Automated analysis of 3D echocardiography

Stralen, M. van

6

Automated left ventricular volume estimation in 3D echocardiography using active appearance models

ASSESSMENT OF LEFT VENTRICULAR (LV) functional parameters, such as LV volume, ejection fraction and stroke volume, from real-time 3D echocardiography (3DE) is labor intensive and subjective, because in current analyses it requires input from the user. Automating these procedures will save valuable time in the analysis and will remove interobserver variability.

We investigated a fully automatic segmentation approach for the left ventricle in real-time 3D echocardiography, based on active appearance models (AAMs). AAMs were built with end-diastolic images from 54 patients. We evaluated generalization capabilities of the shape and texture model and matching performance of the AAM using regular and Jacobian tuning matching algorithms in various scenarios.

The generalization of the shape model was good, comparable to a model containing 97% of the total modeled variation. The generalization of the texture model was moderate, comparable to a model containing 70% of the variation, which may hamper the AAM matching. In the comparison of the regular and Jacobian tuning matching methods, the latter obtained larger capture ranges and a higher accuracy.

The matching results indicate that fully automatic segmentation of the LV in 3DE using AAMs is feasible. Jacobian tuning matching has shown great potential for segmentation in echocardiograms and will improve the assessment of LV functional parameters.

This chapter is partially based on:
Automatic segmentation of the left ventricle in 3D echocardiography using active appearance models. M. van Stralen, K.Y.E. Leung, M.M. Voormolen, N. de Jong, A.F.W. van der Steen, J.H.C. Reiber, J.G. Bosch. Proc IEEE Int Ultrason Symp 2007; 1480-1483 (© 2007 IEEE) and **Improving 3D active appearance model segmentation of the left ventricle with Jacobian tuning.** K.Y.E. Leung, M. van Stralen, M.M. Voormolen, N. de Jong, A.F.W. van der Steen, J.H.C. Reiber, J.G. Bosch. Proc SPIE Med Imaging 2008; 6914; 69143B (© 2008 SPIE)

6.1 | Introduction

6.1.1 | Goal

Assessment of left ventricular (LV) functional parameters, such as LV volume, ejection fraction and stroke volume, from real-time 3D echocardiography (3DE) is labor intensive and subjective, because in current analyses it requires input from the user. Automating these procedures will save valuable time in the analysis and will remove interobserver variability.

Previous chapters of this thesis have focused on semi-automated approaches for automated 3D segmentation of the left ventricle, and preprocessing steps. The current chapter is devoted to our achievements towards the realization of a fully automated 3D segmentation approach based on so-called active appearance models (AAMs). AAMs hold considerable promise for the difficult task of segmentation in 3D ultrasound, and many research groups have addressed the topic, but the issue has not been solved so far.

6.1.2 | Related work

Previously, various techniques for the automated analysis of the left ventricle have been presented [Angelini et al. 2005; Corsi et al. 2002; Gérard et al. 2002; Hansegård et al. 2007*a*; Kühl et al. 2004; Zagrodsky et al. 2005; Zhu et al. 2007], nevertheless, most of these [Angelini et al. 2005; Corsi et al. 2002; Gérard et al. 2002; Kühl et al. 2004; Zagrodsky et al. 2005] still require manual interaction in the form of some indicated landmarks or manually drawn contours to achieve a proper analysis. Zagrodsky et al. [2005] have presented a fully automatic segmentation approach for the LV in 3DE, which is initialized by registration of a presegmented template with the unseen image. This approach is time-consuming and requires a presegmented template that can be successfully matched to any image to initialize the segmentation method. Furthermore, the evaluation was done on a limited number of subjects and showed significant problems when the ventricle was not fully captured in the imaging volume. Hansegård, Orderud et al. [Hansegård et al. 2007*b*; Orderud et al. 2007] combine an active shape model with a Kalman filter and show promising results. Zhu et al. [2007] attempt to detect the endocardial and the very challenging epicardial border using a maximum-a-posteriori framework which incorporates a statistical speckle model and an incompressibility constraint for the myocardium. Nillesen et al. apply adaptive filtering using image statistics as a preprocessing step to the automated segmentation[Nillesen et al. 2007; Nillesen et al.

2008]. The method lacks a constrictive shape model which hampers the detection.

Extensive research on analysis of time varying 2D echocardographic images has been carried out by Comaniciu et al. [2004] and Jacob et al. [2002]. They integrate temporal, textural information with an adaptive shape model in a Kalman filter approach (see also section 6.5.5.1).

other modalities

Fully automated 3D LV segmentation approaches in other modalities like CT [Zheng et al. 2007] and MR [Kaus et al. 2004] do not translate well to ultrasound. Mostly, these rely strongly on the generic intensity difference of tissue and blood and only apply weak shape continuity constraints on the deformable model (balloons, level sets etc). Due to the significant artifacts in ultrasound (in particular drop outs and clutter), anisotropy and position and orientation dependent intensity characteristics, ultrasound approaches require stronger shape constraints and more localized modeling of appearance.

Proposed approach: active appearance models | 6.1.3

Most segmentation and analysis approaches take a data-driven or bottom-up approach: they derive features from image data and try to fit a model (geometric, patterns, etc.) to these features. In our case, we propose to follow an opposite approach, labeled analysis-by-synthesis.

analysis-by-synthesis

We want to analyze a complex object that has a well-defined topology, but can exhibit a wide range of natural variability in shape and intensity patterns. Suppose we have a way of synthetically generating realistic images and can cover the variability with a limited number of parameter settings; then we can solve the analysis problem by finding the parameters that generate the best-fitting image. An example of such an approach is that of active appearance models.

We aim at fully automatic segmentation of the left ventricle in 3DE using active appearance models (AAMs). AAMs were first introduced by Cootes and Taylor [2001*b*], as extension of active shape models (ASMs). This approach has proven to be successful in various image segmentation tasks, starting with face recognition and later on in medical image segmentation. Bosch et al. [2000] have been exploring the application of AAMs to 2D echocardiography (2DE) and introduced the active appearance motion models (AAMMs) for analysis of 2DE time sequences [Bosch et al. 2002]. Since then, AAMs have been adopted to many medical image segmentation tasks and now have taken an important place in medical image analysis research, see section 6.2.2.

6.2 | Active appearance models

6.2.1 | Basic formulation

AAMs represent the shape and the texture of a certain (part of an) organ (in our case, the left ventricle in 3D ultrasound) as a mean appearance with its eigenvariations, by applying principal component analysis (PCA) on example training data annotated by experts. A short, general description of the AAM framework is given below. A complete description can be found in Cootes and Taylor [2001*a*].

6.2.1.1 | Model generation

training samples

We describe the training samples $i \in \{0...N-1\}$ by their shape $s_i = \{x_0, y_0, z_0, \ldots, x_{S-1}, y_{S-1}, z_{S-1}\}$ containing S corresponding surface points $\{x_\cdot, y_\cdot, z_\cdot\}$ and their texture $t_i = \{g_0, \ldots, g_{T-1}\}$ containing T corresponding image samples $g_\cdot$. By applying PCA on these shape and texture vectors we can describe the shape and texture by their mean shape $\bar{s}$ and texture $\bar{t}$, eigenvector matrix Φ_s and Φ_t and the parameter vector b_s and b_t. This requires a definition of (anatomical) point correspondence between the shapes, aligning them to the same pose and size, and calculating the average shape. This shape alignment step is usually solved through Procrustes alignment [Goodall 1991; Gower 1975]. After that, we can perform PCA on the shapes. By warping all shapes to the average shape, we get voxelwise correspondence over the neighborhood of the shapes, and we can calculate an average texture (voxel set) and apply a PCA on texture as well. This gives us a compact description of both shape and texture:

$$s = \bar{s} + \Phi_s b_s \tag{6.1}$$

$$t = \bar{t} + \Phi_t b_t \tag{6.2}$$

We can combine the shape and texture to model possible correlation between typical shape and texture variations by applying a third PCA on the combined parameter vector

$$b = \begin{pmatrix} W_s b_s \\ b_t \end{pmatrix} \tag{6.3}$$

where W_s corrects for the difference in units between shape and texture, to model the appearance,

$$b = \Phi_c c \tag{6.4}$$

The parameter vector c is often extended with parameters describing the pose in 2D or 3D. Depending on the application and the variability that is allowed in the shape model, most commonly uniform scaling, rotation and translation constitute these pose parameters. The appearance and pose parameters μ are combined in the parameter vector $p^T = (c^T|\mu^T)$.

Model matching | 6.2.1.2

One of the main advantages of an active appearance model over other segmentation strategies is its ability to quickly find a good match to the unseen image. An update strategy is used, which only needs a multiplication of the difference image, the difference between the synthesized and the underlying unseen image, with a precomputed parameter update matrix. AAMs are matched iteratively to unseen data by evaluating the difference between the modeled texture t_m and the corresponding texture in the sample image t_s, the residual vector

$$r(p) = t_s - t_m \tag{6.5}$$

and minimizing $E(p) = r^T r$. Minimization of $E(p)$ is achieved iteratively by trying to minimize $E(p+\delta p)$. The first order Taylor expansion of eqn. 6.5 is

$$r(p+\delta p) = r(p) + \frac{\partial r}{\partial p}\delta p \tag{6.6}$$

where $\frac{\partial r}{\partial p}$ is the Jacobian J. By differentiating $E(p+\delta p)$ to p and equating it to zero, we obtain the RMS solution,

$$\delta p = -Ur(p), \text{ where } U = \left(\frac{\partial r}{\partial p}^T \frac{\partial r}{\partial p}\right)^{-1} \frac{\partial r}{\partial p}^T \tag{6.7}$$

For the derivation of eqn. 6.7 we refer to Cootes and Taylor [2006]. An update of p is thus simply generated by a multiplication of the pseudo-inverse of the Jacobian $\frac{\partial r}{\partial p}$, matrix U. The Jacobian J (and thus also U) is assumed to be constant during the matching process and is estimated once and for all matchings in the model training phase.

Model training | 6.2.1.3

In the model training we learn the relation between each of the model parameters, and the difference image. This is done by perturbing each of the model parameters with predefined step sizes and learning the changes in the difference image that appear. This results in a matrix that describes the relation between model

parameters and the difference image (the Jacobian J). During the matching we use the inverse relation (update matrix U) to update the model parameters by simply multiplying the difference image with the update matrix, which speeds up the matching tremendously, compared to classical optimization approaches. We estimate the Jacobian by perturbing the model's parameters for every training sample t_l ($l = 1 \dots N$) as follows:

estimating the Jacobian

$$\frac{dr_i}{dp_j} = \frac{1}{N} \sum_l \sum_k w\left(\delta c_{jk}\right) \left(r_i\left(p + \delta c_{jk}\right) - r_i\left(p\right) \right) \tag{6.8}$$

where δc_{jk} is the k^{th} perturbation of parameter p_j and $\sum_k w\left(\delta c_{jk}\right) = 1$ for all j. The different perturbations of each parameter p_j are weighed using a weighting function $w(.)$, which is usually a suitably normalized Gaussian weighting function or a uniform weighting function, as in our case.

This estimation assumes that the Jacobian is more or less constant near the global optimum. Various studies have shown that this assumption can be made in varying applications. However, ideally one would want to compute the true Jacobian at each position. This would be a computationally very expensive operation. Lately, Cootes and Taylor [2006] proposed a method that iteratively updates the Jacobian during the matching. We will discuss this technique in section 6.3.2.4.

6.2.2 | Evolution of AAM (organs, modalities, dimensions + history)

The classical AAM as defined by Cootes and Taylor has been extended to a range of applications. Original applications by Cootes and Taylor concentrated on segmentation of 2D images of human faces, with some extensions to various medical imaging subjects, such as analysis of vertebral structure in X-ray images [Roberts et al. 2003; Roberts et al. 2007] and MRI of the brain [Cootes and Taylor 2001*b*].

Interesting work on the analysis of metacarpals in X-ray images was performed by Thodberg [2002], including handling of occlusions. Mitchell, Bosch et al. applied AAMs on cardiac MRI and ultrasound images, first on single 2D cross sections [Bosch et al. 2000; Mitchell et al. 2001*a*], later extended to time series (Active Appearance Motion Models or AAMM [Bosch et al. 2002; Mitchell et al. 2001*b*]). For ultrasound, this involved a nonlinear image intensity normalization to overcome the problem of the highly non-Gaussian gray value distribution in ultrasound. Later on, the first 3D implementation of AAM was realized and applied to end-diastolic 3D cardiac MRI datasets and pseudo-3D ultrasound datasets [Mitchell et al. 2002]. The ultrasound datasets were actually time sequences of 2D 4-chamber images stacked into a 3D block, and represented a cylindrical structure with limited 3D freedom. For 3D ultrasound, therefore, this was merely a proof of principle, not

a full realization of a 3D AAM. At the same time, a 2.5D implementation of AAM was realized by Beichel et al. [2002] for segmentation of the diaphragm in CT images. This application was not a full 3D implementation, since it modeled the object as a set of 2D points with the z-coordinate as an attribute, not as a truly 3D shape. Very interesting work has been realized by Stegmann et al.: application in 2D cardiac MRI [Stegmann et al. 2003], extension to a multi-view cluster-aware AAM on cardiac contrast perfusion MRI sequences ([Stegmann et al. 2005] and to a bi-temporal 3D AAM for automated estimation of the ejection fraction in 3D MRI sequences [Stegmann and Pedersen 2005]. Furthermore, 3D segmentation problems have been mostly approached using multi-plane AAM solutions, i.e. using multiple 2D cross sections, either uncoupled [Üzümcü et al. 2005] weakly coupled [Hansegård et al. 2007*a*] or coupled [Leung et al. 2006*b*; Oost et al. 2006; Stegmann et al. 2005]. Such approaches have the benefit of reduced computational load and complexity of modeling and matching. However, they may pose unrealistic constraints on shape change or 3D motion. Oost et al. [2006] relaxed these constraints by employing a dynamic programming detection step using the AAM segmentation.

AAMs in echocardiography

Inspired by the work of Bosch et al., Hansegård et al. [2007*a*] have shown that active appearance models can be applied with success in triplane echocardiograms. They applied multi-view and multi-frame active appearance models, and compared unconstrained AAMs with AAMs that were constrained by manually placed markers and by dynamic programming (DP). A DP-constrained AAM proved to work best in this setting.

In contrast to our approach, Hansegård et al. used a sparse (triplane) AAM, not a full 3D AAM. Furthermore, only a weak coupling in pose was used, to ensure that scale and vertical position did not deviate much over the three views. Hansegård et al. also employed the nonlinear gray value normalization described earlier by our group.

robustness of AAMs

Several modeling and matching methods have been proposed to generate more robust AAM segmentation results. For example, Gross et al. [2006] developed algorithms to apply AAMs to images of faces with occlusions, by combining their inverse compositional approach with a robust error function. An other robust approach for detecting object pose in stereo images consisted of selecting the appropriate multi-view appearance models and subsequent optimization of the robust error function with a modified Gauss-Newton algorithm [Mittrapiyanuruk et al. 2005]. Beichel et al. [2005] proposed a mean-shift-based method to estimate outlier residuals during the matching process. Their approach was applied to different types of medical images containing large artifacts. Recently, Cootes and Taylor [2006] proposed a new Jacobian tuning method, which allows the model's training matrix to adapt itself to new, unseen images during matching. The method is

supposedly more robust, is comparable with respect to speed with the standard matching method, and requires no extra steps in the model-training phase.

6.2.3 | Motivation for our work

AAMs have successfully been applied to a range of segmentation problems, including ultrasound, where they have been shown to offer significant advantages. However, a true 3D AAM for ultrasound has not been demonstrated yet. This could offer significant benefit over 2D or sparse approaches, and it could be extended in several ways, e.g. into a multi-phase 3D or truly 4D approach, a hybrid approach etc. This formed the motivation of our work, and several of these ideas have been investigated, albeit not always with a definite answer.

6.3 | Methods

6.3.1 | Data acquisition

In this work we used clinical data of 54 patients, acquired with two types of 3DE scanners. 18 of these patients were scanned using the fast rotating ultrasound (FRU) transducer [Voormolen et al. 2006], which was connected to a Vingmed Vivid 5 system (GE Vingmed, Horten, Norway).

FRU transducer

The FRU contains a linear phased array transducer that is continuously rotated around its image axis at high speed, up to 480 revolutions per minute (rpm), while acquiring 2D images. A typical data set is generated during 10 seconds at 360 rpm and 100 frames per second (fps). The images of the left ventricle are acquired with the transducer placed in apical position, and its rotation axis more or less aligned with the LV long axis. A single cardiac cycle in general is not sufficient for adequate coverage of the entire 4D space; therefore, multiple consecutive cycles are merged. The cardiac phase for each image is computed offline using detected R-peaks in the ECG [Engelse and Zeelenberg 1979].

preprocessing

The remaining 36 patients, who were referred for stress echo, were scanned using the Philips Sonos 7500 system (Philips Medical Systems, Andover, Massachusetts, USA), equipped with the X4 xMatrix transducer, placed in apical position. For the AAM modeling the FRU data sets are interpolated to Cartesian voxel sets, using a dedicated interpolation method for sparse irregularly sampled data (chapter 4).

Gaussian subsampling was applied to $1/4^{th}$ of the original resolution for both data types, to reduce speckle in the images.

The data was analyzed using a semi-automatic endocardial border detection method, allowing manual corrections (chapter 2). For the matrix acquisitions, the data was resliced to generate 10 equidistantly sampled long-axis views in all cardiac phases, which is a requirement for the analysis. Full cycle LV endocardial borders were extracted from these analyses and used as the training data sets for the AAM.

semi-automatic analysis

AAM for the left ventricle in 3DE | 6.3.2

Modeling | 6.3.2.1

Statistical analysis of the left ventricular shapes requires a corresponding point distribution for all the training samples. We define this point correspondence based on key landmark points from the semi-automatic analysis. The points are defined in an anatomical coordinate system of the left ventricle. This cylindrical coordinate system is oriented around the long axis (LAX). Near the apex the surface points are defined in a spherical coordinate system oriented around a center at $3/4^{th}$ of the LAX (fig. 6.1a). In the cylindrical part, the surface points are sampled equidistantly along the LAX and over the azimuth angle. For the apical part of the surface, sampling is done equidistantly over the elevation and azimuth angle. We chose to define the shape as such, to easily represent the endocardial surface with a regular sampling. This also avoids the need of performing a Delaunay triangulation on the mean shape. The neighboring samples are defined intrinsically in the sampling, which eases the triangulation. The mean shape, with triangulation, is shown in fig. 6.1b. Instead of using the regular definition of the apex, being the point on the surface which is most distant from the mitral valve center, we used a more stable definition. This stable apex is the intersection of endocardial surface with the axis through the center of gravity (CoG) of the upper quarter of the left ventricle (see fig. 6.1c).

shape model

In total, we typically sample at 30 levels from mitral valve to apex, at 30 angles for each level, together with a point for the apex resulting in 901 points for each 3D shape (fig. 6.1b).

We represent the translation, rotation and scaling of the model by 7 pose parameters: 3 for translation, 3 for rotation and 1 for uniform scaling. For the representation of 3D rotation we studied the use of Euler angles and quaternions [Funda and Paul 1988; Horn 1987]. We chose to represent the rotations using quaternions because of the unambiguous representation of 3D orientations, the orthogonality of the representation and the possibility to convert quaternions to rotation matri-

pose representation

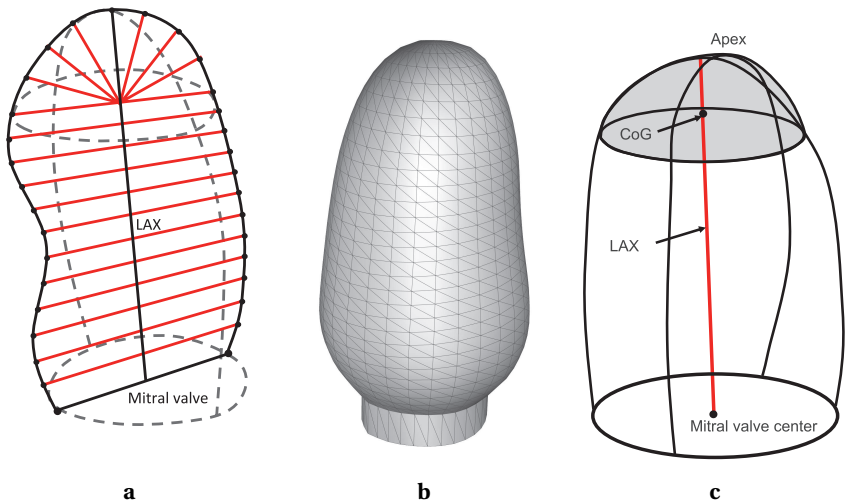

Figure 6.1: *a*) A schematic 2-dimensional representation of the regular cilindrical sampling and the spherical sampling in the apical region. *b*) The mean mesh, with triangulation, of a model containing 54 patients. The bottom ring structure is added for texture sampling. *c*) The redefinition of the long axis and the apex using the center of gravity of the apical region

ces back-and-forth without any loss of precision. To avoid any ambiguity in the quaternion representation $q = \{q_0, q_x, q_y, q_z\}$ and to be able to represent the 3D rotation with 3 quaternion parameters $\{q_x, q_y, q_z\}$, we defined that the first, omitted parameter q_0 is always positive (since $q = -q$). Since the norm of the full quaternion $\|q\|$ is 1, the omitted parameter can be recomputed at any time. Scaling is then represented by an independent parameter s.

Procrustes alignment

In the shape model we want to model only the biological shape variation of the left ventricle. Therefore the presegmented shapes are aligned using a 3D Procrustes alignment [Goodall 1991; Gower 1975]. In this way, the undesirable absolute position and orientation of the shape are removed from the model. These are a consequence of the acquisition procedure, not of any biological phenomena and are therefore not desirable in the shape model.

texture model

The texture sampling has been defined, similarly to the shape model, in the anatomical coordinate system of the left ventricle. We sample the texture radially on the line through the surface points, up to twice the radius of the surface. In this way, the myocardium, part of the right ventricle and a small region outside of the

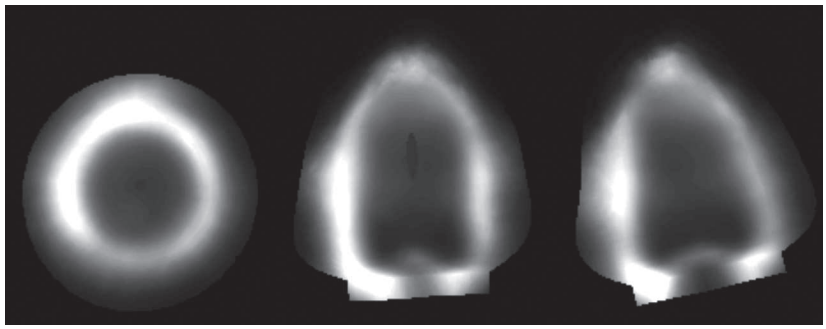

Figure 6.2: Three orthogonal intersections of the model mean. *From left to right*) Short-axis view, 2-chamber view and 4-chamber view (approximately)

heart is also modeled, to enlarge the lock-in region of the AAM.

Using this anatomical definition we can easily adjust the sampling density to be sparse in regions with little information (blood pool) and dense in important regions (near the endocardium). It also eases the warping of a texture to an arbitrary model shape, which speeds up the computation of the residual vector r. For the warping, we apply a trilinear interpolation defined in barycentric coordinates of tetrahedrons. The texture mean is shown in fig. 6.2. fig. 6.3 shows the three most prominent modes of appearance of the model built on 54 patients.

gray value normalization

Since we use a PCA in the modeling of the textures in the AAM, we assume that the texture samples are Gaussian distributed. It is known for ultrasound images that their gray value distribution is non-Gaussian. That is why we apply a non-linear gray value transformation that maps the mean histogram of all the training samples onto a Gaussian distributed histogram with zero mean and unit variance, as introduced by Bosch et al. [2002]. In this procedure, a combined normalized histogram of all the training samples is created. A histogram transformation is defined which transforms this combined histogram into a histogram with a Gaussian gray value distribution. This transformation is then applied to all individual training samples. Subsequently, the training samples are normalized to zero mean and unit variance using the regular gray value normalization used in AAMs.

Training | 6.3.2.2

pose and appearance parameters

In the training procedure we used perturbations δc of {2, 4, 6, 8, 10} mm for translation, and {0.02, 0.04, 0.06, 0.08, 0.1} for the scaling and rotation parameters. A uniform weighting function is used for the weighting of the different perturbations.

For training of the parameter update matrix we use perturbations of {0.2, 0.4,

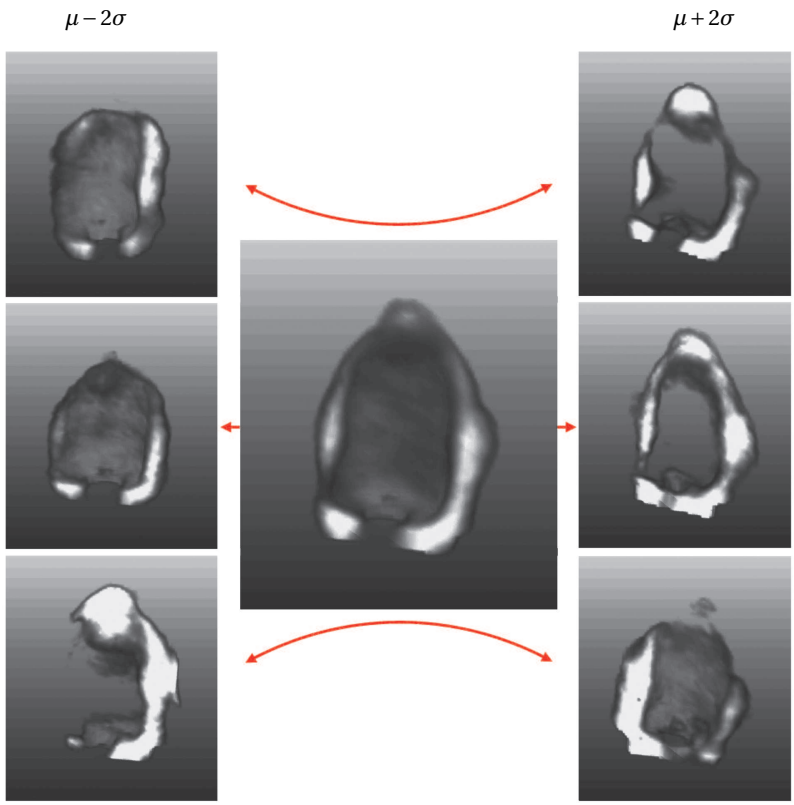

Figure 6.3: 3D renderings illustrating the three most prominent modes of variation from top to bottom with the model mean μ centered

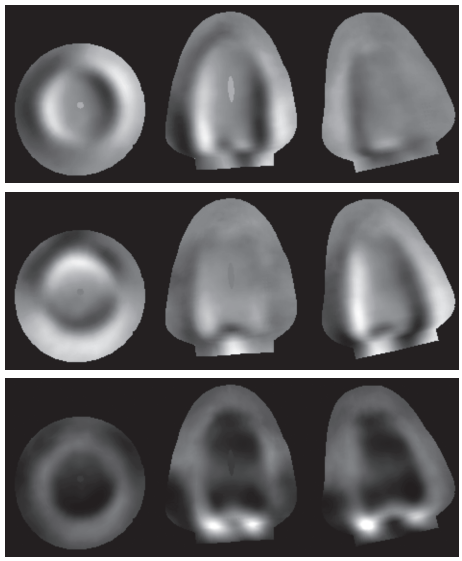

Figure 6.4: Three orthogonal intersections of the column images of the Jacobian matrix J for the three translation parameters. *top to bottom*) translation in x-, y-, and z-direction

0.6, 0.8, 1.0}σ for each of the appearance model parameters. These perturbations are uniformly weighted. The perturbation sizes were determined experimentally and correspond to values reported in literature [Cootes and Taylor 2001*b*; Stegmann et al. 2003].

Evaluations of the training procedure, as described in section 6.2.1.3, can be done by visualizing the columns of the Jacobian matrix, warped as textures to the mean shape. Fig. 6.4 shows the resulting images of the columns of the Jacobian matrix J corresponding to the translation parameters. The column images for translation show the high correspondence to the x-, y-, and z-derivatives of the average image, which is as expected.

Regular matching | 6.3.2.3

For the standard AAM matching the parameter update is generated by multiplication of the current residual $r(p)$ with the update matrix U, as in eqn. 6.7. We employ a linear search along the update vector δp using update steps k to minimize the residual $r(p + k\delta p)$. This linear search can be replaced by more sophisticated

variants if desired [Cootes and Kittipanya-ngam 2002]. Consequently, the parameter vector p is updated according to the optimal update step k:

$$p_i = p_{i-1} + k\delta p \tag{6.9}$$

These two steps are repeated until either the model converges or no improvement in r is found (alg. 6.1).

Algorithm 6.1 Regular AAM matching

p_0, $r_0 = r(p_0)$, $i = 0$
repeat
  $dp_i = -Ur_i$
  for all $k_j \in k$ **do**
    $p_{i+1}(k_j) = p_i + k_j dp_i$
    $r_{i+1}(k_j) = r(p_{i+1}(k_j))$
    $dr(k_j) = |r_i|^2 - |r_{i+1}(k_j)|^2$
  end for
  select $p_{i+1}(k_j)$ and $r_{i+1}(k_j)$ for largest $dr(k_j)$, if $dr(k_j) \geq 0$, else break
  $i = i + 1$
until $i \geq i_{\max}$

6.3.2.4 | Matching with Jacobian tuning

Cootes and Taylor observed that the assumption that the Jacobian is fixed is unsatisfactory, especially if the image to be segmented is significantly different from the model mean [Cootes and Taylor 2006]. Recently, they have proposed a search strategy that updates the Jacobian matrix during each new evaluation of the residual $r(p)$. The algorithm is closely related to the quasi-Newton methods for solving least square problems without derivatives [Broyden 1965]. The essence of the idea is that during the matching process, we apply variations to the parameters and obtain differences in the residuals. The change in residuals as a result of changes in the parameters provides similar information as is obtained in the regression training process, but in this case the information is highly specific for the case under consideration. Therefore, we would like to tune the standard Jacobian to the current case using the information obtained during the search.

patient specific matching

In short, the method uses a set of constraints on the parameter update at the current iteration i, given all previous parameter estimates $(p_0, \ldots, p_i)$ and previous residuals $(r_0, \ldots, r_i)$. The Jacobian J_0 from the training phase provides a regularization term for estimating current updates for the Jacobian matrix J. The updated Jacobian matrix J_i is then used to update the appearance parameters. No addi-

tional line search step is required. A summary of the algorithm is given below; for the original derivation, we refer to Cootes and Taylor [2006].

Consider a set of i observations of parameter differences $dp_k = p_k - p_{k-1}$ and residual differences $dr_k = r(p_k) - r(p_{k-1})$, organized in matrices $X = (dp_1|\ldots|dp_i)$ and $R = (dr_1|\ldots|dr_i)$. We set up i linear constraints on each row j_m of J, assuming that a linear update in the parameters generates a linear change in residuals: $X^T j_m = q_m$, where q_m^T is the m^{th} row of R. Using our trained Jacobian J_0 as a regularizer, we can set up a quadratic function of the form $f(j_m) = \alpha|X^T j_m - q_m|^2 + |j_m - j_{0m}|^2$, where α controls the strength of the regularization and j_{0m} is the m^{th} row of J_0. Differentiating f with respect to j_m and equating to zero leads to an equation for computing a new estimate of J, given the initial estimate from the training set J_0 and all previous parameter updates and residuals:

matching observations

$$(I + \alpha XX^T)J^T = J_0^T + \alpha XR^T \tag{6.10}$$

where I denotes the identity matrix. eqn. 6.10 can be rewritten into a more efficient version, which is then solved iteratively. Let us define three matrices $A = I + \alpha XX^T$, $B = J_0^T + \alpha XR^T$, and $C = B^T B$. By substituting eqn. 6.10 into eqn. 6.7, it can be shown that the optimal parameter update is given by $dp = Ay$, if y is the solution to the linear equation $Cy = -B^T r$. Instead of calculating A, B, and C using their definitions at every iteration, one can show that these matrices can be updated linearly at the current iteration $i+1$ using their values at the previous iteration i:

$$A_{i+1} = A_i + \alpha_i\, dp_i dp_i^T \tag{6.11}$$

$$B_{i+1} = B_i + \alpha_i\, dr_i dp_i^T \tag{6.12}$$

$$C_{i+1} = C_i + \alpha_i\, B_i^T\, dr_i dp_i^T + \alpha_i\, dp_i dr_i^T B_i + \alpha_i^2 |dr_i|^2 dp_i dp_i^T \tag{6.13}$$

This leads to the Jacobian tuning algorithm for AAM matching, alg. 6.2.

The resulting algorithm has only a series of simple linear operations, and can therefore be added straightforwardly to any existing AAM implementation. Note that the matrices A, B, and C are updated every iteration, regardless of the parameter update. It is usually possible to solve the linear equation in alg. 6.2, l.3 using Cholesky decomposition, as C_i is symmetric and (usually) positive definite. As in Cootes and Taylor [2006], we use $\alpha_i = (\delta + |dp_i|^2)^{-1}$, where δ is small, included to avoid numerical instability after small steps.

Algorithm 6.2 Jacobian tuning AAM matching

Initialize p_0, $r_0 = r(p_0)$, $i = 0$, $A_0 = I$, $B_0 = J_0$, $C_0 = B_0^T B_0$
repeat
  Solve $C_i y = -B_i^T r_i$ for y
  $dp_i = A_i y$
  $p_{i+1} = p_i + dp_i$
  $r_{i+1} = r(p_{i+1})$
  $dr_i = r_{i+1} - r_i$
  $z = B_i^T dr_i$
  $A_{i+1} = A_i + \alpha_i dp_i dp_i^T$
  $B_{i+1} = B_i + \alpha_i dr_i dp_i^T$
  $C_{i+1} = C_i + \alpha_i z dp_i^T + \alpha_i dp_i z^T + \alpha_i^2 |dr_i|^2 dp_i dp_i^T$
  if $|r_{i+1}|^2 > |r_i|^2$ **then**
    $p_{i+1} = p_i$
    $r_{i+1} = r_i$
  end if
  $i = i + 1$
until $|dp_i|^2 < \epsilon$ or $i \geq i_{\max}$

6.4 | Experiments and results

6.4.1 | Model generalization

coupling shape and texture

In the proposed AAM a correlation is assumed between the shape and texture of the training samples. That is why the shape and texture model are coupled by an extra PCA on these model parameters. By coupling these models, we might benefit from the correlation between shape and texture, but we also lose some of the variation for the shape and texture that is in the model, since shape and texture are not independent anymore and therefore can only be described together. We investigated the loss of generality by coupling these models.

model truncation

Secondly, since the number of training data sets available for training the AAM is limited, we investigated the degree of generalization that is achieved with these data sets. Therefore, we compare the error that is found when projecting a patient's shape or texture on the model in a leave-one-out (L-1-O) evaluation, with the error that is found when projecting the patients on a model with less variation in the shape or texture model, but with the current patient included. This is done by truncating the eigenvector matrix Φ and parameter vector b, such that the re-

maining eigenvectors account for the desired amount of variation. The number of eigenmodes t is determined by choosing the t largest eigenvalues $\lambda_0 \ldots \lambda_{t-1}$ such that

$$\sum_{i=0}^{t-1} \lambda_i \geq f \sum_{i=0}^{N-1} \lambda_i \tag{6.14}$$

where f is the proportion of the total variation that should be retained and N the total number of eigenvalues. In this way, we can estimate the amount of variation in the L-1-O models, by comparison of the projection errors of the L-1-O models with those of the truncated models. Furthermore, the projection error in L-1-O defines a lower limit on the final segmentation error, when matching to unseen data in a L-1-O evaluation.

Coupled vs. uncoupled models | 6.4.1.1

We evaluated the model generalization of the coupled model versus the independent shape and texture models. The resulting mean projection errors over all L-1-O projections are shown in fig. 6.5. Uncoupling the shape and texture models clearly decreases the projection errors. This is illustrated for the shape model in fig. 6.5. The mean point-to-point (P2P) projection error for the shapes decreases from 1.2 mm for the coupled model to 0.7 mm for the shape model built on 28 patients. A similar decrease in projection can be expected for the texture model. In these experiments we coupled shape and texture models using a weighting factor W_s (eqn. 6.3) that compensates for the difference in units of shape and texture vectors. This weighting factor may also be used to prioritize between shape and texture in the model.

weighting

Comparison with truncated models | 6.4.1.2

We also evaluated the model generalization for the shape and texture models individually and uncoupled, compared to truncated models from all training patients.

The resulting shape projection errors are shown in fig. 6.6. It shows that the shape model generalizes well. For the L-1-0 case with a model containing 53 patients, the model can describe any shape with a mean point-to-point (P2P) distance of 0.41 mm. This corresponds to a projection error of a shape model that contains approximately 97% of the modeled variance (fig. 6.6).

shape

For the generalization of the uncoupled texture models, we compared texture models of raw and normalized textures. For the texture model we express the projection error in mean squared intensity distance (MSD) to the original textures. The texture intensities are normalized to a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. Fig. 6.7 shows that the texture model generalization is

texture

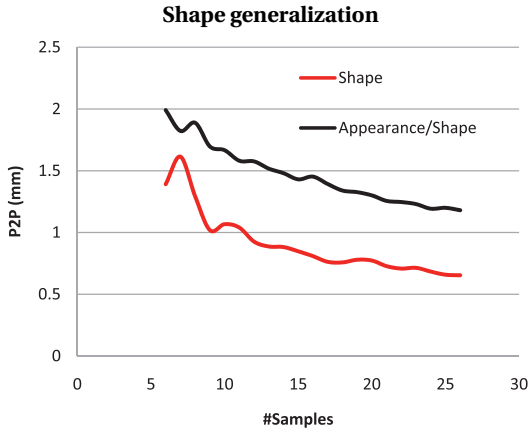

Figure 6.5: Superior generalization levels of uncoupled shape model, w.r.t. the coupled model. Mean projection errors over all the patients on L-1-O models. Point-to-point errors for projection of the training shapes on the shape model of the coupled (black) and uncoupled AAM (red) in L-1-O, for models built on an increasing number of training data sets (x-axis).

weak compared to the shape model, but that the normalized texture model generalizes to a higher level than the raw texture model, comparable to almost 70% and 60% truncated model respectively. Absolute comparison of the projection error is misleading, since the raw texture intensities are also scaled to a distribution with $\mu = 0$ and $\sigma = 1$, while the raw intensities are clearly non-Gaussian and therefore incomparable to the normalized intensities of the normalized texture model.

6.4.2 | Matching evaluation

matching scenarios

Since the Jacobian tuning method allowed the training matrix to adapt to the test image, we hypothesized that the method will have a larger capture range. Therefore, we tested the convergence of both methods: the model was initialized at its ideal pose and appearance in the test image, the appearance and pose parameters were then perturbed randomly in a range of several standard deviations, and subsequently the standard AAM and Jacobian tuning method were applied to match the model to the image. The experiments were first performed using a model describing 100% of the shape and texture variation (scenario A). Next, a model was used

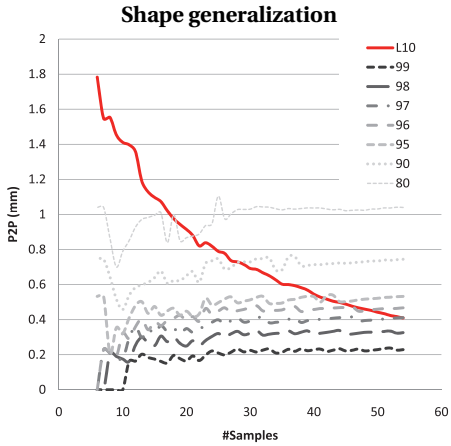

Figure 6.6: Generalization levels of the shape model as a function of the number of training sets. Mean P2P shape projection errors over all the patients on L-1-O models (redline) vs. the errors obtained by projection on truncated models containing 99%, 98%, 97%, 96%, 95%, 90% and 80% of the total shape variation (dashed lines)

which described only 95% of the shape and 75% of the texture variation (scenario B); in the previous experiments this was shown to be an accurate representation of a leave-one-out situation. Models A and B were built and matched on the same training data. A third scenario (C) was considered, in which models were created in leave-5-out fashion, such that five datasets were reserved for matching and the rest was used for training. This resulted in 11 models (with the last model made by leaving out the remaining four patients).

For all the scenarios we also initialized the models at their mean translation and appearance parameters, to evaluate the matching results when no information about the patient is used, to better approximate a real-life matching situation. The optimal parameters were then found using both matching methods.

For the standard algorithm, update steps $k = [1, 1\frac{1}{2}, 2, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}]$ were used. Matching was terminated if $dr(k_j) < 0$ for all steps k_j. As for the Jacobian tuning algorithm, the matching was allowed to continue until $|dp|^2$ was smaller than $\epsilon = 0.01$. For both methods, the matching was stopped if the mean squares of the residual was smaller than 0.001 (in MSD), or if the maximum of 100 iterations was reached.

stop criteria

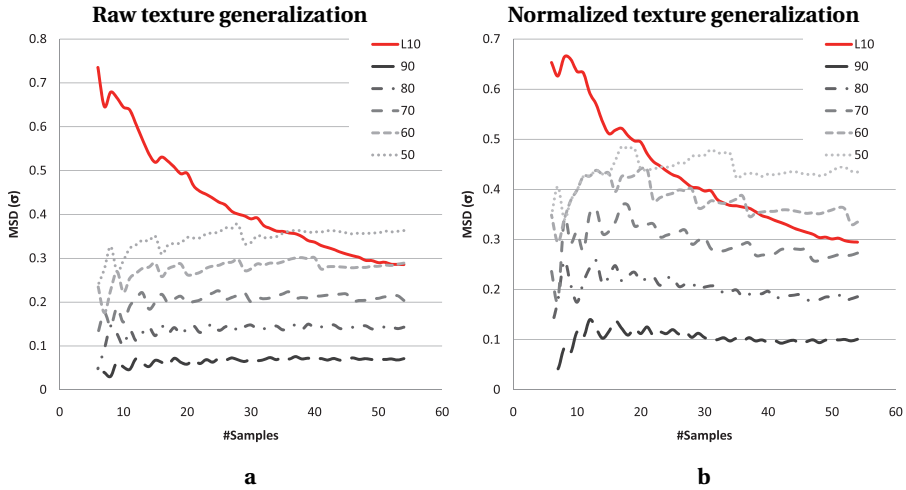

Figure 6.7: Generalization levels of the raw and normalized texture models as a function of the number of training sets. Mean texture projection errors (in MSD) over all the patients on L-1-O models (redline) vs. the errors obtained by projection on truncated models containing 90%, 80%, 70%, 60% and 50% of the total texture variation (dashed lines) *a*) For texture models built on the raw textures. *b*) For texture models built on the normalized textures.
Note that MSD units between raw and normalized texture models cannot be directly compared (see section 6.4.1.2)

6.4.2.1 | Perturbation from ideal parameters

scenario A

Point-to-point errors between the matching results and the manually drawn contours were calculated. With a model describing 100% of the shape and texture variation (scenario A), very low matching errors could be expected. For this experiment, a matching was considered converged if the point-to-point (P2P) error, averaged over the contour, was lower than 1 mm (the largest voxel size). The results revealed that the Jacobian tuning algorithm was superior to the standard algorithm (see fig. 6.8). In this case, 14.2% (69 out of 54*9 = 486) did not converge using the standard matching algorithm, whereas the Jacobian tuning algorithm achieved a 100% convergence rate. Most outliers occurred because the standard algorithm was not able to find an update for all steps k_j during the first iteration, such that the residual was lower than the residual at initialization.

scenario B & C

As for the truncated model (scenario B) and the leave-5-out models (scenario C), a lower spread in errors and much higher accuracy was observed if using the

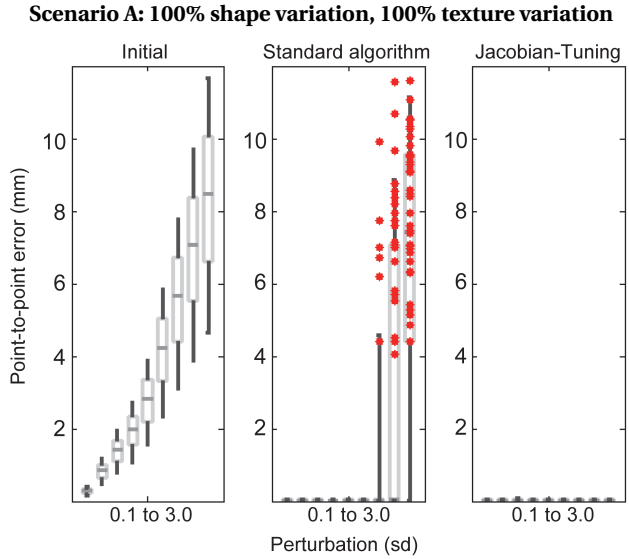

Figure 6.8: Superior matching convergence of the Jacobian tuning algorithm over the standard AAM matching. Perturbation experiments using a model with 100% shape and texture variation, initialized with perturbations of 0.1, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0, 2.5 and 3.0 σ. Boxes indicate 25% and 75% percentiles; whiskers extend to 10% and 90%; reddots indicate cases with a final matching error above 1 mm

Jacobian tuning algorithm, especially when large perturbations were applied (fig. 6.9). Similar results obtained for scenario C are shown in fig. 6.10.

The difference in the results of the Jacobian tuning matching between scenario A and B show the impact of the truncation of the shape and especially the texture model. Just a truncation of the shape model would presumably yield P2P errors that approach those of the projection experiments (fig. 6.6), close to 0.4 mm. However, in the matchings on the truncated model (scenario B) we found much larger errors, around 1.8 mm. This increase can be primarily attributed to the weak truncated texture model.

truncating the models

Initialization at mean parameters | 6.4.2.2

The matching results for initialization at mean translation and mean appearance parameters are given in table 6.1. Significantly lower errors P2P and P2S errors were obtained with the Jacobian tuning algorithm. A segmentation example is shown in

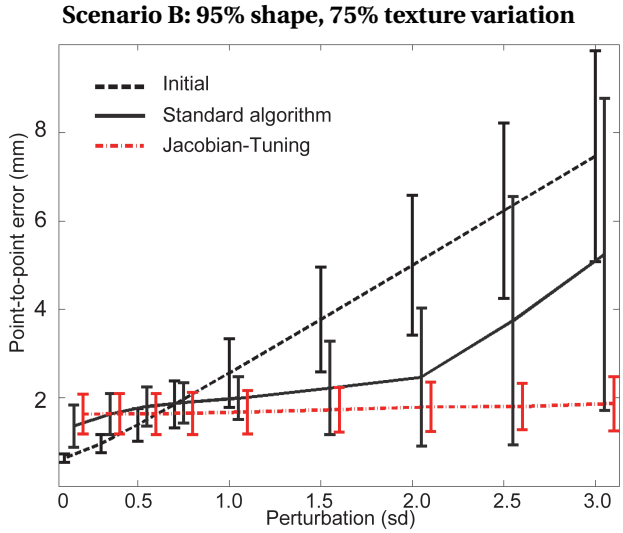

Figure 6.9: Superior matching performance of the Jacobian tuning algorithm (red) over standard AAM matching, compared to initial errors. Perturbation experiments using a model with 95% shape and 75% texture variation, initialized with perturbations of 0.1, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0, 2.5 and 3.0 σ (slightly pulled apart for clearer visualization)

fig. 6.11.

As expected from the previously described perturbation experiments, matching errors again increase from scenario A to B and C. This is due to the diminished level of generalization of the truncated and L-1-O models. Again, especially the weakness of the texture model seems to contribute most to the increased matching errors.

computation time

The computation time required for the AAM matching depends mostly on the number of (allowed) iterations. While there is still much room for optimizations, an iteration takes 1-2 seconds on regular desktop PC (2-3 GHz processor). Compared to the regular AAM matching, the Jacobian tuning takes twice as much time per iteration, but needs less iterations to come to convergence. Thus Jacobian tuning can have a similar computation time if compared to the regular matching.

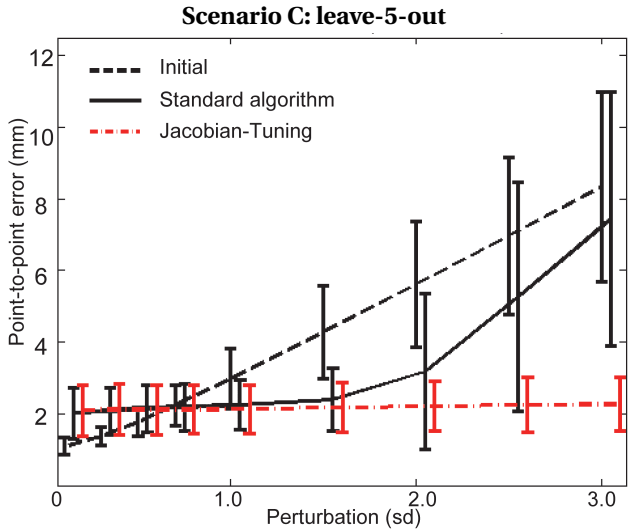

Figure 6.10: Superior matching performance of the Jacobian tuning algorithm (red) over standard AAM matching, compared to initial errors. Perturbation experiments using a leave-5-out model, initialized with perturbations of 0.1, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0, 2.5 and 3.0 σ (slightly pulled apart for clearer visualization)

Discussion | 6.5

General conclusions | 6.5.1

We have successfully developed a fully automatic method for segmentation of the left ventricle in 3D echocardiography based on active appearance models. The method has shown to provide good segmentation within a set of constraints. We have explored and compared variations and extensions of the classical AAM approach, and have charted the current boundaries of applicability in our problem domain. We have shown the generalization capabilities of the derived statistical models, discussed the importance of a proper choice of modeling, training and matching parameters, and showed the considerable added value of enhancements such as the Jacobian tuning matching approach. The required amount of computation time was acceptable for practical applications. Although a number of issues still needs to be tackled, AAMs constitute a very promising approach for automated

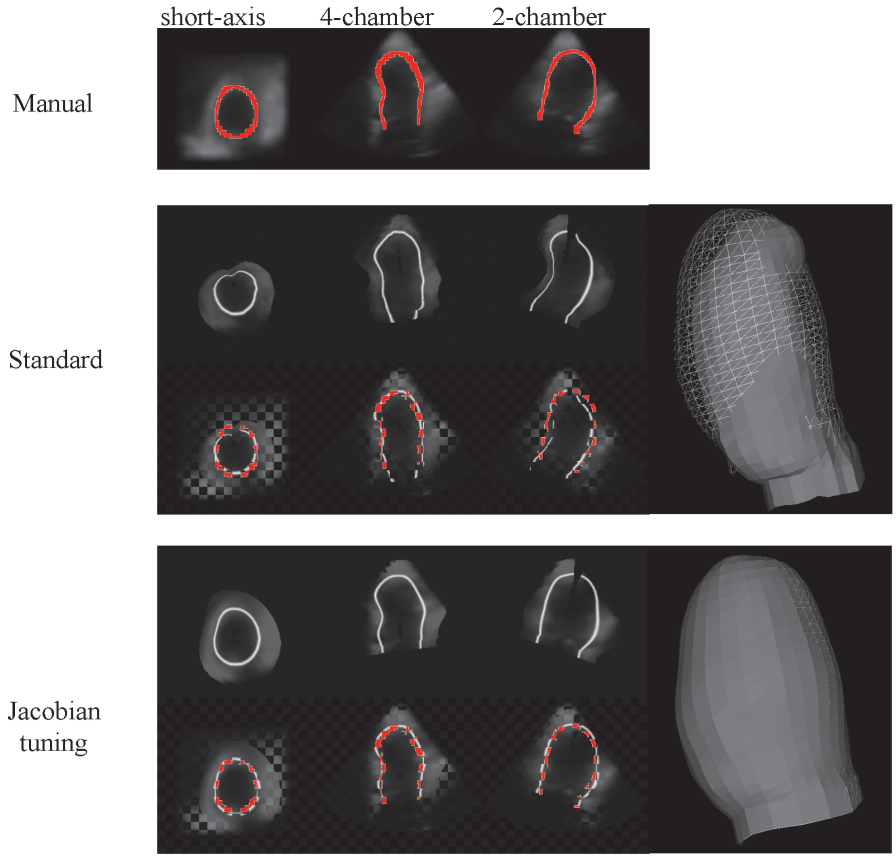

Figure 6.11: Appearance patches and 3D segmentations results using the regular AAM and the Jacobian tuning algorithms. For all segmentations a short-axis slice and two long-axis slices are shown. For the two segmentation approaches the detected model appearances are shown with contours and also a checkerboard image that combines the detected appearance with the original image and manual contour (redline). On the right the detected surface (solid) is shown together with the manual surface (mesh). In this particular case, the manual gold standard is very different from the mean appearance (fig. 6.2). The regular AAM matching has trouble finding the correct segmentation, as opposed to the Jacobian tuning method

Table 6.1: Mean ± standard deviation of P2P matching errors when initialized at mean translation and mean appearance parameters. Also shown are the root-mean-squares intensity error (RMS), in unnormalized intensity units (range [0 255]).
*indicates that Jacobian tuning has statistically significantly better results than standard AAM (paired t-test, $p < 0.05$, $N = 54$)

Measure	Match	Scenario		
		A	B	C
P2P (mm)	Initial	7.5 ± 2.6	7.5 ± 2.6	7.5 ± 2.7
	Regular AAM	2.3 ± 1.1	3.6 ± 1.7	4.4 ± 1.7
	Jacobian tuning	0.06 ± 0.03*	2.9 ± 2.0*	3.9 ± 2.0*
P2S (mm)	Initial	4.4 ± 1.2	4.4 ± 1.2	4.4 ± 1.2
	Regular AAM	2.0 ± 0.6	2.6 ± 0.5	3.0 ± 0.7
	Jacobian tuning	0.06 ± 0.03*	2.2 ± 0.7*	2.8 ± 1.0*
RMS intensity	Initial	5.2 ± 2.5	5.2 ± 2.5	5.2 ± 2.6
	Regular AAM	3.4 ± 1.7	3.6 ± 1.5	3.9 ± 1.7
	Jacobian tuning	2.0 ± 0.8*	3.5 ± 1.6*	3.8 ± 1.7

segmentation of the LV in 3DE that yearns for further substantiation.

We will discuss the advantages of matching using Jacobian tuning with respect to the classical matching approach, the current limitations, related work and conclude with recommendations for further research.

Regular matching vs. Jacobian tuning | 6.5.2

This study demonstrates, among others, the effectiveness of the new Jacobian tuning matching approach in AAM segmentation of the left ventricle in real-time 3D ultrasound images. We showed that the Jacobian tuning algorithm has a larger capture range and higher accuracy than the standard matching algorithm.

It is interesting to see that the outliers in fig. 6.8 are all located above approximately 4mm, suggesting that, below this threshold, it is possible to find the optimal appearance parameters using the standard algorithm. The Jacobian tuning method is much more robust because of its larger capture range, obtaining a 100% success rate for perturbations up to 3 standard deviations (σ) from the ideal parameters. *outliers*

Another interesting observation is the lower bound of 2 mm P2P error for the

truncated and leave-5-out model, which can be achieved for perturbations up to 3 σ using the Jacobian tuning method, whereas the standard algorithm starts to fail around 2 σ perturbation from ideal parameters (fig. 6.9 and 6.10). Of course, these error bounds and perturbation limits are dependent on the amount of variation captured in the model. However, it is clear that the Jacobian tuning method has a much larger capture range than the standard AAM algorithm. This may have significant consequences in matching models to images acquired with different machine settings and transducer equipment. For example, it would be worth experimenting with a model built with Philips data and matched to FRU data. This is a subject of further investigation.

capture range

6.5.3 | AAM matching strategies

Other AAM search algorithms have been reported in the literature, which implement updates to the Jacobian matrix. For example, Batur and Hayes [2005] proposed an algorithm which uses linear updates for the gradient matrix. Their approach is different to this one in the sense that the current parameters of the texture model are used to update the appearance parameters. This is combined with a line search similar to the one in section 6.3.2.3, and matching is stopped if no better residuals are found. This is different in our approach, where the Jacobian can be updated infinitely if desired. The Jacobian tuning method is closely related to the quasi-Newton method for solving least-squares problems without derivatives proposed by Broyden [1965]. More sophisticated approaches were proposed by Xu [1990]. These types of algorithms merit further research and comparison.

6.5.4 | Current bounds of applicability

6.5.4.1 | Rotation matching

No results for perturbations in rotation are reported in this study. In all our evaluations we have seen that adding the rotation parameters to the optimization process raises problems in the optimization. The capture ranges in perturbation experiments including rotation parameters are relatively small compared to capture ranges of other (pose) parameters and the matching accuracy degrades quickly outside these capture ranges for rotation parameters. These limitations could be attributed to several causes.

small capture range

Firstly, we can attribute these problems to the high degree of rotational symmetry of the left ventricle. Both the shape and the texture show only small differences

with respect to rotation around the LV long axis, which makes the objective metric relatively insensitive to these rotations. Therefore, also the training and matching of the AAM show poor performance in characterizing and solving these rotations respectively. *rotational symmetry*

Secondly, the detection of the rotation parameters is likely hampered by typical ultrasound specific image characteristics. Most acquisitions suffer from drop outs as a result of rib shadowing. Current 3D ultrasound equipment employs a footprint that is not small enough to image in between the ribs without resulting in shadows in parts of the image, especially the left ventricular lateral wall. Also, prominent image artifacts are common in the near-field, making it difficult to locate the LV apex precisely. Together, these typical image artifacts challenge the detection of the orientation of the LV. *typical image artifacts*

Furthermore, the myocardial texture may vary considerably throughout the left ventricle, due to misalignment of the acquisition axis with the left ventricular long axis. This is a typical ultrasound specific feature, caused by the angle of incidence of the ultrasound beam to the myocardium. This hampers the detection of the correct rotation with respect to the LAX since this typical change in texture does not correspond to any biological variation in shape. *angle of incidence*

Finally, the inherent nonlinearity of rotation representations may play an important role. From experimentation with Euler-angle representation vs. quaternion representations, we found that quaternions are better behaved with respect to numerical stability, but still the implementation choices might limit the lock-in range for rotations. The optimal representation of rotation angles can be a field for further investigations.

Nonlinear gray value normalization | 6.5.4.2

Another limitation of the current implementation of the AAM for 3DE is the nonlinear gray value normalization. AAM modeling expects a Gaussian distribution of variability, both for point distributions and textures. Ultrasound gray value distributions are known to be non-Gaussian. The nonlinear gray value normalization compensates for the non-Gaussian distribution of the gray values in ultrasound images. It is trained to transform the (linearly) normalized ultrasound histogram into a histogram that approximates a Gaussian distribution. Since this normalization is learned from the set of sample textures, it is limited to normalization of textures that have a comparable normalized histogram to the sample textures. Especially in cases where the AAM is far away from the optimal pose, the sampled texture histogram may be considerably different; e.g. if the scaling factor is too small, the texture contains mainly blood values. *learned normalization*

Alternatives to the gray value normalization scheme merit further research. Im- *preprocessing*

provements may be found in preprocessing steps that try to enhance the textures directly, by emphasizing desired features and masking out undesired ones, such as ultrasound speckle. This is discussed in section 6.5.5.4.

texture modeling

Other improvements may be found in the modeling of the textures. Larsen et al. [2007] describe the intensity information using wavelets and wedgelets, resulting in a more compact representation of the texture model. This might benefit both the distribution of the gray values and the generalization of the texture model. The generalization of the texture model may also be improved by using a cluster-aware AAM that employs clustering of the texture space to divide the non-Gaussian distributed texture model in multiple Gaussian submodels that describe the entire texture variation [Stegmann et al. 2005].

6.5.4.3 | Additional constraints

technical imaging limitations

Current 3D ultrasound imaging systems are not capable of truly real-time imaging of the whole heart. The speed of sound in human tissue is the limiting factor that determines the maximum number of beam that can be sequentially acquired per second. Therefore, images of multiple consecutive cardiac cycles are merged to obtain a high resolution image of the full left ventricle in 3D over the entire cardiac cycle. This multi-beat fusion often generates typical image artifacts, as a result of inter-beat variation, respiratory motion, patient's motion or transducer motion during the acquisition. These artifacts challenge the feasibility of locating the true endocardial border in these images, even by experts.

limited to model variation

AAMs are based on a statistical model of the shape and texture of a certain population of example images with their contours. Training and matching of the AAM relies on correspondence between model parameter variations and the change in the residual image. This relation is learned from the example data sets. Therefore, AAMs are limited to detection of contours which can be described by the statistical variation in the shape model, in images that can be reasonably approximated by the combined shape and texture model. Possible pathological cases, or biological subsets that are not represented in the training population will be detected with limited precision.

manual contours as gold standard

Furthermore, there may be imperfections in the manually segmented example data sets with which the models are built. These imperfections degrade the point correspondence of the shape model and also influence the quality of the texture model. Specifically, the alignment of the rotation with respect to the long axis is problematic. Often the right ventricle's attachment point and the aortic valve are hardly visible in the images; this hampers correct determination of this orientation.

Recommendations | 6.5.5

Temporal domain | 6.5.5.1

In our approach, we have focussed on segmentation of the left ventricle still 3D echocardiography, while much information about the myocardium is in the dynamics of the heart. Therefore, analysis of the temporal domain could greatly enhance also the detection in the single cardiac phases.

Extensions can be made, in parallel to single phase detection, to both the shape and the texture modeling of the AAM. Frangi et al. review various time-varying modeling techniques in cardiac applications [Frangi et al. 2001].

Even more information might be taken from temporal analysis of the texture. This may reveal differences between drop outs and shadowing on one hand, and the lower intensities of the blood pool on the other hand. This distinction is hardly extractable from single phase analyses. Static features in general, are not expected to provide much information about the functional parameters. Integration of temporal features with a dynamic shape model in echocardiography have been proposed by several groups. Comaniciu, Zhou et al. present a tracking framework with an adaptive shape model, where information about measurement uncertainty in the feature detection is combined with shape information using a Kalman filter approach. The framework has been shown to work well in 2D echocardiography [Comaniciu et al. 2004; Zhou et al. 2005]. A similar approach had been proposed by Jacob et al. [2002].

Hybrid matching | 6.5.5.2

The standard AAM matching, as proposed by Cootes and Taylor [2006] solves the update problem by using a fixed update matrix that is estimated at the optimal position. This type of matching assumes that the current model state is close to the optimum and is thus limited to a so-called lock-in range. Also, it only allows updates to the model parameters and is therefore bound to global model updates and to the variation in the model. This usually leads to a globally correct segmentation with minor mismatches. However, local mismatches may even cause the global segmentation to deteriorate.

Jacobian tuning has been shown to have a larger lock-in range and to result in a better fit of the model to the data. However, it is still limited to the variation in the model. Further improvement in local matching of the AAM may be found in a hybrid matching approach, like a combination with (multidimensional) dynamic programming.

6.5.5.3 | Multidimensional dynamic programming

The strength of the regular AAM matching in quickly finding a global solution also has a drawback. Finding a global solution may sometimes result in locally suboptimal segmentations. This may especially occur in pathological cases, which may not be covered by the samples in the model training. Therefore, regular AAM matching could be extended with multidimensional dynamic programming [Üzümcü et al. 2006] (MDP, an extension of dynamic programming (DP) [Amini et al. 1990; Bellmann 1965; Sonka et al. 1999]), to be able to fit the model's appearance more locally to these cases, without harming the global optimization. MDP can be used in this case to locally match the texture, synthesized by the regular matching, to the underlying unseen image. Since MDP is a path search strategy and we aim at detecting an optimal surface in 3D, MDP could be applied in collinear planes (slices) of the 3D volume, intersecting the long axis. This yields a set of 2D contours that constitute the 3D endocardial surface. From this new surface, a pose update can be computed and the shape can be projected onto the shape model to update the appearance model. The use of MDP (instead of regular dynamic programming) allows free movement of the surface points on the plane (or even out of plane) in the MDP update step. This could make the model more easily adapt to rotation errors for example.

We previously applied DP search in a semi-automatic segmentation approach for 3D echocardiography [van Stralen et al. 2005]. Oost et al. [2006] have successfully applied a hybrid approach to multi-phase X-ray angiograms, Hansegård et al. [2007*a*] extended this work to multi-view, multi-phase echocardiography. Therefore, an extension of the 3D AAM with (multidimensional) dynamic programming might be a promising subject for further research.

6.5.5.4 | Preprocessing

Echocardiographic images have a characteristic granular appearance, referred to as speckle [Wells and Halliwell 1981]. Speckle is often seen as undesirable in static (single frame) ultrasound image analysis and therefore we chose to simply reduce the speckle by Gaussian smoothing. More advanced preprocessing techniques have been proposed for speckle reduction. Among them are wavelet-based methods [Achim et al. 2001; Xiao et al. 2004] and anisotropic diffusion methods [Tauber et al. 2004; Yu and Acton 2002]. Nillesen et al. apply adaptive filtering using image statistics as a preprocessing step for automated image segmentation [Nillesen et al. 2007; Nillesen et al. 2008]. This adaptive filter enhances blood-to-tissue contrast and may therefore be used as a preprocessing step in the AAM modeling and matching.

Methods based on anisotropic diffusion may be easily incorporated in the re-

construction process of irregularly distributed ultrasound data, as proposed in chapter 4.

Cardiac phase | 6.5.5.5

A single phase segmentation method provides an important base for detection of left ventricular functional parameters, e.g. by combining it with a tracking approach. In such a framework, one is free to choose any cardiac phase for the initial segmentation.

We studied automated segmentation of the left ventricle in the end-diastolic (ED) phase. This cardiac phase allows clear visualization of the endocardial wall since it can be easily identified by the R-peak in the ECG and it does not suffer from much wall motion. A drawback of the use of the ED phase for LV segmentation is that it is the phase where the LV reaches its largest volume. Therefore, parts of the LV (especially the apex) might not be captured in the field of view or might be shadowed by the ribs or lungs. Also, in ED the endocardial wall is less clearly defined due to possible visibility of individual trabeculae, forming an irregular undulated surface [Mannaerts et al. 2003].

detection in end diastole

An obvious alternative to detection in ED would be single phase detection in ES. The advantages would be that the ventricle is more likely to be fully captured in the image volume and shadowing artifacts might be less prominent. Also, we found in chapter 5 that the acquisition axis is closer to the LV long axis, which eases the initialization. Also, the endocardial border definition might be more clearly defined [Mannaerts et al. 2003]. Difficulties of applying detection in ES might be proper detection of the ES phase from the ECG and the fact that LV geometry in ES may vary more widely than in ED, as a result of various pathological contraction patterns (e.g. the occurrence of 'kissing ventricles').

detection in end systole

Uncoupled AAM | 6.5.5.6

The generalization experiments in section 6.4.1 show that with the current size of the training set, the texture model only generalizes moderately. In cases where only a small training set is available, compared to the intrinsic dimensionality of the data, the detection can be improved by decoupling the shape and texture models. For example, Stegmann and Pedersen [2005] have used these uncoupled appearance models in segmentation of cardiac MR.

A drawback of decoupling the shape and texture models is that correlations between shape and texture modes are no longer modeled. Furthermore, the dimensionality of the model increases. This may hamper the matching process.

Theoretically, uncoupled models are able to describe unseen patients more precisely than coupled models. However, this difference can only be appreciated if the

matching proceeds to a high correspondence with the target image. Only then, the difference between the coupled and uncoupled model can be seen. As long as a good global match is not yet achieved (in pose, shape and texture), the uncoupled model will not be able to tune in great detail and match to the unseen data optimally.

6.5.5.7 | Improved/manual initialization

AAMs follow an iterative matching approach in search for a global optimum. This optimization is not guaranteed to find this global optimum. The quality of the initialization of the AAM influences the final segmentation result. Thus, the better the initialization, the better the final segmentation. There are several options to improve the initial model and to process the given initial data.

automated initialization

The most common way of initializing an AAM is by initialization on the mean position and orientation and the mean model parameters. In chapter 5 we presented a method to improve the initial transformation parameters automatically using a Hough transform and multidimensional dynamic programming. Zagrodsky et al. [2005] use a computationally costly atlas registration as initialization.

manual initialization

Most automated analyses techniques however, are initialized manually by a human observer. Initialization is done either by explicitly annotating the apex, a number of points on the endocardial border or the mitral valve [Corsi et al. 2002; van Stralen et al. 2005], or by indicating the LV position and dimensions by annotations [Angelini et al. 2001; Gérard et al. 2002; Kühl et al. 2004].

Cootes and Taylor [2001*a*] also describe a very elegant way to constrain the AAM, by including priors on the point positions. In this way, manual or automated feature points can be used to guide the AAM search. Hansegård et al. [2007*a*] followed this approach to combine the regular AAM optimization with dynamic programming on multi-plane echocardiography.

6.6 | Conclusions

We investigated the use of AAMs for automated segmentation of the left ventricle in 3D echocardiography. The shape and texture model were based on anatomical reference points, aiming at optimal sampling of the left ventricle image region.

We evaluated the degree of generalization of an appearance model based on 54 patients and found that the variation in shape can be sufficiently described by these 54 samples. L-1-O projection errors on such a model are comparable to projection

errors on a truncated model, containing 97% of the total shape variation. Because of the non-Gaussian distribution of ultrasound intensities, we compared generalization levels of raw and nonlinearly normalized texture models. The normalized texture model showed better generalization than the raw texture model. In contrast to the shape model, the generalization level of the texture models was low: projection errors are comparable to a those of a model that is truncated to 70% of the texture variation. This poses a challenge on the AAM segmentation. Either increasing the training set, or preprocessing the texture to get a more efficient representation of the texture variation should be employed to overcome this limitation.

We evaluated matching of the AAM in different scenarios, including variations in initialization, generalizations levels and matching strategies. Comparison of matching results using full and truncated models showed the impact of the weak generalization of the texture model. This is also reflected in the leave-one-out matching experiments.

The Jacobian tuning algorithm for AAM matching produces quantitatively better segmentations than the regular AAM matching, showing larger capture ranges and a higher accuracy. These experiments demonstrate the effectiveness of an adaptive update matrix during model matching. Given the large variability in ultrasound image appearance, the Jacobian-tuning algorithm has great potential in improving the standard AAM segmentation.

The main source of errors in matching the AAM are in finding correct rotation parameters. We excluded the evaluation of rotation matching from our analyses. Apart from this limitation, which can be circumvented by a proper initialization, 3D AAMs have shown to provide a valuable global segmentation of the left ventricle. Further research on combining the promising AAM approach using local matching strategies, temporal image analysis and improved initialization procedures is necessary to achieve a clinically acceptable automated segmentation approach.

Acknowledgments

The sonographers and clinicians of the Erasmus MC Thoraxcenter, in particular Boudewijn Krenning, Wim Vletter, Jackie McGhie, Attila Nemes, Osama Soliman, Marcel Geleijnse and Folkert ten Cate, are gratefully acknowledged for acquiring the images.