

Unsupervised Motion-compensation of Multi-slice Cardiac Perfusion MRI

M. B. Stegmann^{a,b,*}, H. Ólafsdóttir^a, H. B. W. Larsson^c

^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, H:S Hvidovre Hospital, Kettegaard Allé 30, DK-2650 Hvidovre, Denmark*

^c*Department of Diagnostic Imaging, St. Olavs Hospital, Trondheim University, Olav Kyrres Gate 17, N-7006 Trondheim, Norway*

Abstract

This paper presents a novel method for registration of single and multi-slice cardiac perfusion MRI. Utilising off-line computer intensive analyses of variance and clustering in an annotated training set, the presented method is capable of providing registration without any manual interaction in less than a second per frame. Changes in image intensity during the bolus passage are modelled by a slice-coupled active appearance model, which is augmented with a cluster analysis of the training set. Landmark correspondences are optimised using the MDL framework due to Davies et al. Image search is verified and stabilised using perfusion specific prior models of pose and shape estimated from training data. Qualitative and quantitative validation of the method is carried out using 2000 clinical quality, short-axis, perfusion MR slice images, acquired from ten freely breathing patients with acute myocardial infarction. Despite evident perfusion deficits and varying image quality in the limited training set, a leave-one-out cross validation of the method showed a mean point to curve distance of 1.25 ± 0.36 pixels for the left and right ventricle combined. We conclude that this learning-based method holds great promise for the automation of cardiac perfusion investigations, due to its accuracy, robustness and generalisation ability.

Key words: motion-compensation, registration, cardiac perfusion MRI, active appearance models, minimum description length

* Corresponding author

Email address: mbs@imm.dtu.dk (M. B. Stegmann).

URL: <http://www.imm.dtu.dk/~mbs/> (M. B. Stegmann).

1 Introduction

Within the last decade magnetic resonance imaging (MRI) has proved able to assess myocardial perfusion in an accurate and safe manner, see e.g. Muhling et al. (2003); Larsson et al. (1996). While scanning times have decreased substantially, the amount of manual post-processing renders the method prohibitive to clinical practice. Marking up points of correspondence on the myocardium constitutes a major part of this manual labour, which is essential to ensure compensation of motion during the perfusion sequence. This paper presents a novel approach aiming at replacing this resource-demanding, tedious and error prone task with an automated image analysis method, which provides a structured way of collecting and applying expert knowledge given by medical doctors into a learning-based framework.

Equal to most problems within medical image analysis, registration of perfusion MRI requires a very robust technique. This is partly due to the many factors during the MR image acquisition that deteriorate image quality such as resolution/time trade-offs, field inhomogeneity, susceptibility artefacts, flow artefacts, motion artefacts, partial volume effects, chemical shift artefacts, et cetera. Further, robustness is also required towards the severe biological inter-subject variation as well as variations in scan planning and scanner configuration. Finally, to be clinically viable, a registration technique should be able to deal with pathologies – which in this case mostly include perfusion deficits – without sacrificing accuracy or reliability.

In addition to the above requirements, we believe that methods – insofar possible – should be driven by the characteristics of the data, opposed to having explicitly encoded constraints regarding the properties of the domain problem, e.g. size, position, intensity, et cetera. Consequently, we have based our method on the Active Appearance Models (AAMs) framework by Edwards et al. (1998); Cootes et al. (1998). The philosophy behind AAMs is that computational analyses of annotated trainings sets should yield models capable of automatically annotating new, independent examples. In this paper we analyse the differences between the single-image oriented AAMs and time-series of multi-slice perfusion MRI and propose a framework designed to meet the stated requirements.

The clinical goal of this work is to automate the generation of perfusion maps¹ and ultimately; the actual perfusion measured in ml/(g·min), see e.g. Larsson et al. (1996). However, our method is currently validated using the registration accuracy, since this constitutes the core requirement for obtaining accurately

¹ These include contrast enhancement ratio, transit time, maximum upslope, upslope integral and time-to-peak, et cetera. See e.g. Nagel et al. (2000); Breeuwer et al. (2001b); Hsu et al. (2003); Bracoud et al. (2003).

motion compensated perfusion estimates.

The paper is organised as follows. A brief review of the existing body of work pertinent to image analysis of perfusion MRI is given in Section 2. Section 3 describes the data used for this study. Section 4 treats the methodological backgrounds and describes the proposed framework. Section 5 presents a qualitative and quantitative validation. Section 6 provides a discussion of the obtained results, the merits and demerits of the method, and a view to future work. Finally, Section 7 draws some concluding remarks.

This paper represents a direct extension of the work reported in Stegmann and Larsson (2003).

2 Related Work

In this section, the current status of registration of MR cardiac perfusion images will be considered. Refer to Table 1 for an overview of the methods stated below.

Note that all the following approaches register one slice sequence at a time independent of the remaining slices from the scan. Therefore computation times apply to a single slice sequence, unlike what is presented in this work.

Yang et al. (1997, 1998) use the phase difference of the raw MR data between successive image frames to correct for translational motion. Shape changes are compensated for by using a deformable model. They evaluate their method on 40 perfusion sequences from 20 patients. They compare their results to SPECT data from the same patients. Furthermore, they provide a qualitative evaluation by comparing signal-intensity (SI) curves before and after registration.

Behloul et al. (1997) simplify the problem of myocardial boundary detection by adding a slice from functional MRI to the perfusion sequence. They segment the functional slice by fuzzy clustering and fuzzy inference systems and warp the result to the perfusion sequence. The method requires manual definition of a rectangular region of interest (ROI) on each image of the perfusion sequence. The authors claim that their method can be applied in real-time but unfortunately it lacks validation.

Bidaut and Vallee (2001) use a multi resolution translation/rotation based registration minimising the mean squared differences (MSD) between each image frame and a reference frame. They evaluate their method using 8 sequences of 90 frames from 8 patients with stable coronary artery disease (CAD). The

validation is based on a comparison of image correlation factors in addition to motion reduction measurements for manual (ground truth) and automatic registration. Their computation time is 5 minutes per 100 time frames. Dornier et al. (2003) extend this work by defining cardiac masks with no signal intensity changes throughout the perfusion sequence. Computation time is reduced to 2.5 minutes per 100 time frames. Moreover, they improve the validation by calculating absolute myocardial perfusion and compare to values derived from manually registered sequences (ground truth).

Breeuwer et al. (2001a), Breeuwer et al. (2001b), Spreeuwers and Breeuwer (2001) also use a translation/rotation based registration but with normalised cross-correlation as a similarity measure. After the registration, an exact detection of endocardial and RV boundaries is obtained by region growing on feature images. The epicardial boundaries are detected by a deformable snake model. In Breeuwer et al. (2001a), Breeuwer et al. (2001b), the method is evaluated on 36 perfusion sequences from 12 patients by comparing the calculated perfusion parameters to estimated values from X-ray angiograms of the same patients. They obtain good correspondence for 11 of the patients. Their computation time is less than 4 minutes per slice sequence. Spreeuwers and Breeuwer (2001) evaluate qualitatively on 30 perfusion sequences from 14 scans. They succeed in 26 out of 30 sequences. Their computation time is 25 seconds per 70 time frames.

Gallippi and Gregg (2001) present a statistics based registration method using deformable template matching. The similarity measure depends on local brightness variations and edge directions. They evaluate their method on 12 patients and measure its quality by means of motion reduction and comparison of SI curves prior to and after registration.

Ablitt et al. (2002) use tissue tagging to correct for through-plane distortion on a multi-slice sequence. The in-plane motion in each slice is corrected for using free-form image registration with partial least squares regression (PLSR) deformation learning. They evaluate their method on a set of 8 patients with CAD and 5 healthy individuals in addition to a synthetic data set. The in-vivo data sets include 3 slices with 50 frames per slice. The accuracy of the method is measured relatively by comparing SI curves before and after registration. Gao et al. (2002) do further studies on the PLSR deformation modelling. The method is evaluated using 9 patients with CAD, each including 3 slices and 50 frames per slice, as well as a synthetic dataset. The accuracy of the method is measured by comparing motion reduction for a free-form image registration versus the proposed method.

Gupta et al. (2003), Bansal and Funka-Lea (2002) and Bracoud et al. (2003) use translation based registration methods. Gupta et al. with cross-correlation and the latter two with mutual information (MI) as a similarity measure. None

of the methods compensate for rotation or deformable shape changes. Gupta et al. validate on 10 patients, measuring motion reduction and comparing SI curves. The computation time of 50 frames sequence is 5 seconds but 30 out of 433 frames needed manual post processing. Bracoud et al. evaluate their method on 5 patients (3 slices, 30 frames per sequence) with stable CAD. They compare histograms before and after registration, inspect SI curves and perfusion maps and measure motion reduction. Their method takes 8 minutes for a 30 frame sequence including a manual definition of a rectangular ROI in one image of the sequence.

Spreeuwiers et al. (2002) present a refinement method to optimally place the myocardial borders after manually (or automatically) tracing them. They investigate the effect of displacements of the myocardium in perfusion MRI and use that information to correct the boundaries. They apply their method on 9 image sequences on which the myocardium has been traced, and compare perfusion parameters before and after correction.

3 Data Material

The data material comprises 2500 myocardial perfusion, short-axis, magnetic resonance images (MRI) obtained from ten freely breathing patients with acute myocardial infarction. Five slices of the first 50 sequential frame images each were selected before, during and after the bolus of contrast. The total sequence length acquired was approximately 70 frames. This study is performed using the 2000 images of the latter four slices of this set. The contrast agent was gadolinium diethylenetriaminopentaacetic acid (Gd-DTPA). Registration relative to the heart-cycle was obtained using ECG-triggered acquisition from a whole-body MR unit, Siemens Vision, operating at 1.5 Tesla. Slice acquisition was restarted at every third R-peak, thus providing an approximate frame time of three seconds, depending on the heart rate. MR pulse sequence: inversion recovery turbo-FLASH (fast low-angle shot), matrix size= 128×128 , field of view= 300×300 mm, slice thickness=10 mm, inter slice gap=0 mm, storage bit depth=16 bit. The inversion times for the four slices were TI=598 ms, 209 ms, 792 ms, 404 ms. See Figure 1 for an example multi-slice frame from the first patient showing four slices in anatomical order (in basal-apex direction from left to right). Notice the severe variation in contrast due to the inversion time difference. All slice images in the following sections will also be shown as a row of single images. The left digit in each slice shown in Figure 1 denotes the frame number. The right digit denotes the order in which each slice was acquired, see the inversion times above.

Table 1. Overview of registration methods for MR cardiac perfusion images.

Reference	Method	Similarity measure	Data	Validation method	Breathing	Computation time	Fully auto-matic
Yang et al. (1997)	Translation based registration, deformable model	image forces	20 patients (2 sequences, 100 frames)	SI curves, comparison to SPECT data	Yes	-	No
Behloul et al. (1997)	Segmentation of functional MRI mapped to perf. seq.	-	-	-	-	Real-time	No
Bidaut and Vallee (2001)	Translation/rotation based registration	MSD	8 patients (8 seq. 90 fr.)	Motion reduction, image correlation factors	-	5 min pr. 100 fr.	Yes
Dornier et al. (2003)	Translation/rotation based registration	Mean squared difference	8 patients (100 fr.)	Motion reduction, comparison of ground truth versus automatic SI-curves and quantitative perfusion parameters	-	2.5 min pr. 100 fr.	Yes
Breeuwer et al. (2001a), Breeuwer et al. (2001b)	Translation/rotation based registration, region growing, snake model	Normalised cross correlation	12 patients (36 seq., 60 fr.)	Quantitative comparison of perf. parameters	No	4 min pr. seq.	Yes
Spreeuwiers and Breeuwer (2001)	Translation/rotation based registration, region growing, snake model	Normalised cross correlation	14 patients (30 seq., 70 fr.)	Visual inspection	No	25 sec pr. 70 fr.	Yes
Gallippi and Gregg (2001)	Deformable template matching	Brightness, edge directions	12 patients	Motion reduction, SI curves	Yes	-	Yes
Bansal and Funka-Lea (2002)	Translation based template matching using edge and intensity information	Mutual information	-	-	-	-	Yes
Ablitt et al. (2002)	Free-form image registration with PLSR deformation model	Cross correlation	8 patients (24 seq., 50 fr.), 5 healthy (15 seq., 50 fr.), synthetic data	SI curves	Yes	-	Yes
Gao et al. (2002)	PLSR Deformation model	-	9 patients (27 seq., 50 fr.), synthetic data	Motion reduction	Yes	-	Yes
Gupta et al. (2003)	Translation based registration	Cross correlation	10 patients (46 slices, 433 frames in total)	Motion reduction, SI curves	No/yes ^a	5 sec pr. 50 fr.	No
Bracoud et al. (2003)	Translation based registration	Mutual Information	5 patients (15 seq., 30 fr.)	Histograms, SI curves, perf. maps, motion reduction	Yes	8 min pr. 30 fr.	No
Spreeuwiers et al. (2002)	Sensitivity analysis on perf. assessment	-	9 seq.	Comparison of perf. parameters	-	-	-

^a Breath-holding while comfortable, free-breathing after that.

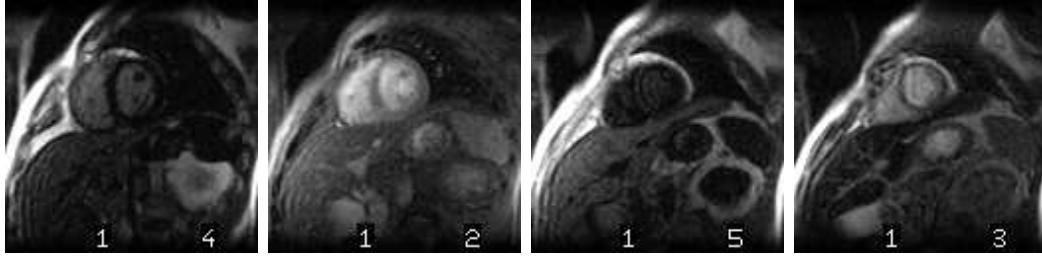


Fig. 1. Four slice images of patient 1 prior to bolus arrival (frame 1). Inversion time, $TI=598$ ms, 209 ms, 792 ms, 404 ms.

4 Methods

4.1 Myocardial Perfusion Imaging

Developments in MR-technology during the past decade have made it possible to acquire physiological information about dynamic processes in the human body. As an example of such, myocardial perfusion imaging encompasses assessment of myocardial perfusion at rest and during stress (e.g. pharmacological). By injecting a bolus of a paramagnetic contrast substance the myocardial perfusion mechanism can be quantified. Thus, perfusion MR qualifies as an essential instrument in the assessment of ischemic heart diseases. As the contrast agent tags the blood stream and amplifies the MR signal by a shortening of the T_1 relaxation time, areas of the myocardium served by diseased arteries show a delayed and attenuated response. Acquisition is carried out dynamically and registered to the heart cycle using ECG-triggering. Images are typically acquired from one or more short-axis slices every n -th heart-beat, where through-plane resolution is traded for temporal resolution and vice versa. If the acquisition time-window is sufficiently short (typically < 40 seconds), breath-hold can be used to remove respiration artefacts.

We further mention that imperfect ECG-triggering can give rise to depiction of erroneous heart-phases, in which tissue correspondences can be partially shattered, due to the long-axis movement of the left ventricle during the heart cycle.

In summary, the most prominent sources of variation in shape and appearance found in perfusion MRI – aside from general MR acquisition artefacts – include: biological inter-subject variation, scan planning, contrast passage, contrast uptake, respiration and inaccuracies in ECG-triggering. Additionally, when images are analysed on the basis of manual annotations, this mark-up process introduces inter- and intra-subject variability.

4.2 Active Appearance Models

Active Appearance Models (AAMs) Edwards et al. (1998); Cootes et al. (1998) were introduced as a learning-based method for registration and interpretation of face images. These models are built based on a set of order-less annotated images. By being a generic approach for image registration and image interpretation, medical applications were soon to follow. Refer to Stegmann et al. (2003) for a summary of medical AAM applications.

Formally, AAMs establish a compact parameterisation of object variability, as learned from a representative training set. The modelled object properties are usually shape and pixel intensities. The latter is henceforward denoted *texture*. From these quantities new images similar to the training set can be generated. Objects are defined by marking up each example with points of correspondence (i.e. landmarks) over the training set either manually, or by semi- to completely automated methods (see e.g. Davies et al. (2002b)). Using a learning-based optimisation strategy, AAMs can be rapidly fitted to unseen images, thus providing image registration and interpretation.

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, Gower (1975). Texture examples are warped into correspondence using a piece-wise affine warp, 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 $\bar{\mathbf{s}}$ and $\bar{\mathbf{t}}$ denote the corresponding sample means. New instances are now generated by adjusting the principal component scores, \mathbf{b}_s and \mathbf{b}_t in

$$\mathbf{s} = \bar{\mathbf{s}} + \Phi_s \mathbf{b}_s \quad , \quad \mathbf{t} = \bar{\mathbf{t}} + \Phi_t \mathbf{b}_t \quad (1)$$

where Φ_s and Φ_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

$$\mathbf{b} = \begin{bmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_t \end{bmatrix} = \begin{bmatrix} \mathbf{W}_s \Phi_s^T (\mathbf{s} - \bar{\mathbf{s}}) \\ \Phi_t^T (\mathbf{t} - \bar{\mathbf{t}}) \end{bmatrix}. \quad (2)$$

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 trans-

form,

$$\mathbf{b} = \Phi_c \mathbf{c} = \begin{bmatrix} \Phi_{c,s} \\ \Phi_{c,t} \end{bmatrix} \mathbf{c}, \quad (3)$$

obtaining the combined appearance model parameters, \mathbf{c} , that generate new object instances by

$$\mathbf{s} = \bar{\mathbf{s}} + \Phi_s \mathbf{W}_s^{-1} \Phi_{c,s} \mathbf{c} \quad , \quad \mathbf{t} = \bar{\mathbf{t}} + \Phi_t \Phi_{c,t} \mathbf{c}. \quad (4)$$

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 $\mathbf{p} = [t_x \ t_y \ s \ \theta]^\top$ where t_x , t_y and θ denotes in-plane translation and rotation, and s denotes the shape size.

Using a least-squares criterion the model is matched to an unseen image by an iterative updating scheme based on a fixed Jacobian estimate Cootes et al. (2001) or a principal component regression Cootes et al. (1998). For this work we have used the former approach, treated in further detail by Stegmann et al. (2003). In brief, 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{R} denote a fixed update matrix pre-calculated from the training set. Updates are then estimated by

$$[\delta \mathbf{c}^\top \ \delta \mathbf{p}^\top]^\top = \mathbf{R}(\mathbf{t}_{image} - \mathbf{t}). \quad (5)$$

For further details on AAMs refer to Cootes et al. (1998, 2001); Cootes and Taylor (2001).

In our application we mention that due to the inherent representation of texture vectors in AAMs, pose- and shape-compensated images are *directly* obtained by projecting each texture vector into the shape-free reference frame.

Thereby a per-pixel correspondence over the complete perfusion sequence can be obtained, ready to be analysed by a perfusion map, or by a perfusion model, see e.g. Muhling et al. (2003); Larsson et al. (1996).

4.3 Shape Annotation and Landmark Correspondences

One important step towards a robust shape model is to generate a training set of shapes with good landmark correspondence. A semi-automatic procedure involves two steps: i) Extraction of shape contours from the images. ii) Definition of marks at corresponding locations across the set of contours.

The first step is a manual procedure while the landmark correspondences

can be obtained automatically by Minimum Description Length (MDL) shape modelling introduced by Davies et al. (2001).

An annotation tool has been developed in order to achieve the first step, extraction of the shape contours. The tool enables the user to extract the desired object by marking points on the contour using a mouse or a digitiser. We use a *WACOM intuos²* tablet digitiser, which both gives faster and more accurate annotation than using a mouse. The cardiac shape includes two anatomical landmarks, the points where the RV border meets the epicardial, called anterior junction and inferior junction as shown in Figure 2.

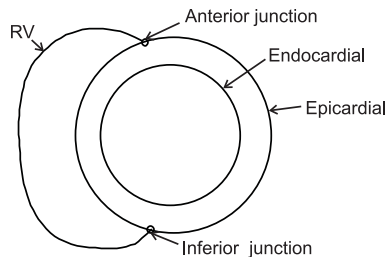


Fig. 2. Anatomical landmarks on the cardiac shape: Anterior junction and inferior junction between right and left ventricle.

The anatomical landmarks are included in the manual process, i.e. the user defines those while annotating. After placing points on the object, the contour itself is given by an interpolating cubic spline between the points.

After this manual step of the procedure, MDL optimisation gives a reparameterisation, representing an optimal point correspondence, of the given shape contours. The MDL approach is based on minimising the description length, a quantity adopted from information theory. The description length in this case is the cost of transmitting the PCA coded model parameters in addition to the transmission cost of the encoded data values. This means that MDL balances the complexity of the model against how well the model fits the data. For further details of MDL shape modelling see Davies (2002) and Davies et al. (2002a).

We have used the freely available MDL implementation from Thodberg (2003). Roughly, his approach is to slide points on one shape at a time according to an adaptive step length, align the set of shapes, do PCA of the set and calculate the description length. For each move, the description length is compared to the current best one. Based on comparison, this move is either accepted or rejected.

MDL shape modelling cannot be applied directly to the cardiac shape, since it is not possible to parameterise it. Furthermore, the fixed points (here, the anatomical landmarks) can give rise to problems since they are not end points. The approach used here was to split the shape into five open fixed-end sub-

contours. The epicardial contour is split at the anatomical landmarks and the endocardial contour is split at pseudo landmarks placed at minimum distance from the anatomical landmarks. This is clarified in Figure 3.

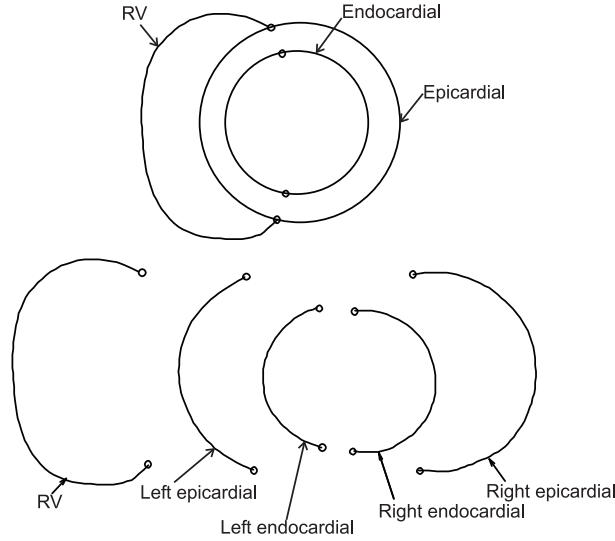


Fig. 3. The cardiac shape split into five open fixed-end sub-contours. The MDL approach is applied to each of the five sub-contour sets.

The MDL approach is applied to each of the five sub-contour sets. The aim is to optimise each sub-contour set locally, but with influence from the full shape. This is done in the alignment step where the full shape (all five sub-contour sets) is included. The PCA and the optimisation itself, i.e. the sliding of points, is performed on the local sub-contour set only.

With the large data set of 2000 shapes, the MDL optimisation process becomes very slow. By noticing that the most shape variation occurs between patients and between slices, the approach used here is to select the last frame of each slice sequence for the MDL optimisation. This reduced the set to 40 shapes (for each sub-contour type) which leads to an acceptable computation time. The output from MDL is a set of parameterisation functions; one for each of the shapes in each sub-contour sets. These parameterisation functions are applied to the corresponding sub-contours from the remaining frames of the sequence. Eventually, the five sub-contour sets are merged again to form the full training set of cardiac shapes.

4.4 Modelling of Perfusion MRI Time-series

Since perfusion MRI sequences differ in structure from the single-image oriented AAMs, this section will discuss the issues of data modelling.

First, one should recognise that the signal variation of the left ventricle (LV)

and the right ventricle (RV) is very small prior to contrast arrival. Due to this lack of image contrast, the standardisation of texture vectors normally used in AAMs would result in severe amplification of scanner noise. Hence, only the texture mean is removed in the following pre-processing of image texture vectors.

Deviations from the assumptions in standard AAMs are not only encountered with respect to texture normalisations. The relationship between shape and texture also differs. It is safe to assume that the process generating shape variation remains stationary throughout a sequence, contrary to the highly non-stationary texture process. Consequently, shape and texture are treated independently by removing the combined PCA, i.e. $\Phi_c = \mathbf{I}$. Treating each perfusion sequence as one observation (as done on cardiac cine MRI in Mitchell et al. (2001)) is not feasible due to the random fluctuations in pose and shape induced by the variation sources mentioned in Section 4.1; whereof respiration dominates.

However, joint modelling of texture – decoupled from changes in shape and pose – for a complete sequence is possible by introducing an annotation of the temporal relationship of each sequence frame. This may not be desirable, as only a few proper temporal landmarks exist (during the actual bolus passage). Further, building sequence models would require far more training examples to model this behaviour properly, compared to a simpler and less constrained frame-based model. Consequently, given the low number of subjects, we will treat each frame in a sequence as an observation.

Circumventing the need for large training sets unfortunately violates a basic assumption in AAMs. Namely, that the variation in texture is well modelled by a single multivariate Gaussian model. Due to the radical changes in intensity during contrast passage and uptake this is clearly not the case. On contrary, five major time intervals can be identified in a perfusion sequence; i) pre-contrast arrival, ii) contrast agent entering the RV, iii) contrast agent entering the LV, iv) contrast agent entering the myocardium, and v) post-contrast bolus.

Since the texture parameters of the texture model shown in Equation 1 are ranked by their variance over the training set, this alleged clustered behaviour should be confirmed by visual inspection of a few highly-ranked parameters from \mathbf{b}_t . Figure 4 shows the three most significant texture parameters, $\{b_{t,i}\}_1^3$, from a texture model built from the ten available perfusion sequences. This figure exhibits a pronounced clustering and thus verifies our concerns. Modelling this multimodal distribution with a multivariate Gaussian is clearly not justified and gives rise to several problems. Most problematic is that the resulting model is not specific to the given problem. Thus, it can easily generate textures that are not plausible to occur during a perfusion bolus passage.

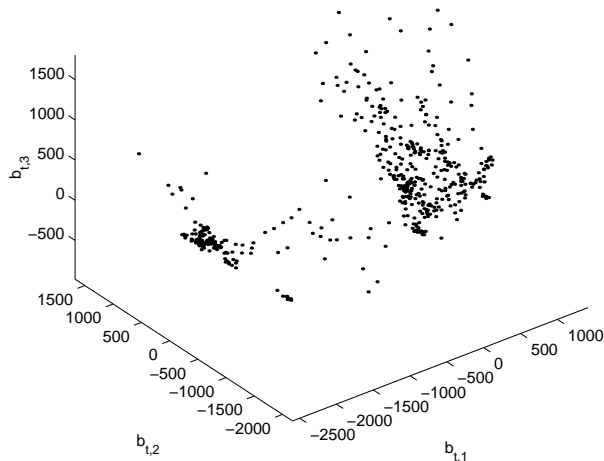


Fig. 4. First, second and third principal component of 500 texture vectors from ten perfusion sequences.

This would in turn lead to potential false positives during the model-to-image matching and will be dealt with in the following.

4.5 Adding Cluster Awareness

To model the distribution of textures we propose an unsupervised learning approach that models texture variation using an ensemble of linear subspaces in lieu of the unimodal linear model employed in AAMs. This ensemble is obtained by clustering and implies that the different time phases in perfusion sequences are not explicitly modelled. Instead the sole structure of the training data determines the partitioning. On contrary, subspaces could have been given by an operator, which identifies different phases of each perfusion sequence. However, to reduce, i) the tedious burden of training set generation, and ii) inter- and intra-observer variability, supervised learning was rejected. Further, we like to leave the option open to evaluate different ensemble sizes in the future when more training data becomes available. This will obviously be very tedious in the case of manual labelling.

Although the machine learning literature offers an abundance of clustering methods, it is generally agreed upon that no *silver bullet* exists. We have chosen a k -means clustering Forgey (1965) combined with a Monte Carlo simulation scheme where several clusterings are carried out, based on different initial random class centres. The final clustering is chosen using a minimax criterion, where the clustering having the smallest maximum distance to the nearest class centre is chosen.

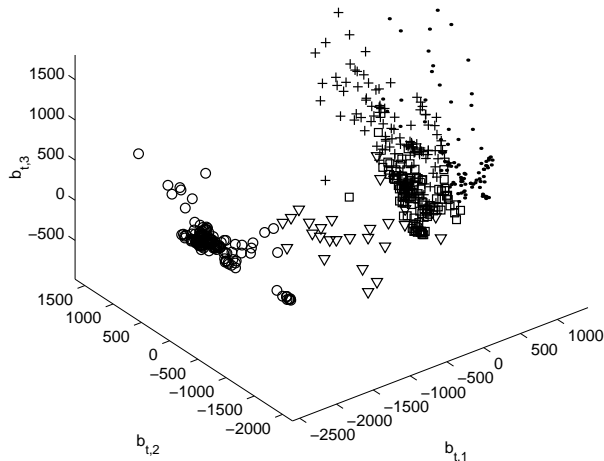


Fig. 5. First, second and third principal component of 500 texture vectors from ten perfusion sequences clustered into five classes.

The obtained clustering using five classes ($k = 5$) of the data set is shown in Figure 5. From this clustering, a set of linear texture subspaces, $\{\Phi_{t,i}\}_{i=1}^k$, is obtained directly by k separate texture PCAs. A corresponding set of model parameter update matrices, $\{\mathbf{R}_i\}_{i=1}^k$, is obtained following the procedure in Cootes et al. (2001) using a displacement scheme specified in Stegmann et al. (2003).

As texture changes in a sequence are deemed to be unrelated to shape changes, building a joint shape model, Φ_s , from all frames in all sequences will yield the best estimate of inter- and intra-subject shape variability. We denote this composition of a single shape model and an ensemble of texture models with associated parameter update matrices for a *Cluster-aware AAM* (CAAM). A CAAM can thus be applied to any domain problem that justifies uncoupled shape and texture eigenspaces and requires piecewise-linear modelling of texture variability.

Fitting a CAAM to unseen images now involves choosing the appropriate texture subspace. As a reasonable choice for k is typically fairly limited, model selection is here performed by exhaustively trying all models and selecting the model producing the best model-to-image fit, subject to a set of constraints given later in this paper. To increase performance during model fitting, model selection could be accomplished by a classification of the texture vector into the set of training classes. However, this has not been tried in the current work.

To choose k prior knowledge can be employed. However, being an optimisation problem in one positive integer variable, a data-driven method should be preferred where k is estimated using cross-validation on the training set. Unfortunately, one should also observe that the number of available training

examples for model estimation within a single class is inversely proportional with k . Due to this fact, we have fixed k to five classes (reflecting the five major intervals mentioned in the previous section) for all experiments as a reasonable compromise, since conclusions regarding the optimal number of classes would remain fragile, due to the limited number of training subjects available.

4.6 *Modelling Object Interfaces*

To add notion about the interface between the heart and the surrounding tissue et cetera, we include intensity samples in the proximity of the heart into the texture model.

This is carried out by sampling landmark normals relative to the current shape size, similar to the approach by Active Shape Models, Cootes et al. (1995). We denote these normals *whiskers*. In our application, this will include the LV/lung interface and LV-RV/abdomen interface into the texture models and thus require these to be present in the unseen image to provide a good model-to-image fit. We have weighted whisker samples so that they constitute one third of the texture variation in all of the subsequent experiments.

This approach has earlier been shown to have a positive, and significant, impact on the registration accuracy, in a case study on corpus callosum registration in brain MRI, Stegmann et al. (2004).

4.7 *Multi-slice Shape Modelling*

Up to this point an observation has been associated to a single image. However, cardiac perfusion MRI is typically acquired as multi-slice images in a spatial arrangement. This enables regional assessment of the perfusion mechanism both in plane as well as through plane. Modelling such slice images independently will per se disregard any spatial and/or intensity coherence between slices. Conversely, a joint modelling of each set of slices will provide a constrained basis, in which slices with well-defined appearance will restrict slices with diffuse appearance from diverging.

Coupled AAMs were first explored by Cootes et al. (2000) for interpretation of multi-view face images. Later Mitchell et al. (2001) coupled time-sequences of cardiac MRI, an approach which was also pursued in Lelieveldt et al. (2001); Sonka et al. (2001); Bosch et al. (2001a,b, 2002). Recently, long and short axis MRI, and end-diastolic and end-systolic angiograms were fused by a similar approach by Lelieveldt et al. (2003) and Oost et al. (2003), respectively.

How coupled AAMs should be constructed is application specific. For our application the following approach was taken. Let a shape consisting of N points in two dimensions be defined by a $2N$ vector,

$$\mathbf{s} = [x_1 \ y_1 \ x_2 \ y_2 \ \dots \ x_N \ y_N]^\top. \quad (6)$$

A multi-slice observation consisting of a stack of C slices is then composed of the set of shapes; $\{\mathbf{s}_i\}_1^C$ by simple concatenation of the unmodified slice image shape coordinates prior to normalisation et cetera,

$$\mathbf{s}_{frame} = \left[\mathbf{s}_1^\top \ \mathbf{s}_2^\top \ \dots \ \mathbf{s}_C^\top \right]^\top. \quad (7)$$

Any subsequent shape analysis is left unchanged. Image sampling is carried out using the appropriate slice for each related subpart of the combined shape vector. This produces a model where inter-slice pose relations are modelled by the shape changes, which is clearly desirable for our application. Should inter-slice differences in pose not be constrained; then slice concatenation should be applied to the Procrustes tangent space shape coordinates instead. This will thus introduce the need for $(C - 1) \times 4$ extra pose parameters into the model (in the 2D case).

4.8 Estimating and Enforcing Pose and Shape Priors

The fact that changes in pose and shape are unrelated to the change of texture is highly useful for initialising and constraining the model fitting process in each frame. Further, it can validate the final registration results as well as the intermediate results for each texture class. Consequently, if it is possible to obtain reliable estimates of the shape and pose variation in a subpart of the sequence these can be used as constraints in the remaining part of that sequence. This is the case in the latter part of a bolus passage where the contrast agent has been washed out of the RV and LV, only leaving the subtle changes stemming from the perfusion mechanism in the myocardium. Hence, we propose to estimate prior distributions of pose and shape from the latter part of a perfusion sequence of P frames.

Let κ , γ , D_{max} denote a set of user-selectable, dimensionless, constants controlling the influence of the priors. Then, let Σ denote the dispersion matrix of the sequence pose parameters and let σ denote the standard deviations of the sequence shape parameters. How these two quantities are estimated is treated in Appendix A. Further, let \mathbf{F}_t denote the t -th frame. Let the set of frames $\{\mathbf{F}_t\}_{t=1}^{S-1}$ denote the unstable period, and the frames $\{\mathbf{F}_t\}_{t=S}^P$ denote the stable period. A formal scheme for exploiting these priors can then be formulated as shown in Algorithm 1.

Algorithm 1 Sequence Prior Augmented CAAM Search

Require: $S, \kappa, \gamma, D_{max}, \Sigma$ and σ

- 1: $\bar{\mathbf{p}} =$ initialisation
 - 2: $\bar{\mathbf{b}}_s =$ initialisation
 - 3: **for** $t = P$ down to S **do**
 - 4: $\{\mathbf{p}_t, \mathbf{b}_{s,t}\} =$ CAAM search started at $\{\bar{\mathbf{p}}, \bar{\mathbf{b}}_s\}$ in \mathbf{F}_t
 - 5: $\bar{\mathbf{p}} = \frac{1}{P-t+1} \sum_{j=t}^P \mathbf{p}_j$
 - 6: $\bar{\mathbf{b}}_s = \frac{1}{P-t+1} \sum_{j=t}^P \mathbf{b}_{s,j}$
 - 7: **end for**
 - 8: **for** $t = S - 1$ down to 1 **do**
 - 9: $\{\mathbf{p}_t, \mathbf{b}_{s,t}\} =$ Constrained CAAM search started at $\{\bar{\mathbf{p}}, \bar{\mathbf{b}}_s\}$ in \mathbf{F}_t
(constrain using $\kappa, \gamma, D_{max}, \Sigma, \bar{\mathbf{b}}_s, \sigma$)
 - 10: **end for**
-

During CAAM search in the unstable period, pose and shape priors are used to stabilise parameter updates by limiting the maximal update step. To simplify notation, the time index t is omitted from this point on. Let Σ_{ij} denote the element in the i -th row and j -th column of Σ , and let p_i denote the i -th element of \mathbf{p} . Pose parameter updates can now be constrained using the following simple clamping approach:

$$\delta p_i = \begin{cases} \text{sign}(\delta p_i) \kappa \sqrt{\Sigma_{ii}} & \text{if } |\delta p_i| > \kappa \sqrt{\Sigma_{ii}} \\ \delta p_i & \text{otherwise.} \end{cases} \quad (8)$$

The constant κ acts thus as a clamping constant given in units of standard deviations of pose variation as estimated from the training sequences. As an alternative to this hyper cuboid constraint we could use a projection onto the hyper ellipsoid given by an eigenanalysis of Σ . We have used $\kappa = 0.5$ in all experiments as a reasonable compromise between damping and number of updates until convergence.

Likewise, we also exploit the prior knowledge of sequence shape variation, σ , (as obtained from the training set) in a similar clamping approach,

$$b_{s,i} = \begin{cases} \bar{b}_{s,i} + \text{sign}(b_{s,i} - \bar{b}_{s,i}) \gamma \sigma_i & \text{if } |b_{s,i} - \bar{b}_{s,i}| > \gamma \sigma_i \\ b_{s,i} & \text{otherwise,} \end{cases} \quad (9)$$

where γ denotes the maximally accepted distance from the mean in units of standard deviations. We have used $\gamma = 3$ in our experiments, which accounts for more than 99 percent of the probability mass of a univariate Gaussian distribution.

After ended CAAM search in the unstable period, pose prior enforcement is

determined by testing the Mahalanobis distance to the pose distribution:

$$\{\mathbf{p}, \mathbf{b}_s\} = \begin{cases} \{\bar{\mathbf{p}}, \bar{\mathbf{b}}_s\} & \text{if } D_{max}^2 < (\mathbf{p} - \bar{\mathbf{p}})^\top \Sigma^{-1} (\mathbf{p} - \bar{\mathbf{p}}) \\ \{\mathbf{p}, \mathbf{b}_s\} & \text{otherwise} \end{cases} \quad (10)$$

Again, under the assumption of normally distributed pose parameters, we have chosen to use $D_{max} = 3$. Hence, implausible pose solutions are discarded and replaced with the maximum likelihood of the prior; the mean configuration. Such solutions are flagged and the system operator can be prompted for manual assistance. This pose prior enforcement is also employed during texture model selection so that solutions failing the above condition are excluded from the competition of being the selected for the current frame. This step is essential, since highly deviating pose configurations easily can produce model-to-image fits that will outperform the true positives.

4.9 Model Initialisation

The initial pose, shape and texture parameter configuration of the CAAM, in frame P , is determined by exploiting the convergence radius of the AAM search. Within this radius the AAM will converge to a plausible fit. Parameters with variation outside this radius over the training set are semi-exhaustively searched, with grid spacing less than twice the radius, by starting the AAM search at each grid point. Each search result represents a potential initialisation candidate. To add robustness this is implemented in a candidate scheme, where several candidates compete in an evolutionary process to become the initial configuration. See Stegmann (2001) for the details of this algorithm.

Similar to the normal CAAM search, initialisation is performed exhaustively for all texture models and subsequently choosing the model providing the best fit.

5 Experimental Results

To evaluate the proposed method an implementation was made and subsequently validated using the data set described in Section 3. The implementation was done in C++ as an extension of our AAM framework (see Stegmann et al. (2003)) and executed on a 2.4 GHz Pentium PC. The stable period was set manually to the last 25 frames of each sequence, i.e. $S = 26$, except for patient 1 who had a delayed bolus passage, $S = 36$ was used. (This parameter can presumably easily be set automatically using the extremum of the temporal image derivative.) Due to the limited data set size, a leave-one-out cross

Table 2

Mean shape errors (shown in pixels).

	Mean	Std.dev.	Median	Min	Max
Pt.pt.	2.81	0.71	2.68	1.51	6.69
Pt.crv.	1.25	0.36	1.17	0.66	3.79

Table 3

Shape errors for each patient (shown in pixels).

Patient	1	2	3	4	5	6	7	8	9	10
Pt.pt.										
Mean	2.50	2.81	3.45	3.23	2.68	2.53	2.62	3.26	2.75	2.26
Std.dev.	0.46	0.71	0.44	0.83	0.41	0.73	0.63	0.39	0.90	0.35
Median	2.45	2.74	3.45	2.99	2.62	2.40	2.52	3.28	2.59	2.21
Min	1.51	1.76	2.53	2.21	1.90	1.52	1.73	2.49	1.74	1.73
Max	3.50	4.56	4.39	5.13	3.72	5.34	5.14	4.06	6.69	3.72
Pt.crv.										
Mean	1.29	1.07	1.62	1.31	1.03	1.18	1.35	1.21	1.35	1.10
Std.dev.	0.23	0.15	0.35	0.22	0.21	0.48	0.39	0.16	0.54	0.24
Median	1.27	1.03	1.68	1.26	0.97	1.05	1.27	1.18	1.21	1.06
Min	0.78	0.85	0.86	1.03	0.75	0.66	0.98	0.94	0.80	0.83
Max	1.82	1.50	2.20	1.83	1.69	2.98	3.26	1.67	3.79	2.21

validation was carried out. This resulted in ten test scenarios where a model was built on nine examples and tested on the independent example that was left out. Each model was automatically initialised in the P -th frame using the approach described above.

Two performance benchmarks were calculated for each model landmark: *pt.pt.*, which measures the Euclidean distance between corresponding landmarks of the model and the ground truth; and *pt.crv.*, which measures the shortest distance to the ground truth curve. These are referred to as *landmark errors* in the following. When these are grouped into mean values for a shape, then they are referred to as *shape errors*. The initialisation strategy successfully located the heart in all patients, but number nine. The CAAM for this patient was subsequently manually initialised and included on equal terms with the remaining nine.

Table 2 summarises the shape errors for all patients, while Table 3 expands on these numbers to show errors for each patient. Computation timings for the initialisation and the image search are given in Table 4. Summerising performance into a single measure, Table 2 shows a mean point to curve distance of 1.25 ± 0.36 pixels.

Shape errors versus frame number are depicted in Figure 6 and 7 as mean values with error bars and as box plots, respectively. Notice the consistent increase of the shape error during the unstable period. Plotting shape errors versus patient number in Figure 8 shows approximately 14 apparent search

Table 4
Timings (shown in seconds).

	Sequence time	Mean frame time
Initialisation	34.0	0.7
CAAM search	9.9	0.2
Total	43.9	0.9

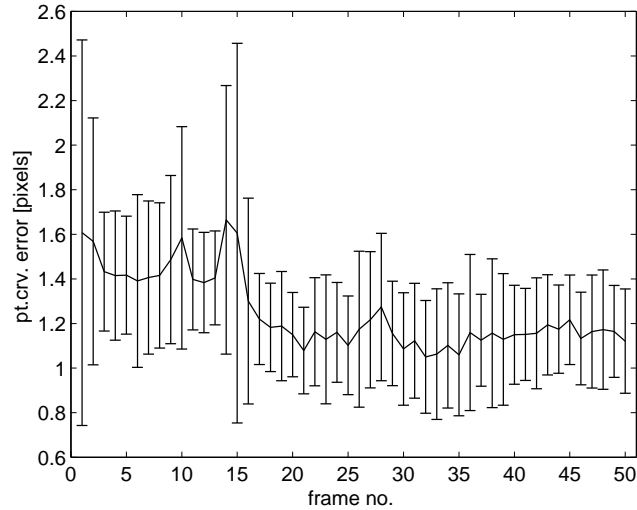


Fig. 6. Pt.crv. shape errors for each frame. Error bars are one std. dev.

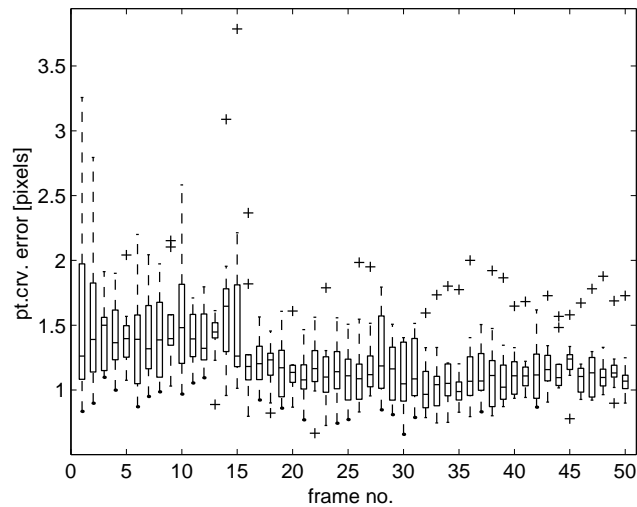


Fig. 7. Box plots of the pt.crv. shape errors for each frame (IQR=1.5).

failures out of 500 frames ($14/500=2.8\%$). From Figure 6 it is seen that all of these occurred during the unstable period. As indicated by circles in Figure 8, seven of these were detected and flagged by Equation 10. For patients 1–3 the replaced mean estimator was close to the ground truth shape, contrary to patient 9.

To give a qualitative impression of the accuracy Figure 9 and 10 show regis-

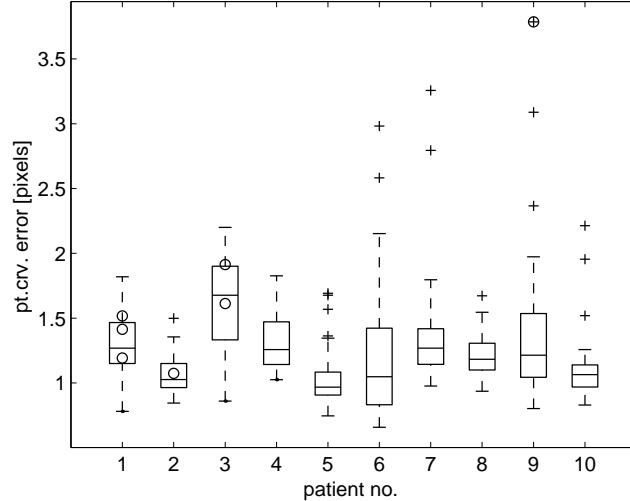


Fig. 8. Box plots of the pt.crv. shape errors for each patient (IQR=1.5). Pose prior enforcements are shown using circles.

tration results before, during and after the bolus passage for patient 10 and 8, respectively. Notice the dark region showing a severe perfusion deficit present in the anterior and inferior septal wall of patient 8 in Figure 10.

To complete the quantitative overview, Figure 11 shows the distribution of pt.crv. landmark errors in addition to their cumulative frequency. It is seen that approximately 80 percent of the landmark errors are below two pixels. Finally, Figure 12 illustrates the spatial distribution of the landmark errors by showing circles with radii proportional to the mean landmark error on the mean shape. Errors on RV dominate together with errors on the LV endocardial contour in proximity of the papillary muscles, i.e. near the anterior and inferior lateral wall, of the first and third slice.

6 Discussion

Registration of perfusion MRI remains a challenge to medical image analysis. However, we believe that the presented method holds promise for the future, although conclusions based on the limited data material will remain fragile. This optimism is partly due to the ability to produce useable results on pathological cases – all patients were diagnosed with acute myocardial infarction – from MRI scans exhibiting an image quality below average compared to recent developments in MR cardiac imaging.

Registration results were produced without manual interaction in nine out of ten cases, due to the described initialisation method. This strategy failed for patient 9, for whom the heart was poorly centred in the slice field-of-view.



Fig. 9. Registration results for patient 10 before, during and after bolus passage.

Further, and more importantly, slices were shifted heavily towards the apex compared to the remaining nine patients, which obstructed a good fit of the LV-RV model in the initialisation process. We consider this patient an outlier and accept the observed limitations.

In general, less than three percent model fit failures were observed judged from Figure 8. All occurred prior to, or during the bolus passage. Half of these were not detected by Equation 10. We consider this performance convincing, although it indicates that registrations based on this limited training data

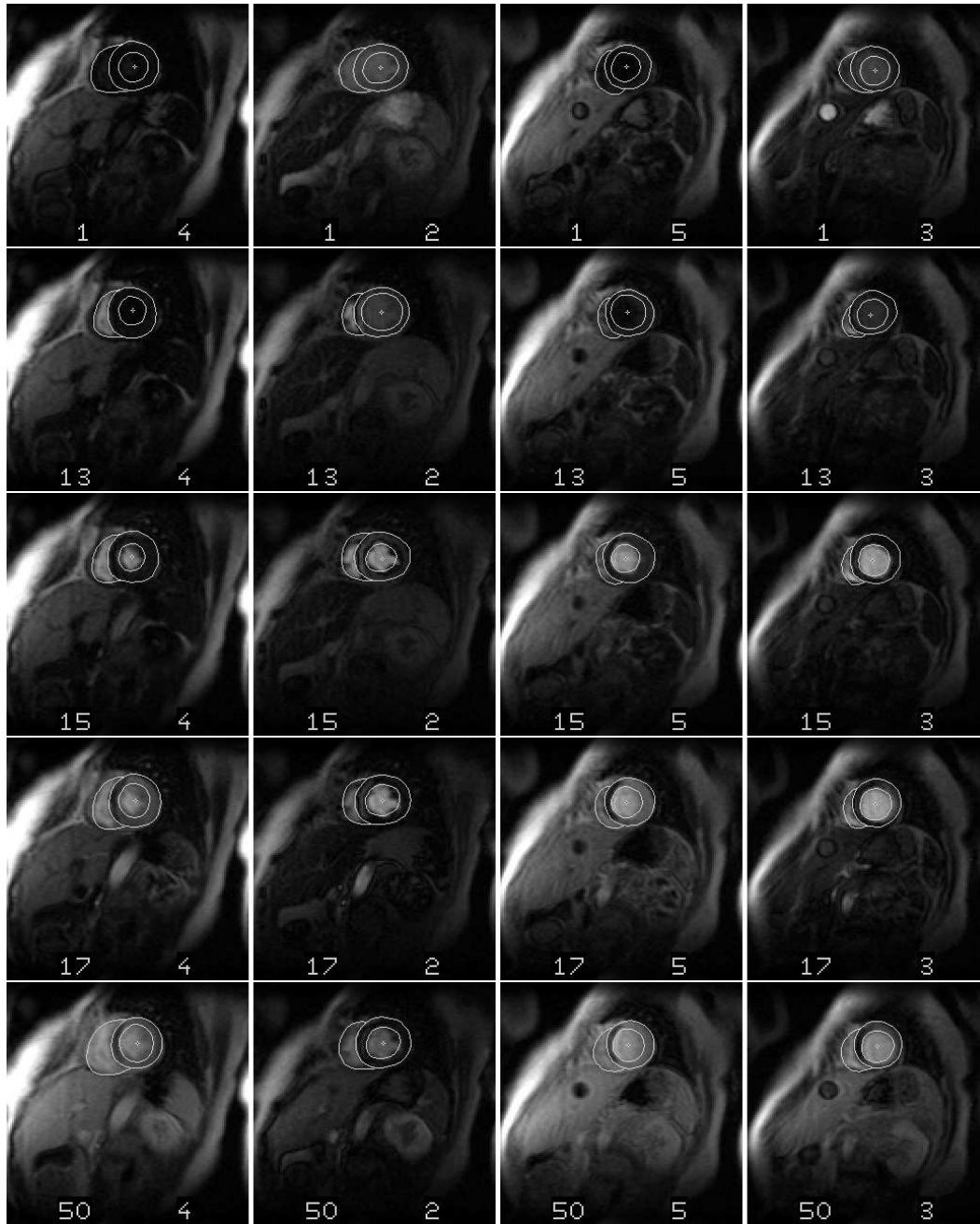


Fig. 10. Registration results for patient 8 before, during and after bolus passage. should be reviewed by the operator, prior to making judgements upon automatically generated perfusion maps. The operator workload involved herein corresponds approximately to performing manual annotation of a single frame. In summary, a mean point to curve distance of 1.25 ± 0.36 pixels was obtained by our method.

Figure 12 reveals that a substantial part of the landmark error stems from imperfections in the localisation of the right ventricle. LV registration errors are consequently smaller than the reported averages in the previous section.

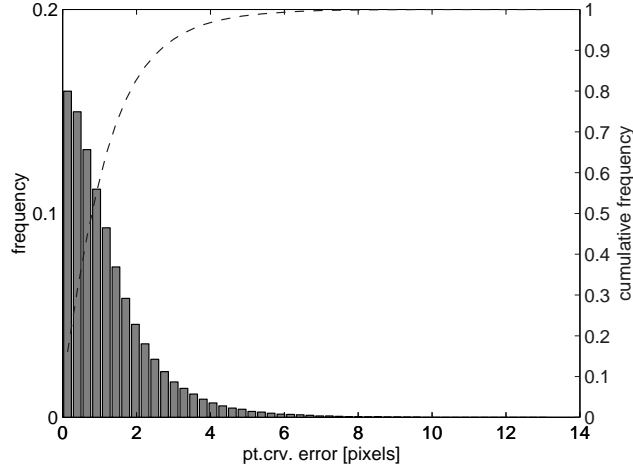


Fig. 11. Distribution of pt.crv. landmark errors for all frames.

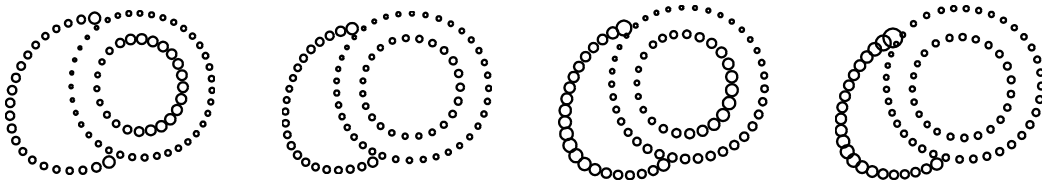


Fig. 12. Mean pt.crv. landmark errors plotted as circles on the mean shape of each slice.

Although the inclusion of the RV adds specificity to the presented model, a hierarchical approach using a separate LV model initialised from the converged RV-LV model is straightforward and should be well worth to pursue.

Since only weak assumptions regarding sequence length and bolus passage position are present in the method, this is easily adapted to arbitrary-length perfusion sequences. The method only requires a coarse indication of the bolus passage position and a stable period of a reasonable number of frames after the bolus, the latter being a standard requirement in perfusion MRI. As stated earlier, the former should be possible to obtain automatically from the extremum of the temporal image derivative.

We believe that the presented method has a noteworthy set of merits compared to other approaches. i) It is based upon a well-understood and well-described framework; the Active Appearance Models, which also has a freely available implementation. ii) It has met the robustness requirements crucial when dealing with perfusion MRI. iii) The implementation showed that results could be produced within a reasonable time frame. iv) Imperfect ECG-triggering is handled by allowing the model to perform complex deformations. This further enables detection of mis-triggered slices and the possibility of subsequently excluding these due to the movement of the heart during the cardiac cycle. This could be an issue for patients with arrhythmia. v) Although current

MR pulse sequences do not allow for multi-slice perfusion sequences without ECG-triggering, this can also be handled by the deformation properties of our method. Finally, vi) To let our method generalise to new data, very few assumptions concerning the data content have been made. Except for a few scalar parameters – which are dimensionless indices relating to the statistics of the actual data – all values are estimated from training data, rather than being explicitly coded into the method. We believe this to be a very fruitful approach, as the method easily adapts to new expert knowledge given by medical doctors. Knowledge, that typically already exists in the form of manually annotated training data from previous studies.

However, the above virtues come at a price. Image annotation is not avoided altogether. A training set needs to be carefully crafted in a similar process to the one this method is aiming at eliminating. However, we have taken steps to alleviate this part significantly by introducing spline-based curve fitting and MDL shape modelling into the annotation process. Further, even if a method requires no training data, a validation set still needs to be established. In summary, we do not consider this issue a major Achilles heel. Secondly, this framework is built upon assumptions of correspondence. Consequently, papillary muscles, which are present in some cases and seemingly not in others, cannot be modelled and thus provides a limit on the level of detail obtainable by such methods. This implies that shape models working on data sets of similar quality only apply to the clinical convention of including papillary muscles into the blood pool and thereby obtaining a much simpler geometry.

Additionally, compared to other approaches given in Table 1 our method has the advantage of modelling all slices simultaneously. Furthermore, our computation times are well competitive, keeping in mind that they apply to modelling of four slices instead of one. Some of the approaches only correct for translational motion, disregarding rotation and deformable shape changes. Our deformable modelling approach gives an accurate registration – a dense point correspondence through the whole sequence. Further, we have given both qualitative and quantitative validation of our method involving comparison with manually defined (ground truth) myocardial boundaries. This is a clear advantage to most of the other approaches, which rely on relative comparison of parameters derived from the raw and the registered data.

Compared to our previous work in Stegmann and Larsson (2003) several extensions have been introduced. A larger data set containing multiple slices from patients, compared to single slices from controls, has been obtained. Image annotation and shape models are substantially improved using spline interpolation and MDL optimised shape models. Slice coherence is modelled compared to single slice-oriented models. Models are utilising the full dynamic range of the MR scanner compared to the reduced 8-bit representation avail-

able previously.² Automated model initialisation is introduced to avoid the need for any manual interaction. Cardiac interfaces to the upper abdomen and the lung are now included into the texture model to improve specificity.

The initialisation part aside, all presented results here are produced reasonably rapidly, albeit not as rapid as in the previous work. This has several reasons. Models are substantially larger as they encompass several slices and include the cardiac interfaces as well. Images have twice the bit depth, which means more memory transfer, less CPU cache consistency, et cetera. Finally, there has been put far less emphasis on code optimisation compared to the previous work. We anticipate that substantial improvements can be achieved on this level alone.

We intend to expand the current work in the following areas. While the need for clustered texture models is obvious, the presented results still leave room for improvement. This is also true for the brute force model selection used, which makes search time (approximately) linear in the number of classes. Model initialisation could be improved by sampling the pose distribution of the training set and by using a multi-resolution implementation, which both will decrease the initialisation time substantially. Further, the claims regarding automatic bolus detection should be proven correct. But most importantly, a clinical validation on a much larger training set should be carried out to demonstrate the true capabilities of the presented statistical approach and its alleged adaptive properties. In addition, it should be explored to what extent LV sub models – as mentioned earlier – will improve the accuracy of myocardium registration. Performance evaluation should be extended to include the clinical end goal; the various generated perfusion maps. Consequently, estimated perfusion parameters should be compared as obtained from automatic versus hand-registered sequences.

7 Conclusion

We have described a novel, data-driven method for registration and motion-compensation of multi-slice cardiac perfusion MRI. Preliminary validation of the method on ten patients with acute myocardial infarction showed promising results. In nine cases results were obtained without manual interaction. Despite evident perfusion deficits and varying image quality in the limited training set our method yielded a mean point to curve distance of 1.25 ± 0.36 pixels for the left and right ventricle combined. We conclude that this learning-based method holds great promise for the automation of cardiac perfusion investigations, due

² Reconstructed MR data was only available to the earlier study in an 8-bit storage format.

to its accuracy, robustness and generalisation ability.

A Estimation of Σ and σ

The dispersion matrix of the pose parameters, Σ , and the standard deviations of the shape parameters, σ , can be estimated from the stable period. However, in order to obtain more reliable estimates we propose to use the training sequences, since many additional samples are available. To avoid confusion of inter- and intra-sequence variability, we filter out sequence-specific information (such as mean shape size) prior to the estimation. Now, let R denote the number of sequences, each containing P frames. Further, let \odot denote the Hadamard, i.e. element-wise, product. Then the estimation of Σ and σ can be specified as shown in Algorithm 2.

Algorithm 2 Estimation of Pose and Shape Variation

- | | | |
|----|--|---|
| 1: | for $i = 1$ to R (for each training sequence) do | |
| 2: | $\mathbf{p}_t \leftarrow \mathbf{p}_t \odot (\frac{1}{\bar{s}} [1 \ 1 \ 1 \ \bar{s}]^T) \ \forall \ t$ | Normalise pose w.r.t. mean size |
| 3: | $\bar{\mathbf{p}}_i = \frac{1}{P} \sum_{t=1}^P \mathbf{p}_t$ | Estimate pose seq. mean |
| 4: | $\Sigma_i = \frac{1}{P-1} \sum_{t=1}^P (\mathbf{p}_t - \bar{\mathbf{p}}_i)(\mathbf{p}_t - \bar{\mathbf{p}}_i)^T$ | Estimate pose dispersion matrix |
| 5: | $\bar{\mathbf{b}}_{s,i} = \frac{1}{P} \sum_{t=1}^P \mathbf{b}_{s,t}$ | Estimate shape mean |
| 6: | $\sigma_i = \sqrt{\frac{1}{P-1} \sum_{t=1}^P (\mathbf{b}_{s,t} - \bar{\mathbf{b}}_{s,i}) \odot (\mathbf{b}_{s,t} - \bar{\mathbf{b}}_{s,i})}$ | Estimate seq. shape std. dev. |
| 7: | end for | |
| 8: | $\Sigma = \frac{1}{R} \sum_{i=1}^R \Sigma_i$ | Calculate pooled estimate for all sequences |
| 9: | $\sigma = \frac{1}{R} \sum_{i=1}^R \sigma_i$ | Calculate pooled estimate for all sequences |
-

Further, to maintain simplicity, normalisation with respect to rotation has not been included in the above algorithm. However, if sequences differ substantially in orientation, this should be taken into account when estimating t_x and t_y . Finally, during CAAM search, appropriate normalisation and de-normalisation using the stable period mean size (and orientation) of the current sequence must be carried out.

References

- Ablitt, N., Gao, J., Gatehouse, P., Yang, G. Z., 2002. Motion decoupling and registration for 3D magnetic resonance myocardial perfusion imaging. In: International Conference Computational Science - ICCS 2002. Vol. 2331. Springer-Verlag, pp. 285–294.
- Bansal, R., Funka-Lea, G., 2002. Integrated image registration for cardiac MR perfusion data. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002. 5th International Conference. Proceedings, Part I (Lecture Notes in Computer Science Vol.2488), 659–66.

- Behloul, F., Roux, J. P., Unterreiner, R., 1997. Automatic boundary detection of the left ventricle and measurement of local myocardial perfusion in MRI. *Computers in Cardiology 1997*, 145–148.
- Bidaut, L. M., Vallee, J. P., 2001. Automated registration of dynamic MR images for the quantification of myocardial perfusion. *Jour. Magn. Reson. Imaging* 13 (4), 648–655.
- Bosch, H. G., Mitchell, S. C., Lelieveldt, B. P. F., Nijland, F., Kamp, O., Sonka, M., Reiber, J. H. C., 2001a. 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.
- Bosch, J. G., Mitchell, S. C., Lelieveldt, B. P. F., Nijland, F., Kamp, O., Sonka, M., Reiber, J. H. C., 2001b. Fully automated endocardial contour detection in time sequences of echocardiograms by active appearance motion models. *Computers in Cardiology*, 93–96.
- Bosch, J. G., Mitchell, S. C., Lelieveldt, B. P. F., Nijland, F., Kamp, O., Sonka, M., Reiber, J. H. C., 2002. Automatic segmentation of echocardiographic sequences by active appearance motion models. *Medical Imaging, IEEE Transactions on* 21 (11), 1374–1383.
- Bracoud, L., Vincent, F., Pachai, C., Canet, E., Croisille, P., Revel, D., 2003. Automatic registration of MR first-pass myocardial perfusion images. In: *Functional Imaging and Modeling of the Heart, FIMH 2003*. Vol. 2674 of *Lecture Notes in Computer Science*. Springer Verlag, Lyon, France, pp. 215 – 223.
- Breeuwer, M., Quist, M., Spreeuwers, L., Paetsch, I., Al-Saadi, N., Nagel, E., 2001a. Towards automatic quantitative analysis of cardiac MR perfusion images. *International Congress Series* 1230, 967–973.
- Breeuwer, M., Spreeuwers, L., Quist, M., 2001b. Automatic quantitative analysis of cardiac MR perfusion images. *Proceedings of the SPIE - The International Society for Optical Engineering* 4322 (1-3), 733–42.
- Cootes, T. F., Edwards, G. J., Taylor, C. J., 1998. Active appearance models. In: *Proc. of European Conf. on Computer Vision 1998*. Vol. 1407 of *Lecture Notes in Computer Science*. Springer, pp. 484–498.
- Cootes, T. F., Edwards, G. J., Taylor, C. J., 2001. Active appearance models. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 23 (6), 681–685.
- Cootes, T. F., Taylor, C. J., 2001. *Statistical Models of Appearance for Computer Vision*. Tech. report, University of Manchester.
URL <http://www.isbe.man.ac.uk/bim/>
- Cootes, T. F., Taylor, C. J., Cooper, D. H., Graham, J., 1995. Active shape models – their training and application. *Computer Vision and Image Understanding* 61 (1), 38–59.
- Cootes, T. F., Wheeler, G. V., Walker, K. N., Taylor, C. J., 2000. Coupled-view active appearance models. *BMVC 2000*, 52–61 vol.1.
- Davies, R. H., 2002. *Learning shape: Optimal models for analysing natural variability*. Ph.D. thesis, Department of Imaging Science and Biomedical

- Engineering, University of Manchester, Stopford building, University of Manchester.
- Davies, R. H., Cootes, T. F., Taylor, C. J., 2001. A minimum description length approach to statistical shape modelling. *Information Processing in Medical Imaging. 17th International Conference, IPMI 2001. Proceedings (Lecture Notes in Computer Science Vol.2082)*, 50–63.
- Davies, R. H., Twining, C. J., Cootes, T. F., Waterton, J. C., Taylor, C. J., 2002a. 3D statistical shape models using direct optimisation of description length. In: *Proc. ECCV 2002. Vol. 2352 of LNCS. Springer*, pp. 3–20.
- Davies, R. H., Twining, C. J., Cootes, T. F., Waterton, J. C., Taylor, C. J., 2002b. A minimum description length approach to statistical shape modelling. *Medical Imaging, IEEE Transactions on* 21 (5), 525–537.
- Dornier, C., Ivancevic, M. K., Thevenaz, P., Vallee, J.-P., 2003. Improvement in the quantification of myocardial perfusion using an automatic spline-based registration algorithm. *Journal of Magnetic Resonance Imaging* 18 (2), 160–168.
- Edwards, G. J., Taylor, C. J., Cootes, T. F., 1998. Interpreting face images using active appearance models. In: *Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition. IEEE Comput. Soc*, pp. 300–5.
- Forgey, E., 1965. Cluster analysis of multivariate data. *Biometrics* 21, 768.
- Gallippi, C. M., Gregg, E. T., 2001. Automatic image registration for MR and ultrasound cardiac images. *Information Processing in Medical Imaging. 17th International Conference, IPMI 2001. Proceedings (Lecture Notes in Computer Science Vol.2082)*, 148–54.
- Gao, J., Ablitt, N., Elkington, A., Yang, G. Z., 2002. Deformation modelling based on PLSR for cardiac magnetic resonance perfusion imaging. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002. 5th International Conference. Proceedings, Part I (Lecture Notes in Computer Science Vol.2488)*, 612–19.
- Gower, J. C., 1975. Generalized Procrustes analysis. *Psychometrika* 40, 33–50.
- Gupta, S. N., Solaiyappan, M., Beache, G. M., Arai, A. E., Foo, T. K. F., 2003. Fast method for correcting image misregistration due to organ motion in time-series MRI data. *Magnetic Resonance in Medicine* 49 (3), 506–14.
- Hsu, L., Rhoads, K. L., Aletras, A. H., Arai, A. E., 2003. Surface coil intensity correction and non-linear intensity normalization improve pixel-resolution parametric maps of myocardial MRI perfusion. In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003, 6th Int. Conference, Montréal, Canada. LNCS 2879. Springer*, pp. 975–976.
- Larsson, H. B. W., Fritz-Hansen, T., Rostrup, E., Søndergaard, L., Ring, P., Henriksen, O., 1996. Myocardial perfusion modeling using MRI. *Magnetic Resonance in Medicine* 35, 716–726.
- Lelieveldt, B. P. F., Mitchell, S. C., van der Geest, R. J., Bosch, H. G., Sonka, M., Reiber, J. H. C., 2001. Time continuous segmentation of cardiac MR images using active appearance motion models. In: *Computer Assisted Radiology and Surgery, CARS 2001, International Congress Series. Vol. 1230.*

- Elsevier Science, pp. 961–966.
- Lelieveldt, B. P. F., Uzumcu, M., van der Geest, R. J., Reiber, J. H. C., Sonka, M., 2003. 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.
- Mitchell, S., Lelieveldt, B., Geest, R., Bosch, H., Reiber, J., Sonka, M., 2001. Time continuous segmentation of cardiac MR image sequences using active appearance motion models. In: *Proc. of SPIE Medical Imaging 2001*. Vol. 1. pp. 249–256.
- Muhling, O., Jerosch-Herold, M., Nabauer, M., Wilke, N., 2003. Assessment of ischemic heart disease using magnetic resonance first-pass perfusion imaging. *Herz* 28 (2), 82–89.
- Nagel, E., Al-Saadiand, N., Fleck, E., 2000. Cardiovascular magnetic resonance: Myocardial perfusion. *Herz* 25 (4), 409–416.
- Oost, E., Lelieveldt, B. P. F., Koning, G., Sonka, M., Reiber, J. H. C., 2003. Left ventricle contour detection in X-ray angiograms using multi-view active appearance models. *SPIE Medical Imaging 2003 5032 I*, 394–404.
- Sonka, M., Lelieveldt, B. P. F., Mitchell, S. C., Bosch, J. G., van der Geest, R. J., Reiber, J. H. C., 2001. Active appearance motion model segmentation. *Proceedings Second International Workshop on Digital and Computational Video*, 64–68.
- Spreeuwers, L., Breeuwer, M., 2001. Automatic detection of the myocardial boundaries of the right and left ventricle in MR cardio perfusion scans. *Proceedings of SPIE - The International Society for Optical Engineering 4322 (3)*, 1207–1217.
- Spreeuwers, L. J., Wierda, F., Breeuwer, M., 2002. Optimal myocardial boundary estimation for MR cardio perfusion measurements using sensitivity analysis. *Computers in Cardiology, 2002*, 197–200.
- Stegmann, M. B., 2001. Object tracking using active appearance models. In: *Proc. 10th Danish Conference on Pattern Recognition and Image Analysis, Copenhagen, Denmark*. Vol. 1. DIKU, pp. 54–60.
- Stegmann, M. B., Davies, R. H., Ryberg, C., feb 2004. Corpus callosum analysis using MDL-based sequential models of shape and appearance. In: *International Symposium on Medical Imaging 2004, San Diego CA, SPIE*. SPIE, pp. 612–619.
- Stegmann, M. B., Ersbøll, B. K., Larsen, R., 2003. FAME – a flexible appearance modelling environment. *IEEE Trans. on Medical Imaging* 22 (10), 1319–1331.
- Stegmann, M. B., Larsson, H. B. W., 2003. Motion-compensation of cardiac perfusion MRI using a statistical texture ensemble. In: *Functional Imaging and Modeling of the Heart, FIMH 2003*. Vol. 2674 of LNCS. Springer, Lyon, France, pp. 151–161.
- Thodberg, H. H., 2003. Minimum description length shape and appearance models. In: Taylor, C., Noble, J. A. (Eds.), *Proceedings of Information Processing in Medical Imaging (IPMI 2003)* ., Lecture Notes In Computer

- Science (LNCS) 2732. Springer, pp. 51 – 62.
URL <http://www.imm.dtu.dk/pubdb/p.php?2973>
- Yang, G. Z., Burger, P., Panting, J., Gatehouse, P. D., Rueckert, D., Pennell, D. J., Firmin, D. N., 1998. Motion and deformation tracking for short-axis echo-planar myocardial perfusion imaging. *Medical Image Analysis* 2 (3), 285–302.
- Yang, G. Z., Gatehouse, P. D., Panting, J., Burger, P., Pennell, D. J., Firmin, D. N., 1997. Motion analysis for magnetic resonance myocardial perfusion imaging. *Image Processing and Its Applications, 1997.*, Sixth International Conference on 2, 838–842 vol.2.