OPEN-SOURCE 4D STATISTICAL SHAPE MODEL OF THE HEART FOR X-RAY PROJECTION IMAGING

Mathias Unberath[⋆†‡], *Andreas Maier*[⋆‡], *Dominik Fleischmann*[¶], *Joachim Hornegger*[⋆‡], *Rebecca Fahrig*[†]

[⋆]Pattern Recognition Lab, FAU Erlangen-Nürnberg
[†]Radiological Sciences Laboratory, Stanford University
[‡]Graduate School in Advanced Optical Technologies, FAU Erlangen-Nürnberg
[¶]Diagnostic Radiology, Stanford University

ABSTRACT

We present an open-source 4D (3D+t) statistical shape model of the heart developed as numerical phantom for cone-beam CT simulation. The training set consists of surface meshes from 20 ten-phase-CT angiography data sets extracted using automatic registration-based segmentation. Incorporating 90% of the training set variation, the model exhibits a generalization ability of $5.00 \pm 0.93\,mm$ and specificity of $7.30 \pm 0.97\,mm$. The model is implemented in CONRAD, an open-source simulation and reconstruction framework. We provide all algorithms and high resolution projection data on the project's homepage. To the best of our knowledge, this is the first open-source 4D statistical shape model of the heart.

***Index Terms*—** Simulation, numerical phantom, XCAT, cone-beam, CT, open-source

1. INTRODUCTION

The long acquisition times of about 5 s in rotational angiography using C-arm systems requires the incorporation of the hearts dynamics into the reconstruction algorithms [1, 2]. Meaningful evaluation and comparison of such algorithms requires projection data from computerized 4D phantoms as ground truth is typically unavailable in clinical cases.
The 4D extended cardiac-torso (XCAT) phantom is the most frequently used numerical cardiac phantom up to date. The underlying 3D model is based on the Visible Human male and female anatomical data sets [3], extended to 4D using deformations obtained from one tagged cardiac MRI and one respiratory-gated CT data set from normal male patient volunteers [3]. As a consequence its clinical relevance suffers when changing clinical parameters, such as the ejection fraction (EF), or when modeling pathologies.
In order to tackle this problem, we develop an open-source 4D statistical shape model (SSM) of the whole heart for the simulation of cone-beam CT angiography.

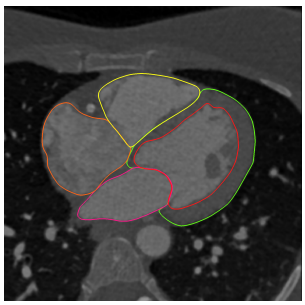
2. METHODS

2.1. Training set generation

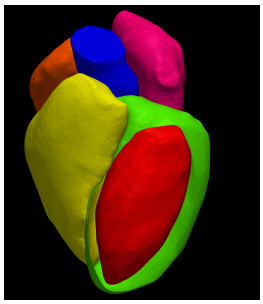
We use a registration-based segmentation approach to automatic whole heart segmentation in order to preserve point correspondence among samples [4]. Surface meshes of the aorta, atria, ventricles, and left-ventricular myocardium of an atlas segmentation are propagated to unseen images using transforms obtained by registration. The atlas segmentation was performed on a retrospectively ECG-gated, four chamber enhanced CT angiography data set of a 45 year old female at best diastolic phase (78%). The image volume had $512 \times 512 \times 241$ pixels with a spacing of $0.29 \times 0.29 \times 0.5\,mm$. Images are then rigidly and non-rigidly registered to the atlas in order to obtain mappings for the surface meshes, following a strategy similar to Rueckert *et al.* [5]. The complete segmentation algorithm consisted of the following steps:

1. Rigid registration using a similarity transform and mutual information [6]
2. Multi-resolution non-rigid registration on four levels using B-spline transforms and mutual information [7]
3. Propagation of atlas surface meshes

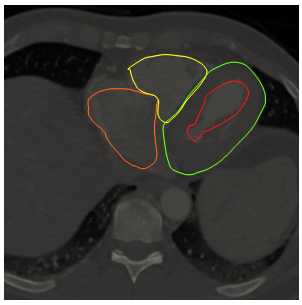
An axial view and surface rendering of the atlas segmentation together with representative example segmentations of a male patient's data set at end-systole, and a female patient's data set at end-diastole are shown in Fig. 1. Using the strategy described above we segmented 20 ten phase, four-chamber enhanced CT angiography data sets from nine male (59.56±25.10 years old) and eleven female patients (59.56±25.10 years old) with EFs ranging from 28.65% to 66.12% (52.13±9.11%). An expert study not described here was conducted to assess segmentation accuracy of 20 representative volumes. Samples not included in the study were inspected visually to ensure plausible results.



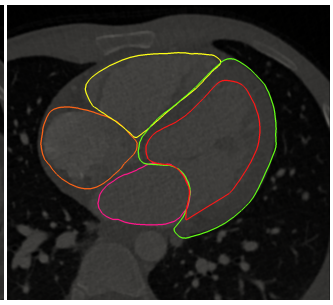
(a) Atlas: axial view



(b) Alas: surface rendering



(c) Example 1: axial view



(d) Example 2: axial view

Fig. 1. Axial view of the atlas segmentation in (a), and corresponding surface rendering in (b). Axial views of representative segmentations of a male patient's data set at end-systole, and a female patient's data set at end-diastole are shown in (c) and (d), respectively.

2.2. 3D model building

We build a 3D SSM describing inter-subject variation at each of the ten cardiac phases. First the m training shapes $\boldsymbol{x}_i^{(p)}$, $i = 1, \cdots, m$, $p = 1, \cdots, P$ at cardiac phase p are aligned in a common reference coordinate system using Generalized Procrustes Analysis, an iterative procedure consisting of translating, rotating, and scaling of each input shape to the iteratively updated sample consensus using orthogonal Procrustes analysis [8]. We rescale the shapes after alignment as heart size is part of anatomical variation.
Subsequently we extract the $c^{(p)}$ principal modes of variation among the training samples by solving the Eigenvalue equation [9]:

$$S^{(p)}\Phi_k^{(p)} = \lambda_k^{(p)}\Phi_k^{(p)}, \tag{1}$$

where $S^{(p)} = \frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{x}_i^{(p)} - \bar{\boldsymbol{x}}^{(p)}\right)\left(\boldsymbol{x}_i^{(p)} - \bar{\boldsymbol{x}}^{(p)}\right)^T$ is the covariance matrix and $\bar{\boldsymbol{x}}^{(p)} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{x}_i^{(p)}$ is the consensus. The amount of modes is determined such that the cumulative variance $\left(\sum_{i=1}^{c^{(p)}}\lambda_i^{(p)}\right) / \left(\sum_{j=1}^{m}\lambda_j^{(p)}\right)$ exceeds 90% of training set variation. Using the resulting SSM $\{\boldsymbol{x}^{(p)}, \Phi^{(p)}\}$ shapes can be expressed as a weighted sum:

$$\boldsymbol{y} = \bar{\boldsymbol{x}}^{(p)} + \Phi^{(p)}\boldsymbol{\beta}^{(p)} + \varepsilon, \tag{2}$$

where $\Phi^{(p)} = \left(\Phi_1^{(p)}, \cdots, \Phi_{c^{(p)}}^{(p)}\right)$, $\boldsymbol{\beta} \in \mathbf{R}^{c^{(p)}}$ are weights, and ε is variation not explained by the model. Projecting shape $\boldsymbol{x}_i^{(p)}$ onto the Eigenmodes Φ_k yields its principal components:

$$\boldsymbol{\beta}_i^{(p)} = \Phi^{(p)\,T}\left(\boldsymbol{x}_i^{(p)} - \bar{\boldsymbol{x}}^{(p)}\right). \tag{3}$$

2.3. Dynamic model

In order to describe valid contractive behavior we concentrate the principal components $\boldsymbol{\beta}_i^{(p)}$ of shape i at all phases p into one dynamic shape vector $\boldsymbol{\beta}_i = \left(\boldsymbol{\beta}_i^{(1),T}, \cdots, \boldsymbol{\beta}_i^{(P),T}\right)$. We then perform principal component analysis as described in Eq. 1 on the dynamic shape vectors $\boldsymbol{\beta}_i$ yielding a SSM of principal components $\{\bar{\boldsymbol{\beta}}, \boldsymbol{\rho}\}$ with Eigenvalues μ. This procedure is related to the method proposed by Bosch *et al.*. They concentrated shape sequences into one dynamic shape and performed principal component analysis [10]. The use of such strategies allows for a compact interface for dynamic shape generation as a whole cardiac cycle $\boldsymbol{y}$ is represented by one principal component vector $\boldsymbol{\delta}$:

$$\begin{aligned}\left(\boldsymbol{y}^{(1)}, \cdots, \boldsymbol{y}^{(P)}\right) &= \left(\bar{\boldsymbol{x}}^{(1),T}, \cdots, \bar{\boldsymbol{x}}^{(P),T}\right) \\ &+ \left(\Phi^{(1),T}, \cdots, \Phi^{(P),T}\right)\left(\bar{\boldsymbol{\beta}} + \boldsymbol{\rho}\,\boldsymbol{\delta}\right).\end{aligned} \tag{4}$$

Linear interpolation is used to allow for continuous sampling of the cardiac cycle.

2.4. Evaluation

We investigated compactness, generalization ability and specificity of the SSM. Generalization ability is the capability of a SSM to represent unseen instances of the object class and is assessed using leave-one-out tests. We evaluated the generalization ability for models containing 90% (95%) of the variation present in the training set. We built SSMs for each cardiac phase with one training shape excluded and calculated the error $\|\varepsilon_i^{(p)}\|_2^2$ for each phase and training set using Eq. 2 and 3. Model specificity measures the validity of newly generated instances. Specificity of the static (dynamic) model was calculated by randomly generating H new instances using Eq.2, where the weights $\beta_k^{(p,h)} \in \left[-\lambda_k^{(p)}/2, \lambda_k^{(p)}/2\right]$ $\left(\boldsymbol{\delta}_b^{(h)} \in [-\mu_b/2, \mu_b/2]\right)$, $h = 1, \cdots, H$ were determined randomly. The generated test samples were then compared to all samples in the training set. Model specificity can be defined as the average distance of H uniformly distributed, randomly generated shape instances to their nearest member among all training shapes. We used $H = 1000$ ($H = 100$) test samples to compute the specificity of the models. The model is intended as numerical phantom for X-ray projection imaging. We generated forward projections of the model with a contrasted left ventricle using a ray-casting forward projector implemented in CONRAD [11]. The detector had

620×480 pixels with an isotropic spacing of $0.62\,mm$. The source to detector distance was $1198\,mm$ and the detector to isocenter distance was $780\,mm$.

3. RESULTS

The model was evaluated using the strategies explained in Section 2.4. In order to incorporate 90% (95%) of training set variation the 16 (18) largest principal modes of variation were needed in all static models $\{\bar{\boldsymbol{x}}^{(p)}, \Phi^{(p)}\}$, and the dynamic model $\{\bar{\boldsymbol{\beta}}, \boldsymbol{\rho}\}$. The assessment of generalization ability yielded $5.00 \pm 0.93\,mm$ $(4.89 \pm 0.90\,mm)$. Using the 16 largest modes of variation resulted in a specificity of $7.18 \pm 0.45\,mm$ $(7.30 \pm 0.97\,mm)$. Shape changes at end-diastole induced by variation of the first two principal components of the dynamic model are shown in Fig. 2. Volume

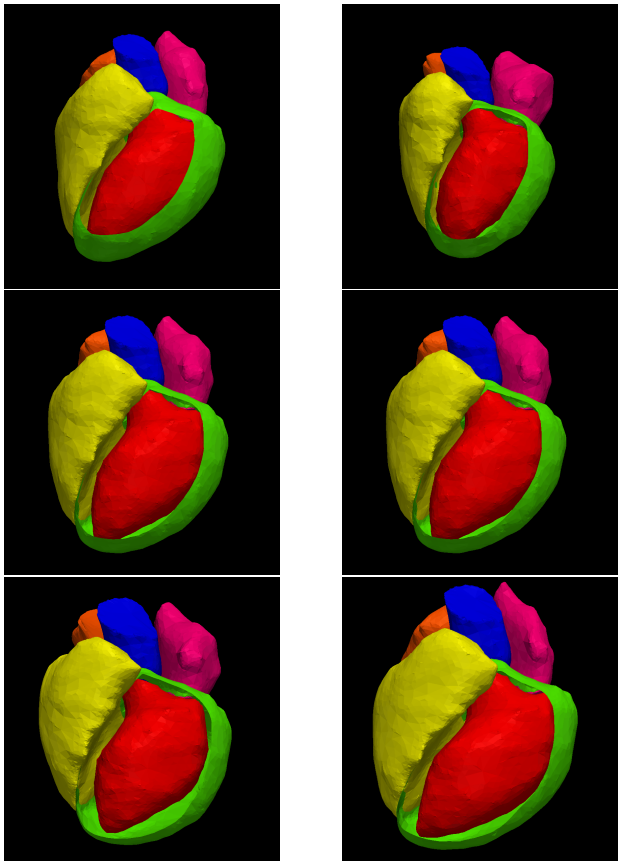
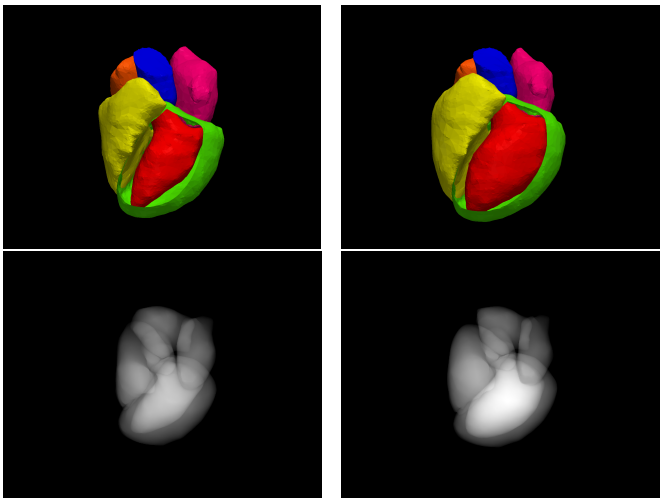


Fig. 2. Shape changes due to variation of the two largest principal components of the dynamic model at 0% phase in decreasing magnitude of variance μ from left to right. The weight $\delta_b = -\mu_b/2$ for the top row, $\delta_b = \mu_b/2$ for the bottom row, and $\delta_b = 0$ in the middle.

renderings of the model at representative phases during the cardiac cycle together with the corresponding X-ray projection image are shown in Fig. 3. The model's source code as well as high resolution forward projections of dynamic heart



(a) (b)

Fig. 3. Volume renderings and forward projections of the mean shape of the dynamic SSM at end-systole in (a), and end-diastole in (b).

Gender:	f	f	m	m
Age:	52	81	56	62
EF /%:	54	62	56	46

Table 1. Predefined model parameters.

scenes simulated with real trajectories are available on the project's homepage: *conrad.stanford.edu*.

4. DISCUSSION AND CONCLUSION

We constructed SSMs of the cardiac anatomy (descending aorta, four chambers, and left ventricular myocardium) at ten phases during the cardiac cycle. The training set consisted of 20 ten phase, four-chamber enhanced CT angiography data sets of nine male and eleven female patients between 23 and 92 years with an average EF of $52.13 \pm 9.11\,\%$. A dynamic model was obtained by concentrating the principal components of each training shape at all phases into one vector and performing principal component analysis on those component vectors. This procedure allowed for a compact representation of cardiac dynamics while guaranteeing valid behavior.

Assessment of the model's compactness showed that the 16 (18) largest modes were sufficient to explain 90% (95%) of training set variation. This corresponds to a dimensionality reduction of only 20% (10%) indicating large variations among the training samples. Large variation is desired in order to express as much diversity as possible, however, it can be problematic if the landmarks do not follow a normal distribution, as the estimation of the allowable shape domain may become unobvious [9].

The computation of generalization ability yielded an average vertex offset of $5.00 \pm 0.93\,mm$ for the model containing

90% of the variation present in the training set. Considering the large dispersion between individual training samples suggested by the compactness of the model generalization ability was found to be rather low. This result may indicate that the training shapes, while being largely diverse, satisfactorily represent the target population.

Specificity was similar for both the static and dynamic cardiac models, yielding numeric values of $7.18 \pm 0.45\,mm$ and $7.30 \pm 0.97\,mm$, respectively. Specificity was significantly worse than generalization ability (c.f. $5.00 \pm 0.93\,mm$, $p = 0.013$). A possible explanation can be found when investigating the testing procedure. Weights for the uniformly distributed, randomly generated test samples were confined to an interval of $\pm\lambda/2$ in all tests, allowing for a broad spectrum of instances. Due to the low amount of training shapes, the space of valid instances is only sparsely populated, leading to an increased expectation value of average distance to the nearest member of the training set and, consequently, specificity. Therefore, verification of newly generated instances remains difficult with the current training set. Invalid shapes and dynamics, however, may lead to false conclusions when using the model as numeric phantom. We therefore provide predefined shape vectors extracted from the training set (see Table 1) as well as projection data. We expect improvements in all quantitative measures for models built from a larger training set, strongly suggesting the acquisition of additional data.

Recently proposed statistical shape models of the heart employed bilinear models to incorporate inter-subject and temporal variation into a single model [12]. This allows for independent tuning of anatomical and temporal variation at the cost of larger parameter vectors. We used statistical analysis of the principal component vectors to provide an interface for dynamic shape generation using few parameters. Our approach assumes that equal anatomies at rest exhibit similar contractive behavior. However, inter-subject and temporal variation neither is independent nor is it deterministic. The commitment to a certain method is a trade-off between model flexibility and ease of use. In the present case the advantages of simplified shape generation prevailed.

The algorithms used for model construction are freely available and provide an open-source basis for statistical shape model generation of arbitrary shapes.

5. REFERENCES

[1] C. Schwemmer, C. Rohkohl, G. Lauritsch, K. Müller, and J. Hornegger, "Residual Motion Compensation in ECG-Gated Interventional Cardiac Vasculature Reconstruction," *Physics in Medicine and Biology*, vol. 58, no. 11, pp. 3717–3737, 2013.

[2] K. Müller, A. Maier, Y. Zheng, Y. Wang, G. Lauritsch, C. Schwemmer, C. Rohkohl, J. Hornegger, and R. Fahrig, "Interventional heart wall motion analysis with cardiac C-arm CT systems," *Physics in Medicine and Biology*, vol. 59, no. 9, pp. 2265–2294, 2014.

[3] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Medical Physics*, vol. 37, no. 9, pp. 4902–4915, 2010.

[4] M.l Spiegel, D. Hahn, V. Daum, J. Wasza, and J. Hornegger, "Segmentation of kidneys using a new active shape model generation technique based on non-rigid image registration," *Computerized Medical Imaging and Graphics*, vol. 33, pp. 29–39, 2009.

[5] S. Ordas, E. Oubel, R. Leta, F. Carreras, and A. F. Frangi, "A statistical shape model of the heart and its application to model-based segmentation," in *Medical Imaging*. International Society for Optics and Photonics, 2007, pp. 65111K–65111K.

[6] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellyn, and W. Eubank, "Nonrigid multimodality image registration," in *Medical Imaging*. International Society for Optics and Photonics, 2001, pp. 1609–1620.

[7] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Non-rigid registration using free-form deformations: application to breast MR images," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 8, pp. 712–721, 1999.

[8] C. Goodall, "Procrustes methods in the statistical analysis of shape," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 285–339, 1991.

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[10] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance motion models," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 11, pp. 1374–1383, 2002.

[11] A. Maier, H. G. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, et al., "Conrad – a software framework for cone-beam imaging in radiology," *Medical Physics*, vol. 40, no. 11, pp. 111914, 2013.

[12] C. Hoogendoorn, N. Duchateau, D. Sanchez-Quintana, T. Whitmarsh, F. M. Sukno, M. De Craene, K. Lekadir, and A. F. Frangi, "A high-resolution atlas and statistical model of the human heart from multislice CT," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 1, pp. 28–44, 2013.