Face Recognition Across Illumination

Saleh Aly
Department of Intelligent Systems
Kyushu University
744, Motooka, Nishi-ku, Fukuoka, 819-0395
aly@limu.is.kyushu-u.ac.jp

Alaa Sagheer
Department of Intelligent Systems
Kyushu University
744, Motooka, Nishi-ku, Fukuoka, 819-0395
alaa@limu.is.kyushu-u.ac.jp

Naoyuki Tsuruta
Department of Electronics Engineering
and Computer Science, Fukuoka University
8-19-1, Nanakuma, Jonan, Fukuoka 814-0180
tsuruta@tl.fukuoka-u.ac.jp.co.jp

Rin-Ichiro Taniguchi
Department of Intelligent Systems
Kyushu University
744, Motooka, Nishi-ku, Fukuoka, 819-0395
rin@limu.is.kyushu-u.ac.jp

Abstract

Illumination variation in face images is one of the most difficult problems in face recognition system. The performance of self-organizing map-based face recognition system is highly degraded when the illumination of test images differ from training images. Illumination normalization is a way to solve this problem. Both global and local image enhancement methods are studied in this paper. Local histogram equalization method is highly improving the recognition accuracy of the CMU-PIE face database.

Keywords—Face recognition, illumination, SOM

1 Introduction

Important practical applications of automatic face recognition have made it a very popular research area in the last three decades [1]. Human-machine interface is considered as one of these applications. For example, the interaction between robots and human requires a fast and accurate identification of the person identity. In spite of the expanding research in the field of face recognition, many problems are still unsolved.

Recently, more researchers focus on robust face recognition which is invariant to pose, expression, illumination variations. Illumination variation is still a challenging problem in face recognition research area especially for appearance-based methods [2]. The same person can appear much differently under varying lighting conditions. Varieties of approaches have been proposed to solve the problem, and they can be classified into three categories: preprocessing and normalization [3, 4], invariant feature extraction [5], and face modeling [6].

Self-organizing map (SOM) [7] is a famous unsupervised neural network. In this work, SOM is used for feature extraction and dimensionality reduction. Two different shapes of SOM maps including sheet, and cylinder shapes are used to learn face manifold. However, the performance of SOM is highly degraded when the lighting condition of the test images differs from that of training images. Therefore, it is proposed to improve its accuracy using illumination normalization techniques.

To improve the performance of SOM, several illumination normalization techniques are investigated. According to the processing strategy, these methods can be categorized into two classes. Global methods in which the whole image is processed at once. On the other hand, region-based methods process small regions from the image separately. In this work, two well-known image enhancement methods are used. Gamma intensity correction (GIC) which normalize the overall image intensity to a given intensity level and histogram equalization (HE) which reassigns the intensity values of pixels in the input image such that the output image contains a uniform distribution of intensities. Both unnormailzed and normalized face images are projected on the 2-dimensional space spanned by the two greatest eigenvectors using principal component analysis (PCA); the structure of the projected manifold under different lighting directions is analyzed, moreover, the separability among different classes is explored.

The remainder of the paper is organized as follows: In section 2, the learning algorithm of self-organizing map neural network is discussed. Section 3 mainly describes global and local illumination normalization methods. Face database, experimental results, and conclusions are given in the following three sections.

2 Self-organizing Map

Consider the set of training images $\chi = \{\boldsymbol{x}_i, 1 < i < M\}$, each image belong to N-dimensional space. SOM is usually represented as a neural network sheet or map whose units, usually called nodes or neurons, become tuned to different input vectors $\boldsymbol{x}_i$. A weight vector $\boldsymbol{w}_j$, sometimes called reference or codebook, is associated with each neuron j and the map weight vectors are given by $\mathcal{W} = \{\boldsymbol{w}_j, 1 < j < N\}$; such that $N < M$. In each training step, the following two steps are repeated for each input sample $\boldsymbol{x}_i$.

1. Find the best matching neuron $\boldsymbol{c}$ using a similarity measure between the input and all

the map's neurons. This step name is *winner-take-all* (*WTA*), where $\boldsymbol{c}$ is the desired *winner* and should satisfy:

$$\|\boldsymbol{x_i} - \boldsymbol{w}_c\| = \min_j (\|\boldsymbol{x_i} - \boldsymbol{w}_j\|) \tag{1}$$

2. Update the weight vector of the winner **c** and also all its topological neighborhood in the map towards the prevailing input according to the rule:

$$\boldsymbol{w}_j(t+1) = \boldsymbol{w}_j(t) + h_{cj}(t)[\boldsymbol{x_i}(t) - \boldsymbol{w}_j(t)] \tag{2}$$

$$h_{cj}(t) = \alpha(t).\exp\left(\|r_c - r_j\| \big/ 2\sigma^2(t)\right) \tag{3}$$

where $h_{cj}(t)$ is the neighborhood kernel function around the *winner* $\boldsymbol{c}$ at time t, $\alpha(t)$ is the learning rate and is decreased gradually toward zero and $\sigma^2(t)$ is a factor used to control the width of the neighborhood kernel.

The SOM codebook has the following characteristics:

1. The probability Distribution Function (PDF) of the codebook is a good approximation for the PDF of the training data.
2. The topographic order of the training data is preserved in the codebook, even if the dimensionality of the SOM is smaller than that of training data.

The second characteristic means that similar facial features are mapped to nearby positions in the feature map. This ordering takes place automatically without external supervision based on only the internal relations in the structure of the input patterns and the coordination of the neuron activities through the lateral connections among the neurons.

3 Normalization Methods

Illumination variation of face patterns is extremely complex due to varying texture reflectance properties, face shape, and type and distance of lighting sources. Hence, in such a general setup, it is difficult to learn. However, most of the variations can be described by dominant principal components, this motivates the illumination subspace analysis using PCA.

The original and the normalized image faces are projected on the 2-dimensional space spanned by the two greatest eigenvectors using principal component analysis (PCA). Figure 1 shows the distribution of Yale B database face images [6] for two subjects under 64 lighting conditions. It is clear that the manifold of these subjects is highly overlapped. Thus, the discrimination between different subjects will be a very hard task.

3.1 Histogram Equalization (HE)

Histogram equalization is most widely used method

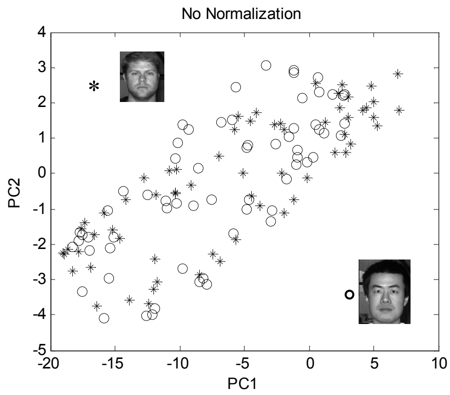

Figure 1: Unnormailzed face images distributions under 64 illumination conditions along the first two principle components.

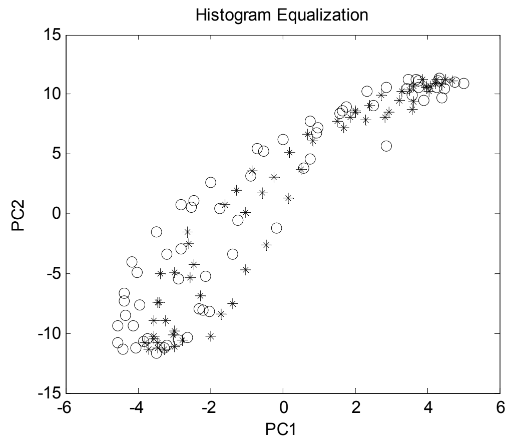

Figure 2: Normalized Face images distributions using histogram equalization method.

to enhance biased contrast image that some pixels are concentrated on a narrow range of the pixel intensity. The result is obtained using cumulative density function of the image as a transfer function. The result of this process is that the histogram becomes approximately constant for all gray values. As shown in Figure (2), the face distribution after applying HE algorithm is highly regularized.

3.2 Gamma Intensity Correction (GIC)

Gamma correction is a technique commonly used in the field of computer graphics. It concerns how to display an image accurately on a computer screen. Images that are not properly corrected can look dark. Gamma correction can control the overall brightness of an image by changing the Gamma parameter. The gamma transform of an image is a pixel transform in which the output and input are related by exponentiation

$$f\,(I\,(x\,,y\,)) = I\,(x\,,y\,)^{\gamma} \tag{4}$$

Depending on the value of γ the output image is darker or brighter. In GIC, the image is gamma transformed as to best match a canonically illuminated image $I_C(x, y)$. Predefine a canonical face image, I_C, which should be lighted under some normal lighting condition To find the value of γ the following equation must be solved

$$\gamma = \arg\min_{\gamma^*} \sum_{x,y} [I(x,y)^{\gamma^*} - I_C(x,y)]^2 \qquad (5)$$

This is a nonlinear optimization problem, and it can be solved using Golden section search algorithm.

As shown in Figure (3), face distribution using GIC normalization algorithm still overlapped and face data still have bad separation.

3.3 Region-based HE and GIC

It is obvious that both HE and GIC are global transforms over the whole image area. Therefore, they fail when side lighting exists. To solve this problem, region-based method is proposed to process the face images based on different local regions. That is, performing HE or GIC in some predefined face regions in order to alleviate the highlight, shading and shadow effect caused by the unequal illumination. As we know, face image is symmetrical around the vertical axis and contains different parts around the horizontal axis; therefore, it is proposed to partition it into four regions. As shown in figures (4, 5), region based normalization methods increase the separation between different subjects, thus the recognition accuracy will be improved.

4 Face Database

CMU-PIE face database [8] is available for studying pose, illumination, and expression problems in face recognition. There are 68 individuals under 43 different lighting and 3 different facial expressions for 13 poses. Since this work mainly deals with the illumination problem, frontal images under varying lighting conditions are selected, these which includes the images under 21 different directional flashes. Example images of one person in frontal pose are shown in Figure (6). The images are divided into four subsets according to the angle that the light source direction makes with the camera axis— Subset 1 (f06~f09, f11, f12, f20), Subset 2 (f05, f10, f13, f14, f19, f21), Subset 3 (f04, f15, f18, f22), and Subset 4 (f02, f03, f16, f17).

5 Experimental Results

In this experiment, subset 1 is used as the training set (gallery images) and other subsets are used for testing (probe images), all images are rescaled to the size of 48x48 in order to accelerate the computation time. All cropped images in the database are photometrically normalized using histogram equalization, gamma

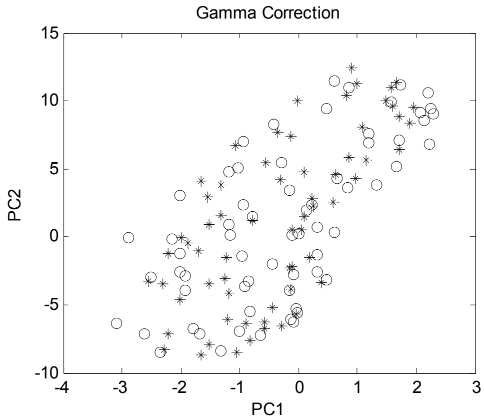

Figure 3: Normalized Face images distributions using Gamma Intensity Correction method.

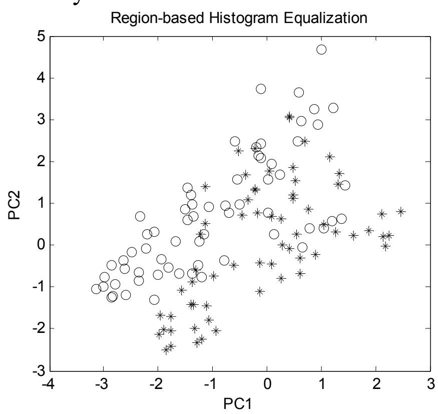

Figure 4: Normalized Face images distributions using Region-based histogram equalization method

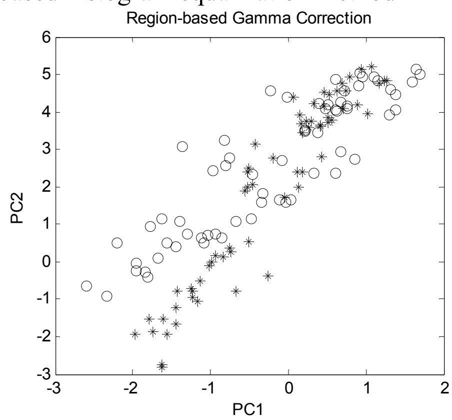

Figure 5: Normalized Face images distributions using Region- based Gamma Correction.

correction, and region-based methods.

Figure (7) shows the canonical face image used to calculate the γ coefficient of the GIC method. This image is the mean image of the training set enhanced with histogram equalization. In the region-based methods, the face images divided into four regions, and these regions are normalizing with the same algorithm of global HE or GIC.

Figure (8) shows results of applying different normalization methods for one subject in the CMU-PIE

Subset 1

Subset 2.

Subset 3.

Subset 4.

Figure 6: sample images of an individual divided into four subsets.

(a) (b)

Figure 7: (a) Canonical face image (b) Facial four regions.

face database. Region-based normalization methods increase the contrast of the image rather than global normalization methods. It is clear that the quality of the image is improved and feature parts such as eyes, nose and mouth are made more clearly.

The local normalization method has the disadvantage that the output is not necessarily realistic. However, in the problem at hand, the objective is not to have a realistic image but to obtain a representation of the face that is invariant to illumination, while keeping the information necessary to allow a discriminative recognition of the subjects. With this idea in mind, it makes sense to use local illumination normalization methods for this type of application.

The normalized gallery images are used to train two different shapes of SOM, sheet and cylindrical. Each gallery image is represented by the index of the winner neuron; in this experiment 2-dimensional map is used for training. The winner of the normalized probe image is considered as the feature vector of the face image, this vector is compared with all gallery feature vectors. Nearest neighbor classifier based on the Euclidean distance is employed for classification. Figure (9) shows the recognition accuracy of applying all normalization methods to process all images before training and testing. The recognition accuracy for each subset and the accuracy for all subsets are calculated. It is clear that RHE improve the accuracy of recognition.

By studding the performance of two different shapes of SOM. Cylinder map shape gives good results than the sheet map shape, because of the circular shape of the face manifold in the image space.

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

Figure 8: the processed images after applying different illumination normalization methods for one image in the CMU-PIE face database. (a) Unnormalized (b) GIC (c) RGIC (d) HE (e) RHE

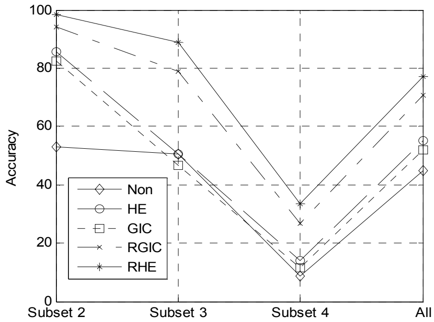

Figure 9: Recognition rate comparisons of different illumination normalization methods on CMU-PIE Face Database using cylinder SOM map.

6 Conclusion

The experimental results on CMU-PIE face database reveal a number of interesting points,

1. SOM performance highly degrades when the test sample lighting condition is different from the training samples condition.
2. Region-based illumination normalization outperforms the global normalization methods.
3. The cylinder shape of SOM is better than the sheet shape because face manifold of the training images tend to be circular.

References

[1] W. Y. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips. “Face recognition: A literature survey”, UMD Technical Report CAR-TR-948, 2000.

[2] S. Z. Li, A. K. Jon, “Handbook of Face Recognition”, Springer, 2005.

[3] Bo Du; Shiguang Shan; Laiyun Qing; Wen Gao, “Empirical comparisons of several preprocessing methods for illumination insensitive face recognition” , IEEE International conference on Acoustic, speech, and signal processing , vol.2,pii/981-ii/984, 2005.

[4] 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. On Systems, Man, and Cypernetics-part B: Cypernetics, vol. 36, No. 2, April 2006.

[5] P.N. Belhumeur, J. Hespanha and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE PAMI, Vol. 19, pp. 711-720, 1997.

[6] A. Georghiades, D. Kriegman, and P. Belhumeur.” From few to many: Generative models for recognition under variable pose and illumination” IEEE Trans. on Pattern Analysis and Machine Intelligence, 40:643--660, 2001.

[7] T. Kohonen, “Self-Organizing Maps”, Springer, 1997.

[8] Sim, T.; Baker, S.; Bsat, M. “The CMU pose, illumination, and expression database”, IEEE Trans. Pattern Anal. Mach. Intell., vol.25, no. 12, pp. 1615-1618, 2003