THREE-DIMENSIONAL FACE RECOGNITION: AN EIGENSURFACE APPROACH

Thomas Heseltine, Nick Pears, Jim Austin

Advanced Computer Architecture Group, Department of Computer Science, The University of York

ABSTRACT

We evaluate a new approach to face recognition using a variety of surface representations of three-dimensional facial structure. Applying principal component analysis (PCA), we show that high levels of recognition accuracy can be achieved on a large database of 3D face models, captured under conditions that present typical difficulties to more conventional two-dimensional approaches. Applying a range of image processing techniques we identify the most effective surface representation for use in such application areas as security, surveillance, data compression and archive searching.

1. INTRODUCTION

Despite significant advances in face recognition technology, it has yet to achieve the levels of accuracy required for many commercial and industrial applications, mainly due to the inaccuracies caused by the environmental circumstances under which images are captured. Variation in lighting, facial expression and orientation all significantly increase error rates, making it necessary to maintain consistent conditions between query and gallery images for the system to function adequately. However, this approach eliminates some of the key advantages offered by face recognition: a passive biometric in the sense that it does not require subject co-operation.

The use of 3D face models is motivated by a number of factors. Firstly, by relying purely on geometric shape, rather than colour and texture information, we render the system invariant to lighting conditions. Secondly, the ability to rotate a facial structure in three-dimensional space, allowing for compensation of variations in pose, aids those methods requiring alignment prior to recognition. Finally, additional discriminating information is captured, when compared with 2D systems. As an example, eye separation can be recovered from both sets of data, but nose depth can only easily be recovered from three-dimensional data. We do recognise, however, that two-dimensional colour-texture information provides a rich source of discriminatory information, which is forfeit if three-dimensional data alone is used. Therefore, the focus of this paper is to determine the ability of three-dimensional data alone to form the basis of a face recognition system, as compared to 2D systems. Additional research can then identify methods of reintroducing normalised two-dimensional texture data in order to reduce error rates further.

We investigate the use of facial surface data, taken from 3D face models, as a substitute for the more familiar two-dimensional images. We take a well-known method of face recognition, namely the eigenface approach described by Turk and Pentland [1] and adapt it for use on three-dimensional data. Testing a range of surface representations and distance metrics, we identify the most effective methods of recognising faces using three-dimensional surface structure.

In order to test this method of face recognition, we require a large database of 3D face models. However, until recently, 3D capture methods have been slow and cumbersome, requiring the subject to remain perfectly still. For these reasons, three-dimensional face recognition has remained relatively unexplored, when compared to the wealth of research focusing on two-dimensional face recognition. Although some investigations have experimented with 3D data [2, 3, 4], they have had to rely on small test sets of 3D face models or used generic face models to enhance two-dimensional images prior to recognition [5, 6, 7]. However, this research demonstrates that the use of three-dimensional information has the potential to improve face recognition well beyond the current state of the art. With the emergence of new 3D capture equipment, the population of a large 3D face database has now become viable and being undertaken at The University of York as part of a project facilitating research into three-dimensional face recognition technology [8].

2. RELATED WORK

In this section, we discuss previous research exploring the possibilities offered by three-dimensional geometric structure to perform face recognition. To date, the majority of research has focused on two-dimensional images, although some have attempted to use a-priori knowledge of facial structure to enhance these existing two-dimensional approaches. For example, Zhao and Chellappa [5] use a generic 3D face model to normalise facial orientation and lighting direction in two-dimensional images. Using estimations of light source direction and pose, the 3D face model is aligned with the two-dimensional face image and used to project a prototype image of the frontal pose equivalent, prior to recognition by linear discriminant analysis. Recognition accuracy on the test set is increased from approximately 81% (correct match within rank of 25) to 100%. Similar results are witnessed in the Face Recognition Vendor Test [9], showing that pose correction using Romdhani, Blanz and Vetter's 3D morphable model technique [6] reduces error rates when applied to the FERET database.

Blanz, Romdhani and Vetter [7] take a comparable approach, using a 3D morphable face model to aid in identification of 2D face images. Beginning with an initial estimate of lighting direction and face shape, Romdhani et al iteratively alter shape and texture parameters of the morphable face model, minimising difference to the two-dimensional image. These parameters are then taken as features for

0-7803-8554-3/04/$20.00 ©2004 IEEE.

identification, resulting in 82.6% correct identifications on a test set of 68 people.

Although these methods show that knowledge of three-dimensional face shape can aid normalisation for two-dimensional face recognition systems, none of the methods mentioned so far use actual geometric structure to perform recognition. Whereas Beumier and Acheroy [2] make direct use of such information, testing various methods of matching 3D face models, although few were successful. Curvature analysis proved ineffective, and feature extraction was not robust enough to provide accurate recognition. However, Beumier and Acheroy were able to achieve reasonable error rates using curvature values of vertical surface profiles. Verification tests carried out on a database of 30 people produced equal error rates (EER) between 7.25% and 9.0% on the automatically aligned surfaces and between 6.25% and 9.5% when manual alignment was used.

Hesher et al [10] take a similar approach to our base method, using PCA of range images and euclidean distance to perform recognition. Matching 37 range images produces a correct identification rate of 94%, when training is performed on the gallery set. However, it is not demonstrated how successful the system is when the training and test set are disjoint and no other surface representations are tested.

Chua et al [4] take a different approach, applying non-rigid surface recognition techniques to the face structure. An attempt is made to identify and extract rigid areas of facial surfaces, creating a system invariant to facial expression. The characteristic used to identify these rigid areas and ultimately distinguish between faces is the point signature, which describes depth values surrounding local regions of specific points on the facial surface. The similarity of two face models is computed by identifying and comparing a set of unique point signatures for each face. Identification tests show that the probe image is identified correctly for all people when applied to a test set of 30 depth maps of 6 different people.

Another method, proposed by Gordon [3], incorporates feature localisation. Using both depth and curvature information extracted from three-dimensional face models, Gordon identifies a number of facial features, including head width, nose dimensions and curvatures, distance between the eyes and eye width. These features are evaluated using fisher's linear discriminant, determining the discriminating ability of each individual feature. Findings show head width and nose location are particularly important features for recognition, whereas eye widths and nose curvatures are less useful. Recognition is performed by means of a simple euclidean distance measure in feature space. Several combinations of features are tested using a database of 24 facial surfaces taken from 8 different people, producing results ranging from 70.8% to 100% correct matches.

3. THE 3D FACE DATABASE

As mentioned previously, there is little three-dimensional face data publicly available at present and nothing towards the magnitude of data required for development and testing of three-dimensional face recognition systems. Therefore, we introduce a new database of 3D face models, collected at The University of York, as part of an ongoing project to provide a publicly available 3D Face Database of over 5000 models [8]. The 3D

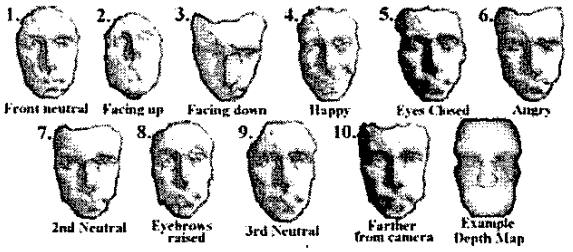


Figure 1. Face models taken from the UOY 3D face database

models are generated using a 3D camera, which operates on the basis of stereo disparity of a high-density projected light pattern. For the purpose of these experiments, we will be using a subset of the 3D face database, acquired during preliminary data acquisition sessions. This set consists of 330 models taken from 100 people under the ten conditions shown in figure 1.

During capture no effort was made to control lighting conditions. In order to generate face models at various head orientations, subjects were asked to face reference points positioned roughly 45° above and below the camera, but no effort was made to enforce a precise angle of orientation.

3D face models are orientated to face directly forwards using our orientation normalisation algorithm (not described here) before being converted into depth maps. The database is then separated into two disjoint sets: the training set consisting of 40 depth maps (type 1, figure 1) and a test set of the remaining 290 depth maps. Both sets contain subjects of various race, age and gender and nobody is present in both the training and test sets.

4. SURFACE REPRESENTATIONS

In previous work we have shown that the use of image processing techniques can significantly reduce error rates of two-dimensional face recognition methods [11, 12], by removing unwanted effects caused by environmental capture conditions. Much of this environmental influence is not present in the 3D face models, but pre-processing may still aid recognition by making distinguishing features more explicit. We test a number of surface representations, which may affect recognition error rates, derived by pre-processing of depth maps, prior to both training and test procedures, as described in table 1.

5. DEFINING SURFACE SPACE

We define surface space by application of PCA to the training set of facial surfaces, taking a similar approach to that described by Turk and Pentland [1] and used in previous investigations [11, 12]. Consider our training set of facial surfaces, stored as orientation normalised 60x105 depth maps, represented as vectors of length 6300. We begin by reducing dimensionality to a practical value, while maximising the spread of facial surfaces within the subspace, by application of PCA to the training set of M (40) depth maps $\{\Gamma_1, \Gamma_2, \ldots \Gamma_M\}$, computing the covariance matrix,

$$C = \frac{1}{M}\sum_{n=1}^{M} \Phi_n \Phi_n^T \qquad A = [\Phi_1 \Phi_2 \Phi_3 \ldots \Phi_M]$$

$$= AA^T \qquad \Phi_n = \Gamma_n - \Psi$$

$$\Psi = \frac{1}{M}\sum_{n=1}^{M} \Gamma_n$$

Where Φ_n is the difference of the *nth* depth map from the average ψ. Eigenvectors and eigenvalues of the covariance

Figure 2. The average surface (left) and first eight eigensurfaces

matrix are calculated using standard linear methods. The resultant eigenvectors describe a set of axes within the depth map space, along which most variance occurs and the corresponding eigenvalues represent the degree of this variance along each axis. The M eigenvectors are sorted in order of descending eigenvalues and the M'(40) greatest eigenvectors chosen to represent surface space. We term each eigenvector an eigensurface, displayed as range images in figure 2.

6. VERIFICATION OF FACIAL SURFACES

Once surface space has been defined, we project any face into surface space by a simple matrix multiplication, using the eigenvectors calculated from covariance matrix C.

$$\omega_k = u_k^T(\Gamma - \Psi) \; for \; k = 1 \ldots M'.$$

Where u_k is the *kth* eigenvector and ω_k is the *kth* weight in the vector $\Omega^T = [\omega_1, \omega_2, \ldots \omega_{M'}]$. The vector Ω is taken as the 'face-key' representing a person's facial structure in surface space and compared by either euclidean or cosine distance metrics.

$$d_{euclidean} = \|\Omega_a - \Omega_b\| \qquad d_{cosine} = 1 - \frac{\Omega_a^T \Omega_b}{\|\Omega_a\| \|\Omega_b\|}$$

In addition, we can also divide each face-key by its respective eigenvalues, prior to distance calculation, removing any inherent dimensional bias and introducing two supplementary metrics, the Mahalanobis distance and weighted cosine distance. An acceptance (facial surfaces match) or rejection (surfaces do not match) is determined by applying a threshold to the calculated distance. In order to evaluate the effectiveness of the face recognition methods, we compare each of the 290 surfaces in the test set with every other surface (41,905 verification operations). False acceptance rates (FAR) and false rejection rates (FRR) are calculated as the percentage of incorrect acceptances and rejections after applying a threshold. Varying the threshold produces a series of FAR, FRR pairs, which plotted on a graph produces an error rate curve (figure 3), from which the EER (where FAR equals FRR) is taken as a single comparative value.

7. RESULTS

Results are presented as error rate curves and bar charts of EERs (figure 3). The results clearly show that dividing by eigenvalues to normalise vector dimensions prior to distance calculations, significantly decreases error rates for both euclidean and cosine distance, with the Mahalanobis metric providing the lowest EER for the depth map system. The EERs produced show that surface gradient representations provide the most distinguishing information, with horizontal derivatives giving the lowest EERs of all, using the weighted cosine distance metric. In fact, the weighted cosine distance returns the lowest EER for the majority of surface representations, excepting a few particular cases. However, the most effective surface representation seems to be dependent on the distance metric used for comparison.

8. CONCLUSION

We have shown that a well-known two-dimensional face recognition method can be adapted for use on three-dimensional face models. Tests have been carried out on a large database of three-dimensional facial surfaces, captured under conditions that present typical difficulties when performing recognition. The error rates produced from the three-dimensional baseline system (19.1% EER using euclidean distance) are notably lower that those gathered in similar experiments using two-dimensional images (25.5% EER) [12]. Although a more direct comparison is required, using a common 2D/3D test database, in order to draw any quantitive conclusions, initial results suggest that three-dimensional face recognition has distinct advantages over conventional two-dimensional approaches.

Experimenting with a number of surface representations, we have discovered that facial surface gradient is more effective for recognition than depth and curvature representations. In particular, horizontal gradients produce the lowest error rates, seeming to indicate that horizontal derivatives provide more discriminatory information than vertical. Another advantage is that gradients are likely to be more robust to inaccuracies in the alignment procedure, as the derivatives will be invariant to translations along the Z-axis.

Curvature representations do not seem to contain as much discriminatory information as other surface representations. We find this surprising, as second derivatives should be less sensitive to inaccuracies of orientation and translation along the Z-axis. However, this could be a reflection of inadequate 3D model resolution and high noise content.

Testing four distance metrics has shown that the choice of method for face-key comparisons has a considerable affect on resulting error rates. The euclidean and cosine measures seem tailored to specific surface representations, suggesting that some create a surface space in which between-class deviation is predominantly angular, whereas others produce more radial deviation. It is also evident that dividing each face-key by respective eigenvalues, normalising dimensional distribution, usually improves results for both euclidean and cosine distances. This indicates that the distribution along one surface space dimension is not necessarily proportional to its discriminating ability and that face-keys become more discriminative when all dimensions are weighted evenly. However, this is not the case for some surface representations with higher EERs, suggesting that these representations incorporate only a few dominant useful components, which become masked when weighted evenly with the majority of less discriminatory components.

Error rates of the optimum 3D eigensurface system (12.7% EER) are substantially lower than the best two-dimensional systems (20.4% EER and 17.8% EER) tested under similar circumstances in our previous investigations [11, 12]. Although we recognise the differences between these experiments (most notably the lack of a common 3D/2D test set), the results do show that geometric face structure is useful for recognition when used independently from colour and texture and capable of achieving high levels of accuracy. Given that the data capture method produces models invariant to lighting conditions and provides the ability to recognise faces regardless of pose, makes this system particularly attractive for use in security and surveillance applications. However, more testing is required to identify the limitations of the eigensurface method, although one obvious issue is the system's reliance on accurate orientation normalisation. A better approach would be to generate a surface representation that was invariant to orientation.

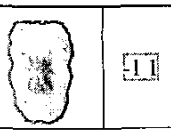
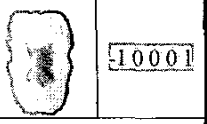
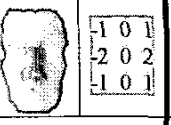
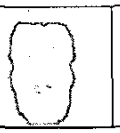
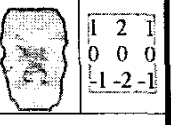
Horizontal Grad.	Horizontal Grad. Large	Horizontal Curvature	Sobel Magnitude	Sobel X	Curvature
					
Applies the 2x1 kernel to compute horizontal derivative	Horizontal gradient over a greater horizontal distance	Applies sobel X twice to calculate the 2nd horizontal derivative	The magnitude of Sobel X and Y combined.	Application of the horizontal sobel derivative filter	The magnitude of the vertical and horizontal curvatures
Vertical Gradient	**Vertical Grad. Large**	**Vertical Curvature**	**Laplacian**	**Sobel Y**	**Curve Type**
					
Applies the 1x2 kernel to compute the vertical derivative	Vertical gradient over a greater vertical distance	Applies sobel Y twice to calculate the 2nd vertical derivative	An isotropic measure of the second spatial derivative	Application of the vertical sobel derivative filter	Segmentation of the surface into 8 discreet curvature types

Table 1. Brief descriptions of surface representations with the convolution kernels used.

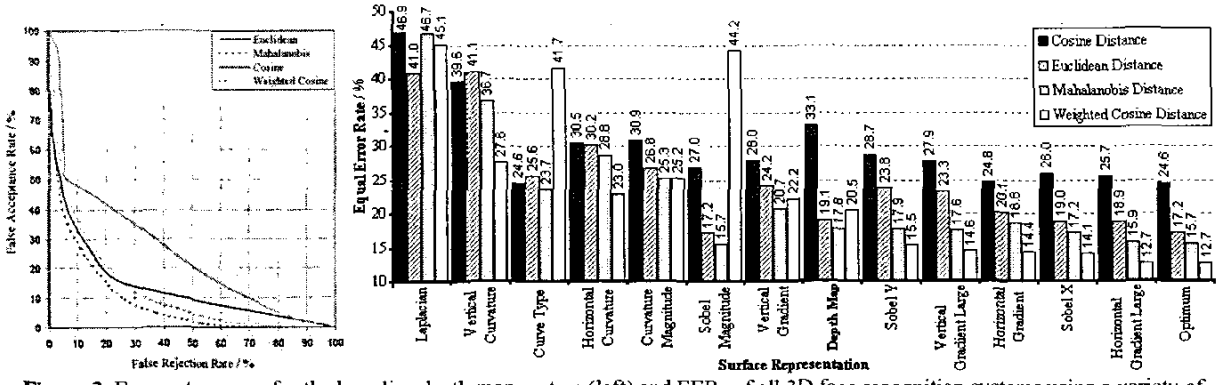


Figure 3. Error rate curves for the base line depth map system (left) and EERs of all 3D face recognition systems using a variety of surface representations and distance metrics (right)

11. REFERENCES

1. M. Turk, A. Pentland, "Face Recognition Using Eigenfaces" Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 586-591 (1991)

2. C. Beumier, M. Acheroy, "Automatic 3D Face Authentication" Image and Vision Computing, vol. 18, no. 4, pp. 315-321 (2000)

3. G. Gordon, "Face Recognition Based on Depth and Curvature Features" Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 108-110 (1992)

4. C. Chua, F. Han, T. Ho, "3D Human Face Recognition Using Point Signature" Proc. 4th IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 233-238 (2000)

5. W. Zhao, R. Chellappa, "3D Model Enhanced Face Recognition" Proc. Int. Conf. on Image Processing (2000)

6. S. Romdhani, V. Blanz, T. Vetter, "Face Identification by Fitting a 3D Morphable Model using Linear Shape and Texture Error Functions" The European Conf. on Computer Vision (2002)

7. V. Blanz, S. Romdhani, T. Vetter, "Face Identification across Different Poses and Illuminations with a 3D Morphable Model" Proc. of the 5th IEEE Conf. on AFGR (2002)

8. "The UOY 3D Face Database" www.cs.york.ac.uk/~tomh

9. P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, J. Bone, "FRVT 2002: Overview and Summary" http://www.frvt.org/FRVT2002/documents.htm, (2003)

10. C. Hesher, A. Srivastava, G. Erlebacher, "Principal Component Analysis of Range Images for Facial Recognition" Proc. CISST (2002)

11. T. Heseltine, N. Pears, J. Austin, "Evaluation of image pre-processing techniques for eigenface-based face recognition" Proc. 2nd Int. Conf. on Image and Graphics, SPIE vol. 4875 pp. 677-685 (2002)

12. T. Heseltine, N. Pears, J. Austin, "Face Recognition: A Comparison of Appearance-based Approaches" Proc. 7th Int. Conf. on Digital Image Computing: Techniques and Applications, vol. 1 pp. 59-69 (2003)