

Comparison of Default Patient Surface Model Estimation Methods

Xia Zhong¹, Norbert Strobel², Markus Kowarschik², Rebecca Fahrig²,
Andreas Maier^{1,3}

¹Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg,
Erlangen, ²Siemens Healthcare GmbH, Forchheim, Germany

³Erlangen Graduate School in Advanced Optical Technologies (SAOT)

xia.zhong@fau.de

Abstract. A patient model is useful for many clinical applications such as patient positioning, device placement, or dose estimation in case of X-ray imaging. A default or a-priori patient model can be estimated using learning based methods trained over a large database. Different methods can be used to estimate such a default model given a restricted number of the input parameters. We investigated different learning based estimation strategies using patient gender, height, and weight as the input to estimate a default patient surface model. We implemented linear regression, an active shape model, kernel principal component analysis and a deep neural network method. These methods are trained on a database containing about 2000 surface meshes. Using linear regression, we obtained a mean vertex error of 20.8 ± 14.7 mm for men and 17.8 ± 11.6 mm for women, respectively. While the active shape model and kernel PCA method performed better than linear regression, the results also revealed that the deep neural network outperformed all other methods with a mean vertex error of 15.6 ± 9.5 mm for male and 14.5 ± 9.3 mm for female models.

1 Introduction

Many medical applications even in the field of X-ray imaging using C-arm angiography systems, can benefit from the use of a default patient model. Such a model can be generated using only a limited set of patient parameters (or measurements). Although these default patient models cannot be expected to fit perfectly to the patient, their accuracy is much better than what can be achieved with a stylized model. In fact, if additional sensor data is available, a default patient model can be used as initialization for further model refinement. When a learning-based method is used to estimate a priori patient models, a large database is needed for the training process. The accuracy of the estimation approach depends not only on quality and the quantity of the available data but also on different estimation methods. In this paper, we evaluate different statistical estimation approaches and propose a deep neural network based approach. These methods are implemented and evaluated on a database comprising surface

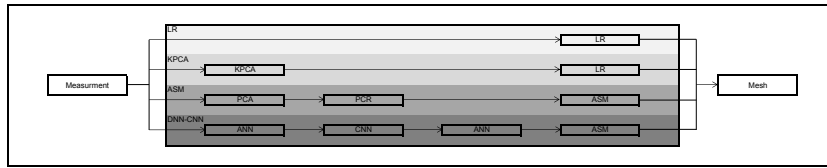


Fig. 1. Pipeline of evaluated methods, from top : linear regression (LR), kernel PCA (KPCA), active shape model (ASM) comprising principle component analysis (PCA) and principle component regression (PCR), and finally deep neural network (DNN) made up of a convolutional neural network (CNN) and an artificial neural network (ANN).

models and associated patient meta data such as gender, height, and weight, to highlight the performance of different approaches. This paper is structured as follows. First, we introduce the related work in the field of shape representation learning and shape estimation. Second, we provide descriptions of the implemented methods. Afterwards, we evaluate on all of the different methods and compare the results. We wrap up by discussing the proposed deep neural network approach for our particular application.

1.1 Related Work

One can put our approach into prior context in two ways. Seen from one direction, our approach is related to methods which learn the shape representation in a low dimensional space. In general, this is a difficult problem as changes in pose between models need to be accounted for as well as differences in body shape. The most widely used solution for this task is the Shape Completion and Animation for People (SCAPE) method [1]. This approach assumes that shape and pose are uncorrelated and solves the two problems separately at first and jointly afterwards. Hasler et al. [2] encodes in the shape representation further information, e.g., height and joint angle and trains a combined representation of shape and pose simultaneously. Based on the SCAPE method, Pishchulin et al. [3] proposed an improved method by incorporating mesh sampling into the training loop. The second set of methods revolves around patient shape generation base on different input parameters. Seo et al. [4] introduced a framework to generate a surface model using different input parameters. Rather than using parameters, Wuhrer et al. [5] uses anthropological measurements e.g., arm length, waist circumference as an input to both initialize and refine the surface model estimate.

2 Materials and Methods

In this paper, we implemented and evaluated different methods for default patient model estimation using patient gender, height, and weight as input. The

approaches can be summarized in three steps: data transform, regression, and shape reconstruction. During the data transform step, we map the input data into a different space for the successive regression step. The mapping function can be linear, e.g., principal component analysis (PCA), or non-linear. In this case, a kernel function or artificial neural network (ANN) may be applied. The regression step maps the transformed data to the patient model itself or its low dimensional representation. The reconstruction step, in turn, maps the representation of the patient model to its original space. In this paper, we evaluated four different methods for patient model estimation. An overview of our approaches is shown in Fig. 1. All these methods are learning based methods, and a database is used for training. The data comprised male and female surface mesh data and corresponding measurements. We can summarize the estimation method using the equation

$$\hat{\mathbf{x}} = \mathcal{F}(\mathbf{X}, \mathbf{g}) \quad (1)$$

where $\hat{\mathbf{x}}$ denotes the estimated patient model, \mathbf{g} is the measurement of the patient, and \mathbf{X} refers to the collection of surface meshes stored in the database. The function \mathcal{F} is the regression function that needs to be found. It depends on the database contents.

2.1 Linear Regression

Linear regression (LR) is the most straightforward method for solving the regression problem. Let $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_k, \dots, \mathbf{g}_N] \in \mathbb{R}^{K \times N}$ denote the measurement matrix combining every measurement \mathbf{g}_i for surface mesh \mathbf{x}_i . Similarly the surface mesh matrix is defined as $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_k, \dots, \mathbf{x}_N] \in \mathbb{R}^{M \times N}$. The variable N refers to the number of meshes, M denotes the number of single mesh dimensions, e.g., the number of vertices, and K is the dimension of the measurements. In this application, $K \ll M$. The LR introduces a regression matrix $\mathbf{A} \in \mathbb{R}^{M \times K}$ where

$$\mathbf{A}\mathbf{G} = \mathbf{X} \quad (2)$$

This problem can be solved simply by using singular value decomposition (SVD) as $\mathbf{A} = \mathbf{X}\mathbf{G}^+$. Then the estimation patient model $\hat{\mathbf{x}}$ with given \mathbf{g} is

$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{g} \quad (3)$$

2.2 Kernel Principal Component Analysis (PCA)

The kernel PCA method tries to improve the result of linear regression by introducing feature mapping using a kernel. In this case, the input measurements are mapped to a feature space using kernel PCA. The feature space is mapped to meshes using linear regression as introduced in the previous subsection. We use Gaussian and linear kernels for the kernel PCA. The reason for feature mapping is that the measurements, e.g., height and weight, may not be linearly related to the mesh vertex positions. Take the weight, for example. An integral is needed

to relate it to the vertex positions. Using a linear or Gaussian kernel function Φ the estimation can be fomulated as

$$\hat{\boldsymbol{x}} = \boldsymbol{A}\Phi(\boldsymbol{g}) \quad (4)$$

2.3 Active Shape Model

The motivation to use an active shape model (ASM) is to learn a joint subspace between shape and measurement. To this end, we assume that shape and measurements are correlated and that they can be described in the same subspace. By using the method proposed in [6], the measurement vector \boldsymbol{g} and the shape \boldsymbol{x} can be described as

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{Q}_s \boldsymbol{c} \quad (5)$$

$$\boldsymbol{g} = \bar{\boldsymbol{g}} + \boldsymbol{Q}_m \boldsymbol{c} \quad (6)$$

where \boldsymbol{c} is the low dimensional representation of both shape and measurement, and $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{g}}$ denote the average shape and measurement. The matrix \boldsymbol{Q}_s and \boldsymbol{Q}_m are the corresponding modes of variation of shape, and measurement respectively. Except the low dimensional representation \boldsymbol{c} , all other parameters are derived from the database by solving a matrix decomposition problem [7]. The estimation equation of the method is

$$\hat{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \boldsymbol{Q}_s \boldsymbol{Q}_m^+ (\boldsymbol{g} - \bar{\boldsymbol{g}}) \quad (7)$$

2.4 Deep Neural Network

The default model estimation problem can also be approached using a deep neural network (DNN) approach. The advantage of using a DNN is that we can learn a potentially non-linear mapping between measurements and shape. We propose a different neural network topology here than the general regression neural networks [8] as we try to estimate a high-dimensional output (mesh model) using a low-dimensional input (height, weight). Instead of regressing to the mesh model directly, we estimate a low dimensional representation of the mesh model. The low dimensional representations we used here is the ASM parameter \boldsymbol{c} trained as described above. Other low dimensional representations e.g. bottleneck features can also be used. We applied patch normalization to all the layers for robust training.

An illustration of the network architecture is shown in Fig. 2. This network takes height and weight as input and uses an ANN to expand the input vector. The output of the ANN is then converted to an image and we use a convolution neural network (CNN) layout (alternating convolution and max pooling layer) in the following layers. The result of the CNN is being flattened into the ANN layout and to the output layer. The combination of CNN followed by ANN is well established as a robust method for multi-label classification [9] of images. The proposed network also benefits from this structure and the evaluation showed

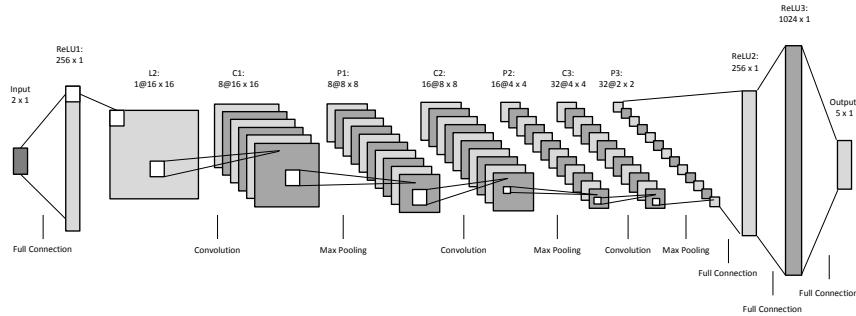


Fig. 2. An illustration of the proposed regression network topology. Input parameters are height and weight. Two DNNs were set up, one for meshes associated with women, another one for meshes representing men.

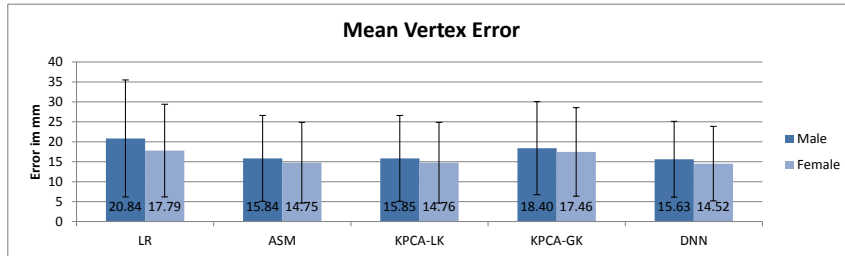


Fig. 3. The mean vertex error for male and female models using (from left to right): linear regression (LR), active shape model (ASM), kernel PCA with linear kernel (KPCA-LK), kernel PCA with Gauss kernel (KPCA-GK), and deep neural network (DNN).

that this architecture works well for our regression problem. Estimating a surface model using a DNN can be formulated as

$$\hat{x} = \bar{x} + Q_s \mathcal{F}^{\text{DNN}}(g) \tag{8}$$

In Eq. 8, the symbol \mathcal{F}^{DNN} represents a trained network mapping function.

3 Evaluation and Results

For evaluation, we used a database comprising surface meshes and corresponding meta data (measurement data) such as height and weight. In all, there were 865 male data sets and 1063 female data sets. In the evaluation, we randomly picked ten percent of the data as the test set. The remaining 90% were used for training. The error measure is the mean vertex error distance in mm between the estimated surface mesh and the original surface mesh associated with the corresponding height and weight. Female and male surface meshes were treated separately. The results of the different methods are shown in Fig. 3.

From Fig. 3 we learn that linear regression had the largest mean vertex error with 20.8 ± 14.7 mm for male and 17.8 ± 11.6 mm for the female surface model. Kernel PCA with Gaussian kernel improved the result for male and female model to 18.4 ± 11.7 mm and 17.5 ± 11.1 mm, respectively. The active shape model and kernel PCA with linear kernel lead to almost identical results with 15.8 ± 10.8 mm for male and 14.7 ± 10.1 mm for female models. The deep neural network outperformed all the methods, albeit slightly, with a mean vertex error of 15.6 ± 9.5 mm for male and 14.5 ± 9.3 mm for female models.

4 Discussion

Although the DNN method outperformed the linear method by introducing a non-linear mapping, the gain obtain from using the deep learning method was limited. One of the possible reasons can be that the size of the data set was too small for deep learning to reveal its full potential. For our particular application, it is difficult to expand the number of data as in other image based problem, where the training data can be expanded by rotating or deforming the images. In our case, we could deform surface meshes, but the corresponding changes in height and weight would be unknown.

Acknowledgments We gratefully acknowledge the support of Siemens Healthineers, Forchheim, Germany. We also thank Siemens Corporate Technology for providing the avatar database. Note that the concepts and information presented in this paper are based on research, and they are not commercially available.

References

1. Anguelov D, Srinivasan P, Koller D, et al. SCAPE: Shape Completion and Animation of People. *ACM Trans Graph.* 2005 Jul;24(3):408–416.
2. Hasler N, Stoll C, Sunkel M, Rosenhahn B, Seidel HP. A Statistical Model of Human Pose and Body Shape. *Comput Graph Forum.* 2009;28(2):337–346.
3. Pishchulin L, Wuhrer S, Helten T, et al. Building Statistical Shape Spaces for 3D Human Modeling. *CoRR.* 2015;.
4. Seo H, Magnenat-Thalmann N. An Automatic Modeling of Human Bodies from Sizing Parameters. In: *Proceedings of the 2003 Symposium on Interactive 3D Graphics;* 2003. p. 19–26.
5. Wuhrer S, Shu C. Estimating 3D human shapes from measurements. *Mach Vis Appl.* 2013;24(6):1133–1147.
6. Edwards GJ, Lanitis A, Taylor CJ, Cootes TF. Statistical models of face images improving specificity. *Image Vis Comput.* 1998;16(3):203 – 211.
7. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models – Their Training and Application. *Comput Vis Image Underst.* 1995 Jan;61(1):38–59.
8. Specht DF. A general regression neural network. *IEEE Trans Neural Netw.* 1991 Nov;2(6):568–576.
9. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: *Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Adv Neural Inf Process Syst;* 2012. p. 1097–1105.