

Illumination Normalization in Face Recognition Using DCT and Supporting Vector Machine (SVM)

¹ Yun-Wen Wang (王詠文), ² Wen-Yu Wang (王文昱), ² Chiou-Shann Fuh (傅楸善)

¹ Graduate Institute of Electronics Engineering,
National Taiwan University, Taiwan

² Graduate Institute of Biomedical Electronics and Bioinformatics,
National Taiwan University, Taiwan
E-mail: w32017@gmail.com , wenyuwang826@gmail.com

ABSTRACT

Face recognition has become a popular identification technique to perform human identity verification. By using the feature extraction methods and dimensionality reduction techniques in the pattern recognition applications, a number of facial recognition systems have been produced with distinct measure of success. However, variation in the illumination of face images makes it hard to recognize faces correctly in real life. As a result, this paper provides a pre-processing method on each face images. Also, Principle Component Analysis (PCA) and the SVMs are used to extract feature vector exactly and to classify them into a group accurately.

Keywords Face Recognition, Illumination Normalization, DCT (Discrete Cosine Transformation), SVM (Support Vector Machine), PCA (Principle Components Analysis)

1. INTRODUCTION

Face recognition has received a great deal of attention from the scientific and industrial communities over the past several decades owing to its wide range of applications in information security and access control, law enforcement, and so on. Due to its relevance for many applications, face recognition is well-studied. However, making recognition more reliable under uncontrolled lighting conditions is one of the most important challenges for practical face recognition systems. In this study, we would focus on pre-processing for difficult lighting conditions. As a result, the preproccession of face recognition can be partitioned as three parts: face normaliztion, feature extration, and feature classification.

In this study, we used DCT, Discrete Cosine Transformation, for face normaliztion on illumination as image preproccession. We adjucted the coeddicients of DCT to eliminate the effect of difficult and different lighting conditions, and evaluate the quality of face images. After modifying images to well-lit conditions,

we performed PCA, Principal Component Analysis, for feature extraction and SVM, Support Vector Machine, for feature classificaiton respectively. In the rest part of this section, we would discuss these three parts in this section respectively.

1.1 Face Normalization

Face normalization is a crucial issue in face recognition. After detecting facial images, the images should be preprocessed for further analysis. In face recognition, we have a training set of images which contain feature (Eigen) vectors of different person's images as well as a number of images of the same person keeping different expressions on the face and images taken from different angles. To handle the different conditions, before feature extraction, it is important to normalize each image.

The methods proposed for face recognition mostly can only work under the well-lit condition. However, most cases in real life, images are not well-lot, t is crucial to reduce the lighting effect for the methods doing preprocess [13]. In this study, we would focus on the method of illumination normalization in face normalization.

1.2 Feature Extraction

We used Principal Component Analysis (PCA) for feature extraction. PCA projects the face image onto a feature space that spans the significant variations among known face images. Then we can get eigenvectors. Eigenvectors here are significant features, or Principal Components, of each face image, and we called them "Eigenfaces".PCA method is extensively used for face recognition because it not only reduces the dimensionality of the image, but also retains the variations in the image data.



Fig. 1: Example of Eigenfaces.

1.3. SVM Classification

SVM, Support Vector Machines, is a kind of kernel methods [6]. Kernel algorithms, like Gaussian kernels, map data from an original space into a higher dimensional feature space using some non-linear transformations, in which the optimal decision surface (hyperplane) is constructed. It seems that it would be more difficult for high-dimensional computation, however, computation of inner product (kernel function) in the feature space exists. Using kernel functions, the feature space needs not be computed explicitly.

SVM separates p -dimensional data using $p-1$ dimensional decision surface to maximize the margin of the data sets. The margin is defined as the minimal distance of a sample to the decision surface. The distance of the hyperplane from the nearest appearance of the individual data sets should be as large as possible. We would illustrate the idea via the Fig.2. The solid line is the hyperplane. Moreover, the dash lines parallel to the solid line contain the support vectors.

SVM can model complex, real-world problems such as text and image classification, hand-writing recognition, and bioinformatics and biosequence analysis. Moreover, SVM performs well on data sets that have many attributes, even if there are very few cases on which to train the model. There is no upper limit on the number of attributes; the only constraints are those imposed by hardware. Traditional neural nets do not perform well under these circumstances.

In this work, we utilized LIBSVM [3] developed by C. C. Chang et al. as our SVM library. According to the features of our databases, we used linear model to classify features. After feature extraction, we divided randomly our database (2432 images) into training set and testing set with ratio 1:4, i.e. we have 486 cases in training set and the rest are in testing set.

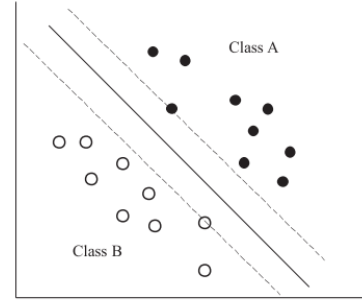


Fig. 2: An example of SVM.

The organization of the paper is as follows: The existing preprocessing methods are presented in section 2. Concise description of proposed preprocessing method in section 3. The experiments results and discussion is presented in section 4 and finally the paper is concluded in section 5.

2. RELATED WORK

Illuminant variation have drawn lots of attention for years. Many approaches are proposed to cope up with the illuminant problems. The existing method toward to illuminant variant can be roughly categorized as: transformation of images with variable illumination to a canonical representation, extracting illumination invariant features, modeling of illumination variation and utilization of some 3-D face models whose facial shapes and albedos are obtained in advance [9].

Histogram equalization [4] is the commonly used approach for photometric normalization. In this, the histogram of the pixel intensities in the resulting image is flat. In homomorphic filtering approach, the logarithm of the equation of the reflectance model is taken to separate the reflectance and luminance. Based on the assumption that the illumination varies slowly across different locations of the image and the local reflectance changes quickly across different locations, a high-pass filtering can be performed on the logarithm of the image to reduce the luminance part, which is the low frequency component of the image, and amplify the reflectance part, which corresponds to the high frequency component.

Short et al. [10] compared five photometric normalization and was proved that histogram equalization helped in every case. Chen et al. [2] employed DCT to compensate for illumination variation in the logarithm domain. The uneven illumination is removed in the image reconstructed by inverse dct after a number of DCT coefficients corresponding to low frequency are discarded.

Few researches emerged on the features of the face image that are invariant to variation in light. This has been achieved by extracting only those features that are not affected by variations in lighting conditions. Gradient faces [9], 2D Gabor Filter [4], DCT coefficients [10], LBP Feature [11] are some of the representations of image. Different methods of passive

approach using the Discrete Cosine Transform under the illumination condition are reviewed here.

Shermina et al. [7] proposed a methodology on illumination correction by adjusting the low-frequency components to lower the effect of luminance. Du et al [8] modified the low-frequency components and high-frequency components as well to make face recognition robust to illuminant variations.

3. METHOD

3.1 Preprocessing/Illumination Enhancement

This section describes our illumination enhancement method. The preprocessing runs before feature extractions and contains a series of stages shown in Fig. 3.

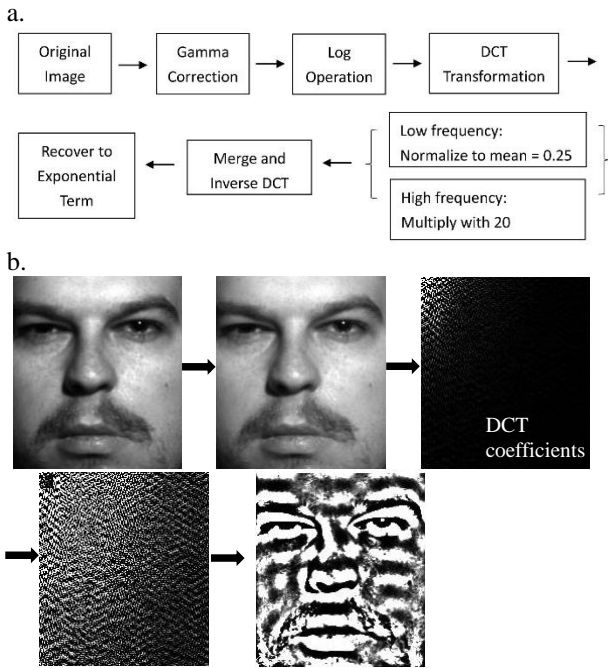


Fig.3: The block diagram of proposed illumination normalization method

In details, the stages are as follows.

3.1.1. Gamma Correction function

Gamma correction function is a nonlinear function which replaces gray-level with new intensity I^α , where α is a user-defined parameter. This step enhances the local dynamic range of the image in dark or shadowed regions while compressing it in bright regions. Moreover, different α are chosen according to different light conditions.

	Set1	Set2	Set3	Set4
α	0.8	1.12	1.12	1.15

3.1.2. Logarithm Transform

The input image (I_f) is the virtual frontal face image. Generally the illuminated face image $I_f(x, y)$ can be considered as the product of reflectance $R(x, y)$ and luminance $L(x, y)$ as shown in Eq. (1).

$$I_f(x, y) = R(x, y) L(x, y) \quad (1)$$

Taking logarithm transform on Eq. (1), we get,

$$\log I_f(x, y) = \log R(x, y) + \log L(x, y) \quad (2)$$

The linear equation obtained by taking logarithm transform on Eq. (1) shows that adding the logarithm transform of reflectance and the logarithm transform of luminance will yield the logarithm transform of the illuminated image. Logarithmic transform is frequently employed in image processing, to expand the values of the dark pixels. While the reflectance component lies mostly in the higher frequency band, the luminance component lies mostly in the low frequency band. Thus, the ‘low-pass version’ of the illuminated images can be taken as the approximate luminance component. The stable facial features are represented by the reflectance component of an illuminated face image under unstable lighting conditions.

3.1.3. Illumination Normalization using DCT

Here, we would use DCT (Discrete Cosine Transformation) for illumination normalization. The DCT is a useful and popular algorithm in image processing. DCT would transform signals in spatial domain into frequency domain. In an $M \times N$ block image, the total energy remains the same in the $M \times N$ blocks, but the energy distribution changes with most energy being compacted to the low-frequency coefficients. The Direct Current (DC) coefficient, which is located at the upper left corner, holds most of the image energy and represents the proportional average of the $M \times N$ blocks. The remaining $((M \times N) - 1)$ coefficients denote the intensity changes among the block images and are referred to as Alternating Current (AC) coefficients.

By adding a compensation term in low-frequency coefficients, which compensates for the non-uniform illumination, the uniform luminance component can be attained from the original image. The following procedure computes the compensation term. First, an $r \times c$ image is rebuilt from the low frequency components of DCT, and its mean m is computed.

Then, those low frequency components of DCT are used and the $\log L(x, y)$ is estimated. Single pixels that are “dark” are indicated by negative value and single pixels that are “bright” are indicated by positive value. We halve the difference between the pixel values and mean value in Eq. (3).

$$\text{Log}L'(x, y) = m + c(x, y) = m + 0.25 \times (m - \text{log}L(x, y)) \quad (3)$$

Also, when under poor illuminations, the high-frequency features become more important. As a result, in addition to compensating low-frequency part, we try to accentuate the remaining high-frequency part by multiplying a scalar 20. The face image after processing can be shown in Fig. 4

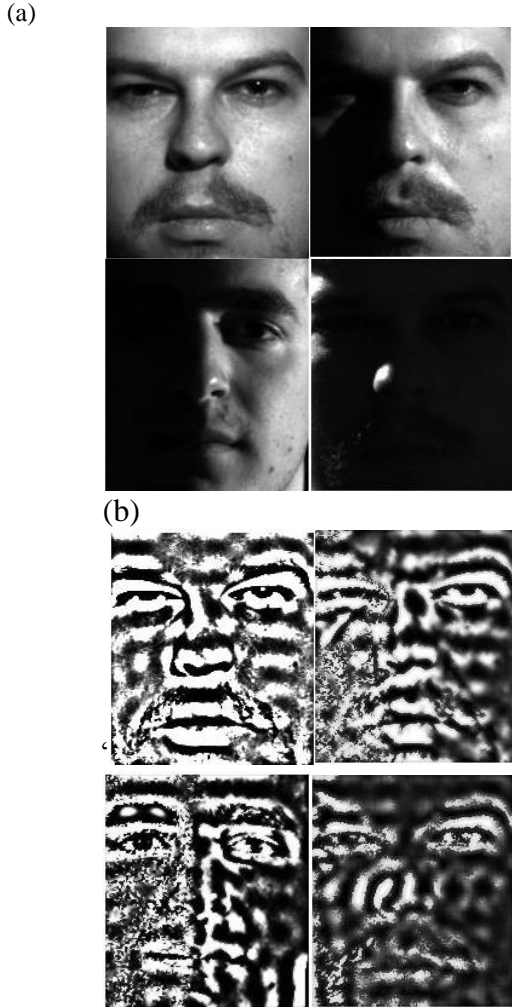


Fig.4: (a) The original images
(b) After Illumination Normalization

4. RESULTS AND DISCUSSION

4.1 Face Datasets

To evaluate the effectiveness of the proposed method, face images with large illumination variations are used.

We used Yale Face Database B Cropped as our database with different lighting conditions. There are 38 subjects under 64 different lighting conditions for one pose in our database (2432 images in total). In our work, we divided the original images of each subject into 4 subsets according to the mean intensity of each image before further image processing. By calculating the quartiles of 64 images of each subject, the images of each subject would be divided into 4 subsets according to 3 quartiles. In one subject, the quartiles of the mean intensity of each subject are 44.8, 78.7, and 112.2 respectively. Therefore, there are 16 images in a subset of an individual and 608 images in total in a subset. The example images of each subset are shown in Fig. 5.

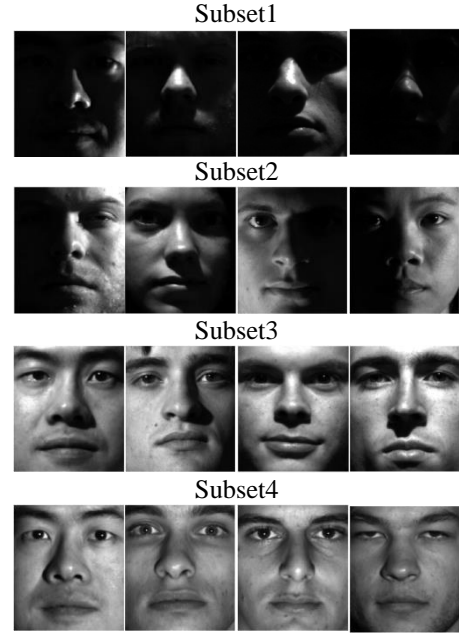


Fig. 5 The subset of images.

4.2 Result Comparisons

We compare the accuracy with our proposed method and other 6 existing methods.

In Table 1, Methods 1, 2, and 3 are DCT-based algorithm. Method 1 is the method proposed by Shan Du et al. [8]. In Method 1, they also modified the coefficients of DCT. They accentuated the high-frequency part by multiplying the coefficients by 50 and attenuated the low-frequency components by set the coefficients 0. Methods 2 and 3 are the approaches proposed by V. P. Vishwakarma et al. [12]. Here, we used the different C_{resc} they defined in the paper. The results below are consistent with their work, which mentioned that the higher C_{resc} , the higher accuracy. Other abbreviations are as follows.

Log: Using logarithm transformed image.

HE: Using Histogram-Equalized image.

RAW: Using the original image to classify faces.

Besides Table 1, we also present a histogram for the comparison in Fig. 6.

Table 1: The accuracy comparison (unit: %).

Test : Training Ratio =1:4

	Set1	Set2	Set3	Set4
Our Method	96.23	84.77	84.57	89.88
Method1	94.87	79.63	79.22	80.66
Method2 ($C_{resc}=56$)	92.37	73.1	78.4	76
Method3 ($C_{resc}=20$)	84.19	64.4	63.79	64.07
Log	96.05	64.91	54.39	56.14
Transformation	96.05	64.91	54.39	56.14
HE	96.05	64.91	54.39	56.14
RAW	87.04	54.41	37.86	22.59

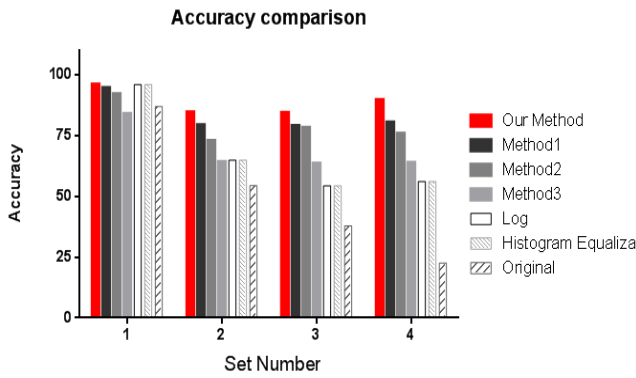


Fig. 6: The Histogram Accuracy Comparison (unit: %).

It is noteworthy to discuss the accuracy values in Set2 and Set3 are lower than Set4, since we divided the database by different intensity means. In Set2 and Set3, their intensities might be higher, but might be lit differently on the right and left sides of the faces. In Set1 and Set4, the light condition is more balanced than Set2 and Set3, therefore after illumination enhancement, the accuracy in Set4 is higher.

5. CONCLUSION

In this paper, a novel image pre-processing for face recognition is proposed. The method deals with the illumination variations by modifying the DCT coefficients. We eliminate the luminance effect by attenuating the values of low-frequency components and increasing the coefficients of remaining high-frequency components in logarithm domain. Moreover, we conduct the experiment on the Cropped Yale B face database. Finally, we showed our method outperforms other existing method. However, the present research shows that the effect of shadowing is not eliminated. In our future work, we will focus on the elimination/reduction of shadowing effect.

REFERENCES

- [1] S. Asteriadis, N. Nikolaidis, and I. Pitas, "Facial Feature Detection Using Distance Vector Fields," *Pattern Recognition*, Vol. 42, Issue 7, pp. 1388-1398, 2009.
- [2] W. Chen, M. J. Er, and S. Wu, "Illumination Compensation and Normalization For Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, Vol.36, no.2, pp.458,466, April 2006
- [3] C. C. Chang, and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol. 2, Issue 3, pp. 27:1-27:27, 2011.
- [4] D. H. Liu, K. M. Lam, and L. S. Shen, "Illumination Invariant Face Recognition," *Pattern Recognition*, Vol. 38, Issue 10, pp. 1705-1716, 2005
- [5] D. Matthias, G. Juergen, and F. Gabriele, "Real-Time Facial Feature Detection Using Conditional Regression Forests," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, pp. 2578-2585, 2012.
- [6] K. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An Introduction to Kernel-Based Learning Algorithms," *Neural Networks, IEEE Transactions on*, vol.12, no.2, pp.181,201, Mar 2001
- [7] J. Shermina, "Illumination Invariant Face Recognition Using Discrete Cosine Transform and Principal Component Analysis," *ICETECT*, pp. 826-830, 2011.
- [8] S. Du, M. Shehata, W. Badawy, and C. A. Rahman, "Eliminating illumination effects by discrete cosine transform (DCT) coefficients' attenuation and accentuation" *Proceedings of SPIE conference*, California, USA, Vol.8661, pp. 86610H-1- 86610H-11, 2013.
- [9] K. R. Singh, M. A. Zaveri, and M. M. Raghuwanshi, "Illumination and Pose Invariant Face Recognition: A Technical Review", *International Journal of Computer Information Systems and Industrial Management Applications*, Vol.2, pp.029-038, 2010
- [10] J. Short, J. Kittler, and K. Messer, "A Comparison of photometric Normalization Algorithm for Face Verification," *In Proc. Int'l conf. AFGR*, pp. 254-259, 2004.
- [11] X. Y. Tan and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions," *IEEE Transactions on Image Processing*, Vol. 19, No. 6, pp. 1635-1650, 2010.
- [12] V. P. Vishwakarma, S. Pandey, and M. N. Gupta, "A Novel Approach for Face Recognition Using DCT Coefficients Re-scaling for Illumination Normalization," *ADCOM*, pp. 535-539, 2007.
- [13] X. X. Zhu and D. Ramanan, "Face Detection, Pose Estimation, and Landmark Localization in the Wild," *Computer Vision and Pattern Recognition (CVPR)* Providence, Rhode Island, June 2012.