Pattern Recognition Letters 30 (2009) 856–860

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Face detection for video summary using illumination-compensation and morphological processing

Jae-Ung Yun, Hyung-Jin Lee, Anjan Kumar Paul, Joong-Hwan Baek *

Department of Information and Telecommunications Engineering, Korea Aerospace University, 200-1 Hwajun-dong, Goyang, Republic of Korea

ARTICLE INFO

Article history:
Available online 3 May 2009

Keywords:
Video summary
Face detection
Illumination-compensation
Morphological processing

ABSTRACT

This paper presents a simple and robust face detection algorithm that can be utilized to video summary. Because the characteristic of the face can be changed from unexpected condition like as shadows or lighting, we firstly process the illumination-compensation for maintaining face components. We then analyze color-based face region in *YCbCr* space to obtain the skin color. Also, we apply the morphological processing to improve the detection performance and reducing the number of false detected face regions. This process finds the candidates in face region and removes noise from the non-face region. For localization of the face region, we make a proper face ratio based on golden ratio. We evaluate our algorithm in the various genres. Experimental results demonstrate the effectiveness of our face detection algorithm that leads 96.7% in precision ratio on the average. The proposed method is applicable to video summary because of these high performances with low complexity.

1. Introduction

With the fast growth of multimedia information, content-based video analyses such as video summary and retrieval have attracted many researchers in recent years (Li et al., 2004). Now a days, video contents of drama, movie or news are easy to access due to the availability of high-speed wireless communication networks and high computing ability of the systems. As the volume of contents increases sharply, it is difficult to access some particular content in video when a user is interested. This is because the user may not know the location or time stamp of his/her interested actor (as for example) or some specific event in the video. This is the reason why video summary is necessary for convenient retrieval process and face detection is a crucial step in these video summary applications (Kim et al., 2004).

The purpose of face detection is to detect and locate the face in given images, and its performance has great effect on face tracking, face recognition and actor-based video summary (Jin et al., 2004). It also has potential applications in surveillance and census systems. The great face detection system can provide valuable insight on how one might approach other similar feature and pattern detection problems (Sung and Poggio, 1998).

Although face detection has various benefits, we need to consider some problems. The human faces can either be present or absent from a face. And we also think about lighting effects from external conditions like as camera fading and shadows. These systems use the techniques such as principal component analysis (PCA), geometric analysis, neural network, motion information, statistical modeling methods, SVM based method and color-based method. Commonly motion information and skin color are useful methods for detecting face although it is hard to get high performance using one color-based approach for face detection.

In this paper, we present a robust face detection method through the illumination-compensation algorithm using *YCbCr* space and morphological processing as supplement that is able to remove noise and can help to find clear face. Using some attribute of face, this gives the information about face of interest (FOI) that helps increasing the performance on face databases. Fig. 1 shows our face detection system.

This paper is organized as follows. Firstly we introduce the illumination-compensation algorithm to detect even under the unclear brightness conditions in Section 2. Section 3 describes the skin tone analysis and shows the processing of FOI to improve detection ratio. Localization of face is discussed in Section 4. Section 5 presents the detection results of our algorithm on several face databases. Conclusion and future works are given in Section 6.

2. Illumination-compensation technique

The skin tone color depends on the illumination conditions like an artificially created light inside, strong sunlight as brightness cases and like the shadow, fading effects as dark cases. In this paper, we firstly introduce the compensation method to remove the illumination effects. The illumination-compensation algorithm shown in Fig. 2 consists of 4 steps.

* Corresponding author. Fax: +82 2 3159 9969.
E-mail addresses: semodamo2@kau.ac.kr (J.-U. Yun), libae@kau.ac.kr (H.-J. Lee), akpaul@kau.ac.kr (A.K. Paul), jhbaek@kau.ac.kr (J.-H. Baek).

0167-8655/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2009.04.010

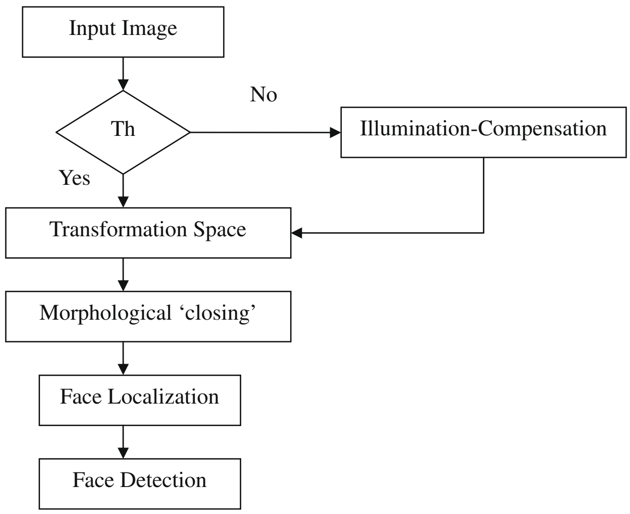

Fig. 1. Flow diagram of the face detection system.

We use a 20×20 window to compute the minimal value which generates 16×12 minimal brightness plane (the size of the original image: 320×240) shown in Fig. 3c. These brightness planes are resized by bilinear-interpolation method in Fig. 3c. At last, the illumination-compensation images are further normalized by applying histogram equalization.

The output intensity levels, *out _level*, are obtained by the following transformation on input levels:

$$out\_level = \int_0^x p_x(w)dw \tag{1}$$

where, w is the integral parameter and $p_x(w)$ denotes the probability density function (*PDF*).

$$out\_level_j = \sum_{i=1}^{j} p_x(x_i) = \sum_{i=1}^{j} n_i/n \tag{2}$$

For $j = 1,2,3,\ldots,L$, where out_level_j is the output intensity values corresponding to value x_j in the input image (Fig. 3d).

The illumination-compensation can be utilized alternatively when the image is too bright or dark. We use the "reference value" to decide the criteria of irregular conditions. If the average values of the components are smaller (<260) or larger (>550) than some threshold values in the image, the illumination-compensation carries into execution before skin tone detection. In case of uniform distribution of luminance, the compensation step may be skipped. Fig. 3 illustrates that brightness or darkness in images can be changed to detect face more precisely.

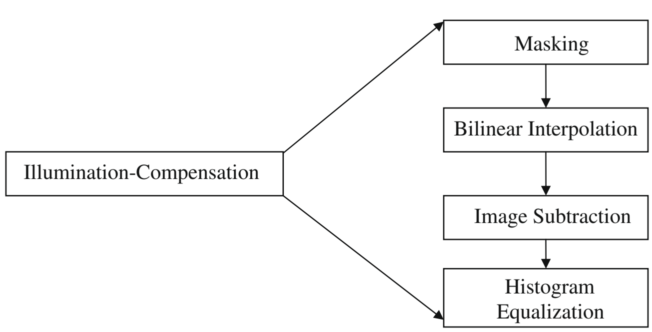

Fig. 2. Illumination-compensation processes.

3. Skin tone detection

The first step of the skin tone detection method is the transformation of color space from *RGB* to *YCbCr* space because *RGB* is not necessarily the most efficient representation of color due to the HVS (Human Visual System) is less sensitive to color than to luminance. After transformation, we extract the skin tone and carry out the morphological processing for face region detection.

3.1. Skin color modeling

Skin color modeling requires choosing an appropriate color space (Hsu and Mohamed, 2002). Comparison of nine different color spaces for face detection, the tint-saturation-luma (TSL) space provides the best results for two kinds of Gaussian density models (Terrillon et al., 2000). Nevertheless, we adopt the *YCbCr* space because the *RGB* color space contains color information that is less sensitive to *HVS* than luminance. The colors are equally important and the luminance is present in all three color components in *RGB* color space. On the other hand, the *YCbCr* space can represent a color image more efficiently by separating the luminance from the color information (Richardson, 2003). The equation of transformation between *RGB* to *YCbCr* is given as the following Eq. (3):

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & 0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3}$$

We collect 100 and 200 samples for training skin tone without and with the illumination-compensation, respectively. They are from various genres such as drama, movie, news and actor images and the sizes are 258×228 and 96×86 pixels. Statistical distributions of *CbCr* components of the training data are shown in Fig. 4a and b. Fig. 4a shows the *Cb* and *Cr* histogram without the illumination-compensation process, and Fig. 4b is with the illumination-compensation. Based on *Cb* and *Cr* information, we can detect the skin tone region shown in Fig. 5.

3.2. Morphological processing

To improve the efficiency of face detection, we apply the morphological process as a combination of dilation and erosion. The operation of dilation and erosion are fundamental to morphological processing and are used most often in various combinations in practical processing such as the opening, closing and hit-or-miss (Gonzalez et al., 2004). We utilize the closing method to remove the noise from face region and non-face regions.

The closing of A by B, as denoted below, is a dilation followed by an erosion:

$$A \cdot B = (A \oplus B)\Theta B \tag{4}$$

Like opening, morphological closing smoothes the contours of objects. Unlike opening, however, it generally joins narrow breaks, and fills holes smaller than the structured element. The process is illustrated in Fig. 6. Using the closing method, we can remove the noise and make the face region detection process more robust.

4. Face localization

The face candidate regions are found by using closing algorithm as mentioned in Section 3.2. As we have mentioned above, the purpose of our research is to achieve high detection efficiency using low computational complexity. The eye and mouth regions are the most important features for face detection and recognition. However, we locate the face directly using the area of FOI and simple geometric features. This task can be divided into (3) subtasks:

Fig. 3. Examples of the illumination-compensation. (a) Original images. (b) Y component images. (c and d) Bilinear interpolated images by the minimal value in 20×20 window. (e) Illumination-compensated images.

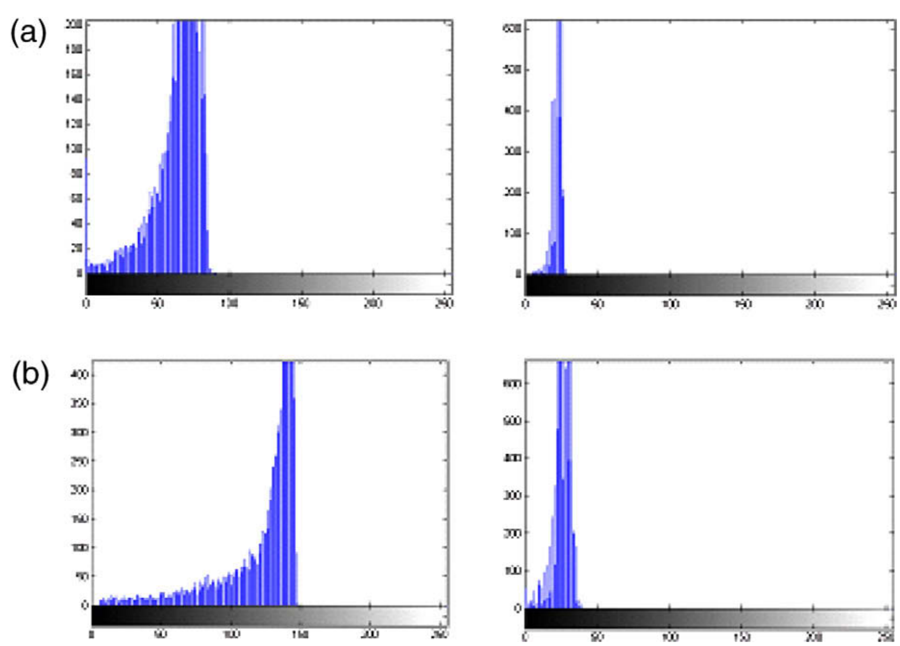

Fig. 4. *Cb* and *Cr* histograms for skin tone detection. (a) Without the illumination-compensation. (b) With the illumination-compensation.

(1) pre-processing, (2) calculation of the area, and (3) face region detection. At first, for the localization of face, we fill holes in the candidate regions in the pre-processing stage.

A hole is a set of pixels in the binary image with similar color of the background; but, originally, holes do not belong to the background. In second step we calculate the area of filled region and then obtain the centroid of that area. It can give us important information for determining the face region sizes.

The face region detection step draws boundary of face using geometrical feature 'golden ratio'. The face has a general ratio between width and height. Therefore, we utilize this information to make a face boundary. For an ideal face, the golden ratio is denoted as Eq. (5):

$$\varphi = Height/Width = \left(1+\sqrt{5}\right)\Big/2 \approx 1.618033989 \tag{5}$$

Fig. 5. Results of the skin tone region detection. (a) Original images. (b) Results.

Fig. 6. Morphological image processing. (a) Original images. (b) Skin tone detected images. (c) Results of the closing process.

Since the face database in our research contains both frontal and side face images, we make an empirical modified face ratio from

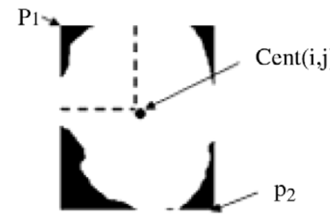

Fig. 7. New golden radio.

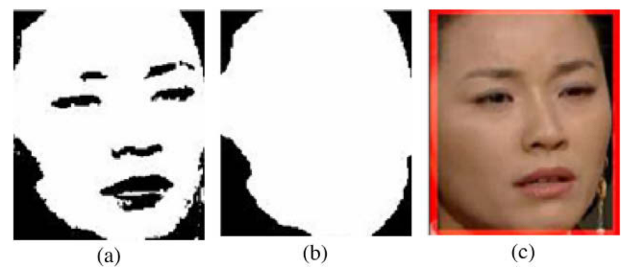

Fig. 8. Face detection result. (a) Candidate of face image. (b) Localization step. (c) Final result.

golden ratio, height:width = 1.28:1. The empirical experiments were performed by 30 images randomly chosen from dataset (see Fig. 7)

$$\begin{aligned} P_1(i,j) &= Cent\left(i-\sqrt{area}/2, j-\sqrt{area}/2\right)\\ P_2(i,j) &= Cent\left(i+\sqrt{area}/2, j+\sqrt{area}/2\right)\\ area &= H\cdot W \end{aligned} \tag{6}$$

where, $Cent(i,j)$: centroid of the face.

P_1: start point of the face.
P_2: end point of the face.

Table 1
Experimental result of the face detection.

Genre	Total number of faces	Without compensation		With compensation	
		False	Precision (%)	False	Precision (%)
Drama	1502	94	94.1	56	96.4
Sitcom	1636	98	94.3	49	97.1
Movie	1996	108	94.8	53	97.4
News	350	43	89.1	19	94.9
Actor	177	22	88.9	15	92.1
Total	5661	365	93.9	192	96.7

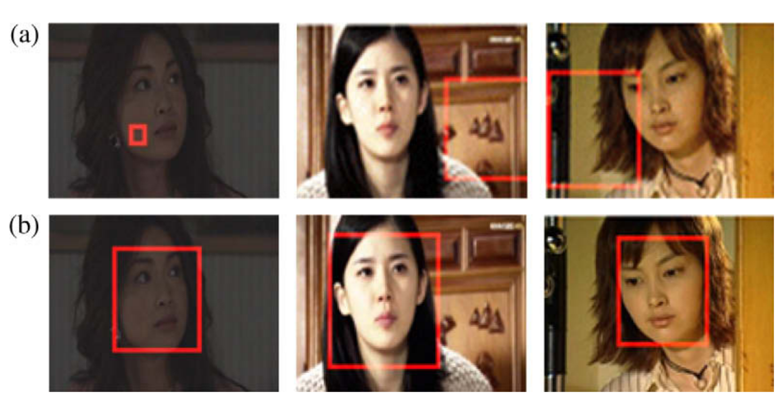

Fig. 9. Face detection examples: (a) without, (b) with illumination-compensation method.

H,W: height and width of the face.
area: area of the face region.
P_1 and P_2 are important points to obtain face line, and the resulting region will be a rectangle size of $H \times W$. The Fig. 8 shows the processes of face region detection.

5. Experimental results

Our approach is applied to detect faces in face images. This approach can be used for video summary. We apply our approach to various genres (news, drama, movie and actor images from web). Also, we make an experiment with same size of images (320×240) to evaluate the performance of the algorithm for face detection.

We test our algorithm with five databases consisting of total 5661 images and get the precision ratio of 96.7% on the average. The precision ratio is obtained by the following equation:

$$precision = F_T/(F_T + F_F) \times 100(\%) \tag{7}$$

where, F_T: total number of faces, F_F: number of false detected faces.

Through this research, we make a more useful and robust face detection system compared with previously work that did not consider about illumination problem. More detailed experimental results are reported in Table 1 and the example of the improve results are shown in Fig. 9 marked with red[1] rectangle. Due to the lighting, the faces in the Fig. 9a images are not detected well. However, these effects are removed by the illumination-compensation method.

In this approach, we check the average value of luminance in the whole image. If it is outside of the threshold range, i.e., the image is too dark or too bright we apply illumination-compensation in addition. It reduces the effect of illumination. So, if we use illumination-compensation for extreme illumination condition, we can increase the precision rate by 2.8%. Table 1 shows the experimental results of the face detection with precision ratios.

6. Conclusion and future works

In this paper, a color-based face detection algorithm is proposed in *YCbCr* color space. We use the illumination-compensation method and closing algorithm to detect the face region robustly. If the average luminance is very high or very low in the image, we apply the illumination-compensation method to reduce the lightening effect. Closing algorithm helps to remove noise from the non-face regions and also defines the face region in the image. Thus, it helps to overcome the difficulty of face detection.

Therefore, our algorithm quickly locates a face region compared with other geometric methods such as detect eye or mouth map. This is because we can find the centroid of certain face region fast and directly. Although our approach is simple and has low complexity, the performance is better than other methods. Also, these advantages can be used for efficient actor-based video summary.

In our future work, we will investigate more useful detection method in various situations such as in multiple face cases. And we need a research about the face recognition method for retrieving after video summary.

References

Gonzalez, R.C., Woods, R.E., Eddins, S.L., 2004. Digital Image Processing Using MATLAB. Prentice Hall.
Hsu, R.L., Mohamed, A.M., 2002. Face detection in color images. IEEE Trans. Pattern Anal. Machine Intell. 24 (5).
Jin, H., Liu, Q., Lu, H., 2004. Face detection using one-class-based support vectors. In: IEEE FGR'04.
Kim, J.S., Lee, S.T., Baek, J.H., 2004. An efficient face region detection for video indexing. In: Proceedings of the First International Computer Engineering Conference, ICENCO. pp. 311–316.
Li, Y., Narayanan, S.S., Kuo, C.C., 2004. Content-based movie analysis and indexing based on audio-visual cues. IEEE Trans. Circ. Syst. Video Technol. 24 (8).
Richardson, Iain E.G., 2003. Video Codec Design. John Willy and Sons, Ltd.
Sung, K.K., Poggio, T., 1998. Example-based learning for view-based human face detection. IEEE Trans. Pattern Anal. Machine Intell. 20 (1).
Terrillon, J.C., Shirazi, M.N., Fukamachi, H., Akamatsu, S., 2000. Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In: IEEE Conf. on Face and Gesture Recognition, pp. 54–61.

[1] For interpretation of colour in Fig. 9, the reader is referred to the web version of this article.