한국산학기술학회논문지
Vol. 8, No. 2, pp. 267-272, 2007

Detection of Facial Feature Regionsby Manipulation of DCT's Coefficients

Boo-Hyung Lee[1*] and Jang-Ryeol Ryu[2]

DCT 계수를 이용한 얼굴 특징 영역의 검출

이부형[1*], 류장렬[2]

Abstract This paper proposes a new approach for the detection of facial feature regions using the characteristic of DCT(discrete cosine transformation) thatconcentrates the energy of an image into lower frequency coefficients. Since the facial features are pertained to relatively high frequency in a face image, the inverse DCT after removing the DCT's coefficients corresponding to the lower frequencies generates the image where the facial feature regions are emphasized. Thus the facial regions can be easily segmentedfrom the inversed image using any differential operator. In the segmented region, facial features can be found using face template. The proposedalgorithm has been tested with the image MIT's CBCL DB and the Yale facedatabase B. The experimental results have shown superior performance under the variations of image size and lighting condition.

Key words : Facial Feature, Face Region, DCT, IDCT, IDCTed image

요 약 본 논문에서는 DCT계수의 특성을 이용하여 조명조건이나 얼굴의 크기에 무관하게 얼굴특징영역을 검출하기 위한 새로운 방법을 제안한다. 일반적으로 영상을 DCT변환하면 영상의 에너지가 저주파영역에 집중되는 특성을 가지나 얼굴 특징요소들은 얼굴영상에서 비교적 고주파 성분들을 포함하고 있기 때문에 저주파에 해당되는 DCT계수들의 일부를 제거한 후 역변환을 취하면 얼굴특징영역이 강조된 영상을 얻을 수 있다. 따라서, 본 논문에서는 DCT변환된 영상으로부터 저주파 계수의 일부를 제거하여 얼굴특징요소 후보들을 추출한 후 템플릿을 적용하여 얼굴특징요소 영역을 결정한다. 얼굴특징요소 영역이 결정되면 얼굴특징요소 추출 알고리즘을 적용하여 눈, 코, 입을 구별한다. 제안된 알고리즘을 MIT의CBCL DB와 Yale facedatabase B 에 적용하여 실험하였다. 실험결과 DCT변환된 영상에서 저주파 일부의 계수를 제거한 후 얼굴 특징영역을 검출했을 경우 그렇지 않은 영상보다 영상의 크기와 조명조건의 변화에 무관하게 인식성능이 향상됨을 알 수 있었다.

1. Introduction

Many researchers have proposed various approaches for face region detection problem, which should be accomplished prior to face recognition process. Recently, Yang[1] has classified these approaches into four categories: knowledge-based, feature invariant, template matching and appearance-based approaches. Among these approaches, feature invariant ones have been used in popular, where the face region is determined by searching the facial features such as eyes, eyebrows, nose and mouth, as well as skin color and textures which exist in any face regardless of sex, age, and race[2~5]. One difficulty of using these approaches is that detection of the facial features is significantly influenced by lighting condition and image size. To solve this difficulty, many researchers have introduced various methods. Basri[6] and Ramamoorthi[7] have respectively represented convex Lambertian objects using 9-D linear

[1]Dep. of Computer Engineering, Kongju National University
[2]Dep. of Electrical & Electronic Engineering, Kongju National University
*Corresponding author :Boo-Hyung Lee(BHL1998@kongju.ac.kr)

subspace and shown that the images of convex Lambertian objects acquired under various lighting conditions can be approximated by the harmonic images spanned by 9 eigen-images. These methods have shown much improved performance in recognizing faces under irregular lighting conditions. However, they suffered from increased computations due to complexity of the algorithm to generate harmonic images and from the difficulty in applying to small images since the facial features cannot be extracted easily when image size is small. To enhance the performance in the small image cases, Baek [8] has proposed the method of using major color components to detect the face region and LDA(linear discriminant analysis) technique for determining its faceness. Although it showed its efficiency in detection of small faces, it also suffered from variation of face color.

To solve such problems of the aforementioned approaches as the performance deterioration depending on the image sizes, lighting conditions, and face colors, this paper proposes a new method of detecting the facial feature regions using the characteristic of DCT(discrete cosine transformation) that DCT concentrates the energy of the image on the lower frequency coefficients. Since the facial features occupy the higher frequency regions in a face image, the facial feature regions will remain by taking the inverse DCT with eliminating the lower frequency coefficients of DCT. In the remained facial regions, the boundaries of facial features will be emphasized since the DC component is eliminated in the region. The experimental results have shown that the proposed algorithm works robust on the variation of image size and lighting condition.

2. Detection of Face Regions using the Characteristic of DCT

2.1 DCT

DCT is one of the transforms, such as DFT (Discrete Fourier Transform), which map the signal in time domain into frequency domain. Differently from other transforms whose eigen-functions are complex ones including imaginary terms, DCT uses the real functions as the eigen-functions as given in Eq. (1) and concentrates the energy of an image into the coefficients of lower frequency components.

$$F(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \cos\left(\frac{\pi(2x+1)u}{2N}\right)\cos\left(\frac{\pi(2y+1)v}{2N}\right) \quad (1)$$

where $f(x,y)$ is the input image, and its coefficients $\alpha(u)$ and $\alpha(v)$ are given as follow:

$$\alpha(u) \text{ or } \alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u \text{ or } v = 0 \\ \sqrt{\frac{2}{N}} & \text{for } u \text{ or } v \neq 0 \end{cases} \quad (2)$$

The coefficient of the lowest order in DCT $F(0,0)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)$ represents the DC component which is the average of the image. As the order increases, the coefficient represents the higher frequency component. Fig. 1 is a face image and the result of DCT applied to a face. It can be found easily from Fig. 1that only the coefficients of lower order have the large values and thus the image canbe approximated with only the coefficients of lower order. This is why DCT is popularly used for image compression.

Fig.1 An face image and it's DCT image

2.2 Inverse DCT with removing the coefficients of lower order

When an input image is small or lighting condition is irregular, any edge operator cannot properly detect the boundaries of the facial features since the boundary is not clear or false edges can be detected. To overcome these problems, this paper takes the DCT of an input image and than applies the inverse DCT with eliminating the coefficients of lower orders corresponding to the lower frequencies. In the result image, the facial feature regions are emphasized and the boundaries of the facial features can be detected easily.

In Eq. (3) describing the inverse DCT, $F(u, v)_{rem}$ is the set of the DCT coefficients where the coefficients corresponding to the DC and lower frequency components are eliminated.

$$f(x,y)_{rem} = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\alpha(u)\alpha(v)F(u,v)_{rem} \cos\left(\frac{\pi(2x+1)u}{2N}\right)\cos\left(\frac{\pi(2y+1)v}{2N}\right) \quad (3)$$

Fig.2 shows the result imagesobtained by applying Eq. (1) and (3) in sequence to three different face images.

Original face image	IDCTed Image with removing DCT coefficients			
	~ $F(1,1)$	~ $F(2,2)$	~ $F(3,3)$	~ $F(4,4)$
				
				

Fig.2. The original and resultimages of IDCT with removing the lower frequency DCT coefficient

In the figure, the first column shows the case where the DCT coefficients from $F(0, 0)$ to $F(1, 1)$ are eliminated and the inverse DCT is followed. The second one does the case where the DCT coefficients from $F(0, 0)$ to $F(2, 2)$ are eliminated, the third one and the fourth one respectively show the cases where the DCT coefficients from $F(0, 0)$ to $F(3, 3)$ and $F(4, 4)$ are eliminated. It can be found in the figures that the relatively brighter regions include the facial features such as eyes, nose, and mouth. Since the histogram of the result image has a complete bimodal shape as the example given in Fig. 3

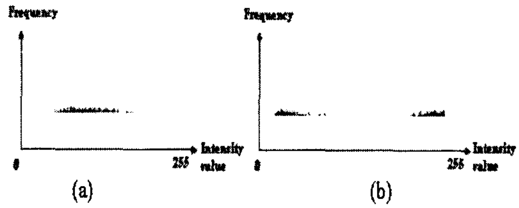

Fig.3. Comparison of the histograms of the original image $F(x, y)$ in (a) and the processed image $F(x, y)_{rem}$ in (b)

2.3 Extraction of facial features in the facial feature regions

Once the IDCTed image is obtained with eliminating the DCT coefficients of lower order, its similarity with the face template is tested. The face template given in Fig. 4(d) is constructed by calculating the average image of the IDCTed face images of the real face images. In order to construct the face template, the face images included in the database of MIT CBCL have been used. From the face images in the DB such as given in Fig. 4(a), the facial feature regions are segmented as given in Fig. 4(b) and the proposed process is applied to obtain the IDCTed image as given in Fig. 4(c). It is experimentally found that the facial feature regions are most clearly segmented in the IDCTed images when the DCT coefficients from $F(0, 0)$ to $F(3, 3)$ are removed. Fig. 4(d) shows the face template finally obtained by the experiments.

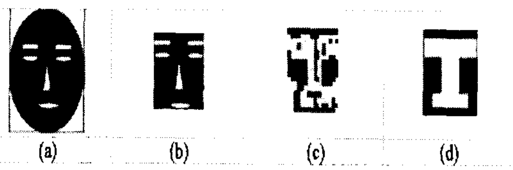

Fig.4. Construction of the face template from the IDCTed images of the real face image

If the facial feature regions fit the face template, then the facial features are extracted in those regions. In the IDCTed images, the brightness of facial feature regions is tilted to the higher range as can be seen in Fig. 4(b). For efficient detection of the facial features, the images are stretched over the whole range.

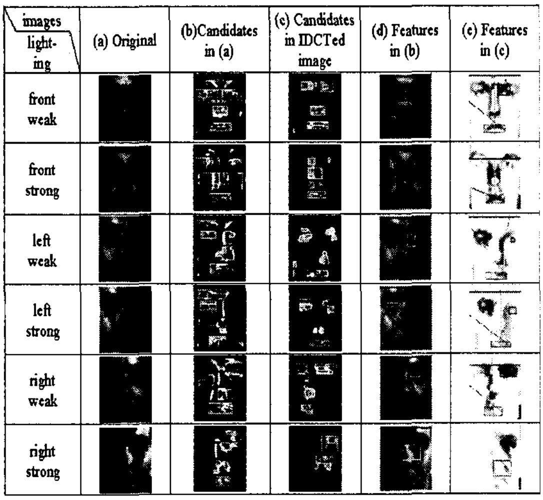

Fig.5. Construction of the face template from face images.

Fig. 5 shows the facial features extracted from the original image and the IDCTed image, using the algorithm provided by Kim and Hahn[9]. The algorithm segments the facial feature regions into three regions for eyes, nose and mouth using the horizontal histogram of the edge image. Then in each region, the candidatesfor the facial features are extracted by using the vertical histogram in the region. The facial features are selected among the candidates based on the topology of the facial features.

In Fig. 5, the first two images are obtained with weak and strong front light respectively, the next two images are obtained with weak and strong left light respectively, and the next two images are obtained with weak and strong right light respectively. The first column shows the input images. The second and third columns show the candidates for facial features such as eyes, nose and mouth extracted from the input image and the IDCTed image, respectively, using the Kim's algorithm. The fourth and the fifth column show the facial features finally selected among the candidates in (b) and (c) respectively. As can be compared with the images in (d) and (e), the facial features appear more clearly when the proposed algorithm is used.

In the case of using the original images, the algorithm extracts too many candidates for the facial features to select the genuine ones. For the front weak image case, although the genuine facial features are included in the candidates, the algorithm cannot select them since the other sets of candidates satisfy the topology condition of the facial features. However, when using the IDCTed image, the unnecessary candidates are not appeared in the image and thus the facial features are selected more clearly and correctly. This is well illustrated by the pictures in the first and second rows.

When the light is polarized as shown in the last four rows, the features in the shadowed area cannot be extracted in most cases. Especially when the shadow is dark, even the candidates cannot be obtained at all by using the original images as shown in the fourth and sixth rows. However, it is shown that some of the features are recovered in the IDCTed images. By experiments, the proposedalgorithm can detect the facial features successfully in most cases, except the strongly shadowed images as shown in the last row.

3. Experiments and Discussion

The performance of the proposed algorithm has been tested with two sets of face databases, MIT CBCL's DB and Yale facedatabase B. MIT CBCL's DB provides face and nonface images of 16*16 size, and Yale facedatabase B provides face images of 16*16 size collected under various lighting conditions. Using these two DBs, the performance is measured in three viewpoints. The first one is how accurately it discerns faces and nonfaces, the second one is how accurately it detects faces acquired under various lighting conditions, and the third one is how well it applies to the variation of image size.

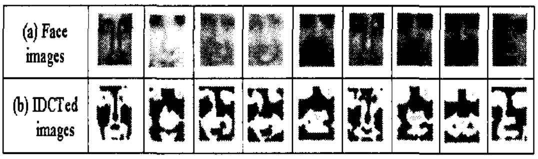

Fig.6. Example images of face included in the MIT CBCL's DB and their IDCTed image

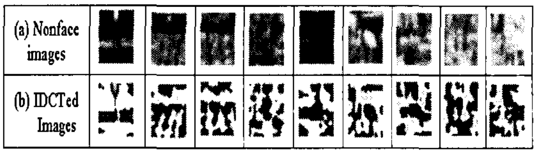

Fig.7. Example images of nonface included in the MIT CBCL's DB and their IDCTed images

The accuracy of discerning the faces and nonfaces is tested using the MIT CBCL's face DB. Fig. 6 shows the examples of face imagesand their IDCTed images. By testing their similarity with the face template whether it is larger than 0.5, they are decided possible faces. The nonfaces whose example images are given in Fig. 7 have the similarities less than 0.5. The results of the classification of face and nonface using the provide algorithm are summarized in Table 1.

Table 1 shows that the success rate of discerning the faces and nonfaces when using MIT CBCL's face DBnot having lighting variations is 100%. Accordingly, we can see that the proposed algorithm have ability to separateedges having facial features and non-edges robust.

Table 1. Results of classification of the faces and nonfaces in the MIT CBCL's face DB xample images of nonface included in the MIT CBCL's DB and their IDCTed images

Image Class	Number of Images	Decision		Success Rate(%)
		Face	Nonface	
Face	2429	2429	0	100
Nonface	4548	0	4548	100
Total	6977	2429	4548	100

The effects of the lighting conditions on the performance of the proposed algorithm are tested with the Yale facedatabase B. The DB includes total 630 images of 10 persons. A quarter of these imagesare shadowed strongly by a polarized light. Fig. 8 shows the images of one person taken under 10 different lighting conditions. Among these, 620 images have been used for the experiment to match the numbers of images taken under the left and the right light directions.

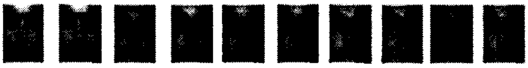

Fig.8. Example images of the same person taken under different lighting conditions

Table 2. Results of face feature detection in the images acquired under different lighting conditions

Image Size	Number of Face Images	Decision		Success Rate(%)
		Face	Nonface	
16*16	620	480	140	77.4

Table 2 shows the results of the decision algorithm applied to the Yale facedatabase B. The results show that the facial features are successfully detected in most images except in the strongly shadowed images. Since the success rate is under 50% when using the original images, it is conformed that the proposed algorithm is more robust on the variation of lighting conditions.

The robustness of the algorithm on the variation of image sizes is conformed by applying the algorithm to the enlarged images of MIT CBCL's DB. The images were enlarged to 32*32 and 64*64 and the same performances were obtained with the 16*16 images.

4. Conclusions

This paper has proposed a new method of detecting the facial feature regions using the characteristic of DCT that it concentrates the energy of the image on the coefficients of lower order. By removing those coefficients and taking the inverse DCT, a new face image is obtained from which the facial features are emphasized even the lighting condition is so bad that the conventional algorithms cannot find the facial features. The performance of the proposed algorithm has been tested with two sets of face DBs, MIT CBCL's face DB and Yale facedatabase B. The experimental results have shown that the algorithm accurately discerns the faces and nonfaces, and that it accurately determines the faceness even when the lighting condition is so bad that a part of face is not clearly shown.

References

[1] M. Yang, D. J. Kriegman, N. Ahuja , "Detecting Faces in Images : A Survey", IEEE Trans. on Pattern Analysis & Machine Intelligence, 24(1), pp. 34-58, 2002.

[2] S. Zhang and Z. Liu, "A Real-Time Face Detector", 2004 IEEE International Conference on Systems, Man and Cybernetics, 3, pp. 2197-2202, 2004.

[3] G. Wei and I.K. Sethi, "Omni-face detection for video/image content description", Proceedings of the 2000 ACM workshops on Multimedia, pp. 185-189, 2000.

[4] V. Manian and R. Vasquez, "Approaches to color and texture based image classification", SPIE JOE, 41, pp. 1480-1490, 2002

[5] S. Hong, S. Lee, B. Lee, S. Pyun, and S. Park, "Facial Image Recognition using Neural Networks and Characteristic Extraction based on DCT", Journal of IEEK, vol. 38-4, pp. 122-129, Dec. 2001.

[6] R. Basri and D. W. Jacobs : Lambertian Reflection and Linear subspaces, IEEE Trans. on Pattern Analysis & Machine Intelligence, 25(2), pp. 218-232, 2003

[7] R. Ramamoorthi and P. Hanrahan : On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object, Journal of the Optical Society of America, 18(10), pp. 2448-2459, 2001.

[8] K. Baek, H. Jang, Y. Han, and H. Hahn : Efficient small face detection in surveillance images using major color component and LDA scheme, CIS 2005, LNAI 3802, pp. 285-290, 2005

[9] S. Kim, H. Hahn, S. Lim, and H. Cha : Block Based Face Detection Scheme Using Face Color and Motion Information, Journal of Korea Fuzzy Logic and Intelligent Systems, 13(4), (2003) 461-468

Bool-Hyung Lee [Regular Member]

- Feb.,1983:Dep. of Electronic-Engineering, Soongsil University (B.S)
- Aug.,1989:Dep. of Electronic-Engineering, Soongsil University (M.S)
- Feb.,1998:Dep. of Electronic-Engineering, Soongsil University (Ph.D)
- Mar.,2006~:Currently, Professor of Dep. of Computer Engineering, Kongju National University

<Interesting Field>
3-D Object Recognition, Image Processing, 3-D Graphics and Robotics, Character Recognition

Jangl-Ryeol Ryu [Regular Member]

- 1982 : Dep. of Electronic-Engineering, Inha University (B.S)
- 1985: Dep.of ElectronicEngineering, Chungnam National University (M.S)
- 1994: Dep.of ElectronicEngineering, Chungnam National University (Ph..D)
- 2006~: Currently, Professor of Dep. of Electrical & Electronic Engineering, Kongju National University

<Interesting Field>
Semiconductor & Materal, Semiconductor Device structure, IC Design