http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (1): 183-187, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.183.187

A Fast Face Detection Scheme Utilizing HSV-based Skin Color Model with K-L Transform

[1]Zheng-Liang Huang, [1]Zhe-Ming Lu,[1]Fa-Xin Yu and [2]Yong Zhang
[1]School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China
[2]ATR National Defense Technology Key Laboratory, College of Information Engineering, Shenzhen University, Shenzhen, 518060, China

Abstract: Fast face detection is very essential for many applications such as video surveillance and human computer interface. This study presents a fast face detection method using the HSV-based skin color model under non-constrained scene conditions. First, a skin color model that combines the HSV color space with K-L transform-based skin color filter is designed. Then, skin regions in the input image are detected and face candidates based on the spatial arrangement of these skin patches are generated. Finally, the mouths in all of the candidate human face regions are verified by a mouth detector. Experimental results show that the proposed scheme is not only fast and efficient, but also robust to head rotation.

Key words: Face detection, HSV color space, skin color model, K-L transform

INTRODUCTION

Face detection is a useful technology that has been successfully applied in various fields such as video surveillance, human computer interface and biometric. Because of complicated background, various view angles, illumination conditions and facial expressions, automatic human face detection from digital images is a hard and challenging task. In fact, face detection is a two-class (face and non-face) classification task, thus many techniques that were originally designed for face recognition have also been used to detect faces. However, the complexity of these techniques is high and thus they are not suitable for real-time face detection, especially for images with large variations.

Over the past two decades, a number of face detection schemes have been presented (Yang *et al*., 2002; Hsu *et al*., 2002; Pattanasethanon and Savithi, 2012). In brief, face detection techniques can be roughly classified into three categories, i.e., neural network based, feature-based and color-based schemes. Neural network based schemes (Le, 2011) train the neural network based on a number of training data including facial and non-facial data, then the trained neural network can distinguish faces from non-facial parts. Feature-based algorithms (Chiang *et al*., 2003) exploit extracted features to detect faces, such as the shape of the face and the locations of eyes, nose and mouth. Color-based techniques (Zou and Kamata, 2010) adopt skin color models to detect faces. As one of the most significant physical surface features, the skin color features embody more information than gray-scale features, which are suitable for real-time applications. Although, the skin colors differ from person to person and race to race, they can be closely clustered in color spaces, because the difference in the color composition of human skin is small (Hjelmasa and Low, 2001). However, since the color of facial pixels is sensitive to different illumination conditions (Zhang *et al*., 2008), it is hard to achieve stable detection performance. Therefore, skin color based face detection schemes have limitations in practice. Thus, many researchers have been focusing on highly robust detection schemes that are insensitive to illumination changes (Tian and Zhuang, 2004; Yong *et al*., 2004; Lei *et al*., 2002). However, all these methods are very time-consuming and they are hard to be used in real-time applications efficiently. For many applications such as human-machine interaction and video surveillance, the most urgent property is real-time detection and tracking of objects. Because the distribution of human skin-color is irrelative to the scale and rotation of the object, the skin-color modeling-based face detection scheme has become a popular method in rough location of human faces nowadays (Yang *et al*., 2002).

To realize fast face detection under complex conditions, this study presents a face detection scheme that is able to deal with a wide range of variations under non-constrained scene conditions in color images. The main idea of the proposed scheme is to model skin colors by combining the skin color model in the HSV color space with the K-L transform based skin color filter and then extract the mouth location features to verify each face

Corresponding Author: Zhe-Ming Lu, School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China

candidate. The rest of this study is organized as follows. First, some related works on skin color models are introduced. Next, the proposed face detection scheme is introduced. Then, some experiments are tested to evaluate the effectiveness of the proposed scheme. Finally, conclusions are drawn.

RELATED WORK

Skin color modeling based on RGB space: Skin color models are mathematical models utilized to characterize the distribution of skin colors. Before generating skin color models, it is essential to select a proper color space. There are many color spaces. Some are proposed for color coding, such as RGB, LUV, LAB, XYZ and YUV. Some are designed for computer graphics, such as YIQ, HSV and HIS.

Since, many color spaces can be directly transformed from the RGB space, it is necessary to analyze the skin color distribution in the RGB space. In general, the three components of a skin pixel satisfy the inequality R>G>B. Detailed statistical analysis on a large number of skin pixels based on the RGB space has been made in (Tian and Zhuang, 2004). In that study, 107 non-overlapping regions containing 3,420,408 skin pixels are segmented from 94 photos taken under standard illuminations for Europeans, Asians and Native Americans. The statistical results show that, for skin pixels, the mean values of the R, G, B components and lightness are 235.3, 195.9, 176.6 and 205.5, respectively, while the standard deviation values of the R, G, B components and lightness are 15.7, 26.2, 33.2 and 22.9, respectively. Therefore, for skin pixels, the distribution of their RGB components complies with certain inherent laws and restraint relationships.

However, the RGB components cannot characterize the chromatic features perceived by the human visual system, such as lightness, hue and saturation. It is proved that the RGB color space is unsuitable for constructing an effective skin color model due to the high correlation among the three components. To improve the detection performance of skin color models, this study adopts other color space that has a nonlinear relationship with the RGB space.

Skin-color modeling based on YCbCr space: Zhang *et al.* (2008) proposed the skin color model based on the YCbCr space. Since, skin color information often moves to some directions from its original color due to the light source or acquisition equipment, the input image of size M×N should be processed as follows to avoid color deviation:

Step 1: Calculate the brightness information of each pixel:

$$k(i,j) = 0.299r(i,j) + 0.587g(i,j) + 0.114b(i,j) \quad (1)$$

Step 2: Sort k (i,j) in the ascending order to get a sequence $k_1, k_2, \ldots, k_{M\times N}$

Step 3: Compute the average value k_a over the first five percent of pixels in the obtained sequence as follows:

$$k_a = \sum_{i=1}^{M\times N\times 0.05} k_i \quad (2)$$

Step 4: Set the average value as the reference white and adjust all pixels with this brightness value to R = 255, G = 255 and B = 255, while adjusting the remained pixels based on the proportion $255/k_a$

After above steps, the new Cb and Cr values can be calculated for each pixel. Then, the contour of the skin color model based on the YCbCr space can be described using the following circle:

$$\frac{(x-1.60)^2}{25.39^2} + \frac{(y-2.41)^2}{14.03^2} = 1 \quad (3)$$

where, x and y are defined by the following transform:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos 2.53 & \sin 2.53 \\ -\sin 2.53 & \cos 2.53 \end{bmatrix} \begin{bmatrix} C_b - 109.8 \\ C_r - 152.02 \end{bmatrix} \quad (4)$$

Based on above two equations, whether a pixel belongs to the skin color space or not can be judged based on the obtained values x and y. If the point (x, y) is within the circle, then the pixel falls in the skin color space. Otherwise, it is not.

PROPOSED FACE DETECTION SCHEME

Proposed skin color model based on HSV space: The first stage of the proposed scheme is to locate the potential facial areas in the image based on skin chrominance information, given that such information can strongly reduce the search space. It can be found that the human skin has the following characteristics (Chang *et al.*, 1994): (1) Skin colors are clustered only in a small area in the color space; (2) The skin color appearances differ more in brightness than in chroma. That is, the difference between skin colors is mainly caused by the difference between their intensities.

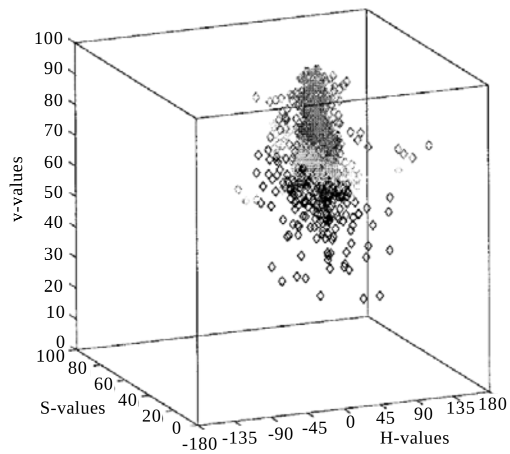

Fig. 1: Distribution of skin color samples in the HSV space (Garcia and Tziritas, 1999)

As discussed in the previous section, the RGB color space is not suitable to quantize the object's color feature, because R, G and B values are correlative to each other. Thus, this study considers another suitable space named HSV. The HSV (Hue, Saturation, Value) space is adopted mainly in computer graphics and is considered by many researchers due to its ability of intuitively using and it is also closed to the manner how an artist actually mixes colors. A lot of skin color patches have been used in order to approximate the skin color subspace in the HSV space. In Fig. 1, the skin color distribution is illustrated in the HSV space (Garcia and Tziritas, 1999).

In the HSV space, the projection onto the HS plane is usually exploited to classify skin colors, which is performed by setting appropriate thresholds to Hue and Saturation components, respectively. Based on these thresholds, the segmentation results are mainly affected by variations in lighting conditions. Thus, in Reference (Garcia and Tziritas, 1999), the shape of the skin color space subspace is directly estimated from the HSV color space. In Reference (Garcia and Tziritas, 1999), a set of planes have been estimated by successive adjustments based on segmentation results and six equations were reported to define the bounding planes in the HSV color space. In the experiments, it is found that the segmentation results are quite equivalent between using V-constrained Hue bounding plane and setting appropriate thresholds to Hue. However, the latter is simpler. In this study, an improved skin-color modeling algorithm is given as follows:

$$\begin{cases} S \geq 10 \\ V \geq 40 \\ S \leq -H - 0.1V + 110 \\ S \geq 0.08(100 - V)H + 0.5V \\ 2 \leq H \leq 43 \end{cases} \tag{5}$$

However, it is impossible for every skin color model to remove all the non-skin color pixels while clustering most of the skin color pixels, this study uses another skin color filter, i.e., the K-L skin color filter, which performs the K-L transform on color coordinates and then establishes a K-L based skin color coordinate system. The K-L transform is one of the best orthogonal transforms used in image compression, which has also been successfully used in the face recognition field. It is reported that the K-L transform has preferable clustering characteristic. This study refers to the K-L skin-color coordinate as shown by Zhou *et al.* (2001). After the K-L transform, the following transform equation can be obtained:

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} = \begin{bmatrix} 0.666 & 0.547 & 0.507 \\ -0.709 & 0.255 & 0.657 \\ 0.230 & -0.797 & 0.558 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6}$$

Due to different bases of the two models, they rarely agree on a false detection of skin pixel. So this study combines (1) and (2), which can eliminate most false detections. Geometry constraints are also considered to reduce the searching space. Regions that do not contain the face structure are removed.

Face detection based on feature matching: Here, the face detection approach is performed in a coarse-to-fine manner. Next this study converts the color image to a binary image and then segments potential face regions. Due to noises and distortions, the binary image may contain non-contiguous skin-color regions and the edge of the binary image may be unclear. In order to improve the detection quality, this study employs non-linear filters to remove noises. The erosion and dilation operators are utilized to remove small regions and fill holes. After region segmentation and merging, all face candidate regions are found. It is obvious that not all detected candidate regions are face regions. Obviously, classification based only on color information is not sufficient to guarantee the perfect segmentation. In order to remove these false face regions, a verification step is needed.

This study chooses the mouth as a feature for verification and only considers the area covered by a face mask (Hsu *et al.*, 2002) that is built by enclosing the grouped skin-tone regions with a pseudo convex hull. The color of mouth region contains stronger red components and weaker blue and green components than other facial regions. Considering the approaches by Yao *et al.* (2000) and Niu *et al.* (2003), this study constructs the mouth mapping as follows:

$$U = -0.147 * R - 0.289 * G + 0.436 * B \tag{7}$$

$$V = 0.615 * R - 0.515 * G - 0.1 * B \quad (8)$$

$$\varphi = tg^{-1}(\frac{V}{U}) \quad (9)$$

Then, the detection model is obtained as follows:

$$\begin{cases} \varphi \in [75^\circ, 102^\circ] \\ U \in [-10, 8] \\ V \in [15, 32] \end{cases} \quad (10)$$

In general, the localization of the upper boundary of the face candidate is very precise. However, the localization of the lower boundary is not so precise. So this study constrains the lower boundary by the mouth localization. The candidate region is resized to keep the upper boundary fixed and let its width less than 1.5 times that of height. In this way, the unwanted neck region and other false regions can be cut.

EXPERIMENTAL RESULTS

In order to demonstrate the effectiveness of the proposed approach, this study carried out extensive experiments. These experiments are performed on a PC with 2 GHz CPU and 1GB RAM using VC+6.0. The test images are various sized 24bits BMP images. First, the effectiveness of the proposed skin color model is tested as shown in Fig. 2. From Fig. 2, it can be seen that the proposed skin color model can correctly obtain the skin colored regions.

Secondly, the effectiveness of the whole face detection scheme is tested. From the experiments, this study finds that, for a 250×160 sized 24 bits image, the average CPU time of the proposed face detection scheme for a still image is about 0.1 second. In addition, the proposed algorithm can detect multiple faces with variations in color, position, scale, orientation and facial expression. Furthermore, this study compared the performance of the proposed system with the face detection system developed by the Fraunhofer Institute, Germany, which uses the contours of the face, eyes, eyebrows and nose for face detection. As expected, their system is sensitive to the pose. In contrast, the proposed system is able to locate faces correctly in any pose and orientation in real time. Fig. 3 illustrates some results. The original images in Fig. 3 are taken from the Internet. Notice from these results that the system has ability to deal with the cases such as the head rotations, these images containing frontal, near-frontal and profile face views of different sizes.

Fig. 2(a-b): (a) Original images and (b) Skin color filtered images

Fig. 3(a-g): Some typical face detection results of the proposed scheme

CONCLUSION

Face detection is a challenging task in real-life application. This study has proposed a scheme for human face detection in color images under non-constrained scene conditions, such as the existence of various illumination and complicated backgrounds. The proposed method checks skin regions over the whole input image and then obtains face candidates according to the spatial arrangement of skin patches. The proposed method utilizes the mouth maps to recognize the face candidates. The main purpose of the proposed scheme is to design a system which can distinguish the faces from non-faces in real time, thus it can be used for face identification and retrieval from image and video databases.

ACKNOWLEDGMENTS

This study is supported by the National Natural Science Foundation of China under grant 61171150 and 61070208 and also supported by Zhejiang Provincial Natural Science Foundation of China under grant R1110006 and Special Fund on Strategic New Industry Development of Shenzhen under grant JCYJ20120817163934173.

REFERENCES

Chang, T.C., T.S. Huang and C. Novak, 1994. Facial feature extraction from color image. Proceedings of the 12th International Conference on Pattern Recognition, Volume 2, October 9-13, 1994, Jerusalem, pp: 39-43.

Chiang, C.C., W.K. Tai, M.T. Yang, Y.T. Huang and C.J. Huang, 2003. A novel method for detecting lips, eyes and faces in real time. Real-Time Imaging, 9: 277-287.

Garcia, C. and G. Tziritas, 1999. Face detection using quantized skin color regions merging and wavelet packet analysis. IEEE Trans. Multimedia, 1: 264-277.

Hjelmasa, E. and B.K. Low, 2001. Face detection: A survey. Comput. Vision Image Understand., 83: 236-274.

Hsu, R.L., M. Abdel-Mottaleb and A.K. Jain, 2002. Face detection in color images. IEEE Trans. Pattern Anal. Mach. Intell., 24: 696-706.

Le, T.H., 2011. Applying artificial neural networks for face recognition. Adv. Artif. Neural Syst., Vol. 2011. 10.1155/2011/673016

Lei, M., Y.J. Zhang and J.Y. Huo, 2002. A skin color detecting algorithm under variant light condition. Comput. Eng. Appl., 38: 123-125.

Niu, D.J., Y. Zhan and S.L. Song, 2003. Research and implementation of real-time face detection, tracking and protection. Proceedings of the International Conference on Machine Learning and Cybernetics, Volume 5, November 2-5, 2003, Zhenjiang, China, pp: 2765-2770.

Pattanasethanon, P. and C. Savithi, 2012. Human face detection and recognition using Web-Cam. J. Comput. Sci., 8: 1585-1593.

Tian, W. and Z.Q. Zhuang, 2004. Self-adaptive skin color detection based on HSV color space. Comput. Eng. Appl., 40: 81-85.

Yang, M.H., D.J. Kreigman and N. Ahuja, 2002. Detecting faces in images: A survey. IEEE Trans. Pattern Anal. Mach. Intell., 24: 34-58.

Yao, H.X., M.B. Liu, W. Gao, X.T. Fan, H.M. Zhang and Y.J. Lu, 2000. Method of face locating and tracking based on chromatic coordinates transformation of color images. Chin. J. Comput., 23: 158-165.

Yong, Y.P., M. Li and B.Z. Yuan, 2004. Face detection in varying light condition. Signal Proc., 20: 101-107.

Zhang, D., B. Wu, J. Sun and Q. Liao, 2008. A face detection method based on skin color model. Proceedings of the 11th Joint Conference on Information Sciences, December 15-20, 2008, Shenzhen, China, pp: 1-5.

Zhou, Z., Y. Han and Y. Gao, 2001. First face location based on KL skin-color. Comput. Eng., 27: 109-111.

Zou, L. and S. Kamata, 2010. Face detection in color images based on skin color models. Proceedings of the IEEE 10th Region Conference on TENCON, November 21-24, 2010, Fukuoka, pp: 681-686.