Bilinear Models for Spatio-Temporal Point Distribution Analysis: Application to Extrapolation of Whole Heart Cardiac Dynamics

Corné Hoogendoorn, Federico M. Sukno, Sebastián Ordás and Alejandro F. Frangi
Computational Imaging Lab
Information and Communication Technologies Department
Pompeu Fabra University
Passeig de Circumval·lació 8, E08003 Barcelona, Spain
http://www.cilab.upf.edu

Networking Center on Biomedical Research - CIBER-BBN, Barcelona, Spain
http://www.ciber-bbn.es

Abstract

In this work we introduce the usage of bilinear models as a means of factorising the shape variation induced by subject variability and the contraction of the human heart. We show that it is feasible to reconstruct the shape of the heart at a certain point in the cardiac cycle if we are given a small number of shapes representing the same heart at different points in the same cycle, using the bilinear model. Depending on pathology and the ratios between healthy and pathological hearts in the training set, RMS reconstruction errors measured between 1.39 and 16.58 millimetres, with a median of 6.79 and 90th percentile of 9.95 millimetres.

1. Introduction

The use of statistical models of shape has established itself as a popular approach to image analysis problems, in the domain of both natural and medical image analysis. Along the way, much research has been devoted to the development of various types of shape models, as well as to solving problems arising from the construction of such models. Many of those models, however, are essentially spatial models, and extension to the spatio-temporal domain is not as trivial or trivially justifiable as the extension from d to $d + 1$ spatial dimensions.

The analysis of Point Distribution Models (PDM's) in shape space received a significant amount of attention from the mid-1980's to the early 1990's [6, 14, 15, 18]. The most renowned result from this work is the emergence of the Principal Components Analysis (PCA)-based statistical shape model from Cootes *et al*. [7], who apply PCA to the covariance matrix of their data set in order to extract a set of orthogonal variations of the sampled points, which became well-known as the modes of variation of the shape class.

Attempts to extend linear shape models to the spatio-temporal domain have been made before, for example by Hamarneh and Gustavsson [12], in whose work each sample used for PCA consists of an entire sequence of observations of the same object sampled throughout the temporal exposure window. Mitchell *et al*. [20] and Bosch *et al*. [3] employ the Active Appearance Model analogue of this type to segment the endocardium in echocardiograms and MR sequences, respectively. Another approach to add the element of time to 3-dimensional segmentation using statistical shape models was presented by Montagnat and Delingette [21]. After building a PCA-based model treating all subjects and all phases as separate samples, they employ a scheme which segments the cycle as one single object, rather than treating each frame as a separate entity. During segmentation, temporal constraints are introduced in the optimisation to limit the differences between segmentations of subsequent frames to reasonable values.

Statistical models of cardiac left ventricular deformation only were constructed by Chandrashekara *et al*. [4], using data from one single subject. A comparable approach to modelling respiratory motion of the liver was taken by Blackall *et al*. [2].

To the best of our knowledge, this is the first time that dynamic shapes are statistically decomposed by acknowledging the fact that individual and temporal variations are different sources of variability. Before, either time was treated as if it does behave in the same way as space [3, 20, 12], or the application of the model was equipped with constraints to limit the first-order derivative of shape points over time [21]. By creating a spatiotemporal model of cardiac dy-

978-1-4244-1631-8/07/$25.00 ©2007 IEEE

namics that decouples individual and temporal variations, we can extrapolate cardiac phases from the statistical model even when they are not available from the individual measurements.

In biometrics, the separation of two (independent) processes that account for variability is a well-known problem which has led to the introduction of bilinear models by Tenenbaum and Freeman in [24], which are discussed more thoroughly in [25]. Dubbing the two sources of variability *style* and *content*, these names can be assigned freely depending on which is most natural given a specific problem. Aside from the examples used in [25], the literature contains examples of the separation of

- face identity and facial expression [1]
- location and content for sparse coding of natural images [11]
- emotion and speech content [5]
- gait (walking characteristics) and viewing conditions [13, 16]
- pairs from the set {identity, action, viewpoint} [8].

The remainder of this paper is organised as follows: we introduce the concept of bilinear models in Section 2, and build bilinear models which factor out subject and temporal variability of the shape of the human heart. Retraining the model with incomplete cardiac cycles is done in Section 3, after which the shapes from the remaining cardiac phases are extrapolated. The extrapolation results are presented in Section 4, followed by a discussion and future research directions in Section 5. We conclude this work with Section 6.

2. Bilinear Models

A bilinear model is a two-factor model which is linear in either factor when the other one is kept constant:

$$\mathbf{y} = \mathbf{aWb} \tag{1}$$

where $\mathbf{y}$ is the observation, $\mathbf{a}$ and $\mathbf{b}$ are parameterisation vectors defined by the factors, and $\mathbf{W}$ is a constant matrix governing the interaction between the factors.

Each element y_k^{sc} of a K-sized observation $\mathbf{y}^{sc}$ can be described by a bilinear model as

$$y_k^{sc} = \sum_{i=1}^{I}\sum_{j=1}^{J} w_{ijk} a_i^s b_j^c. \tag{2}$$

$\mathbf{W}$ is a matrix which forms a mapping from style and content space into observation space. Each $\mathbf{w}_{ij}$ is a K-sized base observation, much akin to the eigenface [26] and eigenshape [7]. In the case of sound recognition, one could think of them as voice harmonics that need to be combined to form a certain sound in a certain accent. The $\mathbf{a}$ and $\mathbf{b}$ vectors provide the information on how to combine those base observations. For the time being, we will adhere to the original nomenclature and call these the style and content vectors, respectively (hence the s and c superscripts).

The choice of bilinear models stems from the idea that variations in a set of observations are the consequence of the variation of two independent factors. The examples used to illustrate the usefulness of bilinear models in [25] call upon analysing the way we manage to recognise known characters, people or phonemes in a font or under viewing circumstances not observed before, or uttered in an accent not heard before. Somehow, we know the invariants of that character, person, or phoneme, and in an observation we can recognise those irrespective of the variations induced by the circumstances.

2.1. Asymmetric Models

A combining matrix $\mathbf{A}$ can be the result of contracting $\mathbf{aW}$ into a single matrix, leading to an asymmetric model:

$$y_k^{sc} = \sum_{j=1}^{J} a_{jk}^s b_j^c \tag{3}$$

where

$$a_{jk}^s = \sum_{i=1}^{I} w_{ijk}^s a_i^s. \tag{4}$$

This is useful when the mapping from content to observation is dependent on style. The J observations of size K each that make up $\mathbf{A}$ are then style-specific base observations that can be mixed using the parameterisation in $\mathbf{b}$. In essence, the model has then become a unilinear model. It is possible to invert the roles of style and content, and build a different asymmetric model from the same training data.

2.2. Symmetric Models

The symmetric model is the original bilinear model as already presented in (2):

$$y_k^{sc} = \sum_{i=1}^{I}\sum_{j=1}^{J} w_{ijk} a_i^s b_j^c,$$

where the mapping from style and content space to observation space, $\mathbf{W}$, is dependent on neither style nor content. The elements of $\mathbf{W}$ are base observations that look like eigenfaces [26] but do not represent an orthogonal basis like eigenelements. The base observations can then be mixed using the $\mathbf{a}$ and $\mathbf{b}$ parameterisations to form any element from the training set, and other parameterisations can be used to construct observations of new style and/or new content.

2.3. Factoring Individual and Temporal Variability

2.3.1 Representation of Training Data

The construction of the bilinear models as presented in [25] assumes a vector representation of the observations. This can easily be achieved by employing a surface representation as is commonly used for the construction of PDM's like those in [7]. In d-dimensional Euclidean space, each training shape is annotated with n points – landmarks – of anatomical correspondence throughout the training set. The set of point coordinates is then concatenated to form an (nd)-dimensional shape vector. This equals a single point in an (nd)-dimensional shape space.

After annotation and vectorisation of our observations, we have IJ (nd)-dimensional shape vectors: I frames per time sequence, J subjects (with one sequence each), n landmarks in d dimensions. As is the case in [12], we construct our observation matrix by 'stacking' all frames for one object onto each other, such that we obtain a $Ind \times J$ shape matrix. This is the starting point for the construction of both the asymmetric and the symmetric bilinear shape models.

Next up is the modeling, which consists of minimising the squared reconstruction error between the original observations and the approximation the model will provide. Denoting the approximation of $\mathbf{y}^{sc}$ as

$$\hat{\mathbf{y}}^{sc} = \begin{cases} \mathbf{A}^s\mathbf{b}^c & \text{if the model is asymmetric} \\ \mathbf{a}^{s^T}\mathbf{W}\mathbf{b}^c & \text{if the model is symmetric} \end{cases}, \quad (5)$$

we minimise

$$E = \sum_{s=1}^{S}\sum_{c=1}^{C} \|\mathbf{y}^{sc} - \hat{\mathbf{y}}^{sc}\|^2. \quad (6)$$

2.3.2 Asymmetric Training

The training of an asymmetric model has a closed-form solution if the number of observations is (nearly) equally distributed over the style and content classes [25]. As the matrix with training data $\mathbf{Y}$ is the result of the product $\mathbf{AB}$, it suffices to compute the Singular Value Decomposition (SVD) $\mathbf{Y} = \mathbf{USV}^T$. Then, $\mathbf{B}$ can be defined as the first J rows of $\mathbf{V}^T$, while $\mathbf{A}$ will be defined as the first J columns of $\mathbf{US}$.

2.3.3 Symmetric Training

The training of the symmetric model requires the notion of the vector transpose, which is explained in Appendix A. This time, an iterative method is required to minimise E. To this end, first $\mathbf{A}$ and $\mathbf{B}$ are computed. Upon convergence of those computations, $\mathbf{W}$ is computed using the results.

Analogously to the case of $K = 1$, the simplified model equation $\mathbf{Y} = \mathbf{AWB}$ can be rewritten as $\mathbf{Y} = \left(\mathbf{W}^{\mathrm{VT}}\mathbf{A}\right)^{\mathrm{VT}}\mathbf{B}$ as well as $\mathbf{Y}^{\mathrm{VT}} = \left(\mathbf{WB}\right)^{\mathrm{VT}}\mathbf{A}$. This leads to two equations that are familiar from the training of the asymmetric model:

$$\left(\mathbf{YB}^{\mathrm{T}}\right)^{\mathrm{VT}} = \mathbf{W}^{\mathrm{VT}}\mathbf{A} = \mathbf{USV}^{\mathrm{T}} \quad (7)$$

and

$$\left(\mathbf{Y}^{\mathrm{VT}}\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{VT}} = \mathbf{WB} = \mathbf{USV}^{\mathrm{T}}. \quad (8)$$

By iterating over these equations, starting from an initial estimate of $\mathbf{B}$ using the SVD of $\mathbf{Y}$, convergence towards the real $\mathbf{A}$ and $\mathbf{B}$ is guaranteed [17]. Upon convergence of the computation of $\mathbf{A}$ and $\mathbf{B}$, what is left is the computation of $\mathbf{W}$:

$$\mathbf{W} = \left(\left(\mathbf{YB}^{\mathrm{T}}\right)^{\mathrm{VT}}\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{VT}}. \quad (9)$$

3. Experiments

We performed experiments to verify the ability of bilinear modelling to capture the dynamics of the shape of the beating heart. To this end we constructed models factorising the temporal and inter-subject variations. Then, the models were used for extrapolating the learnt dynamics to subjects not present in the training set.

3.1. Data

Our test data consists of 45 full hearts acquired consecutively, with closed surface representations for each of five subparts: left and right ventricular myocardium, left and right atrial myocardium, and the fibrous skeleton. The population is distributed as 60% healthy and asymptomatic subjects, 20% subjects with Coronary Artery Disease (CAD) without a history of Myocardial Infarction (MI), and 20% subjects with CAD and a history of MI. The average age of the subjects was 58 years with a standard deviation of 8 years. Of all subjects, 56% were men.

The imaging was performed using a Toshiba Aquilion 64 multi-slice CT system (Toshiba Medical Systems, Japan) with 64 detector rows. Between 80 and 100 ml of contrast material (Xenetic 350) was administered at an injection rate of 5 ml/s. The rotation time of the scanner was, depending on the subject's heart rate, between 400 and 500 ms, and image reconstruction was performed on a Vitrea post-processing workstation (Vital Images, USA). The resulting dataset consisted of 15 image volumes (temporal phases) obtained using retrospective ECG gating [22] with voxel dimensions of $0.4 \times 0.4 \times 2.0mm^3$ per subject. Figure 2 shows the gating sequence over the cardiac cycle, albeit that the ECG used is schematic. Temporal relationships are only preserved with respect to the R-R interval. Overlaid on this sequence are the Root Mean Square (RMS) differences between subsequent shapes, using an annotation scheme as explained hereafter. The value after the 15th phase reflects the RMS difference between phases 15 and 1.

The data set was then used for the annotation as presented in [9], resulting in 26230 points per volume. The point set sampled on the average shape was triangulated using Amira 3.0 (Mercury Computer Systems, USA) to facilitate visualisation of this shape, all shapes in the training set, and all shapes generated using the bilinear model.

With n_{L} landmarks describing our shapes, we denote the point set in $\mathbb{R}^3$ equivalent to shape vector $\mathbf{y}$ as the $n_{\mathrm{L}}\times 3$ matrix $\mathbf{X}$. The first phase $\mathbf{X}^{i1}_{\mathrm{original}}$ of each cycle i was translated to position its center of gravity at the origin. Then, it was rescaled to have unit norm. Subsequently, Procrustes alignment [10] was performed on these shapes, which, for each subject i, yielded a translation $\mathbf{t}_i$, rotation $\mathbf{R}_i$ and isotropic scaling s_i. Then, for each phase j,

$$\mathbf{X}^{ij}_{\mathrm{aligned}} = s_i \times \mathbf{X}^{ij}_{\mathrm{original}}\mathbf{R}_i + \mathbf{t}_i. \tag{10}$$

This way, the spatial relationships within the cycles were preserved.

3.2. Experiment pipeline

We performed 2-fold, 3-fold, 5-fold and 9-fold experiments, splitting our data set into subsets of 22, 15, 9 and 5 subjects, respectively. Models were then built using all but one of the subsets, after which the remaining subset was used as test data. For each of the subjects from the test subset, five consecutive phases were used to derive the subject parameterisation using the constructed model. The remaining ten phases were then approximated and the reconstruction errors recorded. Errors were grouped by phase and model dimensionality, and by phase set.

3.3. Extrapolation

What we wish to do amounts to extrapolation, which is only one of several applications of bilinear models, as is shown in [25]. For an incomplete set of data in a new style $\mathbf{Y}_{\{\mathbf{i_1}\ldots\mathbf{i_n}\}}, n < I \wedge \forall j,k : i_j = i_k \Leftrightarrow j = k$, where I is the number of content classes, the missing elements need to be reconstructed using the characteristics of those contents learned from the training data (other styles), and the style characteristics that are derived from the newly presented observations.

The models we built were symmetric models, but we used them as if they were asymmetric. Knowing which are the phases we have for deriving the subject parameterisations, we used the phase parameterisations from the model associated with these:

$$\mathbf{C}_1^{\mathrm{VT}} = \mathbf{W}^{\mathrm{VT}}\mathbf{A}^{\mathrm{T}}_{\{\mathbf{i_1}\ldots\mathbf{i_n}\}}. \tag{11}$$

What remains is a linear system of equations

$$\mathbf{Y}_{\{\mathbf{i_1}\ldots\mathbf{i_n}\}} = \mathbf{C}_1\mathbf{b}_{\mathrm{new}} \tag{12}$$

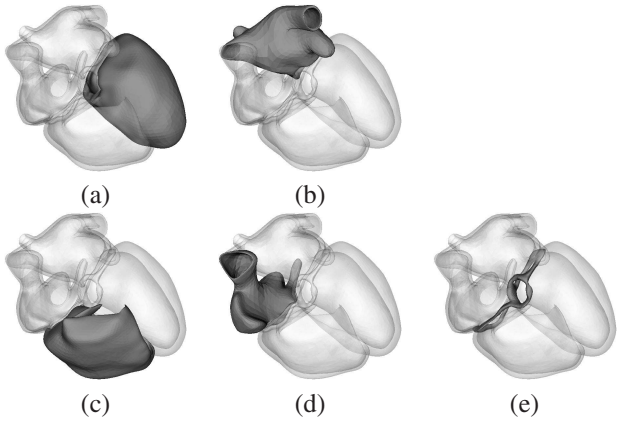


Figure 1. Surface rendering of one of the training hearts. In dark (a) the left ventricle, (b) left atrium, (c) right ventricle, (d) right atrium, and (e) fibrous skeleton.

which, when solved, gives us a parameterisation $\mathbf{b}_{\mathrm{new}}$ of the new subject. This system has many more equations than unknowns (the ratio running into the thousands). Therefore, a least squares approximation is found.

3.4. Reconstruction

Using the resulting $\mathbf{b}_{\mathrm{new}}$ from (12), we approximate the remaining set of phases $\mathbf{Y}_{\{\mathbf{i_{n+1}}\ldots\mathbf{i_I}\}}, \forall j,k : i_j = i_k \Leftrightarrow j = k \wedge \{i_1 \ldots i_n\} \bigcap \{i_{n+1} \ldots i_I\} = \emptyset$ by

$$\mathbf{C}_2^{\mathrm{VT}} = \mathbf{W}^{\mathrm{VT}}\mathbf{A}^{\mathrm{T}}_{\{\mathbf{i_{n+1}}\ldots\mathbf{i_I}\}}, \tag{13}$$

$$\hat{\mathbf{Y}}_{\{\mathbf{i_{n+1}}\ldots\mathbf{i_I}\}} = \mathbf{C}_2\mathbf{b}_{\mathrm{new}}. \tag{14}$$

The reconstruction error is then recorded as the RMS error between $\mathbf{Y}_{\{\mathbf{i_{n+1}}\ldots\mathbf{i_I}\}}$ and $\hat{\mathbf{Y}}_{\{\mathbf{i_{n+1}}\ldots\mathbf{i_I}\}}$ for each of the approximations $\hat{\mathbf{y}}$, which are columns of $\hat{\mathbf{Y}}_{\{\mathbf{i_{n+1}}\ldots\mathbf{i_I}\}}$:

$$\mathrm{RMS}(\hat{\mathbf{y}}) = \sqrt{\frac{\sum_i (\hat{\mathbf{y}}_\mathbf{i} - \mathbf{y}_\mathbf{i})^2}{n_{\mathrm{L}}}}, \tag{15}$$

where n_{L} again is the number of landmarks used to describe each of our shapes. Thus, we deviate slightly from the RMS concept by not averaging over the observation vector size, but over only a third of that. Thus, our results are $\sqrt{3}$ times that of the standard formulation of the RMS, but they will provide us with a better idea of the error in 3-space.

The results are presented for three disjoint sets of phases ($\{1\ldots5\}$, $\{6\ldots10\}$ and $\{11\ldots15\}$) used for retraining (the derivation of a new subject's parameterisation) and thus for the reconstructed groups $\{6\ldots15\}$, $\{1\ldots5, 11\ldots15\}$ and $\{1\ldots10\}$, respectively.

4. Results

In Figure 3, we present the mean RMS error of the worst and best reconstructed phases, as well as the mean over

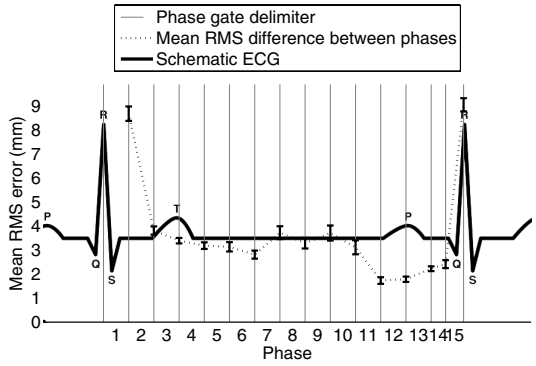


Figure 2. The intervals used in the retrospective ECG gating overlaid on a standard schematic ECG. The plot is the RMS difference between two subsequent shapes, averaged over our 45 subjects.

all phases, against the size of the training sets (and consequently against the number of folds). It can be noted that larger training set sizes produced more accurate reconstructions. Additionally, the tendency of the curves suggests that significant performance improvement could be achieved by further increasing the training set size.

From the same figure, it is possible to conclude that, in general, the worst and best reconstructed phases retain that status irrespective of the training set size. For clarity, the plots of all phases lying in between were removed, but the dramatic transition of phase 8 from worst phase to comparable to average that is observable in Figure 3 (a) is an exception. Thus, the quality of the reconstruction depends to some extent on the reconstructed phase. It also appears that the phases used for retraining influence the success of the reconstruction. This would be clearer had the plots been positioned next to each other, but the reconstructions after retraining with phases 6 through 10 are somewhat better than those using the other two retraining sets.

The fourth factor that has a certain degree of influence is which shape is being reconstructed, both regarding phase, as we pointed out at the hand of Figure 3, and subject. Figure 4 shows the mean approximation error for each subject in the 9-fold experiment. The order of the subjects in the plot follows that of the folds. It is obvious that the shapes of certain subjects turn out to be reconstructed poorly at every phase, regardless of the phase set used for deriving the subject parameters. Thus, the model did not generalise well to these subjects.

Figure 6 shows the RMS errors against the percentage of reconstructions. With the bell curve overlaid constructed using mean and standard deviation of the RMS errors, it gives an idea of the distribution of the error sizes. The median error is 6.79 millimetres, while of all approximations, 90% have errors below 10 millimetres.

Finally, Figure 5 shows the local errors of the reconstructions. The errors are color-coded on the surface representation of the reconstruction, while the wireframe mesh shows the shape that was supposed to be reconstructed. It should be noted that mainly the most complex parts of the heart shape, namely the atrial and valvular sections, appear to be the most difficult to approximate accurately.

5. Discussion and Future Work

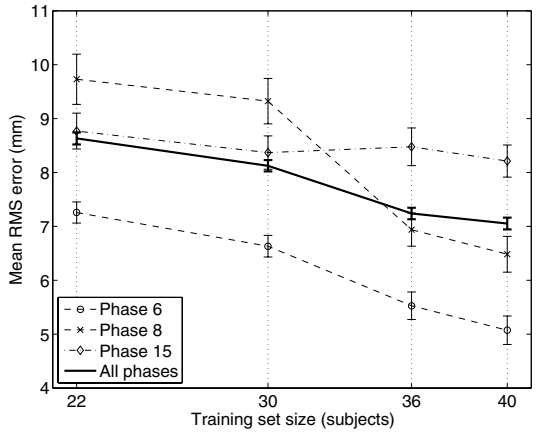
The acquisition scheme as shown in Figure 2 is geared towards the best visibility of the coronary arteries. Therefore, phases 14 and 15 both have a short time span. The bulk of cardiac motion, however, lies just behind the QRS complex, resulting in phases 1, 2 and 3 to be integrated over a time interval with larger spatial changes than the rest. A different temporal sectioning might therefore influence the results.

Secondly, the poor derivation of certain subjects' parameters, as seen in Figure 4, may well be due to the fact that both training and testing data were mixes of healthy and pathological heart shapes. The ratios of these shapes (3:1:1) may have resulted in a bias towards the healthy hearts, resulting in poor approximations of the pathological hearts. While it was out of the scope of this paper, whether the bilinear models are powerful enough to separate these groups may be a topic of further research. Also, whether the results would be better if model and test data consisted of only a single class – healthy, or one specific pathology – is an important issue into the abilities of bilinear models.

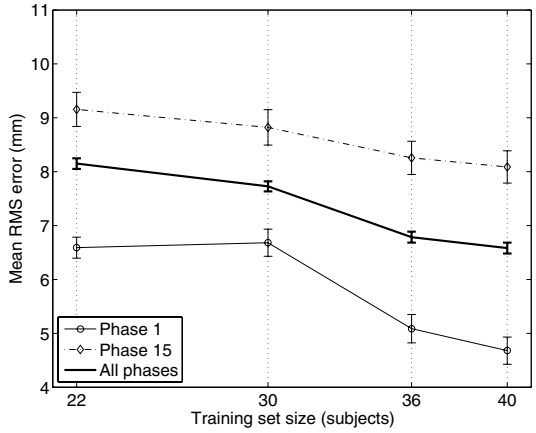
As approximations get worse, they show a tendency to be rotated with respect to the shape they were to form, as can be noted in Figure 5. This could indicate a non-linear variation. A different alignment method, such as Procrustes alignment treating each subject and phase as a different sample, may reduce the approximation errors. However, the relationship between phases of one single subject would need to be preserved in some other way, should this be of importance. Other options would be to pick a most reliable phase, to construct a mean phase, or to treat the entire cycle as one shape during the alignment.

We reported the RMS errors of the reconstructions. While these numbers were not unsatisfactory, they have no clinical application. Future experiments will focus on the prediction of clinical parameters such as Left Ventricular Volume and Ejection Fraction, Wall Thickening and Wall Motion.

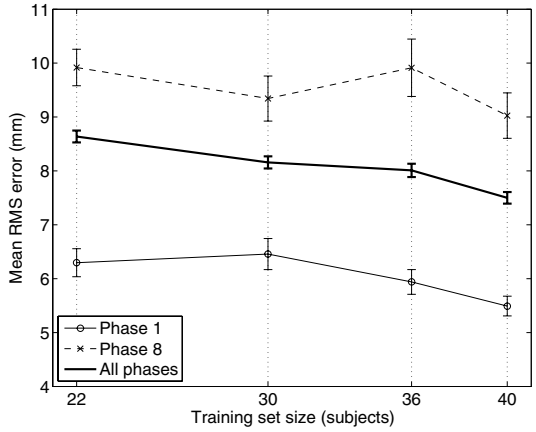
Finally, the reconstruction errors could decrease further if the training set were larger. Assuming that an appropriate number of training samples for PCA-based models of a complex shape, such as the human heart, easily runs into the triple digits [23], there is room for improvement by increasing the number of training samples for the bilinear model as well.



(a)



(b)



(c)

Figure 3. The best, worst and mean RMS reconstruction errors plus and minus the standard error in millimetres, after retraining with (a) phases 1 through 5, (b) 6 through 10, and (c) 11 through 15, using models built with increasing training set sizes.

6. Conclusion

We have shown in this work how to construct bilinear models of the human heart shape, and how to retrain these models with a new subject's data in order to extract its subject parameterisation. Subsequently, we showed that with these parameterisations we could reconstruct the shape of the heart in remaining cardiac phases with a mean RMS error below 7 millimetres, and that 90% of all reconstructions returned an error below 10 millimetres. This suggests that a bilinear factorisation of the heart shape may be appropriate to separate individual and temporal variability.

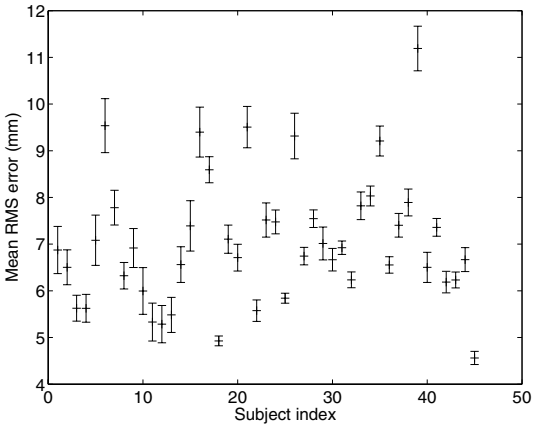


Figure 4. Plot of the mean RMS reconstruction error for each subject. It shows that some subjects have a notoriously bad reconstruction, whereas other subjects' heart shapes were relatively easy to approximate.

Qualitative inspection of the reconstructions in Figure 5 show that even the bad approximations still represent heart shapes. Unlike shape reconstructions made using a PCA-based shape model, this is not trivial, as there is no mean shape, which in itself is a valid instance of the modelled shape class, to which deformations are added, and the constraints usually applied to the deformation of such models (typically a cutoff value in the number of standard deviations) are not present here.

A. Vector Transpose

The training of the symmetric model requires the notion of the vector transpose. Unlike the original application of this term, namely the conversion of a column vector into a row vector and vice versa, the vector transpose we use here is a matrix operation.

Given an $IK \times J$ matrix, where each column was constructed by stacking I K-dimensional column vectors onto each other, the vector transpose [19] of this matrix is a $JK \times I$ matrix. In the case where $K = 1$, the vector transpose is the normal transpose of the matrix. Otherwise, the

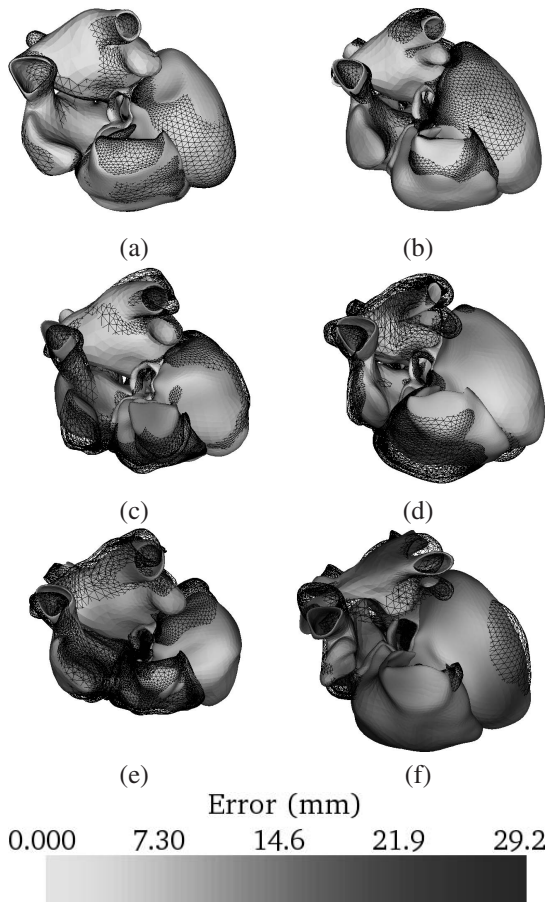


Figure 5. Reconstructions of the hearts (surface rendering) using the highest dimensional model and the hearts that were approximated (wireframe). (a) the best RMS error (1.39 mm), (b) the 25th percentile (5.64 mm), (c) the median (6.79 mm), (d) the 75th percentile (8.23 mm), (e) the 90th percentile (9.95 mm) and (f) the worst RMS error (16.58 mm).

vector transpose looks like

$$\mathbf{X}^{\mathrm{VT}} = \left(\begin{array}{c|c} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ \hline x_{41} & x_{42} \\ x_{51} & x_{52} \\ x_{61} & x_{62} \\ \hline x_{71} & x_{72} \\ x_{81} & x_{82} \\ x_{91} & x_{92} \end{array}\right)^{\mathrm{VT}} = \left(\begin{array}{c|c|c} x_{11} & x_{41} & x_{71} \\ x_{21} & x_{51} & x_{81} \\ x_{31} & x_{61} & x_{91} \\ \hline x_{12} & x_{42} & x_{72} \\ x_{22} & x_{52} & x_{82} \\ x_{32} & x_{62} & x_{92} \end{array}\right) \tag{16}$$

for $K = 3$. A prerequisite is that the size of the vectors is known, but K is usually a user-defined number. Finally, it is easy to see from (16) that the vector transpose operation is invertible: $\left(\mathbf{X}^{\mathrm{VT}}\right)^{\mathrm{VT}} \equiv \mathbf{X}$.

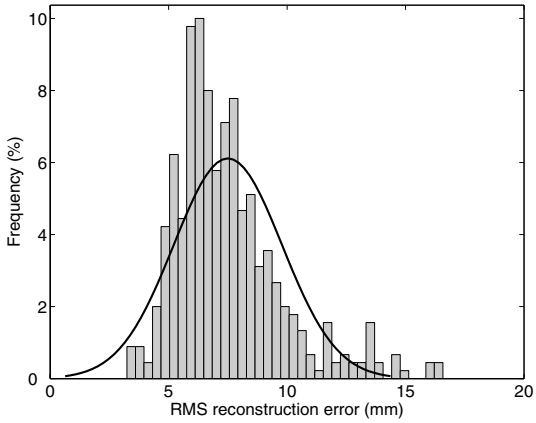


Figure 6. Histogram and bell curve plot of RMS errors over all reconstructions. As each shape occurs twice in the set of shapes to be reconstructed (and once as part of the set of shapes used for the derivation of subject parameters), each shape is represented twice in this histogram.

Acknowledgements

The authors thank R. Leta and F. Carrera from the Cardiology Service, Hospital Santa Creu i Sant Pau, Barcelona, Spain, for the acquisition of the images.

The work of A.F.F. was supported by the Spanish Ministry of Education & Science under a Ramon y Cajal Research Fellowship. This work was partially developed within the framework of the CENIT-CDTEAM Project funded by the Spanish CDTI-MITYC, and also partially supported by MEC TEC2006-03617 and ISCIII FIS2004/40676.

The CILab is a member of the ISCIII CIBER-BBN (CB06/01/0061).

References

[1] B. Abboud and F. Davoine. Bilinear factorization for facial expression analysis and synthesis. *IEE Proceedings – Vision, Image & Signal Processing*, 152(3):327–333, 2004.

[2] J. M. Blackall, A. P. King, G. P. Penney, A. Adam, and D. J. Hawkes. A statistical model of respiratory motion and deformation of the liver. In *Proc. 4th Int. Conf. Medical Image Computing and Computer Assisted Intervention, Utrecht, The Netherlands. Lecture Notes in Computer Science vol. 2208*, pages 1338–1340, 2001.

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

[4] R. Chandrashekara, A. Rao, G. I. Sanchez-Ortiz, R. H. Mohiaddin, and D. Rueckert. Construction of a statistical model for cardiac motion analysis using non-rigid image registration. In *Proc. 18th Int. Conf. Information Processing in Medical Imaging, Ambleside, United Kingdom. Lecture Notes in Computer Science vol. 2732*, pages 599–610, 2003.

[5] E. Chuang and C. Bregler. Mood swings: Expressive speech animation. *ACM Transactions on Graphics*, pages 331–347, 2005.

[6] T. F. Cootes, D. H. Cooper, C. J. Taylor, and J. Graham. Trainable method of parametric shape description. *Image & Vision Computing*, 10(5):289–294, 1992.

[7] T. F. Cootes and C. J. Taylor. Active shape models: 'smart snakes'. In *Proc. British Machine Vision Conference, Leeds, United Kingdom*, pages 266–275, 1992.

[8] F. Cuzzolin. Using bilinear models for view-invariant action and identity recognition. In *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, New York, NY, USA*, pages 1701–1708, 2006.

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

[10] C. Goodall. Procrustes methods in shape analysis. *Journal of the Royal Statistical Society B*, 53(2):285–339, 1991.

[11] D. B. Grimes and R. P. N. Rao. Bilinear sparse coding for invariant vision. *Neural Computation*, 17(1):47–73, 2005.

[12] G. Hamarneh and T. Gustavsson. Deformable spatio-temporal shape models: Extending ASM to 2D+time. In *Proc. British Machine Vision Conference, Manchester, United Kingdom*, pages 13–22, 2001.

[13] E. Hsu, K. Pulli, and J. Popović. Style translation for human motion. *ACM Transactions on Graphics*, 23(3):1082–1089, 2005.

[14] D. G. Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984.

[15] H. Le and D. G. Kendall. The Riemannian structure of euclidean shape spaces: A novel environment for statistics. *Annals of Statistics*, 21(3):1225–1271, 1993.

[16] C.-S. Lee and A. Elgammal. Gait style and gait content: Bilinear models for gait recognition using gait resampling. In *Proc. 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition, Seoul, Korea*, pages 147–152, 2004.

[17] J. R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, 1988.

[18] K. V. Mardia and I. L. Dryden. Shape distributions for landmark data. *Advances in Applied Probability*, 21(4):742–755, 1989.

[19] D. H. Marimont and B. A. Wandell. Linear models of surface and illuminant spectra. *Journal of the Optical Society of America A*, 9(11):1905–1913, 1992.

[20] S. C. Mitchell, B. P. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. Reiber, and M. Sonka. Time-continuous segmentation of cardiac MR image sequences using active appearance motion models. In *Proc. SPIE Medical Imaging, San Diego, CA, USA*, pages 249–256, 2001.

[21] J. Montagnat and H. Delingette. 4D deformable models with temporal constraints: Application to 4D cardiac image segmentation. *Medical Image Analysis*, 9(1):87–100, 2005.

[22] B. M. Ohnesorge, C. R. Becker, T. G. Flohr, and M. F. Reiser. *Multi-slice CT in Cardiac Imaging: Technical Principles, Clinical Application and Future Developments*. Springer-Verlag, 2002.

[23] S. Ordás, E. Oubel, R. Leta, F. Carrera, and A. F. Frangi. A statistical model of the heart and its application to model-based segmentation. In *Proc. SPIE Medical Imaging, San Diego, CA, USA, vol. 6511*, 2007.

[24] J. B. Tenenbaum and W. T. Freeman. Separating style and content. In *Advances in Neural Information Processing Systems, Denver, CO, USA*, pages 662–668, 1996.

[25] J. B. Tenenbaum and W. T. Freeman. Separating style and content with bilinear models. *Neural Computation*, 12:1247–1283, 2000.

[26] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.