

# Shade Face: Multiple Image-based 3D Face Recognition

Ajmal S. Mian

Computer Science and Software Engineering

The University of Western Australia

ajmal@csse.uwa.edu.au

## Abstract

*Three-dimensional face recognition is illumination invariant, however the acquisition process itself is not. In active 3D recognition, multiple images are captured while the face is actively illuminated with different patterns. We propose a 3D face recognition paradigm that bypasses reconstruction and exploits the plethora of information available in multiple images of a person acquired while varying the illumination. Illumination is varied by scanning a horizontal and then a vertical white stripe on the computer screen in front of the subject. Subtracting ambient light leaves images illuminated by the screen from different angles. The contourlet coefficients of the images are calculated at different scales and orientations and then projected to PCA subspace to remove redundancy. The subspace contourlet coefficients of multiple images are stacked to form a global face representation. Sliding windows are used during matching to remove the disparity between the face locations with respect to the screen. The proposed algorithm was tested under varying ambient conditions and compared to a known 3D face recognition technique. Verification results on data from the same subjects show the strength of our algorithm.*

## 1. Introduction

Facial biometrics can be acquired with non-contact sensors such as cameras and 3D scanners. Therefore, biometric devices based on face recognition are socially more acceptable [9] compared to fingerprints for example. On the downside, face recognition is very challenging because of nuisance factors like changing illumination, facial expressions, ornamentation and pose. These factors combined with the low distinctiveness of facial biometrics [9] makes unconstrained face recognition extremely challenging.

Face recognition is a well studied area and various algorithms exist in the literature. Zhao *et al.* [22] provide a detailed survey of face recognition literature and highlight three main categories namely holistic face recognition tech-

niques which match global features of the complete face [17][2], feature-based techniques which match local features of the face [19] and hybrid techniques which use both holistic and local features. The last category resembles the human perceptual system especially when recognition is based on image sequences where a joint representation in spatial and temporal space can be used for identification [22]. A single image provides sufficient spatial information but the temporal dimension defines trajectories of facial features and body motion characteristics which may further assist classification. Temporal information can also be exploited to obtain a 3D shape of the face using SfM techniques. However, an underlying assumption is that the images must contain non-redundant information either due to the relative motion of the camera and the face or the motion of the facial features *e.g.* due to change in facial expression. This implies substantial acquisition times. If multiple images of a face are acquired instantly (say at 25 frames per second) from a single viewpoint, these images would be mostly redundant and the temporal dimension will not have much information.

We will come back to the image sequence based face recognition but first let us talk about an interesting face recognition alternative based on 3D face models. 3D face recognition has gained considerable popularity due to its invariance to illumination because it matches the facial shape as opposed to the reflected light. However, the acquisition process of 3D faces is not illumination invariant. Changes in illumination can have a great impact on the accuracy and completeness of 3D shape data. Dark regions such as eyebrows and specularities can cause missing data or spikes. Moreover, 3D scanners have a very limited depth of field. These problems are discussed in detail by Bowyer *et al.* [4] in their recent survey of 3D face recognition algorithms. They also highlight that the comparison of 2D and multimodal 2D-3D face recognition is usually biased as the latter involves more than one images. They suggest that two gallery images must be matched with two probe images for a correct comparison, overlooking the fact that active 3D shape acquisition usually involves more than two images.

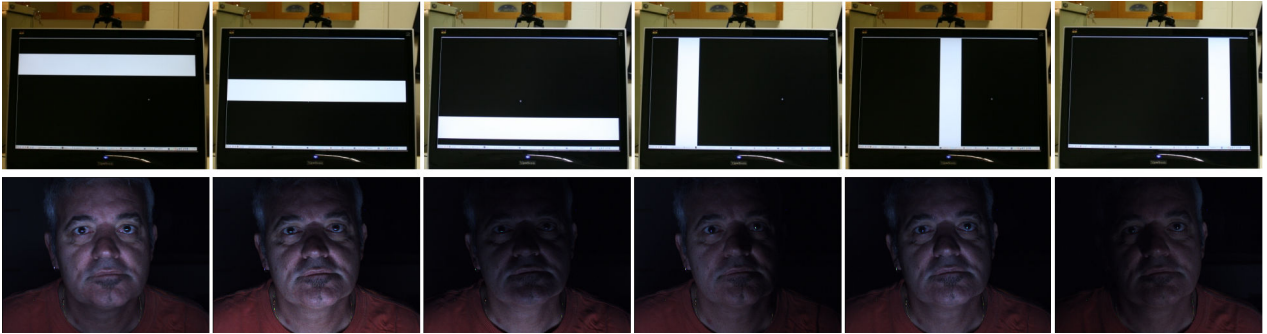


Figure 1. Illustration of the proposed face recognition paradigm. A camera acquires multiple images of a subject while a white stripe is scanned on the screen to vary the illumination.

Passive techniques, such as stereo, cannot provide accurate and dense 3D models therefore, we will focus on active structured light based techniques. These techniques require camera and projector calibration prior to acquisition. Calibration must be repeated every time the relative position of the camera and projector is disturbed. After calibration, a face is actively illuminated with different patterns and its multiple images are acquired. These images are processed to reconstruct the 3D model of the face and invariant features are then extracted from the 3D face to perform classification.

In this paper, we propose a novel 3D face recognition paradigm, *i.e.* Shade Face, which bypasses the reconstruction part and exploits the plethora of information available in multiple images of a person acquired while varying the illumination. We argue that if person identification is the objective, why waste CPU time in reconstructing a 3D face model or impose constraints like the availability of a 3D scanner. With the freedom to vary illumination, perhaps more simple, efficient and yet accurate face recognition can be performed that does not require expensive equipment or cumbersome calibration procedures. Recall our discussion about redundancy in instantly acquired multiple images. In the proposed paradigm, the facial images are acquired instantly but each time varying the illumination therefore, they are non-redundant.

In our setup, illumination is varied by scanning a horizontal and then a vertical white stripe (with black background) on the computer screen in front of the subject. Fig. 1 shows an illustration of our approach. Subtracting ambient light leaves images illuminated by the screen from different angles. The contourlet coefficients [7] of the images are calculated at different scales and orientations and then projected to PCA subspace to remove redundancy. The subspace contourlet coefficients of multiple images are stacked to form a global face representation. Since a subject's position can change with respect to the screen when he/she is first enrolled and then recognized at later stage, sliding

windows are used during matching to remove the disparity between the face locations. The proposed algorithm was tested under varying ambient illuminations and compared to a known 3D face recognition algorithm [14] using data acquired from the same subjects. In the case of Shade Face algorithm, multiple images were acquired using the technique illustrated in Fig. 1 whereas for 3D face recognition, 3D face data was acquired with a Minolta laser scanner. Our results show that Shade Face outperforms the 3D face recognition algorithm [14].

There are many applications for the proposed Shade Face algorithm. It can be used for authenticated login to a computer account without the need for additional hardware. Many other places *e.g.* ATMs and immigration control, where authentication is currently performed using cards and/or PINs already have a screen and camera in place. Shade Face can be deployed at such locations with no additional hardware.

## 1.1. Related Work

The closest work to ours in the existing literature is perhaps the illumination invariant face recognition from NIR (near infra-red) images [5]. In this technique, the face is actively illuminated with NIR LEDs that are coaxial with the NIR camera hence causing frontal illumination. Our method is different from [5] because it actively illuminates the face from different angles. Another difference is that we operate in the visible frequency range using a commercial camera and the computer screen for illumination. However, our proposed technique is generic and can be applied to NIR images as well.

Shape-illumination manifolds [1] have been used to represent a face under changing illumination conditions. This technique finds the best match to a video sequence in terms of pose and then re-illuminates them based on the manifold. This approach is different from ours because it assumes the presence of pose variations *i.e.* images are acquired over a longer duration. Appearance manifolds under

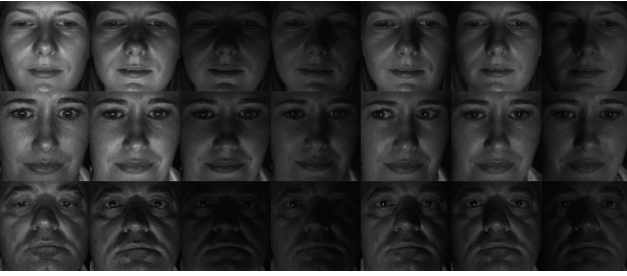


Figure 2. Sample faces after preprocessing.

changing pose were also used by Lee and Kriegman [11] to perform face recognition. In their setup, the subject is instructed to vary his/her facial pose so that the manifolds can be learned. This not only requires more acquisition time but is less friendly (imagine a user instructed by a computer to change pose for logging in) compared to instantly acquiring multiple images while varying the illumination.

Li *et al.* [20] extracted the shape and pose free facial texture patterns from multi-view face images and used KDA for classification. Liu *et al.* [10] perform online learning for multiple image based face recognition. Unlike other techniques they do not use a pre-trained model. Tangelder and Schouten [8] used a sparse representation of multiple still images for face recognition. A common aspect of these and other multiple image or video-based face recognition techniques is that they rely on changes in pose or long term changes to extract additional information which implicitly means longer acquisition times. Moreover, they do not actively illuminate the face but rely on ambient illumination. In contrast, we actively illuminate the face from different angles to instantly acquire non-redundant multiple images for robust face recognition.

Wavelets provide a time-frequency representation of signals. Gabor wavelets have been well studied for face recognition and many variants exist in the literature *e.g.* [19][21][12]. A recent survey of wavelets based face recognition is given in [16]. Wavelets are good at analyzing point (or zero dimensional) discontinuities and are therefore, suitable for analyzing one dimensional signals. Since images are inherently two dimensional, to capture one dimensional discontinuities (such as curves), contourlets [7] have been proposed. The contourlet transform performs multi-resolution and multi-directional decomposition of images allowing for different number of directions at each scale [7]. Contourlet [7] is a relatively new transform and has not been fully explored for face recognition.

## 1.2. Contributions of this Work

The contributions of this paper can be summarized as follows. (1) A novel 3D face recognition paradigm which bypasses the 3D reconstruction part. (2) The use of com-

puter screen to illuminate the face from different angles. (3) A multiple image-based face recognition algorithm using the discrete contourlet transform [7]. (4) Comparison of the proposed Shade Face algorithm with a known 3D face recognition algorithm [14]. (5) Provision of a novel database (which will be made publicly available) comprising multiple images of faces under varying illumination and their corresponding 3D face models acquired with the Minolta 3D scanner. This database is the first of its kind and can be used for comparing multiple image-based and 3D face-based recognition algorithms. It also provides ground truth 3D face data to evaluate the performance of 3D face reconstruction algorithms from single and multiple images.

## 2. Proposed Algorithm

The proposed algorithm *i.e.* *Shade Face*, requires a computer screen and a fixed exposure camera. Many computer screens now come with inbuilt webcams therefore, no additional hardware is required by our algorithm. For a good signal to noise ratio, the subject must not be far from the screen. The output of the camera is displayed on the screen so that the subject can approximately center his/her face in the camera's fov. Image capture is automatically initiated using face detection [18] when the face is correctly positioned (or manually with the press of a button). The screen goes black and a white stripe scans from the top of the screen to the bottom (vertical scan). The thickness of the stripe and scanning speed (*i.e.* step) are chosen as a fraction of the screen size. This is followed by a horizontal scan with a stripe of equal thickness and step size. In our experiments, the stripe was 200 pixels thick and 17 images were captured during vertical scan and 30 during horizontal scan (given the aspect ratio of the screen). A final image is captured in ambient light while the screen is turned off (or black) and subtracted from all other images. All images are converted to gray scale, normalized with respect to scale based on the eye corners and the face region is cropped. Note that the eye corner detection can be accurately performed on the basis of all 47 images given that they are captured instantly *i.e.* no subject movement. See Fig. 2 for sample images.

Let  $v_i$  represent the preprocessed vertically scanned images (where  $i = 1 \dots 17$ ) and  $h_j$  represent the horizontally scanned ones (where  $j = 1 \dots 30$ ). Contourlet coefficients [7] of the images are then calculated at different scales and orientations. Let  $v_i^{s,k}$  represent the vector of contourlet coefficients of image  $v_i$  at scale  $s$  and orientation  $k$ . The contourlet transform is 33% redundant [7] and the images have even greater redundancy because they belong to the same subject or the same class *i.e.* human faces. To remove redundancy, the contourlet coefficients of all training images (belonging to different subjects) calculated at the same scale and orientation are projected separately to the PCA subspaces.

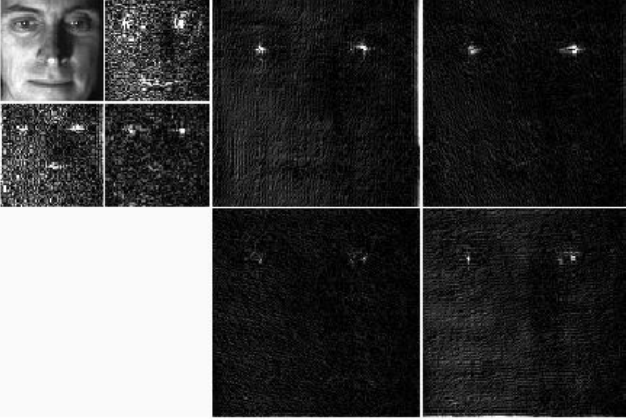


Figure 3. Contourlet coefficients of a face.

Let  $\mathbf{A}^{sk} = [v_{ig}^{sk} \ h_{jg}^{sk}]$  (where  $i = 1 \dots 17$ ,  $j = 1 \dots 30$  and  $g = 1 \dots G$ ) represent the matrix of contourlet coefficients of all 47 images of  $G$  subjects in the training data at the same scale  $s$  and same orientation  $k$ . Each column contains the contourlet coefficients of one image. The mean of the matrix is given by

$$\mu^{sk} = \frac{1}{47G} \sum_{n=1}^{47G} \mathbf{A}_n^{sk}, \quad (1)$$

and the covariance matrix by

$$\mathbf{C}^{sk} = \frac{1}{47G} \sum_{n=1}^{47G} (\mathbf{A}_n^{sk} - \mu^{sk})(\mathbf{A}_n^{sk} - \mu^{sk})^T. \quad (2)$$

The eigenvectors of  $\mathbf{C}^{sk}$  are calculated by Singular Value Decomposition

$$\mathbf{U}^{sk} \mathbf{S}^{sk} (\mathbf{V}^{sk})^T = \mathbf{C}^{sk}, \quad (3)$$

where the matrix  $\mathbf{U}^{sk}$  contains the eigenvectors sorted according to the decreasing order of eigenvalues in the diagonal matrix  $\mathbf{S}^{sk}$ . Let  $\lambda_n$  (where  $n = 1 \dots N$ ) represent the eigenvalues (in decreasing order) then we find  $L$  such that

$$\frac{\sum_{n=1}^L \lambda_n}{\sum_{n=1}^N \lambda_n} \approx 0.95, \quad (4)$$

$$\mathbf{U}_L^{sk} = \mathbf{U}_n^{sk} \text{ where } n = 1 \dots L. \quad (5)$$

Where  $\mathbf{U}_L^{sk}$  contains the first  $L$  eigenvectors of  $\mathbf{C}^{sk}$ . The subspace contourlet coefficients are given by

$$\mathbf{A}_\lambda^{sk} = (\mathbf{U}_L^{sk})^T \mathbf{A}^{sk}. \quad (6)$$

Where  $\mathbf{A}_\lambda^{sk} = [v_{\lambda i}^{sk} \ h_{\lambda j}^{sk}]$ . Note that  $\mathbf{U}_L^{sk}$  and  $\mu^{sk}$  represent the subspace for contourlet coefficients at scale  $s$  and

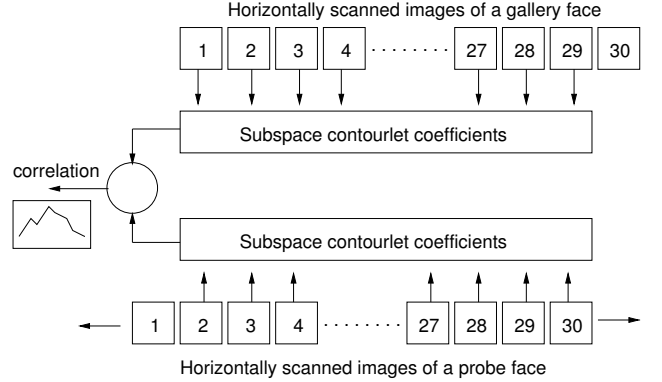


Figure 4. Illustration of sliding window.

orientation  $k$ . Similar subspaces are calculated for different scales and orientations using the training data. In our experiments, we considered two scales and seven orientations along with the low pass sub-band image (see Fig. 3).

The subspace contourlet coefficients of vertical and horizontal scan images are stacked separately to form two global representations of each identity. For example, the global features of an identity  $g$  in the gallery are given by

$$\mathbf{T}_g^v = [v_{\lambda i}^{11} \ v_{\lambda i}^{12} \ \dots \ v_{\lambda i}^{21} \ v_{\lambda i}^{22} \ \dots]^T, \text{ and} \quad (7)$$

$$\mathbf{T}_g^h = [h_{\lambda j}^{11} \ h_{\lambda j}^{12} \ \dots \ h_{\lambda j}^{21} \ h_{\lambda j}^{24} \ \dots]^T. \quad (8)$$

Where each column of  $\mathbf{T}$  corresponds to an image and row corresponds to a specific subspace contourlet coefficient. Recall that  $i = 1 \dots 17$  and  $j = 1 \dots 30$ .

During recognition, similar global features  $\mathbf{P}_\lambda^h$  and  $\mathbf{P}_\lambda^v$  are calculated for the images of a probe face. Note that the subject's position with respect to the screen is not strictly controlled therefore, there may be a shift (vertical and horizontal) in the position when the subject is first enrolled and later being recognized. Consequently, images  $v_i$  and  $h_j$  (for the same values of  $i$  and  $j$ ) may correspond to different lighting conditions of the face images. To overcome this disparity, a sliding window is used while matching the probe to each gallery face. This strategy is similar to matching IrisCodes [6] except that our features are not circular like the IrisCodes and therefore, sliding the window by  $x$  images (or columns) will result in the exclusion of  $x$  feature columns on different sides of the gallery and probe features. Fig. 4 illustrates the sliding window based feature matching technique for horizontal features. Vertical scan features are also matched using the same sliding window approach. Matching is performed using correlation coefficient

$$\gamma = \frac{n \sum \mathbf{T} \mathbf{P} - \sum \mathbf{T} \sum \mathbf{P}}{\sqrt{n \sum (\mathbf{T})^2 - (\sum \mathbf{T})^2} \sqrt{n \sum (\mathbf{P})^2 - (\sum \mathbf{P})^2}}, \quad (9)$$

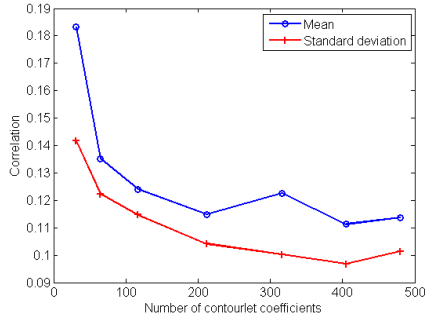


Figure 5. The mean and standard deviation of correlation between different identities decrease as more subspace contourlet coefficients are added in the global representation. Saturation is reached at about 400 coefficients.

where  $\gamma$  is the correlation coefficient and  $n$  is the number of remaining subspace contourlet coefficients inside the sliding window. Given the limited fov of the camera, the window needs to be slid over only a limited number of images (say  $\pm 3$ ) while matching with each gallery identity. This results in a  $7 \times G$  matrix of correlations where each row corresponds to a sliding window position and a column contains the correlations of the corresponding gallery identity with the probe. The window (*i.e.* row) that contains the maximum correlation  $\gamma$  for a gallery face is selected as a row vector of similarity scores of the probe with the gallery identities. Note that this is in contrast with the IrisCodes matching [6] where a different tilt is selected to give a minimum Hamming distance between two IrisCodes. Our approach is not biased towards higher correlations between unrelated faces because we chose the same sliding window (row) position for all matches.

Sliding window is used to match the vertical and horizontal scan features separately resulting in two vectors  $\Gamma^v$  and  $\Gamma^h$  of correlation coefficients. These correlations are combined using a weighted sum rule where the weights are calculated as the difference between the top two correlation coefficients *i.e.* if the best match is close to the next best match, it gets a lower weight. The combined correlation between the probe and gallery identities is calculated as

$$\Gamma = \omega^v \Gamma^v + \omega^h \Gamma^h, \quad (10)$$

where  $\omega^v$  and  $\omega^h$  are the weights of the vertical and horizontal correlation scores. In identification applications, the gallery identity with the maximum combined correlation is declared as the identity of the probe. In authentication, a threshold is used to decide whether the probe should be allowed access or not.

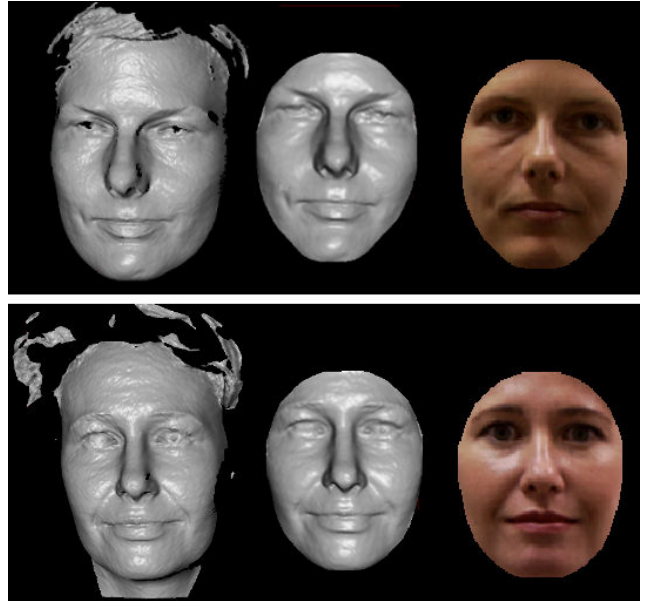


Figure 6. Sample 3D faces (left) acquired with the Minolta laser scanner in our lab. The 3D faces are normalized (middle) along with their texture (right) according to [14].

### 3. Results and Comparison with 3D Face Recognition

The proposed Shade Face approach is completely different and no standard database exists for testing it. Therefore, we generated our own data of multiple face images and their corresponding 3D face models for comparison. The multiple images were acquired with a  $640 \times 480$  camera with fixed exposure while the face was illuminated by a 22" TFT screen with  $280 \text{ cd/m}^2$  brightness and 700:1 contrast ratio. The 3D face models were acquired with the Minolta Vivid 3D scanner. A total of ten subjects participated in our experiments. Each subject was imaged three or more times by Shade Face on different days and under different ambient lighting. This resulted in a total of 1551 images. Out of these, 470 images (one set of vertical and horizontal scan images per subject) were used in training data and to form the gallery while the remaining 1081 images (in sets of 47 images) were used for testing the algorithm *i.e.* probe identities. The Shade Face algorithm was used to calculate the correlation between the probe and gallery identities. All correlation scores were normalized on the scale of zero to one using the min-max rule. Normalization was necessary for comparison with the 3D face recognition algorithm discussed later.

In our first experiment, we studied the effect of the number of subspace contourlet coefficients on the mean and standard deviation of the correlation between different identities which should be ideally very small. Only 470 test

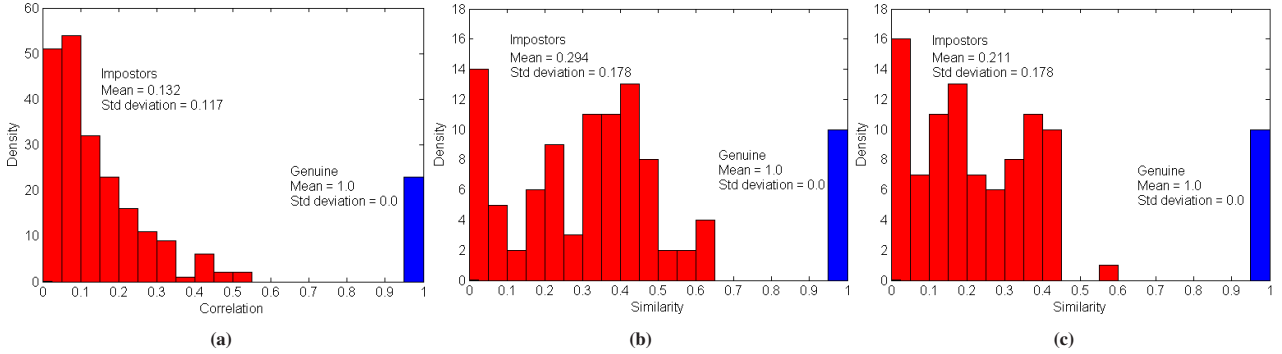


Figure 7. (a) Shade Face algorithm performance. (b) R3D face recognition algorithm [14] performance. (c)  $MMH_a$  multimodal 2D-3D face recognition algorithm [14] performance.

images were used in this experiment. Fig. 5 shows the relationship between the number of subspace contourlet coefficients and the mean and standard deviation of correlation between different identities. The subspace contourlet coefficients were added starting from the smallest scale. Since the mean correlation saturates at about 400 coefficients which corresponds to two scales and seven orientations, we chose this many coefficients for our next experiment.

In the second experiment, we used all the 1081 test images corresponding to 23 imaging sessions of 10 probe identities and matched them with the 10 identities in the gallery. It is important to emphasize that a single set of vertical and horizontal images were used for the recognition of a probe identity. The matching process resulted in 23 genuine correlation scores when a probe was matched with the correct gallery identity and 207 impostor scores when different identities were matched. Fig. 7-a shows the impostor and genuine distributions for our Shade Face algorithm. Due to score normalization, all genuine matches have a correlation of 1 but the distribution of the impostor scores gives us important information about how well the Shade Face algorithm performed. The mean correlation of the impostor distribution is 0.13 indicating that the proposed global features are highly distinctive. The mean impostor correlation (on a normalized scale) can be used as a benchmark to compare two algorithms when both give a 100% identification and verification accuracy.

In the third experiment, we compare Shade Face to an existing 3D and multimodal 2D-3D face recognition algorithm [14]. This algorithm was chosen because of its high identification and verification performance (over 99%) on the largest available 3D face database *i.e.* FRGC [15]. We used our own implementations of the R3D and  $MMH_a$  variants of the algorithm [14]. The R3D algorithm performs region based 3D face recognition. It detects the nose tip, crops the 3D face using a sphere, uniformly resamples it and corrects its pose using PCA. The last two steps are iterated until pose converges. Texture is mapped on the 3D face

during the normalization procedure to get a cropped and pose corrected 2D image of the subject as well. The eyes-forehead and nose regions are then segmented and matched separately using a modified ICP algorithm [3]. The match scores are normalized on a scale of zero to one and combined using the sum rule.

Each subject was scanned twice on different days using the Minolta laser scanner. Fig. 6 shows sample scans of two subjects and their corresponding 3D and 2D faces after normalization. For each subject, the first scan was kept in the gallery and the second was treated as a probe and the R3D algorithm was used to match them resulting in 90 impostor scores and 10 genuine scores. For comparison to the Shade Face algorithm, the similarity scores were normalized on the scale of zero to one using positive polarity (as opposed to negative polarity in [14]). Fig. 7-b shows the impostor and genuine distributions of the R3D algorithm. Note that the mean impostor score is worst than the Shade Face algorithm.

The  $MMH_a$  is a multimodal 2D-3D face recognition algorithm and combines the similarity scores of the R3D algorithm with SIFT [13] and SFR (Spherical Face Representation) based face recognition [14]. A sum rule is used for combining the normalized similarity scores. We gave equal weights to all matching scores and normalized the final scores again on the scale of zero to one for comparison with the previous results. Fig. 7-c shows the performance on the  $MMH_a$  algorithm. As expected,  $MMH_a$  outperforms R3D by achieving lower mean impostor similarity of 0.21 however, its performance still does not match that of the proposed Shade Face algorithm.

## 4. Discussion

The proposed Shade Face algorithm is inspired by 3D scanning technology but uses common computer hardware in an ingenious way. This paper presents a proof of concept rather than being exhaustive in multiple image representa-

tion and testing or critical in setting the best hardware configuration. While, the proposed global features and comparative analysis are sufficient to show the potential of the Shade Face algorithm, there are many improvements that can be done. For example, local features can be added to the global features to perform hybrid face recognition. The hardware setting can be improved by synchronizing the camera trigger with refresh rate of the screen. Faster cameras and screens can further improve the results by avoiding the slightest subject movement. Nevertheless, Shade Face is robust as with the existing unoptimized setup, it outperformed a known 3D face recognition algorithm.

## 5. Conclusion

We presented a novel 3D face recognition paradigm which bypasses the 3D reconstruction part. The proposed approach extracts global features from multiple images of a face acquired while actively varying the illumination. We also introduced the use of the computer screen to illuminate the face from different angles. Many computer screens come with inbuilt cameras therefore, alleviating the need for any additional hardware for our algorithm. We proposed a multiple image-based face recognition algorithm using the discrete contourlet transform [7]. To the best of our knowledge, the contourlet transform has not been used before for face recognition. We compared the proposed algorithm to a known 3D face recognition algorithm [14]. Verification results on images and 3D face models of the same persons show the superiority of our algorithm. We also presented a novel database comprising multiple facial images under varying illumination and their corresponding 3D face models acquired with the Minolta 3D scanner. We plan to grow this database and make it publicly available for researchers to compare multiple image-based and 3D face recognition algorithms. The 3D face data can be used as ground truth to evaluate the performance of 3D face reconstruction algorithms from single and multiple images.

## Acknowledgments

Thanks to M.Do for the Contourlet Toolbox. This research is sponsored by ARC grant DP0881813.

## References

- [1] O. Arandjelovic and R. Cipolla. Face Recognition from Video Using the Generic Shape-Illumination Manifold. In *ECCV*, pages 27–40, 2006. 2
- [2] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Trans. on PAMI*, 19:711–720, 1997. 1
- [3] P. J. Besl and N. D. McKay. A Method for Registration of 3-D Shapes. *IEEE Trans. on PAMI*, 14(2):239–256, 1992. 6
- [4] K. W. Bowyer, K. Chang, and P. Flynn. A Survey Of Approaches and Challenges in 3D and Multi-modal 3D + 2D Face Recognition. *CVIU*, 101:1–15, 2006. 1
- [5] R. Chu, S. Liao, and L. Zhang. Illumination Invariant Face Recognition Using Near-Infrared Images. *IEEE Trans. on PAMI*, 29(4):627–639, 2007. 2
- [6] J. Daugman. Probing the Uniqueness and Randomness of IrisCodes: Results From 200 Billion Iris Pair Comparisons. *Proceedings of the IEEE*, 94(11):1927–1935, 2006. 4, 5
- [7] M. N. Do and M. Vetterli. The Contourlet Transform: an Efficient Directional Multiresolution Image Representation. *IEEE Trans. on Image Processing*, 14(12):2091–2106, 2005. 2, 3, 7
- [8] J. Tangelder and B. Schouten. Learning a Sparse Representation from Multiple Still Images for On-Line Face Recognition in an Unconstrained Environment. In *ICPR*, pages 10867–1090, 2006. 3
- [9] A. K. Jain, A. Ross, and S. Prabhakar. An Introduction to Biometric Recognition. *IEEE Trans. on Circuits and Systems for Video Technology*, 14(1):4–20, 2004. 1
- [10] L. Liu and Y. Wang and T. Tan. Online Appearance Model Learning for Video-Based Face Recognition. In *IEEE CVPR*, pages 1–7, 2007. 3
- [11] K. Lee and D. Kriegman. Online Probabilistic Appearance Manifolds for Video-based Recognition and Tracking. In *CVPR*, volume 1, pages 852–859, 2005. 3
- [12] C. Liu and H. Wechsler. Face Recognition Using Independent Gabor Wavelet Features. In *Audio- and Video-Based Biometric Person Authentication*, pages 20–25, 2001. 3
- [13] D. Lowe. Distinctive Image Features from Scale-invariant Key Points. *International Journal of Computer Vision*, 60(2):91–110, 2004. 6
- [14] A. Mian, M. Bennamoun, and R. Owens. An Efficient Multimodal 2D-3D Hybrid Approach of Automatic Face Recognition. *IEEE Trans. on PAMI*, 29(11):1927–1943, 2007. 2, 3, 5, 6, 7
- [15] 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 *IEEE CVPR*, pages 947–954, 2005. 6
- [16] L. Shen and L. Bai. A Review on Gabor Wavelets for Face Recognition. *Pattern Analysis and Appl.*, 9:273–292, 2006. 3
- [17] M. Turk and A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3:71–86, 1991. 1
- [18] P. Viola and M. J. Jones. Robust Real-Time Face Detection. *IJCV*, 57(2):137–154, 2004. 3
- [19] L. Wiskott, J. M. Fellous, N. Kruger, and C. Malsburg. Face Recognition by Elastic Bunch Graph Matching. *IEEE Trans. on PAMI*, 19(7):775–779, 1997. 1, 3
- [20] Y. Li and S. Gong and H. Liddell. Constructing Facial Identity Surfaces for Recognition. *Int. J. Comput. Vision*, 53(1):71–92, 2003. 3
- [21] H. Zhang, B. Zhang, W. Huang, and Q. Tian. Gabor Wavelet Associative Memory for Face Recognition. *IEEE Trans. on Neural Networks*, 16(1):275–278, 2005. 3
- [22] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face Recognition: A Literature Survey. *ACM Computing Survey*, 35(4):399–458, 2003. 1