21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Illumination Suppression for Illumination Invariant Face Recognition

Aryaz Baradarani, Q.M. Jonathan Wu, Majid Ahmadi
Department of Electrical Engineering, University of Windsor, Ontario, Canada
$\{$*baradar, jwu, ahmadi*$\}$@*uwindsor.ca*

Abstract

This paper describes a multiresolution based method for face recognition under illumination variation. The idea of using the double-density dual-tree complex wavelet transform (DD-DTℂWT) for illumination invariant face recognition is motivated by the structure of the DD-DTℂWT; in addition to the shift-invariance and directionality, the transformation contains more number of wavelets in each level. Assuming that an input image can be considered as a combination of illumination and reflectance, we use a tunable logarithmic function to obtain a representative image. The image is then decomposed into several frequency subbands via DD-DTℂWT. Because the illumination mostly lies in the low-frequency part of the images, the high-frequency subbands are thresholded to construct a mask. Principal component analysis (PCA) and the extreme learning machine (ELM) are used for dimensionality reduction and classification, respectively. Experimental results are presented to illustrate the effectiveness of the proposed method.

1. Introduction

Recently, several interesting approaches have been shown to be effective to the problem of illumination invariant human face recognition. For example, Zhang et al. [12] proposed a discrete wavelet transform based algorithm to reduce the effect of redundant illumination. In [5], Nabatchian et al. have shown that the combination of a lowpass filtering and weighted voting scheme (WVS) can lead to better recognition rates.

It is evident that in the last few years multiresolution analysis has become one of the important and popular tools in pattern recognition and image processing thanks to the development of the discrete wavelet transform (DWT), dual-tree complex wavelet transform (DT-ℂWT), curvelets, double-density discrete wavelet transform (DD-DWT) and double-density dual-tree complex wavelet transform (DD-DTℂWT).

DD-DTℂWT is a recent enhancement over the shift-variance and poor directionality of the DWT in general. Basically, DD-DTℂWT is designed based on the use of two scaling functions and four wavelets at the same time and combines the desired properties of the DD-DWT and DT-ℂWT [6]. Selection of the DD-DTℂWT in this paper is related to the fact that illumination is assumed as the low-frequency component of the input image. As the transformation offers more high-frequency subbands, it is reasonable to accurately suppress high-frequency subbands to keep the low-frequency information. The latter, which can be interpreted as illumination, is to be subtracted from the original image to reduce unwanted illumination effects.

2. Illumination Suppression

2.1. Image Representation

Assuming that the surface orientation does not play a role and faces to be Lambertian surfaces, scene radiance is proportional to the product of the illumination and reflectance [3]. In other words

$$I(x,y) \propto R(x,y).L(x,y) \tag{1}$$

where $R(x,y)$ and $L(x,y)$ represent the reflectance and illumination at (x,y), respectively. A common assumption, based on the relation in (1), has been used in several papers as an approximate equality

$$\begin{aligned} I'_{\log}(x,y) &= \log(I(x,y)) = \log(R) + \log(L) \\ &= R'_{\log}(x,y) + L'_{\log}(x,y) \end{aligned} \tag{2}$$

where $I'_{\log}$, $R'_{\log}$ and $L'_{\log}$ denote the logarithm of the image, reflectance and illumination, respectively.

It should be noted that the logarithm of the luminance is a crude approximation, and therefore, it may partly reduce the illumination effects. In fact, it is difficult to separate R and L under the common assumption, as solving (2) is an ill-posed problem [12]. In order to enhance the recognition rates we have used the following expression that can expand the range of dark pixel

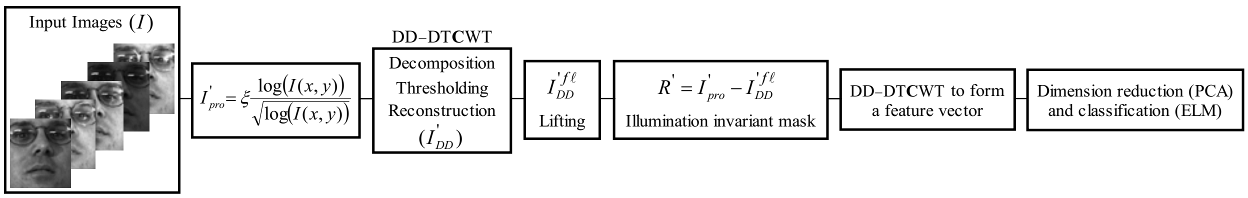


Figure 1. Block diagram of the proposed method.

values and, at the same time, can be controlled by a tuning parameter ξ, that is,

$$I'_{pro}(x,y) = \xi \frac{\log(I(x,y))}{\sqrt{\log(I(x,y))}} . \tag{3}$$

In the next section, it is shown that an efficient multiscale transformation, DD-DT$\mathbb{C}$WT, can facilitate the redundant illumination suppression task.

2.2. DD-DT$\mathbb{C}$WT Subband Filtering

As mentioned above, DD-DT$\mathbb{C}$WT consists of more wavelets than the regular DWT, that is, 32 wavelets in contrast to 3 wavelets in a traditional DWT. Furthermore, the wavelets associated with the DD-DT$\mathbb{C}$WT are directionally-selective and nearly shift-invariant; features can be localized in several directions [6]. Additionally, it fits the problem of extracting low-frequency information by suppressing more number of high-frequency subbands due to the existence of more wavelets.

The idea of using DD-DT$\mathbb{C}$WT is mainly based on preserving the illumination invariant information and at the same time suppressing the redundant data in view of the remarkable properties of the DD-DT$\mathbb{C}$WT. The block diagram of the proposed algorithm is shown in Fig. 1 and the approach can be summarized as follows. For a given image I, we apply (3) to obtain I'_{pro}. The new image I'_{pro} is decomposed into frequency subbands using the DD-DT$\mathbb{C}$WT. Note that unlike the traditional DWT with only three wavelets, DD-DT$\mathbb{C}$WT contains 32 real wavelets. On the other hand, the transformation contains 16 complex wavelets. The high-frequency subbands are thresholded where the value of threshold is fixed to the mean of the minimal coefficients of each row of a subband. The coefficients that are less than threshold remain unchanged and the rest of coefficients are set to zero. The low-frequency subbands and the thresholded high-frequency subbands are sent to the inverse DD-DT$\mathbb{C}$WT to reconstruct I'_{DD}. The obtained image is shifted down by ℓ rows, where ℓ is an integer and $1 \leq \ell \leq 3$, and then subtracted from I'_{pro} to obtain a mask. The lifting step can simply amplify the role of edges in R'_{pro} to improve the quality of the mask. In order to form a feature vector, R'_{pro} is decomposed into frequency subbands via DD-DT$\mathbb{C}$WT to construct the vector. We use PCA and ELM [4] for feature space dimension reduction and classification, respectively. In the next section the effectiveness and performance of the proposed method is shown via experimental results.

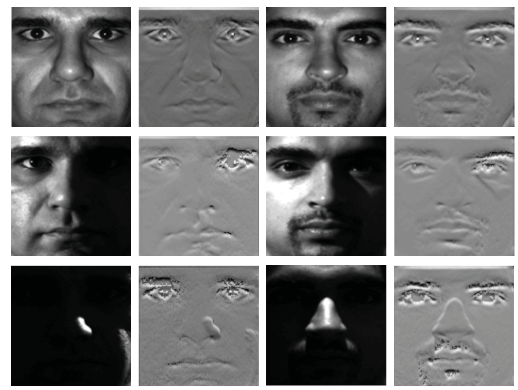
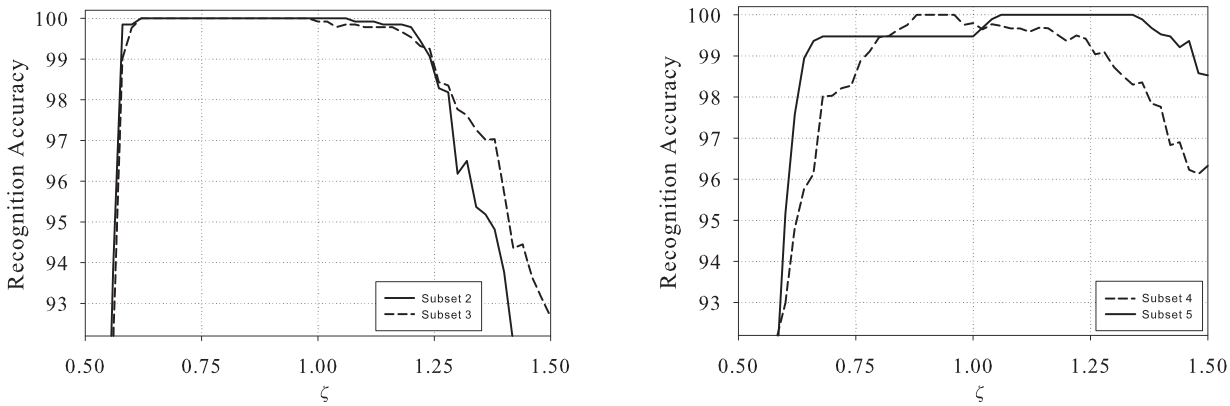


Figure 2. Sample images and the corresponding masks for the Yale B database.

3. Experimental Results

3.1. The Yale B Database

This database contains 5760 images with 9 poses and 64 illumination conditions for each of the 10 subjects in the database [11]. Depending on the angle between the light source and camera axis (α), the database is divided into five subsets. Subset 1 contains 70 images with $\alpha < 12°$. Subsets 2 – 5 contain 120 $\{20° < \alpha < 25°\}$, 120 $\{35° < \alpha < 50°\}$, 140 $\{60° < \alpha < 77°\}$, and 190 $\{\alpha > 78°\}$ images, respectively. All the images of size 192×168 are first eye-aligned, cropped and then resized into 128×128 pixels. In all experiments in this paper, with the Yale B and the CMU-PIE databases, a single-stage DD-DT$\mathbb{C}$WT decomposition is used and the value of ℓ is set to three. The proposed method is applied on images to obtain the masks required for the rest of procedure. Similar to previous approaches, Sub-



(a) Subsets 2 and 3

(b) Subsets 4 and 5

Figure 3. Recognition rate versus ξ for the Yale B database.

Table 1. Yale B: Recognition Rate for Different Techniques (%).

Method	Subset 2	Subset 3	Subset 4	Subset 5
QI [8]	99.30	61.90	34.10	23.30
QIR [8]	100	100	90.60	78.80
Hist. Equz. [7]	100	89.00	55.10	44.40
GIC [7]	100	88.10	39.90	27.50
SQI [10]	100	100	96.40	97.90
LTV+PCA [2]	100	99.17	96.43	92.12
MQI [13]	100	100	100	98.40
Wavelet+PCA [12]	100	100	100	100
II+PCA [5]	100	100	98.60	98.90
II+PCA+WVS [5]	100	100	100	99.47
Our Proposed	100	100	100	100

Table 2. CMU-PIE: Recognition Rate for Different Lighting Conditions (%).

Setting	train images	test images	Total	[5]	Proposed
off	3	18	21	94.85	99.26
on	4	20	24	100	100

set 1 is used for training and the remaining subsets are employed in testing phase. Fig. 2 shows two individuals from the Yale B database. In this figure, column one and three show the two subjects under different illumination conditions, respectively. The obtained masks for each individual are presented in columns two and four, respectively. The average of 50 separate execution of the simulations are given in Table 1 and Fig. 3, and it is seen that the proposed method can reach perfect recognition accuracy and performs reasonably well compared to the approaches in the literature. Recognition rate versus ξ has been shown in Fig. 3 for Subsets 2–5 for the Yale B database, respectively.

Figure 4. Sample images and the corresponding masks for the CMU-PIE database.

3.2. The CMU-PIE Database

The CMU-PIE database consists of 41,368 images from 68 individuals with 13 different poses and 43 illumination conditions [1]. Similar to [5], only frontal images have been used in our experiments. All images in this database are eye-aligned and resized into 96×96 pixels and the proposed algorithm is applied to the images to extract the corresponding masks. For the two lighting conditions, lights-off and lights-on [9], the images {i08, i11, i20} and {I06, I08, I11 I20} of each individual are used for training, respectively. The remaining images are considered for testing. Fig. 4 presents some randomly selected individuals from the lights-off and lights-on settings, respectively. Recognition rates versus ξ have been shown in Fig. 5 for the two lighting conditions. The result are given in Table 2 and it is seen that the proposed method can significantly improve the recognition accuracy compared to one of the recent approaches [5] in the literature.

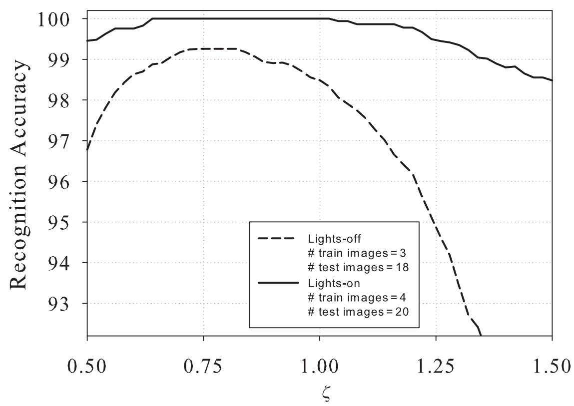


Figure 5. Recognition rate versus ξ for the CMU-PIE database.

3.3. Conclusions

This paper proposes an efficient multiresolution based method, for the problem of illumination invariant face recognition, that relies on illumination suppression via frequency subband filtering. It is assumed that an image can be represented as a combination of illumination and reflectance. Using a parameter to control the representation, illumination is initially enhanced. The method utilizes the double-density dual-tree complex wavelet transform that possesses three main advantages; directional-selectivity, shift-invariance, and extra number of wavelets. The transformation facilitates redundant illumination effects removal task by thresholding high-frequency subbands to preserve low-frequency information. For dimensionality reduction and classification, the principal component analysis (PCA) and the extreme learning machine (ELM) have been used, respectively. Several experiments haven been carried out to evaluate the effectiveness of the approach and the comparison suggests that the results obtained by employing the proposed method are competitive with the recent results in the literature. Further investigation may include sensitivity analysis and robustness in presence of noise, resolutions changes of facial image, and reduction in the number of training images.

Acknowledgment

The work is supported in part by the Canada Research Chair program and the Natural Sciences and Engineering Research Council of Canada.

References

[1] S. Baker, T. Sim, M. Bsat. The CMU pose, illumination, and expression database. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(12):1615–1618, 2003.

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

[3] B.K.P. Horn. *Robot Vision*. MIT Press, 1986.

[4] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew. Extreme learning machine: theory and applications, *Neurocomputing*, 70(1-3):489–501, 2006.

[5] A. Nabatchian, E. Abdel-Raheem, M. Ahmadi. Illumination invariant feature extraction and mutual information based local matching for face recognition under illumination variation and occlusion, *Pattern Recognition*, 44(10-11):2576–2587, 2011.

[6] I.W. Selesnick. The double-density dual-tree discete wavelet transform, *IEEE Trans. Signal Processing*, 52(5):1304–1314, 2004.

[7] S. Shan, W. Gao, B. Cao, D. Zhao. Illumination normalization for robust face recognition against varying lighting conditions, *IEEE Intl. Workshop Anal. Modeling Faces Gest.*, 157–164, 2003.

[8] A. Shashua, T.R. Raviv. The quotient image: class-based re-rendering and recognition with varying illuminations, *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(2):129–139, 2001.

[9] J.R.D. Solar, J. Quinteros. Illumination compensation and normalization in eigenspace-based face recognition: a comparative study of different pre-processing approaches, *Pattern Recognition Letters*, 29(14):1966–1979, 2008.

[10] H. Wang, S.Z. Li, Y. Wang. Face recognition under varying lighting conditions using self quotient image, *IEEE Intl. Conf. Auto. Face Gest. Recognition*, 819–824, 2004.

[11] A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(6):643–660, 2001. The Yale B Face Database, Available [online] at: `cvc.yale.edu/projects/yalefacesB/yalefacesB.html`.

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

[13] Y. Zhang, J. Tian, X. He, X. Yang. MQI based face recognition under uneven illumination, *Springer Advances in Biometrics*, 4642:290–298, 2007.