Technical University of Denmark

DTU

Estimation of shape model parameters for 3D surfaces

Erbou, Søren Gylling Hemmingsen; Darkner, Sune; Fripp, Jurgen; Ourselin, Sébastien; Ersbøll, Bjarne Kjær

DTU Library
Technical Information Center of Denmark

ESTIMATION OF SHAPE MODEL PARAMETERS FOR 3D SURFACES

Søren G. H. Erbou[a,b], Sune Darkner[a], Jurgen Fripp[b], Sébastien Ourselin[b,c], Bjarne K. Ersbøll[a]

[a]DTU Informatics, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark.
[b]The Australian e-Health Research Centre, CSIRO ICT Centre, Brisbane, Australia.
[c]Centre for Medical Image Computing, University College London, UK.

ABSTRACT

Statistical shape models are widely used as a compact way of representing shape variation. Fitting a shape model to unseen data enables characterizing the data in terms of the model parameters. In this paper a Gauss-Newton optimization scheme is proposed to estimate shape model parameters of 3D surfaces using distance maps, which enables the estimation of model parameters without the requirement of point correspondence. For applications with acquisition limitations such as speed and cost, this formulation enables the fitting of a statistical shape model to arbitrarily sampled data. The method is applied to a database of 3D surfaces from a section of the porcine pelvic bone extracted from 33 CT scans. A leave-one-out validation shows that the parameters of the first 3 modes of the shape model can be predicted with a mean difference within [-0.01,0.02] from the true mean, with a standard deviation less than 0.34.

Index Terms— Image shape analysis, Image registration, Biomedical image processing, Optimization methods, X-ray tomography

1. INTRODUCTION

Statistical shape models (SSM) [1] are often used to characterize unseen shapes in terms of model parameters which can be used for classification or regression. In order to estimate the shape model parameters of an unseen shape, full point correspondence is usually needed to be able to project the shape into the parameter space. Obtaining full point correspondence might not be possible in some applications due to restrictions such as acquisition and computation time, dosage (CT) and cost. In such applications there is a need for registration of dense models to incomplete data and for parameter estimation of the unseen shape.

In [2] van Assen *et al.* proposed a method for fitting a dense model to sparse data. Model points near sparse data points are propagated onto void areas using a Gaussian kernel in order to achieve (pseudo-) correspondence making it possible to estimate model parameters. The framework is applied to segmentation of cardiac MRI data and different sparsity schemes are tested. Rajamani *et al.* [3] formulated an algorithm for matching a SSM to sparse digitized points to create patient specific models for pre-operative planning. A linear system of equations is solved to obtain a least squares fit of the model to the digitized points. A Mahalanobis distance based regularization term and M-estimator based weighting of the digitized points are included in the matching algorithm. Due to the nature of their applications both papers focus on the reconstruction error but not on how well the actual model parameters are estimated.

In this paper the focus is on how accurate the model parameters can be estimated. For applications where the parameters are used for classification or regression it is important to quantify how trustworthy this estimate is. An iterative Gauss-Newton optimization algorithm is proposed for fitting a SSM to unseen data using sampled distance maps. It is investigated on dense data, without requiring point correspondence and in future papers the effect of reducing the amount of data, i.e. increasing sparsity, will be investigated. Several authors, e.g. Golland *et al.* [4], have proposed representing shapes using distance maps. This results in a more dense model which would not be feasible in the present application.

The proposed method is applied to a SSM of a porcine bone structure which will be used in a slaughterhouse robotic tool. In this specific application the model parameters are interesting as they can be used to obtain a relation with specific quality measures of the carcasses. The method is also applicable in a range of biomedical applications.

2. METHODS

2.1. Statistical shape models

SSM's were proposed by Cootes *et al.* as a compact way of describing shape variation in a data set [1]. Let the n shapes, or in our case 3D surfaces, be represented by k corresponding 3D points, each arranged in a $3k$ vector $\mathbf{s}$. The idea is to formulate a parameterized model of the form $\mathbf{s} = M(\mathbf{b})$ describing the variation seen in the data, where $\mathbf{b}$ is a vector of shape parameters of the model M. To exclude the effects of translation, rotation and scaling, a generalized Procrustes alignment is performed before constructing the shape model [5]. The sample mean ($\bar{\mathbf{s}} = 1/n \sum_{i=1}^{n} \mathbf{s}_i$) and sample covariance matrix ($\mathbf{C} = 1/(n-1) \sum_{i=1}^{n} (\mathbf{s}_i - \bar{\mathbf{s}})(\mathbf{s}_i - \bar{\mathbf{s}})^T$) are then computed. Since the original parameter space is usually much larger than the number of observations ($3k >> n$) applying principal components analysis (PCA) on the covariance matrix is an obvious choice for dimensionality reduction. The PCA determines the main axes (eigenvectors ϕ_i) of variation of the data and sorts them according to the amount of variation they describe (eigenvalues λ_i). The model can then be formulated as a perturbation of the mean shape:

$$\mathbf{s} = \bar{\mathbf{s}} + \Phi \mathbf{b} \qquad (1)$$

where Φ is the matrix composed of the eigenvectors ϕ_i.

The model parameters of a new aligned shape s' can be obtained by projecting it into the parameter space,

$$\mathbf{b}' = \Phi^T (\mathbf{s}' - \bar{\mathbf{s}}). \qquad (2)$$

This is only possible if all the k points of the aligned shape are available.

2.2. Optimization algorithm

The Lucas-Kanade algorithm for image registration was originally formulated using image intensities [6, 7] and was typically applied within fields such as stereo vision and motion analysis. It is an iterative Gauss-Newton optimization algorithm. In the following the parameter estimation of the SSM is considered as a constrained registration problem, thus the Lukas-Kanade approach can be applied. This allows parameter estimation without all the points being available.

Let I be the signed distance map of the input surface and T the signed distance map of the template surface, with regions of the template surface that are within the input surface having negative distance. Furthermore let $\mathbf{x} = (x, y, z)^T$ be a vector of sample points in the distance maps, $\mathbf{p}$ be a set of parameters and $\mathbf{W}(\mathbf{x}; \mathbf{p})$ a warp of $\mathbf{x}$ with $\mathbf{p}$. The objective function to be minimized is the sum of squared differences between the warped I, and T,

$$\sum_{\mathbf{x}} [I(\mathbf{W}(\mathbf{x}; \mathbf{p})) - T(\mathbf{x})]^2. \tag{3}$$

This can be formulated iteratively with incremental updates of $\mathbf{p}$ using a Gauss-Newton scheme [6],

$$\sum_{\mathbf{x}} [I(W(\mathbf{x}; \mathbf{p} + \Delta\mathbf{p})) - T(\mathbf{x})]^2. \tag{4}$$

For each step the parameters are updated,

$$\mathbf{p} \leftarrow \mathbf{p} + \Delta\mathbf{p}, \tag{5}$$

and this procedure is repeated until convergence. Linearizing by performing a first order Taylor expansion of eq. (4) results in,

$$\sum_{\mathbf{x}} [I(\mathbf{W}(\mathbf{x}; \mathbf{p})) + \nabla I \frac{\partial \mathbf{W}(\mathbf{x}; \mathbf{p})}{\partial \mathbf{p}} \Delta\mathbf{p} - T(\mathbf{x})]^2. \tag{6}$$

where ∇I is the gradient of I evaluated at $\mathbf{W}(\mathbf{x}; \mathbf{p})$ and $\frac{\partial \mathbf{W}(\mathbf{x};\mathbf{p})}{\partial \mathbf{p}}$ is the Jacobian of the warp. Solving for $\Delta\mathbf{p}$ reveals,

$$\Delta\mathbf{p} = \mathbf{H}^{-1} \sum_{\mathbf{x}} \left[\nabla I \frac{\partial \mathbf{W}(\mathbf{x}; \mathbf{p})}{\partial \mathbf{p}} \right]^T [T(\mathbf{x}) - I(\mathbf{W}(\mathbf{x}; \mathbf{p}))] \tag{7}$$

where H is the Gauss-Newton approximation to the Hessian,

$$\mathbf{H} = \sum_{\mathbf{x}} \left[\nabla I \frac{\partial \mathbf{W}(\mathbf{x}; \mathbf{p})}{\partial \mathbf{p}} \right]^T \left[\nabla I \frac{\partial \mathbf{W}(\mathbf{x}; \mathbf{p})}{\partial \mathbf{p}} \right]. \tag{8}$$

Since I consists of distances to the surface to be registered, the gradient (∇I) corresponds to inward pointed normals of the surface evaluated at $\mathbf{W}(\mathbf{x}; \mathbf{p})$. If the surface is moved along the inward pointed normal (away from the sample $\mathbf{x}$) the distance I increases.

The warp $\mathbf{W}(\mathbf{x}; \mathbf{p})$ can be any type, e.g. rigid, affine or non-rigid transformation with corresponding parameters $\mathbf{p}$ [7, 8]. In our case the warp is $\mathbf{W}(\mathbf{x}; \mathbf{p}) = \mathcal{C}(\mathbf{x}, \mathbf{s})$. $\mathcal{C}(\mathbf{x}, \mathbf{s})$ is the set of points in the warped surface $\mathbf{s}$, eq. (1), that are closest point to each sample in $\mathbf{x}$. These are obtained using a kd-tree. The parameters $\mathbf{p}$ are the first t shape model parameters weighted by the square root of their corresponding eigenvalues, normalizing $\mathbf{p}$ to standard deviations ($\mathbf{p} = (b_1/\sqrt{\lambda_1}, ..., b_t/\sqrt{\lambda_t})^T$) off the mean in model parameter space.

∇I is computed as the negative mean of the normals of the faces connected to each point in $\mathbf{W}(\mathbf{x}; \mathbf{p})$.

$$\nabla I_i = -\frac{1}{f_i} \sum_{k=1}^{f_i} \mathbf{v}_k, \tag{9}$$

where f_i is the number of faces connected to the ith vertex/point and $\mathbf{v}_k$ is the outwards normal of the kth connected face. Using the angle weighted normals would likely give a better estimate of the mean curvature at the vertices, but it would be computationally more expensive. The Jacobian of the warp for the ith sample in $\mathbf{x}$ is composed from the respective counterparts in the eigenvectors of the model.

$$\frac{\partial \mathbf{W}(\mathbf{x}_i; \mathbf{p})}{\partial \mathbf{p}} = \begin{pmatrix} \phi_{x_i,1}\sqrt{\lambda_1} & \cdots & \phi_{x_i,t}\sqrt{\lambda_t} \\ \phi_{y_i,1}\sqrt{\lambda_1} & \cdots & \phi_{y_i,t}\sqrt{\lambda_t} \\ \phi_{z_i,1}\sqrt{\lambda_1} & \cdots & \phi_{z_i,t}\sqrt{\lambda_t} \end{pmatrix} \tag{10}$$

In the above formulation the sample vector $\mathbf{x}$ can be constructed arbitrarily, (within a sensible range from the surface) and the optimization algorithm will seek to minimize the rms error between the distance maps. It is therefore possible to estimate the model parameters without having full point correspondence.

To estimate $\mathbf{p}$ requires an initial estimate, which in this case is $\mathbf{p} = \mathbf{0}$ which corresponds to the mean shape of the shape model. By applying equations (7, 9, 10, 8 & 5) we obtain a new estimate of $\Delta\mathbf{p}$ that minimize a first order estimate of a quadratic surface to the parameter space. If this is far from the global optimum, the estimate will be inaccurate. To rectify this a line search is applied at each iteration if the full step did not reduce the cost function. This is initialized with a small step size, which doubles until one step before the cost function starts to increase, which ensures a reasonable tradeoff between computations and optimum step size.

To improve speed several papers, e.g. [7, 8], propose to formulate the Lukas-Kanade algorithm in the inverse compositional way making it possible to pre-compute several steps, especially the Hessian in eq. (8). This is beneficial if the sample vector $\mathbf{x}$ is very large, which it is not in the present application. Furthermore applying the inverse compositional algorithm would require the use of surface information from the template image, which may not be feasible to acquire.

2.3. Validation

The validation is performed in a leave-one-out scheme (LOO), where the model parameters for each surface i are estimated using a SSM constructed using all but the ith surface. The true set of parameters are found by projecting the aligned surface into the parameter space of the model using eq. (2). The actual parameter estimates and the rms errors (point-to-point) are then compared. Only absolute distances less than 5 mm are included when computing the rms error in order to reduce the effect of outliers which can occur due to missing and non-corresponding regions.

3. DATA

The method is applied to a data set consisting of 33 cases of 3D surfaces from a section of the porcine pelvic bone. Implicit surfaces are extracted from CT scans using radial basis functions [9] from which the surfaces are reconstructed as triangular meshes. The surfaces of the bone of interest are disconnected from the skeleton by planes and therefore have two open ends. Furthermore they would have genus 1 topology, i.e. topological similar to a torus, if the ends were closed.

Correspondence is obtained using the iterative closest point (ICP) algorithm [10] using a similarity transform (translation, rotation and scaling) extended with a point-to-surface step determined by a search along the direction of the estimated vertex normal to the other surface. A reference shape is constructed by initially performing registration of a specific shape to the other shapes, then

625

computing a new reference shape as the mean, performing registration of this to all shapes until convergence of the mean shape. The reference shape consists of 3815 vertices and 7397 faces.

Choosing the number of modes (t) to include in the model is a tradeoff between including the general intrinsic variation of the data and excluding noise. Applying parallel analysis (PA) as suggested by Horn [11], results in 7-9 modes of variation to be included for the different LOO models. PA only includes modes that contain more variation than can be explained by noise, i.e. modes with intrinsic variation less than the noise level are excluded.

Since we investigate the parameter estimation ability of the algorithm for this specific data set, the sample vector $\mathbf{x}$ is composed of the full point set of the surface to register the model to. $T(\mathbf{x})$ is $\mathbf{0}$.

4. RESULTS

Figure 1 shows the first 3 modes of variation (rows) in one of the LOO models, perturbed -3 std. (left column) and +3 std. (right column). The middle column is the mean shape, with the first 3 and 7 modes containing 45% and 65% of the variation in the data, respectively.
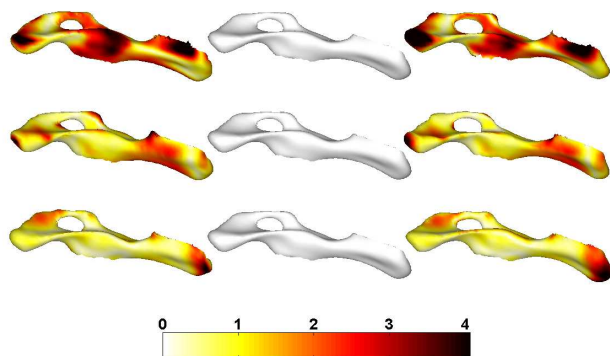

Fig. 1. Three principal modes of variation (rows) cover 25%, 12% and 8% respectively of the total variation in the data. Left and right columns denote -3 and +3 std. from the mean shape (middle column). The color coding indicates the absolute distance in mm to the mean shape.

When estimating shape model parameters errors accumulate through modes and it might not be possible to estimate more than the first few of modes. If the estimate of the first mode is incorrect, the other modes try to compensate in terms of reducing the rms error of the point-to-point distance. With that in mind 3 schemes are reported, one where all the model parameters are estimated in a combined optimization (comb.) and 2 sequential schemes where only the last mode is estimated, fixing the previous modes to the estimated value (seq-est), and to the true value (seq-true). The latter scheme is included for comparison even though the true parameters would not be accessible in an application. Still it gives an indication of the error levels that should be expected.

In the 3 schemes the true pose is used as initialization and parameter estimates and rms errors compared when including from 1 to 7 modes. The true pose is defined as the pose obtained from the ICP registration, applying the similarity transform. The effect of not having the true pose will be investigated in future work.

Modes	Comb.	Seq-est	Seq-true
1	0.01 (0.33)	0.01 (0.33)	0.01 (0.33)
2	-0.07 (0.34)	-0.01 (0.34)	0.02 (0.35)
3	-0.04 (0.56)	0.02 (0.32)	-0.01 (0.32)
4	-0.07 (0.76)	0.03 (0.53)	0.04 (0.52)
5	0.07 (0.58)	0.04 (0.49)	0.01 (0.43)
6	-0.21 (0.97)	-0.07 (0.42)	-0.06 (0.47)
7	-0.55 (1.08)	-0.17 (0.98)	-0.16 (0.76)

Table 1. Difference between estimated and true parameter values for the three schemes when estimating 1 through 7 modes. Mean and (std.) are reported.

Figure 2 and table 1 show the mean and std. of the difference between the estimated and true values of the parameter estimates in the LOO validation. For the majority of cases the mean is within ± 0.1 std. from the true value, which seems reasonable. The std. of the difference is more interesting, since it gives an idea of how far off most of the estimates are. When estimating 1 or 2 modes the three schemes approximately have the same std. of the difference, less than 0.35. When estimating 3-6 modes the two sequential schemes have a std. within 0.53, while the std. of the combined scheme only is within 0.76 when estimating 3-5 modes. This shows that the parameter estimates for the combined scheme try to balance each other out, resulting in less reliable estimates.
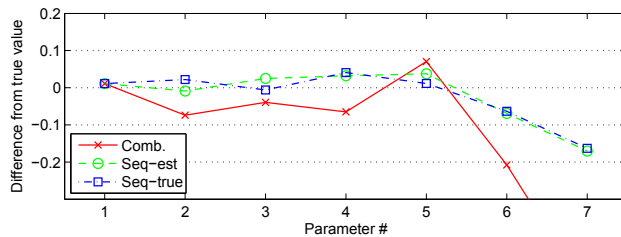
(a) Mean.
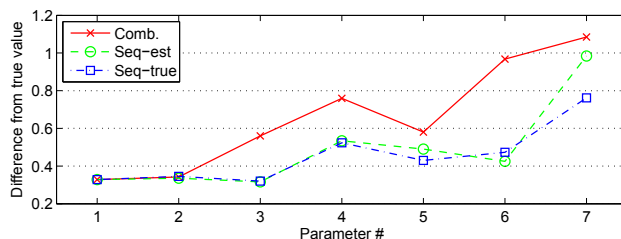
(b) Std.

Fig. 2. Difference between estimated and true parameter values for the three optimization schemes.

Figure 3 shows the mean rms errors for the three parameter estimation schemes and for the true set of parameters. For the initial set of parameters, the mean rms error is 1.85 mm. It shows that the more parameters that are estimated in the combined scheme, the more likely the optimization is to converge to a local minima, where the rms error is significantly higher than for the true set of parameters. For the seq-est scheme the difference is within 0.01 mm for 1-5 modes. This indicates that the seq-est scheme is preferable to the combined scheme, revealing similar rms values as for the true parameter values when using 1-5 modes. The std. of modes higher than

the third mode are substantially higher than for the first three modes, so including more than 3 modes should depend on the acceptable level for the std. of the difference. For the present application it is suggested to include 3 modes.

Figure 4 shows the mean and std. of the rms error for each surface point plotted on the mean surface, when 3 modes of variation are used. The error is nicely distributed over the central parts of the shape, with the main errors located at either end and along the edge of the top right part of the hole. The mean rms error is 1.50 mm and the mean std. is 0.59 mm.
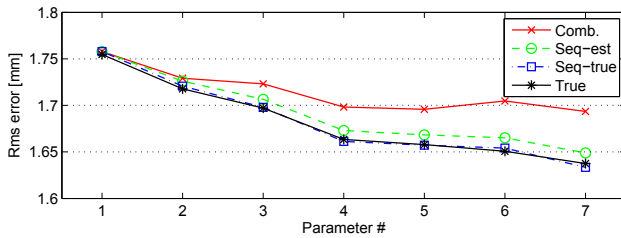

Fig. 3. Mean rms error of leave-one-out validation for different number of modes included, the std. is 0.12 mm. The mean value of the initial rms error is 1.85 mm.
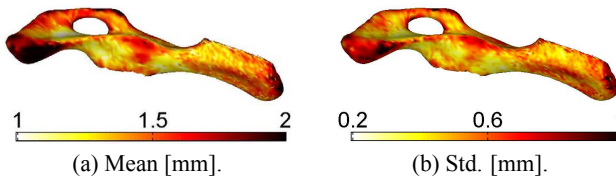

(a) Mean [mm].　　　(b) Std. [mm].

Fig. 4. Mean (a) and std. (b) of rms error, in mm, for each point plotted on the mean shape and reconstructed using 3 modes. The mean rms error on the surface is 1.50 mm. and the mean of the std. is 0.59 mm.

5. CONCLUSION

An iterative Gauss-Newton algorithm is applied to estimate statistical shape model parameters for unseen data. The optimization is driven by sampling in distance maps, which can be done arbitrarily. This enables the estimation of model parameters without the need of full point correspondence, which would be needed if the aligned shape were to be projected into the parameter space.

For applications where it is of interest to fit a dense model to sparse data in order to estimate model parameters e.g. for classification or regression purposes this method can be applied. This is important in applications where time, dosage, cost etc. are issues.

Leave-one-out validation of parameter estimates for a data set of 3D surfaces of a specific porcine bone shows that the three primary modes of the statistical shape model can be estimated with a mean difference between [-0.01,0.02] std. and with a std. of the difference within 0.34. This is done using a sequential estimation scheme where each parameter is estimated sequentially with previous parameters fixed. If 6 modes are estimated the upper limit of the std. of the difference increases to 0.53 and the mean difference is between [-0.07,0.04] std. The difference between the true and the estimated rms error is below 0.01 mm. The rms error decreases

when increasing the number of modes, but the parameter estimates are only reliable enough for the first 3 modes in the present application.

The sample vector that drives the optimization can be arbitrarily defined, which is a topic for future work, along with the introduction of weights on the cost function. Other topics to investigate include applying more localized models, i.e. ICA-models, varimax rotation or sparse PCA.

The statistical shape model will be used in a slaughterhouse robotic tool and the parameters are of interest as they can be used to obtain a relation with specific quality measures of the carcasses.

7. REFERENCES

[1] 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.

[2] H. C. van Assen, M. G. Danilouchkine, A. F. Frangi, S. Ordás, J. J. M. Westenberg, J. H. C. Reiber, and B. P. F. Lelieveldt, "SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data," *Medical Image Analysis*, vol. 10, pp. 286–303, 2006.

[3] K. T. Rajamani, M. A. Styner, H. Talib, G. Zheng, L. P. Nolte, and M. A. G. Ballester, "Statistical deformable bone models for robust 3D surface extrapolation from sparse data," *Medical Image Analysis*, vol. 11, pp. 99–109, 2007.

[4] P. Golland, W. E. L. Grimson, M. E. Shenton, and R. Kikinis, "Small sample size learning for shape analysis of anatomical structures," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2000, vol. 1935, pp. 72–82.

[5] J. C. Gower, "Generalized Procrustes analysis," *Psychometrika*, vol. 40, pp. 33–50, 1975.

[6] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, 1981, pp. 674–679.

[7] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 221–255, 2004.

[8] M. Vester-Christensen, S. G. Erbou, S. Darkner, and R. Larsen, "Accelerated 3D image registration," in *SPIE International Symposium on Medical Imaging*, 2007, vol. 6512, (651201).

[9] G. Turk and J. F. O'Brien, "Shape transformation using variational implicit functions," *Computer Graphics*, vol. 33, pp. 335–342, 1999.

[10] P.J. Besl and N.D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, February 1992.

[11] J. L. Horn, "A rationale and test for the number of factors in factor analysis," *Psychometrika*, vol. 30, pp. 179–185, 1965.