Twins 3D Face Recognition Challenge

Vipin Vijayan [1], Kevin W. Bowyer [1], Patrick J. Flynn [1], Di Huang [2], Liming Chen [2],
Mark Hansen [3], Omar Ocegueda [4], Shishir K. Shah [4], Ioannis A. Kakadiaris [4]

Abstract

Existing 3D face recognition algorithms have achieved high enough performances against public datasets like FRGC v2, that it is difficult to achieve further significant increases in recognition performance. However, the 3D TEC dataset is a more challenging dataset which consists of 3D scans of 107 pairs of twins that were acquired in a single session, with each subject having a scan of a neutral expression and a smiling expression. The combination of factors related to the facial similarity of identical twins and the variation in facial expression makes this a challenging dataset. We conduct experiments using state of the art face recognition algorithms and present the results. Our results indicate that 3D face recognition of identical twins in the presence of varying facial expressions is far from a solved problem, but that good performance is possible.

1. Introduction

We conduct a study on the performance of state of the art 3D face recognition algorithms on a large set of identical twins using the 3D Twins Expression Challenge ("3D TEC") dataset. The dataset contains 107 pairs of identical twins and is the largest dataset of 3D scans of twins known to the authors.

Recently, there have been some twin studies in biometrics research. Phillips *et al.* [1] assessed the performance of three of the top algorithms submitted to the Multiple Biometric Evaluation (MBE) 2010 Still Face Track [2] on a dataset of twins acquired at Twins Days [3] in 2009 and 2010. They examined the performance using images acquired on the same day, and also using images acquired a year apart (i.e., where the face images acquired in the first year were used as gallery images and the face images acquired in the second year as probe images). They also examined the performance with varying illumination conditions and expressions. They found that results ranged from

approximately 2.5% Equal Error Rate (EER) for images acquired on the same day with controlled lighting and neutral expressions, to approximately 21% EER for gallery and probe images acquired in different years and with different lighting conditions.

Sun *et al.* [4] conducted a study on multiple biometric traits of twins. They found no significant difference in performance when using non-twins compared to using twins for their iris biometric system. For their fingerprint biometric system, they observed that the performance when using non-twins was slightly better than using twins. In addition, their face biometric system could distinguish non-twins much better than twins.

Hollingsworth *et al.* [5] examined whether iris textures from a pair of identical twins are similar enough that they can be classified by humans as being from twins. They conducted a human classification study and found that people can classify two irises as being from the same pair of twins with 81% accuracy when only the ring of iris texture was shown to them.

Jain *et al.* [6] conducted a twins study using fingerprints. They found that identical twins tend to share the same fingerprint class (fingerprints are classified into whorls, right/left loops, arches, etc.) but their fingerprint minutiae were different. They concluded that identical twins can be distinguished using a minutiae-based automatic fingerprint system, with slightly lower performance when distinguishing identical twins compared to distinguishing random persons.

To date, there have been no studies conducted in 3D face recognition that focused mainly on twins. The only 3D face recognition study known to the authors that mentioned twins was Bronstein *et al.* [7], where they tested the performance of their 3D face recognition algorithm on a dataset of 93 adults and 64 children which contained one pair of twins, and stated that "our methods make no mistakes in distinguishing between Alex and Mike".

2. The Dataset

The Twins Days 2010 dataset was acquired at the Twins Days Festival in Twinsburg, Ohio [3]. Phillips *et al.* [1] provides more details about the overall dataset. It contains 266 subject sessions, with the 3D scans in the dataset containing two scans: one with a neutral expression and another with a smiling expression. There were 106 sets of identical twins, one set of triplets, and the rest were non-twins.

---

[1] Department of Computer Science and Engineering, University of Notre Dame. 384 Fitzpatrick Hall, Notre Dame, IN 46556, USA
{vvijayan, kwb, flynn}@nd.edu

[2] Université de Lyon, CNRS, Ecole Centrale Lyon, LIRIS UMR 5205, 69134, Ecully, France

[3] Machine Vision Lab, DuPont Building, Bristol Institute of Technology, University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK

[4] Computational Biomedicine Lab, Department of Computer Science, University of Houston, 4800 Calhoun Road, Houston, TX 77004, USA
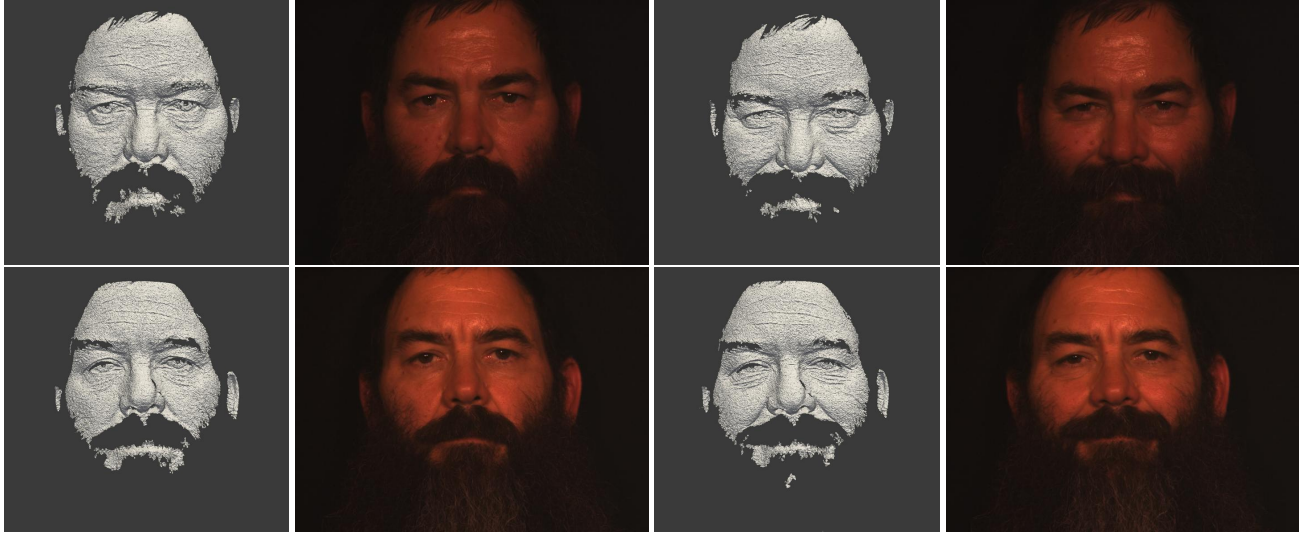
Figure 1: Images of two twins acquired in a single session. The top row shows the images obtained from one twin and the bottom row, the other twin. The left two images contain the neutral expression. The right two are of the smiling expression. (The texture images were brightened to increase visibility in this figure.)

Three pairs of twins came in for two recording sessions and the other twins for only a single session. The twins in this database declared themselves to be identical twins; no tests were done to prove this.

The experiments in this paper use the "3D TEC" subset of the Twins Days dataset, which consists of 3D face scans of 107 pairs of twins (two of the triplets were included as the 107th set of twins) and only the scans acquired in the first session were used for each subject. To our knowledge, this is the only dataset of 3D face scans in existence that has more than a single pair of twins. For information on obtaining the 3D TEC dataset, see [8].

The scans were acquired using a Minolta VIVID 910 3D scanner [9] in a controlled light setting, with the subjects posing in front of a black background. For each pair of twins, their neutral and smile images were taken in a 5 to 10 minute window of time.

The Minolta scanner acquires a texture image and a range image of 480×640 resolution. The telephoto lens of the Minolta scanner was used since it gives a more detailed scan. The distance of the scanner from the subject was approximately 1.2 m. A scan using the telephoto lens contains 70,000 to 195,000 points for the Twins 2010 dataset, with an average of 135,000 points.

3. Algorithms

We describe the four algorithms employed in this study. Table 1 shows the performance of these algorithms on the FRGC v2 [10] dataset.

3.1. Algorithm 1

Faltemier *et al.* [11] performed Iterative Closest Point (ICP) using an ensemble of 38 spherical regions and fused the match scores to calculate the final score. McKeon [12] has a number of optimizations over Faltemier *et al.* which include: (i) the symmetric use of the two point clouds and score fusion on the results, (ii) score normalization of the match scores, and (iii) weighting the scores for the regions. Algorithm 1 is a variation of McKeon. The major difference is the preprocessing step where the face is first roughly aligned using the symmetry plane estimation method described in Spreeuwers [13], and the image is then aligned to a reference face using ICP.

Each region in the ensemble is created by selecting a point in the probe image a certain offset from the origin, and then cropping out all points a certain distance away from the selected point. The nose-tip is set as the origin. Each region of the probe image is matched using ICP against the entire gallery image. The alignment errors for each region are taken to be the region's distance scores.

The scores are then fused as a linear combination of the region's distance scores. The integer weights for the linear combination are trained against FRGC v2 using a greedy algorithm that maximizes TAR at 0.1% FAR.

Let $E(p_1, p_2) = E_{SFSW}(p_1, p_2)$, be the match score of point clouds p_1 and p_2. The ICP algorithm is not symmetric, which means that $E(p_1, p_2) \neq E(p_2, p_1)$ for almost all cases. The two scores are fused using the minimum rule: $E_{min}(p_1, p_2) = \min(E(p_1, p_2), E(p_2, p_1))$.

The match scores are then normalized in two ways. First, the match scores are normalized such that the normalized

score is

$$E_{pkn}(p, g_k) = \frac{E_{min}(p, g_k)}{\sum_{j=1, j \neq k}^{N} \frac{E_{min}(g_j, g_k)}{N-1}} \quad (1)$$

where p is a probe image, g_k are the gallery images, and N is the number of gallery images.

Then we perform min-max normalization over the resulting match score from the first normalization, E_{pkn}, so that the final match score is

$$E_{minmax}(p, g_k) = \frac{E_{pkn}(p, g_k) - \min(V_p)}{\max(V_p) - \min(V_p)} \quad (2)$$

where $V_p = [E_{pkn}(p, g_1), E_{pkn}(p, g_2), ..., E_{pkn}(p, g_N)]$.

If we normalize against the gallery using E_{minmax} for verification, then we would have to match the probe against all images in the gallery. This would be very slow if we use only a single processor. Thus we show the performance of two variations of Algorithm 1: one using the distance scores from E_{pkn} and the second using E_{minmax}.

3.2. Algorithm 2

Algorithm 2 consists of two main steps: intermediate facial representation and Scale Invariant Feature Transform (SIFT) based local matching. If local features are directly extracted from smooth facial range images, it leads to a limited number of local features, or features with low discriminative power. To solve this problem, intermediate facial representation is used to highlight local shape changes of 3D facial surfaces in order to improve their distinctiveness. In this paper, we evaluated three types of intermediate facial maps: Shape Index [14], extended Local Binary Patterns [15] as well as Perceived Facial Images [16]. Figure 2 shows examples of these facial maps. The three types of facial maps are described below.
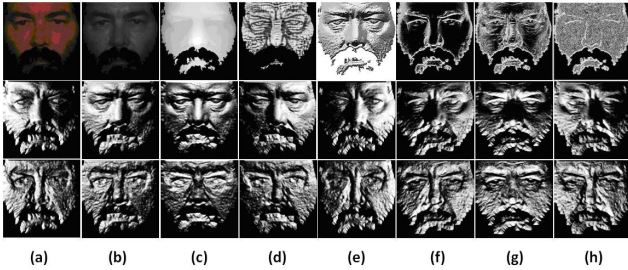

Figure 2: Some examples of intermediate facial representation. The first row contains (a) original RGB image; (b) grayscale texture image; (c) original range image; (d) SI map; (e)-(h) eLBP maps of different layers. The second row contains eight PFIs of quantized orientations of facial range image. The third row contains eight PFIs of quantized orientations of facial texture image.

Shape Index (SI) [14] was first proposed to describe shape attributes. For each vertex p of a 3D facial surface,

its SI value can be calculated using

$$S(p) = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \quad (3)$$

where k_1 and k_2 are the maximum and minimum principal curvatures respectively. Based on the SI values of all the vertices, we can produce the SI map of a given facial surface.

In the Extended Local Binary Pattern (eLBP) [15] approach, a set of multi-scale eLBP maps are generated to represent a given facial range image. eLBP maps consist of four layers. Layer 1 is LBP, which encodes the gray value differences between neighboring pixels into a binary pattern. eLBP also considers their exact value differences and encodes this information into Layers 2 to 4. The eLBP maps are generated by regarding the eLBP codes of each pixel as intensity values. As the neighborhood size of the given pixel changes, multi-scale eLBP maps are formed.

Perceived Facial Image (PFI) [16] aims at simulating the complex neuron response using a convolution of gradients in various orientations within a pre-defined circular neighborhood. Given an input facial image I, a certain number of gradient maps $L_1, L_2, \cdots, L_o$, one for each quantized direction o, are first computed. Each gradient map describes gradient norms of the original image in an orientation o at every pixel. The response of complex neurons is then simulated by convolving its gradient maps with a Gaussian kernel G, and the standard deviation of G is proportional to the radius value of the given neighborhood area R; i.e., $\rho_o^R = G_R * L_o$.

The purpose of the Gaussian convolution is to allow the gradients to shift in a neighborhood without abrupt changes. At a certain pixel location (x, y), we collect all the values of the convolved gradient maps at that location and form the vector $\rho^R(x, y)$, which is a response value of complex neurons for each orientation o. So, $\rho^R(x, y) = [\rho_1^R(x, y), \cdots, \rho_O^R(x, y)]^t$ where $o = 1..O$.

The vector $\rho^R(x, y)$ is further normalized to a unit norm vector $\underline{\rho}^R(x, y)$, which is called response vector. Thus, a new Perceived Facial Image (PFI), J_o, is calculated where $J_o(x, y) = \underline{\rho}_o^R(x, y)$.

After the three types of intermediate facial representations are computed, a SIFT-based matching process [17] is used to find robust keypoints from the facial representations. We expect there to be more correlated keypoints between facial maps of the same subject than those of different subjects. Furthermore, since SIFT has good tolerance to moderate pose variations and all the data in the 3D TEC dataset are nearly frontal scans, we did not perform any registration in preprocessing. All parameter settings of intermediate facial representations are presented in detail in [14, 15, 16].

In addition, SI maps and eLBP maps are mainly proposed for 3D facial range images, while PFIs can be either

applied to facial range or texture images as done in Huang *et al.* [16] for 3D face recognition using shape and texture. Therefore, in this paper, we also tested the performance based on 2D PFIs with SIFT matching for comparison.

3.3. Algorithm 3

Algorithm 3 converts the 3D image to a surface normal representation, then discards data with less discriminatory power and resizes the image. It then matches the images using the Euclidean distance of the variance of the remaining surface normals.

Surface normals have been shown to lend themselves well to face recognition tasks [18]. We convert the depth maps of 3D images to surface normal representations, applying median smoothing and hole filling to reduce noise.

Unnikrishnan [19] conceptualized an approach similar to face caricatures for human recognition. In this approach, only those features which deviate from the norm by more than a threshold are used to uniquely describe a face. Unnikrishnan suggested using features whose deviations lie below the 5th percentile and above the 95th percentile, thereby discarding 90% of the data. In a similar vein, the algorithm that we present here is based on what we call the "Variance Inclusion Criterion". We can use the surface normal variance at each pixel location as a distance measure between images. If a pixel shows a large variance across the dataset, then it can be used for recognition (assuming that variance within the class or subject is small). Therefore, the standard deviation of each pixel is calculated over all the images in the gallery. Whether or not a particular pixel location is used in recognition depends on whether or not the variance is above a pre-determined threshold.

Another key step of this algorithm is resizing the image. Sinha *et al.* [20] summarized a number of findings indicating that humans can recognize familiar faces from very low resolution images. We resize the surface normal maps to 10×10 pixels before applying the Variance Inclusion Criterion to get the number of pixels used for recognition down to just over 60 pixels. The reason for choosing this value is due to experimentation on frontal and neutral expression subsets of the FRGC v2 and Photoface [21] datasets. In these experiments it was found that when retaining only 64 pixels for FRGC v2 data and 61 pixels for Photoface data, rank-one recognition rates of 87.75% and 96.25% were achieved respectively (a loss of only 7% and 2% respectively from the baseline). This is taken as an indication that the high variance pixel locations contain disproportionately more discriminatory information than low variance pixel locations.

Considering the two expressions used between gallery and probe images in the 3D TEC dataset, it was felt that the most variance would occur around the mouth region and bottom half of the face. Therefore, we only performed the

variance analysis on the top half of the face.

Additional pre-processing is performed by aligning all the images to the median left and right lateral canthus and nose tip coordinates for the dataset. A tight crop around the facial features is then applied to remove areas in a straight-forward way that can be occluded by hair. Euclidean distance is used for classification.

It is envisaged that this algorithm be used as a means of pruning the search space due to its computational efficiency before applying more rigorous algorithms.

3.4. Algorithm 4

The UR3D algorithm proposed by Kakadiaris *et al.* [22] consists of three main steps: (i) the 3D facial meshes are aligned to a common reference Annotated Face Model (AFM), (ii) the AFM is deformed to fit the aligned data, and (iii) the 3D fitted mesh is represented as a three-channel image using the global UV-parameterization of the AFM. The benefit of representing the 3D mesh as a multi-channel image is that standard image processing techniques can be applied directly to the images. In this approach, the full Walsh wavelet packet decomposition is extracted from each band of the geometry and normal images and a subset of the wavelet coefficients are selected as the signature of the mesh. The signature can be compared directly using a weighted L^1 norm. Recently, Ocegueda *et al.* [23] presented an extension to UR3D that consists of a feature selection step that reduces the number of wavelet coefficients retained for recognition, followed by a projection of the signatures to a subspace generated using Linear Discriminant Analysis (LDA). The feature selection step was necessary because the high dimensionality of the standard UR3D signature made it infeasible to apply standard algorithms for LDA. However, by using the algorithm proposed by Yu and Yang [24], we can directly apply LDA to the original UR3D metric. We found that applying LDA to the original signature yields slightly better results. We will use this variation of the UR3D algorithm in our experiments. We used the frontal, non-occluded facial meshes from the Bosphorus dataset developed by Savran *et al.* [25] as the training set for LDA.

4. Experimental Design

We arbitrarily label one person in each pair of twins as Twin A and the other as Twin B and perform verification and identification experiments using the four different gallery and probe sets shown in Table 2.

Case I has all of the images with a smiling expression in the gallery and the images with a neutral expression as the probe. Case II reverses these roles. This models a scenario where the gallery has one expression and the probe has another expression. In the verification scenario, both the match and non-match pairs of gallery and probe images

Algorithm	Rank-1 RR	VR (ROC III)
Alg. 1	98.0%	98.8%
Alg. 2 (SI)	91.8%	85.8%
Alg. 2 (eLBP)	97.2%	95.0%
Alg. 2 (Range PFI)	95.5%	90.4%
Alg. 2 (Text. PFI)	95.9%	
Alg. 3	87.8%	
Alg. 4	97.0%	97.0%

Table 1: Rank-one recognition rates and verification rates (TAR at 0.1% FAR) of the algorithms on the FRGC v2 dataset. For recognition, the first image acquired of each subject is in the gallery set and the rest of the images are probes. For the ROC III verification experiment, the gallery set contains the images acquired in the first semester and the probe set contains the images in the second semester.

No.	Gallery		Probe	
I	A Smile,	B Smile	A Neutral,	B Neutral
II	A Neutral,	B Neutral	A Smile,	B Smile
III	A Smile,	B Neutral	A Neutral,	B Smile
IV	A Neutral,	B Smile	A Smile,	B Neutral

Table 2: Gallery and probe sets for cases I, II, III, and IV. "A Smile, B Neutral" means that the set contains all images with Twin A smiling and Twin B neutral.

will have different expressions. In the identification scenario, theoretically the main challenge would be to distinguish between the probe's image in the gallery and his/her twin's image in the gallery since they look similar.

Case III has Twin A smiling and Twin B neutral in the gallery with Twin A neutral and Twin B smiling as the probe. Case IV reverses these roles. This models a worst case scenario in which the system does not control for the expressions of the subject in a gallery set of twins. In the verification scenario, the match pairs would have opposite expressions like in Cases I and II but the non-match pairs that are of the same pair of twins would have the same expression. In the identification scenario, theoretically the main challenge would be to distinguish between the probe's image and his/her twin's image in the gallery. This is more difficult than Cases I and II since the probe's expression is different from his/her image in the gallery but is the same as his/her twin's image in the gallery.

5. Results and Discussion

We evaluate performance using the following characteristics: True Accept Rate at 0.1% False Accept Rate (TAR at 0.1% FAR), Equal Error Rate, and Rank-1 Recognition Rate. Figures 3, 4, and 5 show the Receiver Operating Characteristic (ROC) curves of the verification experiments for Algorithms 1, 3, and 4.
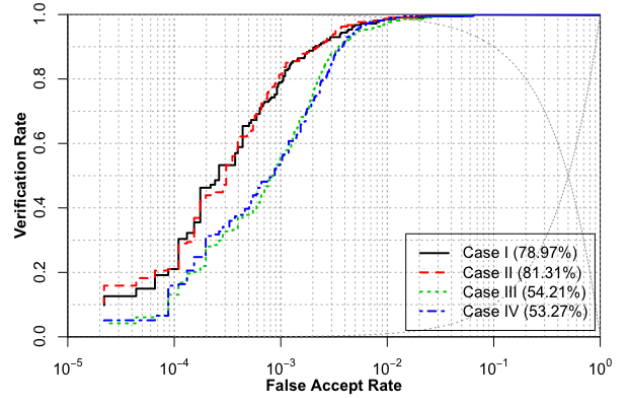

Figure 3: ROC curves of the four cases for Algorithm 1. The legend shows TAR at 0.1% FAR.
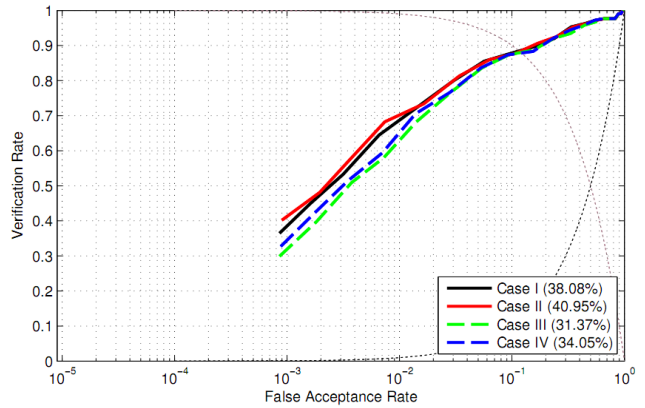

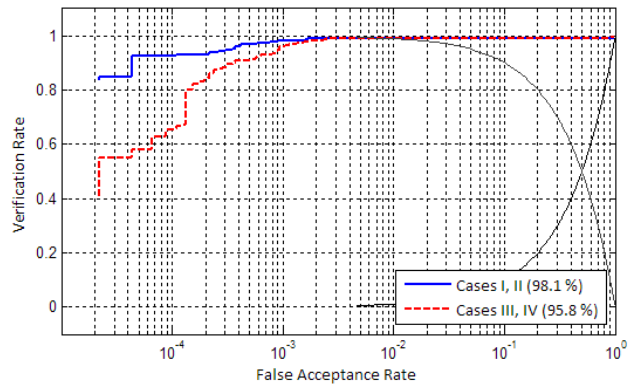Figure 4: Verification performance of Algorithm 3.

Figure 5: Verification performance of Algorithm 4.

In the first two of our four cases, all subjects are enrolled with a 3D face scan that has one expression, and all recognition attempts are made with the other expression. Thus, the difference in expression between enrollment and recognition is the same for all subjects. In these two cases, we find that 3D face recognition accuracy for twins exceeds 90% for

Algorithm	True Accept Rate			
	I	II	III	IV
Alg. 1 (E_{pkn})	79.0%	81.3%	54.2%	53.3%
Alg. 1 (E_{minmax})	**99.5%**	97.7%		
Alg. 2 (SI)	91.1%	89.7%	83.2%	81.8%
Alg. 2 (eLBP)	94.4%	95.3%	79.0%	78.0%
Alg. 2 (Range PFI)	93.5%	94.4%	68.7%	69.2%
Alg. 2 (Text. PFI)	96.7%	96.7%	93.0%	93.5%
Alg. 3	38.1%	41.0%	31.4%	34.1%
Alg. 4	98.1%	**98.1%**	**95.8%**	**95.8%**

Table 3: TAR at 0.1% FAR of the algorithms. All results are not available for Alg. 1 (E_{minmax}) due to duplicate match scores.

Algorithm	Equal Error Rate			
	I	II	III	IV
Alg. 1 (E_{pkn})	1.2%	1.0%	1.4%	1.1%
Alg. 1 (E_{minmax})	**0.2%**	**0.5%**	1.3%	0.9%
Alg. 2 (SI)	2.7%	3.7%	4.2%	4.5%
Alg. 2 (eLBP)	3.7%	3.3%	4.2%	4.2%
Alg. 2 (Range PFI)	4.1%	2.8%	4.7%	4.6%
Alg. 2 (Text. PFI)	2.7%	2.8%	3.3%	2.8%
Alg. 3	11.6%	11.8%	12.0%	12.2%
Alg. 4	0.8%	0.8%	**0.8%**	**0.8%**

Table 4: Equal Error Rate of the different algorithms.

Algorithm	Rank-1 Recognition Rate			
	I	II	III	IV
Alg. 1 (E_{pkn})	93.5%	93.0%	72.0%	72.4%
Alg. 1 (E_{minmax})	94.4%	93.5%	72.4%	72.9%
Alg. 2 (SI)	92.1%	93.0%	83.2%	83.2%
Alg. 2 (eLBP)	91.1%	93.5%	77.1%	78.5%
Alg. 2 (Range PFI)	91.6%	93.9%	68.7%	71.0%
Alg. 2 (Text. PFI)	95.8%	96.3%	91.6%	92.1%
Alg. 3	62.6%	63.6%	54.2%	59.4%
Alg. 4	**98.1%**	**98.1%**	**91.6%**	**93.5%**

Table 5: Rank-one recognition rates.

most of the algorithms. In the last two of the four cases, the facial expression differs between the twins' enrollment images and also between their images for recognition. In these cases, 3D face recognition accuracy ranges from the upper 60% to the lower 80%, except for Algorithm 2 (Texture PFI), which makes use of the texture information, and Algorithm 4. An exception is Algorithm 3, which showed reasonable performance on the FRGC v2 and Photoface [21] datasets but vastly degrades in performance on the 3D TEC dataset.

Why do some algorithms perform very well on this dataset while others don't? Algorithm 3, for example, discards a large amount of data by resizing and uses thresh-

olded Euclidean distance which is a fairly simple classification method. Algorithm 1, on the other hand, discards almost no data: it matches using the original point cloud that was scanned after some standard processing. The results also show a stark difference in the performances in Cases I and II compared to III and IV for some of the algorithms. This difference could demonstrate how well an algorithm deals with different expressions.

The 3D TEC dataset contains only "same session" data, meaning that there is essentially no time lapse between the image used for enrollment and the image used for recognition. Phillips *et al.* [1] examined the performance of 2D images of twins and found that results ranged from approximately 2.5% EER for images acquired on the same day with controlled lighting and neutral expressions, to approximately 21% EER for gallery and probe images acquired in different years and with different lighting conditions. Therefore, any performance estimates from this data are biased to exceed those that can be expected in any practical application.

This work is a collaboration by four research groups. The dataset was acquired and the evaluation framework defined by the Notre Dame group. Each of the groups collaborating on this work independently ran their own algorithm on the dataset and provided their results and the description of their algorithm. The final version of the paper was subject to edits by all co-authors.

6. Conclusion

3D face recognition continues to be an active research area. We have presented results of different state of the art algorithms on a dataset representing 107 pairs of identical twins with varying facial expressions, the 3D Twins Expression Challenge ("3D TEC") dataset. These algorithms have previously been reported to achieve good performance on the FRGC v2 dataset, which has become a de facto standard dataset for evaluating 3D face recognition. However, we observe lower performance on the 3D TEC dataset. The combination of factors related to the facial similarity of identical twins and the variation in facial expression makes for an extremely challenging problem.

The 3D TEC Challenge is smaller and therefore computationally simpler than the FRGC v2 Challenge. It combines a focus on fine discrimination between faces and handling varying expressions. There have been claims in the literature of 3D face recognition algorithms that can distinguish between identical twins. To our knowledge, this is the first time that experimental results have been reported for 3D face recognition involving more than a single pair of identical twins. The results demonstrate that 3D face recognition of identical twins in the presence of varying facial expressions remains an open problem.

7. Acknowledgements

References

[1] P. J. Phillips, P. J. Flynn, K. W. Bowyer, R. W. Vorder Bruegge, P. J. Grother, G. W. Quinn, and M. Pruitt, "Distinguishing identical twins by face recognition," in *Proc. FG*, Santa Barbara, CA, USA, Mar. 2011.

[2] P. J. Grother, G. W. Quinn, and P. J. Phillips, "Report on the evaluation of 2D still-image face recognition algorithms," in *NIST Interagency/Internal Report (NISTIR) - 7709*, 2010.

[3] "Twins days." [Online]. Available: http://www.twinsdays.org/

[4] Z. Sun, A. A. Paulino, J. Feng, Z. Chai, T. Tan, and A. K. Jain, "A study of multibiometric traits of identical twins," in *Proc. SPIE, Biometric Technology for Human Identification VII*, vol. 7667, Orlando, FL, USA, Apr. 2010, pp. 1–12.

[5] K. Hollingsworth, K. Bowyer, and P. Flynn, "Similarity of iris texture between identical twins," in *Proc. CVPR Workshop on Biometrics*, San Francisco, CA, USA, Jun. 2010, pp. 22–29.

[6] A. Jain, S. Prabhakar, and S. Pankanti, "On the similarity of identical twin fingerprints," *Pattern Recognition*, vol. 35, no. 11, pp. 2653–2663, Nov. 2002.

[7] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Expression-invariant 3D face recognition," in *Audio- and Video-based Biometric Person Authentication*, ser. LNCS, no. 2688, 2003, pp. 62–70.

[8] "CVRL data sets." [Online]. Available: http://www.nd.edu/~cvrl/CVRL/Data_Sets.html

[9] "Konica minolta catalogue." [Online]. Available: http://www.konicaminolta.com/instruments/products/3d/non-contact/vivid910/

[10] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the Face Recognition Grand Challenge," in *Proc. CVPR*, San Diego, CA, USA, Jun. 2005, pp. 947–954.

[11] T. C. Faltemier, K. W. Bowyer, and P. J. Flynn, "A region ensemble for 3D face recognition," *IEEE Trans. on Info. Forensics and Security*, vol. 3, no. 1, pp. 62–73, Mar. 2008.

[12] R. McKeon, "Three-dimensional face imaging and recognition: A sensor design and comparative study," Ph.D. dissertation, University of Notre Dame, 2010.

[13] L. Spreeuwers, "Fast and accurate 3d face recognition," *Int. J. Comput. Vision*, vol. 93, pp. 389–414, Jul. 2011.

[14] D. Huang, G. Zhang, M. Ardabilian, Y. Wang, and L. Chen, "3D face recognition using distinctiveness enhanced facial representations and local feature hybrid matching," in *Proc. BTAS*, Washington D.C., USA, Sep. 2010, pp. 1–7.

[15] D. Huang, M. Ardabilian, Y. Wang, and L. Chen, "A novel geometric facial representation based on multi-scale extended local binary patterns," in *Proc. FG*, Santa Barbara, CA, USA, Mar. 2011, pp. 1–7.

[16] D. Huang, W. Ben Soltana, M. Ardabilian, Y. Wang, and L. Chen, "Textured 3D face recognition using biological vision-based facial representation and optimized weighted sum fusion," in *Proc. CVPR Workshop on Biometrics*, Colorado Springs, CO, USA, Jun. 2011 (In press).

[17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, pp. 91–110, Nov. 2004.

[18] B. Gökberk, M. O. Irfanoglu, and L. Akarun, "3D shape-based face representation and feature extraction for face recognition," *Image and Vision Computing*, vol. 24, no. 8, pp. 857–869, 2006.

[19] M. K. Unnikrishnan, "How is the individuality of a face recognized?" *Journal of Theoretical Biology*, vol. 261, no. 3, pp. 469–474, 2009.

[20] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell, "Face recognition by humans: Nineteen results all computer vision researchers should know about," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 1948–1962, Nov. 2006.

[21] S. H. Zafeiriou, M. Hansen, G. Atkinson, V. Argyriou, M. Petrou, M. Smith, and L. Smith, "The photoface database," in *Proc. CVPR Workshop on Biometrics*, Colorado Springs, Colorado, USA, Jun. 2011, pp. 161–168.

[22] I. A. Kakadiaris, G. Passalis, G. Toderici, M. N. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis, "Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach," *IEEE Trans. PAMI*, vol. 29, no. 4, pp. 640–649, Apr. 2007.

[23] O. Ocegueda, S. K. Shah, and I. A. Kakadiaris, "Which parts of the face give out your identity?" in *Proc. CVPR*, Colorado Springs, CO, USA, Jun. 2011, pp. 641–648.

[24] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data with application to face recognition," *Pattern Recognition*, vol. 34, pp. 2067–2070, Oct. 2001.

[25] A. Savran, N. Alyuz, H. Dibeklioglu, O. Celiktutan, B. Gokberk, B. Sankur, and L. Akarun, "Bosphorus database for 3D face analysis," in *Proc. First COST 2101 Workshop on Biometrics and Identity Management*, Roskilde, Denmark, May 2008, pp. 47–56.