Low Dimensional Surface Parameterisation with Applications in Biometrics

Wei Quan, Bogdan J. Matuszewski, Lik-Kwan Shark, Djamel Ait-Boudaoud
Applied Digital Signal and Image Processing Centre
University of Central Lancashire, Preston PR1 2HE, United Kingdom
{wquan,bmatuszewski1,lshark,dait}@uclan.ac.uk

Abstract

This paper describes initial results from a novel low dimensional surface parameterisation approach based on a modified Iterative Closest Point (ICP) registration process which uses vertex based Principal Component Analysis (PCA) to incorporate a deformable element into registration process. Using this method a 3-D surface is represented by a shape space vector of much smaller dimensionality than the dimensionality of the original data space vector. The proposed method is tested on both simulated 3-D faces with different facial expressions and real face data. It is shown that the proposed surface representation can be potentially used as feature space for a facial expression recognition system.

1. Introduction

Surface representation, modelling and manipulation have been for many years of great interest to computer graphics community, where they have been developed and used primarily in computer aided design as well as for efficient representation and rendering of generic scenes. More recently these techniques have been adopted by computer vision community to help solve problems related to reverse engineering, image base rendering, human appearance modelling , volume segmentation and object recognition. This paper proposes a novel approach to low dimensional surface parameterisation based on principal component analysis (PCA) and iterative deformable surface registration. The low dimensional surface parameters can be considered as intrinsic features for 3-D surface classification including facial expression recognition.

The objective of surface modelling is to model the data by different surface representations. Surface representation methods can be mathematically classified into three categories: implicit surfaces, explicit surfaces and parametric surfaces [12]. Implicit surfaces are often used in computer vision and computer graphics to model evolving in-terfaces. Explicit surface are axis dependent and cannot adequately represent multiple-valued functions. In comparison the parametric surfaces, such as radical basis functions, spherical harmonics functions, Bezier and B-spline functions, are extremely flexible, axis independent, and can represent multiple-valued functions.

The purpose of registration of 3-D surfaces is to find transformation which aligns two or more surfaces. According to the choice of transformation model the registration can be classified as rigid or deformable. Based on the similarity criterion, the registration can be: feature, point, model or global similarity based. Feature-based methods attempt to register surfaces based on a set of features which provide a compact description of the surface shape, such as principal curvatures, local peaks, pit and saddle points [5]. However, most of the feature-based methods can only work when the deformation between surfaces is very small or there is no deformation at all. Point-based methods register surfaces using relatively dense data sets. An example of such method is the iterative closest point (ICP) method [6] used to register surfaces by iteratively minimizing a global cost function. Chui et al [8] modified the ICP methods by using a non-rigid transformation model, such as affine and thin-plate spline, instead of rigid body transformation, to achieve the deformable surface registration. The basic idea of model-based methods is to physically model an object by using internal and external virtual forces [11]. It is often used to track deformations of surfaces over time. Global similarity methods use the spin image [10] or shape context [4] but they are not well suited for surface registration with large deformation.

The method proposed in this paper uses a shape space parametric representation to model surface deformations. The shape space is constructed using principal component analysis (PCA) computed for typical surfaces represented in a training data set. The shape space is closely linked to so called Point Distribution Model (PDM) which is pioneered by Cootes et al [9]. A modified ICP is used to calculate correspondence, for dense set of points defined for each surface, and subsequent surface registration with deformations

defined by the constructed shape space. The authors consider the use of the shape space to incorporate a deformable element into the ICP registration as an original contribution. The remainder of the paper is organised as follows: Sections 2 briefly describes the surface shape space; Sections 3 describes the modified ICP and the registration procedure; This is followed by evaluation of the method in Section 4.

2. Surface Model

In order to achieve a low dimensional surface parameterisation, the statistical surface model is built first. The model is learnt from a training data set of 3-D simulated surfaces representing human faces generated by $FaceGenModeller^{\circledR}$ [13]. Using a standard PCA [3], each 3-D surface can be approximately represented in a low dimensional feature or shape vector space [7] instead of the original high dimensional data vector space. Given a training data set of N surfaces, $Q_i(i = 1, \ldots, N)$, each containing M points on its surface $Q_i \in R^{3M}$, where Q_i are vertices of the $i - th$ surface, and assuming that vertices on each surface Q_i are in correspondence. The first step of PCA is to construct the mean shape vector $\overline{Q}$ for the training data set.

$$\overline{Q} = \frac{1}{N} \sum_{i=1}^{N} Q_i \qquad (1)$$

Let C be defined as the covariance matrix calculated from the training data set

$$C = \frac{1}{N} \sum_{i=1}^{N} (Q_i - \overline{Q})(Q_i - \overline{Q})^T \qquad (2)$$

By building a matrix X of "centered" shape vectors with $Q_i - \overline{Q}$ as the $i - th$ column of matrix X, the covariance matrix C is defined as

$$C = X \cdot X^T \qquad (3)$$

where matrix C has $3M$ rows and columns. Since the number of surfaces N in the training data set is smaller than the size of data vector, the eigen decomposition of matrix C is performed by finding eigen decomposition of matrix $C' = X^T X$. The first N largest eigenvalues $\lambda_i(i = 1, \ldots, N)$ and eigenvectors $p_i(i = 1, \ldots, N)$ of the original covariance matrix C are then determined respectively by

$$\lambda_i = \lambda'_i \qquad (4)$$

$$p_i = \frac{X \cdot p'_i}{|X \cdot p'_i|} \qquad (5)$$

where λ'_i and p'_i are respectively eigenvalues and eigenvectors of matrix C'. By using these eigenvalues and eigenvectors, each shape in the training data set can be approximately represented using a linear model of the form

$$\hat{Q}_i = Pb_i + \overline{Q} \qquad (6)$$

where the so called "Shape Matrix" $P = [p_1, \ldots, p_K]$ is $3M \times K$ matrix of K eigenvectors associated with the first K largest eigenvalues, "modes of variation", of matrix C. The shape space vector $b_i(i = 1, \ldots, N)$ controls the contribution of each mode of variation in the approximation surface [7].

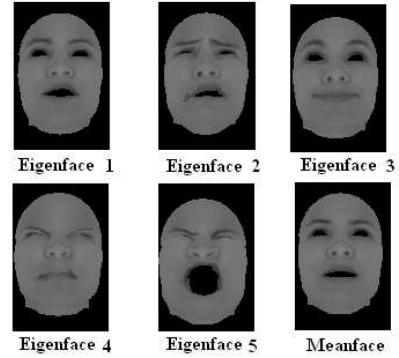**Figure 1. First 20 training faces**

Figure 2. 5 eigenfaces and mean face

In order to obtain a correct statistical model, points on each surface from the training data set must be in correspondence. This is essential, since incorrect correspondence can either introduce too much variation or lead to illegal instances of the model. As $FaceGenModeller^{\circledR}$ provides the correspondence information for the training-faces, therefore its estimation for simulated data set is

not needed. The database consisted of 100 3-D faces with various expression from the same individual. Figure 1 shows the first 20 faces in the database produced by $FaceGenModeller^{®}$. Figure 2. shows the mean face and five other faces obtained as a superposition of the mean face and five largest eigenfaces. These five largest eigenfaces capture about 98.5% of the variation of the facial structure in the database

3. Modified ICP Registration

The ICP was first proposed by Besl and McKay [6], and it is most widely used point-based surface matching algorithm. In the proposed method, a modified ICP registration process is used to establish the correspondence between a new surface and the statistical model built using the method described previously. Two major phases are involved in the modified ICP, which are initial and fine alignment.

3.1. Initial Alignment

In this step, the classical ICP method is used to achieve initial alignment between the new surface and the mean shape of the statistical model. The whole procedure iteratively refines the alignment by alternately choosing corresponding points and finding the best similarity transformation that minimizes a cost function based on the distance between the corresponding points. The cost function is defined by

$$E = \sum_{i=1}^{N} \|p_i' - (SRp_i + T)\|^2 \qquad (7)$$

where $p_i'(i = 1, \ldots, N)$ is the corresponding points in the new surface which are associated with the mean shape of the statistical model, and $p_i(i = 1, \ldots, N)$ is the mean shape of statistical model (here, p_i' and p_i are considered as 3×1 vectors). R is a 3×3 rotation matrix, T is a translation vector (3×1 column matrix) and S is a scaling factor. According to the algorithms in [2][14], R, T and S are calculated as follow:

1. From the point sets, $\{p_i\}$ and $\{p_i'\}$ ($i = 1, \ldots, N$), compute the mean vectors, $\overline{p}$ and $\overline{p}'$, using

$$\overline{p} = \frac{1}{N} \sum_{i=1}^{N} p_i \qquad (8)$$

$$\overline{p}' = \frac{1}{N} \sum_{i=1}^{N} p_i' \qquad (9)$$

2. Calculate q_i and q_i' ($i = 1, \ldots, N$) using

$$\overline{q} = p_i - \overline{p} \qquad (10)$$

$$\overline{q}' = p_i' - \overline{p}' \qquad (11)$$

3. Calculate the matrix H using

$$H = \sum_{i=1}^{N} q_i q_i'^T \qquad (12)$$

4. Find the SVD of H

$$H = USV^T \qquad (13)$$

5. Calculate

$$X = VU^T \qquad (14)$$

6. Compute the rotation matrix using

$$R = USV^T \qquad (15)$$

$$S = \begin{cases} I & \textbf{If } det(X) = +1 \\ diag(1, \ldots, 1, -1) & \textbf{If } det(X) = -1 \end{cases} \qquad (16)$$
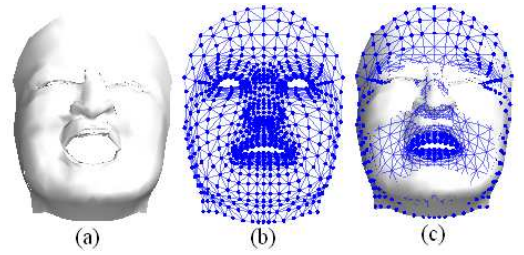
7. Find the translation vector and scaling factor using

$$T = \overline{p}' - R\overline{p} \qquad (17)$$
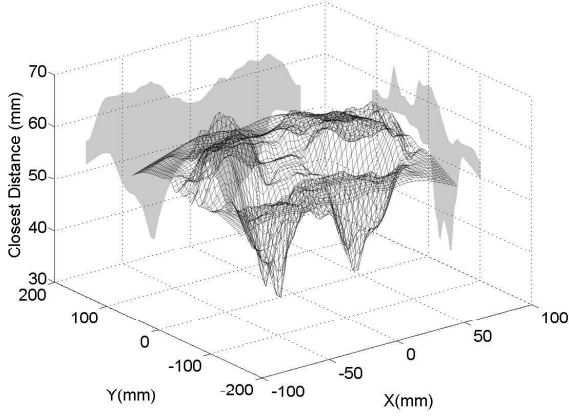
$$S = \frac{tr(PP'^T R)}{tr(PP^T)} \qquad (18)$$

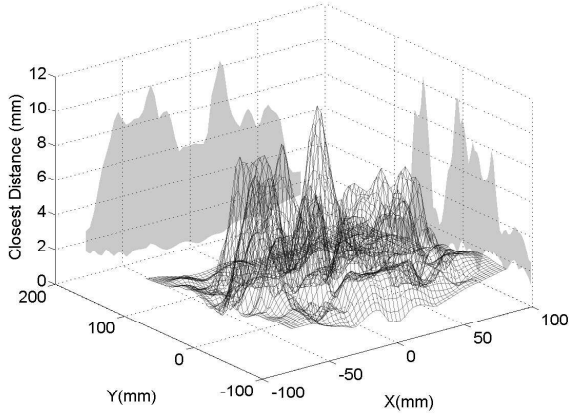where $P = [p_1^T, \ldots, p_N^T]^T$ and $P' = [p_1'^T, \ldots, p_N'^T]^T$ are $3 \times N$ matrix.

The result of the initial alignment of a new face and the mean shape is shown in Figure 3. Figure 3(a) shows the shape of the new surface which has extreme fear expression, and the mesh in the middle of the figure indicates the mean shape of the statistical model. Figure 3(c) illustrates the new surface and transformed mean shape. Although they are locally misaligned, due to different facial expressions (see the mouth and eyes regions), they are globally well aligned.

(a)   (b)   (c)

Figure 3. Result for initial alignment: (a) new face, (b) mean Shape, (c) alignment of mean shape and new face

Figure 4. Closest distance distribution before alignment

Figure 5. Closest distance distribution after alignment

The closest points distance distribution before and after the initial alignment are shown in Figure 4 and Figure 5, respectively. It can be seen that the maximum closest distance between the new surface and mean shape is reduced from around $60mm$ (prior to the initial alignment) to $10mm$ (after initial alignment), and the average closest distance is reduced from $54mm$ to $6mm$.

3.2. Fine Alignment

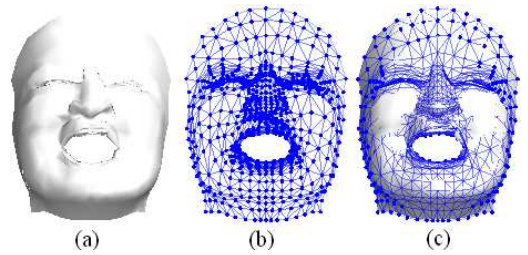Due to the existing deformation between the new surface and the mean shape, the classical ICP can only provide an initial alignment. The fine alignment performs least squares projection of the new surface onto the low-dimensional shape space. The ICP and the least squares projection are preformed iteratively in turn till the alignment error can not be reduced any more. Effectively this process is equivalent to correspondence search. The details of the fine alignment algorithm are explained below.

1. Starting with the initial alignment, the similarity transformation and the corresponding points between new data and the mean shape are calculated by classical ICP.

2. $Q \in R^{3M}$ is a vector of M 3-D points representing new face which is projected onto the shape space. The mean vector $\overline{Q}$ and shape matrix P are obtained from Equation 1.The new shape space vector b is calculated using the least-squares method.
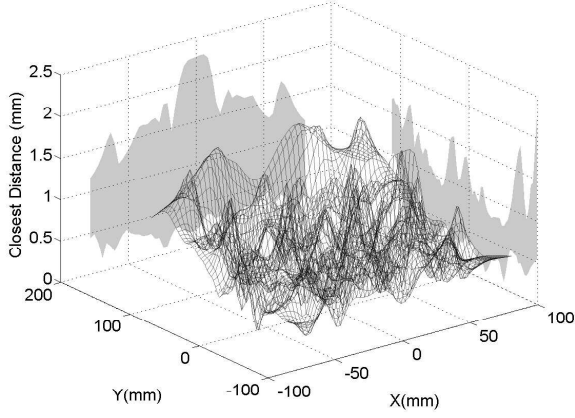
$$b = P^T(Q - \overline{Q}) \qquad (19)$$

3. A deformed new surface is generated using Equation 6.

4. Repeat above steps until the rate of change for RMS distance of closest points falls below a preset threshold τ.

The fine alignment achieves the deformable registration between the new surface and the mean shape which is shown in Figure 6. It shows that the mean shape is deformed in Figure 6(b) and well aligned with the new surface not present in the training set in Figure 6(c). The corresponding distribution of the closest point distance is shown in Figure 7. After fine alignment maximum closest distance is just above $2mm$ with average closest distance of $1.24mm$.

(a)   (b)   (c)

Figure 6. Result for fine alignment: (a) new face, (b) aligned mean shape, (c) alignment of mean shape and new face

Figure 7. Closest distance distribution for fine alignment

4. Experiment and Results

Two types of data are used in the experiments. The first one is the simulated data generated by $FaceGenModeller^{®}$ containing various facial expression for the same individual. The second is the measured data of a face with neutral expression captured by a $3dMD^{®}$[1] 3-D scanner.
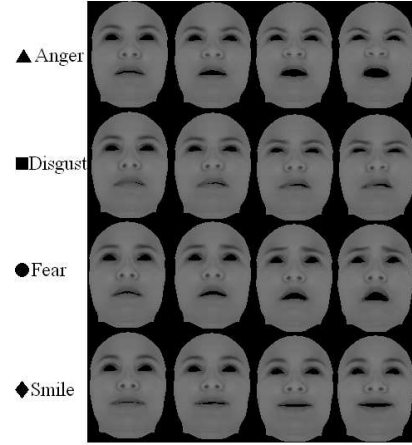
4.1. Simulated Data

The simulated test data consists of 16 faces from the same individual and divided into 4 different groups of expression which are anger, disgust, fear and smile. Figure 8 illustrates these various expressions. The statistical model used is described in Section 2, the eigenfaces and the mean face are shown in Figure 2. After modified ICP registration, the shape space vector of each face is given. Each shape space vector contains 5 coefficients $b_1, \ldots, b_5$. In order to graphically represent the obtained shape space vectors, the first 3 coefficient b_1, b_2 and b_3 are plotted in Figure 9 and Figure 10. It can be seen that the first three elements of the shape space vectors for different expressions occupy well defined disjoint parts of the space and are approximately located on the same line defined by the expression type. This is an indication that shape space vectors can be potentially used for classification of facial expressions.
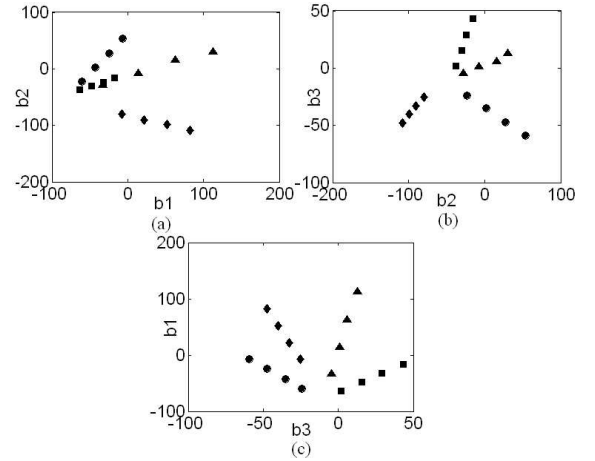
4.2. Measurement Data

In order to test that the proposed method can also be used with real measurement data, a human face with neutral expression has been captured using the $3dMD$ scanner. The
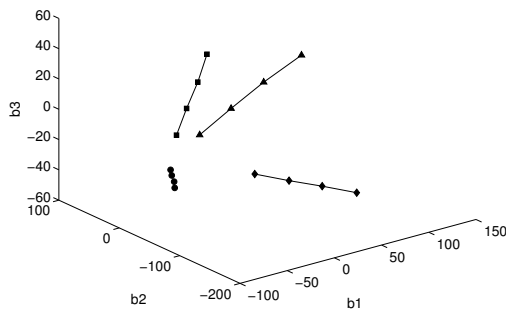
acquired data is shown in Figure 11(a). This time, the statistical model is built using 100 faces of different individuals with all having neutral facial expressions. As before the data is generated by the $FaceGenModeller^{®}$ software. Figure 11 illustrates the result of applying the modified ICP to real data. Figure 11(e)~(i) show the spatial distribution of points based on the value of the corresponding closest distance. The average closest distance from the registered mean shape to the measurement data is around $6mm$. The average closest distance between the mean shape and the initially aligned real data is about $19mm$.
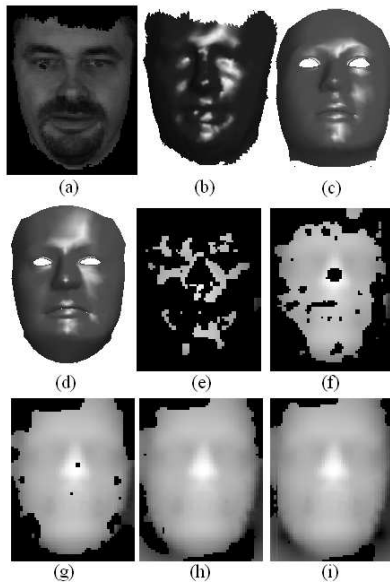

Figure 8. Simulated test faces

Figure 9. Representation of the first three elements of the shape space vector for test faces from Figure 8(a) b_1 **Vs** b_2**, (b)** b_2 **Vs** b_3**, (c)** b_3 **Vs** b_1

Figure 10. All three elements of vector b from Figure 9 displayed on 3D graph

Figure 11. Results for real measurement data: (a) texture of real data, (b) surface of real data, (c) mean shape, (d) registered mean shape, (e) closest distance$\leq 2mm$, (f) $\leq 4mm$, (g) $\leq 6mm$, (h) $\leq 8mm$, (i) $\leq 10mm$

5. Conclusions

Initial results from a novel low dimensional surface parameterizations method have been described. The method is based on combination of the iterative closest point (ICP) method and projections onto the shape space defined by the principal component analysis (PCA). The proposed method is capable of efficiently representing a surface using a small number of coefficients in a shape space vector rather than a large number of data points. The shape space vector is shown to provide the features that can be used for classification of facial expression. Further improvements of the method include building a hierarchical system for face recognition where firstly face type is decided upon and subsequently the facial expression is searched using shape space extracted from the facial expression database constructed for a specific face type. Additionally instead of using directly data points (surface vertices) to build the face space, the use of control points of the B-spline surface representation in construction of the shape space will be investigated.

References

[1] 3dMD, www.3dMD.com. *3dMD 3D Scanner*, 2006.
[2] K. S. Arun and T. S. Huang. Least-square fitting of two 3-d point sets. *IEEE TVGC.*, 9(5):698–700, 1987.
[3] A. T. Basilevsky. *Statistical Factor Analysis and Related Methods: Theory and Applications.* Wiley Interscience, 1994.
[4] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. PAMI.*, 24(4):509–522, 2002.
[5] P. J. Besl and R. C. Jain. Invariant surface characteristics for 3d object recognition in range images. *Comput. Vision Graph. Image Process*, 33(1):33–80, 1986.
[6] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Trans. PAMI.*, 14(2):239–256, 1992.
[7] A. Blake and M. Isard. *Active Contours.* Springer-Verlag Berlin and Heidelberg, 1998.
[8] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst.*, 89(2-3):114–141, 2003.
[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models - their training and application. *Comput. Vis. Image Underst.*, 61(1):38–59, 1995.
[10] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Trans. PAMI*, 21(5):433–449, 1999.
[11] H. Qin and D. Terzopoulos. D-nurbs: A physic-based framework for geometric design. *IEEE TVGC.*, 2(1):85–96, 1996.
[12] D. F. Rogers. *A Introduction to NURBS with Historical Perspective.* Academic Press, 2001.
[13] Singular Inversions, www.facegen.com. *FaceGen Modeller*, 2003.
[14] S. Umeyama. Least-square estimation of transformation parameters between two point patterns. *IEEE TVGC.*, 13(4):376–380, 1991.