Illumination-Adaptive Face Detection and Facial Feature Extraction

Se-Hwan Kim and Yo-Sung Ho
Kwangju Institute of Science and Technology
1 Oryong-dong Puk-gu, Kwang-ju, 500-712, Korea

ABSTRACT

Under moderate illumination conditions, human face can be easily detected by various algorithms. However, the performance of face detection algorithms deteriorates dramatically when the illumination is too high or too low because the skin color changes depending on the reflected light from the face surface. In this paper, we propose a new algorithm for adaptive human face detection under different illumination conditions. Based on the fact that the saturation component of the face color changes inversely proportional to the luminance component within some restricted range, we model the saturation component by the Cauchy probability density function (pdf). We detect facial features using the geometrical relationship of eyes, nose and mouth.

Keywords: Illumination-Adaptive Face Detection, Feature Extraction, Cauchy pdf, LHS Color Space

1. INTRODUCTION

Since human face is a good feature for personal identification, a lot of studies have been done to detect faces in images. Furthermore, as it is necessary to prove the personal identity using computers, importance of detecting human faces has been increased. Face detection methods are applicable to various areas, such as passport, credit card, driver's license and surveillance.

However, it is difficult to find the skin color to locate the human face in most situations. The problem results from several factors that are influenced by changes of illumination, background and camera. Although several methods have been proposed to eliminate the effect of those changing factors, there is no paradigm that solves this problem completely.

Most of all, illumination variation has been proved to be the most difficult operation to cope with because the pixel intensity values are directly modified by changes of illumination intensity, direction, and color; moreover, in a non-linear way. Although the performance of a face detector can work well in normal illumination conditions, if the illumination decreases or increases to a certain degree, its performance deteriorates drastically. In addition, when the illumination condition changes over a period of time, the variation should also be taken into consideration. Due to these problems, it is not easy to detect human faces satisfactorily in various circumstances. Understanding the role of illumination is, thus, essential for resolving the face detection problem.

A well-known face detection approach is to transform the image from the RGB color space to the HSV color space. Based on the fact that the hue, saturation and value of skin color are confined in some restricted ranges, they can determine the face region [1]. There is, however, a problem of illumination changes because they fixed threshold values.

In other approach, they define the chromatic color space, where the chromatic colors (r,g), known as pure colors in the absence of brightness, are represented by normalization. The skin color distribution is represented by the Gaussian model and the final face region is determined using estimated parameters. This approach also has difficulties in detecting the face region in low illumination conditions [2].

In another approach, the RGB color space is transformed into the YCbCr color space. In the CbCr space excluding the luminance component, an ellipse which includes the face region is examined and this ellipse is used as a threshold for face detection [3]. However, the face detection rate decreases dramatically in low illumination conditions.

In this paper, we propose a new algorithm that can detect the human face adaptively in various illumination conditions. The proposed algorithm consists of two parts. The first part is to detect the skin color region in the LHS color space. We propose an illumination-adaptive skin filter to characterize the skin color distribution of different people under different lighting conditions. We change the threshold value adaptively according to the luminance distribution of the skin color.

The second part is for feature extraction. We threshold the input image to extract the face features, such as eyebrows, eyes, nose and mouth, based on the luminance distribution of the input image. The geometrical constraints of those features are used to develop a face-like structure by combining these candidate features. We employ evaluation functions to judge whether the given set of blocks are appropriate for facial features or not.

2. FACE DETECTION

2.1 PREPROCESSING

As a preprocessing, we use the opening operation, one of the morphological transformations, to remove the illumination effect on the face surface, especially in the high illumination conditions. Although the reflection effect rarely happens in the low illumination conditions, high illumination influences the face surface, resulting in modifying the distribution of pixel intensity values on the face surface. We can reduce this phenomenon by taking the morphological opening operation.

2.2 THE LHS COLOR SPACE

One of the essential issues that one encounters when working with color images is the choice of a color space. In this paper, we choose the LHS color space, which is the same as the HSV color space except the luminance component, because luminance and saturation values have the unique relationship. Strickland et al. and Ledley et al. derived equations by dividing the Maxwell plane into three regions and calculating angles for each region [4].

$$L = 0.299R + 0.587G + 0.114B \quad (1)$$

$$H = \begin{cases} \mathrm{acos}\left\{\dfrac{[(R-G)+(R-B)]/2}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\right\} & (if\ G \geq B) \\ 2\pi - \mathrm{acos}\left\{\dfrac{[(R-G)+(R-B)]/2}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\right\} & (if\ G \leq B) \end{cases} \quad (2)$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)] \quad (3)$$

When the luminance varies ideally, i.e., R, G and B components vary by the same amount, the hue does not change. However, the saturation changes depending on the luminance as follows:

$$\Delta S = \frac{3(3\min(R,G,B) - R - G - B)\sigma}{(R+G+B)(R+G+B+3\sigma)} \quad (4)$$

where σ denotes the difference between luminance values in the ideal case. Consequently, the hue is not affected by the shifting operation. However, the saturation is changed when the luminance is modified by shifting. Since (2R-G-B), (2G-B-R) and (2B-R-G) are all negative, ΔS is negative if σ is positive, or vice versa. Therefore, the saturation decreases when the luminance is increased by shifting the RGB values. Similarly, the saturation increases when the luminance is decreased by shifting.

2.3 THRESHOLD DECESION IN THE LS SPACE

We examine detection regions for the original face images. By detection regions, we mean threshold values that are determined experimentally based on the subjective decision. If we select the detection regions under nine illumination conditions for one person, we can obtain Figure 1. We can also get similar results for other people through the same process.

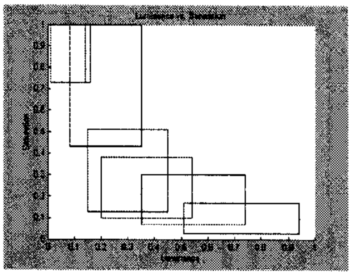

Figure 1. Detection regions based on subjective decision

As shown in Figure 1, detection regions reveal a relationship between luminance and saturation. In order to derive their relationship, we set two points in each rectangle in Figure 1. One point is located at the upper-right side, distant from the right side boundary by one-sixth of the horizontal length and from the upper side boundary by one-sixth of the vertical length, respectively. The other point is located at the exactly opposite side of the first point with respect to the center point in each rectangle. These two points in each rectangle are used to delineate the upper bound and the lower bound of the saturation depending on the luminance. In Figure 2, '*' denotes the upper-right position and '+' denotes the lower-left position.

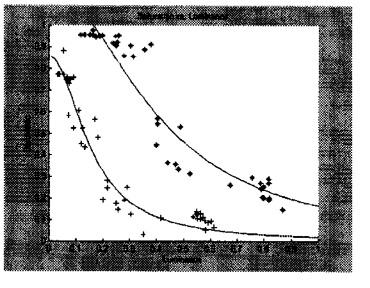

Figure 2. Upper bound and lower bound

We use the Cauchy pdf model to represent the distribution shown in Figure 2 because of the two properties of the Cauchy function.

(a) The Cauchy function has a very steep curve in the middle part of the range from zero to one.

(b) The Cauchy function has a saturation curve near both ends of the range. This means that even though the luminance decreases continuously, the saturation does not increase any more near zero. A similar phenomenon happens near one.

We apply a Levenberg-Marquardt algorithm, a well-known nonlinear regression method, to fit the bounding curves to the Cauchy functions. Resulting equations for the upper and lower bounds are as follows [5]:

Upper bound $$f_X(x)\_upper = \frac{0.601/\pi}{x^2 + 0.405^2} \quad (5)$$

Lower bound $$f_X(x)\_lower = \frac{0.066/\pi}{x^2 + 0.156^2} \quad (6)$$

Figure 2 shows the two corresponding curves.

2.4 FACE REGION DETECTION

In the face detection stage, we make use of the upper and lower bounds based on the Caucy pdf model and the hue component. We judge each pixel as a face region only if its saturation, which is corresponding value to the luminance, stays between the upper and lower bounds.

There is another component we should consider for face region detection. Being different from the saturation and luminance components, the hue component is not very sensitive to illumination variations. Thus, we set the hue range from -36° to 36° that is sufficient to meet the requirement for face detection. Therefore, the final face region is determined by examining the upper and lower bounds in the LS space and the hue component of each pixel. We select the intersection area of these three components as the face region.

2.5 REFINEMENT AND VERIFICATION

The extracted skin color pixels, which are represented by a binary image, usually contain noises that either appear as small holes due to the undetected facial features or appear as objects of the skin-like color in the background. In order to remove such noise points and obtain connected skin clusters, we apply a 5×5 median filter. Furthermore, we count pixels of each cluster, and small clusters whose total pixels is less than a certain ratio of the total image pixels are discarded.

As the final step of face detection, the face region is verified using the shape information of the face. At this stage, the oval shape of the face is approximated by an ellipse. We access how well the connected component C is approximated by its best-fit ellipse E. For that purpose, the following measure V is evaluated:

$$V = \frac{\sum_{(x,y)\in E}(1-b(x,y)) + \sum_{(x,y)\in(C-E)} b(x,y)}{\sum_{(x,y)\in E} 1} \quad (7)$$

where

$$b(x,y) = \begin{cases} 1 & if (x,y)\in C \\ 0 & otherwise \end{cases} \quad (8)$$

denoting the indicator function of C. V determines the distance between the connected component and the best-fit ellipse by counting the holes inside of the ellipse, and the points of the connected component which are outside of the ellipse. The ratio V of the number of false points to the number of points of the interior of the ellipse is calculated. Based on the threshold on this ratio, connected components that are well approximated by their best-fit ellipse are selected as face candidates.

Based on the computed elliptical parameters, we can reduce the number of face candidates. This is done by applying to each ellipse decision criteria concerning its orientation and the relationship between the major and minor axes. Subsequently, the selected face candidates are verified by searching for facial features inside of the connected components.

3. FACIAL FEATURE EXTRACTION

3.1 VALLEY AREA EXTRACTION

Figure 3 depicts the histogram distribution of the luminance values within a rectangle, which includes most of the face region for all nine illumination levels. Each curve represents the histogram distribution for one image. It seems that there are only six histogram distributions even though nine histogram distributions are displayed. This fact reflects the saturation phenomena at both ends.

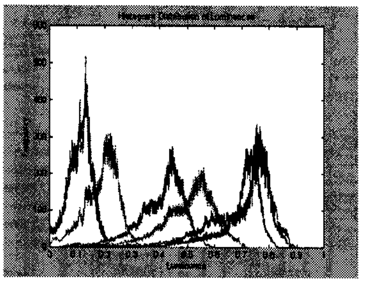

Figure 3. Histogram distribution of luminance

When an image is captured, the average luminance value of the image is proportional to the amount of illumination at that time. Thus, the luminance in the middle range reflects the average illumination. However, even if the amount of illumination diminishes continuously near zero, the luminance value does not decrease any more. The same phenomenon happens near one.

We can characterize each curve of the histogram distribution into three parts. The first one is the peak point around the center of each curve. The second one is the steep slope, and the final one is the gentle slope at the left side of the center. Indeed, the gentle slope indicates regions that exclude the face region within the centered rectangle. In other words, this region indicates eyebrows, eyes, nose, mouth and shadow which is created by facial features.

Figure 3 reflects implicitly that the luminance of facial features is also dependent on illumination. Thus, in the valley area extraction, this should be taken into consideration. The valley has the following properties.

(c) The valley has lower intensity than other neighboring areas.
(d) The valley can be found by evaluating the difference in the intensity value.
(e) The valley exists within the facial region.

Therefore, we determine the threshold value for valley area extraction as the value subtracted from the mean value by 1.75 times of the variance.

Although we have found the valley areas corresponding to the features, the feature blocks belonging to the same facial feature are still separated from each other. Therefore, we need to group these separate blocks by a feature grouping algorithm that is processed with the following properties. The merging radius depends on the size of a feature block under consideration. In other words, the larger the block size is, the smaller the merging radius is. Two feature blocks to be merged must lie on the same axis if one or both of the block sizes is larger than the maximal block size [6].

3.2 FEATURE MATCHING

After the labeling and grouping process, all feature blocks are regarded as facial feature candidates. The line passing through the centers of both eyes is called the base line. D_{eye} denotes the distance between two eyes. $D_{eyebrow}$ represents the distance from the center of eyebrow to the base line. D_{nose} and D_{mouth} denote the distances from the centers of nose and mouth to those of eyes, respectively. A_{nose} and A_{mouth} denote the angles between the baseline and the lines passing through the centers of eyes and nose, mouth, respectively.

The distances $D_{eyebrow}$, D_{nose} and D_{mouth} in the geometrical face model are estimated as $0.30D_{eye}$, $0.78\ D_{eye}$ and $1.12\ D_{eye}$, respectively. The angles A_{nose} and A_{mouth} are estimated as $63.2°$ and $49.1°$, respectively. Feature matching is based on the properties of each block, such as the center point, the number of pixels, the orientation of each block, and the lengths of semimajor and semiminor axes. We adapt the following evaluation functions [7]:

$$E_{Eye} = \exp[-1.2\times\{(l_1-l_2)^2+(l_1+l_2-1)^2+(\theta_1-\theta)^2+(\theta_1-\theta)^2\}] \quad (9)$$

$$E_{Reb} = \exp\left[-4.8\times\left(\frac{d_{Reb}-D_{eyebrow}}{D_{eye}}\right)^2\right] \quad (10)$$

$$E_{Mouth} = \exp[-1.2\times\{(rd_{mouth}-1)^2+(ld_{mouth}-1)^2 + (ra_{mouth}-1)^2+(la_{mouth}-1)^2\}] \quad (11)$$

$$E_{Nose} = \exp[-1.2\times\{(rd_{nose}-1)^2+(ld_{nose}-1)^2 + (ra_{nose}-1)^2+(la_{nose}-1)^2\}] \quad (12)$$

$$E = 0.5E_{Eye} + 0.2E_{Mouth} + 0.1E_{Reb} + 0.1E_{Leb} + 0.1E_{Nose} \quad (13)$$

where rd_{mouth} and ld_{mouth} are lengths between a feature block and the left, right eyes divided by D_{mouth}. ra_{mouth} and la_{mouth} are angles between a feature block and the left, right eyes with respect to each eye divided by A_{mouth}. The similar process is also applied to nose parameters. Finally, we use the overall evaluation function.

4. EXPERIMENTAL RESULTS

Figure 4 shows the detection procedure. Figure 4(a) is the original image. Using the proposed method in the LHS color space, we obtain the H-thresholded image and the LS-thresholded image. If we take the intersection area, we can find the face region as shown in Figure 4(d). However, Figure 4(d) contains the noise components due to the undetected facial feature region and the face-like color region in the background. To remove these noise components, we apply the refinement and verification processes and obtain Figure 4(e). The final detected face region is shown in Figure 4(f). Using the variable valley area threshold, we can obtain Figure 4(g). Based on the geometrical relationship, we can find the facial features, as shown in Figure 4(h).

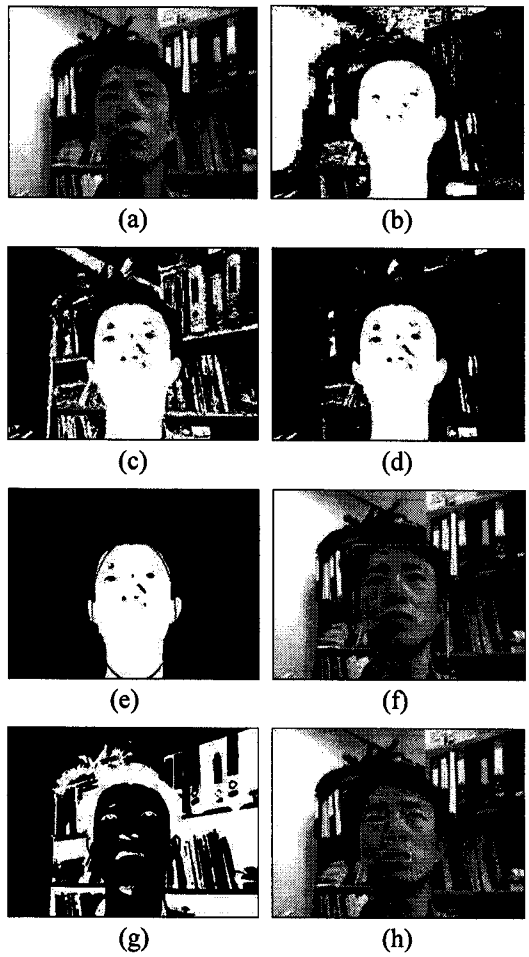

Figure 4. Face detection procedure
(a) Original image (b) H-thresholded image
(c) LS-thresholded image (d) Intersection area
(e) Refined and verified region (f) Detected face region
(g) Valley areas (h) Final result

Figure 5 shows experimental images for one person under three different illumination conditions. Each column represents the same illumination condition and each row represents the results obtained by the same algorithm. Figure 5 demonstrates that the proposed algorithm has better performance than the others, especially in the low illumination condition. In Figure 5, we cannot distinguish any significant differences between the results by the proposed algorithm and other algorithms in the moderate and high illumination conditions. However, the proposed algorithm detects the face region more accurately compared to other algorithms in the low illumination condition.

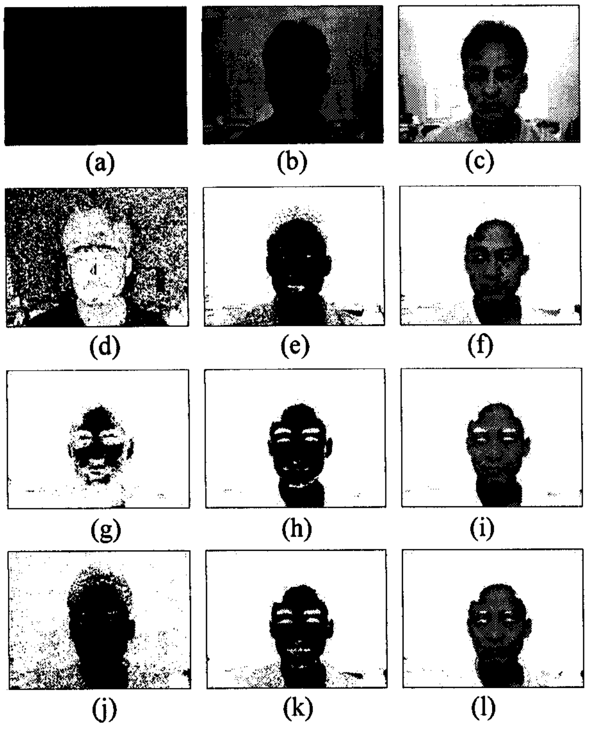

Figure 5. Performance comparison of face detection under three illumination conditions
(a-c) Original images (d-f) rgb based method
(g-i) YCbCr color based method (j-l) Proposed method

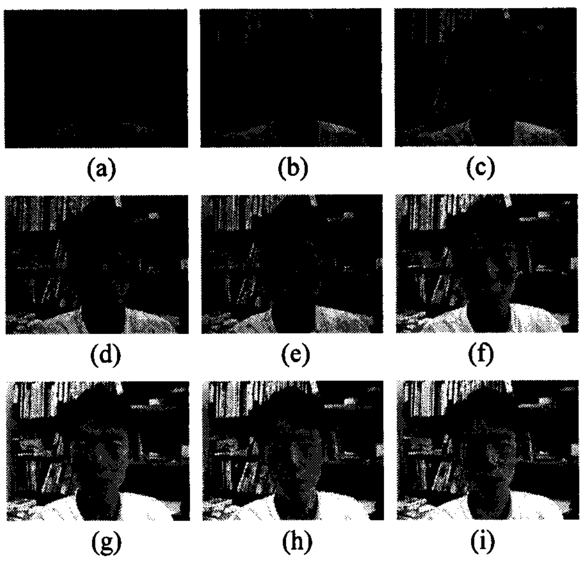

Figure 6. Feature extraction

Figure 6 shows simulation results under nine different illumination conditions. Each figure represents extracted features in different illumination conditions. Figure 6(a) and Figure 6(i) show the results in the most dark and the most bright illumination conditions, respectively. Although a few features have been missed, the simulation results demonstrate this algorithm can detect the face region and facial features under several illumination conditions.

5. CONCLUSIONS

In this paper, we have developed a new approach for adaptive human face detection by removing the effect of illumination changes. Using the face that the luminance is inversely proportional to the saturation, we can detect a human face in the LHS color space under varying illumination conditions. We use the Cauchy pdf model to preserve the relationship between these two color components. We have also exploited the geometrical relationship of the facial features for the feature matching process. Because the geometrical relationship is almost the same for most people, the distances and angles of features are used as evaluation functions. Experimental results demonstrate the effectiveness of the proposed algorithm in detecting human face and facial features under various illumination conditions.

ACKNOWLEDGMENTS

This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optic Networks (UFON) Research Center at Kwangju Institute of Science and Technology (K-JIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

REFERENCES

[1] K. Sobottka and I. Pitas, "Extraction of Facial Regions and Features Using Color and Shape Information," *IEEE International Conference on Pattern Recognition*, pp. 421-425, Aug. 1996.

[2] J. Yang and A. Waibel, "A Real-Time Face Tracker", *IEEE Workshop on Applications of Computer Vision*, pp. 142-147, 1996.

[3] K. H. Kim, S. H. Kim, O. M. Choi and B. H. Lee, "Efficient and Automatic Face Detection Using Skin-tone and Shape", *The Institute of Electronics of Korea Conference*, pp. 575-578, 1999.

[4] C. C. Yang and J. J. Rodriguez, "Efficient Luminance and Saturation Processing Techniques for Color Images", *Journal of Visual Communication and Image Presentation*, pp. 263-277, Sept. 1997.

[5] S. H. Kim and Y. S. Ho, "Illumination-Adaptive Face Detection and Facial Feature Extraction in Color Images", *Korean Signal Processing Conference*, pp. 1097-1100. 1999.

[6] S. H. Jeng, H. Y. Mark Liao, Y. T. Liu, and M. Y. Chern, "An Efficient Approach for Facial Feature Detection using Geometrical Face Model", *IEEE International Conference on Pattern Recognition*, pp. 426-430, 1996.

[7] C. D. Yeom and D. S. Jeong, "Face Detection Using The Geometrical Relationships Between Facial Components", *Korean Signal Processing Conference*, vol. 12, no. 1, pp. 243-246. 1999.