A COMPONENT-BASED DCT/LDA FACE RECOGNITION METHOD FOR CHARACTER RETRIEVAL IN TV PROGRAMS

Byunghee Jung, Myunghwan Ha, Heejung Kim, Keunsoo Park, Heunjin Lee and Whoiyul Kim**

Technical Research Institute, Korean Broadcasting System
E-Mail: {bhjung, mhha, hjkim, keunspark}@kbs.co.kr

Division of Electrical Engineering, Hanyang University*
E-Mail: {hjlee, wykim}@vision.hanyanag.ac.kr

ABSTRACT

As the digital multimedia services such as watching TV programs or movies through the Internet are more widely used, searching the shot or scene that includes special characters becomes a more useful function to both users and program-producing staffs. For character-based retrieval in TV programs, face detection and recognition are necessary. To be applicable to real character retrieval service, these face detection and recognition methods should provide the accuracy as well as real-time execution. For this, we have already proposed a fast face recognition method that uses DCT/LDA method. But there needs a more robust method under the various changes of facial conditions such as light, pose and facial expression, because TV programs contain those changes of conditions frequently. In this paper, we propose an advanced face recognition method based on DCT/LDA method that is component-based and processes scored ranking matching for final decision. Through the experimental result, we show that the proposed method improves the accuracy compared with the DCT/LDA method.

1. INTRODUCTION

The multimedia retrieval services on the Internet became popular owing to the technical improvement of the computer software and hardware and the enhancement of the network performance. Among the various multimedia retrieval services, retrieving a particular shot or scene that includes a special character is useful to both users and program-producing staffs. The face detection and recognition methods are necessary for this character-based retrieval. Owing to the large volumes of video, these methods should satisfy the requirements of speed and accuracy for automatic processing. However, the recognition rate of faces in a video is influenced by the various changes of facial conditions such as light, pose and facial expressions. To improve the accuracy of the face recognition method that is applicable to the real character retrieval service, the robust method under the change of facial conditions are necessary.

We have already proposed a fast face recognition method based on DCT(Discrete Cosine Transform)/LDA(Linear Discriminant Analysis) method [1]. The DCT/LDA method applies the DCT coefficient values of gray-scaled face image’s low frequency region to the LDA. The performance of this method is altered by the change of facial conditions. Therefore, to reduce the influence of the facial conditions, improvement is required.

In this paper, we propose an advanced face recognition method based on DCT/LDA method that produces better performance under the various facial conditions. The method uses component-based DCT/LDA method. That is, we extract DCT/LDA features of whole face images and face components. Then, in the similarity comparison phase, we compare the features using the weighted component-based similarity metric. The weighting factor is automatically computed by the correlation coefficient of each face component. In the final decision phase, we use the scored ranking matching rather than minimum distance matching. We can recognize a character that is located in the high-ranked positions frequently by this scored ranking matching even though the recognized face is not located in the first ranking.

The rest of this paper is organized as follows. In section 2, we review the already proposed DCT/LDA face recognition method. Then, we propose an advanced DCT/LDA face recognition method that is component-based in section 3. Finally, we present the performance of our method in section 4 and discuss the conclusion and the future work in section 5.

2. DCT/LDA METHOD

For character retrieval in TV programs, there needs face detection as well as face recognition. In this section, we review the already proposed our face detection and recognition method.

2.1. Face Detection Method for TV Program

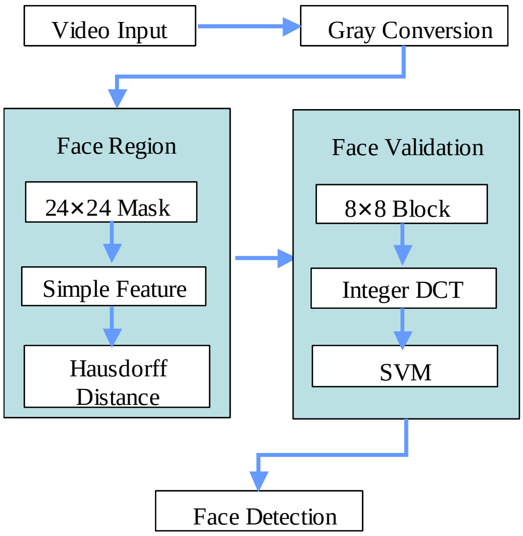

Fig. 1 Face Detection Process

As shown in Fig. 1, the face detection consists of twp steps: face candidate region detection and detected face validation. After detecting the face region using simple feature and Hausdorff distance, we validate the correctness of the region. Simple features are patterns of light and dark part in the face region. In the face validation, to reduce the computing time, we use IntegerDCT as an input for a SVM. The size of the input image is 352x240 and the mask image is 24x24. The mask image moves by two pixels and the input image reduces by 1.2 scales. (Any faces larger than 24x24 can be detected) After computing the IntergerDCT of the detectable 3340 images, the result is used as an input of the SVM. The SVM finally validates the face region. The average computing time of face detection is 62 ms/f (on Pentium-IV 17.Ghz), i.e. real-time detection in our experiments.

2.2. DCT/LDA Face Recognition Method

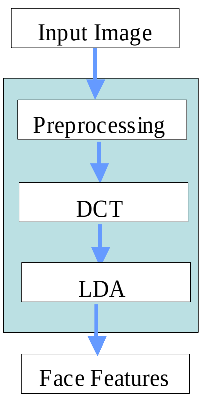

Fig. 2 Face Recognition Process

LDA is the face recognition method that projects the face features on a discrimination plane and recognizes the face. The LDA is applied in the form of PCA/LDA, a sequential combination of PCA and LDA [2]. PCA/LDA presents high performance, but PCA requires much computing power for computing the basis vector and the basis vector changes according to the training data sets. To alleviate these problems, we proposed DCT/LDA method. As shown in Fig. 2, DCT/LDA method applies the DCT as an input of the LDA rather than PCA. This approach reduces computing time, because it does not pass the training process to obtain the basis vector in PCA. Also this DCT/LDA method has an advantage of maintaining the fixed basis vector independent of the training data sets [2]. Our face recognition method works as follows. Firstly, we normalize the luminance factor by dividing the difference between the image's luminance values and the mean values into the variance. After the preprocessing, secondly, we apply the DCT and extract the DCT coefficients. To reduce the influence of illumination change, we exclude the DC element and obtain only the low frequency DCT coefficients using the raster scan. The DCT coefficients are inputs of the LDA. Thirdly, we obtain the basis vector via LDA training and then obtain DCT/LDA face feature values using this basis vector. Finally, the face recognition is the process of DCT/LDA values comparison. The minimum Euclidean distance is the result of the face recognition [3].

3. COMPONENT-BASED DCT/LDA METHOD

The performance of face recognition method that uses only global face as features can be highly affected by the change of light, pose and facial expression [4]. The reason is that the whole feature value is changed even though a part of face is changed by the change of facial conditions. In the other hand, using the face components such as eye, nose, and mouth can reduce the influence of the change of facial condition to the small component region [5]. Therefore, we uses face components as features in addition to the global face.

3.1. Extraction of Face Component Features

Among the several face components, we exclude mouth because the mouth is too changeable compared to the other face components [6]. Therefore, we use right eye, left eye and nose for our face recognition. In our approach, we use the gray-scaled face image that is normalized by the location of eyes, i.e, (45x56). In this normalized face image, we define the location of each face component as shown in Table. 1.

Face Component	(x_1, y_1)	(x_2, y_2)
Whole Face	(6,2)	(2,55)
Right Eye	(9, 15)	(25, 30)
Left Eye	(22, 38)	(15, 30)
Nose	(15, 18)	(31, 36)

Table. 1 Location of Face Component

Fig.3 shows the feature extraction process of our component-based DCT/LDA method. After preprocessing that normalizes the light condition, we extract DCT/LDA values of global face, right, left eye and nose. The final feature value to be used for face recognition is generated by the concatenation of each DCT/LDA feature value.

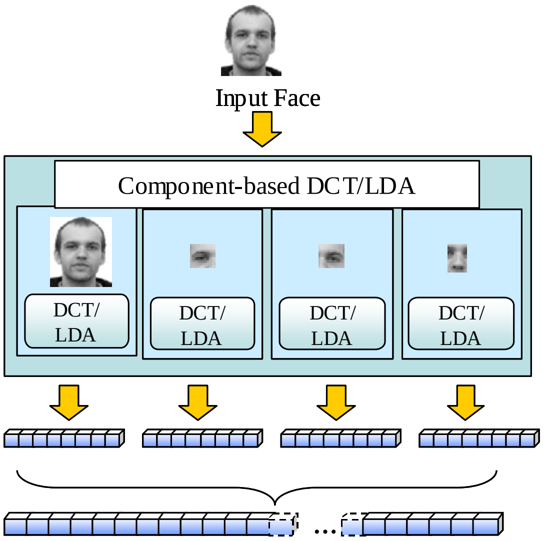

Fig. 3 Component-based Feature Extraction Process

3.2. Weighted Component-based Similarity

When the facial conditions change, certain face components are more affected than the others. For example, if the light is thrown from left direction, then the left eye is affected more than right eye. Therefore, in similarity comparison phase, there needs a metric that reflect this difference. For more affected components we assign low weighting factor. And we assign high weighting factor for less affected components. This method elevates the recognition rate by decreasing the influence of change of condition. As a method for computing the weighting factor, we use correlation coefficient between the image of input face and mean face. The mean face is the average of face images that is used in the training phase. Fig. 4 shows an example of computing left eye's weighting factor.

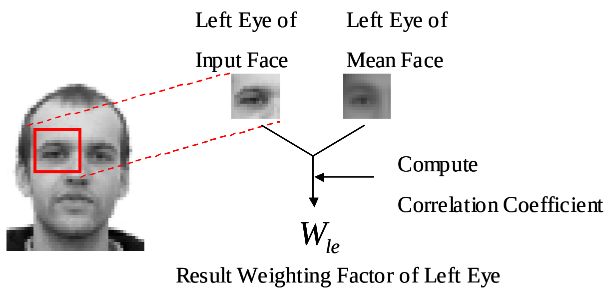

Fig. 4 Example of Computing Weighting Factor

In the example of Fig.4, the correlation coefficient W is yielded by Formula (1). The Formula (1) is the same in the case of left eye, right eye, nose and whole face.

$$W_i = \frac{\sum_x \sum_y \left[f_i(x,y) - \bar{f}_i\right]\left[m_i(x,y) - \bar{m}_i\right]}{\left\{\sum_x \sum_y \left[f_i(x,y) - \bar{f}_i\right]^2 \sum_x \sum_y \left[m_i(x,y) - \bar{m}_i\right]^2\right\}^{1/2}} \qquad (1)$$

$i \in$ {whole face(h), left eye(le), right eye(re), nose(n)}
$f(x, y)$: Face Component Image of Input Face
$\bar{f}$: Average of $f(x, y)$
$m(x, y)$: Face Component Image of Mean Face
$\bar{m}$: Average of $m(x, y)$

As presented in Formula (2), the similarity comparison metric is derived using the result of Formula (1). When comparing a face, we compare the input face with the faces in the faceDB.

$$d(\mathbf{V}_t, \mathbf{V}_r) = W_h \sum_{k=0}^{N} (v_{t,k} - v_{r,k})^2 + W_{le} \sum_{k=N+1}^{2N} (v_{t,k} - v_{r,k})^2 + W_{re} \sum_{k=2N+1}^{3N} (v_{t,k} - v_{r,k})^2 + W_n \sum_{k=3N+1}^{4N} (v_{t,k} - v_{r,k})^2 \qquad (2)$$

$\mathbf{V}_t$: Feature Value of Input Face
$\mathbf{V}_r$: Feature Value of Faces in the FaceDB
W_i: Weighting Factor
N: Number of Feature Values

3.3. Scored Ranking Matching

When we finally recognize a face after the comparison phase, we generally choose a minimum distance face. However, if the correct face is not the minimum distance face and it appears many times in the high-ranked positions, the face cannot be recognized. For this case, we apply the scored ranking matching to the result of similarity comparison phase. This scored ranking matching method assigns a particular score to the faces that are arranged in increasing order of similarity comparison metric. In this paper, the score is determined by the experiment so that the total score becomes low, as the weighted component-based similarity distance is larger. So the recognized face is the face that has the highest score. As shown in Fig. 5, as the face appears more times in high-ranked positions, the possibility of being recognized becomes higher. Our scored ranking method includes the minimum distance matching method.

4. PERFORMANCE EVALUATION

To evaluate the performance of our component-based DCT/LDA face recognition method, we compare our method with the DCT/LDA method. The data set is selected from 1500 images of Alkom data sets. These Alkom data sets are already used in the MPEG-7 face recognition descriptor's experiment. The Alkom data sets consist of 46×56 sized gray images and 15 images per person. The test images are normalized according to the position of the eye and the illumination and pose changes exist [7].

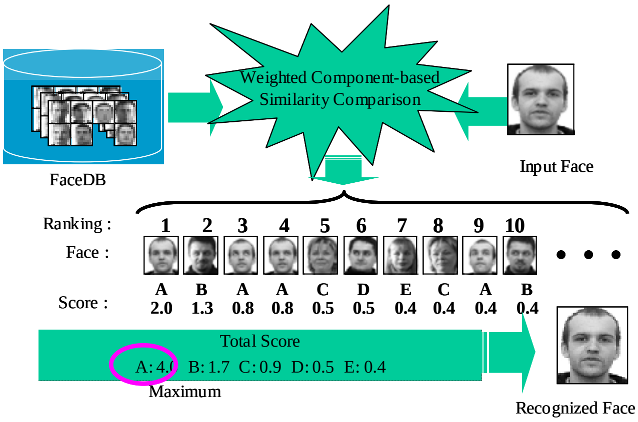

Fig. 5 Scored Ranking Process

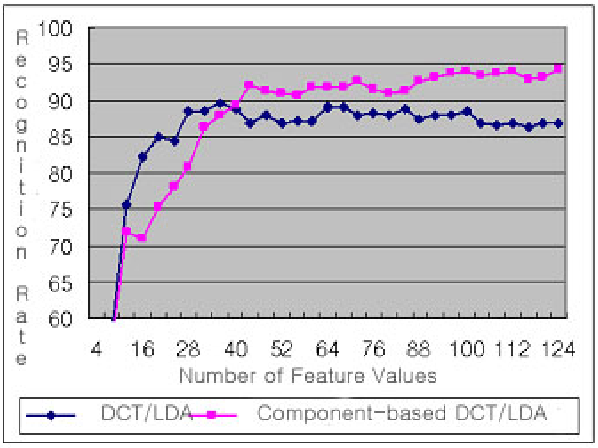

Fig. 6 Performance of Component-based DCT/LDA

Fig. 6 shows the experimental result when 50% data of Alkom data sets are trained and the rest 50% data are used for test. We changes the number of feature values of each component from 1 to 31. In case that the number of final feature values is less than 40, DCT/LDA shows better performance. This is because the feature values for the face component is not sufficient. So when we assign more than 10 feature values to each face component, the proposed component-based DCT/LDA shows better performance than DCT/LDA. Table 2. shows the comparison of recognition rate between the improved performance of scored ranking matching and the minimum distance matching.

Matching	Recognition Rate (%)
Minimum Distance Matching	93.75
Scored Ranking Matching	94.06

Table 2. Performance of Scored Ranking Matching

5. CONCLUSION

As the video delivery service on the Internet becomes popular, the retrieval service of shots or scenes that include a special character becomes more necessary. For character-based retrieval in TV programs, face detection and recognition are necessary. To be applicable to real character retrieval service, these face detection and recognition methods should be robust under the various changes of facial conditions such as light, pose and facial expressions. In this paper, we propose a component-based DCT/LDA face recognition method that improves the recognition rate of DCT/LDA method regardless of facial conditions. Our method uses the weighted component-based similarity comparison method in similarity comparison phase and the scored ranking matching in final decision phase. Our experiments show that the proposed method produces better performance compared with the DCT/LDA method.

There are many research issues to do further. For example, developing an automatic score assigning algorithm and building up a client/server model for a real character retrieval service are those. To develop an automatic score assigning method, the statistic approach can be applied.

6. REFERENCES

[1] B. H. Jung, M. H. Ha, H. J. Kim, K. S. Park and W. Y. Kim, " Fast Face Recognition System for Character Retrievla In TV Programs", Proceedings of International Workshop on Advance Iamge Technology" (to be published in January, 2004).

[2] W. Zhao, R. Chellappa, and A. Krishnaswamy, "Discriminant analysis of principal components for face recognition," In Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, pp. 336-341, 1998.

[3] Ziad M.Hafed and Martin D. Levine, "Face Recognition using the discrete cosine transform," International Journal of Computer Vision, Vol. 43, No.3, pp.167-188, July 2001.

[4] H. W. Kim, T. K. Kim, J. H. Lee, W. J. Hwang, S. C. Kee, "Component-based LDA Face Descriptor," ISO/IEC JTC1/SC21/WG11 M8243, Fairfax, May 2002.

[5] B. Heisele, P. Ho, and T. Poggio, "Face Recognition with Support Vector Machines: Global versus Component-based Approach," In Proc. IEEE International Conference on Computer Vision, 2001

[6] B. Moghaddam, A. Pentland, "Face Recognition using View-based and Modular Eigenspaces," Automatic Systems for the Identification and Inspection of Humans, SPIE Vol. 2277, July 1994

[7] M. Bober ed., "Description of MPEG-7 Visual Core Experiments," ISO/IEC JTC1/SC29/WG11 N4925, October 2002.