Conditional Shape Models for Cardiac Motion Estimation

Coert Metz[1,*], Nora Baka[1,2,*], Hortense Kirisli[1,2], Michiel Schaap[1],
Theo van Walsum[1], Stefan Klein[1], Lisan Neefjes[3], Nico Mollet[3],
Boudewijn Lelieveldt[2,4], Marleen de Bruijne[1,5], and Wiro Niessen[1,4,**]

[1] Dept. of Rad. and Med. Informatics, Erasmus MC, Rotterdam, The Netherlands
[2] Leiden University Medical Center, Leiden, Netherlands
[3] Dept. of Radiology and Cardiology Erasmus MC, Rotterdam, The Netherlands
[4] Delft University of Technology, The Netherlands
[5] University of Copenhagen, Denmark

Abstract. We propose a conditional statistical shape model to predict
patient specific cardiac motion from the 3D end-diastolic CTA scan. The
model is built from 4D CTA sequences by combining atlas based seg-
mentation and 4D registration. Cardiac motion estimation is, for exam-
ple, relevant in the dynamic alignment of pre-operative CTA data with
intra-operative X-ray imaging. Due to a trend towards prospective elec-
trocardiogram gating techniques, 4D imaging data, from which motion
information could be extracted, is not commonly available. The predic-
tion of motion from shape information is thus relevant for this purpose.
Evaluation of the accuracy of the predicted motion was performed using
CTA scans of 50 patients, showing an average accuracy of 1.1 mm.

1   Introduction

Coronary angioplasty is an often applied procedure to reopen narrowed or oc-
cluded coronary arteries. Real time X-ray visualization guides these interven-
tions, providing information about the morphology of the patent vessels. For
difficult interventions, such as reopening chronic total occlusions, integration of
a pre-operative CTA acquisition is expected to improve the intervention result.
Hereto, a correct alignment of the CTA image with the 2D X-ray projection
images is required, which consists of both a rigid alignment to determine the
pose and orientation of the data and a non-rigid alignment to compensate for
cardiac motion. A patient specific cardiac motion prior can be derived from 4D
retrospectively gated CTA data [1]. However, the trend in cardiac CT acquisition
is towards prospective electrocardiogram (ECG) gating techniques to decrease
the effective patient dose. In these situations, the derivation of dynamic infor-
mation from CTA data is often not possible, as only one phase of the cardiac

---

* Both authors contributed equally to this work.

cycle is imaged. Moreover, it is often not known before the CTA acquisition if a patient will undergo coronary angioplasty, making selective dynamic imaging for a subset of patients difficult.

The purpose of this work is to develop and evaluate a method to predict patient specific cardiac motion from a single 3D CTA image. To this end a 4D statistical shape model is built from 50 ECG-gated dynamic CTA images by combining 3D multi-atlas registration and 4D registration. Motion is estimated for an unseen patient by first segmenting the cardiac structures in the acquired 3D CTA scan. Subsequently we derive the most probable motion given the segmentation and the statistical model using a conditional Gaussian distribution.

Statistical shape models have been frequently applied for 3D image segmentation [2]. The use of active shape models for dynamic image segmentation has been investigated, for instance in the work of Ordas et al., who built a single 3D model with training data from multiple ECG phases [3]. However, this method cannot be used for the cardiac motion prediction problem, as a single 3D shape model of the heart is built, without distinguishing between inter-patient and intra-patient variability. Models built from 4D landmark positions can overcome this problem. Such a 4D statistical model was, for instance, built by Perperidis et al. for segmenting the left ventricle, right ventricle, and myocardium, but was not used for motion prediction [4]. Hoogendoorn et al. built a bilinear model for the extrapolation of cardiac motion, assuming that the motion of the heart is independent of its shape [5]. We, in contrast, build statistical shape models of the shape and motion of the heart without assuming their independence and evaluate its applicability for the prediction of cardiac motion by conditioning the motion model on single time point shape information.

2   Methods

The statistical model is built using multi-atlas based segmentation and 4D registration. The structures of interest are the aorta (Ao), endocardium left ventricle (endoLV), epicardium left ventricle (epiLV), right ventricle, left atrium (LA) and right atrium (RA). An atlas landmarking procedure ensures anatomical landmark correspondence. Cardiac motion is predicted by conditioning the motion model on the landmark points of the shape in the 3D end-diastolic image. These steps are explained in more detail in the following sections.

2.1   Statistical Shape Model

The K 4D CTA sequences that are used to build the statistical models [6] consist of T time points, each of which contains 3D shapes represented by n landmark points. Landmark correspondence is ensured both along the sequence and across the training set (see Sect. 2.2). First, a base time point b is determined, which is defined as the time point in the cardiac cycle for which the 3D reconstructions are available. In this work, we use the end-diastolic time point for this purpose. Subsequently, we represent the 4D sequence as a 3D shape in time point b, and

a motion sequence given by the vectors describing the displacements from this time point to all time points in the sequence. Accordingly, two separate models are built: a 3D statistical shape model (at time point b) and a 4D statistical motion model. Because the models are created for prediction purposes, the alignment (translation, rotation and isotropic scaling) of the shapes in the training sequences is performed using the landmarks of the shape in the base time point. This is in agreement with the alignment of the segmented end-diastolic target shape to the model. Assuming a Gaussian shape and motion distribution, PCA analysis is performed to extract the main modes of variation resulting in the following shape and motion model

$$\mathbf{s}_b \approx \bar{\mathbf{s}}_b + \mathbf{\Phi}\mathbf{p}, \qquad \text{and} \quad \mathbf{m} \approx \begin{bmatrix} \mathbf{s}_1 - \mathbf{s}_b \\ \vdots \\ \mathbf{s}_T - \mathbf{s}_b \end{bmatrix} = \bar{\mathbf{m}} + \mathbf{\Psi}\mathbf{q} \qquad (1)$$

where $\mathbf{s}_i$ is the shape of the i-th time point of a sequence, b denotes the base time point, and $\mathbf{m}$ is the motion vector. The mean shape and motion is denoted with $\bar{\mathbf{s}}_b$ and $\bar{\mathbf{m}}$, while $\mathbf{p}$ and $\mathbf{q}$ are the shape and motion parameter vectors respectively. The models were reduced to explain 95% of the total variance.

2.2   Model Construction

Atlases and Atlas Point Correspondence. 3D segmentation of the cardiac structures at one time point in the cardiac cycle is achieved by multi-atlas registration using $J = 8$ atlases [7]. We used `elastix` for all registration procedures in this work [8]. The following procedure was applied to the atlases to obtain automatic anatomical landmark correspondence:

1. Non-rigid B-spline registration of all atlas pairs to each other, resulting in $J *$ J transformations denoted as $\mathbf{T}_{ij}$ where i and j indicate the atlas numbers.
2. Determination of the mean transformation for every atlas: $\overline{\mathbf{T}}_i = \frac{1}{J}\sum_j \mathbf{T}_{ij}$ and subsequent transformation of the manual annotations to the mean space.
3. Creation of a signed distance map SDM_{is} for every atlas i and structure s in the atlas (a) and mean (m) space and averaging of the signed distance maps for all structures in the mean space: $\overline{\mathrm{SDM}}_s^m = \frac{1}{J}\sum_i \mathrm{SDM}_{is}^m$.
4. Creation of a mean surface with approximately equally distributed landmark locations for every structure by extracting the surface at the zero level set of $\overline{\mathrm{SDM}}_s^m$. The number of landmarks used was 15835 (Ao: 1316, endoLV: 2723, epiLV: 4106, LA: 2034, RA: 2079, RV: 3577).
5. Transformation of the obtained mean surface landmarks back to the atlas spaces using $\overline{\mathbf{T}}_i^{-1}$ and projection of the landmark points of the mean surfaces onto the annotated surfaces by non-rigidly registering their SDMs.

Landmark Determination in 4D CTA Sequences. 3D segmentation of the heart structures in the end-diastolic time point of the K 4D CTA sequences is performed by:
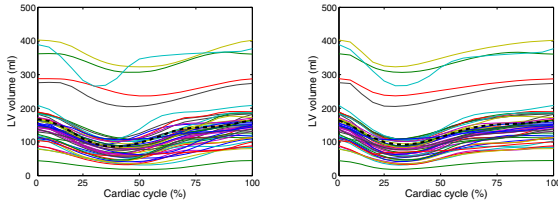
Fig. 1. Left ventricular volume during the cardiac cycle before and after alignment of the volume curves. Each colored line represents a patient and the dashed black line represents the reference curve.

1. Non-rigid B-spline registration of all atlases to the end-diastolic CTA image.
2. Transformation of the surface landmarks from the J atlases to the target image using the resulting transformations.
3. Combination of the J atlas landmark sets into one landmark set by application of the mean shift algorithm to the landmark coordinates [9]. This ensures a robust average of the corresponding landmark points and reduces the influence of possible registration errors.

This procedure results in a set of landmarks describing the cardiac structures at end-diastole. Subsequently, the segmentation of the cardiac structures in all time phases is achieved by determining the cardiac motion for each patient using 4D B-spline registration on the 4D CTA image. During registration, the variance of intensity values at corresponding spatial locations is minimized. Only deformations in the spatial domain are allowed and the deformation is forced to be periodic.

By combining the segmented shapes and the deformations resulting from the 4D registration procedure we determine 4D landmark sets for every patient. To ensure temporal correspondence between the 4D landmark sets, an additional alignment step is performed to align the moment of contraction and relaxation during the cardiac cycle between patients. To this end, time curves of left ventricular volume are obtained from the K 4D segmentations. These curves are subsequently aligned by a 1D non-rigid group-wise registration approach in which the cross-correlation between the curves is optimized. After this alignment, new landmark positions at twenty regular intervals of the cardiac cycle are determined by applying spline interpolation between corresponding landmark points in time. Volume graphs before and after alignment are shown in Fig. 1.

2.3   Conditional Model

We derive the motion given the shape of the cardiac structures at one time point in the cardiac cycle by assuming a Gaussian distribution of the shape vectors $\mathbf{s}_b \in S$ and motion vectors $\mathbf{m} \in M$ respectively. Their combined distribution $P(M, S)$ is a normal distribution with mean

$$\mu = \begin{bmatrix} \bar{\mathbf{m}} \\ \bar{\mathbf{s}}_b \end{bmatrix} \quad \text{and variance} \quad \Sigma = \begin{bmatrix} \Sigma_{MM} & \Sigma_{MS} \\ \Sigma_{SM} & \Sigma_{SS} \end{bmatrix}, \tag{2}$$

with $\mathbf{\Sigma}_{SS}$ and $\mathbf{\Sigma}_{MM}$ the covariance matrix of shape and motion respectively, and $\mathbf{\Sigma}_{SM}$ the covariance matrix between shape and motion. The most likely motion $\hat{\mathbf{m}}$ given a single shape $\mathbf{s}_b^*$ is estimated as the mean of the conditional probability function $P(M|\mathbf{s}_b^*)$ [10]:

$$\hat{\mathbf{m}} = \bar{\mathbf{m}} + \mathbf{\Sigma}_{MS}\mathbf{\Sigma}_{SS}^{-1}(\mathbf{s}_b^* - \bar{\mathbf{s}}_b) \tag{3}$$

However, as the dimensionality of the 3D shape vectors is larger than the number of training shapes, the covariance matrix $\mathbf{\Sigma}_{SS}$ becomes singular, and cannot be inverted. Therefore, conditioning is performed on the shape and motion model parameters after applying PCA rather than on the shape and motion vectors themselves. As the mean parameter vectors are zero, Equation 3 simplifies to:

$$\hat{\mathbf{q}} = \mathbf{\Sigma}_{qp}\mathbf{\Sigma}_{pp}^{-1}\mathbf{p}^*, \quad \text{where} \quad \mathbf{p}^* = \mathbf{\Phi}^T(\mathbf{s}_b^* - \bar{\mathbf{s}}_b), \tag{4}$$

with $\hat{\mathbf{q}}$ being the parameter representation of the estimated motion $\hat{\mathbf{m}}$. The generation of plausible motion is ensured by scaling $\hat{\mathbf{q}}$ such that it always lies within ± 3 standard deviations of the motion model. The final conditional sequence is computed from the known 3D shape $\mathbf{s}_b^*$ and the derived motion $\hat{\mathbf{m}} = \bar{\mathbf{m}} + \mathbf{\Psi}\hat{\mathbf{q}}$.

3   Experiments and Results

3.1   Imaging Data

Eight 3D diastolic CTA reconstructions were used as the atlas images and 50 retrospectively gated 4D CTA images were used for building and evaluating the 4D model. All scans were acquired with Siemens CT scanners and 20 time points were reconstructed for the 4D CTA scans. Note that the variety of the data is large as healthy subjects are not imaged because of radiation regulations.

3.2   Cardiac Motion Prediction

The accuracy of the motion prediction was evaluated in leave-one-out experiments by applying the conditional model described in Sect. 2.3 on end-diastolic segmentations, which are most often available in clinical practice. The accuracy of the resulting 4D shapes was determined by computation of the average root mean squared point-to-surface distance per structure and time point. The results are compared to a projection of the complete sequence onto the model space. This projection represents the best possible reconstruction of an entirely known 4D input sequence given the training data. Note, that this represents a practical limit for accuracy and that the 4D information is not available for the motion estimation. Also, a replication of the end-diastolic segmentations for all time points was used as a worst-case estimate for benchmarking the errors of the conditional shape estimates. Graphs of the accuracy over time are shown in Fig. 2. Table 1 lists the root mean squared point-to-surface distance for both the replication and the conditioning averaged over all time points and over the

Table 1. Mean and standard deviation of root mean squared point-to-surface distance in mm. for replication (assuming no motion) and the proposed method. Values are listed for all time points and for the minimal left ventricular volume time point (30%).

| Structure	Ao	endoLV	epiLV	RV	LA	RA
Replication (all)	1.6 (1.1)	2.2 (1.8)	1.7 (1.4)	1.7 (1.3)	1.6 (1.0)	1.2 (0.9)
Conditioning (all)	0.9 (0.6)	1.1 (0.8)	0.9 (0.6)	0.9 (0.7)	1.0 (0.6)	0.8 (0.5)
Replication (at 30%)	3.2 (1.0)	5.1 (2.0)	4.0 (0.9)	3.7 (0.8)	2.8 (0.8)	2.5 (0.6)
Conditioning (at 30%)	1.4 (0.6)	2.0 (0.7)	1.7 (0.5)	1.5 (0.5)	1.4 (0.5)	1.2 (0.5)
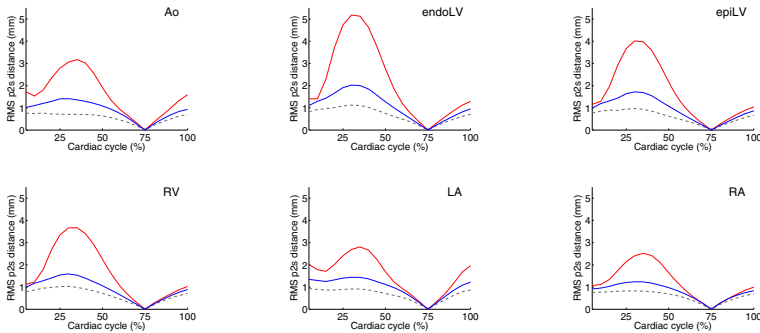
Fig. 2. Average root mean squared (RMS) point-to-surface (p2s) distance over the cardiac cycle. RMS values are averaged over all patients. The red line indicates the replication error. The grey dotted line shows the error for projecting the entire sequence onto the model space and the blue line shows the error when conditioning is used.

time point at which the left ventricular volume was minimal (30%). We expect the shapes at this time point to differ the most from the end-diastolic shapes used for the motion estimation. Fig. 3 shows a color coding of the distances for the left and right ventricle, and atrium at end-systole and end-diastole for two randomly selected patients.

3.3   Model Generalization Ability and Training Set Size

An experiment was conducted to assess the dependency between the accuracy of the model and the training set size. The available training sequences were grouped randomly in subgroups, each containing 5 sequences. Subsequently, for every tested training set size all possible combinations of groups were taken as training set, and a random left out group was used for testing. Fitting was performed by projecting the motion of the 4D segmentation of a CTA sequence on the subspace of the motion model, and back to the original space to derive the motion vectors for these parameters. The resulting motion was applied to
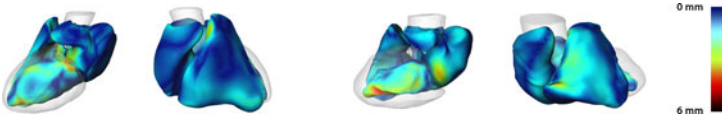
Fig. 3. Two views of end-systolic predicted surfaces for two randomly selected patients. Color coded are the point-to-surface distances from the predicted shape to the segmented shape. The Ao and epiLV are for visualization purposes not color coded.
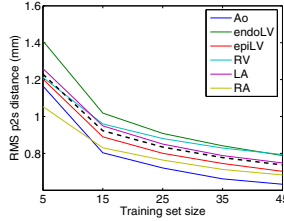
Fig. 4. Root mean squared point-to-surface (p2s) distance in mm. for every structure with respect to training set size

the atlas segmented shape, and compared with the actual 4D segmentations. Fig. 4 shows the results.

4   Discussion and Conclusion

The experiments show that the conditional shape model is able to predict the motion of the heart with an average accuracy of around 1.0 mm (see Table 1). This represents a large improvement compared to shape replication, which is the only option when only a 3D CTA scan is available (see Fig. 2). Accuracy is measured by the point-to-surface error, but depending on the application of interest, the point-to-point error can be used as well. For the here presented experiments, point-to-point errors were on average larger, but showed very similar trends.

Furthermore, the remaining errors when projecting the entire sequence onto the model subspace and the results of the generalization ability experiment show that the size of the training set is most probably not sufficient. Therefore, the effect of adding more training shapes on the accuracy will be investigated in the future. The errors at the left of the graph are related to the magnitude of the motion found by the registration approach. Errors induced by landmark propagation are currently not taken into account and subject to future work.

In the current work no assumptions were made about the relation between cardiac shape and motion, in contrast to the method proposed by Hoogendoorn et al., in which shape and motion variations are explicitly decoupled [5]. Although the motion of the heart depends for a large amount on its electrophysiology, which is not apparent from its shape, we think that shape can still predict cardiac

motion when both the shape and motion of the heart are affected by disease (e.g. due to cardiac remodelling after myocardial infarction). Motion changes related to the shape at other time points in the cardiac cycle are not taken into account.

The conditioning framework described in this paper exploits any available correlation between shape and motion given in the training sequences, and naturally reduces to the mean motion in case no correlation is present. Due to the small training set size we were not able to draw any conclusions with respect to this correlation. We plan to investigate this topic in detail as part of future work. For the conditioning 95% of the statistical shape model variance was retained, however the effect of the value of this parameter on the prediction results should still be investigated.

In conclusion, we presented and evaluated a method that uses conditional shape models to predict cardiac motion from 3D cardiac shape. The results suggest that it can reliably predict the motion of the heart. Our goal is to use this model for the dynamic alignment of pre-operatively acquired CTA images with intra-operative X-ray imaging, which should be applied and evaluated in future work.

References

1. Metz, C., et al.: Patient specific 4D coronary models from ECG-gated CTA data for intra-operative dynamic alignment of CTA with X-ray images. In: Med. Image Comput. Assist Interv. Int. (2009)
2. Heimann, T., Meinzer, H.P.: Statistical shape models for 3D medical image segmentation: a review. Med. Image Anal. 13(4), 543–563 (2009)
3. Ordas, S., Oubel, E., Leta, R., Carreras, F., Frangi, A.: A statistical shape model of the heart and its application to model-based segmentation. In: Proc. SPIE Med. Imaging, vol. 6511 (2007)
4. Perperidis, D., Mohiaddin, R., Rueckert, D.: Construction of a 4D statistical atlas of the cardiac anatomy and its use in classification. In: Med. Image Comput. Assist. Interv. Int., vol. 8, pp. 402–410 (2005)
5. Hoogendoorn, C., Sukno, F., Ordás, S., Frangi, A.: Bilinear models for spatio-temporal point distribution analysis. Int. J. of Comput Vision 85(3), 237–252 (2009)
6. Cootes, T., Cooper, D., Taylor, C., Graham, J.: Active shape models - their training and application. Comput. Vision and Image Understanding 61(1), 38–59 (1995)
7. Kirisli, H., et al.: Fully automatic cardiac segmentation from 3D CTA data: a multiatlas based approach. In: Proc. SPIE Med. Imaging (2010)
8. Klein, S., Staring, M., et al.: `elastix`: a toolbox for intensity based medical image registration. IEEE T. Med. Imaging 29(1), 195–205 (2010)
9. Comaniciu, D., Meer, P.: Mean Shift: A robust approach toward feature space analysis. IEEE T. Pattern Anal. 24(5), 603–619 (2002)
10. de Bruijne, M., Lund, M., Tankó, L., Pettersen, P., Nielsen, M.: Quantitative vertebral morphometry using neighbor-conditional shape models. Med. Image Anal. 11(5), 503–512 (2007)