ISBN 978-93-84422-22-6
Proceedings of 2015 International Conference on Image Processing, Production and Computer Science (ICIPCS'2015)
Istanbul (Turkey), June 3-4, 2015 pp. 326-332

Face Recognition under Difficult Lighting Conditions and Occlusions

Huda M.S. Algharib
University of Ankara Yildirim Beyazit, Turkey
Huda_selem@yahoo.com

Abstract: *In recent years face recognition has received substantial attention from researchers in biometrics, pattern recognition, and computer vision communities. At least two reasons account for this trend: the first is the wide range of commercial and law enforcement applications, and the second is the availability of feasible technologies in recent years of research. Many existing methods in face recognition area perform well under certain conditions, but still facing challenging with illumination changes and occlusions. This paper attempts to deal with the above challenges by combining robust illumination normalization techniques with powerful feature extraction method.*

Keywords *- face recognition, illumination normalization, occlusion, Wavelet Analysis, Wavelet-based image de-noising, feature extraction, RLDA, small sample size (SSS) problem, Cosine distance.*

1. Introduction

Face recognition (FR) has been an active research topic in the fields of computer vision and biometrics over the past several decades [1-3] due to its numerous potential applications, such as human-computer interfaces, access control, security and surveillance, e-commerce, entertainment, and so on. However, many existing methods in face recognition area perform well under certain conditions, but still facing challenges with illumination changes and occlusions. The Illumination variation and occlusions are the most significant factors limiting the performance of face recognition. In Illumination variation, several images of the same person appear to be dramatically different under different illumination, the performance of most existing face recognition methods is highly sensitive to illumination variation, and will be severely degraded if the training/testing face is exposed to severe lighting variations. On the other hand, facial occlusion introduces more intra-class variations which makes the within-class scatter matrix is singular and thus the task of recognizing faces becomes more difficult. A face recognition system can confront occluded faces in real world applications very often due to use of accessories, such as scarf or sunglasses, hands on the face, the objects that persons carry, and external sources that partially occlude the camera view. Several studies have been conducted in order to address this problem [10-14]. To overcome the above challenges simultaneously, a combination of robust preprocessing and feature extraction is proposed. The preprocessing consists of two steps , in the first step of this preprocessing, the face images are enhanced with Histogram equalization, in the second step, the enhanced image is de-noised with Haar wavelet de-noising filter at level two of decomposition, and then robust feature extraction method called Regularized Linear Discriminant Analysis "RLDA" is used to overcome the singularity caused by occlusion. Finally, classify the extracted feature vectors by calculating the cosine distance between them.

2. Pre-Processing

This section describes the preprocessing method. The method combines image enhancement and wavelet de-noising normalization techniques. applying histogram equalization prior to the denoising model would result in improving the image's contrast and redistributing the pixel intensities equally which adds to images variability needed to successfully discriminate between different subjects. The wavelet denoising model to obtain an illumination invariant representation of the facial image

2.1 Histogram equalization

Histogram equalization (HE) [15] is one of the most useful contrast enhancement schemes. When an image's histogram is equalized, image pixel values are mapped to uniformly distributed pixel values, as much as possible. the idea of equalizing a histogram is to stretch and redistribute the original histogram using the entire range of discrete levels of the image, in a way that an enhancement of image contrast is achieved. The most commonly used histogram normalization technique is histogram equalization where one attempts to change the image histogram into a histogram that is constant for all brightness values. This would correspond to a brightness distribution where all values are equally probable. For image I (x, y) with discrete k gray values histogram is defined by (i.e. the probability of occurrence of the gray level *i* is given by):

$$p(i) = \frac{n_i}{N} \qquad (1)$$

Where $i \in 0,1...K-1$ grey level and N is total number of pixels in the image. Transformation to a new intensity value is defined by:

$$I' = \sum_{i=0}^{K-1} \frac{n_i}{N} = \sum_{i=0}^{K-1} p(i) \qquad (2)$$

The above expression defines the mapping of the pixels' intensity values from their original range, for example the 8-bit interval 0-255, to the domain of [0,1]. Thus, to obtain pixel intensity values in the original range, the result has to be rescaled. A visual example of the effect of histogram equalization on the appearance of facial images is shown in Fig 1.

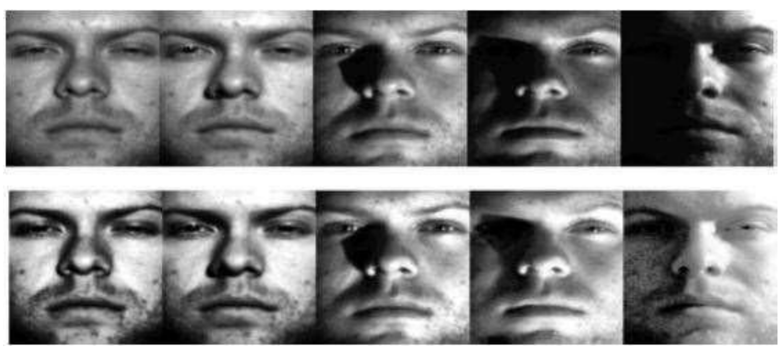

Fig 1. Impact of histogram equalization: original images (upper row), histogram equalized images (lower row)

As shown, histogram equalization improves the contrast of the facial images, but as seen on the last image of the lower row also greatly enhances the background noise. The demonstrated property of contrast enhancement makes histogram equalization is important stage in image processing.

2.2 Wavelet-based image denoising

The wavelet-based image denoising approach [16] is exploited to obtain an illumination invariant representation of the facial image. The technique starts with the imaging model of the retinex theory given by the following imaging model:

$$I(x, y) = R(x, y)L(x, y) \qquad (3)$$

Where $L(x, y)$ is the luminance and $R(x, y)$ is the reflectance, and can separated them by taking the logarithm of the image, which results in the following expression:

$$LogI(x, y) = LogR(x, y) + LogL(x,y) \qquad (4)$$

Under the assumption that the key facial features (R) are high frequency phenomena equivalent to "noise" in the image denoising model, we estimate the luminance $L(x, y)$ by the wavelet denoising model and then compute the illumination invariant reflectance $R(x,y)$ in the imaging model. Let us denote the wavelet coefficient of the input image $X(x,y) = DWT(LogI(x,y))$, where DWT stands for the 2D DWT operator; and, similarly, let $Y(x,y) = DWT(LogL(x,y))$ denote the matrix of wavelet coefficients of the luminance. The estimate of the luminance in the wavelet domain Y(x,y) is then obtained by modifying the detail coefficients of X(x,y) using the so-called soft thresholding technique and keeping the approximation coefficients unaltered as following:

$$Y_s(x,y) = \begin{cases} X_s(x,y) - T, & if \ X_s(x,y) \geq T \\ X_s(x,y) + T, & if \ X_s(x,y) \leq -T \\ 0, & if \ \left| X_s(x,y) < T \right| \end{cases} \tag{5}$$

where $X_s(x,y)$ denotes one of the three sub-bands $s \in \{LH, HL, HH\}$ generated by the detail DWT, $Y_s(x,y)$ stands for the corresponding soft thresholded sub-band and *T* represent a predefined threshold. Once, all three detail coefficient sub-bands have been thresholded, they are combined with the unaltered approximate coefficient sub-band to form the denoised wavelet coefficient matrix Y(x,y). The estimate of the luminance in the spatial domain is ultimately obtained by applying the inverse DWT to the wavelet coefficients in Y(x,y), and can be used to compute the illumination invariant reflectance.

3. Feature Extraction

Feature extraction is the task of reducing the high dimensional training data to a set of features to investigate properties (morphological, geometric etc.) of the data [4,5]. Features are used by recognition approaches to differentiate between faces of different identities. Feature extraction techniques should extract robust features of the input data which make distinct pattern classes separate from each other. RLDA [9] is a robust feature extraction that able to solve significant problems such as SSS and Occlusion that usually occurs in FR systems.

3.1 A regularized LDA: R-LDA

It is well-known that the applicability of linear discriminant analysis (LDA) to face recognition often suffers from the so-called ‘‘small sample size’’ (SSS) problem arising from the small number of available training samples compared to the dimensionality of the sample space. In addition, the recognition performance of LDA is often severely degraded by the presence of occlusion. Therefore, A regularized LDA method was introduced to address these issues. This method can be summarized as following:

Given a face image matrix A of size *mxn*, we construct a vector representation by concatenating all the columns of A to form a column vector *x* of dimension *mxn* . Given a set of training vectors x_i*; i= 1, 2, ..., M* for all persons.

$$X = [x_1, x_2, ..., x_M] \tag{6}$$

The matrix *X* is composed of *C* classes; in each class we have *mi* individuals.

$$\sum_{i=1}^{C} m_i = M \tag{7}$$

Then, we express the between-class scatter matrix S_b as:

$$S_b = \Phi_b^T . \Phi_b \tag{8}$$

With:

$$\Phi_b = [\Phi_{b,1}, ..., \Phi_{b,C}], \Phi_{b,i} = \sqrt{m_i}(\mu_i - \mu) \tag{9}$$

Where

$$\mu_i = \frac{1}{m_i} \sum_{j \in i} x_j, \ \mu = \frac{1}{M} \sum_{i=1}^{M} x_i = \frac{1}{M} \sum_c m_c \mu_c \tag{10}$$

Next, we determine the d eigenvectors of $\phi_b^T.\phi_b$ denoted as $E_d=[e1,...e_d]$, where $d<=C-1$. After that, we calculate the first d most significant eigenvectors U_d of S_b and their corresponding eigenvalues (Λ_b) by:

$$U_d = \Phi_b E_d, \Lambda_b = U_d^T S_b U_d \tag{11}$$

Let $H = U_d \Lambda_B^{-1/2}$. We find the eigenvectors of $H^T S_W H$ denoted as $P=[p_1,...p_d]$ sorted in increasing eigenvalue order. S_w is the within-class scatter matrix which is defined as:

$$S_w = \sum_{i=1}^{C}\sum_{j=1}^{m_i}(x_j-\mu_i)(x_j-\mu_i)^T \tag{12}$$

Then, we choose the first D eigenvectors of P (D≤d). Let P_D and Λ_w are the obtained eigenvectors matrix and their corresponding diagonal eigenvalues matrix, respectively. In RLDA the projection vector W are computed as follow:

$$W = HP_D(\eta I + \Lambda_W)^{-1/2} \tag{13}$$

I is the $(D\times D)$ identity matrix $0\le\eta\le1$ is a regularization parameter. In general, we choose $\eta=1$. Finally, we project the training vectors as follow:

$$Y = W^T X \tag{14}$$

Given a test face image Test, first we transform it on a column vector T, then use Eq. (14) to get the feature vector $T_P = W^T T$, then we calculate the similarity between vectors using Cosine distance.

4. Experimental Results

4.1 Database

Since this proposed face recognition system mainly deals with illumination problem, I used Extended Yale-B face database. The Extended Yale-B face database provides an excellent testing platform for extreme variation in illumination [6]. It consists of still images, in frontal pose, for 38 subjects, each having 64 images captured under different illumination conditions.

4.2 Experiment 1: Recognition without occlusion

In this experiment, the ideal images (subset1) were used as reference images to form the gallery set and all the remaining images were used for testing. Table.1, shows the results compared with another existing methods that handle illumination variation problem [7].

Illumination preprocessing methods

Feature extraction	HE [21]	SSR [22]	SQI [23]	LTV [24]	TT [25]	HE+ Wavelet Denoising
PCA [26]	31.4	51.6	58.2	20.3	50.2	-
LDA [27]	54.8	55.5	72.6	78.0	71.6	-
LBP [25]	62.2	58.7	57.9	52.8	90.3	-
LGBP [28]	95.9	98.4	99.1	92.1	98.9	-
RLDA	-	-	-	-	-	**99.4**

Table.1 Recognition results by different methods on the Extended Yale B database

Experimental results show that the very high recognition rate achieved by RLDA with the proposed preprocessing method and much higher than the methods listed in the comparative study.

4.3 Experiment 2: Recognition with random block occlusion

In this experiment, I have tested the robustness of the proposed method to random block occlusion. As in [8], Subsets 1 and 2 of the Extended Yale B database are used for training and Subset 3 for testing. Each testing sample will be inserted an unrelated image as block occlusion, and the blocking ratio is from 10% to 50% as illustrated in Fig.2.

10% 20% 30% 40% 50%

Fig.2. Occlusion ratio

Table.2, shows the results compared with another existing methods that handle occlusion problem.

Occlusion ratio	10%	20%	30%	40%	50%
SRC [17]	**100%**	99.8%	98.5%	90.3%	65.3%
CRC_RLS [20]	99.8%	93.6%	82.6%	70.0%	52.3%
GSRC [18]	**100%**	**100%**	**99.8%**	96.5%	87.4%
RSC [19]	**100%**	**100%**	**99.8%**	96.9%	83.9%
HE +Wavelet denoising/RLDA	**100%**	**100%**	99.3%	**98.5%**	**95.8%**

Table.2 Recognition results by different methods on the Extended Yale B database with various random occlusion ratio

As shown above, We can see that when the block occlusion ratio is low, all methods can achieve good recognition accuracy; when the block occlusion ratio increases, the accuracy of my proposed method can still has good results Up to more than 90% even with 50% block occlusion.

4.4 Experiment 3: Recognition with key facial Occlusion

In this experiment, the occlusions cover important facial parts like eyes, nose and mouth. The key facial parts are blackened out (likening them to occlusion). Fig 3. shows sample face images with occlusion. this kind of occlusion (key facial occlusion) results in loss of discriminative information, especially in the case of upper face occlusion the eye region, which is known to be very discriminative, this explains the low performance significantly for face recognition systems in the case of key facial occlusion.

Fig 3. Key Facial Parts Occulded

The tests have been done on more challenging subset namely, subset5 with 23.4% occlusion ratio of the key facial landmarks, whereas the subset1 was selected as training images. the results in table 3. shows the system works well under this kind of occlusion also.

Method	Recognition Rate%
HE+ Wavelet denoising/RLDA	98.5

Table.3 Recognition results by proposed methods on the Extended Yale B database with key facial landmarks occluded 23.4% of the image.

4.5 Experiment 4: Recognition with small sample size

In this experiment, the proposed method tested with small training set (one sample per person), and all the remaining images (2166 image) were used for testing, this makes the process of face recognition more challenging, Although many existing face recognition methods have achieved success in real application, but usually fail to work if only one training sample per person is available due to the lack of information to predict the variations in the test sample. However, the results in Table.4 show my proposed methods work well even small sample size scenario.

METHOD	Without Occlusion	With 50% Occlusion
HE+ wavelet Denoising/RLDA	95.4	88.4

Table.4 Recognition results Based on small sample size

5. Conclusion

The main purpose of any face recognition system is to retrieve face images which are very similar to a specific class of face images in a large database. The retrieved face images can be used for many applications, such as monitoring system in airport, visual surveillance, criminal face verification in police office, extracting specific faces from the web, and photo management. Although, many existing methods in face recognition area perform well under certain conditions, but still facing challenging with illumination changes and occlusions. The aim of this research was the development of method to deal with the above challenges and to support a face recognition application. Experiments were performed based on the Extended Yale B facial databases, The results show that very high recognition rates were achieved by RLDA and Cosine distance with the proposed preprocessing method; even if there are random occlusions in the testing images. The method exploited the wavelet denoising model to obtain an illumination invariant representation of the facial image in the wavelet domain. applying histogram equalization prior to the denoising model would result in improving the image's contrast and redistributing the pixel intensities equally. this procedure adds to images variability needed to successfully discriminate between different subjects. to get enhanced-de-noised image, histogram equalization is used for image enhancement then the de-noising are applied. I have demonstrated the effectiveness of the proposed method compared with other methods that handle illumination variation and occlusions. the proposed method gives the best results among them.

6. Future work

This research offers possibility of creating a real time system because it is not computationally complex. Therefore, design Online face recognition system is an important future step. Online face recognition is a real-time recognition task which aims to recognize incoming faces in video frames. Similar to face recognition based on static images, online face recognition can be implemented by using a machine learning-based approach. The goal of online face recognition is to identify the class membership for an unknown facial pattern given some known facial patterns in video frames. Another issue that we can take into account in further research is the recognition of unknown faces, which would allow that the images of the unknown face could be automatically stored in the training set for the purpose of subsequent recognition of the unknown face. Furthermore, the system will be tested on other databases larger and include facial expressions.

7. References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. ACM Computing Survey, 35(4): 399-458, 2003.
http://dx.doi.org/10.1145/954339.954342

[2] S. Z. Li and A. K. Jain. Handbook of Face Recognition (Second Edition). Springer, 2011.
http://dx.doi.org/10.1007/978-0-85729-932-1

[3] R. Jafri and H. R. Abrania. A survey of face recognition techniques. Journal of Information Processing Systems , 5(2): 41-68, 2009.
http://dx.doi.org/10.3745/JIPS.2009.5.2.041

[4] Amir Globerson, Naftali Tishby, Isabelle Guyon, and Andr Elisseeff . “Sufficient dimensionality reduction”. Journal of Machine Learning Research, 3:2003, 2003.

[5] W. K. Pratt. Digital Image Processing. John Wiley & Sons, Inc., 2001. http://dx.doi.org/10.1002/0471221325

[6] K. Lee, J. Ho, and D. Kriegman, “Acquiring linear subspaces for face recognition under variable lighting,” IEEE Trans. Pattern Analysis & Machine Intelligence, vol. 27, no. 5, pp. 684–698, 2005. http://dx.doi.org/10.1109/TPAMI.2005.92

[7] H.Han, S.Shan , X.Chen, W.Gao, “A comparative study on illumination preprocessing in face recognition”, Pattern Recognition, 46 (2013) 1691–1699. http://dx.doi.org/10.1016/j.patcog.2012.11.022

[8] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. IEEE Trans. Pattern Analysis and Machine Intelligence, 31(2): 210–227, 2009. http://dx.doi.org/10.1109/TPAMI.2008.79

[9] J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, “Regularization studies of linear discriminant analysis in small sample size scenarios with application to face recognition”, Pattern Recognition Letters, 26 (2005) 181–191. http://dx.doi.org/10.1016/j.patrec.2004.09.014

[10] Tan, X., Chen, S., Zhou, Z.H., Zhang, F.: Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble. IEEE Transactions on Neural Networks 16(4) (2005) 875-886 http://dx.doi.org/10.1109/TNN.2005.849817

[11] Park, B.G., Lee, K.M., Lee, S.U.: Face recognition using face-ARG matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(12) (2005) 1982-1988 http://dx.doi.org/10.1109/TPAMI.2005.243

[12] S. Fidler, D. Skocaj, and A. Leonardis, “Combining Reconstructive and Discriminative Subspace Methods for Robust Classification and Regression by Subsampling,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 3, pp. 337-350, Mar. 2006. http://dx.doi.org/10.1109/TPAMI.2006.46

[13] Wright, J., Ganesh, A., Yang, A., Ma, Y.: Robust face recognition via sparse representation. Technical report, University of Illinois, USA (2007)

[14] Jia, H., Martinez, A.M.: Face recognition with occlusions in the training and testing sets. In: Proceedings of IEEE Int'l. Conf. on Automatic Face and Gesture Recognition. (2008). http://dx.doi.org/10.1109/afgr.2008.4813410

[15] S. M. Pizer and E. P. Amburn, “Adaptive histogram equalization and its variations,” *Comput. Vis. Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987. http://dx.doi.org/10.1016/S0734-189X(87)80186-X

[16] Zhang, T., Fang, B., Yuan, Y., Tang, Y.Y., Shang, Z.,Li, D., Lang, F.: Multiscale Facial Structure Representation for Face Recognition Under Varying Illumination, Pattern Recognition, vol. 42, no. 2, pp. 252-258, 2009 (wavelet denoising). http://dx.doi.org/10.1016/j.patcog.2008.03.017

[17] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. IEEE Trans. Pattern Analysis and Machine Intelligence, 31(2): 210–227, 2009. http://dx.doi.org/10.1109/TPAMI.2008.79

[18] M. Yang and L. Zhang, “Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary,” *Proc. European Conf. Computer Vision*, 2010.

[19] M. Yang, L. Zhang, J. Yang, D. Zhang. Robust sparse coding for face recognition. In CVPR, 2011. http://dx.doi.org/10.1109/cvpr.2011.5995393

[20] L. Zhang, M. Yang, X. Feng. Sparse representation or collaborative representation: Which helps face recognition? In ICCV, 2011.

[21] Gonzalez R.C., Woods R.E. “Digital Image Processing. 2nd edition”, Prentice Hall, 2002.

[22] D.J. Jobson, Z. Rahman, and G.A. Woodell, “Properties and Performance of a Center/Surround Retinex,” IEEE Trans. Image Processing, vol. 6, no. 3, pp. 451-462, Mar. 1997.

[23] H. Wang, S.Z. Li, Y. Wang, Face recognition under varying lighting conditions using self quotient image, in: Proceedings of the Automatic Face and Gesture Recognition, 2004, pp. 819–824.

[24] T. Chen, X. S. Zhou, D. Comaniciu, and T. S. Huang, “Total variation models for variable lighting face recognition,” IEEE Trans. Pattern Anal. Mach. Intell., 28 (9):1519–1524, 2006.

[25] Tan, X., Triggs, B., “Enhanced local texture feature sets for face recognition under difficult lighting conditions,” IEEE Trans. Image Process. 19 (6):1635-1650, 2010.

[26] M. Turk, A. Pentland, Eigenfaces for recognition, Journal of Cognitive Neuroscience 3 (1991) 71–86.

[27] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (1997) 711–720.

[28] W. Zhang, S. Shan, W. Gao, X. Chen, H. Zhang, Local gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition, in: Proceedings of the International Conference on Computer Vision, 2005, pp. 786–791.