Bi-temporal 3D Active Appearance Models with Applications to Unsupervised Ejection Fraction Estimation

Mikkel B. Stegmann[a,b], Dorthe Pedersen[b]

[a]Informatics and Mathematical Modelling, Technical University of Denmark,
Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark

[b]Danish Research Centre for Magnetic Resonance,
Copenhagen University Hospital Hvidovre, Kettegård Allé 30, DK-2650 Hvidovre, Denmark

ABSTRACT

Rapid and unsupervised quantitative analysis is of utmost importance to ensure clinical acceptance of many examinations using cardiac magnetic resonance imaging (MRI). We present a framework that aims at fulfilling these goals for the application of left ventricular ejection fraction estimation in four-dimensional MRI. The theoretical foundation of our work is the generative two-dimensional Active Appearance Models by Cootes et al., here extended to bi-temporal, three-dimensional models. Further issues treated include correction of respiratory induced slice displacements, systole detection, and a texture model pruning strategy. Cross-validation carried out on clinical-quality scans of twelve volunteers indicates that ejection fraction and cardiac blood pool volumes can be estimated automatically and rapidly with accuracy on par with typical inter-observer variability.

1. INTRODUCTION

Cardiovascular disease (CVD) continues to be the world's leading cause of death.[27] Consequently, efforts to improve on cardiovascular diagnosis and therapy are not only desirable, but imperative to ensure the public health and a realistic economy within the health sector.

The recently Nobel Prize awarded magnetic resonance imaging (MRI) technique has been firmly established as the gold standard for cardiac imaging. As such, it accurately depicts cardiac structure, function, perfusion and myocardial viability with a capacity unmatched by any other single imaging modality.[22] However, while MRI is a widely accepted method for cardiovascular examinations within research, its use in clinical practice is limited.[39] One major reason for this disparity is the often prohibitive amount of manual interaction required to extract functional indices, see e.g.[26, 36, 42]

This work addresses the crucial problem of estimating global cardiac function in the form of left ventricular (LV) volume, mass and ejection fraction (EF)*. A prerequisite for doing so, is an accurate delineation of the myocardium from surrounding tissue, air and blood as shown in a single short-axis cardiac MRI slice in Figure 1 (left). Compiling such planar delineations yields the two three-dimensional endocardial (inner) and epicardial (outer) surfaces shown in Figure 1 (right) at the end-diastolic phase.

Delineating the myocardium in clinical quality MRI is often not a simple task since many factors during the MR image acquisition deteriorate image quality. These include resolution/time trade-offs, field inhomogeneity, susceptibility artefacts, flow artefacts, motion artefacts, partial volume effects, chemical shift artefacts, et cetera. The data treated in this paper is no exception. Consequently, models with weak priors – such as local edge detection – may succeed in some, limited parts of the myocardium, but are most likely to utterly fail. Only strong priors can redeem this.

In the following, we will apply global models of shape and image appearance constrained to synthesise plausible instances of the myocardium and the left ventricular blood pool. Solutions will thus be imputed by these models in areas exhibiting weak image evidence, e.g. at parts of the blood/muscle interface. This is analogous to the typical image annotation process where the operator integrates knowledge about the current image slice,

---

Corresponding author is M. B. Stegmann, E-mail: mbs@imm.dtu.dk, Web: http://www.imm.dtu.dk/~mbs/.
*The proportion of blood in the left ventricle that is ejected with each heartbeat.
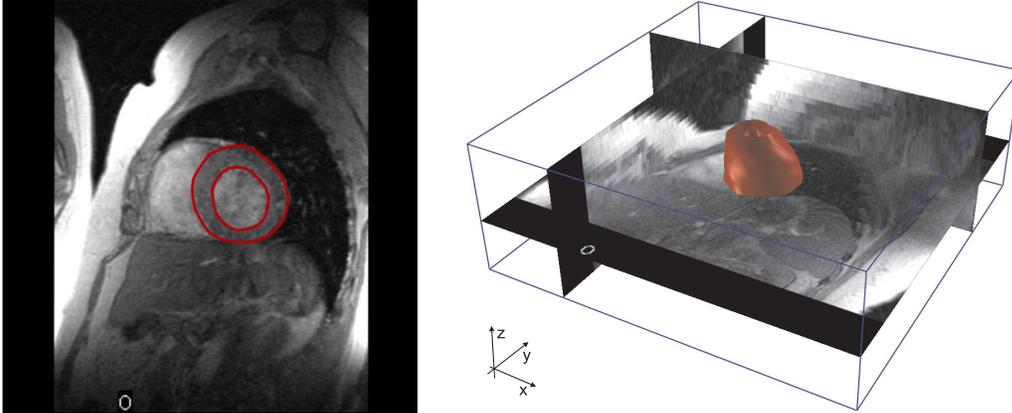
Figure 1: Example delineation of the left ventricle in cardiac MRI (left). Four-dimensional cardiac cine image (shown at end-diastole). Ground truth endo- and epicardial surfaces shown as shaded meshes (right).

not only from the neighbouring slices, but also from previously annotated patients. In fact, these global prior models are built from representative examples containing solutions to the problem at hand; delineation of the myocardium. This powerful concept of learning-based models has recently gained much attention in the image analysis literature due to its virtues of being general computational frameworks that – once trained – remain specific to the domain problem. One widely-used two-dimensional realisation of such a learning-based model is the class of Active Appearance Models (AAMs),[13, 23] which were recently extended to three dimensions.[42] The work presented in the following continues this development with extensions to bi-temporal, three-dimensional Active Appearance Models with applications to estimation of left ventricular volumes and ejection fraction from four-dimensional cardiac cine MRI.

2. DATA MATERIAL

Dynamic, volumetric magnetic resonance images in four dimensions were acquired from twelve healthy, obese (*body mass index* > 33 kg/m^2) subjects using a whole-body MR unit (Siemens Impact) operating at 1.0 Tesla. The pulse sequence was an ECG-triggered, breath-hold, fast low angle shot (FLASH) cinematographic pulse sequence. Matrix=256×256, field of view=263×350 mm, slice thickness=6 mm, inter slice gap=0 mm, phase time=55 ms. The spatio-temporal volumes were $(x, y, z, phases) = 256 \times 256 \times 23 \times 12$ voxels on average. The endocardial (endo) and epicardial (epi) contours of the left ventricle were annotated at the end-diastole (ED) and the end-systole (ES) by manually placing landmarks along the myocardial border, see Figure 1 (left).

3. METHODS

The methodological overview presented in this section is given in two parts. First, the Active Appearance Models (AAMs) are presented using broad terms, without touching the question of dimensionality or the domain problem in itself. Then specific details regarding modelling of bi-temporal cardiac data are given; with emphasis on issues pertinent to the extension of AAMs to higher dimensions than two.

3.1. Active Appearance Models

Interpretation by synthesis has been shown to be a very powerful approach to image analysis. On a coarse level, this approach consists of two parts; i) a mathematical model, which is able to mimic the image formation process, and ii) a search regime, which is able to match this model to an image by letting the image synthesised by the model – in some sense – be "similar" to the unseen image.

For such a strategy to be useful, certain requirements must be met for both parts. The model must either encode specific features of interests and/or contain parameters that reflect interesting latent variables of the objects in question. For example this could be point features locating interesting structures, and a variable related to the age of the object. Additionally, this model should only be capable of synthesising valid image instances of the

object class in question. Matching such a model to an unknown image would thus draw inference about these properties. The matching procedure should provide a meaningful model-to-image match in a reasonable amount of time, dependent on the application, or alternatively reject the presence of the sought-after object.

Active Appearance Models (AAMs)[13, 23] represent one method that has gained considerable attention in the literature, seeking to meet the above requirements. The encoded properties of the model are correlated movement of landmark points and latent texture variables controlling image appearance normalised for shape changes given by landmark points. In essence a fruitful marriage between the ideas of Eigenface Models,[57] Point Distribution Models[15] and Active Shape Models[17] with some ingenious refinements. All these techniques originated in the early nineties. Since their introduction in 1998, AAMs have been refined and adapted in many aspects, see e.g.[14, 16] A recent summary of AAM refinements and applications within medical imaging can be found in.[53]

Formally, AAMs establish a compact parameterisation of object variability, as learned from a representative training set. Objects are defined by marking up each example with points of correspondence (i.e. landmarks) over the training set either by hand, or by semi- to completely automated methods. Using a learning-based optimisation strategy, AAMs can be rapidly fitted to unseen images.

Variability is modelled by means of a Principal Component Analysis (PCA), i.e. an eigenanalysis of the dispersions of shape and texture. Let there be given Q training examples for an object class, and let each example be represented by a set of N landmark points and M texture samples. Shape examples are aligned to a normalised common mean using a Generalised Procrustes Analysis.[30] Texture examples are warped into correspondence by forcing landmarks to match, intensity normalised, and subsequently sampled from this *shape-free* reference. Typically, this geometrical reference frame is the Procrustes mean shape. Let $\mathbf{s}$ and $\mathbf{t}$ denote a synthesised shape and texture and let $\overline{\mathbf{s}}$ and $\overline{\mathbf{t}}$ denote the corresponding sample means. New instances are now generated by adjusting the principal component scores, $\mathbf{b}_s$ and $\mathbf{b}_t$ in $\mathbf{s} = \overline{\mathbf{s}} + \mathbf{\Phi}_s \mathbf{b}_s$ and $\mathbf{t} = \overline{\mathbf{t}} + \mathbf{\Phi}_t \mathbf{b}_t$ where $\mathbf{\Phi}_s$ and $\mathbf{\Phi}_t$ are eigenvectors of the shape and texture dispersions estimated from the training set. To obtain a combined shape and texture parameterisation, $\mathbf{c}$, the values of $\mathbf{b}_s$ and $\mathbf{b}_t$ over the training set are combined into

$$\mathbf{b} = \left[\begin{array}{c} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_t \end{array} \right] = \left[\begin{array}{c} \mathbf{W}_s \mathbf{\Phi}_s^{\mathrm{T}} (\mathbf{s} - \overline{\mathbf{s}}) \\ \mathbf{\Phi}_t^{\mathrm{T}} (\mathbf{t} - \overline{\mathbf{t}}) \end{array} \right]. \tag{1}$$

A suitable weighting between pixel distances and pixel intensities is carried out through the diagonal matrix $\mathbf{W}_s$. To recover any correlation between shape and texture the two eigenspaces are usually coupled through a third PC transform,

$$\mathbf{b} = \mathbf{\Phi}_c \mathbf{c} = \left[\begin{array}{c} \mathbf{\Phi}_{c,s} \\ \mathbf{\Phi}_{c,t} \end{array} \right] \mathbf{c}, \tag{2}$$

obtaining the combined appearance model parameters, $\mathbf{c}$, that generate new object instances by $\mathbf{s} = \overline{\mathbf{s}} + \mathbf{\Phi}_s \mathbf{W}_s^{-1} \mathbf{\Phi}_{c,s} \mathbf{c}$ and $\mathbf{t} = \overline{\mathbf{t}} + \mathbf{\Phi}_t \mathbf{\Phi}_{c,t} \mathbf{c}$. The object instance, $\{\mathbf{s}, \mathbf{t}\}$, is synthesised into an image by warping the pixel intensities of $\mathbf{t}$ into the geometry of the shape $\mathbf{s}$ and applying the current pose parameters; translation, rotation, and shape size denoted by $\mathbf{p}$.

An AAM is matched to an unseen image using a least-squares criterion by an iterative updating scheme, which is based on a fixed Jacobian estimate,[14] or originally, a principal component regression.[13] For this work we have used the former approach, treated in further detail elsewhere.[53] In short, this allows model parameters, $\{\mathbf{c}, \mathbf{p}\}$, to be updated based on the difference between the model texture, $\mathbf{t}$, and image texture, $\mathbf{t}_{image}$, the latter being the image texture covered by the model. Let $\mathbf{\Psi}$ denote a fixed update matrix pre-calculated from the training set. Model parameter updates are then estimated by $[\, \delta\mathbf{c}^{\mathrm{T}} \ \delta\mathbf{p}^{\mathrm{T}} \,]^{\mathrm{T}} = \mathbf{\Psi}(\mathbf{t}_{image} - \mathbf{t})$.

For further details on AAMs refer to.[13, 14, 16]

3.2. Bi-temporal Cardiac Modelling in Three Dimensions

Originally, AAMs operated in the two-dimensional domain only. However, a study employing a truly three-dimensional AAM has recently been presented[35, 42, 43] ([42] being the most comprehensive reference). Semi three-dimensional applications have been described in.[3, 4, 10]

While bearing resemblance to the work on 3D cardiac AAMs presented by Mitchell et al.,[42] this study differs in several aspects. Partly, due to the coupled nature of the application and partly due to differences in design
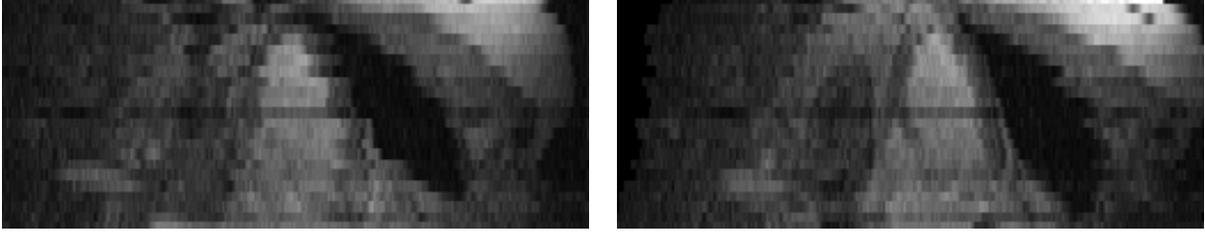
Figure 2: Unsupervised correction of respiratory induced slice displacements. Original yz-plane of patient 4 (left). Corrected yz-plane of patient 4 (right).

and usage of auxiliary methods. We have sought to make this explicit in the following with some mathematical details placed in appendices for completeness. Reproduction of the presented method should be possible from this reference and a detailed description of the 2D AAM as referenced in Section 3.1.

3.2.1. Correction of Respiratory Induced Slice Displacements

Prior to landmarking and subsequent analysis, 4D cardiac cine MRI should optimally be corrected for motion induced by respiration during the acquisition process. This motion is due to the 4D image being constructed from a set of 2D slice image time-series, each having a potentially different inspiration depth, which causes a displacement of the upper abdomen and the heart.

This correction can be carried out by unsupervised image analysis. Statistical models of shape and appearance in 2D are employed to localise the left and right ventricle in a central slice between the apex and base. The segmented region is subsequently propagated through all slices in the apex-basal direction using a non-rigid registration process, thus providing an estimate of the inter-slice translation due to variation in inspiration depth for each breath-hold. The details of this method alongside a quantitative and qualitative validation can be found in.[52] In summary, we found that it was possible to markedly suppress respiratory motion artefacts without manual interaction using less than ten seconds of CPU time for each 4D image. An example result is shown in Figure 2 for patient 4. Notice the obvious reconstruction of cardiac and upper abdominal structures in the left half of Figure 2 (right), while the respiratory-displacements have been "transferred" to the structures not affected by respiration, i.e. the right half of Figure 2 (right).

3.2.2. Point Correspondences on the Myocardium

At the heart of shape modelling lies the requirement of *correspondence*. Which, in the present case, translates to a demand of tissue correspondences on the myocardium; a near landmark-less structure. Due to the simple geometry of the endocardial and epicardial contours, inter-slice correspondences were established by arc-length resampling. The starting point was placed at the anterior junction between the left and right ventricle. Sampling density was chosen to let the endocardial and epicardial surfaces be approximately equally dense on average. The number of sampling points was held constant for all slices. Subsequently, z-axis correspondence was established by linear interpolation of corresponding points between slices. All volumes were resampled to have the average number of slice contours. This inter-slice correspondence enabled a simple conversion from the set of contours to endocardial and epicardial surfaces in a meshing process connecting corresponding points to form strips of triangles, for which surface normals could be estimated.

This crude approach to landmarking, is justified by the combination of scan type (short-axis), the simple geometry of the myocardium, and finally by the low-frequency contours obtained using the common clinical practice (see e.g.[42, 47]) of excluding papillary muscles and ventricular trabeculae from the endocardial contour. More sophisticated approaches to cardiac landmarking were presented in[25, 28] and.[40] Another recent landmarking method based on minimisation of description length[19, 20] is also of interest to this problem, although it currently only handles three-dimensional objects with topology equivalent to a sphere.

3.2.3. Coupling End-diastole and End-systole

To model coherence between the two cardiac phases of interest; the end-diastole and end-systole, a coupled AAM representation is chosen. This was first explored in[18] for interpretation of multi-view face images. Later[41]

coupled cardiac time-sequences, an approach which was also pursued in[9, 11, 12, 37, 50] and for long- and short-axis MRI in.[38] However, most similar to our approach is the recent AAM work presented in,[45, 46] where two-dimensional, end-diastolic and end-systolic angiograms were coupled in shape and appearance, but not in pose due to the nature of the data material. However, in our case, the main objective is to constrain the differences in pose by encoding these into the shape model. This is accomplished by a simple concatenation shown below.

Let a shape consisting of N points in three dimensions be defined by a $3N$ vector,

$$\mathbf{s} = [\ x_1\ y_1\ z_1\ x_2\ y_2\ z_2\ \ldots\ x_N\ y_N\ z_N\ ]^\mathsf{T}. \tag{3}$$

One combined patient ED-ES observation is then composed of three such shapes; $\mathbf{s}_{endo_{ED}}$, $\mathbf{s}_{epi_{ED}}$ and $\mathbf{s}_{endo_{ES}}$ by concatenation prior to Procrustes alignment and tangent space projection,

$$\mathbf{s}_{subject} = \left[\begin{array}{c} \mathbf{s}_{endo_{ED}} \\ \mathbf{s}_{epi_{ED}} \\ \mathbf{s}_{endo_{ES}} \end{array} \right]. \tag{4}$$

Likewise, the associated image texture is composed from samplings of two 3D images; the volume spanned by the blood pool and the myocardium from the end-diastolic phase, $\mathbf{t}_{ED}$, and the blood pool of the end-systolic phase, $\mathbf{t}_{ES}$. The combined texture vector then becomes $\mathbf{t} = [\ \mathbf{t}_{ED}^\mathsf{T}\ \mathbf{t}_{ES}^\mathsf{T}\ ]^\mathsf{T}$.

3.2.4. Procrustes Alignment

Building a model of shape variability as in AAMs, requires per se elimination of effects stemming from translation, rotation and isotropic scaling. This is accomplished by an Ordinary Procrustes Analysis (OPA) bringing two shapes into alignment, or by a Generalised Procrustes Analysis (GPA) aligning a set of shapes to their common mean. Unlike the impression given in the earlier 3D and semi-3D AAM work,[10, 42, 43] well-behaved solutions to both problems do exist in two and three dimensions formulated in terms of orthonormal matrices, see.[2, 21, 32, 55]

In summary, closed-form OPA solutions exist in 2D and 3D, while a closed-form GPA solution only exists in 2D. In the current work, we employ OPA in 3D based on a singular value decomposition (SVD) of the inter-shape correlation matrix, avoiding the quaternion formulation used in[10, 42, 43] altogether. We obtain the GPA solution by a simple iterative scheme, see e.g.[8, 16] The mathematical details of OPA in k dimensions are given in Appendix A.

3.2.5. Texture Correspondences

Two-dimensional image warping using a triangular mesh is a fast – and often sufficient – solution to the ill-conditioned problem of establishing a dense planar correspondence from two sparse sets of corresponding points. Typically, a Delaunay triangulation of one of the point sets is used to form this mesh of two-dimensional simplexes. See e.g.[16, 51] for the details of this warping process that also can be accelerated using contemporary graphics hardware.[53]

In three dimensions, the Delaunay decomposition produces three-dimensional simplexes, called tetrahedra. Alternatively, these tetrahedra can be produced by a dedicated hand-tailored method as in done in.[42] To avoid this tedious process, we employ the freely available 3D Delaunay implementation TetGen[49] to produce a tetra-hedral decomposition of the mean shape. This ensures a rapid and unsupervised model building process, which is adaptive to the mean geometry, $\overline{\mathbf{s}}$, encountered in a given domain problem.

Let $\mathbf{b}_i = [\ \alpha\ \beta\ \gamma\ \delta\ ]^\mathsf{T}$ denote the barycentric coordinates† of a tetrahedron given by $\mathbf{T}_i$. Then, let $\mathcal{T} = \{\mathbf{T}_i\}_1^R$ denote the set of R tetrahedra in the mean configuration (vertices incident to the mean shape) and let $\mathcal{T}' = \{\mathbf{T}_i'\}_1^R$ denote a corresponding set of deformed tetrahedra. Using the notation of Appendix B, warping a point from $\mathcal{T}$ to $\mathcal{T}'$ can then be given by

$$\mathbf{p}_i' = \mathcal{W}(\mathbf{p}_i) = \mathcal{W}(\mathbf{T}_i\mathbf{b}_i) = \mathbf{T}_i'\mathbf{b}_i. \tag{5}$$

Consequently, during model generation, a set of texture sample positions of the mean tetrahedral configuration are chosen, $\{\mathbf{p}_i\}_1^M$, and their corresponding barycentric coordinates, $\{\mathbf{b}_i\}_1^M$, are determined (using Appendix B)

---

†See Appendix B.

and stored. Texture sample positions for arbitrary tetrahedral configurations during AAM search are now obtained by simple matrix-vector multiplications.

In the current work, texture samples are distributed equidistantly with a spacing of one voxel within the axis-aligned bounding box containing the mean shape. This involves solving the point-in-tetrahedra problem; a task that can be substantially sped up by a spatial search tree of the tetrahedra. However, being an off-line process, a (slow) linear search using an axis-aligned bounding box test for each tetrahedron, followed by the exact test in Appendix B, was sufficient in our case.

A computationally more demanding alternative to the above could be to choose the three-dimensional warp field that minimises the bending energy (the integral over $\mathbb{R}^3$ of the squares of the second derivatives). This solution is often referred to as thin-plate splines, see e.g.[7]

Lastly, notice that in the bi-temporal case, two separate warping functions must be established from two tetrahedral decompositions; one for end-diastole and one for end-systole.

3.2.6. Texture Sampling

As stated above, all elements in $\{\mathbf{p}_i\}_1^M$ remain fixed and – to ease synthesis of model textures – the values of these are chosen to take on integer values forming a regular lattice. Elements of $\{\mathbf{p}_i'\}_1^M$ will thus typically be irrational numbers for which no image data exist. Consequently, an interpolation method is required. For computational efficiency, this work uses the commonly used tri-linear interpolation (the volumetric equivalent of bi-linear interpolation in 2D).

Alternatively, a more optimal reconstruction could be obtained by sinc interpolation, which unfortunately exhibits excessive computational demands compared to tri-linear interpolation. However, it was recently shown that the major error component from premature truncation of the sinc kernel can be compensated for by using a renormalisation strategy.[56] Alternatively, sophisticated adaptive filtering strategies can be taken. See e.g. the impressive results on interpolation of anisotropic voxel data from computed tomography in.[59]

3.2.7. Proximity Modelling

In medical images, organs and structures typically have a very homogenous appearance. Since AAMs only model the appearance spanned by the triangular (2D) or tetrahedral (3D) mesh, no information regarding the surroundings is modelled if the outer landmarks are placed exactly on the object boundary. This is naturally typically the case, as this is the desired structure the model should infer knowledge about.

The two above issues lead to models, which can show minima in the model-to-image cost function when the outer surface (or contour in 2D) is erroneously *inside* the object of interest. In the present case, the epicardial surface could be between the blood pool and the lung. Therefore, we apply a simple scheme for modelling the texture variation of the object proximity. This is accomplished by adding surface normals outwards from the shape. These normals are denoted *whiskers* and added implicitly during texture sampling with a scale relative to the current shape size. Texture samples obtained by sampling along whiskers are now concatenated to the texture vector of the conventional AAM texture samples.

3.2.8. Texture Model Pruning

AAMs build on assumptions of normally distributed data. Consequently, the seemingly absence and presence of papillary muscles are per se not well modelled in the texture model. Therefore we propose a pruning strategy, which excludes these problematic texture sample positions. In lieu of tedious manual pruning; a "suitability" measure, $\mathcal{S}(t_i)$, is used to obtain such pruning.

Optimally, $\mathcal{S}$ should measure the non-Gaussianity; a striking resemblance to a problem occurring in the estimation of independent components in ICA (Independent Component Analysis). In[33] this is solved by exploiting a central result from information theory, namely that a Gaussian distributed random variable exhibits maximal entropy among all random variables of equal variance. This leads to the concept of *negentropy*, that[33] presents computationally feasible estimators for, which are far more robust than the traditional kurtosis measure.

However, due to the small sample size of our data (twelve samples per element in $\mathbf{t}$), texture models are pruned on the basis of variance; $\mathcal{S}(t_i) = \sum_{j=1}^{Q}(t_{i_j} - \bar{t}_{i_j})^2$. Texture positions above a user-selected percentile of $\mathcal{S}$ are now excluded from the texture samples of the training set and omitted in future samplings during image search.

3.2.9. Pose Representation

Pose in three dimensions can be parameterised by seven components,

$$\mathbf{d} = [\ t_x\ t_y\ t_z\ \theta_x\ \theta_y\ \theta_z\ 1-s\ ], \tag{6}$$

denoting spatial translation, rotation and one minus the isotropic scaling, respectively. Having the identity transformation represented by the null vector, this representation is suitable for representing pose displacements in the aligned shape domain employed by AAMs. Let us denote this domain $\mathcal{A}$ and let the image shape domain – where landmarks are given in terms of voxel positions – be denoted by $\mathcal{I}$. A convenient representation of the mapping $\mathcal{I} \mapsto \mathcal{A}$ is

$$\mathcal{Q} = \{\ t_x,\ t_y,\ t_z,\ s,\ \mathbf{R}\ \} \tag{7}$$

where $\mathbf{R} \in \mathbb{R}^{3\times3}$ is an orthonormal matrix rotating a shape in $\mathcal{I}$ to its reference orientation in $\mathcal{A}$ as obtained by OPA. Shape displacement experiments required during the estimation of $\mathbf{\Psi}$ in the AAM building process can now be composed by $\mathbf{d} * \mathcal{Q}$, i.e. a shape displacement in $\mathcal{A}$ and a mapping (fixed for each training shape) to $\mathcal{I}$.

3.2.10. Systole Detection

To render ejection fraction estimation using a bi-temporal 3D model completely unsupervised, the systolic phase needs to be estimated. Due to the flexibility of the model, this can be straightforward estimated by running several searches on different time phases near the central part of the time-series and choose the search result maximising the difference between the endocardial volumes (i.e. the stroke volume);

$$\hat{t}_{systole} = \arg\max_t \left(\text{Volume}\{\mathbf{s}_{endo_0}\} - \text{Volume}\{\mathbf{s}_{endo_t}\}\right). \tag{8}$$

4. IMPLEMENTATION

The presented bi-temporal 3D AAM framework was implemented in standard C++ with help from three major third-party libraries. Singular value decomposition and eigenvalue solver were provided by LAPACK.[1] A fast, architecture-specific BLAS implementation for matrix-vector multiplication was provided by the Intel Math Kernel Library.[34] Three-dimensional Delaunay decomposition was provided by TetGen.[49] In addition to the AAM implementation, a framework for data visualisation in four dimensions were written in C++ using OpenGL.[24] This produced all figures in this article showing MRI data and geometry.

To obtain a reasonable level of performance, aggressive C++ *inlining* has been used; in particular for the warping code. Computational complexity and cache consistency were considered during design and implementation, but neither code nor memory optimisation have been applied in the current implementation.

5. EXPERIMENTAL RESULTS

To assess the ability of the described AAM framework to locate the left ventricle in 4D cine MRI, a series of cross-validation experiments were carried out on a 2 GHz Pentium PC equipped with 1 GB RAM. Due to the small sample size (twelve patients), a leave-one-out validation approach was chosen. Hence, twelve models were built on eleven training examples and the remaining example was used for testing. For each experiment, the bi-temporal AAM searched the end-diastolic phase and five central frames to locate the end-systole. In each of these five searches, the AAM was initialised in a 3×3 grid in the xy slice plane around the mean pose configuration as determined from the (independent) training set. The grid spacing used was 33% of the mean shape width in x and y, respectively. This grid size showed sufficient coverage of the pose variation in the current training set, since at least one of the nine searches converged onto the true location of the heart in all patients. For other data sets, a different grid size may be needed. In total, $12 \times 5 \times 9 = 540$ AAM searches were carried out.

All shape models contained 1560 points, i.e. 4680 variables. To reduce memory requirements and speed up computations, all input images were subsampled in x and y to 128×128, z remained unchanged. Texture models contained approximately 22000 texture samples. The proximity of the left ventricle was modelled by adding whiskers at each landmark with a length of five voxels (measured on the mean shape sized to mean size). Five

Table 1: Unsigned and signed errors for the estimated cardiac volumes and ejection fraction using automatic systole detection.

	Unsigned mean errror	Signed mean error
Endo EDV	6.8±3.7 ml	-0.7±7.9 ml
Endo ESV	3.9±2.5 ml	1.3±4.6 ml
Epi EDV	23.8±17.2 ml	5.6±29.7 ml
EF	2.8±1.9 percent	-1.2±3.3 percent

Table 2: Unsigned and signed errors for the estimated cardiac volumes and ejection fraction using ground truth systole.

	Unsigned mean errror	Signed mean error
Endo EDV	6.5±5.2 ml	2.5±8.1 ml
Endo ESV	3.8±2.3 ml	0.8±4.5 ml
Epi EDV	23.3±24.8 ml	13.1±32.0 ml
EF	2.1±2.0 percent	0.0±3.0 percent

samples for each whisker were acquired during texture sampling. The texture model pruning retained the 70% samples showing the least variance over the training set. Subsequently, shape and texture PCA models were truncated to explain 98% variation in the training set (see e.g.[53]). To gain increased flexibility from the limited training set, shape and texture eigenspaces were not coupled, i.e. $\Phi_c = \mathbf{I}$.

Table 1 shows the signed and unsigned mean errors, between the AAM and the human observer, of the end-diastolic volumes (EDV), end-systolic volume (ESV) and ejection fraction (EF). A more detailed impression of agreement between these two methods for estimation of global functional cardiac indices is given by the Bland-Altman plots[5] in Figure 3 (left). The accuracy obtained in each patient is illustrated in Figure 3 (right).

Qualitative results are given for patient 4 in Figure 4, showing the initial and converged model geometry overlaid on the MRI. Movie sequences illustrating the model convergence can be obtained electronically by request to the corresponding author.
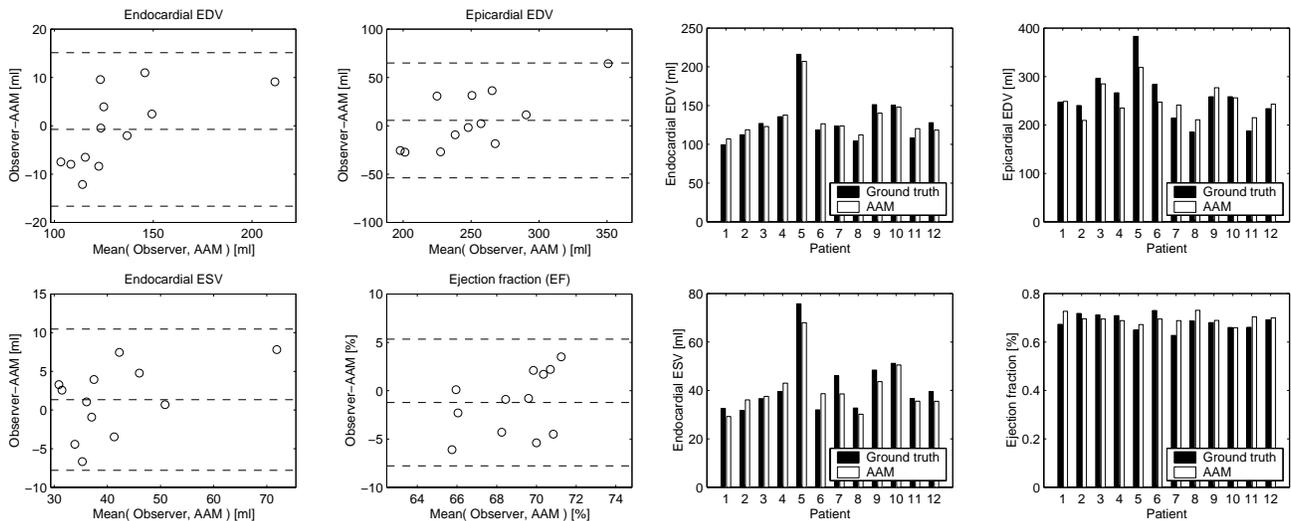

Figure 3: Left: Bland-Altman plots for end-diastolic endocardial volume, end-diastolic epicardial volume, end-systolic endocardial volume, and left ventricular ejection fraction. Limits of agreement are plotted as dashed lines at ±2 standard deviations. Right: Bar plots for end-diastolic endocardial volume, end-diastolic epicardial volume, end-systolic endocardial volume, and left ventricular ejection fraction.
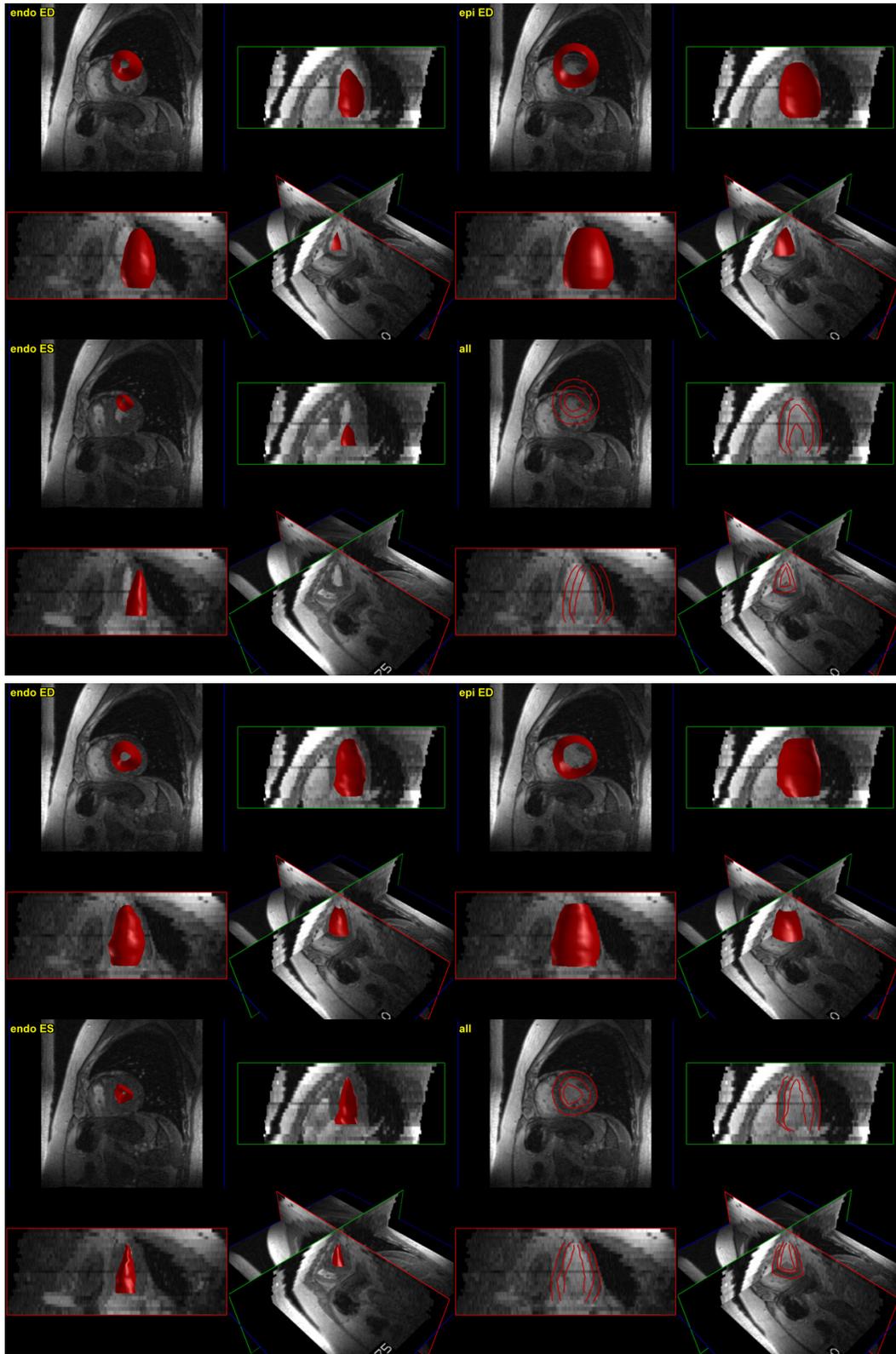
Figure 4: End-diastolic, end-systolic, endocardial and epicardial surfaces shown in the initial (top) and optimal (bottom) model configuration for patient 4 after 17 iterations.

Table 3: Model building and model search timings using automatic systole detection.

	All experiments [seconds]	Per experiment [seconds]
Model building	571.0	47.6±2.4
Model search	217.3	18.1±3.5
Average single phase search	43.5	3.6
Average single search	4.8	0.4

The systole detection determined the ground truth time phase given by the human observer in five cases (42%). In six cases (50%) the neighbouring phase was deemed the systole. In one case (8%) the detected systole and the observer-given differed by two phases. All seven systolic phases that did not match the observer-given phase were detected in a later phase. To assess the potential confounding between systole detection and ventricular volume estimation, a set of leave-one-out experiments using the ground truth systole was carried out. A summary of these results is given in Table 2.

A single leave-one-out experiment peaked with respect to memory usage at 153 MB in this unoptimised implementation. Model building and model search timings are shown in Table 3. The figures in the third and fourth row are obtained by dividing the second row with the number of phase searches (five), and the third row by the number of initialisation grid sites (nine). One image texture sampling took 15 ms on average.

In comparison to Table 2, a similar set of experiments was carried out, but with no pruning of the texture model, no whiskers and using the traditional coupled shape and texture eigenspaces. The unsigned mean errors were: Endo EDV=10.0±9.1 ml; Endo ESV=5.5±4.0 ml; Epi EDV=28.6±23.3 ml; EF 3.2±1.7 percent.

Repeating the experiments in Table 2, but without the described slice correction procedure showed a markedly decrease in accuracy for all figures but the Epi EDV. The unsigned mean errors were: Endo EDV=11.7±9.3 ml; Endo ESV=6.7±5.0 ml; Epi EDV=24.6±17.6 ml; EF 3.1±2.5 percent.

Finally, to prove that full-resolution modelling of the input data was possible, an additional set of experiments was carried out. These were similar to the ones in Table 2, but with no x and y subsampling. Results were comparable, but errors were – although modestly – consistently higher compared to the equivalent subsampled results.

6. DISCUSSION

The simple grid-based initialisation strategy succeeded in localising the heart in all twelve patients. This success may in part be attributed to the relatively small training set showing a moderate variation in pose. For larger training sets, it may be needed to perform leave-one-out analyses to determine the limits of convergence for each pose and model parameter. These, in combination with the observed distribution of pose and model parameters, will then determine the required minimal grid size and spacing to guarantee successful generic initialisation. Additionally, a multi-resolution implementation can be used to speed up this process.

Our experiments showed that the end-systole localisation obtained is acceptable. The results in Table 1 are – by design – biased towards larger end-diastolic volumes and smaller end-systolic volumes. This, and the fact that only five out of twelve end-systolic phases were correctly classified, are still considered acceptable. The end-systolic phases (and thereby the associated volume) were manually determined from a single slice image only. On the contrary, the model was able to search for a global volumetric minimum. Further, considering the coarse temporal sampling and the pronounced bias towards a later phase in the detected systoles, we consider the automatically determined end-systolic phases on equal terms to the manually given.

Quantitatively, we observe that our method handles challenges such as epicardial fat well due to its constrained nature. See e.g. the epicardial contours in Figure 4. However, inspection of the obtained results revealed that determination of the extreme basal and apical slice still poses a challenge as it does for the manual observer. This is primarily due to the anisotropy of the voxels, which only allows for a slice-based annotation. To avoid this, high-definition models can be built based on data from sophisticated acquisition schemes such as[6] where

a much improved coverage of the LV is obtained. Such models could subsequently be used to search sparse, short-axis slice-oriented data similar to the data in this study. We anticipate that results from such a strategy will outperform human performance due to the vastly more detailed prior model.

The obtained volume errors were fair except for one outlier in both size and shape that also showed pronounced image artefacts (patient 5). Epicardial volumes (used for calculating the mass of myocardium) showed a relative error, which was markedly higher than for the endocardial volumes.

Although not directly comparable due to differences in scanners, pulse sequences, subjects, et cetera, the variability of the obtained ejection fraction estimates compares very well to the inter-observer variability in a recent MRI-based ejection fraction study.[47] Further, limits of agreement for end-systolic volumes and end-diastolic volumes are near-identical to those presented in,[44] which concluded on basis of these that a new method for assessing ESV and EDV was accurate and produced results comparable to the established method. Due to this similarity, we see a clear indication that the presented method can replace the current resource-demanding and tedious practice of manual annotation required to estimate cardiac blood pool volumes and ejection fraction.

The usage of texture model pruning, whiskers and decoupling of shape and texture eigenspaces in combination was showed to have a markedly positive influence on the results. So did the correction of respiratory induced slice displacements. However, using full resolution images compared to a decimation in the x and y axes increased the error slightly. This suggests that the regularisation obtained by the subsampling was desirable, i.e. it eliminated more noise structures, than it deteriorated the latent tissue structure. In this conjunction, it may be interesting to apply a more principled and data-driven regularisation using a truncated wavelet texture representation as introduced for AAMs in[60] and further explored in.[54]

Since this work (arguably) represents the first application of 3D AAMs independent from the work by Mitchell and co-workers[42] some comparisons may be of value to the reader to outline differences and extensions. Contrary to the single image 3D AAM in,[42] this work features a bi-temporal 3D AAM applied to 4D images using automated systole detection. The tetrahedral warper employs a general 3D Delaunay decomposition, contrary to application-tailored tetrahedral meshes. Ordinary Procrustes alignment was based on a simple SVD-based solution, contrary to – admittedly elegant, but not so well-known and more involved – quaternion-based alignment. The problem of respiratory motion mentioned in[42] has been addressed and assessed in our data. Non-Gaussian texture appearance caused by papillary muscles has been addressed by a texture model pruning mechanism. A simple method of adding proximity sampling to the texture model was employed, contrary to application-tailored tetrahedra. Running multiple searches, contrary to a dedicated search method initialised the models. Finally, our timings indicate a substantially faster implementation. In,[42] it took 2–3 minutes to optimise a single-image 3D AAM containing approximately 7000 voxel using a 1 GHz PC. In our setting, a 3×3 grid search using a bi-temporal 3D AAM containing approximately 22000 voxels took 3.4 seconds on average using a 2 GHz PC. Although this indicates a significant difference in performance, we emphasise that a direct comparison is not possible, due to the different requirements stemming from different data sets, et cetera.

We hypothesize three major issues that limit the accuracy obtainable from the described approach. Primarily, the shape PCA may not be able to approximate new examples from a training set of only eleven examples. This may also be true for the texture model, which may cause uninteresting minima in the cost function evaluated during the model-to-image fitting process. Judged from the image appearance in the data set, we believe the shape PCA to be the limiting factor. Next is the issue of the extreme basal and apical slice, which is more a data problem than a model problem. As this also pertains to the ground truth solution it may be hard to address and evaluate. In this matter, we have outlined a solution above based on additional image data. Finally, the naïve approach taken to establishment of point correspondences may prove insufficient. In particular, if more detailed models of the myocardium, or local measures such as wall thickness, et cetera, are required in future studies. Fortunately, this topic is well described in the literature and some alternatives are given above.

As a concluding remark; statistical approaches similar to the described method are ultimately limited by the sample size and how representative the training set is. Consequently, hard conclusions regarding the general performance of such a method will remain fragile until more training data becomes available.

7. CONCLUSION

A method for unsupervised estimation of ventricular volumes, masses and ejection fraction from four-dimensional cardiac cine MRI has been presented. The learning-based method establishes statistical models of myocardial shape and appearance from a training set of end-diastolic and end-systolic images, which enables it to rapidly infer global cardiac indices from unseen 4D images contaminated with substantial image artefacts.

Cross-validation of the method carried out on clinical-quality scans of twelve volunteers indicates that ejection fraction and cardiac blood pool volumes can be estimated automatically and rapidly with an accuracy comparable to the typical inter-observer variability.

APPENDIX A. ORDINARY PROCRUSTES ALIGNMENT IN K DIMENSIONS

This appendix serves to demonstrate the validity of the solution to the orthogonal ordinary Procrustes alignment problem for two shapes in k dimensions (in our case $k = 3$). Let $\mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{N \times k}$ denote two shapes with centroids incident with the origin. The optimal rotation – in a least-squares sense – that superimposes $\mathbf{X}_2$ onto $\mathbf{X}_1$, is given by

$$\arg \min_{\mathbf{R}} ||\mathbf{X}_1 - \mathbf{X}_2\mathbf{R}||^2, \tag{9}$$

where $||\mathbf{M}|| = \sqrt{\mathrm{tr}(\mathbf{M}^\mathsf{T}\mathbf{M})}$ denotes the Frobenius norm, and $\mathbf{R}$ is an orthonormal matrix subject to $\det(\mathbf{R}) = 1$. Now consider the following expansion

$$
\begin{aligned}
\mathbf{P} &= (\mathbf{X}_1 - \mathbf{X}_2\mathbf{R})^\mathsf{T}(\mathbf{X}_1 - \mathbf{X}_2\mathbf{R}) \\
&= \mathbf{X}_1^\mathsf{T}\mathbf{X}_1 - \mathbf{X}_1^\mathsf{T}\mathbf{X}_2\mathbf{R} - \mathbf{R}^\mathsf{T}\mathbf{X}_2^\mathsf{T}\mathbf{X}_1 + \mathbf{R}^\mathsf{T}\mathbf{X}_2^\mathsf{T}\mathbf{X}_2\mathbf{R},
\end{aligned}
\tag{10}
$$

and its trace

$$
\begin{aligned}
\mathrm{tr}(\mathbf{P}) &= \mathrm{tr}(\mathbf{X}_1^\mathsf{T}\mathbf{X}_1) - \mathrm{tr}(\mathbf{X}_1^\mathsf{T}\mathbf{X}_2\mathbf{R}) - \mathrm{tr}((\mathbf{R}^\mathsf{T}\mathbf{X}_2^\mathsf{T}\mathbf{X}_1)^\mathsf{T}) + \mathrm{tr}((\mathbf{R}^\mathsf{T}\mathbf{X}_2^\mathsf{T}\mathbf{X}_2\mathbf{R})^\mathsf{T}) \\
&= \mathrm{tr}(\mathbf{X}_1^\mathsf{T}\mathbf{X}_1 + \mathbf{X}_2^\mathsf{T}\mathbf{X}_2) - 2\,\mathrm{tr}(\mathbf{X}_1^\mathsf{T}\mathbf{X}_2\mathbf{R}).
\end{aligned}
\tag{11}
$$

Minimising (9) is thus equivalent to maximising $\mathrm{tr}(\mathbf{X}_1^\mathsf{T}\mathbf{X}_2\mathbf{R})$. This is achieved if, and only if, $\mathbf{X}_1^\mathsf{T}\mathbf{X}_2\mathbf{R}$ is symmetric and positive semi definite. See[55] for a proof.

Consider the following singular value decomposition (see e.g.[29, 48])

$$\mathbf{X}_1^\mathsf{T}\mathbf{X}_2 = \mathbf{U}\mathbf{S}\mathbf{V}^\mathsf{T}, \tag{12}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices (called singular vectors) and $\mathbf{S}$ is a diagonal matrix of singular values. The optimal orthonormal solution to (9) is claimed to be

$$\mathbf{R} = \mathbf{V}\mathbf{U}^\mathsf{T}. \tag{13}$$

It now follows directly that

$$\mathbf{X}_1^\mathsf{T}\mathbf{X}_2\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^\mathsf{T}\mathbf{V}\mathbf{U}^\mathsf{T} = \mathbf{U}\mathbf{S}\mathbf{U}^\mathsf{T} \tag{14}$$

is symmetric and positive semi definite since all eigenvalues – given by $\mathbf{S}$ – are nonnegative.

However, $\mathbf{V}\mathbf{U}^\mathsf{T}$ only holds a rotation matrix when $\mathbf{U}\mathbf{S}\mathbf{U}^\mathsf{T}$ is symmetric and positive definite. In the rare case where not all singular values are nonnegative, the correlation matrix, $\mathbf{X}_1^\mathsf{T}\mathbf{X}_2\mathbf{R}$, can be maximized using a reflection, i.e. $\det(\mathbf{R}) = -1$. A solution to (9) that ensures rotation is

$$\mathbf{R} = \mathbf{V}\boldsymbol{\Delta}\mathbf{U}^\mathsf{T}, \tag{15}$$

where $\boldsymbol{\Delta} = \mathrm{diag}(\,[\,1\,\dots\,1\,\det(\mathbf{V}\mathbf{U}^\mathsf{T})\,]\,) \in \mathbb{R}^{k \times k}$. Refer to[2, 21, 31, 32, 55, 58] for further elaborations.

APPENDIX B. BARYCENTRIC COORDINATES OF A TETRAHEDRON

Let $\mathbf{b} = [\,\alpha\,\beta\,\gamma\,\delta\,]^\mathsf{T}$ denote the barycentric coordinates of a simplex in three dimensions; a tetrahedron. Subject to the two constraints, $1 = \alpha + \beta + \gamma + \delta$ and $0 \le \alpha, \beta, \gamma, \delta \le 1$, these coordinates will uniquely determine any position, $\mathbf{p}$, within the tetrahedron, $\mathbf{T}$ (given by the point set $\{x_i, y_i, z_i\}_1^4$),

$$
\mathbf{p} = \mathbf{Tb} \quad \Leftrightarrow \quad
\begin{bmatrix} 1 \\ x \\ y \\ z \end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 1 & 1 \\
x_1 & x_2 & x_3 & x_4 \\
y_1 & y_2 & y_3 & y_4 \\
z_1 & z_2 & z_3 & z_4
\end{bmatrix}
\begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}.
\tag{16}
$$

Provided that $\mathbf{T}$ has full rank the barycentric coordinates corresponding to the point $\mathbf{p}$ are thus given by $\mathbf{b} = \mathbf{T}^{-1}\mathbf{p}$ and expanded below, where $\det(\mathbf{T})$ is denoted by d

$$
\begin{aligned}
\alpha =\ -\tfrac{1}{d}(\ & -x_3y_2z + x_3yz_2 + yx_2z_4 - z_4xy_2 + z_4xy_3 - y_4x_2z - yx_4z_2 + y_2zx_4 - x_2z_3y + \\
& zx_3y_4 - zx_4y_3 - z_4x_3y + z_3yx_4 - z_3y_4x + xy_2z_3 - xy_3z_2 + x_2zy_3 + y_4xz_2 - \\
& x_4y_2z_3 + x_4y_3z_2 + x_3y_2z_4 - x_3y_4z_2 + x_2z_3y_4 - x_2z_4y_3\ \ ),
\end{aligned}
\tag{17}
$$

$$
\begin{aligned}
\beta =\ \tfrac{1}{d}(\ & -z_1yx_4 - z_1xy_3 + z_1y_4x - z_1x_3y_4 + z_1x_4y_3 + z_1x_3y - x_1z_3y - x_3y_1z + z_4xy_3 + \\
& yx_1z_4 + y_1zx_4 - y_1z_4x - y_4x_1z + zx_3y_4 - zx_4y_3 - z_4x_3y + z_3yx_4 - z_3y_4x + \\
& xy_1z_3 + x_1zy_3 - x_1z_4y_3 + x_1z_3y_4 - x_4y_1z_3 + x_3y_1z_4\ \ ),
\end{aligned}
\tag{18}
\tag{19}
$$

$$
\begin{aligned}
\gamma =\ -\tfrac{1}{d}(\ & -x_2y_4z_1 + x_4z_1y_2 - xz_1y_2 + x_2yz_1 - z_1yx_4 + z_1y_4x + xz_2y_1 - y_4x_1z - y_1z_4x - \\
& y_2zx_4 - x_1yz_2 + x_1y_4z_2 - x_4z_2y_1 + x_1zy_2 - y_4xz_2 - x_1z_4y_2 - yx_2z_4 + yx_4z_2 + \\
& yx_1z_4 + z_4xy_2 + x_2y_1z_4 - x_2y_1z + y_4x_2z + y_1zx_4\ \ ),
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
\delta =\ \tfrac{1}{d}(\ & -xz_1y_2 + z_1xy_3 - x_2y_3z_1 + x_2yz_1 - z_1x_3y + x_3z_1y_2 - x_2y_1z + x_1z_3y - x_3y_2z + \\
& x_3yz_2 + x_3y_1z - x_1z_3y_2 - x_2z_3y - x_1yz_2 + xz_2y_1 + x_1y_3z_2 - x_3z_2y_1 - xy_1z_3 + \\
& xy_2z_3 - xy_3z_2 + x_1zy_2 - x_1zy_3 + x_2zy_3 + x_2y_1z_3\ \ ),
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
d =\ & -x_2y_3z_1 + x_2y_4z_1 - z_1x_3y_4 - x_4z_1y_2 + z_1x_4y_3 + x_3z_1y_2 - x_1z_3y_2 + x_1y_3z_2 - \\
& x_3z_2y_1 + x_2y_1z_3 - x_1y_4z_2 + x_4z_2y_1 - x_2y_1z_4 - x_1z_4y_3 + x_4y_2z_3 + x_1z_3y_4 - \\
& x_4y_3z_2 - x_4y_1z_3 + x_3y_1z_4 - x_3y_2z_4 + x_3y_4z_2 - x_2z_3y_4 + x_2z_4y_3 + x_1z_4y_2.
\end{aligned}
\tag{22}
$$

Degenerate tetrahedra (having four coplanar points) will thus have $d = 0$, i.e. no inverse.

REFERENCES

1. E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide.* Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999. Web page: http://www.netlib.org/lapack/.
2. K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Trans. on Pattern Anal. and Machine Intelligence*, 9(5):698–700, 1987.
3. R. Beichel, G. Gotschuli, E. Sorantin, F. W. Leberl, and M. Sonka. Diaphragm dome surface segmentation in CT data sets: a 3D active appearance model approach. *SPIE Medical Imaging 2002*, pages 475–484, 2002.
4. R. Beichel, S. Mitchell, E. Sorantin, F. Leberl, A. Goshtasby, and M. Sonka. Shape- and appearance-based segmentation of volumetric medical images. *IEEE International Conference on Image Processing*, 2:589–592, 2001.
5. J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 1(8476):307–310, 1986.
6. D. C. Bloomgarden, Z. A. Fayad, V. A. Ferrari, B. Chin, M. G. St. John Sutton, and L. Axel. Global cardiac function using fast breath-hold MRI: Validation of new acquisition and analysis techniques. *Magnetic Resonance in Medicine*, 37(5):683–92, 1997.

7. F. L. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–85, 1989.

8. F. L. Bookstein. Landmark methods for forms without landmarks: localizing group differences in outline shape. *Medical Image Analysis*, 1(3):225–244, 1997.

9. H. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber. Active appearance-motion models for endocardial contour detection in time sequences of echocardiograms. *Proceedings of SPIE - The International Society for Optical Engineering*, 4322(1):257–268, 2001.

10. J. G. Bosch, S. C. Mitchell, B. P. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. Reiber. Fully automated endocardial contour detection in time sequences of echocardiograms by three-dimensional active appearance models. *SPIE Medical Imaging 2002*, pages 452–462, 2002.

11. J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber. Fully automated endocardial contour detection in time sequences of echocardiograms by active appearance motion models. *Computers in Cardiology*, pages 93–96, 2001.

12. J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber. Automatic segmentation of echocardiographic sequences by active appearance motion models. *Medical Imaging, IEEE Transactions on*, 21(11):1374–1383, 2002.

13. T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In *Proc. of European Conf. on Computer Vision 1998*, volume 1407 of *Lecture Notes in Computer Science*, pages 484–498. Springer, 1998.

14. T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.

15. T. F. Cootes and Taylor. Active shape models – 'smart snakes'. In *Proc. British Machine Vision Conf., BMVC92*, pages 266–275, 1992.

16. T. F. Cootes and C. J. Taylor. *Statistical Models of Appearance for Computer Vision*. Tech. report, University of Manchester, 2001.

17. T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham. Training models of shape from sets of examples. In *Proc. British Machine Vision Conf., BMVC92*, pages 9–18, 1992.

18. T. F. Cootes, G. V. Wheeler, K. N. Walker, and C. J. Taylor. Coupled-view active appearance models. *BMVC 2000*, pages 52–61 vol.1, 2000.

19. R. H. Davies. *Learning Shape: Optimal Models for Analysing Natural Variability*. PhD thesis, Department of Imaging Science and Biomedical Engineering, University of Manchester, Stopford building, University of Manchester, 2002.

20. R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. 3D statistical shape models using direct optimisation of description length. In *Proc. ECCV 2002*, volume 2352 of *LNCS*, pages 3–20. Springer, 2002.

21. I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis*. John Wiley & Sons, 1998.

22. J. P. Earls, V. B. Ho, T. K. Foo, E. Castillo, and S. D. Flamm. Cardiac MRI: Recent progress and continued challenges. *Journal of Magnetic Resonance Imaging*, 16(2):111–127, 2002.

23. G. J. Edwards, C. J. Taylor, and T. F. Cootes. Interpreting face images using active appearance models. In *Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pages 300–5. IEEE Comput. Soc, 1998.

24. SGI et al. Opengl. Web page: http://www.opengl.org/.

25. A. F. Frangi. *Three-dimensional model-based analysis of vascular and cardiac images*. PhD thesis, Image Sciences Institute, University Medical Center, Utrecht University, The Netherlands, 2001.

26. 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, 2001.

27. A. F. Frangi, D. Rueckert, and J. S. Duncan. Three-dimensional cardiovascular image analysis. *Medical Imaging, IEEE Transactions on*, 21(9):1005–1010, 2002.

28. A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen. Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling. *IEEE Transactions on Medical Imaging*, 21(9):1151–1166, 2002.

29. G. H. Golub and C. F. van Loan. *Matrix Computations*. Jonhs Hopkins, 1983.

30. J. C. Gower. Generalized Procrustes analysis. *Psychometrika*, 40:33–50, 1975.

31. D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46(3):R1–45, 2001.

32. B. K. P. Horn, H. M. Hilden, and S. Negahdaripour. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society of America A (Optics and Image Science)*, 5(7):1127–35, 1988.

33. A. Hyvarinen and E. Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13(4-5):411–430, 2000.

34. Intel. Math kernel library. Web page: http://developer.intel.com/software/products/mkl/.

35. B. Lelieveldt, R. van der Geest, S. C. Mitchell, J. Bosch, M. Sonka, and J. Reiber. 3D active appearance models: Fully automatic detection of endo- and epicardial contours in short-axis cardiac MR data. In *Proc. International Society of Magnetic Resonance In Medicine, ISMRM*, volume 2, page 1668. ISMRM, 2002.

36. B. P. F. Lelieveldt. *Anatomical models in cardiovascular image analysis.* PhD thesis, Department of Radiology, Leiden University, The Netherlands, 1999.

37. B. P. F. Lelieveldt, S. C. Mitchell, R. J. van der Geest, H. G. Bosch, M. Sonka, and J. H. C. Reiber. Time continuous segmentation of cardiac MR images using active appearance motion models. In *Computer Assisted Radiology and Surgery, CARS 2001, International Congress Series*, volume 1230, pages 961–966. Elsevier Science, 2001.

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

39. D. C. Levin, L. Parker, J. H. Sunshine, and M. J. Pentecost. The practice of radiology – cardiovascular imaging: Who does it and how important is it to the practice of radiology? *American Journal of Roentgenology*, 178(2):303–306, 2002.

40. H. Luo and T. O'Donnell. A 3D statistical shape model for the left ventricle of the heart. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2001, 4th International Conference, Utrecht, The Netherlands*, volume 2208 of *Lecture Notes in Computer Science*. Springer, 2001.

41. S. Mitchell, B. Lelieveldt, R. Geest, H. Bosch, J. Reiber, and M. Sonka. Time continuous segmentation of cardiac MR image sequences using active appearance motion models. In *Proc. of SPIE Medical Imaging 2001*, volume 1, pages 249–256, 2001.

42. S. C. Mitchell, J. G. Bosch, B. P. F. Lelieveldt, R. J. van der Geest, J. H. C. Reiber, and M. Sonka. 3-D active appearance models: Segmentation of cardiac MR and ultrasound images. *IEEE Trans. on Medical Imaging*, 21(9):1167–1178, 2002.

43. S. C. Mitchell, B. P. Lelieveldt, J. G. Bosch, R. van der Geest, J. H. Reiber, and M. Sonka. Segmentation of cardiac MR volume data using 3D active appearance models. In *Proc. of SPIE Medical Imaging 2002*, pages 433–443, 2002.

44. E. Nagel, U. Schneider, S. Schalla, T. Ibrahim, B. Schnackenburg, A. Bornstedt, C. Klein, H. B. Lehmkuhl, and E. Fleck. Magnetic resonance real-time imaging for the evaluation of left ventricular function. *Journal of Cardiovascular Magnetic Resonance*, 2(1):7–14, 2000.

45. C. R. Oost, B. P. F. Lelieveldt, M. Üzümcü, H. Lamb, J. H. C. Reiber, and M. Sonka. Multi-view active appearance models: Application to X-ray LV angiography and cardiac MRI. In *Information Processing in Medical Imaging*, volume 2732 of *LNCS*. Springer, 2003.

46. E. Oost, B. P. F. Lelieveldt, G. Koning, M. Sonka, and J. H. C. Reiber. Left ventricle contour detection in X-ray angiograms using multi-view active appearance models. *SPIE Medical Imaging 2003*, 5032 I:394–404, 2003.

47. A. Pednekar, I. A. Kakadiaris, R. Muthupillai, and S. Flamm. Validation of the automatic computation of the ejection fraction from cine-MRI. In *MICCAI, 6th Int. Conference, Montréal, Canada*, LNCS 2879, pages 987–989. Springer, 2003.

48. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. R. Flannery. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, second edition, 1992.

49. H. Si. Tetgen – A 3D Delaunay tetrahedral mesh generator. Technical Report 4, Weierstrass Institute for Applied Analysis and Stochastics, WIAS, Berlin, 2002.

50. M. Sonka, B. P. F. Lelieveldt, S. C. Mitchell, J. G. Bosch, R. J. van der Geest, and J. H. C. Reiber. Active appearance motion model segmentation. *Proceedings Second International Workshop on Digital and Computational Video*, pages 64–68, 2001.

51. M. B. Stegmann. Active appearance models: Theory, extensions and cases. Master's thesis, Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, 2000. http://www.imm.dtu.dk/∼aam/.

52. M. B. Stegmann. *Generative Interpretation of Medical Images.* PhD thesis, Informatics and Mathematical Modelling, Technical University of Denmark, 2004.

53. M. B. Stegmann, B. K. Ersbøll, and R. Larsen. FAME – a flexible appearance modelling environment. *IEEE Trans. on Medical Imaging*, 22(10):1319–1331, 2003.

54. M. B. Stegmann, S. Forchhammer, and T. F. Cootes. Wavelet enhanced appearance modelling. In *International Symposium on Medical Imaging 2004, San Diego CA, SPIE*, pages 1823–1832. SPIE, 2004.

55. J. M. F. ten Berge. Orthogonal procrustes rotation for two or more matrices. *Psychometrika*, 42:267–276, 1977.

56. N. A. Thacker, A. Jackson, D. Moriarty, and E. Vokurka. Improved quality of re-sliced MR images using re-normalized sinc interpolation. *Journal of Magnetic Resonance Imaging*, 10(4):582–588, 1999.

57. M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *Proc. 1991 IEEE Com. Soc. Conf. on CVPR*, pages 586–91. IEEE Com. Soc. Press, 1991.

58. S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(4):376–380, 1991.

59. C.-F. Westin, J. Richolt, V. Moharir, and R. Kikinis. Affine adaptive filtering of CT data. *Medical Image Analysis*, 4(2):161–177, 2000.

60. C. B. H. Wolstenholme and C. J. Taylor. Wavelet compression of active appearance models. In *Medical Image Computing and Computer-Assisted Intervention, MICCAI*, pages 544–554, 1999.