to Noise Ratio (SNR)

$$\sqrt{\frac{\|s - s'_k\|^2}{\|s\|^2}} \tag{4.34}$$

where $s$ is the shape vector and $s'_k$ is the approximation with $k$ eigenvectors as defined in section 4.6. This measure is not easily related to clinical values, though. Instead, it is more informative to use the mean distance $d_{\text{mean}}$ in pixels of Eq. (4.27), and the maximum distance $d_{\max}$

$$d_{\max} = \max_i \|X(i) - X'_k(i)\| \tag{4.35}$$

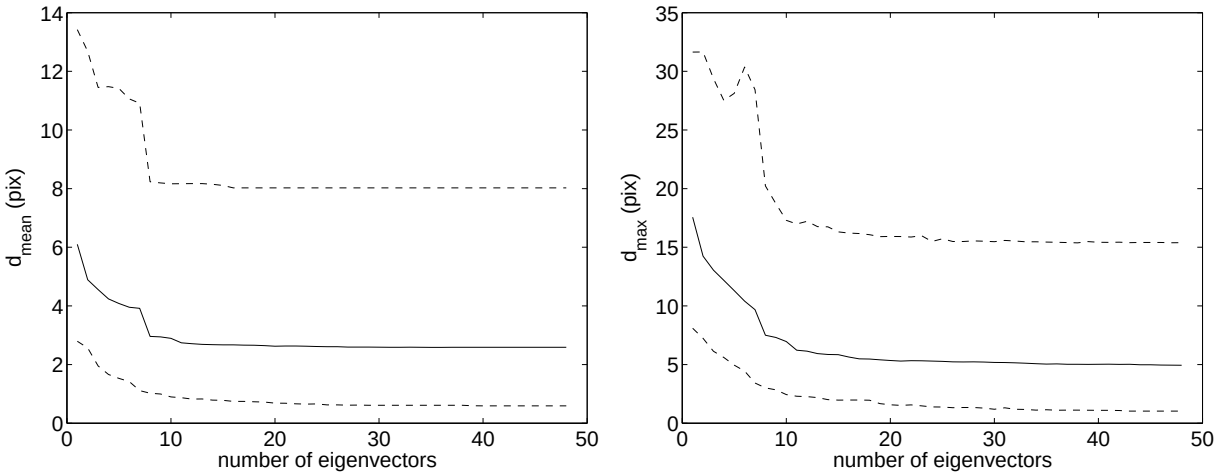
where $P$ is the number of landmarks, and $d_{\max}$ is the *Hausdorff distance*.

The effect of the proposed gaussianisation method on the model dimensionality is illustrated with an example using 21 studies in the 2C plane with different number of frames. The contours of 20 studies were used to generate a PCA shape space (100 variables, 335.2 frames on average per training data set). The data of another patient (100 variables, 16.8 frames on average per testing data set) was used to evaluate the approximation error. Fig. 4.5 suggests that the PCA model is more compact and generalises better when the data is gaussianised. The approximation error plateaus after 10 eigenvectors for non-gaussianised models[5], while it decreases consistently for gaussianised models.

The effect of the gaussianisation method on the multi-dimensional probability distribution can be partly assessed from its effect on the distribution of the data projected on the PCA eigenvectors. From the definition of the PDM in Eq. (4.3), the projections are the shape coefficients $b$

$$b = V^\top (s - \bar{s}) \tag{4.36}$$

[5]While in theory the eigenvalues are an orthonormal basis of $\mathbb{R}^{2P}$, and thus, the approximation error should be zero for $k = 50$, in practice the few eigenvectors with tiny positive or negative eigenvalues can produce numerical errors and are removed from the shape space.




(a) Before gaussianisation, $d_{\text{mean}}$.

(b) Before gaussianisation, $d_{\max}$.

(c) After gaussianisation, $d_{\text{mean}}$.

(d) After gaussianisation, $d_{\max}$.

Figure 4.5: Approximation error for increasing number of eigenvectors $k$. 2D endocardial model. Results obtained from 21 studies in 2C view, with a leave-one-out rota. Solid: Median. Dashed: 95% CI.

If the multi-dimensional probability distribution is Gaussian, the distribution of each projection $b_i$ will be Gaussian too. Deviation from Gaussianity for the data of the previous experiment was evaluated from the skewness and kurtosis (similarly to Appendix E). To get a better idea of the whole deviation from Gaussianity, the incremental skewness deviation $\Delta_{\text{sk}}$ for the $k$-th shape coefficient was computed as

$$\Delta_{\text{sk},k} = \sum_{i=1}^{k} |\text{sk}_i| \tag{4.37}$$

The absolute value in Eq. (4.37) ensures that positive and negative skewness values do not cancel out. If all distributions were Gaussian, $\text{sk} = 0$, and thus, $\Delta_{\text{sk},k} = 0, \forall k$. The graph in Fig. 4.6a shows that, after gaussianisation, the projected distributions are less skewed. The incremental kurtosis deviation $\Delta_{\text{kr}}$ for the $k$-th shape coefficient was computed as

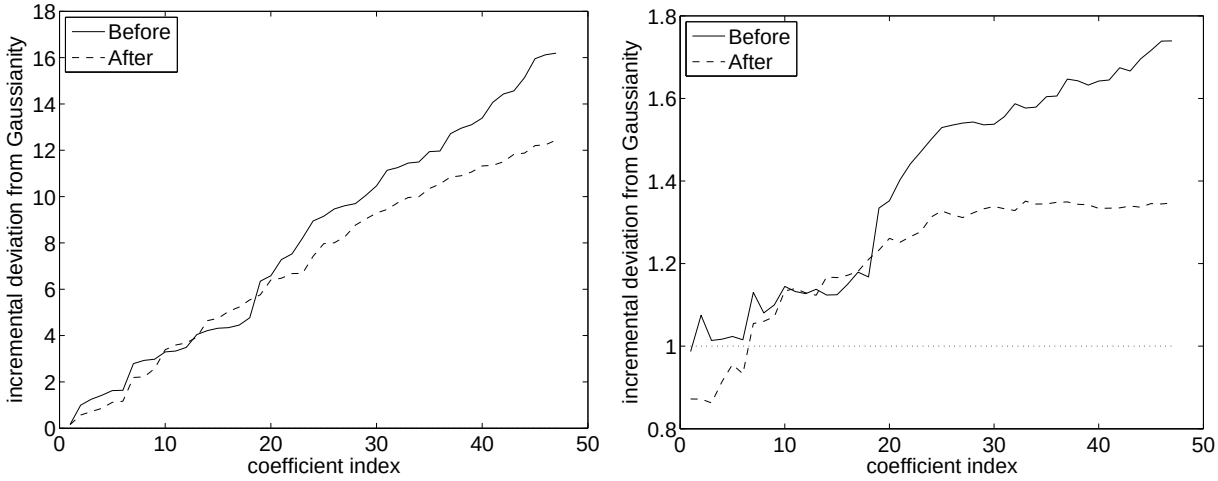
$$\Delta_{\text{kr},k} = \frac{1}{3k} \sum_{i=1}^{k} \text{kr}_i \tag{4.38}$$

If all distributions were Gaussian, $\text{kr}_i = 3$, and thus, $\Delta_{\text{kr},k} = 1, \forall k$. The graph in Fig. 4.6a shows that, after gaussianisation, the projected distributions have a kurtosis closer to the Gaussian's as more eigenvectors are considered.

Assuming that the behaviour of the projections in Fig. 4.6 reflects that of the whole data distribution, these results seem to support the hypothesis formulated above. Gaussianisation of the shape variables $s_i$ makes the underlying multi-dimensional distribution of the data more Gaussian, and this has a noticeable effect on the compacteness of the PCA model, as shown in Fig. 4.5.

## 4.8 Atlas-based deformable template models

Deformable template models (deformable models for short) were introduced in computer vision in the early 1970s, and received renewed interest after the introduction of Active Contour Models (snakes) by Kass et al. [88]. Geometric modelling in computer vision




(a) Skewness. (b) Kurtosis.

Figure 4.6: Gaussianisation effect on the PCA projections. The incremental deviations for (a) skewness and (b) kurtosis are the $\Delta_{\text{sk}}$ and $\Delta_{\text{kr}}$ defined in Eqs. (4.37) and (4.38). Deviations were computed before and after applying the gaussianisation method proposed in this section.

was a mature field by the late 1980s (see the review by Besl [19]). By the mid-1990s deformable models were already intensively used in medical imaging (see the review by McInerney and Terzopoulos [115]). In essence, deformable models combine a geometric model (or template) that outlines an object with a space of variation for that geometry and a mechanism to fit the model to the image data. In medical imaging, segmented objects are usually organs (e.g. lungs), parts of organs (e.g. the LV of the heart or the brain ventricles), vessels (e.g. coronary arteries), tissue, or structures (e.g. cells). A particular type of deformable model, called Active Appearance Model (AAM), and its extensions, are discussed in more detail in section 4.9.

The terminology for deformable models is a bit confusing in the literature. For example, the review by Jain et al. [85] considered deformable template models to include both free-form models (e.g. snakes) and parametric models (e.g. Active Shape Models), while Blake and Isard [21, section 2.2] did not consider snakes as deformable template models. Matthews and Baker [111] pointed out that things get even more confusing for certain types of deformable template models, e.g. linear shape and appearance models, especially

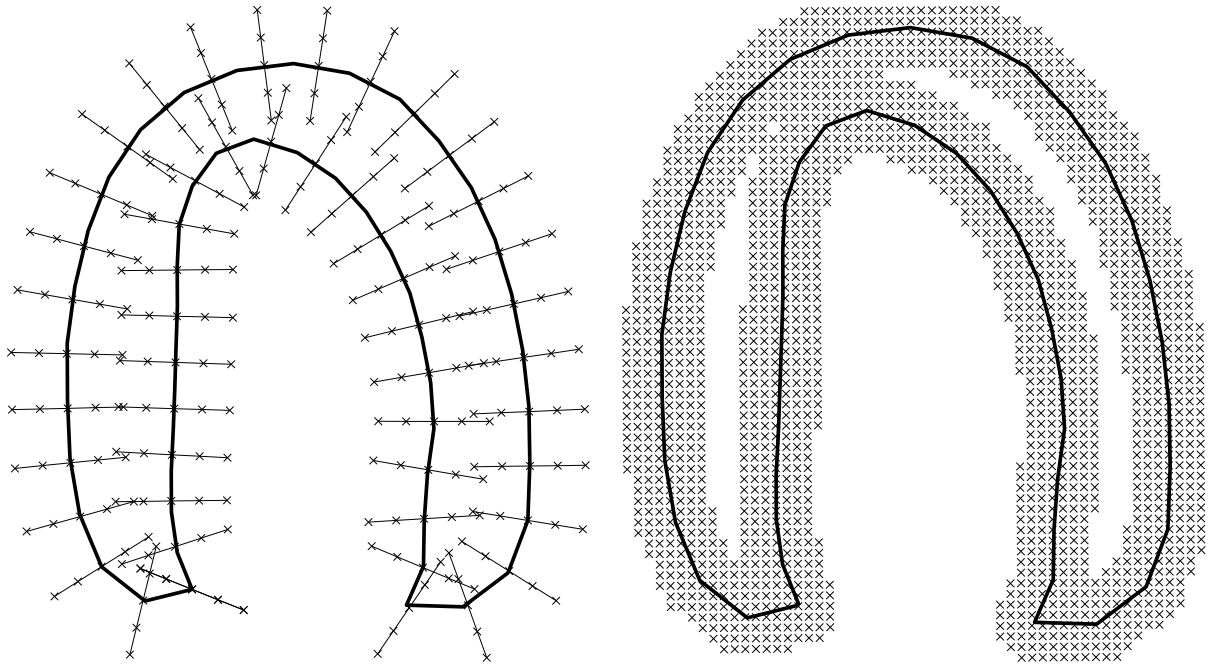
because the fitting algorithm is often included as part of the model. To emphasise the difference between model and fitting algorithm, the latter is studied in Ch. 6.

It is not computationally efficient to evaluate the fitting algorithm on the whole image in each iteration (e.g. [21, Ch. 5]), so a search region is defined within or around the template. It will be convenient to see the search region as a sampling mask or configuration of sampling points, as illustrated by Fig. 4.7. Typically, snakes or ASMs sample the search region along normals to the contour, as in Fig. 4.7a, but a more regular sampling mask can be obtained. For instance, in the original AAM formulation, sampling points were on a regular grid contained within the convex hull of the mean shape [48, 62]. Stegmann [168] noted that if the texture inside the mean shape is homogeneous, then the algorithm 'sometimes tends to lie inside the real object', and called it the *shrinking problem*. Tao et al. [175] independently noted a similar problem with the active contour segmentation formulation. The reason for the shrinking problem is that any patch of texture inside the object will have roughly the same texture as the whole object and this creates many local minima. To overcome this difficulty, Stegmann [168] proposed the Border AAM, which is like a normal AAM but where intensity values are sampled only around the shape, as in Fig. 4.7b. Bosch et al. [26] combined both types of masks, sampling both the inside of the shape and a band around it, but I find more convenient to follow [168].

Depending on the sampling space, deformable template models can be split into two categories: Feature-based deformable templates and template matching.

Feature-based deformable templates pre-process the image with a 'feature' detector, e.g. an edge-detector (e.g. [21]) and use the output to look for saliency in the image, without previous knowledge of the general appearance of the object.

What confusingly enough is called *template matching* or *template tracking* [107, 108, 112], is a deformable model extended with an image of the object we are looking for. For example, a picture of a car is provided, and it has to be tracked in each frame of a video sequence. But to achieve automatic segmentation of the myocardium, it has to be assumed that an initial segmentation of the muscle is not available, thus template matching is not



(a) Active Shape Models (snakes). (b) Border Active Appearance Models.

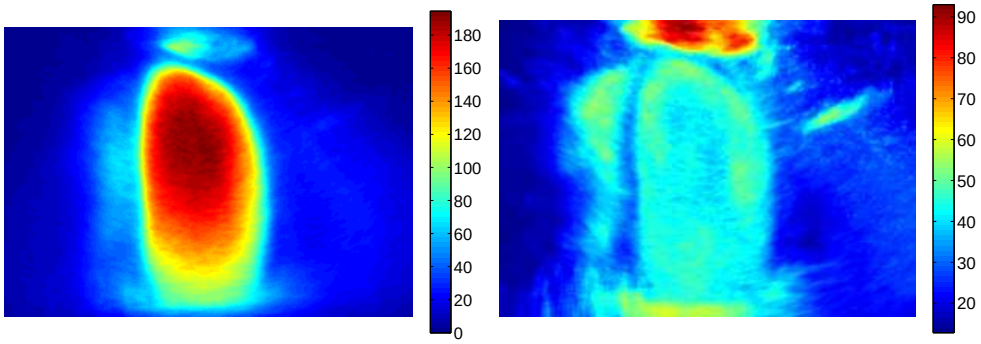
Figure 4.7: Comparison of search region between Active Shape Models and Border Active Appearance Models for a 2C plane. Each '×' is a point sampled in the segmentation algorithm.

possible. Instead, prior information about the target can be provided using a *probabilistic atlas*[6] deformable model. In this approach, an *atlas* or average image is used instead of an actual segmented picture of the target, as the former is a good compromise in terms of prior knowledge when the general anatomy of the organ is known. This is the atlas-based deformable model segmentation method presented in detail in Ch. 6. In the probabilistic atlas, intensity or texture statistics are computed for each sampling point in the mask. This removes the need for materials to have uniform properties all across the image, as statistics are computed locally. For instance, myocardial tissue near the apex and near the base can have different intensity distributions in the atlas. Probabilistic atlases were developed mostly for other modalities, especially 3D brain imaging (see the review by Mazziotta et al. [113]); in the last few years, atlas-based segmentation has been applied, under that name, to cardiac Magnetic Resonance (MR) [100, 104–106]. Atlas-

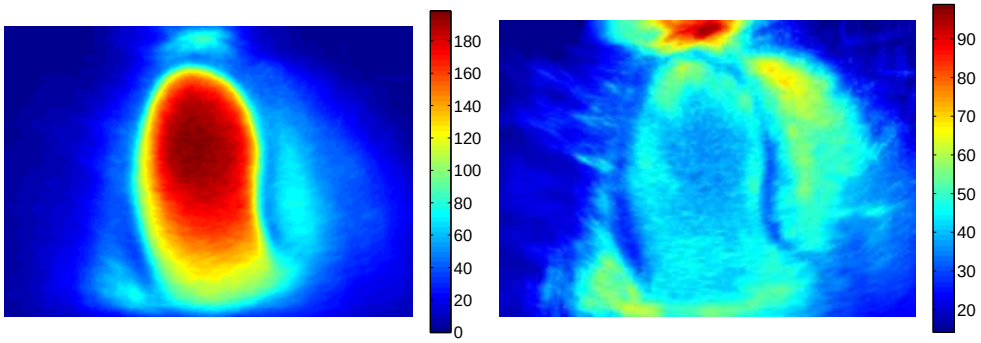
[6]Matthews et al. [112] called *template* to the intensity values. I use *template* for the geometry, and *atlas* for the intensity values.

based segmentation has been applied to fundamental imaging echocardiography too, but under the guise of AAMs. To my best knowledge, no atlas for contrast echocardiography has been published to date. To fill this gap, Fig. 4.8 provides probabilistic atlases for 2C, 3C, 4C and SAX planes. Atlases were computed as in section 4.3, for every pixel in the image, computing the mean and standard deviation values from each histogram. The results are displayed in Fig. 4.8.

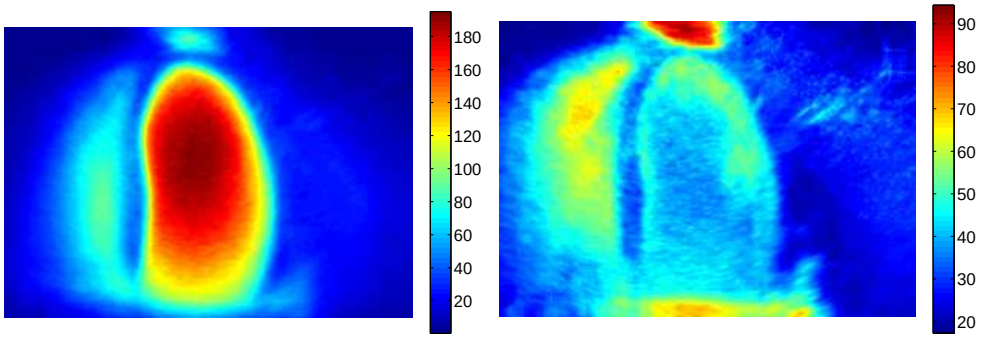
In section 3.3.4, it was noted that endocardial motion is easier to assess than wall thickening. Likewise, most models for echocardiography have been aimed at endocardial border identification (see the reviews by Hammoude [73] and Noble and Boukerroui [131]). It is apparent that the atlases in Fig. 4.8 display good contrast between the blood pool and myocardium, while the external boundary is less conspicuous at best, and unseeable in large regions of the image. These results are in agreement with the conclusions of Ch. 3, that suggested that external boundary segmentation and, thus, wall thickening evaluation, may not be possible using Power Modulation contrast echocardiography. They also cast a shadow of doubt on the reliability of expert hand traced external contours.



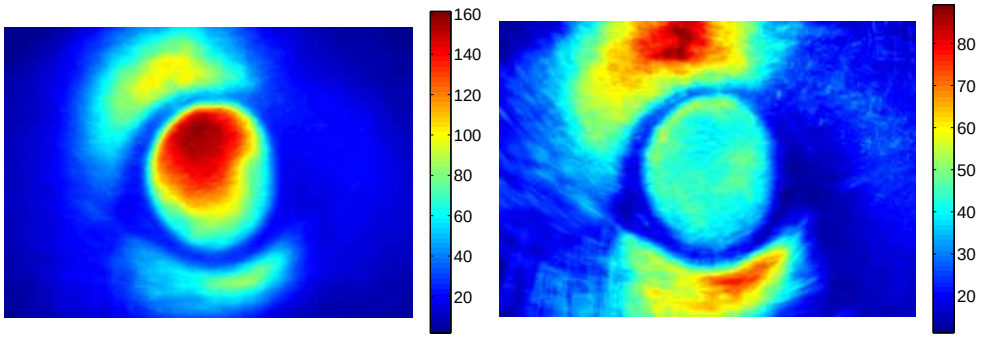

(a) 2C (mean). (b) 2C (std).




(c) 3C (mean). (d) 3C (std).




(e) 4C (mean). (f) 4C (std).




(g) SAX view (mean). (h) SAX view (std).

Figure 4.8: Probabilistic atlases for Power Modulation contrast echocardiography studies. Frames were scaled by a factor 0.75. Grey levels were represented using a colour map for better visual assessment. Atlases computed as in section 4.3, for every pixel in the image.

## 4.9 Active Appearance Models

Active Appearance Models (AAMs) are atlas-based deformable models where both the geometric template and a probabilistic atlas of mean intensities are allowed to vary within a PCA space. They were developed mostly as an independent type of algorithms, although Lorenzo-Valdés et al. [105] and Lapp et al. [100] referred to them as atlas-based deformation models. Deformable models can be forward or inverse depending on the formulation of the least-squares problem they solve. This will be explained in depth in section 6.1, but for the sake of comprehensiveness of this section, it is worth noting that AAMs have been formulated since their inception as inverse algorithms: 'To build a statistical model of the grey-level appearance we warp each example image so that its control points match the mean shape (using a triangulation algorithm)' [48]. In this section, I provide a historical overview of AAM development, as well as details about the model itself. In depth discussion of specific aspects of AAMs is presented in the remaining sections of this chapter.

AAMs are an extension of Appearance Models (AMs) [44, 60, 61, 97–99]. Cootes and Taylor [44] proposed statistical models of shape and texture to link texture sampling points with shape point coordinates, and originally used a Genetic Algorithm for the optimisation process. AAMs were formally defined when Edwards et al. [62] and Cootes et al. [48] proposed 1) sampling the texture from a reference frame instead of the shape configurations, following Craw and Cameron [51], and 2) using iterative model parameter refinement with a prediction matrix computed with multivariate linear regression (MLR) [168] for the fitting algorithm. Cootes et al. [50] replaced the MLR method by a more efficient ad hoc algorithm. It should be noted that although it was not acknowledged in the early AAM literature, 1) is in fact the formulation of an atlas-based segmentation method, as advanced above. Details for 2) were not available in the original publication, but were later provided by Stegmann [170]. For a comprehensive presentation of the classic formulation of AAMs, see Stegmann [168–170, 172]. Matthews and Baker [111] superseded the classic formulation noting that the assumptions of the ad hoc optimisation

algorithm in 2) were incorrect, and proposed an efficient gradient descent alternative that requires shape and texture to be modelled independently. AAMs have been extended from 2D to 2D+t [26], pseudo-3D [14], 3D [122] and 3D+t [171]. Current temporal extensions, and a new approach (also published in Casero and Noble [36]), are discussed in Ch. 5. A summary of deformable models for some key publications relevant to this thesis is presented at the end of this chapter in p. 109, Table 4.3. For pointers to algorithms similar to AAMs and other variations, see Matthews and Baker [111] and Cootes and Kittipanya-ngam [42].

In the AAM literature, *texture* refers to grey level intensity, i.e. only the mean value of the probabilistic atlas is used. Texture vectors $g$ are created by sampling the image with the sampling mask, e.g. Fig. 4.7b, and concatenating the intensity values into a column vector

$$g = [g_1,\ \ldots,\ g_L]^\top \tag{4.39}$$

Procrustes Analysis removes variability from the geometry data $s$ that should not be modelled by the shape space. Similarly, intensity normalisation removes variability from the intensity data that should not be modelled by the texture space. Methods for intensity normalisation of $g$ are discussed in section 4.10.

The main idea behind AAMs is to combine shape and texture variables into an *appearance* vector $a = [s^\top,\ rg^\top]^\top$, where $r$ is an scaling factor. The appearance vector approach will be discussed in section 4.11. The scaling factor will be discussed in section 4.13.

## 4.10 Intervolume intensity normalisation

Cootes et al. [48] proposed minimising the effect of interframe changes in global illumination to improve the performance of AAMs. In ultrasound data, the equivalent of global illumination is due to offset $u_1$ and amplification $u_2$ settings in the machine that affect all pixels $g_i$ in the same way, namely $g_i u_1 + u_2$ (related to TGC and attenuation). Some authors have considered the problem of compensating the attenuation field on the image,

e.g. [27, 80, 189], but as intraframe normalisation rather than interframe normalisation. To compensate for global illumination, Cootes et al. [48] iteratively normalised the intensity vectors $g$ so that the elements of $\bar{g}$ had mean 0 and variance 1. Later publications used the same scheme, e.g. [26, 170]. Similarly, Cootes and Taylor [46] proposed to standardise intensities

$$\hat{g}_i = \frac{g_i - \bar{g}}{\mathrm{std}(\bar{g})} \tag{4.40}$$

so that

$$\mathrm{Var}(\bar{g}) = 1 \tag{4.41}$$

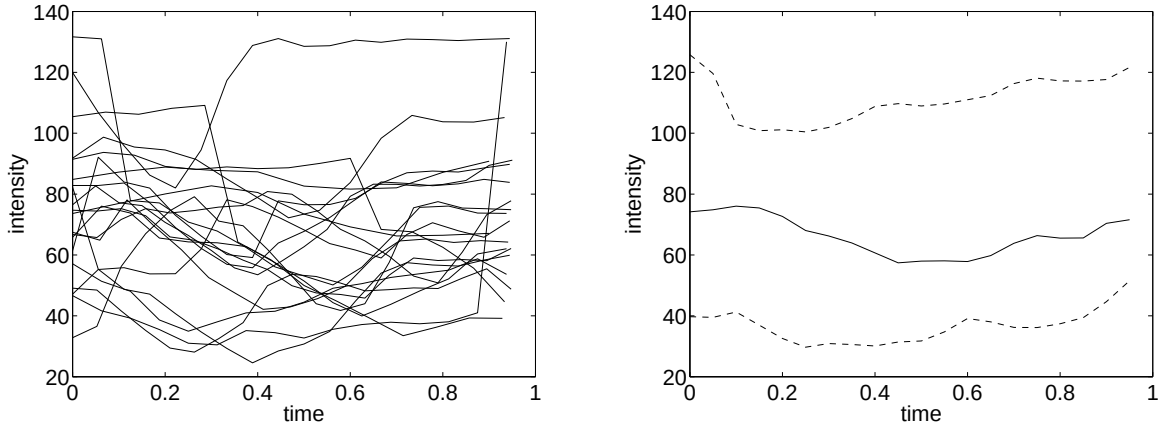
It should be noted that these methods remove important information, though. In the training phase, all intensity vectors correspond to the same anatomical region, so the normalisation is appropriate. However, when the fitting algorithm is running, the above range normalisation removes the mean and standard deviation information from the sampled intensities, i.e. both a sampling of the blood pool or of the outside of the heart will be normalised to similar values. Global illumination normalisation for all frames is problematic too, as end diastole frames have more blood pixels and, thus, are brighter on average. To illustrate this problem, Fig. 4.9a shows the mean intensity curves for the ultrasound window[7] of 20 patients in 2C view. The curves were interpolated to 20 frames. The corresponding median intensity and 90% Confidence Intervals (CI) are displayed in Fig. 4.9b. The median intensity value has a variation of 76.0 - 57.4 = 18.6 grey levels.

In this thesis I compute intensity normalisation on the whole spatio-temporal volume instead of frame by frame, and from the whole ultrasound window instead of the sampling area only. The normalisation parameters used in this thesis are the intensity mean and standard deviation computed for the whole volume, $\mu_{vol,i}$, $\sigma_{vol,i}$. Sampled intensity values

[7] By whole ultrasound area I mean the fan-shaped area with ultrasound signal in the ultrasound frame, as opposed to the black edges where some data like the patient's name is displayed.

$g_i$ are normalised to match the average parameters of the training set, $\bar{\mu}_{vol},\ \bar{\sigma}_{vol}$

$$\hat{g}_i = (g_i - \mu_{vol,i})\frac{\bar{\sigma}_{vol}}{\sigma_{vol,i}} + \bar{\mu}_{vol} \tag{4.42}$$




(a) Mean intensity of each frame for 21 patients.

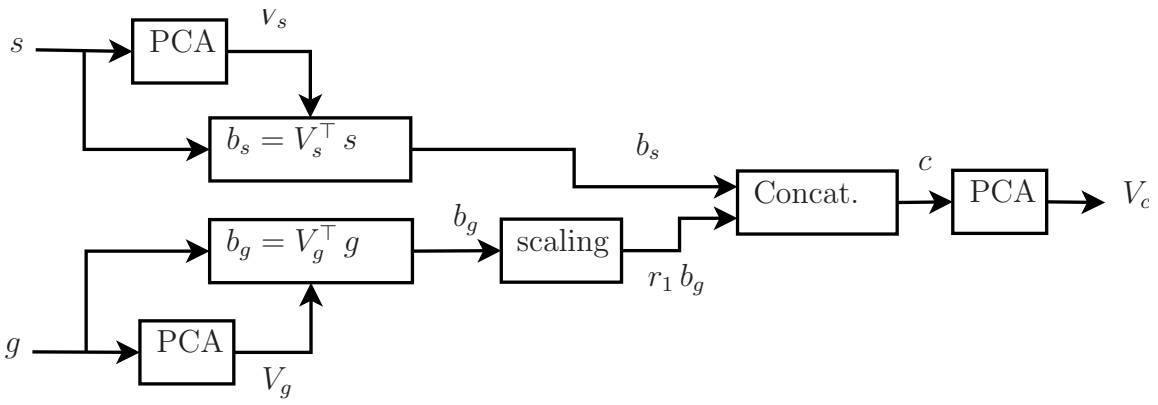
(b) Median (solid) and 90%-CI (dashed) of interpolated mean intensities across all patients.

Figure 4.9: Global intensity of contrast ultrasound 2D+t volumes. 20 baseline patients, 2C view, pixels within ultrasound window only.

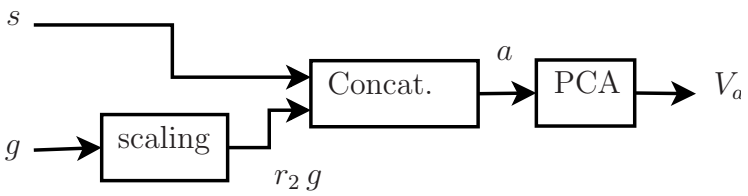
## 4.11 Combining different types of variables

A crucial aspect of AAMs is the combination of different types of variables into the model, traditionally shape coordinates and pixel intensities to produce an *appearance* vector. Edwards et al. [60] proposed combining variables as illustrated in the diagram of Fig. 4.10a; PCA models are computed independently for shapes and intensities. Then the training vectors are projected onto their respective models, intensity coefficients $b_g$ weighted and concatenated to shape coefficients $b_s$, and a new PCA computed. There are two objections to this scheme, that are discussed in this section.

The first objection is whether the combination of PCA models into another PCA model is necessary. Stegmann [168] observed (without proof) 'that another feasible method to obtain the combined model is to concatenate both shape points and texture samples into




(a) Appearance as a combination of PCA models from Edwards et al. [60].




(b) Appearance as a combination of variables, as noted but not used by Stegmann [168].

Figure 4.10: Approaches to compute combined PCA models (for appearance in this case).

one observation vector from the start and then perform PCA on the correlation matrix of these observation. [...] We regard the reason for two separate PCAs as being partly historical'. Despite his observations, Stegmann used Edwards' scheme himself [168, 170], and this is the method consistently used in the literature, even in recent publications, e.g. [74, 103]. Stegmann's alternative follows the diagram in Fig. 4.10b. This method will be used in Ch. 5 to propose a novel spatio-temporal model, so in the rest of this section I show that both approaches produce the same result when all eigenvectors are considered[8].

Let $s$, $g$ be the shape and texture vectors, and $r$ a scaling factor. Computing PCA

[8]A formal proof for the case when the shape and texture spaces are truncated is beyond the scope of this thesis. This notwithstanding, it should be noted that truncating the eigenvector spaces crops the data coordinates, but does not change the main directions of variability. Thus, truncating $V_g$, for example, should be equivalent to reducing $r$, as the corresponding subspace of texture would contribute less variance to the combined model. Testing this hypothesis is left as a future line of work.

models for $s$ and $g$ independently, and concatenating the results, leads to the linear system

$$\begin{bmatrix} s \\ rg \end{bmatrix} = \begin{bmatrix} \bar{s} \\ r\bar{g} \end{bmatrix} + \begin{bmatrix} V_s\ b_s \\ V_g\ rb_g \end{bmatrix} \tag{4.43}$$

where $b_s,\ b_g$ are coefficient vectors. Computing a joint PCA model $V_c$ on the combined coefficient vectors, the following equivalent model is obtained (see e.g. [48])

$$\begin{bmatrix} s \\ rg \end{bmatrix} = \begin{bmatrix} \bar{s} \\ r\bar{g} \end{bmatrix} + \begin{bmatrix} V_s\ V_{c,s} \\ V_g\ V_{c,g} \end{bmatrix} b_c \tag{4.44}$$

where $b_c$ is another coefficient vector. Using simple block matrix operations Eq. (4.44) is equivalent to

$$\begin{bmatrix} s \\ rg \end{bmatrix} = \begin{bmatrix} \bar{s} \\ r\bar{g} \end{bmatrix} + \begin{bmatrix} V_s & 0 \\ 0 & V_g \end{bmatrix} V_c b_c \tag{4.45}$$

As $V_s,\ V_g$ are orthonormal matrices, then $\begin{bmatrix} V_s & 0 \\ 0 & V_g \end{bmatrix}$ is orthonormal too. That is, the 2 independent PCAs are equivalent to a rotation or rotoinversion in appearance space. In general, this orthonormal transformation is not optimal in the Hotelling [77] sense of variance maximisation for the appearance vectors. To find the optimum in the variance maximisation sense, another rotation or rotoinversion is needed, and that is the role of $V_c$. This shows that computing $V_s,\ V_g$ is an unnecessary step, because the multiplication of two orthonormal matrices is another orthonormal matrix, and thus $\begin{bmatrix} V_s & 0 \\ 0 & V_g \end{bmatrix} V_c$ is the unique solution in the variance maximisation sense for the appearance vectors. Instead, the optimal rotation or rotoinversion can be computed in one step as the PCA solution

to the appearance vector $a = [s^\top,\ rg^\top]^\top$ data set, i.e.

$$\begin{bmatrix} s \\ rg \end{bmatrix} = \begin{bmatrix} \bar{s} \\ r\bar{g} \end{bmatrix} + V_a b_a \tag{4.46}$$

where the eigenvectors in $V_a$ are the same as in $\begin{bmatrix} V_s & 0 \\ 0 & V_g \end{bmatrix} V_c$ except perhaps for the sign.

The second objection is more profound, as it argues against the convenience of the combined model itself. Matthews and Baker [111] contended that the combined model precludes orthonormal shape and texture vectors, and that it makes the fitting algorithm less efficient because more parameters are updated in each iteration of the algorithm. This is the approach followed in this thesis.

## 4.12 Covariance vs. correlation matrix

Section 4.5 presented the PCA shape model as the solution to an eigenproblem defined by the covariance matrix. The possibility of using the correlation matrix instead of the covariance matrix has been acknowledged in the literature too, e.g. [165, 168]. The correlation matrix is equal to the covariance matrix once the variables have been standardised. If all variables have the same variance (as it is the case with standardised variables), then PCA looks for high values of the Pearson correlation coefficient $\rho$ [137][9], that indicate linear relationships between variables

$$\rho_{ij} = \frac{\sum_{k=1}^{M}(s_i(k) - \bar{s}_i)(s_j(k) - \bar{s}_j)}{(M-1)\sigma_{s_i}\sigma_{s_j}}, \quad -1 \leq \rho_{ij} \leq 1 \tag{4.47}$$

If standardised variables are weighted up, they get larger loadings on the first eigenvector, i.e. they become more important in the model. When the weights are the standard

[9] Pearson called it the Galton function or coefficient of correlation, denoted by $r$; it is generally accepted that it was Galton the first to propose a measure of correlation in that sense, i.e. [68], although it was Pearson who gave it mathematical rigour.

deviation of each variable, then the usual PCA on the covariance matrix is obtained. In other words, PCA takes into account both the linearity of relationships between variables and their variance.

Computing the correlation matrix is useful to visualise relationships between variables. As an example, Fig. 4.11 displays the value of the correlation matrix for the gaussianised shape coordinates of 21 studies in 2C. The matrix suggests very strong positive linear relationships between coordinates of blocks of 10 consecutive points, which corresponds to points on the same side of the wall boundaries. These results have a physiological interpretation, as displacements of nearby points are expected to be proportional. There is also a strong negative linear relationship between endocardial points on opposite sides of the endocardium. The physiological explanation is that when the cavity is expanding or contracting, both sides of the endocardium have to move in opposite directions. However, computing the shape model from the correlation matrix is not convenient, as information about the relative mobility of landmarks is lost.

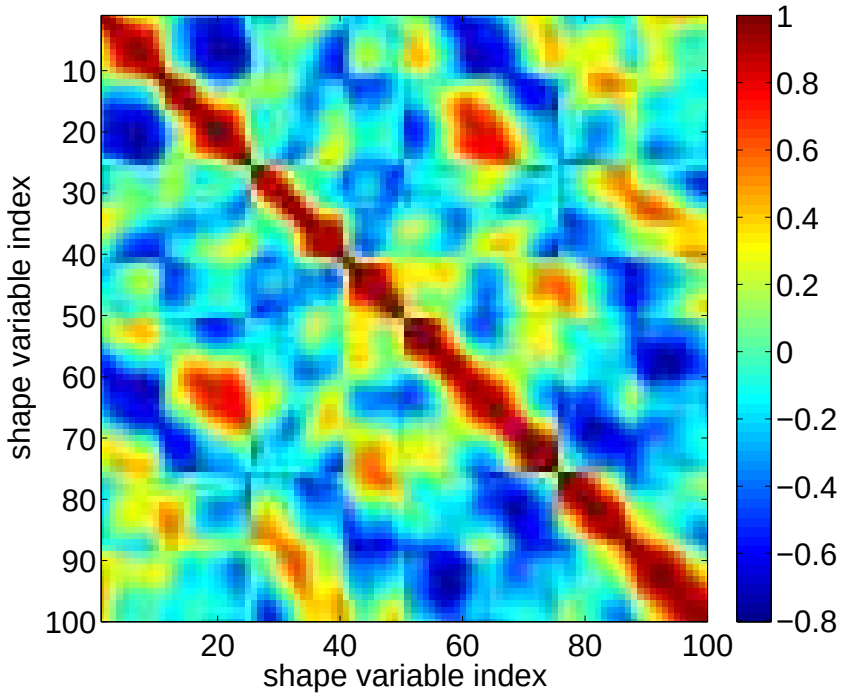



Figure 4.11: Correlation matrix for the gaussianised shapes of 21 studies in 2C. The 50 first variables correspond to $u$-coordinates, and the last 50 to $v$-coordinates. Within each group, 25 are endocardial and 25 epicardial landmarks, all clockwise.

## 4.13 Scaling factor for combination of variables

Once the role of correlation and variance is understood, it is easier to tackle the problem of computing the scaling factor $r$. Apart from linearising the relationships between variables, the other possible way to influence the model is by scaling some variables with respect to the others, so that they contribute more or less variance to be explained by the eigenvectors. This is an analogous situation to a zero sum game, in the sense that if some variables are scaled up, the rest are implicitly scaled down. Variable scaling was proposed in early stages of the AAM [48] to account for the 'difference in units between the shape and grey models'. To compute $r$, each shape coefficient $b_s$ in each training sample was computed and displaced $\Delta b_s$. Then the corresponding change $\Delta g$ was measured, and the $i$-th shape coefficient scaled by

$$r_i = \sqrt{\frac{1}{M}\sum_{j=1}^{M}\left(\frac{\Delta g}{\Delta b_{s,j}(i)}\right)^2} \tag{4.48}$$

Later publications have used the same $r$ [49, 100, 121, 122, 171] or no inter-variable scaling [26]. Cootes and Taylor [46] proposed as a simpler alternative

$$r = \sqrt{\frac{\mathrm{Var}(g)}{\mathrm{Var}(s)}} \tag{4.49}$$

where all $s$ and $g$ variables are aggregated together. They also stated 'In practise the synthesis and search algorithms are relatively insensitive to the choice of [the weighting factors]'. I think that as per the discussion above in this section, the scaling factor $r$ has a central influence in the model, that steers it for a better representation of a group of variables instead of others.

With Eq. (4.48), shape variables can get different scaling factors, which changes the information about the importance of each shape variable relative to each other. Eq. (4.49) prevents this from happening, but it does not take into account the number of variables

in each group. The more variables in a group, the more variance the group contributes for the model to explain. This is especially noticeable when the number of variables is one or more orders of magnitude different, as is the case with shape and texture vectors.

In this thesis, I propose that when variables are combined in a PCA model, then each group of interest should contribute the same variance to the model. For example, suppose that we want to combine shape $s$, texture $g$ and time $t$ variables, then texture and time should be scaled by $r_g,\ r_t$ respectively

$$r_g = \sqrt{\frac{\sum_{i=1}^{n_s} \mathrm{Var}(s_i)}{\sum_{i=1}^{n_g} \mathrm{Var}(g_i)}} \tag{4.50a}$$

$$r_t = \sqrt{\frac{\sum_{i=1}^{n_s} \mathrm{Var}(s_i)}{\sum_{i=1}^{n_t} \mathrm{Var}(t_i)}} \tag{4.50b}$$

Considering that the eigenvalues provide the variance in each eigenvector direction, this is equivalent to the approach proposed by Stegmann et al. [172]

$$r_s = \sqrt{\frac{\sum_{i=1}^{n_g} \lambda_{g,i}}{\sum_{i=1}^{n_s} \lambda_{s,i}}} \tag{4.51}$$

## 4.14 Comparison between AAM and atlas-based segmentation

AAMs were proposed and have been used in the literature as segmentation algorithms for over a decade. Validation results available from the literature have focused on the following issues: convergence of the algorithm after perturbation of the ground truth pose [48, 50, 111, 171], ground truth shape parameters [111]; convergence of the algorithm to acceptable solution [26, 74, 122]; effect of image scale on convergence [48, 169]; error from the algorithm solution to ground truth: landmark/curve/surface distance [42, 48–50, 74, 100, 121, 122, 169], endocardial area [26, 121, 122], endocardial volume [74, 100, 122, 171], epicardial volume [74, 122, 171], EF [26, 74, 171], myocardial mass [122], relative

overlap error: area [15], volume [100]; texture/reconstruction/matching error in single AAM [42, 49, 50, 62] and in AAM vs. a hybrid AAM/ASM [121]; colour vs. grey-scale appearance models [169]; segmentation error vs. intra- and inter-human expert variability [26, 122]; temporal error of border positioning [122]; comparison of fitting algorithms [111].

But to my best knowledge, the dependence of segmentation results on the dimensionality of texture space has not been discussed. Only Stegmann [169] noticed a modest improvement in the landmark error when instead of 28 modes for texture, 9 modes were used, but attributed the results to overfitting of the $95\%V_T$ criterion. Atlas-based segmentation is equivalent to an AAM with 0 texture eigenvectors, so ultimately, considerations about the dimensionality of the texture space lead to the question 'Are AAMs better than atlas-based segmentation algorithms?'. Matthews et al. [112] argued that in a template tracking problem, the target picture needs to be updated along the sequence for better segmentation, and this is possibly the underlying motivation behind the texture space in AAMs too. To discuss this point, it will be convenient to reproduce below the *project-out* approach proposed by Matthews and Baker [111]. Let the AAM least-squares optimisation problem be formulated as

$$p_{\text{opt}},\ b_{g,\text{opt}} = \min_{p,\ b_g} \frac{1}{2} \sum_z \Big( T(z) + V_g(z) b_g - I(W(z;\ p)) \Big)^2 \tag{4.52}$$

The least-squares problem is explained in more detail in section 6.3. Briefly, $p$ are some segmentation (pose and image transformation) parameters, $V_g$ is the matrix of texture eigenvectors, $b_g$ are the texture space coefficients, $z$ are the template pixel coordinates, $T$ is the mean atlas, $I$ is the image, and $W$ is an image transformation and warp. Matthews and Baker [111] showed that Eq. (4.52) is equivalent to

$$\begin{aligned} p_{\text{opt}},\ b_{g,\text{opt}} = \min_{p,\ b_g} \frac{1}{2} \Big\{ & \| T(z) - I(W(z;\ p)) \|^2_{\text{span}(V_g)^\perp} \\ & + \| T(z) + V_g(z) b_g - I(W(z;\ p)) \|^2_{\text{span}(V_g)} \Big\} \end{aligned} \tag{4.53}$$

Hence, the AAM optimisation process can be run in two independent stages: **Stage 1)** the segmentation parameters $p_{\text{opt}}$ are found by projecting out the atlas and image onto $\text{span}(V_g)^{\perp}$ —the orthogonal complement of the texture space— and solving iteratively a standard atlas-based segmentation problem. **Stage 2)** Once the segmentation is solved, the atlas is projected onto the image to obtain $b_{g,\text{opt}}$ in one step. Matthews and Baker [111] confirmed experimentally that the performance of the project-out approach was the same as the classic AAM formulation (e.g. [45]), while being computationally more efficient.

In the rest of this section, I comment on implications of the project-out approach not discussed by Matthews and Baker [111]. The project-out approach indicates, in fact, a trade-off between updating the template and updating the atlas. At one end, $V_g = \emptyset$ corresponds to atlas-based segmentation. That is, the texture space has zero eigenvectors and the atlas is fixed

$$g(z) = T(z) \tag{4.54}$$

Thus, the template is updated in Stage 1, but the atlas does not change, as Stage 2 does not take place. At the other end, $V_g$ is an orthonormal basis. That is, $V_g$ has all the eigenvectors to be square and full rank, and the atlas has absolute flexibility to change and match any image.

$$g(z) = T(z) + V_g(z)b_g \tag{4.55}$$

Hence, the template is not updated, because Stage 1 does not need to take place, as in Stage 2, the atlas changes to exactly match the image under the initial sampling mask. In all intermediate cases ($V_g \neq \emptyset$ and $V_g$ does not have full column rank), Stage 1 updates the template removing the texture model components from the error image, and Stage 2 is not directly relevant for the quality of the segmentation result. (Stage 2 improves the intensity matching error, not the segmentation error.)

To assess the effect of adding texture eigenvectors to an atlas-based segmentation algorithm, first I segmented a study with a fixed atlas and the project-out approach implemented using Matlab's least-squares non linear optimisation function `lsqnonlin()`.

For simplicity, the algorithm optimised the similarity transformation parameters only. When the algorithm stopped, the segmentation error was measured with $d_{\text{mean}}$, between the optimised contour and the expert hand traced contour, as defined in section 4.6.

Then the texture model $V_g$ from 20 patients was computed. Segmentation was run on another patient (leave-one-out approach), but this time the atlas was allowed to vary with the first texture eigenvector from $V_g$. The segmentation error $d_{\text{mean}}$ was computed and written down. This process was repeated for the first 2 texture eigenvectors from $V_g$, then the first 3 texture eigenvectors from $V_g$, and so on until the whole $V_g$ matrix was used.

The experiment was repeated 21 times, each time leaving a different patient out to compute the texture model on all the frames in 2C, 3C, 4C and SAX planes (the frames were scaled by a factor 0.25 to a size of 69x90 pixel). All 4 planes showed similar results, so for clarity, only the graphs for 3C are presented here (the remainder graphs are presented in Appendix D).

Fig. 4.12a shows the median $d_{\text{mean}}$ error and the 95%-CI for all experiments. The results suggest that using a low number of eigenvectors makes the algorithm diverge in many cases. A zoom-in into the graph for the lower number of modes is provided by Fig. 4.12b. It is worth noting that the atlas-based approach (0 texture eigenvector) is always convergent, but adding just 1 or 2 texture eigenvectors significantly worsens the results. If failed segmentations are removed from the results, Figs. 4.12c and 4.12d suggest that for successful segmentations, atlas-based results are best, and they worsen slightly with increasing number of texture eigenvectors. A possible explanation is that the image error term has substantial Principal Components even in the lowest modes of texture space, and so the optimisation algorithm fails when that information is projected out. The area of divergent segmentations in the plot could be attributed to the projected out error leading the parameter updates in directions orthogonal to the correct ones. The posterior plateau area could be explained by misleading information being removed from the image error too. The horizontal lines in Figs. 4.12c and 4.12d outline the median

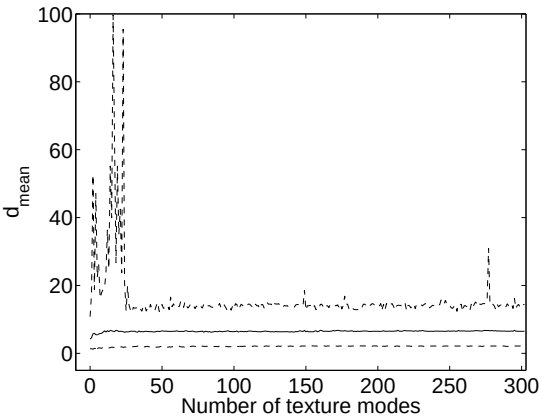
and 95%-CI for the segmentation error at iteration 0 of the algorithm, i.e. for the initial mean contour. The graphs suggest that AAMs achieve segmentation improvement in the plateau area; this would justify their successful use in past publications.

However, the experimental results and discussion above also suggest that simple atlas based segmentation (the graphs in Fig. 4.12 evaluated at 0 texture modes) actually outperforms AAM segmentation. This observation, together with the discussion in section 4.15 about the feasibility of computing texture models, justify that the segmentation problem in Ch. 6 is formulated using an atlas-based deformable template model, instead of an AAM.

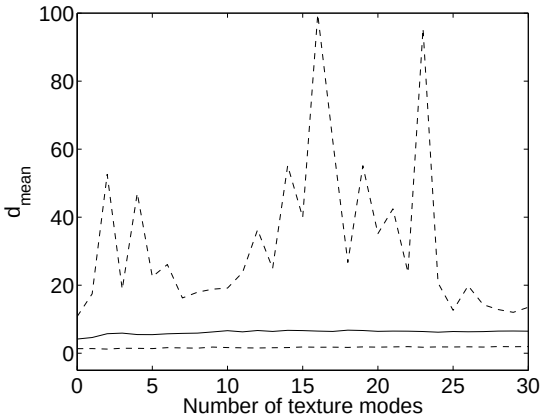
This notwithstanding, my results are limited to specific hard-to-segment ultrasound data and a similarity transformation. Further research is necessary with other implementations and data sets to better assess the performance of AAM compared to atlas-based algorithms. In addition, the similarity transformation is only meant as a first approximation, as shown by the median $d_{\text{mean}}$ in the order of 4 pixels. Optimisation of the shape parameters (which allows for a non-rigid deformation of the template) is left as future work. The formulation of the corresponding segmentation algorithm is presented in Appendix G as preliminary work.

## 4.15 Feasibility of computing PCA shape and texture models

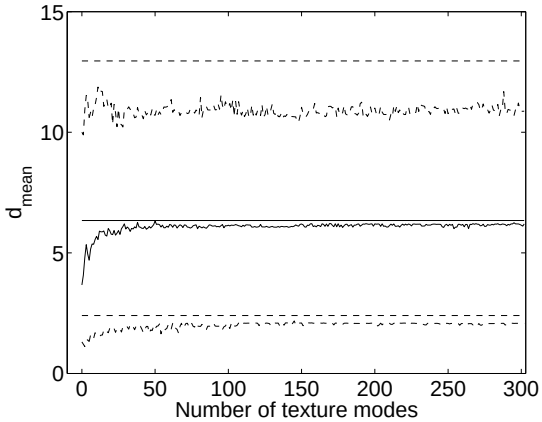
Discussion on the size of the training data set has been quite limited in Point Distribution Model (PDM) literature. This could be attributed to the usual presentation of Principal Component Analysis (PCA), as in section 4.5, where two methods to compute models are available, one for $n \leq M$ and one for $n > M$, where $n$ is the number of variables and $M$ the number of training samples. The latter case —the MDS approach— allows to compute a PCA space when the training data set is too small, but this does not mean that



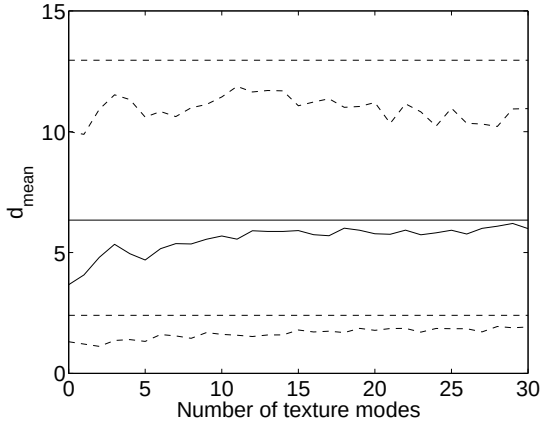

(a) Convergent and divergent segmentations.




(b) Convergent and divergent segmentations. Zoom into (a).




(c) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixel.




(d) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixel. Zoom into (c).

Figure 4.12: Segmentation error vs. dimensionality of texture space. Number of texture modes = number of texture eigenvectors from $V_g$. See main text for details. Mean distance error $d_{\text{mean}}$: Median (solid graph) and 95%-CI (dashed graphs). Leave-one-out 2D model with 21 patients, 3C view. The horizontal straight lines in c) and d) show the median and 95%-CI for the segmentation error at iteration 0.

the computed PCA space is representative of the data. Even though linear models suffer relatively little from the curse of dimensionality problem, Henry et al. [76] estimated that in order to start obtaining significant results with PCA, the number of training samples $M$ should be in principle

$$M > 30 + \frac{n+3}{2} \tag{4.56}$$

where $n$ is the number of variables. For example, for a typical apical plane and 25 endocardial landmarks, $n = 50$, and the data set should have $M \geq 57$ training vectors. For 50 myocardial landmarks, $M \geq 82$. Our contrast DSE database has e.g. $M = 352$ sample vectors in 2C from 21 patients, well above the minimum, although it has been pointed out that the temporal correlation between frames could reduce the effective size of the data set[10]. On the other hand, Mei et al. [116] found that $M = 300$ samples were sufficient to model 3D LV surfaces with $n = 1581$, while Eq. (4.56) suggests $M > 822$; this could indicate intra-landmark correlation that reduces the effective number of variables. Although further work is necessary to address those issues, it seems that our data set size is sufficient for 2D models.

The situation is worse for texture models. Table 4.2 presents typical sizes of the sampling mask in endocardial and myocardial models. Even if the images are resized by a factor 0.25, there are in the order of 2 to 5 times more texture variables to be modelled than training vectors. For images at full size, there are 30 to 90 times more variables than training vectors. The problem of having too large intensity vectors has been acknowledged in the literature before. Yang et al. [190] proposed a method called two-dimensional PCA (2DPCA) to reduce the dimensionality of vectors used to describe whole images, and preserve the neighbourhood structure of the pixels. However, it has been shown that 2DPCA is simply PCA applied to the concatenated rows of all training images [92, 187]. That is, dimensionality is reduced because each column is modelled with a single variable. While there are more samples and less variables, in ultrasound data this approach would

[10] Anonymous reviewer for Casero and Noble [36]. Personal communication.

| | 2C | 3C | 4C | SAX |
|---|---|---|---|---|
| Scale 1.00 | 13,885 | 13,180 | 13,973 | 11,055 |
| Scale 0.25 | 865 | 823 | 877 | 685 |

(a) Mask: 15 pixel distance to endocardium.

| | 2C | 3C | 4C | SAX |
|---|---|---|---|---|
| Scale 1.00 | 26,893 | 25,518 | 27,532 | 16,839 |
| Scale 0.25 | 1,680 | 1,598 | 1,728 | 1,678 |

(b) Mask: 15 pixel distance to endocardium + all pixels inside of cavity.

| | 2C | 3C | 4C | SAX |
|---|---|---|---|---|
| Scale 1.00 | 29,705 | 27,848 | 28,655 | 27,378 |
| Scale 0.25 | 1,842 | 1,734 | 1,796 | 1,717 |

(c) Mask: 15 pixel distance to myocardial contours.

| | 2C | 3C | 4C | SAX |
|---|---|---|---|---|
| Scale 1.00 | 30,556 | 27,981 | 29,065 | 29,238 |
| Scale 0.25 | 1,905 | 1,743 | 1,819 | 1,832 |

(d) Mask: 15 pixel distance to external contour + all pixels within myocardium + all pixels inside of cavity.

Table 4.2: Typical sampling mask sizes, in number of pixels, for 2D echocardiography models. 2 models are considered: Endocardial and myocardial. For each model, 2 sampling masks are considered: A border model with a search region of 15 pixels around the contours (see Fig. 4.7b), and a combined model that also includes all pixels within the convex hull of the contours. At scale 1.00, the image measures 256x360 pixel. At scale 0.25, the image measures 64x90 pixel.

force each input variable to represent different regions of tissue and blood pool.

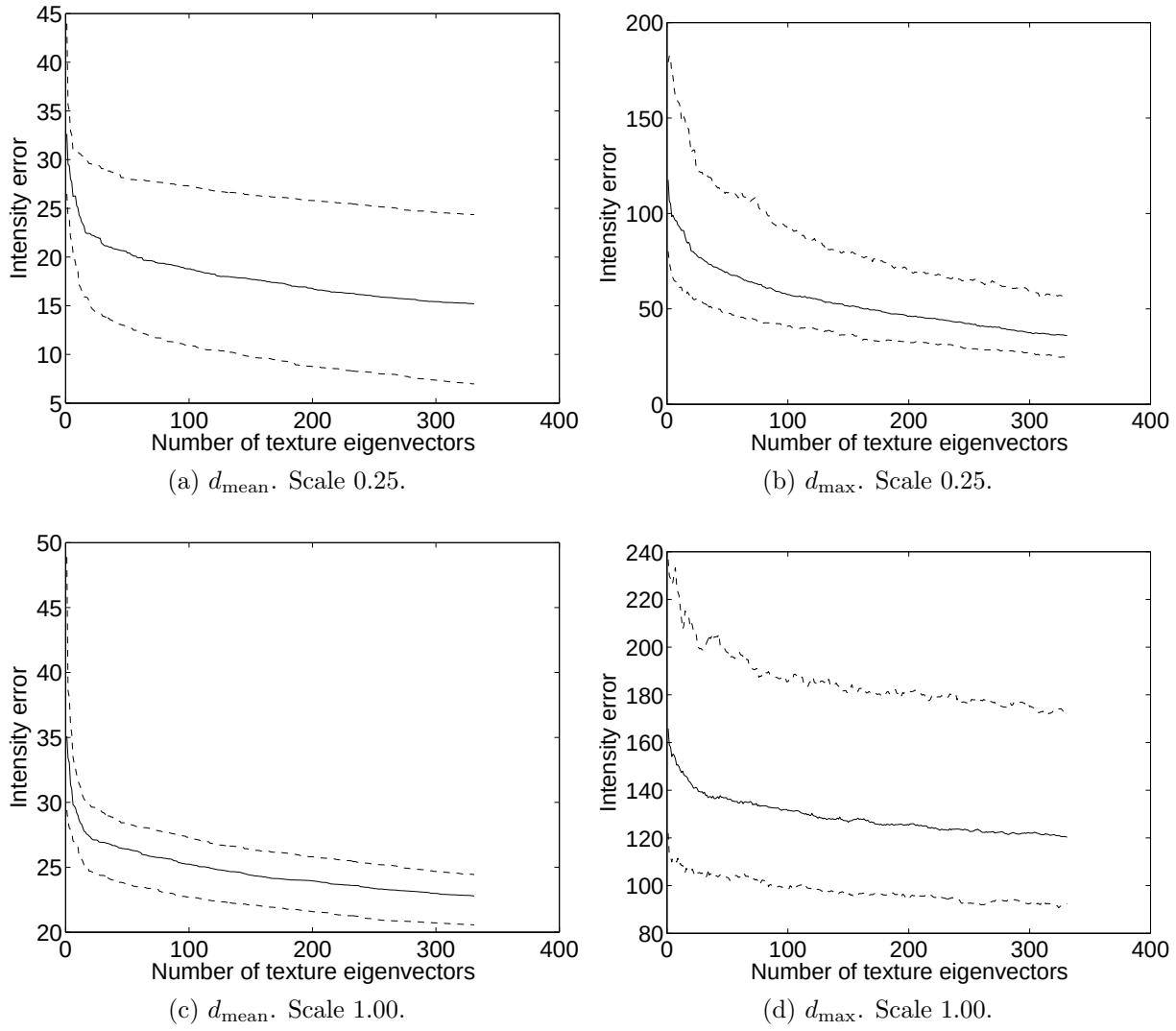
Thus, with the current technology PCA texture modelling requires much larger datasets than what we have. But segmenting a dataset by hand is a very tedious and time-consuming task, as reported in section 2.4. Acquiring this amount of data is not only a matter of time, but also finding a medical expert willing to spend months or years collecting it and doing segmentations. Frangi[11] suggested using an ultrasound simulator to generate synthetic examples to create a large enough data set. To my best knowledge, there is no simulator that can generate contrast echocardiographic data. But even if it was possible to find and hand trace, or simulate, a big enough data set, solving the eigen-problem on a very large matrix would still be an issue. In addition, section 4.3 showed that intensity data is highly non-Gaussian. Gaussianisation of shape data was tackled in section 4.7, but intensity data poses new challenges, as discussed in Appendix E.

To illustrate the problem created by the length of intensity vectors, the experiments of section 4.7 were repeated for intensity data. Similar results were obtained for all 4 planes, so for clarity, only the results for 2C are presented next. Figs. 4.13a and 4.13b show the approximation error $d_{\text{mean}}$ (pixel intensity has no units) decreasing with increasing number of eigenvectors for the image size scaled by a factor of 0.25, and a sampling mask within 15 pixel of the endocardium. As shown in Table 4.2, the intensity vector has approximately 865 variables. Eq. (4.56) requests 464 training vectors, but the leave-one-out data sets have around 290 frames only. Thus, even using all eigenvectors the median approximation error decreases to $d_{\text{mean}} = 15.2$, with median values of $d_{\text{max}} = 35.9$. Figs. 4.13c and 4.13d display the results for the full size image. The training data sets have the same size, but now the intensity vectors have lengths of approximately 13,885 variables. It is worth noting that the median $d_{\text{mean}}$ increases by 50%, from a minimum of 15.2 to 22.8. But the median $d_{\text{max}}$ increases by 231%, from 35.9 to 120.3. This can be explained by higher frequency intensity components in the full size image, and the inability of a model trained

[11] A.F. Frangi. Universitat Pompeu Fabra, Barcelona, Spain. Personal communication. I have been unable to find a published reference implementing this idea.

with too little data to fit them.

To sum up, the results in this section suggest that there are two reasons that make working with PCA texture models unfeasible in practice: the required size for the data set to train the model, and computation of the eigenanalysis on the correspondingly large covariance matrix.




(a) $d_{\text{mean}}$. Scale 0.25.

(b) $d_{\max}$. Scale 0.25.

(c) $d_{\text{mean}}$. Scale 1.00.

(d) $d_{\max}$. Scale 1.00.

Figure 4.13: Approximation error in the texture model. Computed from 21 patient data, leave-one-out, 2C plane. At scale 1.00, the image measures 256x360 pixel. At scale 0.25, the image measures 64x90 pixel. Sampling mask: 15 pixel distance to endocardium.

## 4.16 Summary and conclusions

In this chapter, I argued that segmentation of myocardial boundaries on the contrast echocardiographic cine loops should combine knowledge from 3 sources: texture, geometry and kinetics. This chapter and the following are devoted to the analysis of a framework that combines previous knowledge from those 3 sources of information to solve the segmentation problem. Texture and geometry were studied in this chapter, while kinetics are the main focus of Ch. 5.

Starting with texture, section 4.3 suggested that the distribution of intensity values in an anatomical location has a left-skewed distribution and clipped off tails, with a variance that increases with the median value. The former violates the assumption of Gaussianity for PCA, and the latter makes the segmentation algorithm depend more heavily on the brighter areas the image. This chapter argued against the convenience of using PCA texture models altogether, so gaussianisation was addressed in Appendix E. The study of whether a scaling of intensity values to balance residuals variances improves the segmentation results is proposed as future work.

Section 4.4 presented landmark models as a convenient representation of the geometry of myocardial contours, in particular using interpolating cubic splines control points as landmarks and pseudo-landmarks. This allows to discretise the problem and, thus, represent continuous curves as point configurations. The Least-Squares Fit Generalised Orthogonal Procrustes Analysis was corrected and used to map the LV geometry point configurations to shapes, i.e. point configurations without the location, size and rotational effects.

Section 4.5 presented PCA as a method to compute spaces of shape variability, and noted that the formulation that has been presented in the literature as a computational trick to ease the calculation of spaces for long vectors is, in fact, the formulation of Multidimensional Scaling (MDS), a dual of PCA. This observation will be used to propose a spatio-temporal model in Ch. 5. Section 4.6 introduced a long discussion on dimension-

ality of the shape model, a topic that has received little attention in the literature. In agreement with recent results from other authors, the variance criterion was found to be inadequate for dimensionality selection. A family of data-driven criteria, the Generalised Information Criterion (GIC), was presented and discussed. The results suggested that GIC may be a useful criterion to determine the optimal number of landmarks to represent the data, but at the same time overestimates the practical dimensionality of the shape model. I proposed an anatomical/functional criterion for dimensionality selection below 1.5 pixels. This criterion takes into account that myocardial thickness can typically be as small as 15 pixel in end diastole. Section 4.7 illustrated the poor performance of 2D shape models if the Gaussianity assumption is not met. It was hypothesised that, theoretical limitations notwithstanding, gaussianisation of each input variable using the empirical CDF (and being careful with the tails of the distribution) would gaussianise the underlying data distribution. Experimental results supported this hypothesis, as gaussianisation of shape variables substantially improved the compactness of the 2D shape model.

Section 4.8 gave an overview of atlas-based deformable template models as a framework that integrates a probabilistic atlas (texture) and a template (geometry) with a border search region for segmentation optimisation. For the first time, probabilistic atlases of the mean intensity and standard deviation for Power Modulation contrast echocardiography in the four principal planes were computed. These atlases will be used in the segmentation algorithm of Ch. 6. Further research could evaluate the performance of atlases based on the mean compared to other statistics, e.g. the median or mode.

A historical review and description of so-called Active Appearance Model (AAM) was presented in section 4.9. AAMs were developed as independent algorithms, and they have been shown to use suboptimal ad hoc optimisation methods based on incorrect assumptions. But in the last few years the literature has identified them as an extension to atlas-based deformable models, and their ad hoc optimisation methods have been replaced with standard least-squares gradient descent methods. The following sections studied particular aspects of AAMs in detail. Section 4.10 contended that intensity vectors should not

be standardised, to avoid losing fundamental information, and instead proposed intensity normalisation on the whole cine loop. Section 4.11 proved that computing PCA models on vectors with heterogeneous variables (e.g. shape and texture) produces the same result as the classic approach in AAMs of computing a PCA model for shape, a PCA model for texture, and another PCA model of the combined coefficients. It was also noted that other authors have argued against combined shape and texture models, because it impairs the fitting process. Nonetheless, combined models will be a useful tool for the integration kinetics in Ch. 5. Section 4.12 discussed the role of the correlation matrix in identifying linear relations between variables with a physiological meaning. For example, horizontal coordinates of endocardial landmarks on one side of an apical plane have a strong positive linear correlation amongst them, and a strong negative linear correlation with landmarks on the other side. The observation that linear correlations are weighted by variance estimates in the covariance matrix led to section 4.13, where a scaling factor was proposed with the aim of balancing the total variance contribution of heterogeneous variables in a combined model.

This chapter closed with two critiques of AAMs. First, section 4.14 noted that comparisons between the performance of AAMs and atlas-based segmentation have been missing in the literature. It has been implicitly assumed that AAMs should perform better than atlas-based methods, because a texture space adds flexibility to a fixed probabilistic atlas. My experimental results agreed with published results in that the AAM reduces the segmentation error of the initial contour when it converges, i.e. the AAM *works*. However, experiments and theoretical considerations based on the project-out formulation suggested that, in fact, important information for the segmentation process may be lost when using a PCA texture space. This may be the cause of AAM high rate of divergence, but even when convergence was achieved, the AAM was less accurate and diverged more often than atlas-based segmentation.

And second, section 4.15 discussed the feasibility of computing PCA shape and texture models in practice. While vectors for shape data seemed small enough to be modelled with

a 21 patient data set, texture vectors were 2 orders of magnitude longer at full scale, and 1 order of magnitude if the image size is reduced by a factor of 0.25. Typical approximation errors for the texture model remained substantial even when all eigenvectors were used. Interestingly, maximum approximation errors increased much more sharply that median errors as texture vectors got longer, suggesting that increased resolution in the image adds high frequency components that the model cannot explain.

To sum up, the results in the chapter justify formulating the segmentation problem as an atlas-based deformable template model, rather than an AAM, in Ch. 6. The reasons are better performance of the former, and unfeasibility of computing appropriate PCA texture models for the latter.

| | Kass et al. [88] | Chalana et al. [39] | Cootes and Taylor [43] |
|---|---|---|---|
| Name | Active Contour Model (snake) | Multiple Active Contour Model | Active Shape Model |
| Data preprocessing | None | None | Procrustes alignment of shapes |
| Data dimensionality | 2D/2D+t | 2D+t/3D | 2D |
| Geometry | 2D spline | Surface spline | Fine sampling of the shape |
| Model | Implicit in regularisation | id. | PCA |
| Shape constraints | 1st and 2nd derivative curve regulariser | id. | Truncated PCA model + Mahalanobis distance to training set |
| Sampling mask | Normals to spline | id. | id. |
| Image features | Edges | id. | id. |
| Warp | Displacement along normals to spline | id. | id. |
| Warp direction | $T \mapsto I$ | id. | id. |
| Optimisation type | Local minimum | id. | id. |
| Optimisation variables | Coordinates of spline fine sampling | id. | Pose, shape coefficients |
| Number of modelled variables | $\mathcal{O}(10^2)$-$\mathcal{O}(10^3)$ | id. | $\mathcal{O}(10^2)$ |
| Number of opt. variables | $\mathcal{O}(10^2)$-$\mathcal{O}(10^3)$ | id. | $\mathcal{O}(10)$ |
| Optimisation method | Iterative solution to Euler equations in matrix form | id. | Least squares solution of system of proposed displacements for all points at each iteration |

(a) Shape models.

| | Cootes et al. [48] | Cootes et al. [50] | Matthews and Baker [111] |
|---|---|---|---|
| Name | Active Appearance Model | id. | id. |
| Data preprocessing | Procrustes alignment of shapes | id. | id. |
| Data dimensionality | 2D | id. | id. |
| Geometry | Landmark and pseudo-landmark points | id. | id. |
| Model | Coupled PCA | id. | Decoupled PCA |
| Shape constraints | Truncated PCA model | id. | id. |
| Sampling mask | Uniform grid within mean shape | id. | id. |
| Image features | Normalised pixel intensities (loss of mean intensity information) | id. | Pixel intensities |
| Warp | Piece-wise affine? | Not specified | Piece-wise affine |
| Warp direction | $I \mapsto T$ | id. | id. |
| Optimisation type | Local minimum | id. | Local minimum (pose, shape), least squares (appearance) |
| Optimisation variables | Pose, appearance coefficients | Pose, global illumination, appearance coefficients | Pose, appearance coefficients |
| Number of modelled variables | $\mathcal{O}(10^4)$ | id. | 4 pose, $\approx$ 140 shape?, unspecified for appearance |
| Number of opt. variables | $\mathcal{O}(10^2)$ | id. | 4 pose, 3 shape, 9 appearance |
| Optimisation method | Multilinear regression to solve system of linear relations between intensity errors and model parameters | Ad hoc, solving a least squares system to make the intensity error equal to 0; numerical estimation of the Jacobian matrix | Gauss-Newton Inverse Compositional Algorithm w/o line search (pose, shape), least squares (appearance); analytical computation of Jacobian matrix |

(b) Appearance models.

Table 4.3: Schema of deformable models from a selection of publications, as referenced from section 4.9. id.: Same as in column to the left.

| | Bosch et al. [26] | Rueckert and Burger [148] |
|---|---|---|
| Name | Active Appearance Motion Model | TPS Geometrically Deformable Template (GDT) |
| Data preprocessing | Procrustes alignment of shapes, volume reslicing to same number of frames, Gaussian subsampling of images | none |
| Data dimensionality | 2D+t/3D | 2D+t/3D+t |
| Geometry | Landmark and pseudo-landmark points, concatenated frames | Landmark points |
| Model | Coupled PCA | Implicit in regularisation, solution in $i$-th frame initialises tracking in $i+1$-th frame |
| Shape constraints | Truncated PCA model | Implicit in minimisation of TPS bending energy |
| Sampling mask | Uniform grid within mean shape | Uniform grid on image |
| Image features | Pixel intensities, gaussianised as an aggregate (loss of mean intensity information) | Edges |
| Warp | Not specified | 2D/3D TPS |
| Warp direction | $I \mapsto T$ | $T \mapsto I$ |
| Optimisation type | Local minimum | Quasi global minimum (segmentation) + local minimum (tracking) |
| Optimisation variables | Affine pose, global illumination, appearance coefficients | Coordinates of landmarks, TPS bending energy |
| Number of modelled variables | Not specified; my guess is $\mathcal{O}(10^4)$-$\mathcal{O}(10^5)$ | 32-64 (2D), 512-1024 (3D) |
| Number of opt. variables | Not specified | 32-64 (2D), 512-1024 (3D) |
| Optimisation method | Multilinear regression to solve system of linear relations between intensity errors and model parameters | MAP Simulated annealing (segmentation) + MAP Iterated Conditional Modes (tracking) |

(c) Other.

Table 4.3: Schema of deformable models from a selection of publications (cont.).

# CHAPTER 5

# Temporal extension of cardiac PCA shape models

## 5.1 Background

Echocardiography depends strongly on temporal information for clinical evaluation and segmentation, as was advanced in section 4.2. Main contributions of this thesis are a novel spatio-temporal model for cardiac contours, and an extension to Procrustes alignment that removes pose and subject size variability while maintaining dynamic effects. The shape model and alignment method are proposed and discussed in this chapter, part of which has very recently been published by Casero and Noble [36]. The shape model is proposed in section 5.2. The alignment method is proposed in section 5.3.

PCA shape models have been extended to accommodate temporal information. The position of a 2D contour in a frame can be used to define an area of probability where the contour can be found in the next frame using a Bayesian framework [56]. 2D active contours can be extended to 2D+t introducing a time variable in the energy term [39, 93].

This in practice means that the active contour has a global smoothing term for the cardiac sequence that forces a smooth deformation between frames. 2D+t echocardiography data has been processed as a 3D volume using isotropic (same spread in spatial and temporal coordinates) 3D filters for feature detection [126, 127]. This way, features that appear in consecutive frames produce a stronger response than feature-like artifacts. A Kalman filter formulation can be used to update the coefficients of a PCA shape model [22, 23]. The Kalman filter implements a motion model that makes a prediction on the PCA coefficients based on the 2 previous frames. This approach was used to track myocardial borders in echocardiographic sequences [82–84].

In this thesis, the main interest are other type of models that can be readily integrated within a deformable template scheme for segmentation of a full cardiac cycle. In particular, implicit time extensions were proposed to 2D and 3D PCA models by Bosch et al. [26] and Stegmann and Pedersen [171], in order to extend shape models with temporal information. Such models are implicit because instead of adding a time variable, they are built from the concatenation of shape vectors

$$s_{\text{implicit 2D+t}} = [s_{\text{2D}}^{1\top},\ s_{\text{2D}}^{2\top},\ \ldots,\ s_{\text{2D}}^{F\top}]^{\top} \tag{5.1}$$

where $s_{\text{2D}}^{i}$ is the shape at time $t(i)$. Shapes are obtained from landmark configurations applying a novel extension of Procrustes Alignment proposed in section 5.2. For the moment, it will be assumed that we have a way to remove pose and subject size variability while maintaining dynamic effects, and then, PCA is computed in the usual way. This approach has 3 important shortcomings. First, all cardiac cycles in the data set need to have the same number of frames; considering the variability of heart rates in subjects and sampling rates between studies, this is never going to be the case in practice. Thus, it becomes necessary to interpolate the contours to a fixed number of frames. If the model is then used for segmentation, it needs to be re-interpolated to the number of frames of the new study, or the image data interpolated to the number of frames of the model, a

hard and computationally expensive problem that requires 2D+t volume registration and can introduce new artifacts.

Second, when $F$ frames are stacked together, the size of the data set is reduced by a factor $F$, and the number of variables increases by the same amount. That is, implicit 2D+t models require, in principle, $\mathcal{O}(F^2)$ times more subjects than simple 2D to approximate the data. With $F \approx 16$ in typical studies, this would become effectively infeasible. A computational issue may also arise, even if there is enough data, as the matrices of the eigenproblem can become very large. For example, the contrast DSE database provides 352 frames from 21 patients in 2C. The endocardial shape model has 25 landmarks, i.e. 50 variables. The mean number of frames is approximately 16. If each patient is interpolated to 16 frames, an implicit model would be built from 21 training vectors with 50x16=800 variables each. Section 4.15 discussed the problem of having many more variables than training vectors for texture data, but the argument can be applied to shape data too, as will be demonstrated in the experimental part of this chapter.

Third, implicit 2D+t models assume that consecutively occurring positions of the same landmark are separate independent variables, while it is more realistic and informative to model the variability of each point as a 2-dimensional random variable that changes with time. In the next section, the novel explicit 2D+t cyclic shape model proposed by Casero and Noble [36] to address these issues is presented.

## 5.2 A novel explicit 2D+t cyclic shape model

Looking for an explicit 2D+t model, a 3D model [122] may look like a sensible option, if the third spatial coordinate is replaced by time. But because all the contour points in the same frame share the same value of $t$, this is equivalent to concatenating the same variable $n$ times to the shape vector. It follows that the determinant of the covariance matrix $|C| = 0$, and it is not so straightforward to solve the eigenproblem. To avoid this, Casero and Noble [36] proposed an extended shape vector with a single time variable

$t \in [0,\ 1]$

$$s_{\text{explicit 2D+t}} = [s_{\text{2D}}^{\top},\ rt]^{\top} \tag{5.2}$$

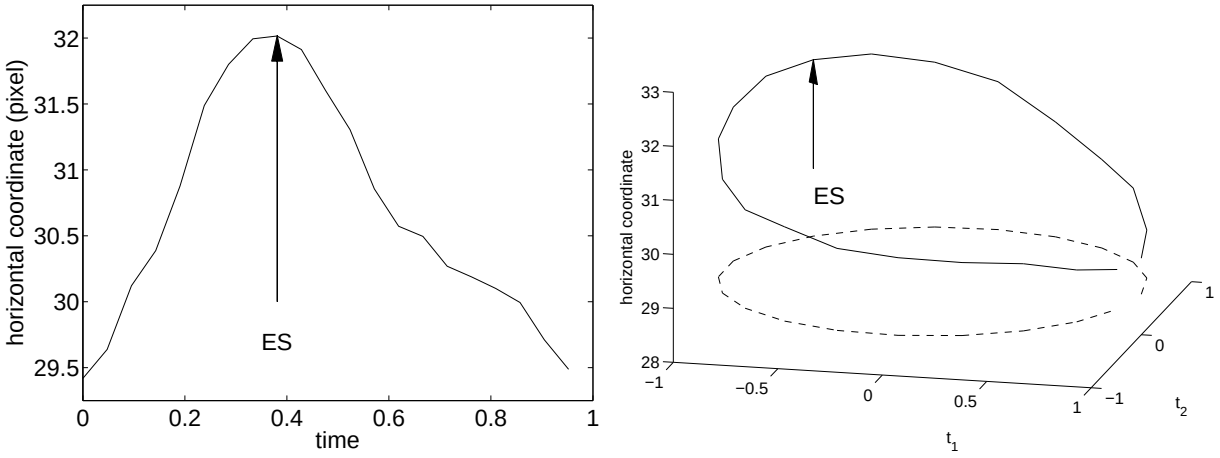
where $r$ is the scaling factor that was discussed in section 4.13. In brief, to define $r$, it should be noted 1) that PCA searches not only for those directions in which relationships between variables are more linear, but also for those with larger variance; and 2) that because there are many more shape than time variables, the model tends to underestimate the temporal effect. A choice of $r$ consistent with these observations is given by Eq. (4.50b), that can be rewritten as

$$r = \sqrt{\frac{\sum_{i=1}^{n/2} \text{Var}(u(i)) + \sum_{i=1}^{n/2} \text{Var}(v(i))}{\text{Var}(t_1) + \text{Var}(t_2)}} \tag{5.3}$$

where $u,\ v$ are the Cartesian coordinates of a landmark, so that the total variance contributed to the model by shape variables is the same as that contributed by time variables, where the variance estimate Var is computed over the sample of size $M$.

The vector in Eq. (5.2) has important shortcomings of its own for cyclic dynamics. Fig. 5.1a illustrates the typical horizontal displacement of a 2D contour point in the middle of the left wall of a 2-chamber view. First, the horseshoe-like curve means that any linear model such as PCA will poorly approximate the relationship between spatial coordinates and time. Second, PCA is dual to linear Multidimensional Scaling (MDS) [70], where the distance matrix is defined by the scalar products between the training vectors, i.e. PCA tries to preserve Euclidean distances between training vectors. In Fig. 5.1a, points near $t = 0$ and $t = 1$ are far apart according to the Euclidean distance for the model; in reality, we know that they are close in the cardiac cycle.

Casero and Noble [36] contended that both the lack of linearity and the distance problem can be tackled with Kernel PCA (KPCA) [154], a non-linear generalisation of PCA (see details in Appendix F). The main idea that they borrowed from KPCA is that shape+time vectors can be mapped to a higher dimensional space in which the relations




(a) Linear time.

(b) Cyclic time. Dashed: projection on a horizontal plane.

Figure 5.1: Empirical horizontal displacement of an endocardial point in the middle of the inferior segment in 2C (see point marked with a '∘' in Fig. 5.5a). Curve computed as the mean of 21 subjects. Time for the cardiac cycle has been normalised to $t \in [0,\ 1]$, with $t = 0$ end diastole. The arrow points to end systole (ES). Coordinate units are pixels.

between variables are linear, and then we can compute PCA in that space. A detailed explanation of KPCA can be found in Appendix F. Casero and Noble [36] proposed the transformation

$$s_{\text{explicit 2D+t}} = [s_{\text{2D}}^{\top},\ rt_1,\ rt_2]^{\top} \tag{5.4a}$$

$$t_1 = \cos(2\pi t) \tag{5.4b}$$

$$t_2 = \sin(2\pi t) \tag{5.4c}$$

In terms of KPCA, data has been embedded in a hyper-cylindrical manifold using the non-linear mapping

$$\phi\left(\begin{bmatrix}\tilde{s}\\ r\tilde{t}\end{bmatrix}\right) = \begin{bmatrix}\tilde{s}\\ r\cos(2\pi t)\\ r\sin(2\pi t)\end{bmatrix} \tag{5.5}$$

While KPCA usually maps the data to a much higher dimensional space and uses MDS and the kernel trick to make computations tractable, Eq. (5.4) only increases the dimen-

sionality by 2, so it is possible to work directly in feature space. Fig. 5.1b illustrates the advantages of the map in Eq. (5.4). First, the curve and the manifold that contains it can be reasonably approximated by an ellipse and a plane, respectively, which suggests a good linear approximation $u \approx \alpha_1 t_1 + \alpha_2 t_2$ for some scalars $\alpha_1,\ \alpha_2$. And second, the points near $t = 0$ and $t = 1$ are now close in Euclidean distance. The PDM of Eq. (4.3) can now be expanded using Eq. (5.4). In centred block matrix form we have

$$\begin{bmatrix} \tilde{s} \\ r\tilde{t}' \end{bmatrix} = \begin{bmatrix} V_{1,1} & V_{1,2} \\ V_{2,1} & V_{2,2} \end{bmatrix} \begin{bmatrix} b' \\ b_r \end{bmatrix} \tag{5.6}$$

where $t' = [t_1,\ t_2]^\top$, $b' = [b_1,\ b_2]^\top$. An explicit relationship between shape and time can be obtained noticing that

$$\tilde{s} = V_{1,1} b' + V_{1,2} b_r \tag{5.7a}$$

$$rt' = V_{2,1} b' + V_{2,2} b_r \tag{5.7b}$$

Substituting $[b_1,\ b_2]^\top$ from Eq. (5.7b) in Eq. (5.7a), and uncentering $\tilde{s}$, the explicit 2D+t shape model can be formulated as

$$s = c + A_b b_r + A_t t' \tag{5.8}$$

where

$$c = \bar{s} - A_t \bar{t}' \tag{5.9a}$$

$$A_t = r V_{1,1} V_{2,1}^{-1} \tag{5.9b}$$

$$A_b = -\frac{1}{r} A_t V_{2,2} + V_{1,2} \tag{5.9c}$$

A common criticism of PCA is that coefficients are linear combinations of all input variables and vice versa, so it is difficult to make medical sense of the model. Methods like

sparse PCA have been devised to trade orthogonality in the eigenspace and uncorrelation in the coefficients for sparsity in the loadings (see e.g. [54, 165] for an overview). That is, each input variable becomes a function of only a subset of coefficients. But sparse models are not much better than normal PCA in terms of medical interpretation. Even if the coupling of variables is smaller with sparse PCA, in both cases modes of variation have to be plotted and then effects identified ad hoc (cf. [165]). Hence, I think that an approach like that of Eq. (5.8) addresses the problem of separating the temporal effect for a cardiac model in a more useful way than sparse PCA (note that both approaches are not incompatible and could be combined too). Eq. (5.8) models cardiac contour deformation as a function of time and/or traditional shape coefficients. While $t' = t'(b_r)$ still holds, changing the values of $b_r$ does not change $t'$ because of the implicit variables $b_1,\ b_2$ changing their values in the background. Effectively, what Eq. (5.8) does is replace the traditional modes of variation 1 and 2 given by $b_1,\ b_2$ by another two modes of variation given by $t_1,\ t_2$.

## 5.3 Alignment of 2D+t configurations

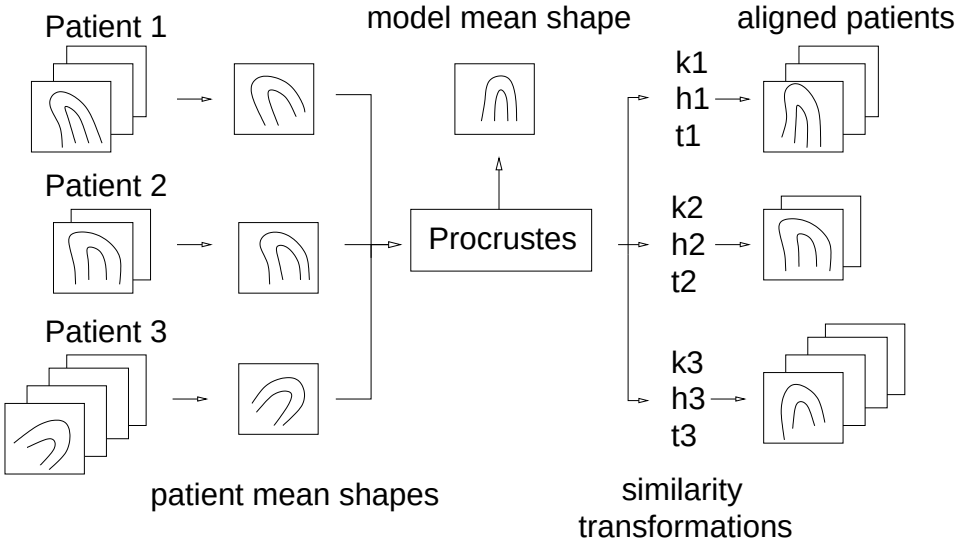



Figure 5.2: Procrustes alignment for 2D+t data. Similarity transformations comprise a scaling $k$, a rotation $h$ and a translation $t$.

Section 4.4 explained Procrustes alignment, an algorithm that removes unwanted (or extrinsic) variability in the data, namely pose and patient size. But if Procrustes alignment is applied to all data frames from all patients stacked together, then important variability due to cardiac dynamics is removed as well. To preserve time variability while still removing pose and patient size variability, Casero and Noble [36] proposed the scheme in Fig. 5.2. For 2D+t Procrustes, each patient data set is reduced to a mean shape, and all mean shapes are aligned with the usual 2D Procrustes method. Then the similarity transformation computed for the $i$-th mean shape is applied to each frame of the $i$-th data set

$$s'_j = k_i \, s_j \, h_i + t_i, \quad j = 1, \ \ldots, \ F \tag{5.10}$$

where $k_i$ is a scaling factor, $t_i$ is a translation row vector and $h_i$ is a rotation matrix with counterclockwise rotation $\theta$

$$h_i = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \tag{5.11}$$

## 5.4 Temporal reparameterisation for model asymmetry

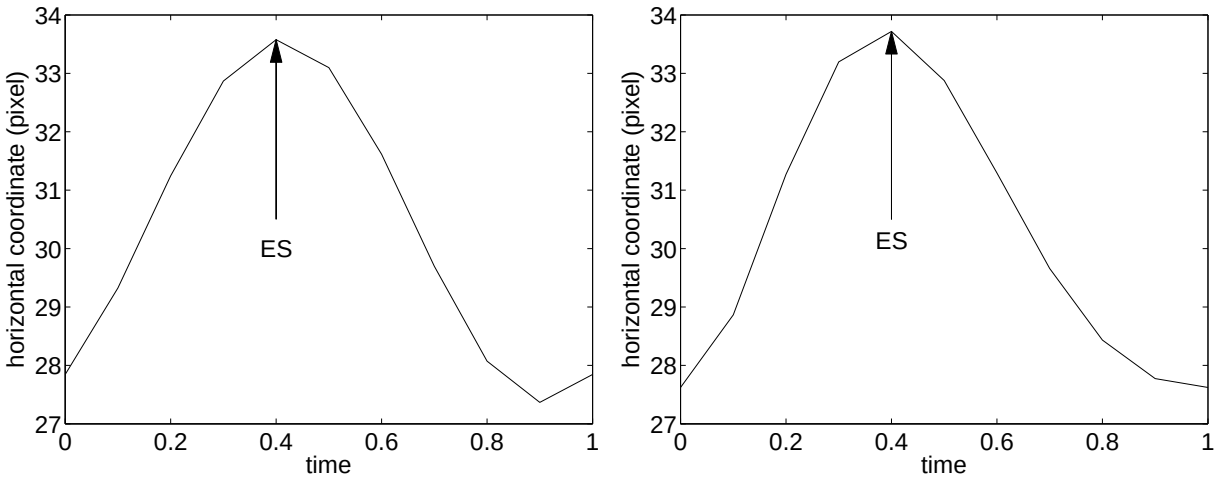
Fig. 5.3a shows that the 2D+t model of the previous section is not able to reflect the asymmetry of the actual contraction, as in Fig. 5.1a. In this section I propose a reparameterisation of the temporal variable that attempts to approximate that physiological asymmetry. Moreover, without the reparameterisation, the myocardium finishes relaxing at roughly $t = 0.9$, as displayed by Fig. 5.3a, and then starts contracting again before end dyastole. The reparameterisation $F_t : [0, \ 1] \mapsto [0, \ 1]$ of the temporal variable can be expressed as

$$\tilde{t}' = \begin{bmatrix} \cos(2\pi F_t(t)) \\ \sin(2\pi F_t(t)) \end{bmatrix} \tag{5.12}$$

The reparameterisation $F_t$ can be defined as a 2nd-order polynomial

$$F_t(t) = p_2 t^2 + p_1 t \tag{5.13}$$

The polynomial coefficients $p_i$ are computed from 3 point correspondences: $t = 0 \mapsto t = 0$, $t_{\mathrm{ES}} \mapsto t = 0.5$ and $t = 1 \mapsto t = 1$. To estimate $t_{\mathrm{ES}}$, first the 2D+t model without reparameterisation is computed. Points on the left and right walls with maximum horizontal displacement are found from the model. The corresponding $t_{\mathrm{ES,left}}$, $t_{\mathrm{ES,right}}$ values are averaged to find $t_{\mathrm{ES}}$. Finally, the 2D+t model is recomputed using the reparameterisation $F_t$. Fig. 5.3b illustrates the approximation to asymmetric cardiac dynamics of the reparameterised model.




(a) Without temporal reparameterisation. (b) With temporal reparameterisation.

Figure 5.3: Modelled horizontal displacement of the endocardial point in Fig 5.1. Computed by changing time in Eq. (5.8) with a fixed $b_r = 0$. ES: End systole.

## 5.5 Visualisation of the model

Fig. 5.4 displays the shape variation in ED ($t = 0$) for the 4 first modes of the 2C endocardial model, for 5 values of the corresponding shape coefficient within $\pm 3\sigma$. Fig. 5.5 displays the temporal behaviour of the model in the 4 principal planes, both for endocar-

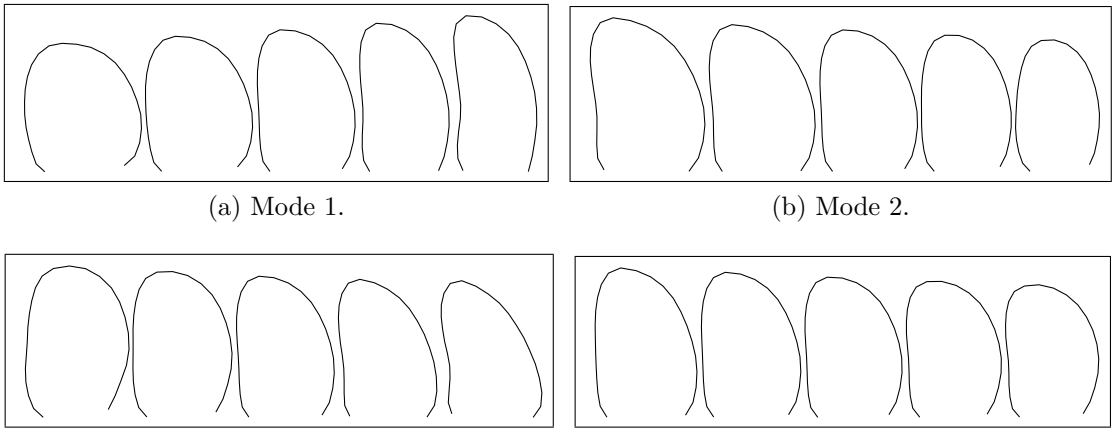
dial and myocardial models. It is worth noting that the LV endocardium presents larger displacements than the external boundary, as expected. Also, that in the SAX plane the endocardium rotates counterclockwise while the external boundary rotates clockwise. This is the expected behaviour that would be caused from myocardial torsion. Thus, the model is apparently in agreement with the physiology, although further research is required to determine whether that is the actual cause. In favour of this hypothesis is the fact that the observed 'torsion' is produced by a model computed from hand traced landmarks by a human expert. That is, the human expert did not only draw the myocardial contours, but also marked one of the ventriculo-septal junctions as a reference landmark on both boundaries. Hence, it is conceivable that the human expert introduced torsion information into the data this way. Against the hypothesis is the fact that there are only 2 such reference points per frame, and the rest of the pseudo-landmarks are extrapolated from them, so it could be an artifact too. A proposed future line of work would be to use Speckle Tracking to track speckle patterns on the myocardium and evaluate torsion. Speckle Tracking was briefly mentioned in section 4.3, but it is beyond the scope of this thesis.

Apical planes also seem to agree in broad terms with the physiology, although Kohl[1] pointed out that a larger vertical displacement of the mitral annulus would be expected in apical planes. Interestingly, the plots illustrate the problem with segment functional heterogeneity that was highlighted in section 3.4.2; that is, different segments have different degrees of expected wall excursion, that will have to be taken into account when building distributions of normal/abnormal functional values.

In this respect, another interesting future line of work would be to compute the models from normal studies only, use them as a mean physiology reference, and test for abnormalities as too large deviations from that reference. It should be noted that the models in Fig. 5.5 were computed from both normal and abnormal studies, assuming that low incidence of regional functional abnormalities would not significantly shift the model away

[1]Dr. Peter Kohl, MD. Department of Physiology, University of Oxford, UK. Personal communication.

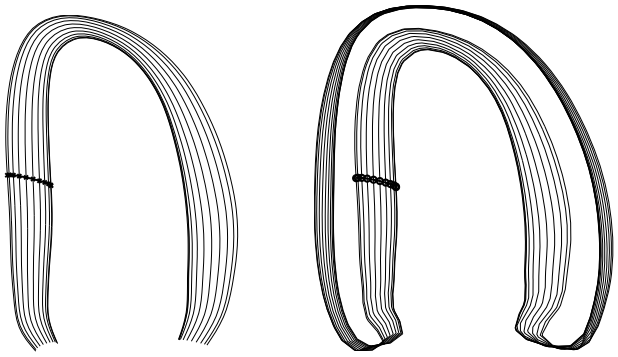
from normal behaviour. This approach suffices for this preliminary analysis, but in order to obtain meaningful clinical results, a much larger database of segmented data would be required.



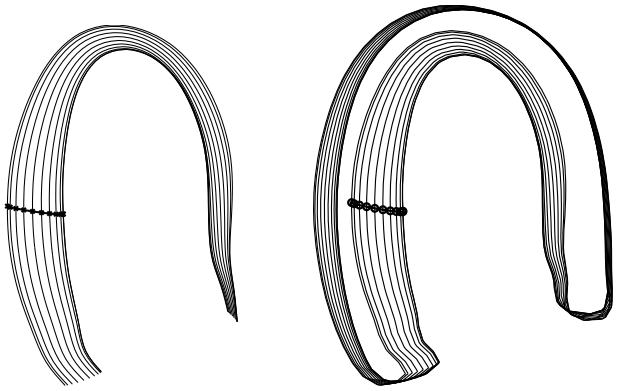
(a) Mode 1. (b) Mode 2.

(c) Mode 3. (d) Mode 4.

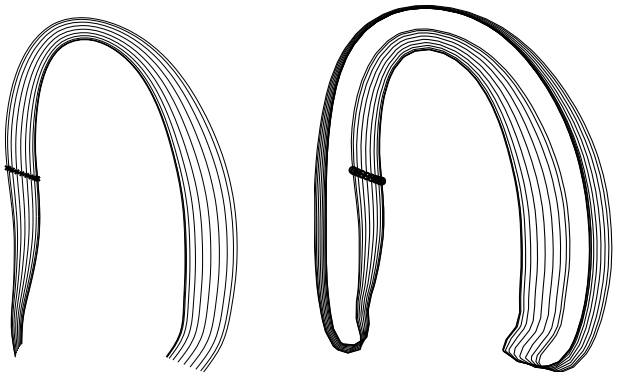
Figure 5.4: Modes of shape variation for 2D+t model in Eq. (5.8), 2C plane. Computed by evaluating the first 4 shape coefficients within $\pm 3\sigma$ in ED ($t = 0$).



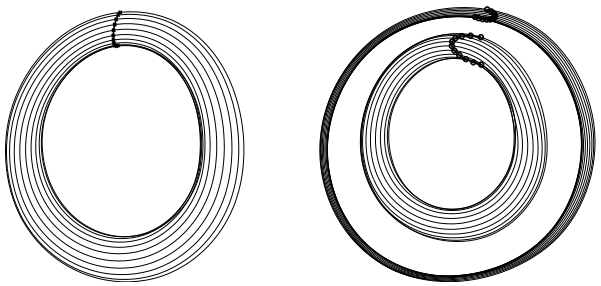
(a) 2C. (b) 2C.



(c) 3C. (d) 3C.



(e) 4C. (f) 4C.



(g) SAX. (h) SAX.

Figure 5.5: Mean shapes of 2D+t endocardial (left column) and myocardial (right column) models, with temporal reparameterisation. 11 contours in half a cardiac cycle, from ED to ES, are plotted for each plane.

## 5.6 Comparison between explicit and implicit models

In order to compare the reparameterised explicit 2D+t model proposed in this chapter to previous implicit models, the experiment in section 4.6 to compute the approximation error of 2D shape models is extended to 2D+t shape models in this section.

Shape models were computed from 20 patients applying a leave-one-out scheme to the 21 patient database. Contours from the remaining patient were aligned and converted to shapes $s$, projected onto shape space, and projected back into input space to obtain $\hat{s}$. For the implicit model

$$b^i = A^{i\dagger}\left(s^i_{\text{implicit}} - \bar{s}^i\right) \tag{5.14a}$$

$$\hat{s}^i_{\text{implicit}} = \bar{s}^i + A^i b^i \tag{5.14b}$$

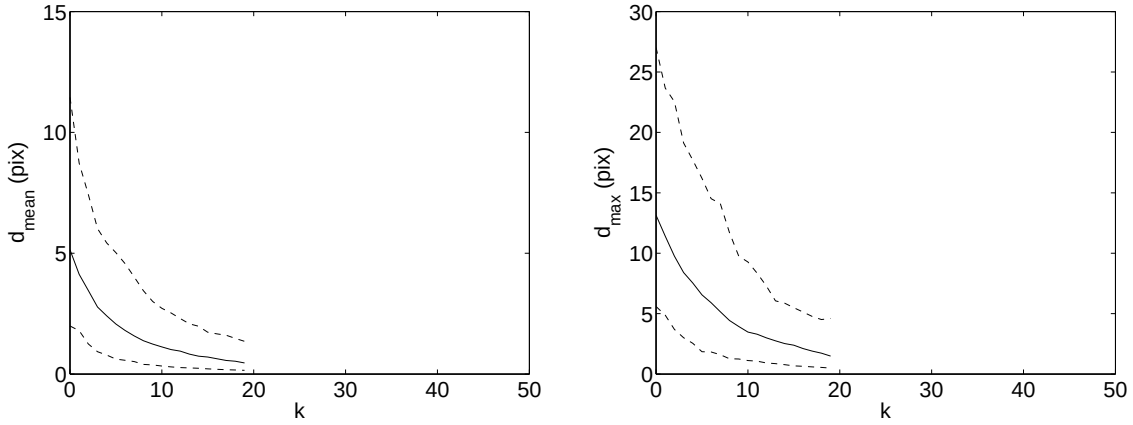
where $b^i,\ \bar{s}^i,\ A^i$ are the vector components and matrix block that correspond to the $i$-th frame $s^i$, and $\dagger$ is the Moore-Penrose pseudo-inverse. For the explicit model

$$b_r = A_b^{\dagger}\left(s_{\text{2D}} - c - A_t t'\right) \tag{5.15a}$$

$$\hat{s}_{\text{2D}} = c + A_b b_r + A_t t' \tag{5.15b}$$

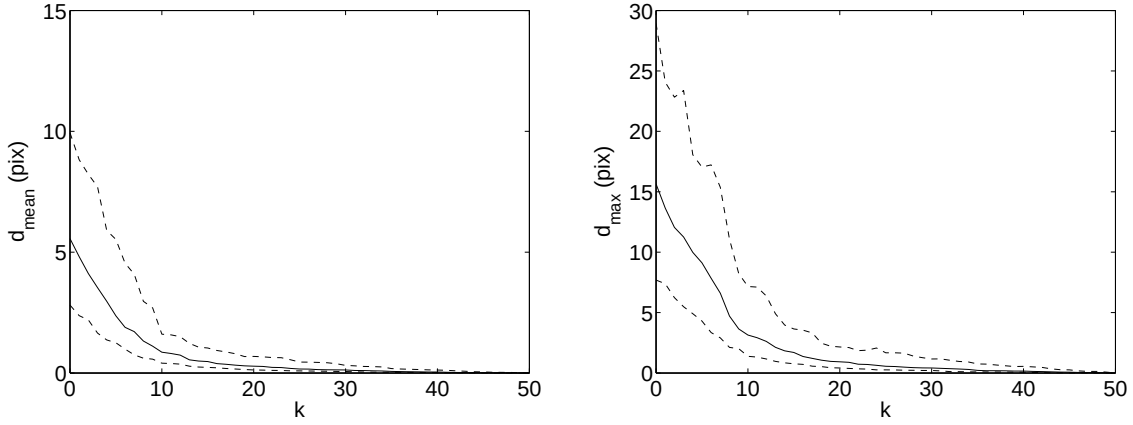
where $\hat{s}_{\text{2D}}$ is the back projection onto shape space. No constraints were imposed on the reconstruction. The approximation error was computed as the mean distance error $d_{\text{mean}}$ defined in Eq. (4.27) and maximum distance error $d_{\max}$ defined in Eq. (4.35). Similar results were obtained for all principal planes. The median and 95%-CI for the 2C view is shown in Fig. 5.6. The graphs suggest that the explicit model is slightly more compact, as the CIs are narrower and the median smaller. The implicit model is limited to a maximum of 20 eigenvectors, because it uses only 20 training vectors. Although not displayed, results for gaussianised data were computed too, but no significant difference was found with respect to non-gaussianised data. This is an interesting outcome, because it suggests that taking into account the temporal component of the data (both in implicit

and explicit models) naturally corrects violations of the Gaussianity assumption that had to be rectified with a gaussianisation method in section 4.7 for the 2D model. Moreover, both temporal models are substantially more compact than a simple 2D model, when compared to Fig. 4.5. From a KPCA point of view, this improvement can be explained by the data linearisation produced by the non-linear mapping of Eq. (5.5) in feature space. Data linearisation was illustrated by Fig. 5.1.




(a) Implicit model ($d_{\text{mean}}$).

(b) Implicit model ($d_{\max}$).




(c) Explicit model ($d_{\text{mean}}$).

(d) Explicit model ($d_{\max}$).

Figure 5.6: Approximation error for 2D+t endocardial models, frame by frame. Model truncated at $k$ shape coefficients. Solid: Median value. Dashed: 95%-CI. Leave-one-out scheme for 21 patients, 2C view.

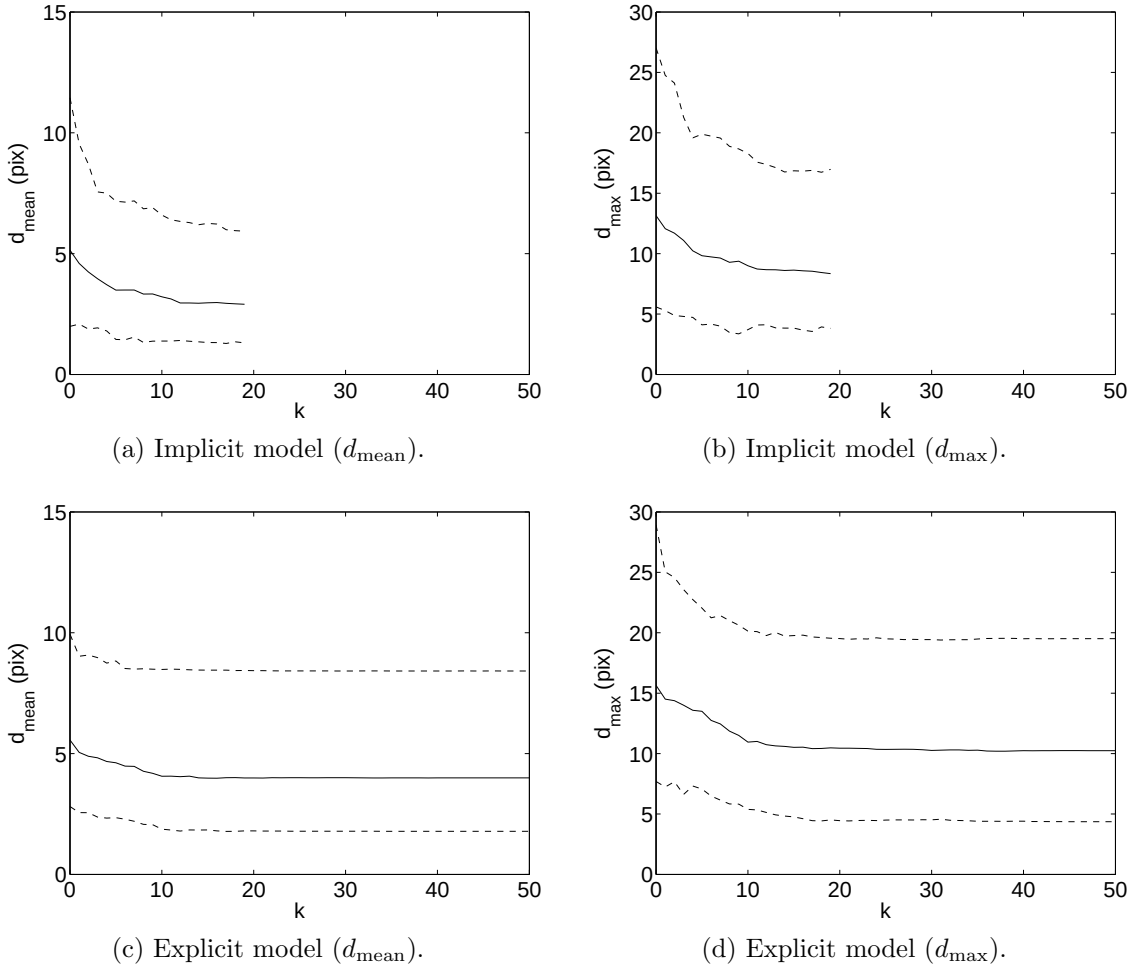
However, the true potential of spatio-temporal models lies on their ability to separate the shape variability (i.e. each patient's heart has an intrinsic shape) from the temporal deformation. This, for example, allows to formulate the segmentation problem as an

optimisation problem with a single shape vector for the whole volume. To evaluate this feature, instead of computing one shape vector per frame, a single shape vector was computed for each patient, i.e. assuming that each patient's heart has an *intrinsic shape* that changes from frame to frame only due to kinetics. For implicit models, the full size version of Eq. (5.14) is used. For explicit models, frames were concatenated to compute a shape vector $b_r$ optimal in the least-squares sense

$$b_r = \begin{bmatrix} A_b \\ \vdots \\ A_b \end{bmatrix}^{\dagger} \begin{bmatrix} s_{2\mathrm{D}}^1 - c - A_t t_1' \\ \vdots \\ s_{2\mathrm{D}}^F - c - A_t t_F' \end{bmatrix} \tag{5.16}$$

Fig. 5.7 shows the median and 95%-CI for the approximation error. In this case, the implicit model gives a slightly better approximation. A possible explanation for this is that the implicit model does not project landmark displacement onto a plane. Another point worth mentioning is that Figs. 5.6 and 5.7 contradicts the hypothesis in Casero and Noble [36] that $\mathcal{O}(F^2)$ more subjects are required to construct an implicit than an explicit model. On the other hand, refinements in the explicit model are still possible, e.g. by using more frequency components. Further research is required to explore this an other improvements, and compare the models with larger data sets or sets with more landmarks.

Another point worth noting is that the approximation error for both implicit and explicit models seems to plateau after 10 eigenvectors, and that maximum distances can be quite large, e.g. in the order of 20 pixel for an image of size 256x360 pixel. There are two possible causes that I think are most likely to be responsible for the error. First, the contrast DSE database is composed of 11 abnormal and 10 normal subjects, but even abnormals have working hearts where only part of the muscle is damaged. Thus, the models are expected to represent mostly healthy hearts, and may have trouble expressing abnormalities. While it would be very interesting to add variables to the model that can




(a) Implicit model ($d_{\text{mean}}$).

(b) Implicit model ($d_{\max}$).

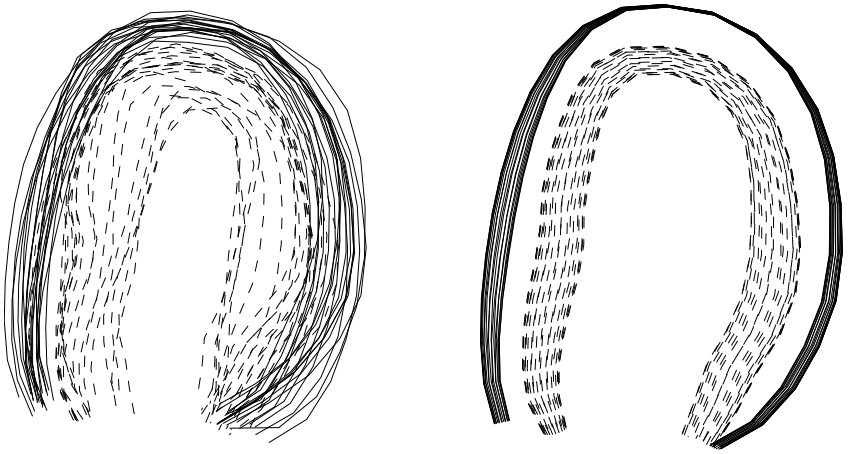
(c) Explicit model ($d_{\text{mean}}$).

(d) Explicit model ($d_{\max}$).

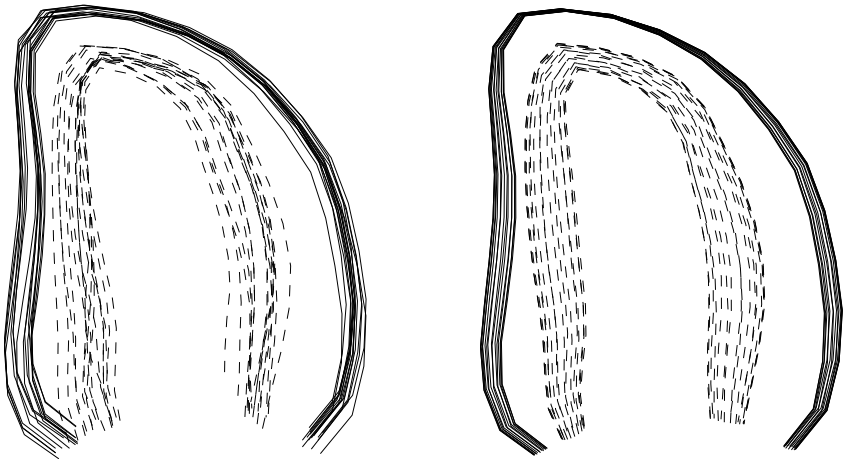
Figure 5.7: Approximation error for 2D+t endocardial models, all frames together. Model truncated at $k$ shape coefficients. Solid: Median value. Dashed: 95%-CI. Leave-one-out scheme for 21 patients, 2C view.

account for reduced contractility or for dyskinesia, the database simply does not have enough data for those studies, as was pointed out in section 3.4.2. Second, some studies are badly imaged and the muscle comes in and out of plane, as noted in section 3.4.1. This was, for example, the case with Patient 038, that has the largest approximation error in 2C. Patient 038 was diagnosed as normal, but because the imaging plane does not remain fixed, the contours change in a way that cannot be explained by the model. This is most visible in the large displacements of the external contour, as shown in Fig. 5.8a. In studies without this effect, the model arguably offers a good approximation to the expert contours, as illustrated by Fig. 5.8b. As future work, the model could be extended to

account for the out of plane effect. However, Becher[2] suggested that badly imaged studies should not be used for medical assessment at all. Moreover, increasing availability of 3D echocardiography machines will render this problem irrelevant, as 3D data does not suffer from the out of plane effect.



(a) Study going out of plane (Patient 038).



(b) Study staying in plane (Patient 005).

Figure 5.8: Out of plane effect. 2C view, both healthy subjects. Left: Expert contours. Right: Best approximation with 2D+t model and single shape vector. Solid: External contour. Dashed: Endocardium.

Finally, a limitation of both implicit and explicit models is that they assume a temporal correspondence between frames. But systolic time depends on many factors, e.g. heart rate and gender [188], and stress levels [183]. While it is tempting to use electrocardiograms to identify different cardiac phases, they are not reliable enough, especially with

[2]Dr. H. Becher, John Radcliffe Hospital, Oxford. Personal communication.

increasing HR and stress. Shekhar et al. [163] proposed doing *temporal alignment*, i.e. end diastole (ED) and end systole (ES) frames where used to partition the cardiac cycle in two intervals. Then each interval was normalised to define a linear temporal correspondence between frames in different studies. To obtain corresponding frames, each interval was resliced using Nearest Neighbour interpolation. Bosch et al. [26] resliced all studies to 16 frames, so that ED and ES frames had indices 1 and 9, respectively. In both approaches, frames were marked by an expert. Another problem is that a temporal linear interpolation will not always be correct. Systole is composed of 2 phases: pre-ejection period (PEP) and left ventricular ejection time (LVET), and for certain cardiopathies, the PEP increases, while the LVET decreases so that the total systolic time interval remains relatively unaltered [188]. Future research should address this issues, and aim for automatic temporal alignment, that according to current standard recommendations should be done using mitral valve motion and cavity size [96, Table 1].

# CHAPTER 6

# Atlas-based deformable model segmentation

## 6.1 Formulation of atlas-based segmentation

A classic approach to atlas-based segmentation, image registration or image alignment is the Lucas-Kanade algorithm [107, 108], that attempts to solve the least-squares optimisation problem

$$p_{\text{opt}} = \min_{p} \frac{1}{2} \sum_{z} \Big( T(z) - I(W(z;\ p)) \Big)^2 \tag{6.1}$$

where $W : \mathbb{R}^2 \mapsto \mathbb{R}^2$ is a function with parameters $p$ that maps point coordinates $z$ in the template image $T$ to the target image $I$, i.e. a *forward* transformation/warp function. Baker and Matthews [5, 6] reviewed variations on the formulation of Eq. (6.1), summarised in the rest of this section. Full details and references to relevant papers can be found in [5], or more extensively in [6]. The Lucas-Kanade algorithm applied gradient descent to image alignment. It is a forward *additive* algorithm, i.e. it solves for an increment $\Delta p$ of the parameters to minimise the first order Taylor expansion of

$$p_{\text{opt}} = \min_{p} \frac{1}{2} \sum_{z} \Big( T(z) - I(W(z;\ p + \Delta p)) \Big)^2 \tag{6.2}$$

The Shum-Szeliski algorithm is a forward *compositional* algorithm, and hence it solves for an incremental warp instead of an increment of the parameters

$$p_{\text{opt}} = \min_{p} \frac{1}{2} \sum_{z} \Big( T(z) - I(W(W(z;\ \Delta p),\ p)) \Big)^2 \tag{6.3}$$

Additive and compositional algorithms are equivalent to a first order approximation [5] and in terms of computational cost [6]. To avoid the huge cost of computing the Hessian at every step of the Lucas-Kanade algorithm, the Hager-Belhumeur algorithm exchanged the role of the template and the image in the Lucas-Kanade formulation, i.e. it is an *inverse* additive algorithm. The Hager-Belhumeur algorithm can only be applied to a very reduced set of warps, though. The least-squares inverse compositional algorithm searches for the optimal parameters

$$p_{\text{opt}} = \min_{p} \frac{1}{2} \sum_{z} \Big( T(W(z;\ \Delta p)) - I(W(z;\ p)) \Big)^2 \tag{6.4}$$

The transformation/warp is updated using

$$W(z;\ p) \leftarrow W(z;\ p) \circ W^{-1}(z;\ \Delta p) \tag{6.5}$$

Baker and Matthews [5] contended that forward/inverse compositional algorithms can only be applied to warps that form semigroups/groups, respectively, under composition. However, one of the properties of semigroups and groups, associativity, is not required by their formulation[1]. Thus, forming a semigroup/group under composition is a sufficient but not a necessary condition for the transformations.

Baker and Matthews [6] proposed an inverse compositional algorithm for (global) affine warps using several standard gradient descent approximations to solve the least-squares problem (see e.g. [132] for extensive details on numerical optimisation methods): Gauss-Newton, Newton, steepest-descent, diagonal Hessian and Levenberg-Marquardt.

[1] S. Baker, Microsoft Research, Redmon, USA. Personal communication.

Matthews and Baker [111] extended the Gauss-Newton approximation to AAMs, combining similarity transformations and piece-wise affine warps. In an unrelated paper, Eriksson and Åström [64] used the Gauss-Newton approximation[2] and thin-plate splines in a registration problem that matched all pixels on an image to all pixels on another image. Without mentioning it explicitly, they used the same formulation as the Lucas-Kanade algorithm to solve the optimisation problem, i.e. a forward additive algorithm. A limitation of these approaches is that they do not include line search, even though it is a standard element of optimisation algorithms (see e.g. [132, sec. 3.5]). Implementation details for the Gauss-Newton method applied to the inverse compositional algorithm using similarity transformations, thin-plate splines and line search are provided in Appendix G. To be able to compare the results of my implementation, line search and stopping criteria were implemented following Matlab optimisation functions. Although beyond the scope of this thesis, it is worth noting that the inverse compositional algorithm has been extended to 3D AAM [1] and generalised to other matching error measures, e.g. normalised correlation and mutual information [28].

## 6.2 Hierarchical schemes

Due to the curse of dimensionality (a term coined by Richard E. Bellman in the context of dynamic programming), the volume of the search space in an optimisation problem increases exponentially with the number of variables, making the search harder. In addition, deformations are in general computationally more expensive and the number of local minima increases with the number of Degrees of Freedom of the model (e.g. [124]). Bergen et al. [17] argued that ignoring high resolution information in the image is not only efficient but necessary, to avoid aliasing of high spatial frequency components. Ashburner [4] reviewed the onus created by large deformations for warps to be diffeomorphic. All these observations translate into solving the problem hierarchically, refining the solution

[2]Eriksson and Åström [64] derived the problem using the Newton approximation, but then simplified the Hessian using the Gauss-Newton approximation.

using transformations with increasing degrees of freedom (*hierarchical deformation*) on images with increasing resolution (*multiresolution*) [17].

Transformations are presented in detail in section 6.3, but a brief overview of hierarchical deformation examples follows. Cootes et al. [47] proposed the Point Distribution Model (PDM), that is more or less flexible depending on the number of eigenvectors, as discussed in section 4.6. Davatzikos et al. [55] used Wavelet Transform coefficients to explain contours at different resolution levels. Shang et al. [161] discretised cardiac 3D surfaces using meshes with increasing number of vertices. Kervrann and Heitz [91] proposed optimising the parameters of a PCA and similarity transformation deformable hand model first, and then refine the segmentation allowing for local displacement of the landmarks under a Gaussian random model. Metaxas et al. [117] proposed to optimise the similarity transformation, and a local deformation based on a Free Form Deformation model, simultaneously, but shifting weight from the former in the first 5-20 steps of the algorithm, to the latter afterwards. Feldmar and Ayache [65] registered pairs of surface segmentations of 3D organs and skeletal regions using rigid, affine and piece-wise affine transformations sequentially. As mentioned above, Matthews and Baker [111] combined similarity transformations and piece-wise affine warps in the inverse compositional algorithm.

Similarly to the schemes above, this thesis splits the transformation $W$ into a similarity transformation and a thin-plate spline warp based on the PCA shape model. Following the notation in Matthews and Baker [111], the combined global transformation and local warp can be expressed as

$$W(z;\ p) = W_G(W_L(z,\ \bar{s},\ s(b));\ q) \tag{6.6}$$

where $p = [q^\top,\ b^\top]^\top$, $W_G$ is the global transform and $W_L$ is the local warp. The global transform $W_G$ is presented in section 6.3. Local warps $W_L$ were used to generate the texture analysis and probabilistic atlases in sections 4.3 and 4.8, respectively. Atlases

are used for segmentation in this chapter, but non-rigid optimisation for segmentation is beyond the scope of this thesis. A discussion and formulation of local warps is provided in Appendix C as a basis for future work.

## 6.3 Similarity transformations

The global transform is Eq. (6.6) is a similarity transformation, i.e. a scaling, a rotation and a translation. Note from section 4.4 that those are precisely the variations removed by Procrustes alignment. The transform $W_G$ is often formulated as

$$W_G(z;\ \alpha,\ \theta,\ t) = \alpha \begin{bmatrix} cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} z + t \tag{6.7}$$

where $\alpha$ is the scaling factor, $\theta$ is the rotation angle and $t$ is the translation vector. However, for atlas-based segmentation the shape-matrix parameterisation is more convenient (e.g. [21, section 4.2], [111])

$$W_G(z;\ q) = z + A_G(z)\, q \tag{6.8a}$$

$$A_G(z) = \begin{bmatrix} z(1) & -z(2) & 1 & 0 \\ z(2) & z(1) & 0 & 1 \end{bmatrix} \tag{6.8b}$$

where $z = [z(1),\ z(2)]^\top$.

## 6.4 Applying the 2D+t model to atlas-based segmentation

The formulation of the atlas-based segmentation problem in section 6.1 involves a 2D template that is optimised to match a 2D image. For simplicity, only similarity transformation optimisation is considered. Solving the optimisation problem separately for every

frame produces poor results, even if different templates are computed from a 2D+t model for each frame, as illustrated by Fig. 6.1. The main reasons are the changing blood flow at the base of the heart while the atlas remains constant, and the presence of artifacts as discussed in section 4.2. It can be hypothesised that the results can be improved if the segmentation algorithm is run on the whole 2D+t cardiac cycle at the same time. In order to test this hypothesis, I propose to extend the formulation of the atlas-based segmentation problem using the explicit 2D+t model proposed in section 5.2

$$p_{\text{opt}} = \min_p \frac{1}{2} \sum_z \left( \frac{1}{F} \sum_{i=1}^{F} T(W(z;\ \Delta p,\ t_i)) - I(W(z;\ p,\ t_i)) \right)^2 \tag{6.9}$$

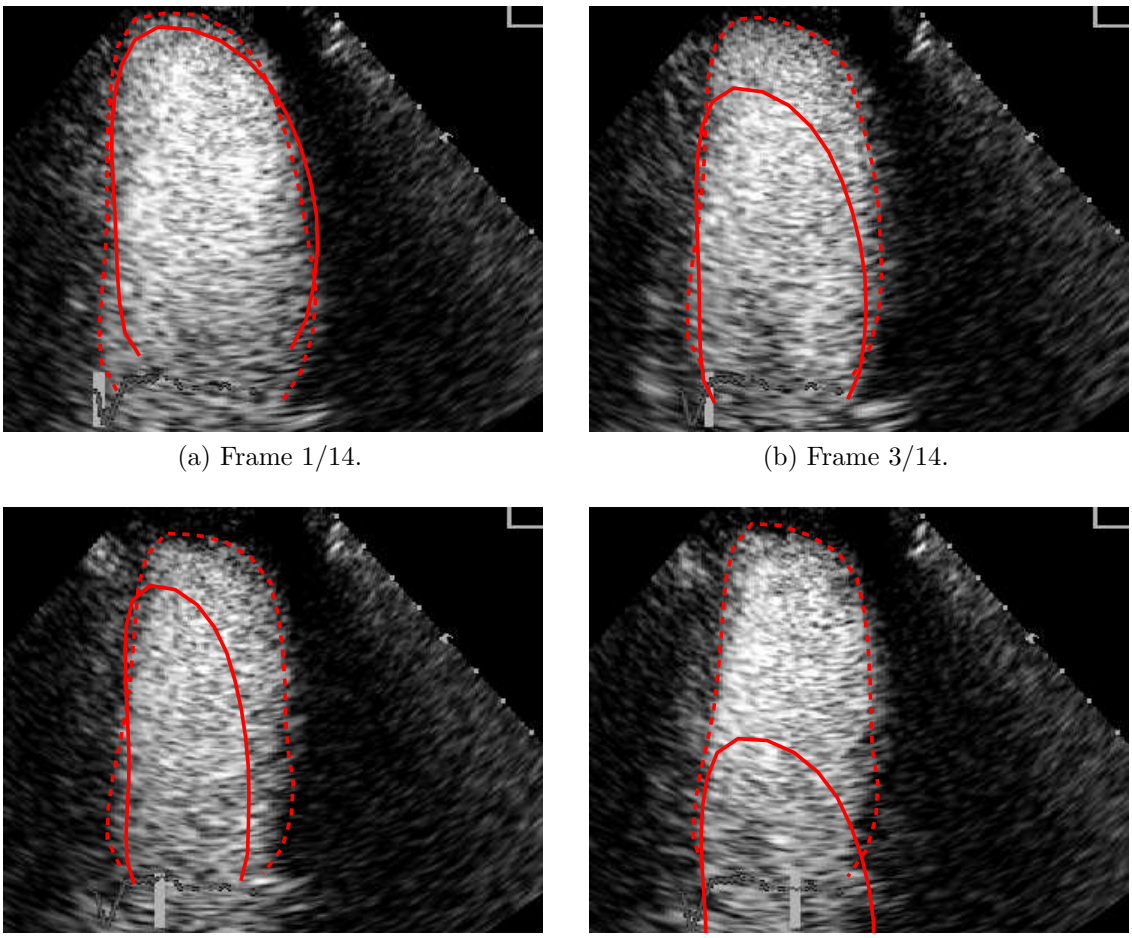
The transformation $W$ now depends on a temporal variable, that allows to generate a different template for each of the $F$ frames with the model in section 5.8, reproduced here for convenience

$$s = c + A_b b_r + A_t \begin{bmatrix} \cos(2\pi F_t(t)) \\ \sin(2\pi F_t(t)) \end{bmatrix}$$

where $s$ is a shape vector, $t \in [0, 1]$ is time normalised in the cardiac cycle, $b$ is a vector of shape coefficients shared by all frames, $F_t$ is the temporal reparameterisation function, and $M_t,\ M_b,\ m_c$ are the model matrices and vector. Eq. (6.9) can be seen as replacing the residual of a sampling mask point by its average residual over all frames. While the mean is a less robust statistic than e.g. the median, it is faster to compute and can be differentiated analytically. The implementation of the algorithm is similar to Algorithm G.1 (p.190), but averaging the results from each frame in each step.

## 6.5 Experimental results

To compare the performance of different approaches quantitatively, shape models and atlases were computed from 20 patients of the 21 in the database with the leave-one-out approach used previously in this thesis. Then, the similarity transformation was optimised



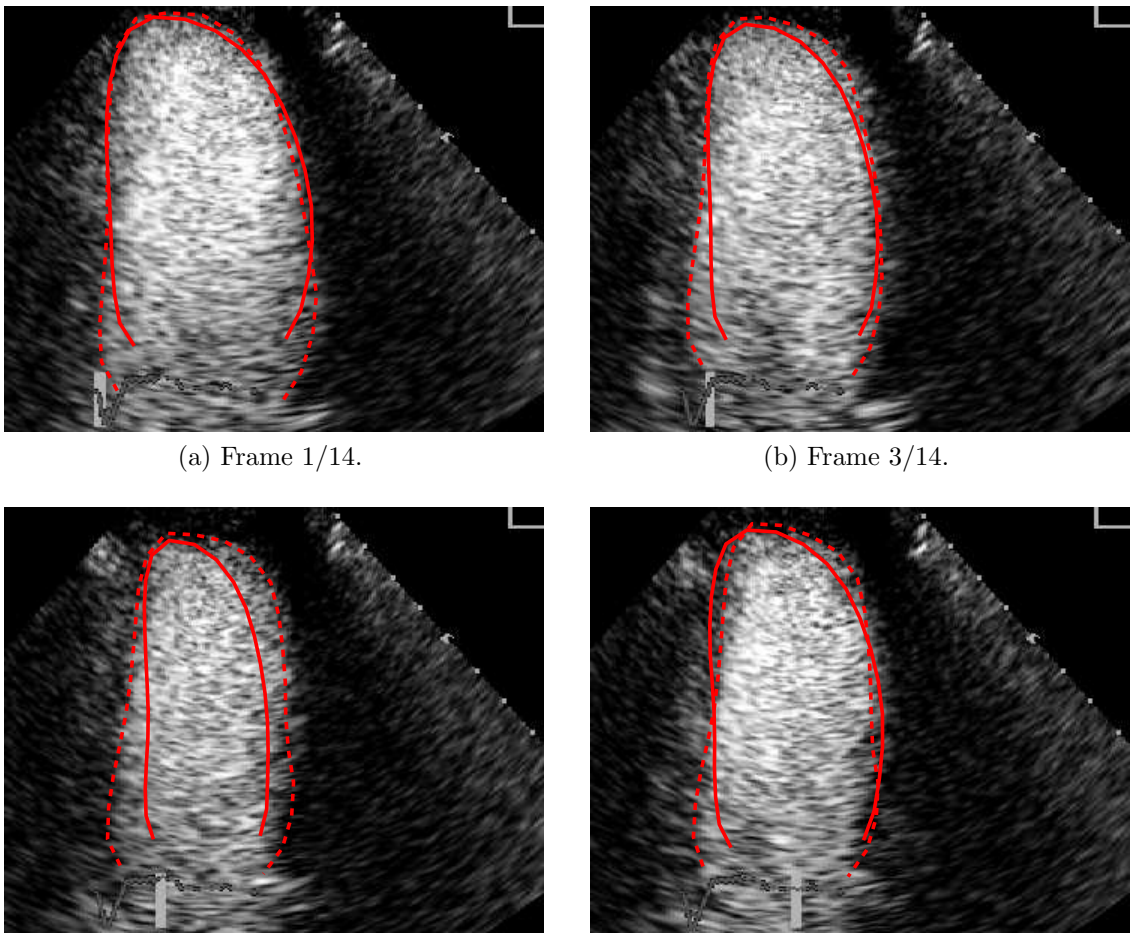
(a) Frame 1/14.

(b) Frame 3/14.

(c) Frame 6/14.

(d) Frame 10/14.

Figure 6.1: Atlas-based segmentation with 2D+t model frame by frame. Dashed: Expert traced contour. Solid: Algorithm optimum. 2C. 20 patient model run on "Patient 018" data (leave-one-out). Gauss-Newton algorithm with line search. Segmentation was optimised on the image scaled down by a factor 0.25 to a size of 64x90 pixel.



(a) Frame 1/14.

(b) Frame 3/14.

(c) Frame 6/14.

(d) Frame 10/14.

Figure 6.2: Atlas-based segmentation with 2D+t model on whole cardiac cycle. Dashed: Expert traced contour. Solid: Algorithm optimum. 2C. 20 patient model run on "Patient 018" data (leave-one-out). Gauss-Newton algorithm with line search. Segmentation was optimised on the image scaled down by a factor 0.25 to a size of 64x90 pixel.

on the remaining patient using 4 methods: 1) Direct minimisation of the least-squares distance to human expert hand-traced contours (this was considered the baseline approximation error for the other 3 methods), whole cardiac cycle; 2) Optimisation frame by frame, using templates from the 2D+t model and the inverse compositional algorithm; 3) Optimisation of whole cardiac cycle, using Matlab's `lsqnonlin` function[3], that implements the Gauss-Newton forward additive algorithm with line search; 4) Optimisation of the whole cardiac cycle, using my implementation of the inverse compositional algorithm extended with line search, as described above. In methods 2), 3) and 4), the similarity transformation variables were scaled by their respective standard deviations estimated from the training data set. The reason is that the performance of the Gauss-Newton algorithm suffers in poorly scaled problems (e.g. [132, Ch. 2]), and the translation variables $q_3,\ q_4$ are $\mathcal{O}(10^2)$ larger than the rotation and scaling variables $q_1,\ q_2$.

Fig. 6.2 illustrates the segmentation improvement achieved by the 2D+t model. This is only a visual example with Patient 018, but leave-one-out experiments were run for all patients (including the ones identified as outliers in section 3.4.1), and the results presented in the rest of this section suggest that the improvement is not accidental.

Table 6.1 displays the percentage of frames with divergent segmentations ($d_{\text{mean}} > 15$ pixel). The frame by frame (ICFbF) method performed worse than the 2D+t inverse compositional (IC) method, in general. Segmenting the whole 2D+t volume instead of each frame separately substantially reduced the number of divergent cases, although further work is necessary to explain the worsening for the 4C plane. The 2D+t forward additive (FA) method outperformed the IC method, and in fact, it converged in virtually all tests.

Table 6.3 provides running times in seconds for each method. For the inverse compositional approaches, two values were computed: the set-up time (pre-processing time), and the iteration time (running time). In the FA method, no pre-processing is possible. The difference between inverse compositional methods was small; the results suggest that

[3]Matlab's `nlsq()` function code was hacked to prevent it from switching to the Levenberg-Marquardt method when the Gauss-Newton method is suboptimal.

it is slightly faster to optimise the whole cycle than optimising each frame on its own. On the other hand, the difference between the inverse compositional and forward additive approaches was large, with the former being 20 to 25 times faster.

Therefore, reducing divergence cases in the IC method to FA levels would be an interesting line of work for the future. Considering the speed up provided by the IC method, divergence could be reduced simply by using multiple initialisations. The speed gain brings segmentation of a whole cardiac cycle to the order of 0.1 to 4.5 seconds, using Matlab code with many loops —known to be inefficient. An average cardiac cycle lasts for 0.86 sec. (for HR=70 beats per min.), so real time segmentation could be within reach[4].

Table 6.5 presents the approximation error values once divergent segmentations were removed from the results. The approximation error was computed as the mean distance between landmarks $d_{\text{mean}}$ defined in Eq. (4.27). The table presents the median followed by the 95% Confidence Interval (CI) in brackets. Images were scaled down by a factor 0.25 to a size of 64x90 pixel. The results suggest that the two inverse compositional algorithm approaches are similar and slightly less accurate than the FA method (0.5 pixel in apical planes and 1 to 2 pixel in SAX, approximately).

To gain some insight into hierarchical schemes, the segmentation evaluation was repeated for the images at full scale. Table 6.6 suggests that $d_{\text{mean}}$ values were 4 times larger, but on images 4 times larger too. Thus, increasing the resolution did not significantly change the segmentation results. Running times seemed to increase linearly with the number of sampling points, and thus, with the square of the resolution factor. This view is supported by the 16-fold increase in running times at full scale. Therefore, full scale images should not be used for similarity transformation optimisation, as accuracy is not improved, despite the substantial increase in running time.

[4]Experiments were run on a computer cluster, with 10 nodes. Each node has two AMD Opteron 265 Dual core (1.8 GHz) processors and 4GB of RAM. Any given node offers 4 queues, each of which can run one experiment at a time. Reported times correspond to one experiment running in one queue.

## 6.6 Summary and conclusions

This chapter posed endocardial segmentation in Power Modulation contrast echocardiography 2D+t volumes as an atlas-based segmentation problem, that can be tackled using a standard Gauss-Newton gradient descent framework. The framework was discussed in terms of standard hierarchical considerations from computer vision. A numerical optimisation method developed in the last few years, called the inverse compositional algorithm, was tested as an alternative to the classic Lucas-Kanade algorithm. The optimisation framework was extended with the 2D+t model proposed in section 5.2 so that whole cardiac cycles can be segmented, as opposed to segmentation of individual frames.

Algorithms were implemented with the same structure, line search and stopping conditions as Matlab optimisation functions, so that results could be compared meaningfully. The results suggest that the forward additive algorithm is more reliable, and slightly more accurate, but 20 to 25 times slower than the inverse compositional approaches. Segmenting the whole cardiac cycle together in fact seems to produce substantially better results than segmenting frame by frame in inverse compositional methods.

It should be noted that segmentation limited to similarity transformations is only valid as a rough approximation to the solution. But if the number of divergent segmentations can be reduced, then the inverse compositional approach could be a reasonable fully automatic first step for real time segmentation. There are limitations to the 2D+t model too. For example, Fig. 6.2 illustrated that the model seems to assume a healthy heart, and thus underestimates the predicted cavity area in end systole for a very hypokinetic LV. Hence, future work should address the introduction of local warps and more flexibility into the model, and advance of which is discussed in Appendix. C. Other improvements could be obtained from replacing the Gauss-Newton method by the Levenberg-Marquardt algorithm.

Finally, no significative performance differences were found between segmentation on full scale images and images scaled-down by a factor of 0.25, while running times increased

| | 2C | 3C | 4C | SAX |
|---|---|---|---|---|
| ICFbF | 8.5 | 5.0 | 6.1 | 35.6 |
| FA | 0.0 | 0.0 | 0.0 | 0.3 |
| IC | 0.0 | 0.0 | 12.7 | 11.5 |

Table 6.1: Divergent segmentations in Gauss-Newton algorithms for similarity transformation. The measure is the percentage of frames with large error segmentation results ($d_{\text{mean}} > 15$ pixel). ICFbF: Inverse compositional frame by frame. FA: Forward additive whole cycle. IC: Inverse compositional whole cycle. Images were scaled down by a factor 0.25 to a size of 64x90 pixel.

| | 2C | 3C | 4C | SAX |
|---|---|---|---|---|
| ICFbF | 6.2 | 6.2 | 10.0 | 49.0 |
| FA | 4.5 | 0.0 | 0.0 | 2.6 |
| IC | 9.9 | 5.5 | 4.5 | 6.8 |

Table 6.2: Divergent segmentations in Gauss-Newton algorithms for similarity transformation. As Table 6.2, but images at full scale were used, with a size of 256x360 pixel. The divergent segmentation criterion was set at $d_{\text{mean}} > 60$ pixel to reflect the change in scale.

by a factor of 16. This could be due to the lack of change in intensity histograms observed in Ch. 4. Following the discussion about hierarchical models, this suggests that all the information relevant for similarity transformation optimisation is contained in the coarser scale, so using the full scale data is not necessary.

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 1.8 | (1.1, 2.1) | 0.9 | (0.5, 3.0) |
| FA | N/A | | 24.7 | (4.9, 756.7) |
| IC | 1.5 | (0.9, 1.8) | 0.6 | (0.3, 1.7) |

(a) 2C

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 1.6 | (1.1, 2.1) | 0.8 | (0.6, 1.4) |
| FA | N/A | | 21.8 | (2.3, 1284.6) |
| IC | 1.3 | (0.8, 1.7) | 0.6 | (0.3, 1.4) |

(b) 3C

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 1.7 | (1.0, 2.2) | 0.9 | (0.5, 1.4) |
| FA | N/A | | 22.1 | (1.8, 905.5) |
| IC | 1.4 | (0.8, 2.0) | 0.6 | (0.3, 4.5) |

(c) 4C

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 1.4 | (0.8, 1.9) | 0.7 | (0.3, 1.3) |
| FA | N/A | | 14.7 | (1.6, 1526.3) |
| IC | 1.2 | (0.6, 1.7) | 0.5 | (0.1, 3.6) |

(d) SAX

Table 6.3: Speed comparison of Gauss-Newton segmentation algorithms for similarity transformation. The measure is time to stop (sec): Median (95% CI). ICFbF: Inverse compositional frame by frame. FA: Forward additive whole cycle. IC: Inverse compositional whole cycle. Images were scaled down by a factor 0.25 to a size of 64x90 pixel.

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 33.8 | (19.3, 40.2) | 13.2 | (5.7, 107.9) |
| FA | | N/A | 415.8 | (17.7, 19138.4) |
| IC | 33.9 | (19.2, 40.6) | 9.7 | (5.7, 127.8) |

(a) 2C

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 31.4 | (19.8, 41.2) | 14.4 | (6.6, 49.1) |
| FA | | N/A | 497.7 | (20.8, 1898.8) |
| IC | 30.3 | (19.2, 41.0) | 8.9 | (4.6, 96.8) |

(b) 3C

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 34.5 | (19.5, 46.5) | 13.6 | (5.4, 38.8) |
| FA | | N/A | 311.8 | (17.2, 5503.8) |
| IC | 34.3 | (19.5, 49.5) | 10.0 | (5.2, 60.6) |

(c) 4C

| | Set-up | | Iterations | |
|---|---|---|---|---|
| ICFbF | 27.6 | (14.1, 40.5) | 7.1 | (3.6, 16.2) |
| FA | | N/A | 579.5 | (163.1, 8564.7) |
| IC | 27.0 | (13.7, 39.8) | 6.1 | (0.9, 124.5) |

(d) SAX

Table 6.4: Speed comparison of Gauss-Newton segmentation algorithms for similarity transformation. As Table 6.3, but images at full scale were used, with a size of 256x360 pixel.

| | 2C | | 3C | | 4C | | SAX | |
|---|---|---|---|---|---|---|---|---|
| Baseline | 1.4 | (0.7, 3.3) | 1.4 | (0.6, 2.6) | 1.2 | (0.7, 2.7) | 1.0 | (0.5, 2.1) |
| ICFbF | 3.1 | (1.3, 8.1) | 3.3 | (1.1, 8.0) | 3.3 | (1.2, 12.1) | 5.7 | (1.4, 14.5) |
| FA | 2.7 | (1.0, 13.8) | 2.8 | (1.1, 7.1) | 2.7 | (0.9, 6.1) | 3.2 | (1.0, 14.0) |
| IC | 3.3 | (1.3, 8.5) | 3.5 | (1.2, 7.8) | 2.9 | (1.1, 5.4) | 4.2 | (1.1, 14.2) |

Table 6.5: Performance comparison of Gauss-Newton segmentation algorithms for similarity transformation. The measure is the approximation error computed as the mean distance between landmarks $d_{\text{mean}}$ (pixel): Median (95% CI). ICFbF: Inverse compositional frame by frame. FA: Forward additive whole cycle. IC: Inverse compositional whole cycle. Frames with errors larger than 15 pixel were removed as divergent segmentations. Images were scaled down by a factor 0.25 to a size of 64x90 pixel.

| | 2C | | 3C | | 4C | | SAX | |
|---|---|---|---|---|---|---|---|---|
| Baseline | 5.5 | (2.8, 13.2) | 5.5 | (2.5, 10.5) | 4.9 | (2.8, 10.9) | 3.9 | (1.9, 8.5) |
| ICFbF | 13.0 | (5.2, 37.3) | 14.0 | (5.2, 46.3) | 12.8 | (4.3, 49.6) | 20.4 | (4.9, 57.7) |
| FA | 10.8 | (4.4, 37.6) | 10.3 | (3.8, 26.9) | 10.7 | (3.8, 43.3) | 13.6 | (5.5, 47.5) |
| IC | 11.8 | (5.7, 27.2) | 11.3 | (5.1, 22.5) | 11.1 | (4.0, 24.2) | 19.2 | (8.7, 57.3) |

Table 6.6: Performance comparison of Gauss-Newton segmentation algorithms. As Table 6.5, but images at full scale were used, with a size of 256x360 pixel.

# CHAPTER 7

# Conclusions and future work

## 7.1 Conclusions

This thesis has explored the opportunities and challenges of 2D+t contrast echocardiography for Left Ventricle (LV) functional analysis, both clinically and within a computer vision deformable template model framework. Similarly to the introduction in Ch. 1, the rest of this section presents a summary of findings and conclusions, organised by chapter for better readability.

### 7.1.1 Data

A database was built with 21 studies of Power Modulation contrast dobutamine stress echo in all 4 principal planes, with clinical variables, human expert hand-traced myocardial contours and visual scoring. The initial decision to include 10 normal and 11 abnormal patients provided a more realistic variety of case studies, but in retrospect, it would have been more useful to have more data to build a normal model, and then study abnormalities as a deviation from the normal. In any case, given the small prevalence of abnormal

segments in the database (each abnormal patient contributes some abnormal segments only), it can be assumed that the models are a rough approximation of normal physiology.

### 7.1.2 Clinical evaluation

Quantification using standard measures of global (Ejection Fraction) endocardial function showed expected values, and good agreement with human expert visual scoring. A possible explanation is that, even though it was hypothesised that perfusion would highlight the muscle, in fact shadows and removal of tissue signal by Power Modulation make the external wall invisible in large regions. From the study of patients with outlying values of Ejection Fraction, the following obstacles for functional analysis were conjectured: shifts in the interrogation plane and the heart going out of plane, heterogeneity of segment displacement, dragging of ischemic tissue by healthy tissue, insufficient LV opacification, and inadequacy of the segment scoring scale for dyskinetic segments. The amount of data was insufficient to draw strong conclusions from local functional analysis. Measures from endocardial wall motion (Fractional Area Change) compared reasonably well to visual scoring, although worse than Ejection Fraction. On the other hand, local function computed from wall thickening was found to be unreliable. It was contended that the lack of external wall visibility may lead to hand tracing errors large enough to preclude Power Modulation contrast echocardiography as an appropriate technique to measure myocardial thickening. Further research is needed to substantiate this claim.

### 7.1.3 Cardiac segmentation and deformable models

It was argued that segmentation of myocardial boundaries on the contrast echocardiographic cine loops should combine knowledge from 3 sources: texture, geometry and kinetics. Texture and geometry were studied in this chapter, while kinetics were left for Ch. 5.

Texture analysis suggested left-skewed distribution of intensities with a variance that

increases with the median value, and clipped off tails on both ends, with little change at a coarser resolution.

Interpolating cubic splines control points as landmarks and pseudo-landmarks were found to be a convenient representation of geometry, and for building Principal Component Analysis (PCA) shape spaces. The variance criterion was found to be inappropriate for dimensionality computation; the Information Criteria family could be used to evaluate the correct number of landmarks for modelling, while I proposed a criterion based on anatomical and physiological considerations as a more appropriate method for computing a dimensionality useful within a hierarchical scheme. A Gaussianisation method was proposed and shown to improve the compactness of the 2D shape model. Statistical atlases (mean and standard deviation intensity values) for Power Modulation contrast echography were computed for the first time. Those atlases illustrate the lack of visibility of the external myocardial bounday, and agree with my previous observation that the external wall is invisible in large regions of the image.

Active Appearance Model (AAM) were discussed as atlas-based deformable models extended with a texture space. Criticism of the components of AAMs (intensity standardisation, combined heterogeneous variables, correlation and covariance matrix, etc.) can be found summarised at the end of the chapter. But it is worth noting here that the AAM was more divergent and less accurate than atlas-based segmentation, and this could be due to important information for the segmentation process being removed by the PCA texture space in the AAM. Experiments also suggested the unfeasibility of computing appropriate PCA texture models for AAMs.

### 7.1.4 Temporal extension of cardiac PCA shape models

A novel spatio-temporal model of cardiac contours was proposed to integrate kinetics into the deformable model. The new explicit model does not require frame interpolation, and was shown to be more compact than previous implicit ones when the shape vector changes

from frame to frame. Results were similar when an *intrinsic shape* was assumed for the whole cardiac cycle, though, an indication that spatial and temporal variability components are not perfectly separated, and that the assumption that time can be expressed with a single frequency $\cos(2\pi t),\ \sin(2\pi t)$ is only a first approximation. Other sources of error were identified as lack of temporal alignment, and out of plane studies.

### 7.1.5 Atlas-based deformable model segmentation

Endocardial segmentation in contrast echocardiography 2D+t volumes was posed as an atlas-based segmentation problem combined with the explicit 2D+t model from the previous chapter, that can be tackled using a standard Gauss-Newton gradient descent framework. Classic forward additive algorithms were compared to the relatively recent inverse compositional approach. The results suggest that the forward additive algorithm is more reliable, and slightly more accurate, but 20 to 25 times slower than the inverse compositional approach, so the inverse compositional approach could be within real-time processing reach. Segmenting the whole cardiac cycle together in fact seems to produce substantially better results than segmenting frame by frame. Optimisation was performed on the similarity transformation alone for simplicity, which only allows for a rough first approximation, though. In addition, the 2D+t method models a normal LV, and this degrades the algorithm's performance when the patient is abnormal. Testing segmentation at full and a coarser scale, the results suggest that all the information relevant for optimisation of the similarity transformation is contained at the coarser scale, so using the full scale data is not necessary.

## 7.2 Further work

By posing the segmentation problem in a standard deformable template model framework, years of development in computer vision and biomedical engineering can be explored for improvements. In this section, a few ideas are advanced to continue the line of work

proposed in this thesis.

There are several specific improvements that are needed for the deformable model. First, the deformable model needs to be extended with a local warp, so that segmentation can be refined for each patient. Some background and preliminary work are provided in Appendices C and G as a starting point. Optimal model dimensionality in terms of the dimensionality of the PCA shape model would be a useful study. Second, extra flexibility needs to be added to the model so that it can express abnormalities. Third, the explicit model currently maps displacement on a single frequency component ($\cos(2\pi t)$, $\sin(2\pi t)$), while Figs. 5.1 and 5.3b suggest that higher frequency terms are required. Fourth, the difference between intrinsic dimensionality and the dimensionality from an anatomical criterion needs to be better understood, specially in terms of the effect of intra- and inter-frame correlation mentioned in section 4.15. This could possibly explain the results from the comparison between implicit and explicit 2D+t models in section 5.6, and help to better predict the data set size as a function of the number of variables. Dimensionality and data set size are problems that could be critical in higher order data sets, i.e. 3D+t echocardiography, as the number of landmarks increases sharply. Finally, more experiments with different data sets and implementations are required to confirm or reject the claims that atlas-based segmentation performs better than Active Appearance Models.

An obvious point of interest is better texture modelling. The mean intensity value is not a sufficient statistic of the skewed distributions shown in section 4.3. Intensities could be gaussianised, e.g. as described in Appendix E, or atlases built from other statistics, e.g. median or mode, as proposed in section 4.16. In addition, texture needs to be modelled with respect to time, as now each point in the sampling mask reflects the average over the whole cardiac cycle. Intensities in points near the base have a strong temporal component as the mitral valve opens and closes. Speckle tracking is an alternative with potential, as the interference pattern is ideally fixed for each anatomical location on the myocardium. An application would be to identify areas in the image without any significant texture. This way, the segmentation algorithm could be modified to ignore residuals in those areas,

and at the same time provide a measure of confidence in the segmentation results, i.e. whether a segment of the contour has been placed on a certain position based on local image information, or by interpolation from adjacent results.

Further clinical work can be undertaken too. A problem highlighted for local function evaluation was segment displacement heterogeneity, that hinders the distinction between normal and abnormal segments. With the 2D+t model, though, it could be possible to correct the wall motion in each segment by a scale factor, and look for a common measure of local function abnormality. Alternatively, the patient's wall motion for each segment could be compared to the value that the 2D+t model assumes to be normal. Both studies would have clinical interest for diagnosis. Another study could confirm the results obtained in this thesis that suggest that the external myocardial boundary cannot be reliably found in Power modulation contrast echocardiography images. For this, hand-tracings from more experts would be necessary, to evaluate the inter- and intra-subject variability, and how it affects local function diagnosis.

# APPENDIX A

## Quamus quadratic approximating splines

This appendix presents in detail the formulation for myocardial Quamus quadratic approximating splines, mentioned in section 2.4, p. 27. Let $\Phi$ be the planar closed oriented space curve or contour that we want to trace, and let $\phi : \mathbb{R} \mapsto \mathbb{R}^2$ be a parameterisation of $\Phi$ in 2D Cartesian coordinates

$$\phi(t) = \big(\phi_1(t),\ \phi_2(t)\big) \tag{A.1}$$

When using quadratic closed approximating splines, $\phi$ is a piece-wise continuous function with $K+1$ pieces and $K$ control points $c_1,\ \ldots,\ c_K$ that lie outside the contour. The $k$-th piece $\phi_k$ is

$$\phi_k(s_k) = \left(\frac{1}{2} - s_k + \frac{1}{2}s_k^2\right) c_a + \left(\frac{1}{2} + s_k - s_k^2\right) c_b + \left(\frac{1}{2}s_k^2\right) c_c \tag{A.2}$$

where $c_a,\ c_b,\ c_c$ are 3 consecutive control points

$$a = k \tag{A.3}$$

$$b = \mathrm{mod}(k, K) + 1 \tag{A.4}$$

$$c = \mathrm{mod}(k + 1, K) + 1 \tag{A.5}$$

$s_k$ is the piece-wise parameterisation variable

$$s_k = \frac{t_k - T_k}{T_{k+1} - T_k}, \quad t_k \in [T_k,\ T_{k+1}],\ s_k \in [0,\ 1] \tag{A.6}$$

where $t_k \in [0,\ T_{K+1}]$ is the spline parameterisation variable and $T = [T_1,\ \ldots,\ T_{K+1}]^T$ is the *knot* vector such that $\phi(T_1) = \phi(T_{K+1})$. The relationship between control points and the knot vector is illustrated in Fig. A.1. Quamus uses uniform spacing for the knot vector

$$T_i = i - 1, \quad 1 \leq i \leq K + 1 \tag{A.7}$$

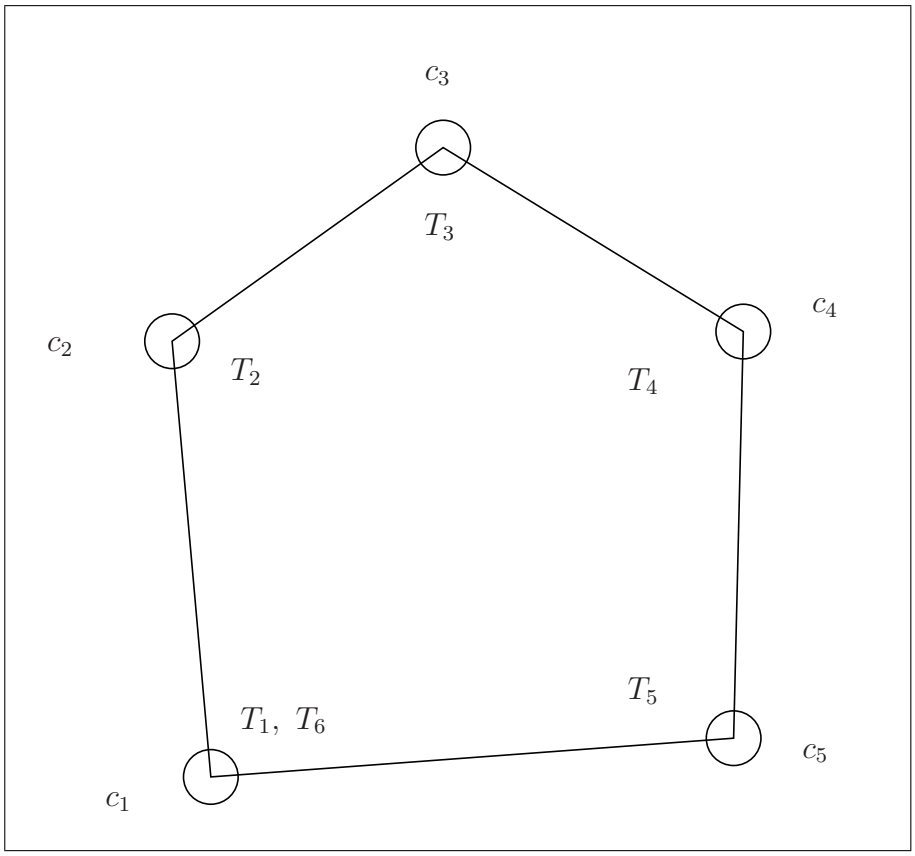



Figure A.1: Correspondence between control points and knot vector for $K = 5$, where $K$ is the number of control points $c$, and $T$ are the spline parameterisation variable knots.

# APPENDIX B

# Generalised Procrustes analysis

This appendix presents pseudo-code for Rohlf and Slice [146] Least-Squares Fit Generalised Orthogonal Procrustes Analysis (LSFGOPA) method, with a correction to avoid oscillations, and a reformulation of some operations to speed them up, as mentioned in section 4.4, p. 64.

## B.1 Corrected LSFGOPA

1. Set convergence tolerance tol = $10^{-4}$

2. Compute consensus as in (4.2)

3. Centre all configurations

$$x_i := x_i - \bar{x} \tag{B.1}$$

4. Normalise all configurations

$$x_i := \frac{x_i}{\|x_i\|} \tag{B.2}$$

5. Chose as consensus the first configuration

$$\bar{x} = x_i \tag{B.3}$$

6. Compute the rotation matrix $H_i$ for each configuration with (B.13)

7. Rotate each configuration

$$x_i := x_i H_i \tag{B.4}$$

8. Update consensus with (4.2)

9. Compute initial Sum of Squares Error (SSE)

$$sse0 = M\big(1 - \mathrm{tr}(\bar{x}\bar{x}^T)\big) \tag{B.5}$$

$$sse = \infty \tag{B.6}$$

10. Set initial weight factors

$$\rho_i = 1 \tag{B.7}$$

11. If $|sse0 - sse| \leq \mathrm{tol}$ then the algorithm has converged, otherwise continue

12. $sse0 = sse$

13. Compute rotation matrix for each configuration and rotate to fit consensus as above

14. Update consensus with (4.2)

15. Compute weight ratio for each configuration

$$\frac{\rho_i^\star}{\rho_i} = \left| \sqrt{\frac{\mathrm{tr}(x_i^\star \bar{x}^T)}{\mathrm{tr}(x_i^\star x_i^{\star T})\mathrm{tr}(\bar{x}\bar{x}^T)}} \right| \tag{B.8}$$

16. Update configurations and weight ratios

$$x_i := x_i \rho_i^\star / \rho_i \tag{B.9}$$

$$\rho_i := \rho_i^\star \tag{B.10}$$

17. Update consensus with (4.2)

18. Compute SSE

$$sse = M\big(1 - \mathrm{tr}(\bar{x}\bar{x}^T)\big) \tag{B.11}$$

19. Go to step 11

## B.2 Rotation matrix

Pseudocode to compute the rotation matrix to fit the configurations $x_i$ to the consensus $\bar{x}$:

1. Compute the Singular Value Decomposition (SVD) of $\bar{x}^T x_i$

$$U \cdot S \cdot V^T = \bar{x}^T x_i \tag{B.12}$$

2. Compute the rotation matrix

$$H_i = V \cdot \mathrm{sign}(S) \cdot U^T \tag{B.13}$$

## B.3 Note on the weight ratio

Rohlf and Slice [146] original formulation of (B.8) was

$$\frac{\rho_i^{\star 2}}{\rho_i^2} = \frac{\mathrm{tr}(x_i^\star \bar{x}^T)}{\mathrm{tr}(x_i^\star x_i^{\star T})\mathrm{tr}(\bar{x}\bar{x}^T)}$$

This can produce complex weight ratios and the algorithm oscillates. With (B.8) this problem is solved. Next I prove that the weighting condition used by [146] still holds

$$\begin{aligned}
\sum_i \mathrm{tr}(x_i^{\star\star} x_i^{\star\star T}) &= \sum_i \mathrm{tr}\left(\frac{\rho_i^{\star 2}}{\rho_i^2} x_i^\star x_i^{\star T}\right) \\
&= \sum_i \frac{\rho_i^{\star 2}}{\rho_i^2} \mathrm{tr}(x_i^\star x_i^{\star T}) \\
&= \sum_i \left(\left|\sqrt{\frac{\mathrm{tr}(x_i^\star \bar{x}^T)}{\mathrm{tr}(x_i^\star x_i^{\star T})\mathrm{tr}(\bar{x}\bar{x}^T)}}\right|\right)^2 \mathrm{tr}(x_i^\star x_i^{\star T}) \\
&= \sum_i \frac{\mathrm{tr}(x_i^\star \bar{x}^T)}{\mathrm{tr}(x_i^\star x_i^{\star T})\mathrm{tr}(\bar{x}\bar{x}^T)} \mathrm{tr}(x_i^\star x_i^{\star T}) \\
&= \sum_i \frac{\mathrm{tr}(x_i^\star \bar{x}^T)}{\mathrm{tr}(\bar{x}\bar{x}^T)} \\
&= \frac{1}{\mathrm{tr}(\bar{x}\bar{x}^T)} \sum_i \sum_j \sum_k x_i(j,k)\bar{x}(j,k) \\
&= \frac{N}{\mathrm{tr}(\bar{x}\bar{x}^T)} \sum_j \sum_k \bar{x}(j,k)\bar{x}(j,k) \\
&= \frac{N\mathrm{tr}(\bar{x}\bar{x}^T)}{\mathrm{tr}(\bar{x}\bar{x}^T)} \\
&= N
\end{aligned} \tag{B.14}$$

# APPENDIX C

## Local warp

Local warps are non-rigid transformation of the image domain. In this thesis, local warps were used to compute texture histograms in section 4.3, p. 58, and probabilistic atlases in section 4.8, p. 80. Local warps are also a component of the segmentation algorithm presented in Ch. 6, p. 129. For simplicity, the algorithm was illustrated with experiments restricted to a similarity transformation. This appendix provides a discussion of two types of local warps: piece-wise affine and thin-plate splines, that can be used as the basis for future work.

Stegmann [168] proposed using a piece-wise affine warp on a Delaunay triangulation, which is locally linear. On the one hand, these warps are very fast if computational geometry algorithms, cached operations and hardware acceleration are used [172]; but on the other hand, they do not scale well to 3D, the deformation field is not continuous, and the Delaunay triangulation may produce triangles outside the shape and long triangles that are undesirable for the piece-wise affine approximation (Fig. C.1). Matthews and Baker [111] found that composition of transformations in the optimisation algorithm using piece-wise affine warps is hard to define. As mentioned in section 4.8, deformable model

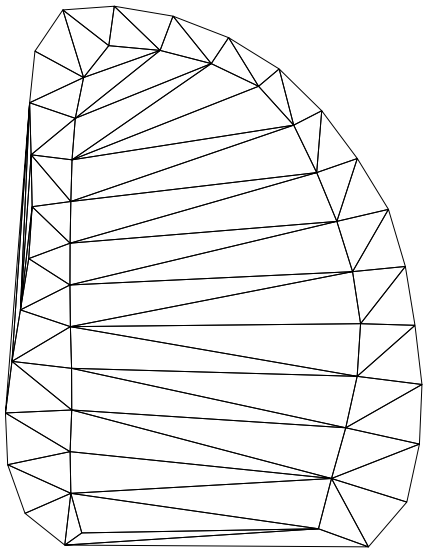


Figure C.1: Delaunay triangulation of 2C template.

segmentation is believed to be more successful if a band around the object is included in the template. Thus, a new set of artificial landmarks needs to be created at a certain distance from the shapes. Stegmann [168] measured distances along the normals to landmark points, but this method is sensitive to errors due to landmark sparsity, regions of high curvature and non-bijectivity of the distance measure between contours (cf. Fig. 4.7a). Stegmann introduced the artificial landmarks into the shape model; but apart from the error sources mentioned above, artificial landmarks are not linearly correlated to or do not describe anatomical landmarks.

Interpolating thin-plate splines (TPSs) were introduced to the field of image registration by Bookstein [25]. A TPS is a function $W(z;\ X_0,\ X)$ where the map between two sets of landmarks $X_0 \mapsto X$ defines the deformation on the coordinates of a data point $z$. In the particular case of template matching using the PDM

$$s(b) = \bar{s} + V\,b \tag{C.1}$$

then $X_0$ is the mean shape set of landmarks. In the following, it will be convenient to alternate between the shape vector $s$ and 2-column matrix $X$ notation for sets of

landmarks

$$s = \begin{bmatrix} X_{11} & \dots & X_{1P} & X_{21} & \dots & X_{2P} \end{bmatrix}^\top \tag{C.2}$$

A formulation of TPS warps appropriate for PDM template matching is

$$W_L(z;\ \bar{s},\ s(b)) = \begin{bmatrix} w_1 & \dots & w_P & a_1 & a_z \end{bmatrix} \begin{bmatrix} u(\|z - \bar{x}_1\|) \\ \vdots \\ u(\|z - \bar{x}_P\|) \\ 1 \\ z \end{bmatrix} \tag{C.3}$$

where $z,\ \bar{x}.,\ w.,\ a_1$ are 2-vectors, $a_z$ is a $(2,\ 2)$ matrix, and $u(r) = r^2 \log r = 1/2 r^2 \log r^2$. The coefficients $w.,\ a.$ of the TPS are computed using

$$\begin{bmatrix} U_{\bar{x}} & 1 & \bar{X}^\top \\ 1 & 0 & 0 \\ \bar{X} & 0 & 0 \end{bmatrix} \begin{bmatrix} w_1^\top \\ \vdots \\ w_P^\top \\ a_1^\top \\ a_z^\top \end{bmatrix} = \begin{bmatrix} x_1^\top \\ \vdots \\ x_P^\top \\ 0 \\ 0 \end{bmatrix} \tag{C.4}$$

$U_{\bar{x}}$ is the matrix where $U_{ij} = u(\|\bar{x}_i - \bar{x}_j\|)$, and $\bar{X}$ is the matrix of source points where the $i$-th column is $\bar{x}_i$. TPSs have the following characteristics:

- Unlike piece-wise affine warps, TPSs define a smooth continuous deformation field.
- TPS formulation is simple (as seen above), and the Jacobian can be computed analytically (see section G.4.2).
- TPSs are diffeomorphic (i.e. differentiable and with a differentiable inverse) as long as they are bijective, i.e. when they do not present folding[1]. Warps that guarantee a

[1]Eriksson and Åström [63, 64] proposed sufficient quadratic constraints to ensure bijection, but it is not clear how to compute the constraints in practice.

diffeomorphism in all cases constitute an active field of research, e.g. [3, 4, 33, 109, 181], but their formulation and computation is complex and beyond the scope of this thesis.

- The inverse of a TPS (provided it exists) is *not* a TPS, and hence TPS warps are not closed under inversion. In particular

$$W_L^{-1}(z;\ \bar{s},\ s) \neq W_L(z;\ s,\ \bar{s}) \tag{C.5}$$

except at the landmarks

$$W_L^{-1}(\bar{s};\ \bar{s},\ s) = W_L(s;\ s,\ \bar{s}) \tag{C.6}$$

This has originated some confusion in the literature. For example, Johnson and Christensen [86] claimed that TPS interpolation 'does not define a consistent correspondence between the two images except at the landmark points'. But, in fact, Johnson and Christensen's [86] critique is misplaced, because TPS are diffeomorphic when there is no folding, and thus have an inverse and define a consistent correspondence between the two images. The problem is that lacking a closed formula for the inverse, a non-linear optimisation problem needs to be solved for each warped point.

- TPS warps are *not* closed under composition. So, even if $W_L^{-1}$ was a TPS, the parameter update would only be an approximation

$$s(p_{i+1}) = W_L(z;\ \bar{s},\ s(p_i)) \circ W_L^{-1}(\bar{s};\ \bar{s},\ s(\Delta p_i)) \tag{C.7a}$$

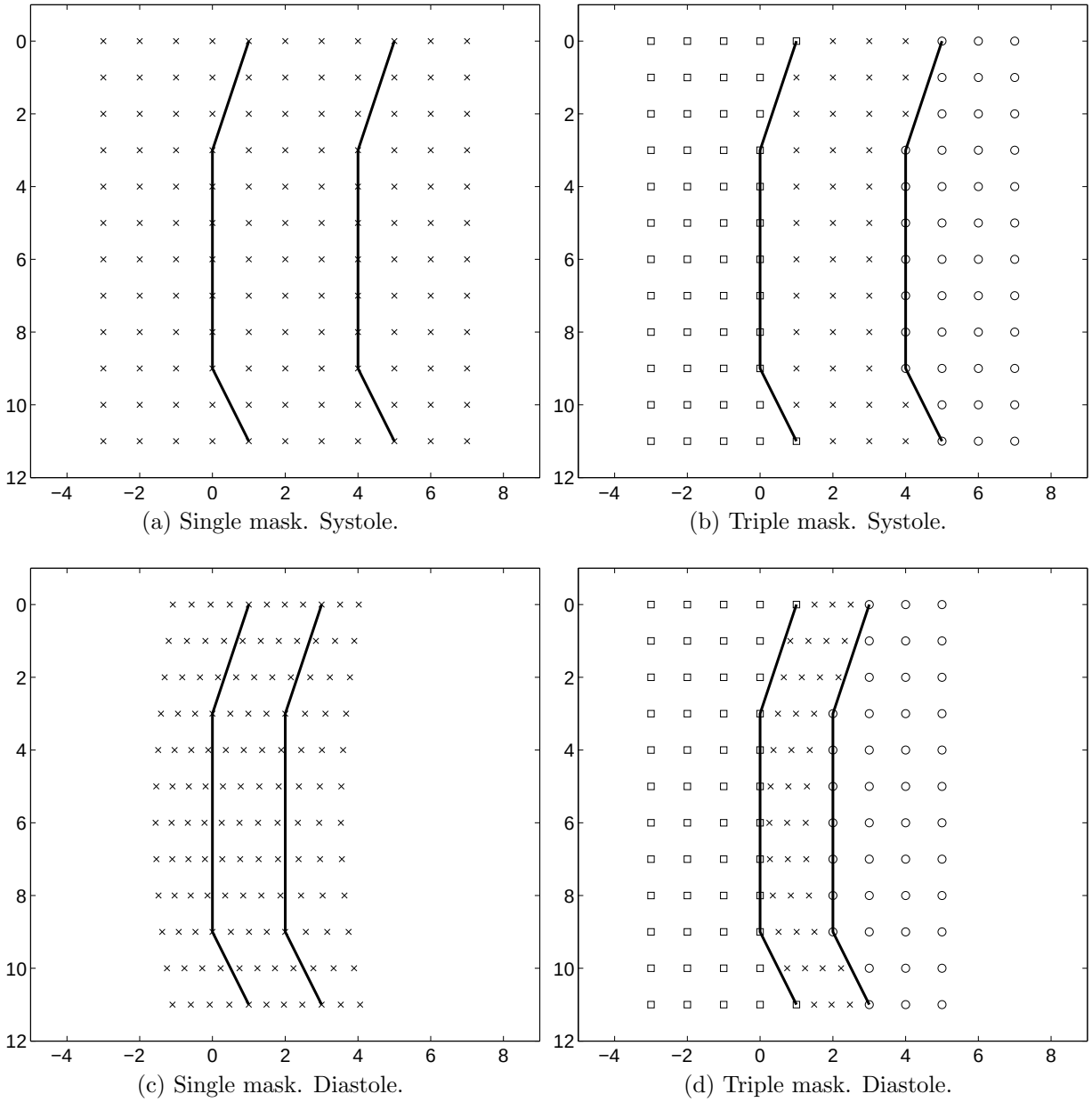
$$W_L(z;\ \bar{s},\ s(p_{i+1})) \approx W_L(z;\ \bar{s},\ s(p_i)) \circ W_L^{-1}(z;\ \bar{s},\ s(\Delta p_i)) \tag{C.7b}$$

- TPSs have global support, as opposed to e.g. linear combinations of B-splines (used to create a Free-Form Deformation model for registration of 2D cardiac tagged-MRI

and CT [79, 117], and 3D+t cardiac MRI [140]), that have compact support. For warps with global support, local landmark displacement impacts the whole image, while for warps with local support, the impact is limited to a neighbourhood. In neither case the deformation caused by the warp necessarily corresponds to physiology. Cardiac segmentation has the added difficulty that whether local or global, landmarks cannot contribute to the warp of every template sampling point in the same way, as shown by Fig. C.2. For warps with global support, an advantage of TPSs over other functions is that they minimise the bending energy of the interpolant, and so the global deformation; see e.g. in [182] that global deformation with GISs is noticeably larger than with TPSs.

- On the other hand, a drawback of TPS's global support is that the system matrix in Eq. C.4 is not sparse. Because of the absence of sparsity, direct methods to compute TPSs require $\mathcal{O}(P^3)$ operations. In addition, computing logarithms in $u$ is slow. Computational complexity has been cited as a reason to not use TPSs (e.g. [168, sec. 5.3.1]). Fast implementation methods have been suggested (e.g. [29, Ch. 7]). One of the most popular is the Fast Multipole Algorithm [11, 29, 143], that reduces the complexity to $\mathcal{O}(P \log P)$ at set-up and $\mathcal{O}(\log P)$ at evaluation, but both with a large constant that makes the method only useful for large numbers of landmarks and data points. My own experimentation with this method suggests that for the data in this thesis the Fast Multipole Algorithm is 1 to 2 orders of magnitude slower than the naive implementation.

- Speed considerations above notwithstanding, the TPS warp formulation in Eqs. (C.3) and (C.4) enables the implementation in Matlab without any loops.



(a) Single mask. Systole.

(b) Triple mask. Systole.

(c) Single mask. Diastole.

(d) Triple mask. Diastole.

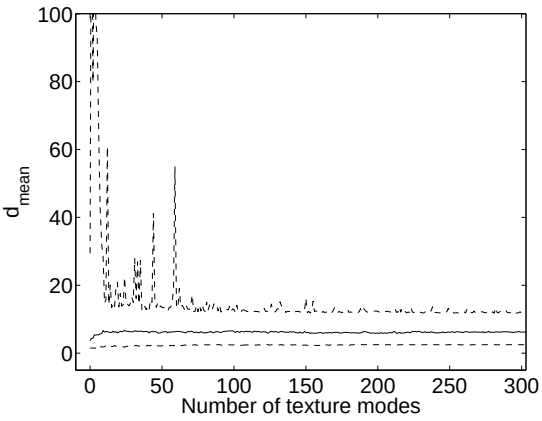
Figure C.2: Need for partition of template sampling points into 3 sets: external (□), myocardial (×) and blood pool (∘). Left column shows deformation defined from both contours for all sampling points. Right column shows 'triple mask', i.e. the warp for □ and ∘ depends only on the epicardium and endocardium, respectively, while × are warped using both contours, as usual.

To illustrate the difference, the endocardium moves towards the epicardium, that remains stationary. Note that there are two problems with the left column: 1) Even though the epicardium is stationary, and in principle so is the outside of the heart, the external sampling points are warped by the endocardium. 2) Sampling points in the blood pool are warped the *wrong* way, i.e. instead of filling the blood pool in diastole, they retreat closer to the endocardium.

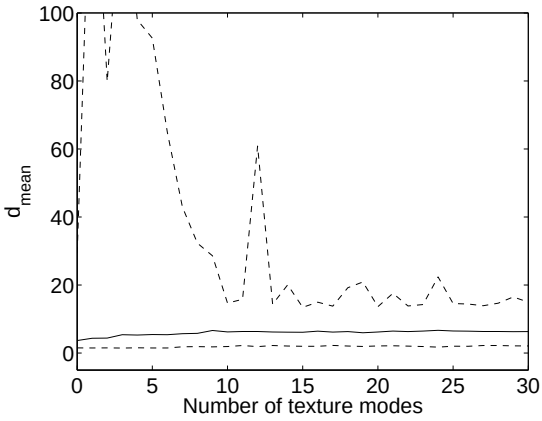
Masks can be easily partitioned in Matlab e.g. using the signed distance function implementation of Persson and Strang [141], and the KD Tree Nearest Neighbour and Range Search Toolbox of Michael [118].

# APPENDIX D

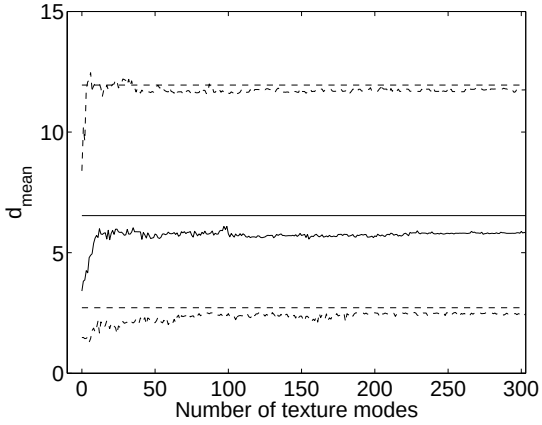
Segmentation error vs. dimensionality figures



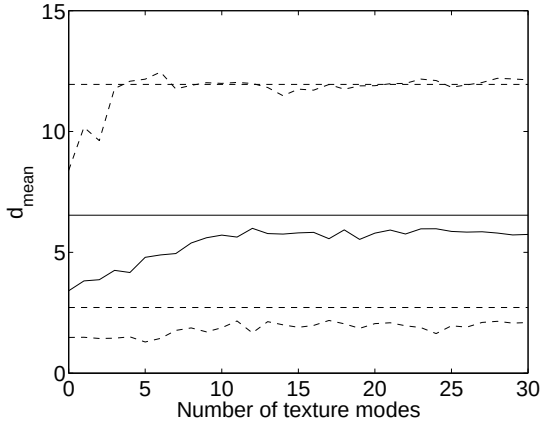

(a) Convergent and divergent segmentations.




(b) Convergent and divergent segmentations (zoom in).

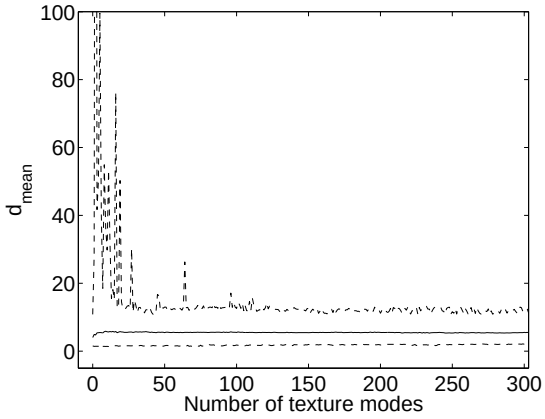



(c) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixels.

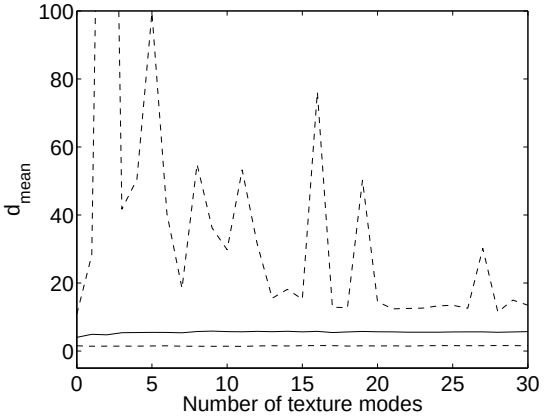



(d) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixels (zoom in).

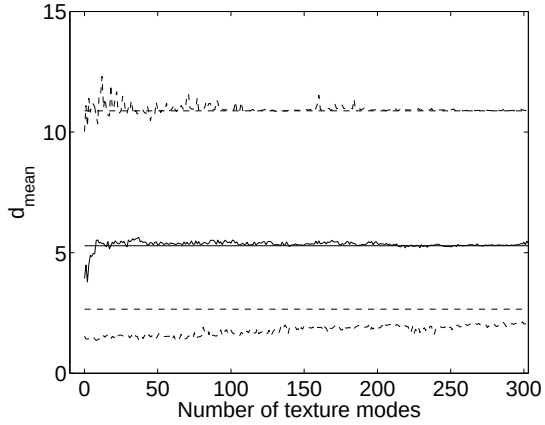
Figure D.1: Segmentation error vs. dimensionality of texture space (see Fig. 4.12), 2C view.



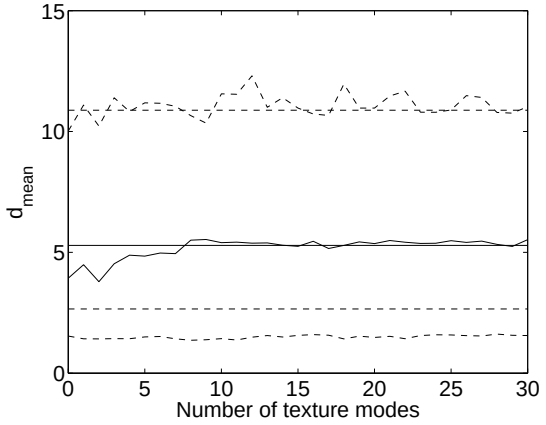

(a) Convergent and divergent segmentations.




(b) Convergent and divergent segmentations (zoom in).

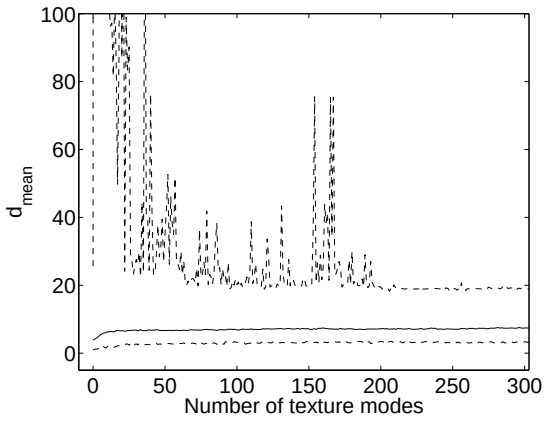



(c) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixels.

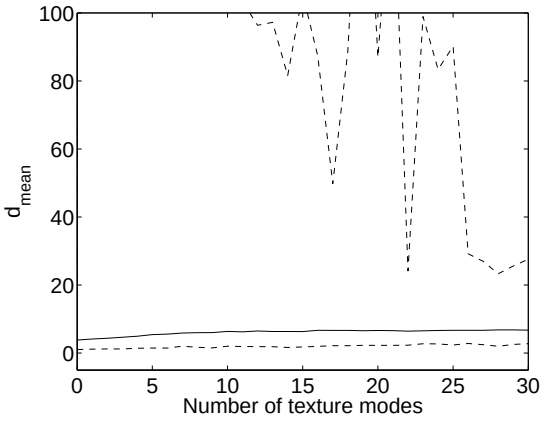



(d) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixels (zoom in).

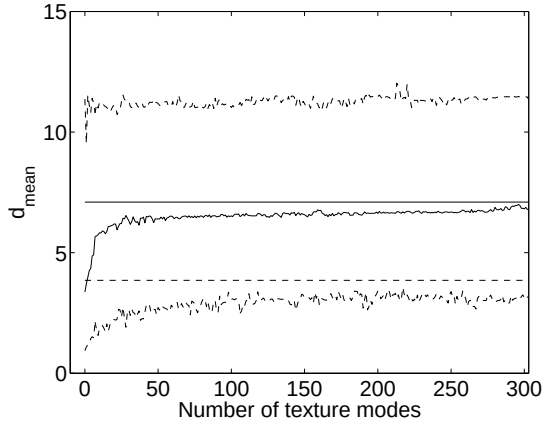
Figure D.2: Segmentation error vs. dimensionality of texture space (see Fig. 4.12), 4C view.



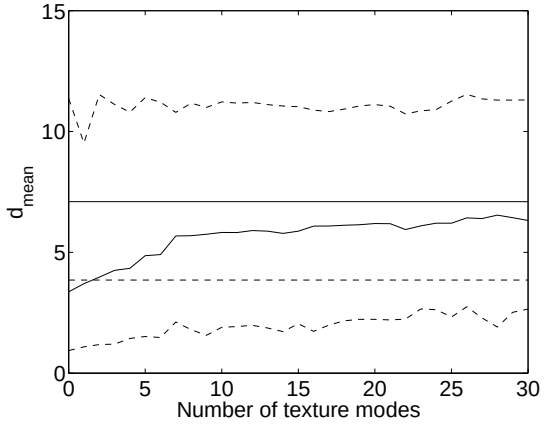

(a) Convergent and divergent segmentations.




(b) Convergent and divergent segmentations (zoom in).




(c) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixels.




(d) Remove failed segmentations, $d_{\text{mean}} > 13.04$ pixels (zoom in).

Figure D.3: Segmentation error vs. dimensionality of texture space (see Fig. 4.12), SAX view.

# APPENDIX E

## Gaussianising intensity vectors

Gaussianisation of intensity values is not used in the segmentation algorithm of Ch. 6, because my current implementation makes optimisation too slow. The method is presented in this appendix, as basis for future work.

Gaussianisation of shape variables for PCA models was presented in section 4.7, p. 76. A problem faced when gaussianising intensity values is dealing with underexposed and saturated areas, i.e. pixels with 0 and 255 values, respectively. Equation (4.30) assumes that $T$ is diffeomorphic, but this condition is not met when intensity values fall outside the dynamic range and are clipped. Fig. E.1a and E.2a illustrate the case when the sampling point falls in an often underexposed region of the image: there is a noticeable discontinuity at $F_g(0)$ that prevents (4.30) from actually gaussianising the data, as displayed in Fig. E.3a.

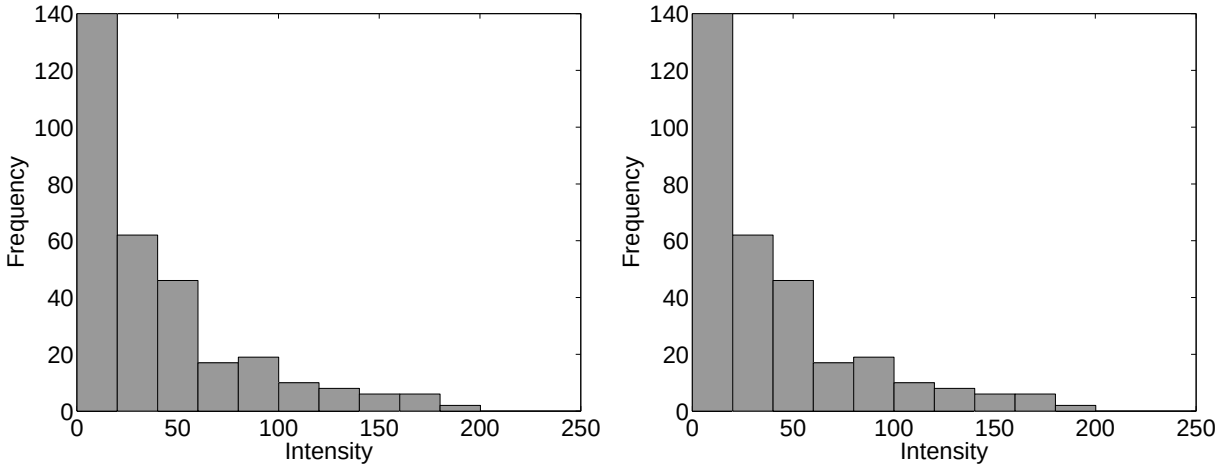
While the information in underexposed and saturated areas is lost and can not be truly recovered, attempts can be made to recover detail in those areas, i.e. to reconstruct the lost texture. As a naive first solution, a Gaussian filter was used to create a blurred copy of the image, and the problematic pixels were replaced with their blurred counterparts.

| | scale 1.00 | | scale 1.0 | | scale 0.25 | |
|---|---|---|---|---|---|---|
| | $\mathrm{sk}_e$ | $\mathrm{kr}_e$ | $\overline{\mathrm{sk}}$ | $\overline{\mathrm{kr}}$ | $\overline{\mathrm{sk}}$ | $\overline{\mathrm{kr}}$ |
| Ideal Gaussian | 0.00 | 3.00 | 0.00 | 3.00 | 0.00 | 3.00 |
| Original images | 1.45 | 4.71 | 1.09 | 4.50 | 0.95 | 4.05 |
| Images with recovered texture | 1.48 | 4.79 | 1.14 | 4.66 | 0.95 | 4.07 |
| Gaussianised original images | 0.38 | 2.54 | 0.23 | 2.66 | 0.07 | 2.74 |
| Gaussianised images with recovered texture | -0.09 | 2.93 | 0.05 | 2.82 | 0.05 | 2.78 |

Table E.1: Effect of gaussianisation.

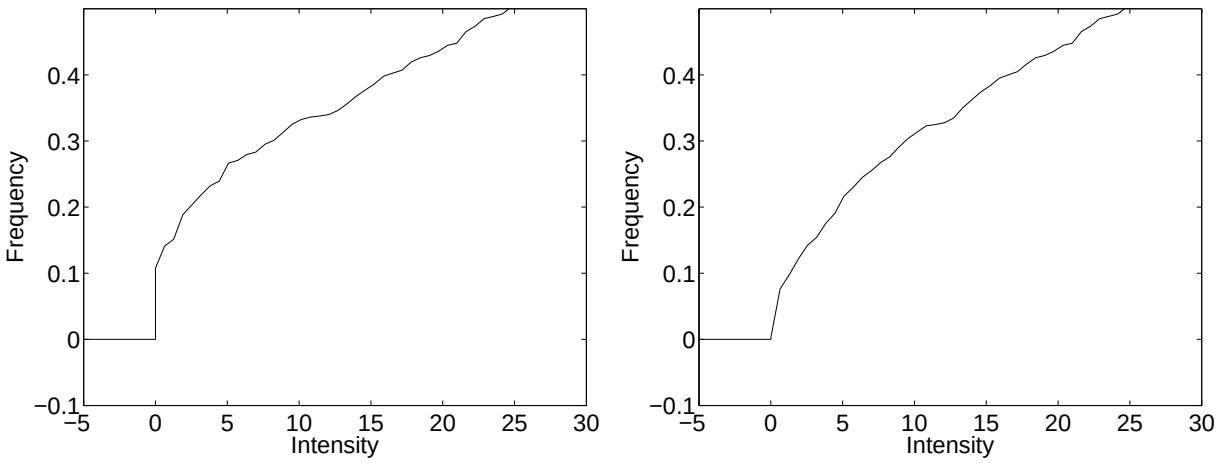
By trial and error and visual inspection, a square Gaussian filter of side $R' = \mathrm{round}(R/19)$ pixel, where $R$ is the number of rows in the image and $\sigma = R'/2$ was found to be a good compromise between removing extreme values while not introducing artifacts from nearby regions.

The recovery effect is barely apparent in the intensity histogram for a sampling point in a dark region Fig. E.1b, because most recovered pixel values will fall within the same histogram bin. It is more conspicuous, though, in the CDF of Fig. E.2b; the discontinuity at the origin has been smoothed out. Fig. E.3b shows the histogram of the Gaussianised recovered data, clearly more symmetric than Fig. E.3a. To quantify the improvement in Gaussianity, Table E.1 shows the skewness ($\mathrm{sk}_e$) and kurtosis ($\mathrm{kr}_e$) for the example above (a sampling point at scale 1 in a dark region), the average ($\overline{\mathrm{sk}}$, $\overline{\mathrm{kr}}$) for all sampling points in the data set at scale 1.00, and the average for all sampling points at scale 0.25.



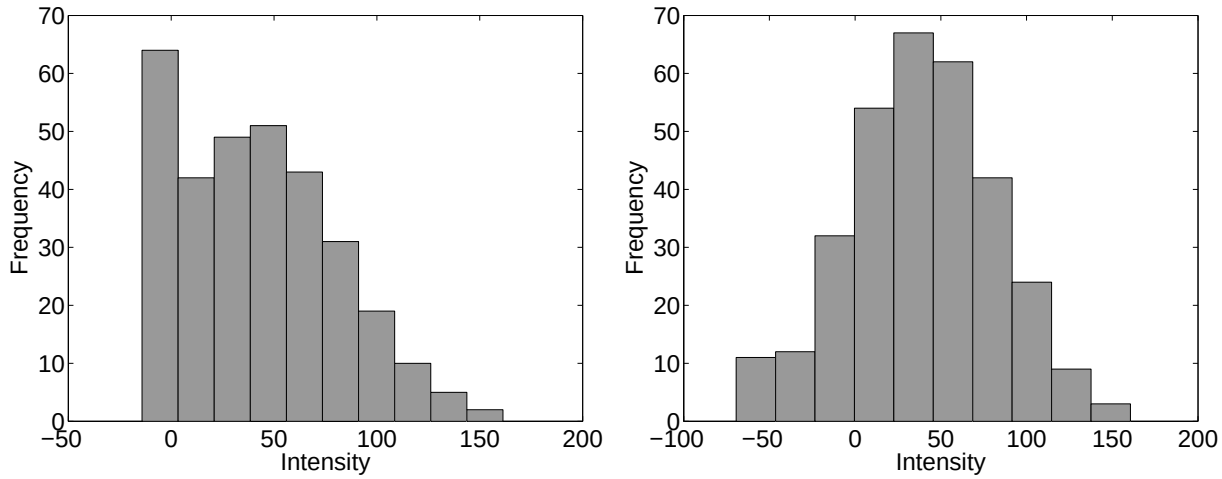

(a) Original images. (b) Images with recovered texture.

Figure E.1: Histograms of intensity values for a sampling point in a dark region.




(a) Original images. (b) Images with recovered texture.

Figure E.2: Detail of empirical CDF of a sampling point from an often underexposed region.




(a) Gaussianised original images.

(b) Gaussianised images with recovered texture.

Figure E.3: Histograms of the data from Fig. E.1 gaussianised.

# APPENDIX F

# Kernel PCA

## F.1 Overview

Kernel PCA (KPCA) is a non-linear generalisation of PCA proposed by Schölkopf et al. [154]. The temporal extension of the shape model proposed in Ch. 5, p. 111, draws heavily from KPCA and Multidimensional Scaling (MDS) ideas, but in order to not break the thesis flow, KPCA is presented in this appendix for reference. In addition, some experiments are run on the approximation error of texture KPCA models. To be able to compute the approximation error, the theory of how to map from input to feature space and vice versa is discussed, also to highlight that it is not a trivial problem.

There are 2 main ideas behind KPCA: First, non-linear relationships between variables in *input space* $\mathbb{R}^n$ can be linearised in *feature space* $\mathcal{F}$ using a non-linear function $\phi : \mathbb{R}^n \mapsto \mathcal{F}$. This is a generalisation of basic data linearisation in Statistics.

Second, even if $\mathcal{F}$ is a very high or infinite dimensional space, there is no need to compute feature vectors explicitly. The reason is twofold: 1) Thanks to the MDS formulation seen in section 4.5, the eigenvectors can be computed from the $(M,\ M)$-*kernel matrix*, or

| | |
|---|---|
| Polynomial | $k(s_i,\ s_j) = (s_i^\top s_j)^q$ |
| Gaussian | $k(s_i,\ s_j) = \exp\left(-\frac{\|s_i - s_j\|^2}{2\sigma^2}\right)$ |
| Inhomogeneous polynomial | $k(s_i,\ s_j) = (s_i^\top s_j + c)^q,\ c \geq 0$ |

Table F.1: Examples of kernels.

matrix of inner products also called *Gram matrix* (e.g. [150]), instead of the usually much larger covariance matrix. 2) The kernel matrix can be computed directly from the input space data thanks to the *kernel trick* [154]

$$\langle \phi(s_i),\ \phi(s_j) \rangle = k(s_i,\ s_j) \tag{F.1}$$

if $k$ is positive-definite (with this property, $k$ is called a *Mercer kernel*). Examples of kernels are shown in Table F.1. The formulation for KPCA is the same as PCA, generalising

$$V = \tilde{S} A' \Lambda'^{1/2} \tag{F.2a}$$

$$\tilde{k}_{i,j} = \tilde{s}_i^\top \tilde{s}_j \tag{F.2b}$$

where $\tilde{k}_{i,j}$ is the $(i,\ j)$ element of the centred *kernel matrix* $\tilde{K}$, to

$$V = \tilde{\Phi} A' \Lambda'^{1/2} \tag{F.3a}$$

$$\tilde{k}_{i,j} = \langle \phi(s_i),\ \phi(s_j) \rangle \tag{F.3b}$$

Hence, PCA can be seen as a particular case of KPCA with a polynomial kernel of order $q = 1$. Note that in general $\tilde{k}_{i,j} \neq k(\tilde{s}_i,\ \tilde{s}_j)$. Matrix $\tilde{K}$ can be computed as[1] [154]

$$\tilde{K} = K - \breve{1}_M K - K \breve{1}_M + \breve{1}_M K \breve{1}_M \tag{F.4}$$

[1] In theory, $\tilde{K}$ is symmetric, but in practice and due to finite precision errors in (F.4) it is not. This produces complex eigenvalues and eigenvectors. A simple solution is $\tilde{K} := (\tilde{K} + \tilde{K}^\top)/2$.

where $\breve{1}_M$ is an $(M, M)$-matrix with each element equal to $1/M$ and $K$ is the kernel matrix with elements $k_{ij} = k(s_i,\ s_j)$. The PDM of Eq. (4.3) is formulated in KPCA as

$$\phi(s) = \bar{\phi}(s) + Vb \tag{F.5}$$

The coefficient vector $b$ can be computed without having to explicitly compute $V$ or $\phi$ noting that $V^{-1} = V^\top$, and using Eq. (F.3a)

$$b = V^{-1}\left(\phi - \bar{\phi}\right)^\top = V^\top \tilde{\phi} = A^\top \tilde{\Phi}^\top \tilde{\phi} = A^\top \begin{bmatrix} \tilde{k}(s_1,\ s) \\ \vdots \\ \tilde{k}(s_M,\ s) \end{bmatrix} \tag{F.6}$$

If $V$ had some columns removed, then the formulation is the same replacing the inverse by the pseudo-inverse. To compute $b$ in terms of uncentred vectors it is useful to write Eq. (F.4) in vector form

$$\tilde{k}(s_i,\ s) = \tilde{\phi}_i^\top \tilde{\phi} = k(s_i, s) - 1_M^\top \Phi^\top \phi - \phi_i^\top \Phi 1_M + 1_M^\top K 1_M \tag{F.7}$$

Using Eq. (F.7) in Eq. (F.6)

$$\begin{aligned} b &= A^\top \left\{ \begin{bmatrix} k(s_1, s) \\ \vdots \\ k(s_M, s) \end{bmatrix} - 1_1\, 1_M \Phi^\top \phi - \Phi^\top \Phi 1_M + 1_1\, 1_M^\top K 1_M \right\} \\ &= A^\top \left\{ \begin{bmatrix} k(s_1, s) \\ \vdots \\ k(s_M, s) \end{bmatrix} - \breve{1}_M \begin{bmatrix} k(s_1, s) \\ \vdots \\ k(s_M, s) \end{bmatrix} - K 1_M + \breve{1}_M K 1_M \right\} \\ &= A^\top \left(\mathbb{I} - \breve{1}_M\right) \left( \begin{bmatrix} k(s_1, s) \\ \vdots \\ k(s_M, s) \end{bmatrix} - K 1_M \right) \end{aligned} \tag{F.8}$$

where $1_M$ is a vector with each element equal to $1/M$.

For the moment, I am going to follow the conventional application of KPCA, using the MDS approach and typical kernels. However, I would like to note that MDS is independent of the idea of linearisation; it is simply a computational trick. In fact, when time is integrated into the shape model in Ch. 5, p. 111, the corresponding non-linear transformation in Eq. 5.5, p. 115 allows to operate directly in feature space.

## F.2 Choosing parameter $\sigma$ for Gaussian kernels

It is generally accepted that for Gaussian kernels the value of the $\sigma$ parameter is important (e.g. [2, 52, 53]). Different proposed estimates for this value are (by [2], [53] and [52], respectively)

$$\sigma = \overline{d_{\mathbb{R}^n,NN}} \tag{F.9a}$$

$$\sigma = \sqrt{\overline{d^2_{\mathbb{R}^n,NN}}} \tag{F.9b}$$

$$\sigma = 1.5\sqrt{\overline{d^2_{\mathbb{R}^n,NN}}} \tag{F.9c}$$

where $d_{\mathbb{R}^n,NN}(s_i)$ is the Euclidean distance of the $i$-th training vector to its nearest neighbour.

The MDS formulation is useful to understand the effect of $\sigma$. The kernel matrix is the collection of inner products between training vectors, which is a measure of the distance or dissimilarity among them. The value of $\sigma$ determines the size of the local neighbourhood; small (large) values of $\sigma$ mean that few (many) training vectors are considered similar. Eq. (F.9a)-(F.9c) use the smallest possible $\sigma$ so that local neighbourhoods are restricted to each training vector and its nearest neighbours.

Lafon et al. [95] argued that different Gaussian kernels should be used for building the eigenvector space and for finding the pre-image[2]: a small $\sigma$ for better learning the

[2]In fact, Lafon et al. [95] studied the out-of-sample extension problem, but Arias et al. [2] showed the close relation with the pre-image problem.

'geometry of the underlying structure of the data set', and a large $\sigma$ to allow the model to generalise. This is in line with the *learning approach* of Bakir et al. [7, 8] that requires 2 kernels, but multi-kernel methods are beyond the scope of this thesis. For single-kernel methods I propose defining $\sigma$ as a compromise between describing local structure and generalising,

$$\sigma_{\text{Casero}} = \frac{\bar{d}_{\mathbb{R}^n}}{\sqrt{-2\ln(0.5)}} \approx 0.8493\overline{d_{\mathbb{R}^n}} \tag{F.10}$$

i.e. 2 vectors at mean distance have a kernel value $k = 0.5$. The effect of $\sigma$ can be discussed in terms of the kernel matrix: Small $\sigma$ values create quasi-identity kernel matrices, while large values create kernel matrices with all values close to 1 (i.e. local structure in the space is lost). This is illustrated in Fig. F.1a, where an apparently small change in $\sigma$ (0.05 to 0.35) completely changes the configuration of the kernel matrix. Fig. F.1b shows the distribution of kernel values for the 4 different $\sigma$ values presented in this section.

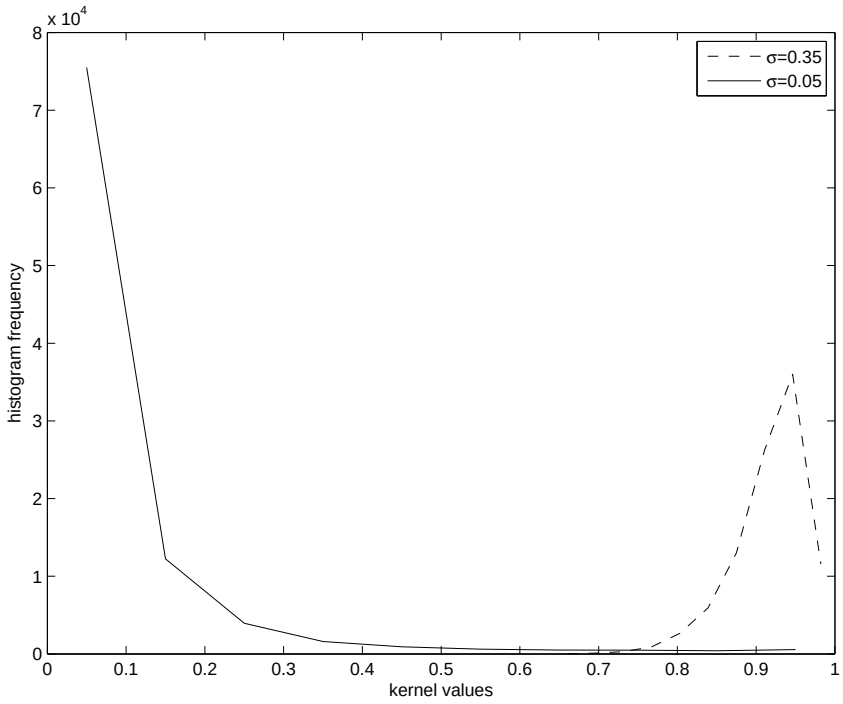
## F.3 The pre-image problem in KPCA

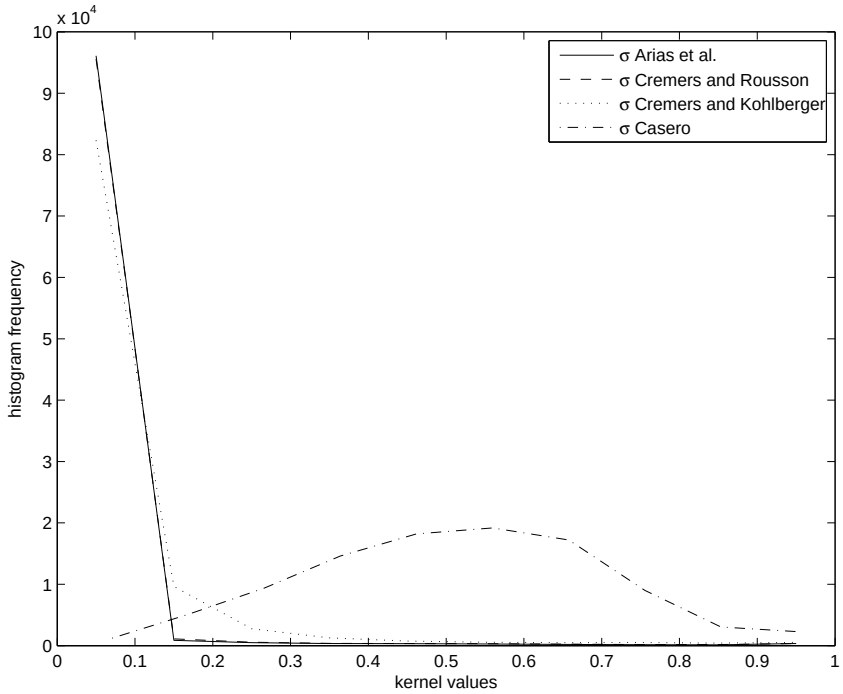
It might now seem trivial to extended PCA-based models using nonlinear functions. However, while each *pre-image* $s$ has a feature coefficient vector $b$, the opposite is not true in general [156, 157], i.e. $\nexists\ s \in \mathbb{R}^n$ such that

$$s = \phi^{-1}(\phi) \tag{F.11}$$

even if $\phi^{-1}$ exists. Eq. (F.11) is known as the *pre-image problem*. Burges [30] proposed relaxing the pre-image problem in the context of Support Vector Machines (SVMs)

$$s_\star \approx \phi^{-1}(\phi) \tag{F.12}$$




(b) $\sigma_{\text{Arias et al.}} = 0.0297$, $\sigma_{\text{Cremers and Rousson}} = 0.0318$, $\sigma_{\text{Cremers and Kohlnerger}} = 0.0478$, $\sigma_{\text{Casero}} = 0.1191$

Figure F.1: Relationship between kernel matrix values and $\sigma$ for Gaussian kernels. 20 patients data set (336 shape vectors with 100 variables), 4C, image size scaled by a factor of 0.25, Gaussian kernel. The curves represent centres of histogram bins created from the values in the respective kernel matrices.

to minimise the distance measure

$$\rho = \|\phi - \phi_\star\|^2 \tag{F.13}$$

Schölkopf et al. [156] proposed changing the distance to minimise to

$$\rho = \|P_M\phi - \phi_\star\|^2 \tag{F.14}$$

where $P_M\phi$ is the projection of $\phi$ onto the subspace spanned by the feature eigenvectors $\{v_i\}_{i=1}^M$. In AAMs, the pre-image problem is formulated as finding $s_\star$ given a feature coefficient vector $b$, i.e. $\phi$ is implicitly expressed as a linear combination of $\{v_i\}_{i=1}^M$, so $\phi = P_M\phi$.

Schölkopf et al. [156] proposed gradient descent to minimise Eq. (F.14) for radial basis function (e.g. Gaussian or spline kernels). Schölkopf et al. [155] and Mika et al. [119] proposed a fixed-point iteration method to implement gradient descent. For Gaussian kernels the iteration is [145]

$$s_\star(t+1) = \frac{\sum_{i=1}^M \tilde{\gamma}_i \exp\left(-\frac{\|s_\star(t)-s_i\|^2}{2\sigma^2}\right) s_i}{\sum_{i=1}^M \tilde{\gamma}_i \exp\left(-\frac{\|s_\star(t)-s_i\|^2}{2\sigma^2}\right)} \tag{F.15}$$

where $\{\tilde{\gamma}_i\}_{i=1}^M$ are the components of the vector $\tilde{\gamma}$. With this method, the objective function is nonlinear and not convex, so it suffers from local minima; besides, Eq. (F.15) tends to be numerically unstable and needs several starting points [2, 8]. But with the approximation $P_M\phi \approx \phi_\star$, Eq. (F.15) can be computed in a single iteration; for instance, for Gaussian kernels [145]

$$s_\star = \frac{\sum_{i=1}^M \tilde{\gamma}_i \left(2 - d^2_{\mathcal{F}}(\phi_\star,\ \phi_i)\right) s_i}{\sum_{i=1}^M \tilde{\gamma}_i \left(2 - d^2_{\mathcal{F}}(\phi_\star,\ \phi_i)\right)} \tag{F.16}$$

where $d^2_{\mathcal{F}}$ is the distance in feature space

$$\begin{aligned} d^2_{\mathcal{F}}(\phi_\star,\ \phi_i) &= \|\phi_\star - \phi_i\|^2 \\ &= \|\phi_\star\|^2 - 2\phi_\star^\top \phi_i + \|\phi_i\|^2 \end{aligned} \tag{F.17}$$

Formulae to compute $d^2_{\mathcal{F}}$ in the literature depend on $s_\star$ [8, 94, 145], which is the unknown variable. To find an expression that can be computed directly from known variables, 3 results are needed. The first one is

$$\begin{aligned} \|\phi_\star\|^2 &= \|\bar{\phi} + Vb\|^2 \\ &= \|\bar{\phi}\|^2 + 2\bar{\phi}^\top Vb + \|Vb\|^2 \\ &= \|\Phi 1_M\|^2 + 2\, 1_M^\top \Phi^\top \tilde{\Phi} Ab + \|b\|^2 \\ &= 1_M^\top \Phi^\top \Phi 1_M + 2\, 1_M^\top \Phi^\top \left(\Phi - \bar{\phi} 1_1^\top\right) Ab + \|b\|^2 \\ &= 1_M^\top K 1_M + 2\, 1_M^\top \Phi^\top \Phi \left(\mathbb{I} - \breve{1}_M\right) Ab + \|b\|^2 \\ &= 1_M^\top K 1_M + 2\, 1_M^\top K \left(\mathbb{I} - \breve{1}_M\right) Ab + \|b\|^2 \\ &= 1_M^\top K \big(1_M + 2\left(\mathbb{I} - \breve{1}_M\right) Ab\big) + \|b\|^2 \\ &= 1_M^\top K \big(2\tilde{\gamma} - 1_M\big) + \|b\|^2 \end{aligned} \tag{F.18}$$

where

$$\tilde{\gamma} = 1_M + (\mathbb{I} - \breve{1}_M) Ab \tag{F.19}$$

The second result is

$$\begin{aligned}
\phi_\star^\top \phi_i &= \left(\bar{\phi} + Vb\right)^\top \phi_i \\
&= \left(1_M^\top \Phi^\top + b^\top A^\top \tilde{\Phi}^\top\right) \phi_i \\
&= \left(1_M^\top \Phi^\top + b^\top A^\top (\mathbb{I} - \breve{1}_M) \Phi^\top\right) \phi_i \\
&= \tilde{\gamma}^\top \Phi^\top \phi_i \\
&= \tilde{\gamma}^\top \begin{bmatrix} k(s_1, s_i) \\ \vdots \\ k(s_M, s_i) \end{bmatrix}
\end{aligned} \tag{F.20}$$

The third result is

$$\|\phi_i\|^2 = \phi_i^\top \phi_i = k(s_i, s_i) \tag{F.21}$$

Using Eq. (F.18)-(F.21) a formula to compute $d_{\mathcal{F}}^2$ for all kernels is obtained

$$\begin{aligned}
d_{\mathcal{F}}^2(\phi_\star,\ \phi_i) =& 1_M^\top K\left(2\tilde{\gamma} - 1_M\right) + \|b\|^2 \\
& - 2\tilde{\gamma}^\top \begin{bmatrix} k(s_1, s_i) \\ \vdots \\ k(s_M, s_i) \end{bmatrix} \\
& + k(s_i, s_i)
\end{aligned} \tag{F.22}$$

The single iteration approximation can be extended to other kernels. For example, polynomial homogeneous kernels [145] ($d_{\mathcal{F}}^2$ can be computed using the same formula Eq. (F.22) for all kernels)

$$s_\star = \sum_{i=i}^{M} \tilde{\gamma}_i \left(\frac{\|\phi_\star\|^2 + k(s_i, s_i) - d_{\mathcal{F}}^2(\phi_\star,\ \phi_i)}{2\|\phi_\star\|^2}\right)^{\frac{d-1}{d}} s_i \tag{F.23}$$

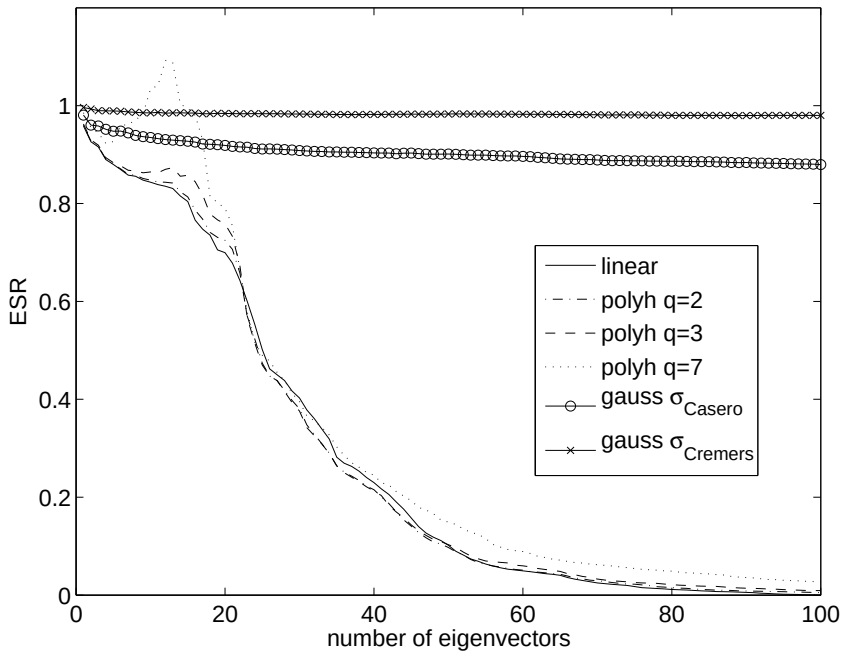
Using Eqs. (F.18) and (F.22) in Eq. (F.23)

$$s_\star = \left(1_M^\top K\left(2\tilde{\gamma} - 1_M\right) + \|b\|^2\right)^{-\frac{q-1}{q}} \sum_{i=i}^{M} \tilde{\gamma}_i \left( \tilde{\gamma}^\top \begin{bmatrix} k(s_1, s_i) \\ \vdots \\ k(s_M, s_i) \end{bmatrix} \right)^{\frac{q-1}{q}} s_i \tag{F.24}$$

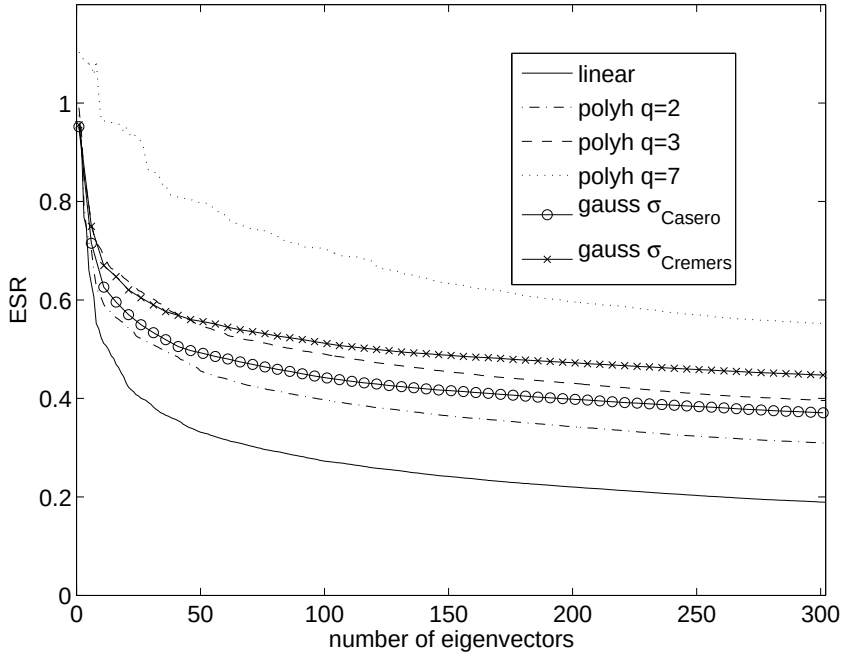
To assess the performance of KPCA, linear PCA was compared to the 2 most popular types of kernel, polynomial and Gaussian. Fig. F.2a and F.2b display the mean leave-one-out approximation error for gaussianised shape, and intensity, respectively, of 21 patients in the 3C view. The approximation error was computed with a similar measure to the one in Eq. (4.7)

$$ESR = \frac{\|s - s'_k\|^2}{\|s\|^2} \tag{F.25}$$

where $s$ is the data vector, $s'_k$ is $s$ projected onto the first $k$ eigenvectors of the shape space and back to input space. The results suggest that polynomial and Gaussian kernels offer no advantage over linear PCA for these data, and in general worsen the performance of the model.




(a) Shape.




(b) Texture, scale 0.2.

Figure F.2: Approximation error for homogeneous polynomial (polyh) and Gaussian KPCA. $\sigma_{\text{Casero}}$ is defined in Eqs. (F.10) and $\sigma_{\text{Cremers}}$ in (F.9c). The ESR is defined in Eq. (F.25).

# APPENDIX G

# Implementation details of the segmentation algorithm

## G.1 Introduction

This appendix presents the implementation details for the Gauss-Newton solution to the inverse compositional algorithm segmentation problem posed in Eq. (6.4), p. 130. The part of the implementation that corresponds to the similarity transformation was used in the experiments of Ch. 6. In addition, the formulation for the similarity transformation plus thin-plate splines scheme is presented, as basis for future work.

## G.2 Gauss-Newton method

The Gauss-Newton solution to a least-squares optimisation problem can be found in numerical optimisation books (e.g. [132]). In the rest of this section, the solution applied to the inverse compositional algorithm formulation is outlined, following [111], as the notation will be useful in subsequent sections. The first order Taylor expansion on $\Delta p$ of the

objective function in Eq. (6.4) is

$$\frac{1}{2}\sum_{z}\left(T(z)+\frac{\partial T}{\partial p}(z;\ 0)\,\Delta p - I(W(z;\ p))\right)^2 \tag{G.1}$$

where the steepest descent image is the $k$ row vector

$$\frac{\partial T}{\partial p}(z;\ 0) = \nabla^\top T\,\frac{\partial W}{\partial p}(z;\ 0) \tag{G.2}$$

and the gradient $\nabla T$ of the template is the 2-vector, for each pixel $z$,

$$\nabla T = \left[\frac{\partial T}{\partial z(1)},\ \frac{\partial T}{\partial z(2)}\right]^\top \tag{G.3}$$

The derivative of Eq. (G.1) with respect to $\Delta p$ is

$$\sum_{z}\frac{\partial T}{\partial p}^\top(z;\ 0)\left(T(z)+\frac{\partial T}{\partial p}(z;\ 0)\,\Delta p - I(W(z;\ p))\right) \tag{G.4}$$

Equating to 0 and rearranging terms, the optimal parameter update or *search direction* at the $i$-th iteration is the 4-vector

$$\Delta p_i = -H^{-1}(0)\sum_{z}\frac{\partial T}{\partial p}^\top(z;\ 0)\,e_{T,I}(z;\ 0,\ p_{i-1}) \tag{G.5}$$

where the residual for a pixel $e_{T,I}(z;\ p_{i-1})$ is

$$e_{T,I}(z;\ 0,\ p_{i-1}) = T(z) - I(W(z;\ p_{i-1})) \tag{G.6}$$

$H$ is the Gauss-Newton approximation to the $(k,\ k)$ Hessian matrix

$$\begin{aligned} H(\Delta p) &\approx \sum_{z}\frac{\partial T}{\partial p}^\top(z;\ \Delta p)\frac{\partial T}{\partial p}(z;\ \Delta p) \\ &= \Upsilon^\top(\Delta p)\Upsilon(\Delta p) \end{aligned} \tag{G.7}$$

| | Pre-computation | Iteration |
|---|---|---|
| $H^{-1}\left(\Upsilon^\top e_{T,I}\right)$ | 0 | $\begin{cases} kL + k^2 & \text{mults.} \\ 2k^2 & \text{sums} \end{cases}$ |
| $\left(H^{-1}\,\Upsilon^\top\right) e_{T,I}$ | $\begin{cases} k^2L & \text{mults.} \\ k^2L & \text{sums} \end{cases}$ | $\begin{cases} kL & \text{mults.} \\ k^2 & \text{sums} \end{cases}$ |

Table G.1: Number of operations (multiplications and sums) to compute $\Delta b$ in Eq. (G.8).

$\Upsilon(\Delta p)$ is an $(L,\ k)$-matrix. The descent direction can be written in matrix form as

$$\Delta p_i = -H^{-1}(0)\Upsilon^\top(0)\, e_{T,I}(0,\ p_{i-1}) \tag{G.8}$$

where abusing the notation, the residual vector for the template is

$$e_{T,I}(0,\ p_{i-1}) = [e_{T,I}(z_1;\ 0,\ p_{i-1}),\ \ldots,\ e_{T,I}(z_L;\ 0,\ p_{i-1})]^\top \tag{G.9}$$

Table G.1 shows that by pre-computing $(H^{-1}\Upsilon^\top)$, $k^2$ multiplication and $k^2$ sum operations are saved at each iteration of the algorithm. The trade-off is an extra $k^2L$ multiplications and $k^2L$ sums at pre-computation time. For a single run of the algorithm, the trade-off is worthy only after $L$ iterations, a very unlikely situation for typical values of $L$. However, pre-computations can be performed off-line and saved for every run of the algorithm; thus, pre-computing $(H^{-1}\,\Upsilon^\top)$ would be advantageous in a clinical scenario.

## G.3 Line search method

Gradient descent optimisation algorithms are usually combined with line search methods. These are methods that compute the step length along the search direction at iteration $i$ of the algorithm. In the additive formulation (e.g. [132, Ch. 3])

$$p_i = p_{i-1} + \alpha_i\, \Delta p_i \tag{G.10}$$

In the inverse compositional algorithm, $p$ is not the $p$ in Eq. (G.10). Thus, to avoid confusion, Eq. (G.10) is reformulated as

$$\Delta \dot{p}_i = \alpha_i \, \Delta p_i \tag{G.11}$$

(Matthews and Baker [111] assumed unit step length, $\alpha_i = 1, \ \forall i$, for their inverse compositional algorithm.) Full featured line search methods are hard to code, so it is recommended to use public domain implementations [132, p. 68]. Following that recommendation, line search in this thesis is based on the cubic polynomial method of the Matlab Optimization Toolbox. The actual implementation in the Optimization Toolbox does not correspond to the official documentation [110, p. 5-12], so implementation details are provided in Appendix I. Below, only one of the scenarios in the full implementation is presented for illustration purposes in Fig. G.1, but first, it is necessary to formulate the objective function and its derivative. Let $f$ be the objective function in Eq. (6.4)

$$f(\Delta p, \ p) = \frac{1}{2} \sum_z \Big( T(W(z; \ \Delta p)) - I(W(z; \ p)) \Big)^2 \tag{G.12}$$

At the $i$-th iteration, the line search tries to optimise

$$\phi(\alpha_i) = f(\alpha_i \, \Delta p_i, \ p_{i-1}) \tag{G.13}$$

In Fig. G.1, the solution at the current iteration is $\phi(0)$, and its directional derivative is always negative $\phi'(0)^\top \Delta p_i < 0$, where the objective function gradient is

$$\begin{aligned} f'(\Delta p, \ p) &= \sum_z \frac{\partial T}{\partial p}(z; \ \Delta p) \Big( T(W(z; \ \Delta p)) - I(W(z; \ p)) \Big) \\ &= \Upsilon^\top(\Delta p) \, e_{T,I}(\Delta p, \ p) \end{aligned} \tag{G.14}$$

The new descent direction $\Delta p_i$ is combined with the previous step length $\alpha_{i-1}$ to produce the new solution $\phi(\alpha_{i-1})$. The line search consists on finding a new step length $\alpha_i$ that

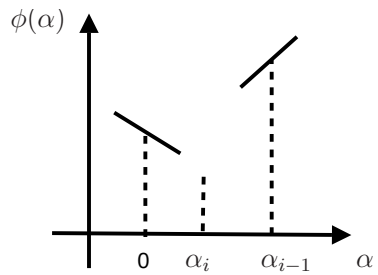



Figure G.1: Case 1 of the cubic polynomial interpolation line search scenarios. $\phi(\alpha) = f(\alpha\,\Delta p_i,\ p_{i-1})$.

provides a better solution. In this scenario, the new step length is $0 \leq \alpha_i \leq \alpha_{i-1}$. So the 5 values $\phi(0),\ \phi(\alpha_{i-1}),\ \phi'(0),\ \phi'(\alpha_{i-1}),\ \alpha_{i-1}$ are used to interpolate a cubic polynomial and find $\alpha_i$ at its minimum. In other scenarios, the new step length may need to be $\alpha_i > \alpha_{i-1}$. For implementation details on all scenarios see Appendix I.

There are alternative methods. Also implemented in Matlab [110, p. 5-15], and explained in e.g. [132, pp. 57–59], the mixed quadratic and cubic polynomial method replaces the computation of a gradient $\phi'(\alpha_{i-1})$ by another function evaluation. However, it will be shown in the following sections that analytical derivatives are available and fast to compute for similarity transformations and TPSs.

## G.4 The Gauss-Newton inverse compositional algorithm using line search

In this section, the Gauss-Newton inverse compositional algorithm proposed by Matthews and Baker [111] is simplified for similarity transformations and then the piece-wise affine warp of [111] is replaced by a TPS warp using a result from Eriksson and Åström [64]. The approximation in [111] to the inverse of the similarity transformation is replaced by an exact solution; an approximation to the inverse of the TPS warp is proposed too. Finally, a standard line search is added to the algorithm. For clarity, both variations of the method are summarised in separate tables in Algorithms G.1 and G.2.

**Algorithm G.1** Gauss-Newton inverse compositional algorithm using similarity transformations and line search

**Nomenclature**

Similarity transformation $W = W_G$

Parameters $p = q$

Number of parameters $k = 4$

**Set-up**

1: Compute template gradient $\nabla T$ using Eq. (G.3)
2: Compute Jacobian matrix $\partial W_G / \partial q$ of the transformation for each pixel using Eq. (G.16)
3: Compute steepest descent images $\partial T / \partial q$ pixel by pixel using Eq. (G.2)
4: Compute $(4,\ 4)$ Hessian matrix $H$ using Eq. (G.7)
5: Compute descent direction matrix $H^{-1}\Upsilon^{\top}$
6: Init transformation parameters, $q_0 = 0,\ \alpha_0 = 1$

$i$ **and** $i+1$ **iteration steps,** $i = 1,\ 3,\ 5,\ \ldots$

1: **repeat**
2: Compute residual vector $e_{T,I}(0,\ q_{i-1})$ as explained in section G.4.3
3: Compute $\phi(0),\ \phi'(0)$ using Eq. (G.40)
4: Compute descent direction $\Delta q_i$ using Eq. (G.8)
5: Compute directional derivative $\phi'(0)^{\top}\Delta q_i$
6: Set $\alpha_i = \alpha_{i-1}$
7: Update transformation parameters $q_i = \mathsf{update\_simil}(q_{i-1},\ \alpha_i \Delta q_i)$ as in Appendix H
8: Compute residual vector $e_{T,I}(0,\ q_i)$ as explained in section G.4.3
9: Compute $\phi(\alpha_{i-1}),\ \phi'(\alpha_{i-1})$ using Eq. (G.40)
10: Set $\Delta q_{i+1} = \Delta q_i$
11: Compute directional derivative $\phi'(\alpha_{i-1})^{\top}\Delta q_{i+1}$
12: Compute new step length $\alpha_{i+1}$ using Appendix I
13: Update transformation parameters $q_{i+1} = \mathsf{update\_simil}(q_{i-1},\ \alpha_{i+1}\Delta q_{i+1})$ as in Appendix H
14: **until** Termination condition in Appendix I

**Algorithm G.2** Gauss-Newton inverse compositional algorithm using similarity transformations, TPS warps and line search

**Nomenclature**

Similarity transformation and TPS warp $W = W_G \circ W_L$

Parameters $p$

Number of parameters $k = k_{\text{PDM}} + 4$

**Set-up**

1: Compute template gradient $\nabla T$ using Eq. (G.3)
2: Compute Jacobian matrix $\partial W/\partial p$ of the transformation/warp for each pixel using Eq. (G.25)
3: Compute steepest descent images $\partial T/\partial p$ pixel by pixel using Eq. (G.2)
4: Compute $(k,\ k)$ Hessian matrix $H$ using Eq. (G.7)
5: Compute descent direction matrix $H^{-1}\Upsilon^{\top}$
6: Init transformation parameters, $p_0 = 0,\ \alpha_0 = 1$
7: Compute inverse of template's TPS energy matrix $W_{L2}$ in (G.21)
8: Compute $g(z)$ for template grid's TPS in Eq. (G.21)

**$i$ and $i+1$ iteration steps,** $i = 1,\ 3,\ 5,\ \ldots$

1: **repeat**
2: Compute residual vector $e_{T,I}(0,\ p_{i-1})$ as explained in section G.4.4
3: Compute $\phi(0),\ \phi'(0)$ using Eq. (G.40)
4: Compute descent direction $\Delta p_i$ using Eq. (G.8)
5: Compute directional derivative $\phi'(0)^{\top}\Delta q_i$
6: Set $\alpha_i = \alpha_{i-1}$
7: Update transformation parameters $p_i = \mathsf{update\_parameters}(p_{i-1},\ \alpha_i \Delta p_i)$ as in Eq. (G.37)
8: Compute residual vector $e_{T,I}(0,\ p_i)$ as explained in section G.4.4
9: Compute $\phi(\alpha_{i-1}),\ \phi'(\alpha_{i-1})$ using Eq. (G.40)
10: Set $\Delta p_{i+1} = \Delta p_i$
11: Compute directional derivative $\phi'(\alpha_{i-1})^{\top}\Delta q_{i+1}$
12: Compute new step length $\alpha_{i+1}$ using Appendix I
13: Update transformation parameters $p_{i+1} = \mathsf{update\_parameters}(p_{i-1},\ \alpha_{i+1}\Delta p_{i+1})$ as in Eq. (G.37)
14: **until** Termination condition in Appendix I

### G.4.1 Evaluate Jacobian of the similarity transformation

The Jacobian of the similarity transformation $W = W_G$ evaluated at a point $z$ on the template is a $(2,\ 4)$-matrix

$$\frac{\partial W_G}{\partial q}(z;0) = \begin{bmatrix} \frac{\partial W_G(1)}{\partial q(1)} & \cdots & \frac{\partial W_G(1)}{\partial q(4)} \\ \frac{\partial W_G(2)}{\partial q(1)} & \cdots & \frac{\partial W_G(2)}{\partial q(4)} \end{bmatrix} \tag{G.15}$$

The term $\partial W_G/\partial q$ is a $(2,\ 4)$-matrix that can be easily computed from Eq. (6.8)

$$\frac{\partial W_G}{\partial q}(z;\ 0) = A_G(z) \tag{G.16}$$

### G.4.2 Evaluate Jacobian of the similarity transformation and TPS warp

The combined similarity transformation and TPS warp $W$ follows the formulation in Eq. (6.6). The Jacobian of the transformation evaluated at a point $z$ on the template is a $(2,\ k+4)$-matrix

$$\frac{\partial W}{\partial p}(z;0) = \left[\frac{\partial W}{\partial q},\ \frac{\partial W}{\partial b}\right](z;0) \tag{G.17}$$

where $W_G$ and $W_L$ are defined in Eqs. (6.8) and (C.3), respectively. The term $\partial W/\partial q = \partial W_G/\partial q$ as in Eq. (G.16). The term $\partial W/\partial b$ is a $(2,\ k)$-matrix computed using the chain rule

$$\frac{\partial W}{\partial b}(z;\ 0) = \frac{\partial W_G}{\partial W_L}\frac{\partial W_L}{\partial s}\frac{\partial s}{\partial b}(z;\ \bar{s},\ s(0)) \tag{G.18}$$

noting that $s(0) = \bar{s}$. From Eq. (6.8) the term $\partial W_G/\partial W_L$ is the $(2,\ 2)$-matrix

$$\frac{\partial W_G}{\partial W_L}(z;\ p) = \begin{bmatrix} 1+q(1) & -q(2) \\ q(2) & 1+q(1) \end{bmatrix} \tag{G.19}$$

Evaluating on the template,

$$\frac{\partial W_G}{\partial W_L}(z;\ 0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{G.20}$$

To compute the $\partial W_L / \partial s$ term, it is convenient to combine Eqs. (C.3) and (C.4) [64]

$$\begin{aligned} W_L(z;\ \bar{s},\ s(b)) &= \begin{bmatrix} x_1 & \dots & x_P & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_{\bar{X}} & 1 & \bar{X}^\top \\ 1 & 0 & 0 \\ \bar{X} & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} u(\|z - \bar{X}_1\|) \\ \vdots \\ u(\|z - \bar{X}_P\|) \\ 1 \\ z \end{bmatrix} \\ &= W_{L1}(b_2)\, W_{L2}(b_1)\, W_{L3}(z;\ b_1) \\ &= X\, g(z) \end{aligned} \tag{G.21}$$

where $g(z)$ is a $P$-vector. Differentiating Eq. (G.21) and evaluating in $p = 0$

$$\frac{\partial W_L}{\partial s}(z;\ \bar{s},\ s(0)) = \begin{bmatrix} g(z)^\top & 0 \\ 0 & g(z)^\top \end{bmatrix} \tag{G.22}$$

Differentiating, the term $\partial s / \partial b$ is

$$\frac{\partial s}{\partial b}(z) = V \tag{G.23}$$

Substituting Eqs. (G.20), (G.22) and (G.23) into Eq. (G.18)

$$\begin{aligned} \frac{\partial W}{\partial b}(z,\ 0) &= \begin{bmatrix} g(z)^\top & 0 \\ 0 & g(z)^\top \end{bmatrix} V \\ &= \begin{bmatrix} g(z)^\top V_1 \\ g(z)^\top V_2 \end{bmatrix} \end{aligned} \tag{G.24}$$

where $V = [V_1^\top \; V_2^\top]^\top$. Substituting Eqs. (G.16) and (G.24) into Eq. (G.17), the Jacobian of the transformation evaluated on the template can be written as

$$\frac{\partial W}{\partial p}(z;0) = \begin{bmatrix} A_G(z), & \begin{bmatrix} g(z)^\top V_1 \\ g(z)^\top V_2 \end{bmatrix} \end{bmatrix} \tag{G.25}$$

The Jacobian matrix has to be computed for all $L$ pixels of the template image. All vectors $g(z_i)$ can be efficiently computed using [64]

$$\begin{bmatrix} U_{\bar{X}} & 1 & \bar{X}^\top \\ 1 & 0 & 0 \\ \bar{X} & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} u(\|z_1 - \bar{x}_1\|) & \dots & u(\|z_L - \bar{x}_1\|) \\ \vdots & \ddots & \vdots \\ u(\|z_1 - \bar{x}_P\|) & \dots & u(\|z_L - \bar{x}_P\|) \\ 1 & \dots & 1 \\ z_1 & \dots & z_L \end{bmatrix} \tag{G.26}$$

where the $i$-th column is $g(z_i)$.

### G.4.3 Compute residuals for similarity transformation

For $W_G$ an alternative formulation is, e.g. [111],

$$W_G(z;\; q) = \begin{bmatrix} 1+q(1) & -q(2) \\ q(2) & 1+q(1) \end{bmatrix} z + \begin{bmatrix} q(3) \\ q(4) \end{bmatrix} \tag{G.27}$$

Template pixel coordinates are mapped onto the image using Eq. (G.27). Image $I$ is sampled at the new coordinates $W_G(z;\; q)$ using bilinear interpolation. The residuals are computed for each pixel in the template using Eq. (G.6).

### G.4.4 Compute residuals for similarity transformation and TPS warp

Target shape $s(b)$ at the current iteration is computed using the PDM in Eq. (C.1). The template grid (i.e. pixel coordinates) is warped using $W_L(z;\ \bar{s},\ s(b))$ in Eq. (G.21). This is a very fast operation, because $g(z)$ was precomputed from $W_{L2}W_{L3}$ in the set up phase of the algorithm for every pixel, and so, the warp is simply a matrix multiplication.

The warped grid is then mapped onto the image using the similarity transform $W_G(z;\ q)$ in Eq. (G.27). Image $I$ is sampled at the new coordinates $W(z;\ p)$ using bilinear interpolation. The residuals are computed for each pixel in the template using Eq. (G.6).

### G.4.5 Update transformation parameters for similarity transformation and TPS warp

The updated transformation was computed in [6] for general diffeomorphic transformations/warps

$$W(z;\ p_i) = W(z;\ p_{i-1}) \circ W^{-1}(z;\ \Delta p_i) \tag{G.28}$$

Matthews and Baker [111] computed a first order approximation to the inverse of the total transformation. In the case of TPSs, the approximation is

$$W^{-1}(z;\ \Delta p) \approx W(z;\ -\Delta p) \tag{G.29}$$

However, it is possible to be more precise. By the rules of composition, the inverse of the transformation is

$$W^{-1}(z;\ \Delta p) = W_L^{-1}(z;\ \bar{s},\ s(\Delta b)) \circ W_G^{-1}(z;\ \Delta q) \tag{G.30}$$

First, the inverse $W_G^{-1}$ of the similarity transformation is computed using Eq. (H.2), so there is no need for the first order approximation of [111]. Second, as the inverse of a TPS

(when it exists) is not a TPS, then $W_L^{-1}$ needs to be approximated. But instead of the first order approximation of [111], the following approximation is convenient for TPSs

$$W_L^{-1}(z;\ \bar{s},\ s(\Delta b)) \approx W_L(z;\ s(\Delta b),\ \bar{s}) \tag{G.31}$$

The approximation is exact at the landmarks, and degrades as data points move away from them. Plugging Eq. (G.31) into Eq. (G.30),

$$W^{-1}(z;\ \Delta p) \approx W_L(z;\ s(\Delta b),\ \bar{s}) \circ W_G^{-1}(z;\ \Delta q) \tag{G.32}$$

The only approximation is for the inverse of the TPS. As TPSs are invariant with respect to similarity transformations, then Eq. (G.32) can be written as[1]

$$W^{-1}(z;\ \Delta p) \approx W_G^{-1}(z;\ \Delta q) \circ W_L(z;\ s(\Delta b),\ \bar{s}) \tag{G.33}$$

The transformation update can be computed plugging Eq. (G.33) into Eq. (G.28), and using again the similarity transformation invariance of the TPS

$$W(z;\ p_i) = W_G(z;\ q_{i-1}) \circ W_G^{-1}(z;\ \Delta q_i) \circ W_L(z;\ \bar{s},\ s(b_{i-1})) \circ W_L(z;\ s(\Delta b_i),\ \bar{s}) \tag{G.34}$$

The inverse of a similarity transformation is a similarity transformation too, i.e. $W_G^{-1}$ represents a scaling, rotation and translation. In addition, the composition of the two similarity transformations is a similarity transformation

$$W_G(z;\ q_i) = W_G(z;\ q_{i-1}) \circ W_G^{-1}(z;\ \Delta q_i) \tag{G.35}$$

[1]Note: The similarity transformation is applied to $z$, $s(\Delta b)$ and $\bar{s}$, while the warp is applied only to $z$.

The updated parameters $q_i$ can be computed exactly as explained in Appendix H (but note the difference of notation). Thus, Eq. (G.34) is reduced to

$$W(z;\ p_i) = W_G(z;\ q_i) \circ W_L(z;\ \bar{s},\ s(b_{i-1})) \circ W_L(z;\ s(\Delta b_i),\ \bar{s}) \tag{G.36}$$

For the parameter update $p_i = [q_i^\top,\ b_i^\top]^\top$ the following notation will be used

$$p_i = \mathsf{update\_parameters}(p_{i-1},\ \alpha_i \Delta p_i) \tag{G.37}$$

The parameter update can be broken down into 3 steps (note the change of notation in Appendix H)

$$q_i = \mathsf{update\_simil}(q_{i-1},\ \alpha_i \Delta q_i) \text{ as in Appendix H} \tag{G.38a}$$

$$s_i = W_L(z;\ \bar{s},\ s(b_{i-1})) \circ W_L(\bar{s};\ s(\alpha_i \Delta b_i),\ \bar{s}) \tag{G.38b}$$

$$b_i = V^\top (s_i - \bar{s}) \tag{G.38c}$$

The TPS parameter update requires 2 warps applied to the mean shape, i.e. a small number of points:

- $W_L(\bar{s};\ s(\alpha_i \Delta b_i),\ \bar{s})$ is a new warp at each iteration, i.e. nor $W_{L2}$ neither $W_{L3}$ can be precomputed and cached; thus, this warp is the principal limiting factor to the maximum number of landmarks in the model

- $W_L(z;\ \bar{s},\ s(b_{i-1})) \circ W_L(\bar{s};\ s(\alpha_i \Delta b_i),\ \bar{s})$ allows $W_{L2}(b_1)$ to be precomputed an cached; as the number of landmarks is low given the previous limitation, computing $W_{L3}$ is not expensive. But the evaluation of $W_{L3}$ can be made even faster by pre-computing distance maps for each template landmark

### G.4.6 Compute step length using line search

The line search method was outlined in section G.3. This section provides implementation details for the inverse compositional algorithm. To compute the step length $\alpha_i$, the following 4 results are needed

$$\phi(0) = f(0,\ p_{i-1}) = \frac{1}{2}\|e_{T,I}(0,\ p_{i-1})\|^2 \tag{G.39a}$$

$$\phi(\alpha_{i-1}) = f(\alpha_{i-1}\Delta p_i,\ p_{i-1}) = \frac{1}{2}\|e_{T,I}(\alpha_{i-1}\Delta p_i,\ p_{i-1})\|^2 \tag{G.39b}$$

$$\phi'(0) = f'(0,\ p_{i-1}) = \Upsilon(0)^\top e_{T,I}(0,\ p_{i-1}) \tag{G.39c}$$

$$\phi'(\alpha_{i-1}) = f'(\alpha_{i-1}\Delta p_i,\ p_{i-1}) = \Upsilon(\alpha_{i-1}\Delta p_i)^\top e_{T,I}(\alpha_{i-1}\Delta p_i,\ p_{i-1}) \tag{G.39d}$$

Computing $e_{T,I}(\alpha_{i-1}\Delta p_i,\ p_{i-1})$ requires the parameter update in Appendix H for similarity transformations, or in Eq. (G.37) when TPSs are also used, so that Eq. (G.39) can be written

$$\phi(0) = f(0,\ p_{i-1}) = \frac{1}{2}\|e_{T,I}(0,\ p_{i-1})\|^2 \tag{G.40a}$$

$$\phi(\alpha_{i-1}) = f(0,\ p_i) = \frac{1}{2}\|e_{T,I}(0,\ p_i)\|^2 \tag{G.40b}$$

$$\phi'(0) = f'(0,\ p_{i-1}) = \Upsilon(0)^\top e_{T,I}(0,\ p_{i-1}) \tag{G.40c}$$

$$\phi'(\alpha_{i-1}) = f'(0,\ p_i) = \Upsilon(\alpha_{i-1}\Delta p_i)^\top e_{T,I}(0,\ p_i) \tag{G.40d}$$

Note that

- The residual $e_{T,I}(0,\ p_{i-1})$ was already computed in line 2 of the iteration step
- As long as the transformation and warp are linear in the shape coefficients, the Jacobian of the transformation is constant and $\Upsilon(0) = \Upsilon(\alpha_{i-1}\Delta p_i) = \Upsilon$. This is the case with similarity transformations and TPS warps, as shown by Eq. (G.16) and (G.24), respectively.

# APPENDIX H

## Composition of similarity transformations

The update transformation parameters $q$ in Algorithm G.1 requires the composition of two similarity transformations. This appendix derives the analytical solution to that problem.

Let $W_G$ be a similarity transformation

$$W_G(z;\ q) = \begin{bmatrix} 1+q_1 & -q_2 \\ q_2 & 1+q_1 \end{bmatrix} z + \begin{bmatrix} q_3 \\ q_4 \end{bmatrix} \tag{H.1}$$

In the rest of this thesis $q(i)$ refers to the $i$-th element of vector $q$, but for clarity, in this appendix the notation $q_i$ is used instead. The inverse of $W_G$ for incremental parameters $\Delta q$ is

$$W_G^{-1}(z;\ \Delta q) = \frac{1}{\left(1+\Delta q_1\right)^2 + \left(\Delta q_2\right)^2} \begin{bmatrix} 1+\Delta q_1 & \Delta q_2 \\ -\Delta q_2 & 1+\Delta q_1 \end{bmatrix} \left( z - \begin{bmatrix} \Delta q_3 \\ \Delta q_4 \end{bmatrix} \right) \tag{H.2}$$

The inverse of a similarity transformation is a similarity transformation $W_G^{-1}(z;\ \Delta q) =$

$W_G(z;\ \Delta q')$ with parameters $q'$

$$q'_1 = \frac{1 + q_1 - (1 + q_1)^2 - q_2^2}{(1 + q_1)^2 + q_2^2} \tag{H.3a}$$

$$q'_2 = \frac{-q_2}{(1 + q_1)^2 + q_2^2} \tag{H.3b}$$

$$q'_3 = \frac{-(1 + q_1)q_3 - q_2 q_4}{(1 + q_1)^2 + q_2^2} \tag{H.3c}$$

$$q'_4 = \frac{q_2 q_3 - (1 + q_1)q_4}{(1 + q_1)^2 + q_2^2} \tag{H.3d}$$

The composition of both functions is another similarity transformation

$$\begin{aligned} W_G(z;\ q'') &= W_G(z;\ q) \circ W_G^{-1}(z;\ \Delta q) \\ &= W_G(z;\ q) \circ W_G(z;\ \Delta q') \end{aligned} \tag{H.4}$$

with parameters

$$q''_1 = (1 + q_1)(1 + \Delta q'_1) - q_2 \Delta q'_2 - 1 \tag{H.5a}$$

$$q''_2 = (1 + q_1)\Delta q'_2 + q_2(1 + \Delta q'_1) \tag{H.5b}$$

$$q''_3 = (1 + q_1)\Delta q'_3 - q_2 \Delta q'_4 + q_3 \tag{H.5c}$$

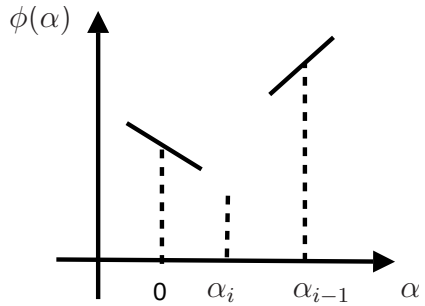
$$q''_4 = q_2 \Delta q'_3 + (1 + q_1)\Delta q'_4 + q_4 \tag{H.5d}$$

# APPENDIX I

# Line search algorithm and termination conditions

This appendix describes in full the cubic polynomial line search method and termination conditions used by the Matlab Optimization Toolbox's (R2006a to R2008a) implementation of the Gauss-Newton algorithm in function `nlsq()`. It should be noted that the official documentation [110, p. 5-12] does not correspond to the actual implementation in `nlsq()`. Line search and termination conditions are used in the Gauss-Newton solution to the inverse compositional segmentation algorithm described in Appendix G, and detailed here for clarity.

## I.1 Line search scenarios

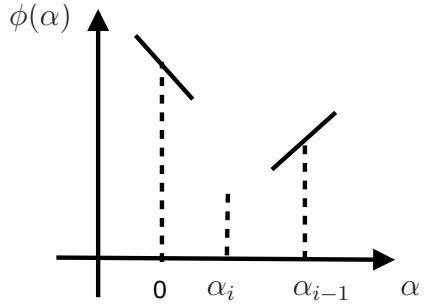



(a) Case 1.

**Require:** $\begin{cases} \phi(0) < \phi(\alpha_{i-1}) \\ \phi'(\alpha_{i-1})^\top \Delta p_{i+1} > 0 \end{cases}$

1: $\alpha_i = 0.9\ \mathsf{cubici1}\big(\phi(\alpha_{i-1}),\ \phi(0),\ \phi'(\alpha_{i-1})^\top \Delta p_{i+1},\ \phi'(0)^\top \Delta p_i,\ \alpha_{i-1}\big)$

(b) Case 1.



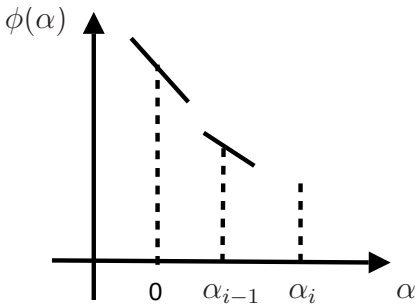

(c) Case 2.

**Require:** $\begin{cases} \phi(0) > \phi(\alpha_{i-1}) \\ \phi'(\alpha_{i-1})^\top \Delta p_i \geq 0 \end{cases}$

1: $\alpha_c = \mathsf{cubici1}\big(\phi(\alpha_{i-1}),\ \phi(0),\ \phi'(\alpha_{i-1})^\top \Delta p_{i+1},\ \phi'(0)^\top \Delta p_i,\ \alpha_{i-1}\big)$
2: **if** $\alpha_{i-1} > 0.9$ **then**
3: $\quad \alpha_i = \min(1,\ |\alpha_c|)$
4: **end if**

(d) Case 2.

Figure I.1: Cubic polynomial interpolation line search scenarios.

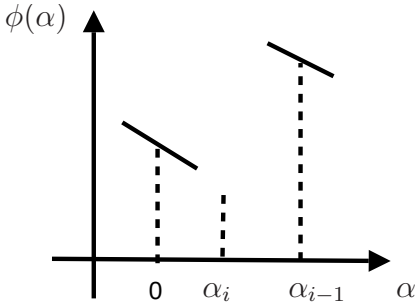



(e) Case 3.

**Require:** $\begin{cases}\phi(0) > \phi(\alpha_{i-1}) \\ \phi'(\alpha_{i-1})^\top \Delta p_i < 0\end{cases}$

1: $\alpha_c = \mathsf{cubici1}\big(\phi(\alpha_{i-1}),\ \phi(0),\ \phi'(\alpha_{i-1})^\top \Delta p_{i+1},$
$\phi'(0)^\top \Delta p_i,\ \alpha_{i-1}\big)$
2: **if** $\alpha_{i-1} > \alpha_c$ **then**
3: $\quad \alpha_c = 2\alpha_{i-1} + 10^{-4}$
4: **end if**
5: $\alpha_i = |\alpha_c|$

(f) Case 3.

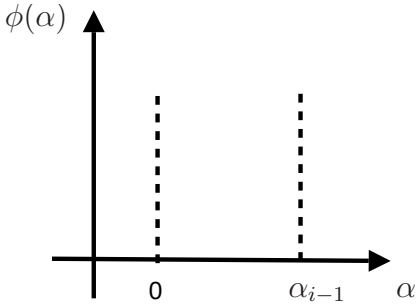



(g) Case 4.

**Require:** $\begin{cases}\phi(0) < \phi(\alpha_{i-1}) \\ \phi'(\alpha_{i-1})^\top \Delta p_i \leq 0\end{cases}$

1: **if** $\alpha_c < 8 \cdot 10^{-8}$ **then**
2: $\quad \alpha_i = -\alpha_c/8$
3: **else**
4: $\quad \alpha_i = \alpha_c/8$
5: **end if**

(h) Case 4.




(i) Case 5.

**Require:** $\phi(0) = \phi(\alpha_{i-1})$

1: No improvement, terminate

(j) Case 5.

Figure I.1: Cubic polynomial interpolation line search scenarios (cont.).

## I.2 Cubic interpolation

Cubic interpolation can be performed with function `cubici1()` of the Matlab Optimization Toolbox. The optimal step length $\alpha_c$ is

$$\alpha_c = \frac{z + w - \phi'(0)}{\phi'(\alpha_{i-1}) - \phi'(0) + 2w} \tag{I.1a}$$

$$z = 3\,\frac{\phi(0) - \phi(\alpha_{i-1})}{\alpha_{i-1}} + \phi'(0) + \phi'(\alpha_{i-1}) \tag{I.1b}$$

$$w = \Re\left\{\sqrt{z^2 - \phi(0)\,\phi(\alpha_{i-1})}\right\} \tag{I.1c}$$

## I.3 Termination conditions

The algorithm terminates in any of the following cases, where $\epsilon_f = 10^{-6}$, $\epsilon_p = 10^{-6}$, $i_{\max} = 100k$ by default

- Directional derivative along search direction less than $\epsilon_f$ and infinity-norm of gradient less than $10\,(\epsilon_f + \epsilon_p)$

$$\text{both conditions} \begin{cases} \phi'(0)^\top \Delta p_i < \epsilon_f \\ \|\phi'(0)\|_\infty < 10\,(\epsilon_f + \epsilon_p) \end{cases} \tag{I.2}$$

- Magnitude of search direction less than $\epsilon_p$

$$\|\Delta p_i\|_\infty < \epsilon_p \tag{I.3}$$

- Number of iterations exceeded

$$i > i_{\max} \tag{I.4}$$

- Step length too small

$$\alpha < 10^{-8} \tag{I.5}$$

For the last case, if the step length becomes too small, function `nlsq()` switches to the Levenberg-Marquardt method.

# BIBLIOGRAPHY

[1] A. Andreopoulos and J.K. Tsotsos. A novel algorithm for fitting 3-D active appearance models: Applications to cardiac MRI segmentation. In *Image Analysis*, volume 3540, pages 729–739. Springer, 2005.

[2] P. Arias, G. Randall, and G. Sapiro. Connecting the out-of-sample and pre-image problems in kernel methods. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 18–23 Jun, Poster 237, Jun 2007.

[3] V. Arsigny, X. Pennec, and N. Ayache. Polyrigid and polyaffine transformations: A new class of diffeomorphisms for locally rigid or affine registration. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2003)*, Lecture Notes in Computer Science, pages 829–837. Springer, 2003.

[4] J. Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38 (1):95–113, 2007.

[5] S. Baker and I. Matthews. Equivalence and efficiency of image alignment algorithms. In *Procs. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001 (CPVR 2001)*, volume 1, pages I–1090–I–1097, 2001.

[6] S. Baker and I. Matthews. Lukas-Kanade 20 years on: A unifying framework. Part

1: The quantity approximated, the warp update rule, and the gradient descent approximation. *International Journal of Computer Vision*, 56(3):221–255, 2004.

[7] G.H. Bakir, J. Weston, and B. Schölkopf. Learning to find pre-images. In S. Thrun, L.K. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems*, volume 16, pages 449–456. MIT Press, 2004.

[8] G.H. Bakir, B. Schölkopf, and J. Weston. On the pre-image problem in kernel methods. In Camps-Valls et al. [34], chapter 12.

[9] J.C. Bamber and C. Daft. Adaptive filtering for reduction of speckle in ultrasonic pulse-echo images. *Ultrasonics*, 24(1):41–44, 1986.

[10] A.M. Baumberg and D.C. Hogg. Learning flexible models from image sequences. In *European Conference on Computer Vision (ECCV)*, pages 299–308, Stockholm, May 1994.

[11] R.K. Beatson and G.N. Newsam. Fast evaluation of radial basis functions: I. *Computers & Mathematics with Applications*, 24(12):7–19, 1992.

[12] H. Becher and P.N. Burns. *Handbook of contrast echocardiography. Left ventricular function and myocardial perfusion*. Bonn, Germany, and Toronto, Canada, 2000. URL `http://www.sunnybrook.utoronto.ca/EchoHandbook/`. Available for free download.

[13] H. Becher and R. Olszewski. Current indications for contrast echocardiography. *European Cardiovascular Disease*, 2:44–48, 2007.

[14] R. Beichel, S. Mitchell, E. Sorantin, F. Leberl, A. Goshtasby, and M. Sonka. Shape- and appearance-based segmentation of volumetric medical images. In *Proc. 2001 International Conference on Image Processing*, volume 2(7–10), pages 589–592, Oct 2001.

[15] R. Beichel, H. Bischof, F. Leberl, and M. Sonka. Robust active appearance models and their application to medical image analysis. *IEEE Transactions on Medical Imaging*, 24(9):1151–1169, 2005.

[16] B.D. Beleslin, M. Ostojic, J. Stepanovic, A. Djordjevic-Dikic, S. Stojkovic, M. Nedeljkovic, G. Stankovic, Z. Petrasinovic, L. Gojkovic, and Z. Vasiljevic-Pokrajcic. Stress echocardiography in the detection of myocardial ischemia. head-to-head comparison of exercise, dobutamine, and dipyridamole tests. *Circulation*, 90(3):1168–1175, 1994.

[17] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Computer Vision — European Conference on Computer Vision (ECCV'92)*, volume 588, pages 237–252, 1992.

[18] J. Bermejo, J. Timperley, R.G. Odreman, M. Mulet, J.A. Noble, A.P. Banning, R. Yotti, E. Pérez-David, J. Declerck, H. Becher, and M.A. García-Fernández. Objective quantification of global and regional left ventricular systolic function by endocardial tracking of contrast echocardiographic sequences. *International Journal of Cardiac Imaging*, 124(1):47–56, 2008.

[19] P.J. Besl. Geometric modeling and computer vision. *Proceedings of the IEEE*, 76 (8):936–958, 1988.

[20] V.K. Bhatia and R. Senior. Contrast echocardiography: Evidence for clinical use. *Journal of the American Society of Echocardiography*, 21(5):409–416, 2008.

[21] A. Blake and M. Isard. *Active contours: The application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion*. Springer, London, 1998.

[22] A. Blake, R. Curwen, and A. Zisserman. A framework for spatio-temporal control

in the tracking of visual contours. *International Journal of Computer Vision*, 11(2): 127–145, 1993.

[23] A. Blake, M. Isard, and D. Reynard. Learning to track the visual motion of contours. *Artificial Intelligence*, 78:179–212, 1995.

[24] F.L. Bookstein. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, 1(2):181–242, 1986.

[25] F.L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11 (6):567–585, Jun 1989.

[26] J.G. Bosch, S.C. Mitchell, B.P.F. Lelieveldt, F. Nijland, O. Kamp, and M. Sonka. Automatic segmentation of echocardiographic sequences by active appearance motion models. *IEEE Transactions on Medical Imaging*, 21(11):1374–1383, 2002.

[27] D. Boukerroui, J.A. Noble, M.C. Robini, and J.M. Brady. Enhancement of contrast regions in sub-optimal ultrasound images with application to echocardiography. *Ultrasound in Medicine & Biology*, 27(12):1583–1594, 2001.

[28] R. Brooks and T. Arbel. Generalizing inverse compositional image alignment. In *Procs. of the 18th International Conference on Pattern Recognition (ICPR'06)*, volume 2, pages 1200–1203, 2006.

[29] M.D. Buhmann. *Radial basis functions: Theory and implementations*. Cambridge University Press, 2003.

[30] C.J.C. Burges. Simplified support vector decision rules. In *13th International Conference on Machine Learning*, pages 71–77, 1996.

[31] L.J. Busse, T.R. Crimmins, and J.R. Fienup. A model based approach to improve the performance of the geometric filtering speckle reduction algorithm. In *Procs. of Ultrasonics Symposium 1995*, volume 2, pages 1353–1356, 1995.

[32] E.G. Caiani, E. Toledo, P. MacEneaney, K.A. Collins, R.M. Lang, and V. Mor-Avi. Objective assessment of left ventricular wall motion from cardiac magnetic resonance images. *Computers in Cardiology*, 31:153–156, 2004.

[33] V. Camion and L. Younes. Geodesic interpolating splines. In *Proceedings of the Third International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 2134 of *Lecture Notes In Computer Science*, pages 513–527, London, UK, 2001. Springer.

[34] G. Camps-Valls, J.L. Rojo-Alvarez, and M. Martinez-Ramon, editors. *Kernel Methods in Bioengineering, Signal and Image Processing*. Idea Group Publishing, Hersey, PA, USA, 2007.

[35] S. Carstensen, U. Hoest, L. Kjoeller-Hansen, K. Saunamäki, D. Atar, and H. Kelbaek. Comparison of methods of fractional area change for detection of regional left ventricular dysfunction. *International Journal of Cardiac Imaging*, 16(4):257–266, 2000.

[36] R. Casero and J.A. Noble. A novel explicit 2D+t cyclic shape model applied to echocardiography. In *Procs. of Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2008*, volume 5241, pages 527–534. Springer, 2008.

[37] M.D. Cerqueira, N.J. Weissman, V. Dilsizian, A.K. Jacobs, S. Kaul, W.K. Laskey, D.J. Pennell, J.A. Rumberger, T. Ryan, and M.S. Verani. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: A statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *Circulation*, 105(4):539–542, 2002.

[38] V. Chalana and Y. Kim. A methodology for evaluation of boundary detection algorithms on medical images. *IEEE Transactions on Medical Imaging*, 16(5):642–652, 1997.

[39] V. Chalana, D.T. Linker, D.R. Haynor, and Y. Kim. A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 15(3):290–298, Jun 1996.

[40] M.D. Cheitlin, W.F. Armstrong, G.P. Aurigemma, G.A. Beller, F.Z. Bierman, J.L. Davis, P.S. Douglas, D.P. Faxon, L.D. Gillam, T.R. Kimball, W.G. Kussmaul, A.S. Pearlman, J.T. Philbrick, H. Rakowski, and D.M. Thys. *ACC/AHA/ASE 2003 Guideline Update for the Clinical Application of Echocardiography. A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (ACC/AHA/ASE Committee to Update the 1997 Guidelines for the Clinical Application of Echocardiography).* American College of Cardiology Foundation and American Heart Association, 2003. URL `http://www.acc.org/clinical/guidelines/echo/index.pdf`. American College of Cardiology Web Site.

[41] S.S. Chen and R.A. Gopinath. Gaussianization. In *Proc. Advances in Neural Information Processing Systems (NIPS'2000)*, volume 13, Denver, 2000.

[42] T.F. Cootes and P. Kittipanya-ngam. Comparing variations on the Active Appearance Model algorithm. In *Procs. of the British Machine Vision Conference (BMVC)*, pages 837–846, 2002.

[43] T.F. Cootes and C.J. Taylor. Active shape models - 'Smart snakes'. In *Procs. of the British Machine Vision Conference (BMVC)*, pages 266–275. Springer, 1992.

[44] T.F. Cootes and C.J. Taylor. Modelling object appearance using the grey-level surface. In E. Hancock, editor, *British Machine Vision Conference*, pages 479–488, York, UK, Sep 1994.

[45] T.F. Cootes and C.J. Taylor. Statistical models of appearance for medical image analysis and computer vision. In *Proc SPIE Medical Imaging 2001*, pages 236–248, San Diego, CA, 2001.

[46] T.F. Cootes and C.J. Taylor. Statistical models of appearance for computer vision. Technical report, University of Manchester, 2001. URL `http://www.isbe.man.ac.uk/~bim/`.

[47] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Training models of shape from sets of examples. In *Procs. of the British Machine Vision Conference (BMVC)*, pages 266–275, Berlin, 1992. Springer.

[48] T.F. Cootes, C.J. Edwards, and C.J. Taylor. Active appearance models. In H. Burkhardt and B. Neumann, editors, *5th European Conference on Computer Vision*, volume 2, pages 484–498, Berlin, 1998. Springer.

[49] T.F. Cootes, G.V. Wheeler, K.N. Walker, and C.J. Taylor. Coupled-view active appearance models. In M. Mirmedhi and B.T. Thomas, editors, *Proceedings of the British Machine Vision Conference*, Sep 2000.

[50] T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681–685, Jun 2001.

[51] I. Craw and P. Cameron. Parameterising images for recognition and reconstruction. In *British Machine Vision Conference*, pages 376–370, London, UK, 1991. Springer.

[52] D. Cremers and T. Kohlberger. Probabilistic kernel PCA for statistical shape models with applications to image segmentation. In Camps-Valls et al. [34], chapter 15.

[53] D. Cremers and M. Rousson. Efficient kernel density estimation of shape and intensity priors for level set segmentation. In J.S. Suri and A. Farag, editors, *Parametric and Geometric Deformable Models: An application in Biomaterials and Medical Imagery*. Springer, May 2007.

[54] A. d'Aspremont, L. El Ghaoui, M.I. Jordan, and G.R.G. Lanckriet. A direct formu-

lation for sparse PCA using semidefinite programming. *ArXiv Computer Science e-prints*, pages 1–15, Jun 2004.

[55] C. Davatzikos, X. Tao, and D. Shen. Hierarchical Active Shape Models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, 22(3):414–423, 2003.

[56] J.M.B. Dias and J.M.N. Leitão. Wall position and thickness estimation from sequences of echocardiographic images. *IEEE Transactions on Medical Imaging*, 15 (1):25–38, Feb 1996.

[57] R.J. Dickinson. A computer model for speckle in ultrasound images: Theory and application. In A.F. Metherell, editor, *Acoustical Imaging*, volume 10, pages 115–129. Plenum, 1983.

[58] I.L. Dryden and K.V. Mardia. *Statistical Shape Analysis*. John Wiley & Sons, 1998.

[59] J.G. Dumesnil, E.L. Ritman, R.L. Frye, G.T. Gau, B.D. Rutherford, and G.D. Davis. Quantitative determination of regional left ventricular wall dynamics by Roetgen viodeometry. *Circulation*, 50:700–708, 1974.

[60] G.J. Edwards, A. Lanitis, C.J. Taylor, and T.F. Cootes. Statistical models of face images - Improving specificity. In *British Machine Vision Conference*, Edinburgh, UK, 1996.

[61] G.J. Edwards, C.J. Taylor, and T.F. Cootes. Learning to identify and track faces in image sequences. In *Procs. of the British Machine Vision Conference (BMVC)*, 1997.

[62] G.J. Edwards, C.J. Taylor, and T.F. Cootes. Interpreting face images using Active Appearance Models. In *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pages 300–305, Nara, Japan, Apr 1998.

[63] A.P. Eriksson. *Optimization Methods for Large Scale Combinatorial Problems and Bijectivity Constrained Image Deformations*. Phd thesis, Lund University, Sweden, 2008.

[64] A.P. Eriksson and K. Åström. Bijective image registration using thin-plate splines. In *The 18th International Conference on Pattern Recognition (ICPR'06)*, volume 3, pages 798–801, 2006.

[65] Jacques Feldmar and Nicholas Ayache. Rigid, affine and locally affine registration of free-form surfaces. *International Journal of Computer Vision*, 18(2):99–119, 1996.

[66] World Health Organization Regional Office for Europe. European mortality data base (MDB), Jul 2008. URL `http://data.euro.who.int/hfamdb/`. Accessed Oct 2008.

[67] A.F. Frangi, W.J. Niessen, and M.A. Viergever. Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE Transactions on Medical Imaging*, 20(1):2–25, Jan 2001.

[68] F. Galton. Co-relations and their measurement, chiefly from anthropometric data. In *Procs. of the Royal Society*, volume 45, pages 135–145, 1888.

[69] H.J. Gelberg, B.H. Brundage, S. Glantz, and W.W. Parmley. Quantitative left ventricular wall motion analysis: a comparison of area, chord and radial methods. *Circulation*, 59:991–1000, 1979.

[70] J.C. Gower. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3 and 4):325–338, 1966.

[71] J.C. Gower. Generalized Procrustes analysis. *Psychometrika*, 40(1):33–51, Mar 1975.

[72] G. Hamarneh and T. Gustavsson. Combining snakes and Active Shape Models for segmenting the human left ventricle in echocardiographic images. In *Computers in Cardiology*, pages 115–118, 2000.

[73] A. Hammoude. Endocardial border identification in two-dimensional echocardiography images: review of methods. *Computerized Medical Imaging and Graphics*, 22 (3):181–193, May 1998.

[74] J. Hansegård, S. Urheim, K. Lunde, and S.I. Rabben. Constrained active appearance models for segmentation of triplane echocardiograms. *IEEE Transactions on Medical Imaging*, 26(10):1391–1400, Oct 2007.

[75] S.A. Hayat and R. Senior. Safety: The heart of the matter. *European Journal of Echocardiography*, 6(4):235–237, Aug 2005.

[76] R.C. Henry, C.W. Lewis, P.K. Hopke, and H.J. Williamson. Review of receptor model fundamentals. *Atmospheric Environment*, 18:1507–1515, 1984.

[77] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441, 1933.

[78] X. Hu and L. Xu. A comparative investigation on subspace dimension determination. *Neural Networks*, 17(8–9):1051–1059, Oct–Nov 2004.

[79] X. Huang, D. Metaxas, and T. Chen. MetaMorphs: Deformable shape and texture models. In *Procs. Computer Vision and Patter Recognition (CVPR'04)*, volume 1, pages I–496–I–503. IEEE Computer Society Press, 2004.

[80] D.I. Hughes and F.A. Duck. Automatic attenuation compensation for ultrasonic imaging. *Ultrasound in Medicine and Biology*, 23(5):651–664, 1997.

[81] C.M. Hurvich and C.-L. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307, 1989.

[82] G. Jacob, J.A. Noble, M. Mulet-Parada, and A. Blake. Evaluating a robust contour tracker on echocardiographic sequences. *Medical Image Analysis*, 3(1):63–75, Mar 1999.

[83] G. Jacob, J.A. Noble, A.D. Kelion, and A.P. Banning. Quantitative regional analysis of myocardial wall motion. *Ultrasound in Medicine and Biology*, 6(6):773–784, 2001.

[84] G. Jacob, J.A. Noble, C. Behrenbruch, A.D. Kelion, and A.P. Banning. A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography. *IEEE Transactions on Medical Imaging*, 21(3):226–238, Mar 2002.

[85] A.K. Jain, Y. Zhong, and M.-P. Dubuisson-Jolly. Deformable template models: A review. *Signal Processing*, 71(2):109–129, Dec 1998.

[86] H.J. Johnson and G.E. Christensen. Consistent landmark and intensity-based image registration. *IEEE Transactions on Medical Imaging*, 21(5):450–461, 2002.

[87] D. Kaplan and Q. Ma. On the statistical characteristics of log-compressed Rayleigh signals. In *Proc. IEEE Ultrasound Symposium*, pages 961–964, 1993.

[88] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision*, 1(4):321–331, Nov 1988.

[89] D.G. Kendall. The diffusion of shape. *Advances in Applied Probability*, 9(3):428–430, 1977.

[90] D.G. Kendall. Comment to 'Size and shape spaces for landmark data in two dimensions'. *Statistical Science*, 1(2):222–226, 1986.

[91] C. Kervrann and F. Heitz. A hierarchical statistical framework for the segmentation of deformable objects in image sequences. In *Procs. Computer Vision and Pattern Recognition (CVPR '94)*, pages 724–728, 1994.

[92] H. Kong, L. Wang, E.K. Teoh, X. Li, J.-G. Wang, and R. Venkateswarlu. Generalized 2D principal component analysis for face image representation and recognition. *Neural Networks*, 18:585–594, 2005.

[93] D. Kucera and R.W. Martin. Segmentation of sequences of echocardiographic images using a simplified 3D active contour model with region-based external forces. *Computerized Medical Imaging and Graphics*, 21(1):1–21, 1997.

[94] J.T.-Y. Kwok and I.W.-H. Tsang. The pre-image problem in kernel methods. *IEEE Transactions on Neural Networks*, 15(6):1517–1525, Nov 2004.

[95] S. Lafon, Y. Keller, and R.R. Coifman. Data fusion and multicue data matching by diffusion maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1784–1797, Nov 2006.

[96] R.M. Lang, M. Bierig, R.B. Devereux, F.A. Flachskampf, E. Foster, P.A. Pellikka andM.H. Picard, M.J. Roman, J. Seward, J.S. Shanewise, S.D. Solomon, K.T. Spencer, M. St John Sutton, and W.J. Stewart. Recommendations for chamber quantification: A report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *Journal of the American Society of Echocardiography*, 18:1440–1463, 2005.

[97] A. Lanitis, C.J. Taylor, and T.F. Cootes. Automatic tracking, coding and reconstruction of human faces, using flexible appearance models. *IEE Electronics Letters*, 30(19):1587–1588, Sep 1994.

[98] A. Lanitis, C.J. Taylor, and T.F. Cootes. An automatic face identification system using flexible appearance models. *Image and Vision Computing*, 13(5):393–401, 1995.

[99] A. Lanitis, C.J. Taylor, and T.F. Cootes. A unified approach to coding and interpreting face images. In *Procs. of the 5th International Conference on Computer Vision*, pages 368–373, Cambridge, USA, 1995. IEEE Computer Society Press.

[100] R.M. Lapp, M. Lorenzo-Valdés, and D. Rueckert. 3D/4D cardiac segmentation using active appearance models, non-rigid registration, and the Insight Toolkit. In *Procs. of the Medical Image Computing and Computer-Assisted Intervention (MICAII 2004)*, volume 3216 of *Lecture Notes in Computer Science*, pages 419–426. Springer, Sep 2004.

[101] J. Laub and K.-R. Müller. Feature discovery in non-metric pairwise data. *Journal of Machine Learning Research*, 5:801–818, 2004.

[102] E.T.Y. Lee. Choosing nodes in parametric curve interpolation. *Computer-Aided Design*, 21(6):363–370, 1989.

[103] B. Lelieveldt, A. Frangi, S. Mitchell, H. van Assen, S. Ordas, J. Reiber, and M. Sonka. 3D active shape and appearance models in cardiac image analysis. In N. Paragios, Y. Chen, and O. Faugeras, editors, *Handbook of Mathematical Models in Computer Vision*, chapter 29, pages 471–485. Springer, New York, 2006.

[104] X. Lin, A. Young, and B. Cowan. Localization and atlas-based segmentation of the heart from cardiac MR images: Validation with a large clinical trial. In *Procs. of the 2nd International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, pages 2319–2322, 2008.

[105] M. Lorenzo-Valdés, G.I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert. Atlas-based segmentation and tracking of 3D cardiac mr images using non-rigid registration. In *Procs. of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2488, pages 642–650, 2002.

[106] M. Lorenzo-Valdés, G.I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert. Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. In *Procs. of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2878, pages 440–450, 2003.

[107] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Procs. 7th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 674–679, 1981.

[108] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Procs. Imaging Understanding Workshop*, pages 121–130, 1981.

[109] S. Marsland and C.J. Twining. Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images. *IEEE Transactions on Medical Imaging*, 23(8):1006–1020, Aug 2004.

[110] *Matlab Optimization Toolbox 4. User's guide*. The MathWorks, 2008. URL `http://www.mathworks.com/access/helpdesk/help/pdf_doc/optim/optim_tb.pdf`.

[111] I. Matthews and S. Baker. Active appearance models revisited. *International Journal of Computer Vision*, 60(2):135–164, Nov 2004.

[112] I. Matthews, T. Ishikawa, and S. Baker. The template update problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):810–815, 2004.

[113] J.C. Mazziotta, A.W. Toga, A. Evans, P. Fox, and J. Lancaster. A probabilistic atlas of the human brain: Theory and rationale for its development. the international consortium for brain mapping (ICBM). *NeuroImage*, 2(2):89–101, 1995.

[114] M. McCulloch, C. Gresser, S. Moos, J. Odabashian, S. Jasper, J. Bednarz, P. Burgess, D. Carney, V. Moore, E. Sisk, A. Waggoner, S. Witt, and D. Adams. Ul-

trasound contrast physics: A series on contrast echocardiography, article 3. *Journal of the American Society of Echocardiography*, 13(10):959–967, Oct 2000.

[115] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[116] L. Mei, M. Figl, D. Rueckert, A. Darzi, and P. Edwards. Sample sufficiency and number of modes to retain in statistical shape modelling. In *Procs. of Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2008*, volume 5241, pages 425–433. Springer, 2008.

[117] D. Metaxas, T. Chen, X. Huan, and L. Axel. Cardiac segmentation from MRI-tagged and CT images. In *8th WSEAS International Conference on Computers*, Athens, Greece, 2004.

[118] S. Michael. KD tree nearest neighbor and range search Matlab Toolbox. Matlab Central File Exchange, 2008. Available as `http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=7030`.

[119] S. Mika, B. Schölkopf, A. Smola, K.-R. Müller, and M. Scholz G. Rätsch. Kernel PCA and de-noising in feature spaces. In M.S. Kearns, S.A. Solla, and D.A. Cohn, editors, *Procs. of the 1998 Conference, Advances in Neural Information Processing Systems*, volume 11, pages 536–542. MIT Press, 1999.

[120] T.P. Minka. Automatic choice of dimensionality for PCA. In T. Leen, T. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems 2000 (NIPS'2000)*, volume 13, pages 598–604. MIT Press, 2001.

[121] 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 on Left and Right Ventricles in cardiac MR images. *IEEE Transactions on Medical Imaging*, 20(5):415–423, May 2001.

[122] 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 Transactions on Medical Imaging*, 21(9):1167–1178, 2002.

[123] K. Mizushige, W. Furumoto, K. Hirao, Y. Iwado, K. Ohmori, and H. Matsuo. Quantitative evaluation of left ventricular regional wall motion using a real-time wall thickness curve system with two-dimensional echocardiography. *American Journal of Cardiology*, 84:1204–1208, 1999.

[124] Johan Montagnat and Hervé Delingette. 4D deformable models with temporal constraints: application to 4D cardiac image segmentation. *Medical Image Analysis*, 9:87–100, 2005.

[125] M. Mulet Parada. *Intensity independent feature extraction and tracking in echocardiographic sequences*. DPhil thesis, Medical Vision Laboratory, Department of Engineering Science, University of Oxford, 2000.

[126] M. Mulet-Parada and J.A. Noble. Intensity-invariant 2D+T acoustic boundary detection. In B. Vemuri, editor, *Worshop on Biomedical Image Analysis*, Santa Barbara, California, USA, Jun 1998. IEEE Computer Society.

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

[128] S.L. Mulvagh, A.N. DeMaria, S.B. Feinstein, P.N. Burns, S. Kaul, J.G. Miller, M. Monaghan, T.R. Porter, L.J. Shaw, and F.S. Villanueva. Contrast echocardiography: Current and future applications. ASE position paper, 2000. URL `http://www.asefiles.org/contrastechoapps.pdf`.

[129] J. Müller-Nordhorn, S. Binting, S. Roll, and S.N. Willich. An update on regional variation in cardiovascular mortality within Europe. *European Heart Journal*, 29 (10):1316–1326, 2008.

[130] *Digital Imaging and Communications in Medicine (DICOM). Part 3: Information Object Definitions*. National Electrical Manufacturers Association, Rosslyn, USA, 2007. URL `http://medical.nema.org/dicom/2007/07_03pu.pdf`.

[131] J.A. Noble and D. Boukerroui. Ultrasound image segmentation: A survey. *IEEE Transactions on Medical Imaging*, 25(8):987–1010, 2006.

[132] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, 2nd edition, 2006.

[133] R. Olszewski, J. Timperley, S. Cezary, M. Monaghan, P. Nihoyannopoulis, R. Senior, and H. Becher. The clinical applications of contrast echocardiography. *European Journal of Echocardiography*, 8:S13–S23, 2007.

[134] J.E. Otterstad. Measuring left ventricular volume and ejection fraction with the biplane simpon's method. *Heart*, 88:559–560, 2002.

[135] X. Papademetris. *Estimation of 3D Left Ventricular Deformation from Medical Images Using Biomechanical Models*. Ph.d. thesis, Yale University, May 2000.

[136] X. Papademetris, A.J. Sinusas, D.P. Dione, R.T. Constable, and J.S. Duncan. Estimation of 3-D left ventricular deformation from medical images using biomechanical models. *IEEE Transactions on Medical Imaging*, 21(7):786–800, Jul 2002.

[137] K. Pearson. Mathematical contributions to the theory of evolution – III Regression, hereditary and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 187:253–318, 1896.

[138] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.

[139] D.J. Pennell, U.P. Sechtem, C.B. Higgins, W.J. Manning, G.M. Pohost, F.E. Rademakers, A.C. van Rossum, L.J. Shaw, and E.K. Yucel. Clinical indications for car-

diovascular magnetic resonance (CMR): Consensus panel report. *European Journal of Echocardiography*, 25:1940–1965, Nov 2004.

[140] D. Perperidis, R. Mohiaddin, and D. Rueckert. Spatio-temporal free-form registration of cardiac MR image sequences. *Medical Image Analysis*, 9(5):441–456, 2005.

[141] P.-O. Persson and G. Strang. A simple mesh generator in MATLAB. *SIAM Review*, 46:329–345, 2004. URL `http://www-math.mit.edu/~persson/mesh/`.

[142] Z.B. Popovic, C. Benejam, J. Bian, N. Mal, J. Drinko, K. Lee, F. Forudi, R. Reeg, N.L. Greenberg, J.D. Thomas, and M.S. Penn. Speckle tracking echocardiography correctly identifies segmental left ventricular dysfunction induced by scarring in a rat model of myocardial infarction. *American Journal of Physiology - Heart and Circulatory Physiology*, 292(6):H2809–2816, 2007.

[143] M.J.D. Powell. Truncated Laurent expansions for the fast evaluation of thin plate splines. *Numerical Algorithms*, 5:99–120, 1993.

[144] M.A. Quinones, P.S. Douglas, E. Foster, J. Gorcsan, J.F. Lewis, A.S. Pearlman, J. Rychik, E.E. Salcedo, J.B. Seward, J.G. Stevenson, D.M. Thys, H.H. Weitz, W.A. Zoghbi, M.A. Creager, W.L. Winters, Jr, M. Elnicki, J.W. Hirshfeld, Jr, B.H. Lorell, G.P. Rodgers, and C.M. Tracy. ACC/AHA clinical competence statement on echocardiography. *Journal of the American College of Cardiology*, 41(4):687–708, 2003.

[145] Y. Rathi, S. Dambreville, and A. Tannenbaum. Statistical shape analysis using kernel PCA. In E.R. Dougherty, J.T. Astola, K.O. Egiazarian, N.M. Nasrabadi, and S.A. Rizvi, editors, *Procs of SPIE, Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, volume 6064, Feb 2006.

[146] F.J. Rohlf and D. Slice. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, 39(1):40–59, 1990.

[147] L. Rosenberg. Nonnormality of linear combinations of normally distributed random variables. *The American Mathematical Monthly*, 72(8):888–890, Oct 1965.

[148] D. Rueckert and P. Burger. Shape-based segmentation and tracking in 4D cardiac MR images. In *CVRMed-MRCAS'97*, pages 43–52. Springer, 1997.

[149] G. Saon, S. Dharanipragada, and D. Povey. Feature space gaussianization. In *Procs. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04)*, volume 1, pages I–329–I–332, May 2004.

[150] L.K. Saul, K.Q. Weinberger, F. Sha, J. Ham, and D.D. Lee. Spectral methods for dimensionality reduction. In B. Schölkopf, O. Chapelle, and A. Zien, editors, *Semisupervised learning*. MIT Press, Cambridge, USA, 2006.

[151] S.G. Sawada, D.S. Segar, T. Ryan, S.E. Brown, A.M. Dohan, R. Williams, N.S. Fineberg, W.F. Armstrong, and H. Feigenbaum. Echocardiographic detection of coronary artery disease during dobutamine infusion. *Circulation*, 83(5):1605–1614, 1991.

[152] N.B. Schiller. Two-dimensional echocardiographic determination of left ventricular volume, systolic function, and mass. *Circulation*, 84(3):I–280–I–287, 1991.

[153] N.B. Schiller, P.M. Shah, M. Crawford, A. DeMaria, R. Devereux, H. Feigenbaum, H. Gutgesell, N. Reichek, D. Sahn, I. Schnittger, N.H. Silverman, and A. Jamil Tajik. Recommendations for quantitation of the left ventricle by two-dimensional echocardiography. *Journal of the American Society of Echocardiography*, 2(5):385–367, 1989.

[154] B. Schölkopf, A.J. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5):1299–1319, 1998.

[155] B. Schölkopf, P. Knirsch, A. Smola, and C. Burges. Fast approximation of support vector kernel expansions, and an interpretation of clustering as approximation in fea-

ture spaces. In *Informatik Aktuell, Mustererkennung 1998, 20. DAGM-Symposium*, pages 125–132, London, UK, 1998. Springer.

[156] B. Schölkopf, S. Mika, A. Smola, G. Rätsch, and K.-R. Müller. Kernel PCA pattern reconstruction via approximate pre-images. In *Procs. of the 8th International Conference on Artificial Neural Networks*, pages 147–152, Berlin, 1998. Springer.

[157] B. Schölkopf, S. Mika, C.J.C. Burges, P. Knirsch, K.-R. Müller, G. Rätsch, and A.J. Smola. Input space versus feature space in kernel-based methods. *IEEE Transactions on Neural Networks*, 10(5):1000–1016, Sep 1999.

[158] B. Schölkopf, A.J. Smola, and K.-R. Müller. Kernel principal component analysis. In B. Schölkopf, C.J.C. Burges, and A.J. Smola, editors, *Advances in Kernel Methods - SV Learning*, pages 327–352. MIT Press, 1999.

[159] S.L. Sclove. Some aspects of model-selection criteria. In *Proc. first US/Japan conference on the frontiers of statistical modeling*, volume 2: Multivariate Statistical Modeling, pages 37–69. Kluwer, 1994.

[160] R. Senior. Assessment of myocardial ischaemia by stress echocardiography. *Cardiac Imaging*, 2(1):2–5, 1999.

[161] Y. Shang, G. Su, and O. Dössel. Hierarchical 3D shape model for segmentation of 4D MR cardiac images. In *Medical Imaging and Augmented Reality*, volume 4091, pages 333–340. Springer, 2006.

[162] F.H. Sheehan, E.L. Bolson, H.T. Dodge, D.G. Mathey, J. Schofer, and H.W. Woo. Advantages and applications of the centerline method for characterizing regional ventricular function. *Circulation*, 74(2):293–305, 1986.

[163] R. Shekhar, V. Zagrodsky, M.J. Garcia, and J.D. Thomas. Registration of real-time 3-D ultrasound images of the heart for novel 3-D stress echocardiography. *IEEE Transactions on Medical Imaging*, 23(9):1141–1149, Sep 2004.

[164] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America*, 4(3):519–524, 1987.

[165] K. Sjöstrand, M.B. Stegmann, and R. Larsen. Sparse principal component analysis in medical shape modeling. In *International Symposium on Medical Imaging*, volume 6144, San Diego, USA, 2006. The International Society for Optical Engineering.

[166] X. Song, A. Myronenko, and D.J. Sahn. Speckle tracking in 3D echocardiography with motion coherence. In *Procs. of IEEE Conference on Computer Vision and Pattern Recognition 2007*, pages 1–7, 2007.

[167] L.H. Staib and J.S. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11): 1061–1075, 1992.

[168] M.B. Stegmann. *Active Appearance Models: Theory, Extensions & Cases*. Master thesis, Lyngby, Sep 2000.

[169] M.B. Stegmann. Analysis and segmentation of face images using point annotations and linear subspace techniques. Tech report, Technical University of Denmark, Lyngby, Denmark, Aug 2003.

[170] M.B. Stegmann. *Generative interpretation of medical images*. PhD Thesis, Technical University of Denmark, Lyngby, Apr 2004.

[171] M.B. Stegmann and D. Pedersen. Bi-temporal 3D active appearance models with applications to unsupervised ejection fraction estimation. In *Proc. of SPIE, International Symposium on Medical Imaging 2005*, volume 5747, pages 336–350, 2005.

[172] M.B. Stegmann, B.K. Ersbøll, and R. Larsen. FAME - A Flexible Appearance Modeling Environment. *IEEE Transactions on Medical Imaging*, 22(10):1319–1331, Oct 2003.

[173] T. Störk, M. Möckel, O. Danne, H. Völler, H. Eichstädt, and U. Frei. Left ventricular hypertrophy and diastolic dysfunction: Their relation to coronary heart disease. *Cardiovascular Drugs and Therapy*, 9(Supp. 3):533–537, 1995.

[174] J.S. Suri. Computer vision, pattern recognition and image processing in left ventricle segmentation: The last 50 years. *Pattern Analysis & Applications*, 3:209–242, 2000.

[175] Z. Tao, C.C. Jaffe, and H.D. Tagare. Tunnelling descent: A new algorithm for active contour segmentation of ultrasound images. In *Information Processing in Medical Imaging*, volume 2732 of *Lecture Notes in Computer Science*, pages 246–257, 2003.

[176] Z. Tao, H.D. Tagare, and J.D. Beaty. Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images. *IEEE Transactions on Medical Imaging*, 25(11):1483–1491, Nov 2006.

[177] J. Timperley, A.R.J. Mitchell, and H. Becher. Contrast echocardiography for left ventricular opacification. *Heart*, 89:1394–1397, 2003.

[178] J. Timperley, H. Thibault, and H. Becher. The Oxford way of doing contrast echocardiography. Zamorano et al. [192], chapter 14, pages 235–249.

[179] M.E. Tipping and C.M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society, Series B*, 61(3):611–622, 1999.

[180] M.E. Tipping and M. Bishop. Mixtures of probabilistic principal component analyzers. *Neural Computation*, 11(2):443–482, Feb 1999.

[181] C. Twining, S. Marsland, and C. Taylor. Measuring geodesic distances on the space of bounded diffeomorphisms. In P.L. Rosin and D. Marshall, editors, *Procs. of the British Machine Vision Conference (BMVC), Cardiff*, volume 2, pages 847–856, Sep 2002.

[182] C.J. Twining and S. Marsland. Constructing diffeomorphic representations of non-rigid registrations of medical images. In C.J. Taylor and J.A. Noble, editors, *Information Processing in Medical Imaging, 18th International Conference*, volume 2732 of *Lecture Notes in Computer Science*, Ambleside, UK, Jul 2003. Springer.

[183] G.M. van der Hoeven, P.J. Clerens, J.J. Donders, J.E. Beneken, and J.T. Vonk. A study of systolic time intervals during uninterrupted excercise. *British Heart Journal*, 39:242–254, 1977.

[184] O.P. van Driel. On various causes of improper solutions in maximum likelihood factor analysis. *Phychometrika*, 43(2):225–243, 1978.

[185] C.D. von Land, S.R. Rao, and J.H.C. Reiber. Development of an improved centreline wall motion model. In *Procs. of Computers in Cardiology 1990*, pages 687–690, 1990.

[186] A.D. Waggoner, D. Ehler, D. Adams, S. Moos, J. Rosenbloom, C. Gresser, J.E. Perez, and P.S. Douglas. Guidelines for the cardiac sonographer in the performance of contrast echocardiography: Recommendations of the American Society of Echocardiography Council on cardiac sonography. *Journal of the American Society of Echocardiography*, 14:417–420, 2001.

[187] L. Wang, X. Wang, X. Zhang, and J. Feng. The equivalence of two-dimensional PCA to line-based PCA. *Pattern Recognition Letters*, 26:57–60, 2005.

[188] A.M. Weissler, W.S. Harris, and C.D. Schoenfeld. Systolic time intervals in heart failure in man. *Circulation*, 37(2):149–159, Feb 1968.

[189] G. Xiao, M. Brady, J.A. Noble, and Y. Zhang. Segmentation of ultrasound B-mode images with intensity inhomogeneity correction. *IEEE Transactions on Medical Imaging*, 21(1):48–57, Jan 2002.

[190] J. Yang, D. Zhang, A.F. Frangi, and J.-Y. Yang. Two-dimensional PCA: a new ap-

proach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):131–137, Jan 2004.

[191] A.L. Yettram and C.A. Vinson. Geometric modelling of the human left ventricle. *Journal of Biomechanical Engineering*, 101:221–223, 1979.

[192] J.L. Zamorano, M.Á. García Fernández, and S. Kaul. *Contrast Echocardiography in Clinical Practice*. Springer, 2004.