2010 International Conference on Pattern Recognition

Raw vs. Processed: how to Use the Raw and Processed Images for Robust Face Recognition under Varying Illumination

Li Xu
Graduate University,
Chinese Academy of Sciences,
Beijing,China
e-mail:xuli08@mails.gucas.ac.cn

Lei Huang, Changping Liu
Character Recognition Engineering Center,
Institute of Automation, Chinese Academy of Sciences,
Beijing,China
e-mail: {lei.huang,changping.liu}@ia.ac.cn

***Abstract*—Many previous image processing methods discard low-frequency components of images to extract illumination invariant for face recognition. However, this method may cause distortion of processed images and perform poorly under normal lighting. In this paper, a new method is proposed to deal with illumination problem in face recognition. Firstly, we define a score to denote a relative difference of the first and second largest similarities between the query input and the individuals in the gallery classes. Then, according to the score, we choose the appropriate images, raw or processed images, to involve the recognition.The experiment in ORL,CMU-PIE and Extended Yale B face databases shows that our adaptive method give more robust result after combination and perform better than the traditional fusion operators, the sum and the maximum of similarities.**

Keywords-face recognition; illumination; processed image

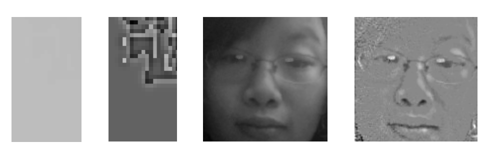
I. INTRODUCTION

The effect of variation in the illumination conditions, which causes dramatic changes in the face appearance, is one of the challenging problems in face recognition. Many well-known algorithms have been developed to tackle this problem. Generally, these approaches can be divided into two groups: modeling and image processing transformation/filtering. The modeling methods, including illumination variation modeling and 3D morphable model, require either lots of training samples or assumptions of light source, which are not applicable to practical application. Thus, we are interested in the second one.

It is well known that the illumination is on the low-frequency components of images. Most methods based on image processing extract illumination invariant by reducing low-frequency components of images. For example, the Multiscale Retinex (MSR) method of Jobson et al. [1] normalized the illumination by dividing the image by a smoothed version of itself. Related methods also include self quotient image (SQI) [2], logarithmic total variation (LTV) [3], and anisotropic smoothing [4] (GB, proposed by Gross &Brajovic). In addition, processing in frequency is also used to extract illumination-invariant. Homomorphic filtering[10] separates slow and fast changes by applying high-pass filter on the flourier spectrum of logarithm of the image. LOG-DCT [11] normalizes illumination by discarding low-frequency DCT coefficients in logarithm domain

Although some of these methods improve the recognition rate in the databases which include extreme illumination changes, they may perform worse when lighting variation is small. We can decompose face difference into three components: intrinsic, transformation difference and noise [7]. When the imaging conditions between training and query data are very different, image filtering can dominantly decrease the energy of transformation difference (illumination) and improve the performance. On the other hand, when the imaging conditions between training and query data are similar, the corresponding energy of transformation difference itself is very low, but processed images will lose facial details in the modified components, and hence decrease intrinsic difference. In this case, image processing will decrease signal-to-noise ratio, and worse the performance (see [6] for a thorough discussion).Some examples are presented in Fig.1. Small changes of the pixels in the top-right of Fig.1 (a) cause sharp changes in Fig.1(b), and the girl's right earring should be thought as noise of the image Fig.1(c) is clearly shown after processing, see Fig.1(d).

In this paper, we add raw grayscale and processed images to the gallery and query sets, and then define a score to determine whether the raw or the processed images are used to recognize.



(a) (b) (c) (d)

Figure 1. LTV filter:(a)(c) raw images; (b)(d) the results of LTV preprocessing on images

II. ALGORITHM

The above analysis make us propose a new way to solve different illumination cases: if the change of illumination condition is small, no normalization is actually needed and hence we only compares the raw images in the gallery and query sets; when the lighting

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.660

variation is obvious, the processed images are compared.

A. Relationship to previous works

To show that our method still works by using different processed images, one raw grayscale image and one processed image using four image processing filters respectively (1) LTV (2) homomorphic (3) GB (4) wavelet-based filter are used to evaluate our method. The four processing methods were designed to weaken low-frequency illumination fields for extracting illumination invariant, and can perform well in the extremely bad illumination condition.

Details of the four methods are as follows:

(1) LTV filter

In the Total-Variation based logarithm of the image (LTV), the illumination invariance v is obtained by $v = \log(I) - u$, where u is the illumination which can be solved using TV-L1 model:

$$u = \arg\min_u \int_\Omega |\nabla u| + \lambda |\log(I) - u| dx$$

(2) Homomorphic filter

In this approach, a high-pass filter is performed to reduce the illuminance part in logarithm of image.

(3) GB filter

The illumination L is estimated by minimizing an anisotropic function over the image region:

$$J(L) = \iint_\Omega \rho(x,y)(L-I)^2 dxdy + \lambda \iint_\Omega (L_x^2 + L_y^2) dxdy,$$

where $\rho(x,y)$ is space varying permeability weight which controls the anisotropic nature of the smoothing. L_x and L_y are the special derivatives of L. So the difference of image I and illumination L is the illumination invariant.

(4) wavelet-based filter

This method uses wavelet denoising model [8] to decompose face image into two parts, where the "noise" part in denoising model is corresponding to illumination invariant.

B. Raw grayscale vs. processed image

In order to choose more appropriate images, raw or processed image, for recognition under some illumination condition, we let the raw images compete with the processed ones.

Let y_0 be the query input corresponding to one of the

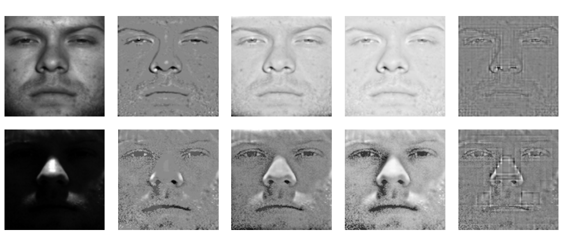


Figure 2. One raw image and four processed images by using (1)LTV (2) Homomorphic (3) GB (4) wavelet-based filter (from left to right)

gallery classes, and $\rho(y_0, X^k)$ a similarity function between the query input and the individual which is the kth best match for y_0. For example, $\rho(y_0, X^1)$ and $\rho(y_0, X^2)$ are the similarities of the query data y_0 and the two gallery individuals most similar to it.

We define a score θ as a relative difference between $\rho(y_0, X^1)$ and $\rho(y_0, X^2)$ to determine the winner, the raw grayscale or processed images.

$$\theta = \frac{1-\rho(y_0, X^1)}{1-\rho(y_0, X^2)} , \quad (1)$$

$$\theta_F = \frac{1-\rho(F(y_0), F(X^1))}{1-\rho(F(y_0), F(X^2))} , \quad (2)$$

where $F(\bullet)$ is one of four illumination-invariant filters which we have already introduced before. In this study, we apply the correlation coefficient between the features of two images for their similarities normalized between 0 and 1. And the value of $\rho(y_0, X^i)$ is set as the max value of similarities between y_0 and all the training images of the individual X^i.

Next, let us consider two extreme cases, when query data and the corresponding gallery data are acquired in (1) the same illumination conditions, and (2)extremely different illumination conditions. Because the similarities in formula 2 are obtained by comparing two processed images which have little relation with varying illumination, we suppose that θ_F have little change in varying lighting.

(1) The same illumination conditions

In this case, the similarity of the query data and the corresponding gallery data is nearly perfect (close to 1.0) and we can write $1-\rho(y_0, X^1) \ll 1-\rho(y_0, X^2)$. Thus, the result of formula 1 will be very small (close to 0), and $\theta < \theta_F$, see Fig.3 (a).

(2) Extremely different illumination conditions

Intrapersonal appearance changes due to illuminetion will be greater than interpersonal variations [9]. Thus, the values of $\rho(y_0, X^1)$ and $\rho(y_0, X^2)$ are very close to each other. Consequently the result of formula 1 will be very large (close to 1.0), we then write $\theta_F < \theta$, see Fig. 3(b).

Thus, the images, the raw or the processed, corresponding to a smaller score θ are used for face recognition. Then, we classify y_0 by assigning it to the class that maximizes the corresponding similarities between the query and training data.

III. EXPERIMENT RESULT

In the experiments, the face images are roughly aligned between different subjects, resized to 64 x 64 .In additions, LDA-based subspace method is used to extract features of face images.

A. Experiments on ORL database

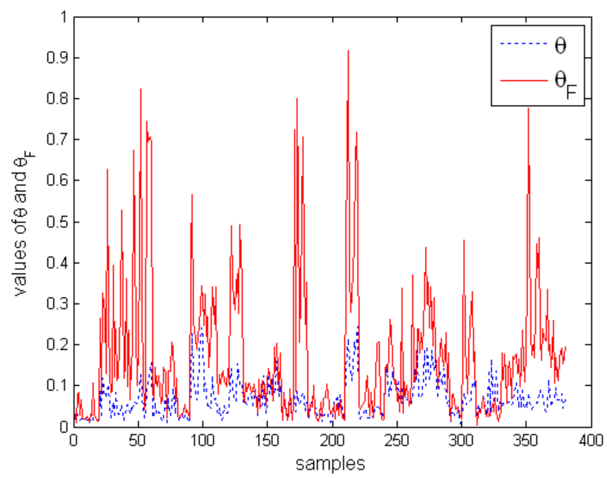
The ORL database contains ten different images of each of 40 distinct subjects, little variation in illumination. For each one in the database we collected five images of the person as training samples and the others as test images.

The blue line in Fig.4(a) shows that the processed images decrease the performance under normal lighting. At the same time, our method maintains the recognition rate.

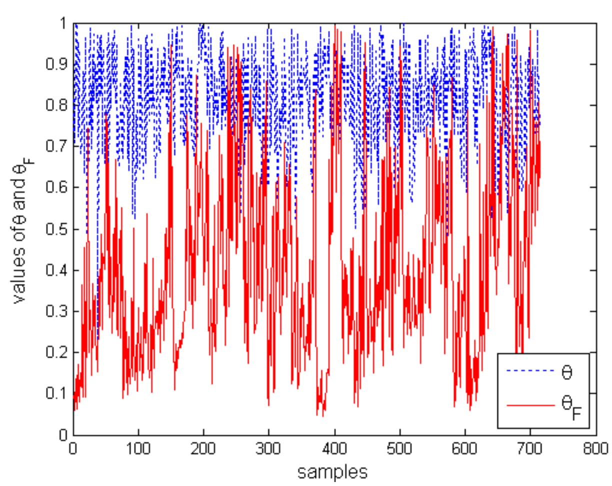
B. Experiments on CMU PIE face database

The CMU PIE face database captured under 13 synchronized cameras, 43 illumination and 4 expessions. In our work, illumination variations are mainly concerned, so the illumination subset (C27) which includes “illum” and “light” subsets is chosen for testing. In this experiment, we choose three images that are little affected by light per subject in “light”subset (f7-f9) for training, and all the images in “illum” subset for testing.

The proposed algorithm has improved the recognition



(a) Subset 2 in Extend Yale-B



(b) Subset 5 in Extend Yale-B

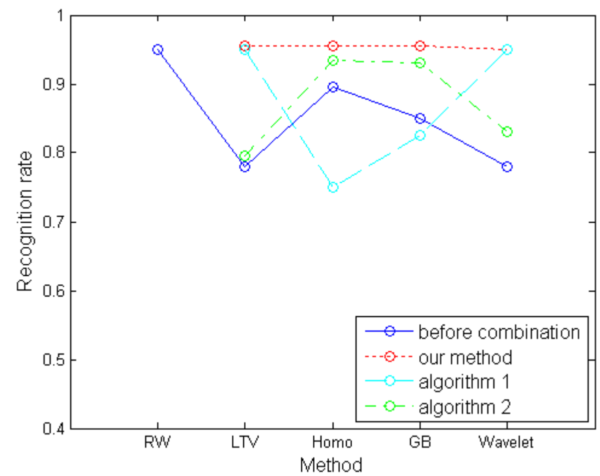
Figure 3. The plots of values of θ and θ_F .The values of θ_F obtained by using wavelet-based filtered data.

the performance on all method combinations (see the red and blue lines in Fig.4 (b)).

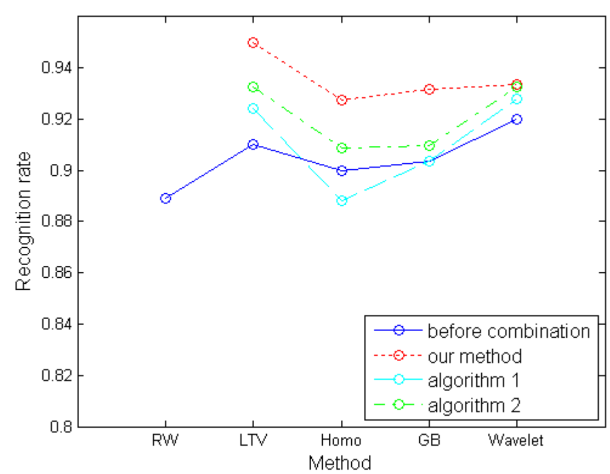
C. Experiments on Extend Yale B

The Extended Yale face database B contains images of 38 individuals in 9 poses and 64 illuminations per pose. We use 64 frontal face images under illumination conditions per person for evaluation. Image are divided into five subsets based on the angle of the light source directions. The five subsets are subset 1(0-12), subset 2(13-25), subset 3(26-50), subset 4(51-77), subset 5(above 78). In this experiment, we use images of subset 1 taken under small illumination conditions as training samples, and the other images from subsets 2 to 5 are used as testing images respectively. The results are given in Table 1. As can be seen, the four processing methods can dramatically reduce the error rate, and the result of our method (R- prefix) has a trend towards better one. Also, it has a reduction in error rates comparing the processing methods in subset 2-4.

Although most of images in subset 5 are almost black



(a) ORL



(b) CMU-PIE

Figure 4.Recognition Rate using different combination methods, and RW means only using raw grayscale images. Algorithm 1 and 2 are introduced in section III-D

TABLE 1. MEAN ERROR (%) ON YALE-B

	RW	LTV	R-LTV	Homo	R-Hom	GB	R-GB	Wavelet	R-Wl
Sub2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sub3	3.62	2.28	1.31	15.02	3.04	3.43	3.04	1.14	1.14
Sub4	47.64	3.51	3.29	5.71	5.49	10.99	9.90	6.81	6.37
Sub5	64.82	5.47	7.55	7.42	8.26	7.85	11.11	6.85	9.78

the clear facial lines (see the second row of Fig.2) have been extracted after processing which made similarities not too small (larger than 0.7). But the similarities are very close to each other due to noises produced by processing. As a result, θ_F may be larger than θ. The larger θ_F account for choosing the raw input images and then introduce error in recognition(see the five line of Table 1). However, we do not use an almost black image for recognition in practice.

D. *Comparing with traditional fusion operator*

We compare our method with two traditional combination methods as follows:

Algorithm1:

$$y_0 \in \arg\max_i (\max(\rho(y_0, X_i), \rho(F(y_0), F(X_i))))$$

Algorithm2:

$$y_o \in \arg\max_i (\rho(y_0, X_i) + \rho(F(y_0), F(X_i)))$$

As seen in Fig 4, these two methods (the cyan and green lines) produce an unstable performance, especially in the ORL database. However, whatever processing method is used or condition of illumination is, our framework is shown more robust for face recognition.

IV. CONCLUSIONS

In this paper, we have presented a novel method for face recognition in varying lighting. We use an adaptive way to exploit the raw grayscale input and the processed imagery for better performance in recognition. Evaluated on three different illumination conditions (ORL, CMU PIE and Extended Yale-B), the proposed framework provides vary promising performance, whether lighting variation is obvious or not.

References

[1] D. Jobson, Z. Rahman. "A multiscale retinex for bridging the gap between color images and the human observation of scenes," IEEE TIP, 6(7):965–976, 1997.

[2] H. Wang, S. Li, and Y. Wang. "Face recognition under varying lighting conditions using self quotient image," In Proc. AFGR'04, 2004.

[3] T. Chan, W. Yin, X. Zhou, D. Comaniciu, and T. Huang. "Total variation models for variable lighting face recognition," IEEE TPAMI, 28(9):1519–1524, 2006.

[4] R. Gross and V. Brajovic. "An image preprocessing algorithm for illumination invariant face recognition," In Proc. AVBPA: 10–18, 2003.

[5] T. Zhang, B.Fang. "Multiscale facial structure represention for face recognition under varying illumination," PR,42: 251-258, 2009.

[6] O.Arandjelovic, R.Cipolla."A methodology for rapid illmination-invariant face recognition using processing filters," Computer Vision and Image Understanding: 159-171, 2009. [7] X. Wang, X. Tang. "Unified subspace analysis for face recognition," In: Proc. IEEE International Conference on Computer Vision, 1:679-686, 2003.

[8] D.L.Donoho, "De-noising by soft-thresholding," IEEE, Trans .Inf. Theory, 41(3):613-627, 1995.

[9] Y. Adini, Y. Moses, S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," IEEE TPAMI, 19 (7) (1997) 721–732.

[10]Rafaek C.Gonzalez, Digital Image Processing Using,Pearson Prentice Hall,inc:2005

[11]W. L. Chen, E. M. Joo and S. Wu. Illumination Compensation and Normalization for Robust Face Recognition using Discrete Cosine Transform in Logarithm domain. IEEE Transactions on Systems, Man and Cybernetics, Part B, 36(2):458～466, 2006.