2D-3D Hybrid Face Recognition Based on PCA and Feature Modelling

Chris McCool, Vinod Chandran, Sridha Sridharan
Image and Video Research Laboratory
Queensland University of Technology,
2 George Street, Brisbane, Australia, 4001, GPO 2434
{c.mccool, v.chandran, s.sridharan}@qut.edu.au

Abstract

Hybrid face recognition, using image (2D) and structural (3D) information, has explored the fusion of Nearest Neighbour classifiers. This paper examines the effectiveness of feature modelling for each individual modality, 2D and 3D. Furthermore, it is demonstrated that the fusion of feature modelling techniques for the 2D and 3D modalities yields performance improvements over the individual classifiers. By fusing the feature modelling classifiers for each modality with equal weights the average Equal Error Rate improves from 12.60% for the 2D classifier and 12.10% for the 3D classifier to 7.38% for the Hybrid 2D+3D clasiffier.

1. Introduction

Face recognition is an area of biometric research that has received great interest in recent times. The Face Recognition Grand Challenge (FRGC) [11] highlighted research directions still left to be explored in face recognition. An area which remains unsolved is the fusion of 2D and 3D face data for recognition; however, it's noted that 2D and 3D face recognition still have their respective challenges.

Face recognition using image data (2D face recognition) is a mature field that has been researched for over 20 years. Recent research has made use of Principal Component Analysis (PCA) [14], the 2D Discrete Cosine Transform (2D-DCT) [10] and morphable models [1], among others. The *de facto* standard for 2D face recognition is *eigenfaces* which is a technique that applies PCA to an image and was introduced to the field by Turk and Pentland [14]. The 2D-DCT based technique proposed in [10] models the 2D-DCT features using Gaussian Mixture Models (GMMs). In this work a background, or global, GMM is used to derive client dependant models and verification is performed using the difference of these two scores (from scoring against the background and client models). The 2002 Face Recognition Vendor Test [12] highlighted that illumination and pose variation remain unsolved problems for 2D face recognition. In addition to this there are still problems associated with expression variation.

Several methods for conducting 3D face recognition have been proposed. Recent work into 3D face recognition has proposed the use of point signatures [16], isometric transformations [3] and PCA [4]. Some recent work has explored the modelling of 3D data, in [5] the registration error from the Iterative Closest Point algorithm was modelled with GMMs. Unlike 2D face recogntion illumination and pose variations do not affect 3D face recognition, however, the problem of expression variation remains largely unsolved. For a more in depth review of 3D face recognition readers are referred to [2].

Fusing the 2D and 3D modalities for face recognition is receiving greater attention because of an increased interest in multi-modal biometric recognition. Combining the two modalities, 2D and 3D, can be broadly divided into data fusion, feature fusion and classifier fusion. This work and much of the current work into fused 2D and 3D face recognition explores methods for classifier fusion; in particular current research has concentrated on performing Nearest Neighbour (NN) classifier fusion. In [4] separate 2D and 3D classifiers were formed using PCA-based features, a hybrid 2D+3D classifier was then formed by fusing the individual classifiers with linear weights based on the Rank scores for identification. A technique using log-Gabor features on 2D and 3D part face images was described in [6]. The log-Gabor features were reduced with PCA and the 2D and 3D classifiers were fused with equal weighted fusion. Despite the ongoing research into hybrid 2D and 3D face recognition, work into fusing feature modelling techniques for 2D and 3D face recognition remains unaddressed.

The work presented in this paper examines the use of feature modelling as a method for improving 2D, 3D and hybrid 2D+3D face recognition. Initial experiments demonstrate that feature modelling is a more effective technique than several standard NN classifiers; this is conducted for

both 2D and 3D modalities. Feaure modelling is conducted using GMMs to describe Intra-Personal (IP) and Extra-Personal (EP) difference vectors. The fusion of the two modalities, 2D and 3D, is approached as classifier fusion and it is shown that weighted summation provides consistent improvements when fusing two feature modelling techniques from different modalities.

The paper is structured as follows. Related work in 3D face recognition and 2D PCA-based face recognition will be presented in Section 2. Following this methods for data normalisation and feature modelling will be described in Sections 3 and 4 respectively. The experimental procedure is outlined in Section 5 and the results are presented in Section 6. Conclusions and future work are then presented and discussed in Section 7.

2. Related Work

Face recognition broadly refers to the tasks of face verification and face identification. Face verification consists of comparing an input image claiming an ID against the stored image of the claimed ID whilst face identification consists of finding the best matching image (in a database of images) given the input image. The experiments conducted in this work perform face verification. A flowchart of the process of verification for 2D face recognition is provided in Figure 1.

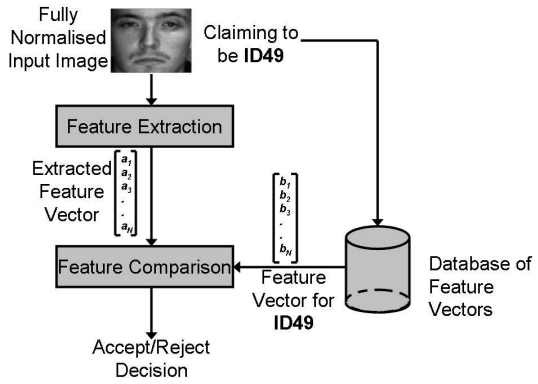

Figure 1. A flowchart describing the verification process for the 2D modality.

For this work PCA was chosen as the feature extraction technique. This has previously been applied to 2D face recognition in [14] and to 3D face recognition in [4]. The feature modelling technique chosen is GMMs which has been used with 2D-DCTs in [10] and is also a standard technique used in speaker recognition [13]. Only one feature extraction technique is being used for this work to ensure that when multi-modal fusion is conducted the complementary information, from the two modalities, is not being derived from the use of alternate feature extraction techniques. This is also the reason for keeping the feature modelling technique constant. For a detailed explanation of GMMs readers are directed to [13].

2.1. PCA-Based Face Recognition Metrics

Turk and Pentland first introduced *eigenfaces* to the field of face recognition in [14]. This technique applies eigenvalue decomposition to the covariance matrix of a training set of M vectorised images. Each image, $\boldsymbol{x}_j$, consists of N pixels and the average image (ψ) is calculated from the M training images. The mean normalised images are represented by,

$$\overline{\boldsymbol{x}}_j = \boldsymbol{x}_j - \psi, \tag{1}$$

and then used to derive the covariance matrix,

$$C = \frac{1}{M}\sum_{j=1}^{M} \overline{\boldsymbol{x}}_j \overline{\boldsymbol{x}}_j^T, \tag{2}$$

upon which eigenvalue decomposition is applied. The eigenvalues found by this procedure represent the variance of the corresponding eigenvectors. Therefore the larger eigenvalues represent those eigenvectors which describe the most variation of the training data.

A review of 2D PCA-based face recognition distance metrics was conducted in [17]. This work found the Mahalanobis metric consistently outperformed L_1, L_2 and cosine metrics. In [4] the Mahalanobis metric outperformed other metrics for 3D face data. A description of each of these metrics can be found below:

- L_1 Metric (Manhattan Distance)

$$d(\boldsymbol{x}, \mathbf{y}) = \|\boldsymbol{x} - \mathbf{y}\|_1 \tag{3}$$

- L_2 Metric (Euclidian Distance)

$$d(\boldsymbol{x}, \mathbf{y}) = \|\boldsymbol{x} - \mathbf{y}\|_2 \tag{4}$$

- Cosine Metric

$$d(\boldsymbol{x}, \mathbf{y}) = \frac{\boldsymbol{x}.\mathbf{y}}{|\boldsymbol{x}||\mathbf{y}|} \tag{5}$$

- Mahalanobis Metric

$$d(\boldsymbol{x}, \mathbf{y}, \mathbf{C}) = \sqrt{(\boldsymbol{x} - \mathbf{y})^T \mathbf{C}^{-1} (\boldsymbol{x} - \mathbf{y})} \tag{6}$$

where ***x*** and **y** are the two vectors to compare and **C** is the covariance matrix derived from the eigenvalues found through PCA. Note that **C** is a diagonalised covariance matrix as eigenvectors are orthogonal. For the experiments conducted in this paper the first 30 eigenvalues and eigenvectors were retained; this is because of the limited data available to conduct feature modelling and the limited impact this had on the performance of the NN classifiers.

3. Data Normalisation

Data normalisation is a key step in any face recognition system. It provides a common basis for comparison of two images (or signals). It was shown in [4] that low resolution 3D face images significantly hinder recognition performance, however the same is not true of 2D face images. For this reason the 2D and 3D face images were retained at different resolutions. The 2D face data had a resolution of 64×64 pixels whereas the the 3D face data had a resolution of 128×128 pixels. For both the 2D and 3D face images only part face images, above the mouth region and below the brow region were used. This was done to reduce the effect of expression variation.

For the 2D face data only in-plane rotations could be recovered. The two eye corners were used to perform the in-plane normalisation and cropping of the images. Registration of the images is based on the eye corners; they reside on the same y-axis and are separated by 64 pixels. Illumination normalisation was conducted by applying local mean window normalisation, a technique previously used in [9]. The images were also mean and standard deviation normalised to provide a common basis for comparison. An example 2D face image which is cropped and fully normalised is provided in Figure 2.

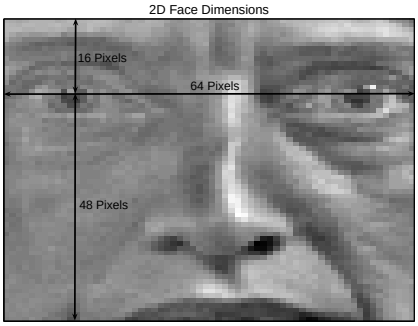

Figure 2. A cropped and fully normalised 2D face image, included in this image are the dimensions defining the position of the registered eye corners within the 64×64 image.

The 3D data consisted of point cloud data on a semi-regular x- and y-grid; there is limited variation along the grid lines. Every image has four landmark points: the right eye corner, left eye corner, nose tip and chin. Using three of these landmark points (right eye corner, left eye corner and chin) the 3D data is normalised for in-plane and out-of-plane rotations. After the data has been normalised for all rotations the valid data is then interpolated onto a regular grid of 128×128 pixels. This regularised data is then further processed with a median filter to reduce the effect of noise; an example output of this procedure is provided in Figure 3.

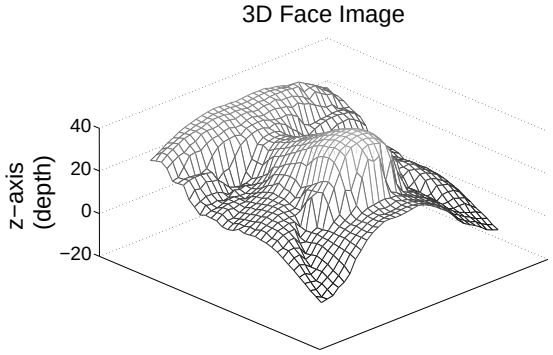

Figure 3. A mesh plot of a cropped and interpolated 3D face image.

In addition to filtering and rotations, the 3D data must also be registered and range normalised. Registration is achieved by placing the eye corners on the same y-gridline separated by 128 pixels. Range normalisation consists of setting the maximum value to 255 (the maximum value for a normal 2D image) and adjusting all other values to be relative to this value. An example of a cropped and fully normalised image is provided in Figure 4.

4. Feature Modelling

For these experiments feature modelling was conducted upon difference vectors; which described two forms of variation Intra- and Extra-Personal. Intra-Personal (IP) variation consists of variations that occur between different images of the same ID whilst Extra-Personal (EP) variation consists of variations that occur between images of different IDs. The IP and EP variations are modelled using separate GMMs.

Since the two classes (forms of variations), Ω_{IP} (IP) and Ω_{EP} (EP), are both described by GMMs this intuitively leads to the discriminant function for classification [7],

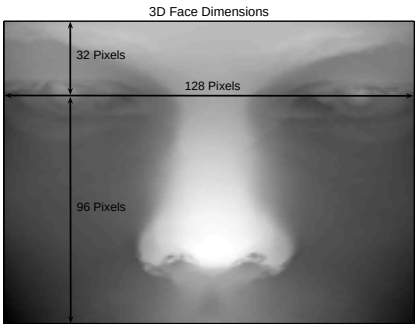

Figure 4. A cropped and fully normalised 3D face image, included are the dimensions defining the position of the registered eye corners within the 128×128 regular grid.

$$g(\boldsymbol{x}) = \ln(\frac{p(\boldsymbol{x} \mid \Omega_{IP})}{p(\boldsymbol{x} \mid \Omega_{EP})}) + \ln(\frac{P(\Omega_{IP})}{P(\Omega_{EP})}) \quad (7)$$

where $p(\boldsymbol{x} \mid \Omega_{IP})$ is the probability that observation $\boldsymbol{x}$ belongs to class Ω_{IP}, $p(\boldsymbol{x} \mid \Omega_{EP})$ is the probability that observation $\boldsymbol{x}$ belongs to class Ω_{EP}, $P(\Omega_{IP})$ represents the probability of class Ω_{IP} and $P(\Omega_{EP})$ is the probability of class Ω_{EP}. If both classes are considered to be equally likely, $P(\Omega_{IP}) = P(\Omega_{EP}) = 0.5$, then Equation 7 simplifies to

$$g(\boldsymbol{x}) = \ln(p(\boldsymbol{x} \mid \Omega_{IP})) - \ln(p(\boldsymbol{x} \mid \Omega_{EP})) \quad (8)$$

which is the form of the discriminant function used in this work.

5. Experiments

Two sets of experiments were conducted. The first set of experiments explored the effectiveness of the feature modelling technique and the second set of experiments fused the best 2D and 3D classifiers to form a hybrid 2D+3D classifier. Also examined was the effect of between-session variation and the impact it would have on recognition performance. All the experiments were conducted on the FRGC v2.0 database [11]. This database was split into two sessions; Session1 consisted of data from Fall2003range and Session2 consisted of data from Spring2004range. The split of this data is discussed in more detail below.

5.1. Data Split

A subset of the FRGC database, the validation set, was used for these experiments; in this set there are 466 IDs with 4007 joint 2D and 3D images taken across two sessions. The two sessions, Session1 and Session2, contained different sets of IDs but there were 249 crossover IDs (IDs that were present in both Session1 and Session2). The data for each session was split into three disjoint sets; one for training, one for tuning and one for testing. The tuning data is used to derive weights for linear fusion and also for evaluating classifiers. To ensure that the Train, Tune and Test sets were disjoint, the IDs were split between the sets with a 2:1:1 ratio respectively. This split is represented in Figure 5 to highlight that crossover IDs used in the Train set for Session1 were present in the Train set for Session2 but, these IDs were not present in any of the Test or Tune sets. To ensure the consistency of the results 6 random splits (Splits A-F) were used.

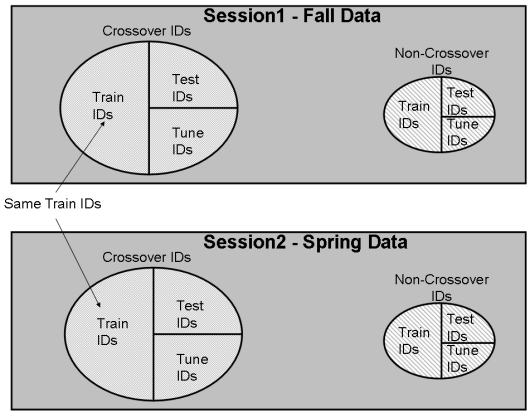

Figure 5. Diagrams providing an overview of how the IDs are split, highlighting that the crossover IDs chosen for the Train, Tune and Test data sets are the same for both Session1 and Session2.

5.2. Usage of Enrolment and Background Data

By describing the feature vectors with difference vectors, the GMM training algorithm is provided with extra observations that are normally unavailable. By forming difference vectors the number of observations is determined by the number of IP permutations possible,

$$\sum_{i=1}^{D} n_i P_2, \quad (9)$$

rather than having the number of images available,

$$\sum_{i=1}^{D} n_i, \quad (10)$$

as the number of observations; where D is the number of training IDs available to conduct training and n_i is the number of images available for the i^{th} ID.

For these experiments the Train set and enrolment data were used to derive the GMMs [15]. The enrolment process consisted of randomly taking P (for these experiments $P = 4$) images for every ID in the set (Tune or Test) with $P + 1$ images; this left at least 1 image, for each enrolled ID, to be used for testing. The IP GMMs were derived globally, to describe all the IDs rather than an individual one, by forming all the permutations of those IDs in the Train set with P images; this was done to avoid any bias in the training scheme. The EP GMMs were derived for each enrolled ID. The P enrolment images were compared against the images from the Train data to form the difference vectors and so derive the GMM. With the limited amount of training data available the GMMs were limited to 4 components.

The P enrolment images meant that for each ID multiple scores were obtained which provide multiple chances for the correct decision to be made. To perform score modelling the best score was chosen as the matching score. For example for the L_1 NN classifier this would mean that

$$L_1\_bestscore = \min_{i=1}^{P}(L_{1k}) \tag{11}$$

would provide the best matching score, where L_{1k} is the k^{th} classifier and P represents the number of enrolment images. This provides limited score modelling without impacting on the effectiveness of the feature modelling technique, which is a key area of investigation in this work.

6. Results and Discussion

The experimental results indicate that the IP-EP GMM is an effective classifier for 3D face recognition but the results for IP-EP GMMs for 2D face recognition are inconclusive. However, fusing feature modelling classifiers for the two modalities, 2D and 3D, demonstrates a significant performance increase as is shown in Figure 6. For both sets of experiments score range normalisation was trialled but the techniques trialled proved detremental to the fusion results; the primary range normalisation technique trialled was mean and standard deviation normalisation. Results in the following sections are presented as Equal Error Rates (EERs); which is where the False Alarm probability is equal to the Miss probability.

6.1. Nearest Neighbour versus Feature Modelling

The effectiveness of feature modelling, using GMMs, against several standard NN classifiers for face verification were analysed. Two sets of experiments were conducted:

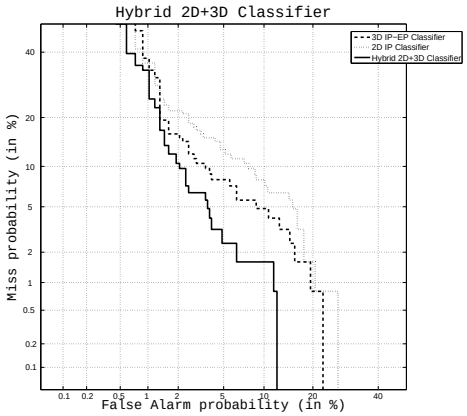

Figure 6. An example Detection Error Trade-off plot highlights the effectiveness of the hybrid 2D+3D face recognition technique when using feature modelling. The results for the DET plot comes from the within-session experiments for SplitE Session1. Other splits and sessions show similar trends.

within-session and between-session. Within-session experiments were tests conducted with Train and Test sets from the same session whereas between-session experiments were tests conducted with the Train and Test sets coming from different sessions.

The within-session experiments, Table 1, indicated that the IP-EP GMM classifier for 2D and 3D face modalities performed as well or better than the best NN classifier; the best NN classifier is chosen to be the optimal classifier from the Tune set. However, there were 3 tests for the 2D modality where the IP GMM classifier outperformed the the IP-EP GMM classifier and further analysis shows that the 2D IP GMM classifier performs as well or better than the best NN classifier in all of the tests.

The experiments on the between-session variability, Table 2, indicated that the IP-EP GMM classifier was not optimal for the 2D face modality. For the 2D modality the IP GMM classifier consistently outperformed the IP-EP GMM classifier and performed as well or better than the best NN classifier. For the 3D face modality the IP-EP GMM classifier performed as well or better than the best NN classifier in all but 2 of the tests. For these tests, where the best NN classifier performed better than the IP-EP GMM classifier, the IP GMM classifier performed better than the IP-EP GMM.

These experiments demonstrate the effectiveness of feature modelling using IP-EP GMMs for PCA feature vectors for the 3D modality and the effectiveness of the IP GMMs for describing the PCA feature vectors for the 2D

		2D			3D		
		Best NN Classifier	IP GMM Classifier	IP-EP GMM Classifier	Best NN Classifer	IP GMM Classifier	IP-EP GMM Classifier
SplitA	Session1	8.33%	8.33%	**7.41%**	8.33%	10.19%	**5.56%**
	Session2	8.70%	**7.79%**	8.70%	14.50%	15.22%	**10.87%**
SplitB	Session1	11.02%	11.02%	**9.32%**	8.47%	9.32%	**5.08%**
	Session2	11.43%	**9.14%**	9.71%	18.29%	16.00%	**9.71%**
SplitC	Session1	7.08%	7.08%	**7.08%**	10.62%	11.50%	**5.31%**
	Session2	12.58%	11.92%	**9.93%**	13.25%	11.92%	**10.60%**
SplitD	Session1	11.90%	9.52%	**7.94%**	9.52%	10.32%	**6.35%**
	Session2	12.33%	**10.27%**	10.96%	14.38%	16.44%	**11.64%**
SplitE	Session1	8.87%	8.87%	**8.87%**	11.29%	10.48%	**6.45%**
	Session2	10.13%	10.13%	**9.49%**	15.19%	14.56%	**12.03%**
SplitF	Session1	9.73%	8.85%	**7.08%**	11.50%	14.16%	**7.96%**
	Session2	12.20%	11.59%	**10.98%**	17.68%	16.46%	**10.37%**

Table 1. Results for Within-Session Variation, Train and Test sets are from the same sessions, presented as EERs. Highlighted are the best 2D and 3D classifiers for each test.

		2D			3D		
		Best NN Classifier	IP GMM Classifier	IP-EP GMM Classifier	Best NN Classifier	IP GMM Classifier	IP-EP GMM Classifier
SplitA	Session1	10.19%	**10.19%**	12.96%	8.33%	12.04%	**8.33%**
	Session2	10.14%	**9.42%**	14.49%	13.04%	13.77%	**11.59%**
SplitB	Session1	15.25%	**14.41%**	21.19%	8.47%	8.47%	**6.78%**
	Session2	16.57%	**16.00%**	21.14%	16.57%	16.57%	**13.71%**
SplitC	Session1	9.73%	**9.73%**	16.81%	10.62%	11.50%	**7.96%**
	Session2	14.57%	**11.92%**	15.89%	14.57%	**13.25%**	14.57%
SplitD	Session1	12.70%	**12.70%**	15.08%	11.11%	11.11%	**9.52%**
	Session2	15.75%	**15.75%**	17.81%	18.49%	17.12%	**14.38%**
SplitE	Session1	12.10%	**11.29%**	17.74%	12.10%	11.29%	**10.48%**
	Session2	14.56%	**14.56%**	20.89%	**14.56%**	15.19%	17.82%
SplitF	Session1	11.50%	**10.62%**	15.04%	11.50%	11.50%	**10.62%**
	Session2	14.63%	**14.63%**	19.51%	**17.07%**	18.29%	19.51%

Table 2. Results for Between-Session Variation, Train and Test are from different sessions, presented as EERs. Highlighted are the best 2D and 3D classifiers for each test.

modality. It is interesting to note that the combination of IP and EP GMMs is effectively fusion of a strong and weak classifier respectively. Although the performance of the 3D IP-EP GMM classifier was reduced when between-session variability was introduced it still provided improved performance against the best NN classifier. However, once between-session variability was introduced for the 2D modality the performance of the IP-EP GMM classifier was significantly reduced such that the IP GMM performed better. The performance degradation in the 2D modality is likely due to variability between the sessions in terms of illumination and pose which significantly affects the performance of any 2D classifier.

6.2. Fused 2D and 3D Classifier

Having determined that the 3D IP-EP GMM and the 2D IP GMM were the most effective methods of modelling the PCA-based features this set of experiments explored possible methods of fusing these classifiers to improve face recognition. The method explored in this work is the sum rule. It was shown in [8] that the sum rule is robust to es-

timation errors and so is a robust method for conducting score fusion. The final hybrid 2D+3D classifier chosen for testing was an equal weighted summation classifier, using 2D IP GMMs and 3D IP-EP GMMs classifiers for fusion.

The results for the within-session experiments, Table 3, demonstrate that this hybrid classifier significantly outperforms each individual modality. The average EER for the within-session variation for the 2D modality is 9.56% and 8.49% for the 3D modality whilst the hybrid 2D+3D classifier has an average EER of 4.78% which is a significant improvement. This improvement continues through to the between-session variation tests, Table 4, where the 2D modality has an average EER of 12.60% and 12.10% for the 3D modality with the hybrid 2D+3D classifier achieving an average EER of 7.38%. This consistent improvement across both modes of testing demonstrates the effectiveness of fusing 2D and 3D feature modelling classifiers for improved face recognition.

7. Conclusions and Future Work

A hybrid 2D and 3D classifier formed using 2D IP GMMS and 3D IP-EP GMMs provides a significantly improved classifier; an average EER of 12.60% for the 2D classifier and 12.10% for the 3D classifier can be fused to form a Hybrid 2D+3D classifier with an average EER of 7.38%. Furthermore it was shown that feature modelling PCA-based difference vectors, describing IP and EP variation, performs as well or better than NN classifiers. It is interesting to note that between-session variability has less impact on the 3D modality than it does on the 2D modality; this is due to the fact that the 3D modality is unaffected by illumination and pose variations which remain unsolved problems with the 2D modality. However, even with this issue the 2D modality provides significant complementary information to the 3D modality, as was demonstrated by the hybrid 2D+3D experiments.

Several avenues of research have been highlighted by this work. First, methods for forming a discriminant IP-EP GMM function robust to between-session variability needs to be explored for both the 2D and 3D modalities. Alternate methods for classifier fusion also need to be investigated such as Support Vector Machines (SVMs) and neural networks. In addition to this other forms of fusion should be investigated; in particular feature level fusion as both modalities, 2D and 3D, have PCA features extracted.

8. Acknowledgements

This research was supported by the Australian Research Council (ARC) through Discovery Grants Scheme, Grant DP0452676,2004-6.

References

[1] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1063–1074, 2003.

[2] K. Bowyer, K. Chang, and P. Flynn. A survey of approaches to three-dimensional face recognition. *Proceedings of the 17th International Conference on Pattern Recognition*, 1:358–361, 2004.

[3] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Expression-invariant 3d face recognition. *Audio- and Video-Based Person Authentication*, pages 62–70, 2003.

[4] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Face recognition using 2d and 3d facial data. *Workshop in Multimodal User Authentication*, pages 25–32, 2003.

[5] J. Cook, V. Chandran, S. Sridharan, and C. Fookes. Face recognition from 3d data using iterative closest point algorithm and gaussian mixture models. *Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization and Transmission*, pages 502–509, 2004.

[6] J. Cook, V. Chandran, S. Sridharan, and C. Fookes. Gabor filter bank representation for 3d face recognition. *Proceedings of Digital Image Computing: Techniques and Applications*, 2005.

[7] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification: 2nd Edition*. John Wiley and Sons, Inc., 2001.

[8] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:226–239, 1998.

[9] S. Lowther, C. McCool, V. Chandran, and S. Sridharan. Improving face localisation using claimed identity for face recognition. *Proceedings of the 3rd Interational Workshop on the Internet, Telecommunications and Signal Processing*, 2004.

[10] S. Lucey. The symbiotic relationship of parts and monolithic face representations in verification. *Proceedings of IEEE Conference of Computer Vision and Pattern Recognition Workshop*, page 89, 2004.

[11] J. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. *Proceedings of IEEE Conference of Computer Vision and Pattern Recognition*, 1:947–954, 2005.

[12] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002: Overview and summary. *IEEE Interational Workshop on Analysis and Modeling of Faces and Gestures*, page 44, 2003.

[13] D. Reynolds. Speaker identification and verification using Gaussian mixture speaker models. *ESCA Workshop on Automatic Speaker Recognition*, pages 27–30, 1994.

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

[15] R. Vogt, J. Pelecanos, and S.Sridharan. Dependence of gmm adaptation on feature post-processing. *Eurospeech*, pages 3013–3016, 2003.

[16] Y. Wang, C. Chua, and Y. Ho. Facial feature detection and face recognition from 2d and 3d images. *Pattern Recognition Letters*, 23:1191–1202, 2002.

		2D IP GMM Classifier	3D IP-EP GMM Classifier	Hybrid 2D IP + 3D IP-EP Classifier
SplitA	Session1	8.33%	5.56%	**3.70%**
	Session2	7.97%	10.87%	**4.35%**
SplitB	Session1	11.02%	5.08%	**3.39%**
	Session2	9.14%	9.71%	**6.29%**
SplitC	Session1	7.08%	5.31%	**2.65%**
	Session2	11.92%	10.60%	**5.30%**
SplitD	Session1	9.52%	6.35%	**3.97%**
	Session2	10.27%	11.64%	**6.85%**
SplitE	Session1	8.87%	6.45%	**4.03%**
	Session2	10.13%	12.03%	**6.96%**
SplitF	Session1	8.85%	7.96%	**4.42%**
	Session2	11.59%	10.37%	**5.49%**

Table 3. Results for Within-Session Variation, Train and Test sets are from the same sessions, of the best 2D, 3D and hybrid 2D+3D classifiers; all results are presented as EERs.

		2D IP GMM Classifier	3D IP-EP GMM Classifier	Hybrid 2D IP + 3D IP-EP Classifier
SplitA	Session1	10.19%	8.33%	**5.56%**
	Session2	9.42%	11.59%	**7.24%**
SplitB	Session1	14.41%	6.78%	**3.39%**
	Session2	16.00%	13.71%	**9.71%**
SplitC	Session1	9.73%	7.96%	**6.19%**
	Session2	11.92%	14.57%	**8.61%**
SplitD	Session1	12.70%	9.52%	**6.35%**
	Session2	15.75%	14.38%	**10.27%**
SplitE	Session1	11.29%	10.48%	**6.45%**
	Session2	14.56%	17.72%	**9.49%**
SplitF	Session1	10.62%	10.62%	**6.19%**
	Session2	14.63%	19.51%	**9.15%**

Table 4. Results for Between-Session Variation, Train and Test sets are from different sessions, of the best 2D, 3D and hybrid 2D+3D classifiers; all results are presented as EERs.

[17] W. S. Yambor, B. A. Draper, and J. R. Beveridge. Analyzing pca-based face recognition algorithms: Eigenvector selection and distance measures. *2nd Workshop on Empirical Evaluation in Computer Vision*, 2000.