Robust face recognition by an albedo based 3D morphable model

Guosheng Hu, Chi Ho Chan, Fei Yan, William Christmas, Josef Kittler
Centre for Vision, Speech and Signal Processing
University of Surrey
Guildford, GU27XH, UK

{g.hu,chiho.chan,f.yan,w.christmas,j.kittler}@surrey.ac.uk

Abstract

Large pose and illumination variations are very challenging for face recognition. The 3D Morphable Model (3DMM) approach is one of the effective methods for pose and illumination invariant face recognition. However, it is very difficult for the 3DMM to recover the illumination of the 2D input image because the ratio of the albedo and illumination contributions in a pixel intensity is ambiguous. Unlike the traditional idea of separating the albedo and illumination contributions using a 3DMM, we propose a novel Albedo Based 3D Morphable Model (AB3DMM), which removes the illumination component from the images using illumination normalisation in a preprocessing step. A comparative study of different illumination normalisation methods for this step is conducted on PIE and Multi-PIE databases. The results show that overall performance of our method outperforms state-of-the-art methods.

1. Introduction

Face recognition has attracted substantial attention due to not only its wide applications in real-world scenarios ranging from handhold devices to surveillance, but also the impressive improvements in performance achieved over the last 30 years [20]. However, face recognition in an unconstrained environment, such as pose and non-uniform illumination, is still very challenging. Many existing solutions have been proposed to cater for such environments.

The existing pose robust face recognition methods fall into two categories: 2D-based and 3D-based. In the first category, poses are either handled in image space or projection space. Markov Random Field [4], local region based matching [6, 9] and stereo matching [10], establishing the local patch correspondence between the pair of face images being matched, are examples of the image space approach. Linear regression, Canonical correlation analysis and Deep Learning exemplify the projection space approaches [11, 17, 36]. These methods project the images

with different poses into latent variable spaces, where a pose-independent representation is obtained.

In the second category, the most well-known work is the 3D morphable model (3DMM) [7, 8]. In this category, the recognition methods are based on (1) model parameter matching, (2) pose normalisation, (3) pose synthesis and (4) filter transformation. The model parameter matching approach [3, 23] first fits the model to face images in the gallery and probe and then the fitted texture and shape parameters are used for matching. The pose normalisation approach [5, 18] is based on warping 2D face images with arbitrary poses to a canonical (frontal) view, so that the traditional 2D face matcher can be applied to face recognition. On the other hand, the pose synthesis approach [19, 21, 35] synthesizes images with different poses for each gallery image. When matching the gallery and probe images, only virtual images with similar pose to the probe will be chosen. In the filter transformation approach [31], feature extraction filters are adapted according to the pose and shape of the face image.

To speed up the 3D-based approaches, [19, 31] proposed a simplified 3DMM focusing only on the shape fitting. One of the shortcomings is that this model cannot handle uncontrolled lighting. Therefore, Gabor filters [31] are applied to overcome this problem. However, a face under non-uniform lighting has shadows that most texture descriptors, such as Gabor, HOG and LBP, cannot handle. In general, illumination normalisation is applied to convert the face image to a more canonical form in which illumination variations are suppressed. Surprisingly, there is no research that studied the effect of the illumination normalisation on face recognition under both illumination and pose changes. In this paper, we propose a novel Albedo Based 3D Morphable Model (AB3DMM), which is built using illumination-normalised images.

The contributions of this paper are as follows.

- In AB3DMM, the illumination component is removed from the image, leaving the face albedo. This image can then be used as input to the AB3DMM fitting that

does not handle the lighting parameters. Thus, the fitting of the AB3DMM becomes easier and more accurate.

- The AB3DMM can flexibly be embedded into other face recognition systems for feature extraction. In this paper, the AB3DMM is used to extract pose- and illumination-invariant features.

- To investigate the robustness of the existing illumination normalisation approaches to pose changes, a comparative study is conducted.

- The recognition rates across pose and illumination on the PIE and MultiPIE databases are significantly improved by our proposed approach over the state of the art.

In this paper, the 3DMM is introduced in Section 2. In Section 3, our proposed system is described. The experiments conducted and the results achieved are discussed in Section 4. Lastly, the conclusions are drawn in Section 5.

2. Related work

By virtue of a fitting process, a 3DMM can recover the face (shape and texture) and scene properties (illumination and camera model) from a single image. This section reviews related work on fitting algorithms.

The original version of 3DMM fitting [7, 8] is based on Stochastic Newton Optimisation (SNO). The performance of SNO, being a gradient-based method, is badly affected by local minima. In addition, there are not many constrains on shape and texture in the process of fitting, so that SNO tends to either over-fit or under-fit.

A linear shape and texture fitting (LiST) [23] strategy was proposed for improving the fitting efficiency. The idea is to update the shape and texture parameters by solving a linear system. However, the illumination and camera parameters are optimised by the Levenberg-Marquardt method, which is gradient-based. Not surprisingly, the illumination and camera parameters tend to get trapped into local minima. The experiments reported in [23] show that the fitting is faster than the SNO algorithm, but with similar fitting accuracy.

A very recent work [3] also improves greatly the optimisation efficiency. It decomposes the optimisation into geometric and photometric parts, which leads to closed-form solutions of the optimisation problems. The reflectance estimation is split into the diffuse and specular reflection estimation. First, the RGB values of the model and input images are projected to a specularity-free space for diffuse light estimation. Then the specularity is estimated in RGB colour space. However, [3] uses the affine camera model, whose representation ability is relatively weak.

A Multi-Feature Fitting (MFF) strategy in [24] is proposed to improve the fitting accuracy. MFF makes use of complementary features extracted from both the input image and 3DMM to constrain the fitting process. Specifically, MFF exploits texture edges, specular highlights and texture constraint features. As the use of these features reduces the chance for the fitting algorithm to be trapped in local minima, MFF achieves a better solution as demonstrated by [24]. However, these complementary features are very computationally expensive.

All the existing 3DMMs try to estimate the proportion of the albedo and illumination contributions in a pixel intensity in the process of fitting, which is a difficult task. Incorrect proportion estimation will cause bad fitting performance. In this work, illumination normalisation methods are used to preprocess the input images to generate the albedo component. It makes the fitting process easier and more accurate.

3. Methodology

The proposed illumination- and pose-robust face recognition using the AB3DMM and the traditional face matcher is presented in this section. Our system consists of training and testing stages. The construction of a 3D face model involving registration and model training is described in Section 3.1.1 and 3.1.2. The testing pipeline consists of model fitting, pose normalisation and face matching which are detailed in section 3.1.3 and 3.2.

3.1. AB3DMM

In contrast to the traditional 3DMM, we first remove the illumination component from the face texture associated with 3D face image training data. The description of AB3DMM is detailed in the following sections.

3.1.1 3D face registration

Our Albedo Based 3D Morphable Model (AB3DMM) is trained using a local 3D face database which contains 3D shape and texture of 150 subjects captured with a 3D scanner. With this database, face registration is conducted to establish dense correspondences between the 3D faces of different subjects, which implies all the 3D faces will be represented by the same number of vertices and organised in the same manner. In this work, the Iterative Multiresolution Dense 3D registration (IMDR) method [22] is used to perform the 3D non-rigid registration. Figure 1 demonstrates the difference of registrations between the traditional 3DMM and our AB3DMM. After registration, the 3D face, presented in a vector form of shape and texture ($\mathbf{s}$ and $\mathbf{t}$) can be conveyed by:

$$\mathbf{s} = (x_1...x_n, y_1...y_n, z_1...z_n)^T \qquad (1)$$

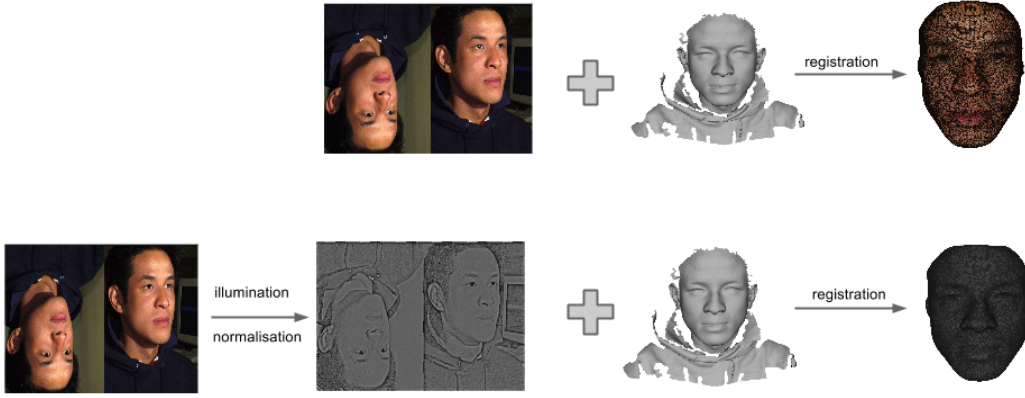$$\mathbf{t} = (g_1, ..., g_n)^T \qquad (2)$$

Figure 1. Registration comparisons between the 3DMM (1st row) and AB3DMM (2nd row). Left to right, 1st row: RGB texture map, shape and registered face. Left to right, 2nd row: RGB and preprocessed texture map, shape, and registered face.

where (x_i, y_i, z_i) is the coordinate of the ith vertex, g_i represents the grey value of the ith vertex, and n denotes the total number of registered vertices.

3.1.2 Model training

Given a set of $\mathbf{s}$ and $\mathbf{t}$, principal component analysis (PCA) is performed on the shape and texture vectors separately to decorrelate the data:

$$\mathbf{s} = \mathbf{s}_0 + \sum_{i=1}^{m-1} \alpha_i \mathbf{s}_i \qquad (3)$$

$$\mathbf{t} = \mathbf{t}_0 + \sum_{i=1}^{m-1} \beta_i \mathbf{t}_i \qquad (4)$$

where m is the size of the face training set, $\mathbf{s}_0$ and $\mathbf{t}_0$ are the mean shape and texture respectively, and $\mathbf{s}_i$ and $\mathbf{t}_i$ are the ith eigenvectors of the shape and texture covariance matrices. In order to constrain the shape and texture coefficients, it is assumed that they have normal distributions:

$$p(\boldsymbol{\alpha}) \sim \exp\left[-\frac{1}{2} \sum_i \frac{\alpha_i^2}{\sigma_{i,s}^2} \right], \qquad (5)$$

$$p(\boldsymbol{\beta}) \sim \exp\left[-\frac{1}{2} \sum_i \frac{\beta_i^2}{\sigma_{i,t}^2} \right], \qquad (6)$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_{m-1})^T$ and $\boldsymbol{\beta} = (\beta_1, ..., \beta_{m-1})^T$ are the shape and texture parameters, $\sigma_{i,s}^2$ and $\sigma_{i,t}^2$ denote the ith eigenvalue of the shape and texture covariance matrices, respectively.

3.1.3 Model fitting

A method for fitting the AB3DMM to 2D images is detailed in this section. The fitting process optimises model parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, so that the appearance reconstructed by the AB3DMM matches that of the illumination-normalised input images. In addition to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, a set of camera parameters which controls the virtual camera is also optimised. The original 3DMM explicitly modelled illumination and optimised the corresponding illumination parameters during its fitting process. In contrast, in this method, the illumination component can be omitted from the AB3DMM, because it fits 2D albedo images.

The first step for image fitting is to calculate the 2D image positions of the corresponding model vertices. The object-centred coordinates of a vertex $\mathbf{x} = (x, y, z)^T$ are mapped to a position $\mathbf{w} = (w_x, w_y, w_z)^T$ in the camera coordinate system by the rigid transformation :

$$\mathbf{w} = \mathbf{R}\mathbf{x} + \boldsymbol{\tau} \qquad (7)$$

where $\mathbf{R} \in \mathbb{R}^{3\times3}$ denotes the Euler rotation matrix, and $\boldsymbol{\tau} \in \mathbb{R}^3$ defines a spatial translation. Then $\mathbf{w}$ is projected on the 2D image plane coordinates $\mathbf{p}=(p_x, p_y)$ by a perspective projection:

$$p_x = o_x + f\frac{w_x}{w_z}, \quad p_y = o_y - f\frac{w_y}{w_z} \qquad (8)$$

where f is the focal length, and (o_x, o_y) is the image-plane position of the optical axis.

With Eq. (7) and (8), the illumination-normalised input image can be aligned and organised in the same manner as the model synthesised image. Based on this alignment, the fitting can be performed. The primary goal of the AB3DMM fitting is to minimise the sum of squared differ-

Table 1. Description of illumination Normalisation

| Symbol	Description	Symbol	Description
SSR	Single scale retinex [16]	GRF	Gradientface[34]
HOM	Homomorphic filtering [15]	DCT	Discrete cosine transform based normalisation [12]
SQI	Self quotient image [29]	LSSF	Large and small scale features normalisation [30]
WA	Wavelet-based normalisation [13]	WD	Wavelet-denoising-based normalisation [33]
WEB	Weberface[28]	MSW	Multi-scale weberface [28]
TT	Tan and Triggs normalisation[27]	DOG	Difference of Gaussians filter
RAW	Without illumination normalisation		
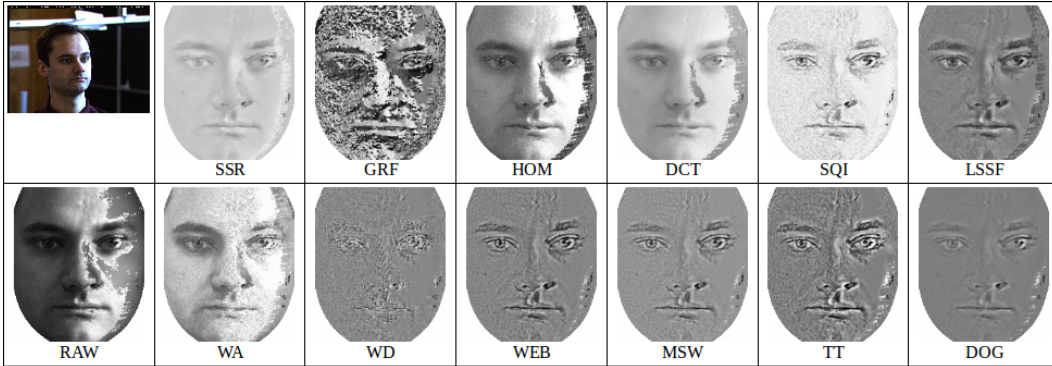
Figure 2. Face images reconstructed by different AB3DMMs.

ences between the synthesised image and the illumination-normalised input image over all pixels. In order to avoid over-fitting, the shape and texture are regularised based on Eq. (5) and (6). Accordingly, the fitting process is formulated as:

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\rho}} \|\mathbf{a}^I(\boldsymbol{\rho},\boldsymbol{\alpha}) - \mathbf{a}^M(\boldsymbol{\beta})\|^2 + \lambda_1 \|\boldsymbol{\alpha}./\boldsymbol{\sigma}_s\|^2 + \lambda_2 \|\boldsymbol{\beta}./\boldsymbol{\sigma}_t\|^2 \tag{9}$$

where $\boldsymbol{\sigma}_s = (\sigma_{1,s}, ...\sigma_{m-1,s})$, $\boldsymbol{\sigma}_t = (\sigma_{1,t}, ...\sigma_{m-1,t})$. $\boldsymbol{\rho}$ is a set of camera parameters: $\{\mathbf{R}, \boldsymbol{\tau}, f\}$, and $./$ is element-wise division, The pixel intensity from an illumination-normalised input image and the AB3DMM are denoted by $\mathbf{a}^I$ and $\mathbf{a}^M$, respectively; $\mathbf{a}^M$ is equal to $\mathbf{t}$ of Eq. (4), λ_1 and λ_2 are two free parameters to weight the shape and texture regularisation terms.

3.2. Face matcher

3.2.1 Pose normalisation

By virtue of the model fitting process, the shape, texture and camera parameters $(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\rho})$ are estimated; thus the occluded pixels in the input image can be reconstructed using the estimated $\boldsymbol{\beta}$. By changing the camera parameters, the input face can be rendered in any given virtual view. In this work, pose normalisation is performed by transforming the input face to the frontal view, which includes most visible facial information. The pixel values of the occluded part are reconstructed by the estimated $\boldsymbol{\beta}$, and those of the visible part are extracted from the illumination-normalised input image, aiming to keep the discriminative facial features.

3.2.2 Face descriptor

After pose normalisation, a face descriptor is extracted from the reconstructed image. We used the local binary pattern histogram (LBPH) [1] and local phase quantisation pattern histogram (LPQH) [2] as descriptors, and the chi-squared distance is applied to measure the distance between gallery and probe descriptors.

4. Experiments

In order to measure the robustness of our framework on face recognition, 12 different illumination normalisation algorithms are evaluated.

4.1. Databases and Protocols

To ensure reproducibility of the experiments and comparability with other methods, we tested our approach on the well-known CMU-PIE [25] and MultiPIE [14] face databases in this work.

A subset of the CMU-PIE database covering both illumination and pose variations is used for evaluation. It is divided into a gallery set containing 68 frontal images of 68 subjects under neutral light, and probe set containing 2,856 images of the same subjects with frontal and side poses under 21 different flashing light directions. The results are summarised by averaging the rank 1 recognition rates under different light directions.

A subset of the MultiPIE database in the first session consisting of 249 subjects with 7 poses from left $45°$ and right $45°$ yaw in steps of $15°$, and 20 illuminations is used.

Table 2. Mean recognition rates of all 68 subjects of the CMU-PIE databases averaged over 22 illumination conditions per pose using a gallery image

	LBP			LPQ		
	front	side	mean	front	side	mean
HOM	97.89	93.20	95.49	99.62	97.67	98.62
WD	96.74	27.94	61.61	98.27	28.49	62.64
GRF	99.55	92.52	95.96	99.94	98.04	98.97
MSW	100.00	88.85	94.31	100.00	99.26	99.62
LSSF	100.00	92.77	96.31	100.00	99.26	99.62
TT	99.62	85.48	92.40	100.00	98.53	99.25
DOG	100.00	93.32	96.59	100.00	99.20	99.59
WEB	100.00	87.93	93.84	100.00	99.26	99.62
SSR	99.68	96.32	**97.97**	100.00	99.39	**99.69**
DCT	100.00	79.23	89.39	100.00	99.02	99.50
WA	92.65	73.16	82.70	98.85	92.65	95.68
SQI	99.87	93.57	96.65	99.87	96.63	98.22
RAW	95.72	88.91	92.24	98.59	96.14	97.34
State-of-the-art-methods						
MFF [24]	98.90	96.10	**97.50**			
Z&S [32]	96.50	94.60	95.50			
3DMM [3]	99.50	95.10	97.30			
3DMM [23]	97.00	91.00	94.00			

The 249 frontal images under neutral illumination are the gallery set and the remaining 34,611 images are the probe set.

4.2. Experimental Set-up

In this study, 12 well-known illumination normalisation algorithms, presented in Table 1, available in the INface toolbox [26], are applied in turn. Each probe image in the testing stage is first illumination-normalised and then fitted to the corresponding AB3DMM for pose normalisation. The reconstructed image is extracted and scaled to a size of 142×120 ($rows \times columns$). The reconstructed images of different illumination normalisations are shown in Figure 2. Then LBP and LPQ operators are applied. The pattern image is then separated into 7×7 non-overlapping regions and the image from each region is summarised by a histogram. The histograms from all the regions are concatenated to form a face descriptor.

4.3. Results of CMU-PIE

In this experiment, pose normalisation using the AB3DMM with 12 different illumination normalisation methods and without any illumination normalisation is studied. The performances of LBP and LPQ face matchers are reported in Table 2. It clearly shows that the performance when using local texture descriptors on the pose normalised image (i.e. RAW+LBP and RAW+LPQ) is higher than that of only using the coefficients of 3DMMs [23, 3]. Comparing the performance of LBP and LPQ systems, the latter is better, but LPQ with illumination normalisation improves the recognition rate even further. Specifically, the best of our systems, SSR+LPQ, achieve 99.69% rank 1 recognition rate, which is around 2% improvement over the best state-of-the art method [24].

4.4. Results of MultiPIE

In this experiment, our LBP and LPQ systems with 11 illumination normalisation methods, excluding WD, are evaluated. Under neutral illumination, Table 4 reports the system performance for different poses and Figure 3 compares our two best performing systems with [36, 18].

The standard face matchers without applying any pose-invariant techniques, i.e., normalised correlation on the raw image (NC) and Eigenface (PCA), are regarded as benchmark systems. Compared with them, our systems and other state-of-the-art methods are significantly better. It also shows that the performance of the LPQ system is better than that of the LBP system. The best of our systems, SSR+LPQ, achieves highest recognition rate. Note that the best of the state-of-the-art methods, RL+LDA, uses training samples for supervised learning. In contrast, our face matcher does not involve any training. It is simple and efficient.

Table 5 reports the average recognition rates over all illumination conditions of 6 different poses and Figure 4 presents the performance of our two best systems. In contrast to Table 4, the performance of all systems, including the state-of-the-art methods, degrades under illumination changes. Under different illumination, the results show that the variance of the performance of different illumination normalisation methods on six different poses is larger than that on just the frontal pose. It shows that face recognition under both illumination and pose changes is very difficult. The best of the state-of-the-art system [36], RL+LDA, achieves 98.4% under only pose change and 74.7% under both pose and illumination changes. The recognition rate of their system drops by around 23%. In contrast, our best system, SSR+LPQ drops by 13% to 86.76%. Table 3 reports the averaged recognition rate over all pose variations of 19 different illumination conditions. We observe that the performance of our best systems degrades under 5 of 19 il-

Table 3. Mean recognition rates of all subjects of the MultiPIE databases averaged over 6 poses per illumination condition using a gallery image

|		00	01	02	03	04	05	06	07
LBP	LSSF	93.31	83.40	77.18	74.56	62.12	41.77	44.11	38.69
	SQI	87.68	78.45	72.82	70.01	59.91	45.45	41.10	36.08
LPQ	SSR	**100.00**	**97.79**	**96.59**	**95.05**	**87.75**	69.08	66.80	61.18
	LSSF	99.80	97.26	94.78	94.24	84.67	69.54	69.08	64.06
Li [18]		51.50	49.20	55.70	62.70	79.50	**88.30**	**97.50**	
RL+LDA [36]		72.80	75.80	75.70	75.70	75.70	75.70	75.70	
		08	09	10	11	12	13	14	
LBP	LSSF	43.37	50.67	67.27	74.63	80.52	83.33	72.56	
	SQI	39.29	48.26	61.31	70.21	75.03	77.04	69.48	
LPQ	SSR	64.06	76.97	**88.69**	**92.70**	**96.32**	96.59	**94.04**	
	LSSF	66.93	76.77	86.01	91.57	95.31	**96.72**	90.63	
Li [18]		**97.70**	**91.00**	79.00	64.80	54.30	47.70	67.30	
RL+LDA [36]		75.70	75.70	75.70	75.70	75.70	75.70	73.40	
		15	16	17	18	19	mean		
LBP	LSSF	64.93	63.25	66.87	75.70	92.84	67.55		
	SQI	61.51	57.56	61.58	67.54	86.75	63.35		
LPQ	SSR	**85.41**	**83.94**	**88.02**	**94.38**	**99.93**	**86.76**		
	LSSF	82.26	**83.94**	85.68	91.77	99.80	86.04		
Li [18]		67.70	75.50	69.50	67.30	50.80	69.32		
RL+LDA [36]		73.40	73.40	73.40	72.90	72.90	**74.77**		
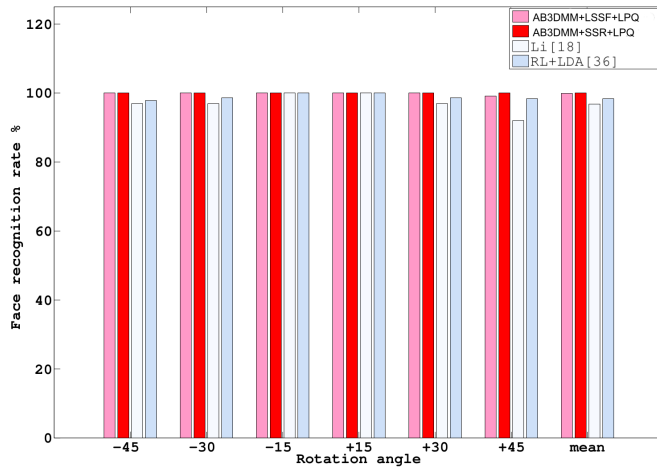
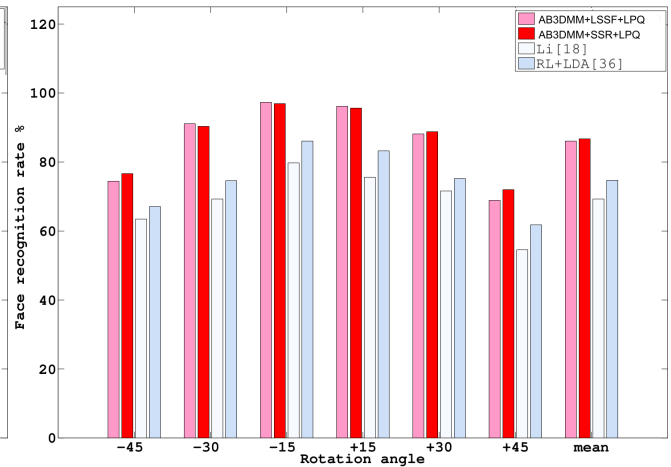Figure 3. Results with pose variations under neutral illumination

Figure 4. Results with pose variations averaging all the illuminations

lumination conditions, but the overall performance of our best system is around 12% better than that of the best of the state-of-the-art [36].

5. Conclusions

We proposed a AB3DMM (Albedo Based 3D Morphable Model) and its fitting strategy, for face recognition under pose and illumination changes. For establishing the AB3DMM, a set of 3D texture-mapping images is preprocessed by illumination normalisation. During testing, the illumination-normalised probe image is fitted by the proposed AB3DMM for pose normalisation. Then the local texture descriptor is extracted from the reconstructed frontal face image. In this paper, a comparative study of 12 different illumination normalisation methods is conducted. The best of our systems, SSR+LPQ, evaluated on the Multi-PIE database with 86.76% recognition rate, outperforms the state-of-the-art method [36].

References

[1] T. Ahonen, A. Hadid, and M. Pietikäinen. Face recognition with local binary patterns. In *ECCV (1)*, pages 469–481, 2004.

[2] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä. Recognition of blurred faces using local phase quantization. In *ICPR*, pages 1–4, 2008.

[3] O. Aldrian and W. A. Smith. Inverse rendering of faces with a 3d morphable model. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(5):1080–1093, 2013.

[4] S. R. Arashloo and J. Kittler. Energy normalization for pose-invariant face recognition based on mrf model image matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(6):1274–1280, 2011.

[5] A. Asthana, T. K. Marks, M. J. Jones, K. H. Tieu, and M. V. Rohith. Fully automatic pose-invariant face recognition via 3d pose normalization. In *ICCV*, pages 937–944, 2011.

[6] T. Berg and P. N. Belhumeur. Poof: Part-based one-vs.-one features for fine-grained categorization, face verification, and attribute estimation. In *CVPR*, pages 955–962, 2013.

[7] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *CGIT*, pages 187–194, 1999.

[8] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(9):1063–1074, 2003.

[9] Z. Cao, Q. Yin, X. Tang, and J. Sun. Face recognition with learning-based descriptor. In *CVPR*, pages 2707–2714, 2010.

[10] C. D. Castillo and D. W. Jacobs. Using stereo matching with general epipolar geometry for 2d face recognition across pose. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(12):2298–2304, 2009.

[11] X. Chai, S. Shan, X. Chen, and W. Gao. Locally linear regression for pose-invariant face recognition. *IEEE Trans. Image Processing*, 16(7):1716–1725, 2007.

[12] W. Chen, M. J. Er, and S. Wu. Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans. Systems, Man, and Cybernetics, B*, 36(2):458–466, 2006.

[13] S. Du and R. Ward. Wavelet-based illumination normalization for face recognition. In *ICIP*, volume 2, pages II–954. IEEE, 2005.

[14] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-pie. *Image and Vision Computing*, 28(5):807–813, 2010.

[15] G. Heusch, F. Cardinaux, and S. Marcel. Lighting normalization algorithms for face verification. *IDIAP-com 05*, 3, 2005.

[16] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell. Properties and performance of a center/surround retinex. *IEEE Trans. Image Processing*, 6(3):451–462, 1997.

[17] T.-K. Kim, J. Kittler, and R. Cipolla. Discriminative learning and recognition of image set classes using canonical correlations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(6):1005–1018, 2007.

[18] A. Li, S. Shan, and W. Gao. Coupled bias-variance trade-off for cross-pose face recognition. *Image Processing, IEEE Transactions on*, 21(1):305–315, 2012.

[19] K. Niinuma, H. Han, and A. Jain. Automatic multi-view face recognition via 3d model based pose regularization. In *BTAS*, pages 1–8, 2013.

[20] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, and W. Worek. Preliminary face recognition grand challenge results. In *FGR*, pages 15–24, April 2006.

[21] U. Prabhu, J. Heo, and M. Savvides. Unconstrained pose-invariant face recognition using 3d generic elastic models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(10):1952–1961, 2011.

[22] J. T. Rodriguez. *3D Face Modelling for 2D+3D Face Recognition*. PhD thesis, Surrey University, Guildford, UK, 2007.

[23] S. Romdhani, V. Blanz, and T. Vetter. Face identification by fitting a 3d morphable model using linear shape and texture error functions. In *ECCV (4)*, pages 3–19, 2002.

[24] S. Romdhani and T. Vetter. Estimating 3d shape and texture using pixel intensity, edges, specular highlights, texture constraints and a prior. In *CVPR*, volume 2, pages 986–993. IEEE, 2005.

[25] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression (pie) database. In *AFGR*, pages 46–51, 2002.

[26] V. Štruc and N. Pavešic. Photometric normalization techniques for illumination invariance. *Advances in Face Image Analysis: Techniques and Technologies, IGI Global*, pages 279–300, 2011.

[27] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In *AMFG*, pages 168–182. Springer, 2007.

[28] B. Wang, W. Li, W. Yang, and Q. Liao. Illumination normalization based on weber's law with application to face recognition. *IEEE Trans. Signal Processing Letters*, 18(8):462–465, 2011.

[29] H. Wang, S. Z. Li, Y. Wang, and J. Zhang. Self quotient image for face recognition. In *ICIP*, volume 2, pages 1397–1400. IEEE, 2004.

[30] X. Xie, W.-S. Zheng, J. Lai, P. C. Yuen, and C. Y. Suen. Normalization of face illumination based on large-and small-scale features. *IEEE Trans. Image Processing*, 20(7):1807–1821, 2011.

[31] D. Yi, Z. Lei, and S. Z. Li. Towards pose robust face recognition. In *CVPR*, pages 3539–3545, 2013.

[32] L. Zhang and D. Samaras. Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(3):351–363, 2006.

[33] T. Zhang, B. Fang, Y. Yuan, Y. Yan Tang, Z. Shang, D. Li, and F. Lang. Multiscale facial structure representation for face recognition under varying illumination. *Pattern Recognition*, 42(2):251–258, 2009.

[34] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu. Face recognition under varying illumination using gradientfaces. *IEEE Trans. Image Processing*, 18(11):2599–2606, 2009.

[35] X. Zhang, Y. Gao, and M. K. H. Leung. Recognizing rotated faces from frontal and side views: An approach toward effective use of mugshot databases. *IEEE Trans. Information Forensics and Security*, 3(4):684–697, 2008.

[36] Z. Zhu, P. Luo, X. Wang, and X. Tang. Deep learning identity preserving face space. In *ICCV*, volume 1, page 2, 2013.

Table 4. Pose Recognition Rates of All Subjects of the MultiPIE under neutral illumination conditions using a single gallery image

		-45	-30	-15	15	30	45	mean
LBP	HOM	99.60	100.00	100.00	100.00	100.00	97.19	**99.46**
	GRF	91.57	100.00	100.00	99.60	99.20	80.72	95.18
	MSW	67.47	97.59	100.00	98.80	91.97	61.04	86.14
	LSSF	83.53	98.80	100.00	100.00	99.20	78.31	93.31
	TT	59.04	93.57	100.00	93.57	93.57	51.00	81.79
	DOG	74.70	98.39	100.00	100.00	99.20	66.67	89.83
	WEB	63.45	96.39	100.00	97.99	94.38	54.22	84.40
	SSR	86.75	99.60	100.00	100.00	98.39	82.73	94.58
	DCT	46.99	91.16	100.00	99.60	77.51	51.00	77.71
	WA	94.78	99.60	100.00	100.00	95.58	86.35	96.05
	SQI	70.68	93.57	99.60	97.99	94.78	69.48	87.68
LPQ	HOM	100.00	100.00	100.00	100.00	100.00	99.20	99.87
	GRF	98.80	100.00	100.00	100.00	100.00	96.79	99.26
	MSW	97.19	100.00	100.00	100.00	99.60	95.98	98.80
	LSSF	100.00	100.00	100.00	100.00	100.00	98.80	99.80
	TT	97.59	100.00	100.00	100.00	99.60	95.18	98.73
	DOG	99.20	100.00	100.00	100.00	99.60	99.20	99.67
	WEB	96.79	100.00	100.00	100.00	99.60	95.98	98.73
	SSR	100.00	100.00	100.00	100.00	100.00	100.00	**100.00**
	DCT	88.76	100.00	100.00	100.00	99.60	91.57	96.65
	WA	100.00	100.00	100.00	100.00	99.60	96.39	99.33
	SQI	93.57	98.80	99.60	99.60	98.80	88.35	96.45
NC [21]		0.40	8.80	15.70	24.90	9.60	1.20	10.10
PCA [21]		1.20	18.50	24.50	31.30	18.10	2.00	15.93
3DGRM+NC [21]		37.10	59.40	75.50	71.10	49.00	45.00	56.18
Li [18]		97.00	97.00	100.00	100.00	97.00	92.00	96.80
RL+LDA [36]		97.80	98.60	100.00	100.00	98.60	98.40	**98.40**

Table 5. Mean Recognition Rates of All Subjects of the MultiPIE Databases Averaged over 20 illumination conditions per Pose using a Gallery image

		-45	-30	-15	15	30	45	mean	front
LBP	HOM	49.80	65.94	76.41	74.12	60.76	47.95	62.50	84.02
	GRF	42.93	69.32	82.53	77.49	63.96	37.71	62.32	93.17
	MSW	28.86	59.46	81.85	73.59	58.15	28.29	55.03	99.09
	LSSF	41.71	73.43	89.88	86.00	71.99	42.31	**67.55**	99.68
	TT	26.06	59.64	84.14	71.49	59.46	27.97	54.79	99.70
	DOG	34.32	69.40	88.49	77.35	70.46	34.42	62.41	99.43
	WEB	26.93	58.05	80.92	73.43	57.79	27.49	54.10	99.43
	SSR	40.54	66.73	86.61	82.21	63.29	37.57	62.82	97.67
	DCT	19.58	47.99	72.63	67.11	37.11	21.85	44.38	89.14
	WA	40.12	52.47	69.96	69.58	49.04	37.89	53.18	80.93
	SQI	38.21	67.75	87.63	80.48	65.78	40.26	63.35	99.09
LPQ	HOM	62.43	78.21	84.66	82.23	73.39	60.64	73.59	94.44
	GRF	65.64	81.77	89.74	89.40	79.42	60.54	77.75	97.55
	MSW	63.35	88.98	97.47	96.12	86.57	59.48	81.99	99.98
	LSSF	74.56	91.08	97.31	96.16	88.21	68.92	86.04	99.92
	TT	63.31	88.67	97.79	96.51	88.39	62.67	82.89	99.98
	DOG	65.38	89.50	98.13	96.87	88.88	64.82	83.93	99.98
	WEB	63.98	89.30	97.29	96.59	88.01	62.21	82.89	99.96
	SSR	76.67	90.38	97.05	95.62	88.82	72.05	**86.76**	99.81
	DCT	40.04	80.18	89.78	88.39	65.00	40.56	67.33	97.44
	WA	55.94	69.94	80.96	79.14	63.80	50.06	66.64	91.00
	SQI	55.84	82.97	95.56	92.25	81.95	57.05	77.60	99.64
Li [18]		63.50	69.30	79.70	75.60	71.60	54.60	69.30	
RL+LDA [36]		67.10	74.60	86.10	83.30	75.30	61.80	**74.70**	