The 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON)
Nov. 5-8, 2007, Taipei, Taiwan

Simple Face-detection Algorithm Based on Minimum Facial Features

Yao-Jiunn Chen, Yen-Chun Lin
Mechanical & Systems Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan, ROC.,
GaryJCHEN@itri.org.tw

***Abstract*—This article presents a algorithm for rapid and accurate face-detection. The algorithm detects human faces by the geometric correlations between locations of faces and hairs. Ranges of skin color are used to figure out possible face regions so as to initially localize the faces; furthermore, probable hair blocks in an image are determined by means of hair color spectrums. Grouped skin and hair blocks decide candidate face areas in light of the geometric relation. The accuracy of the single-face detection is higher than 92% with the frame-rate of 10 fps, if tested with a simple background and sufficient light source. The algorithm can be transferred from a PC to embedded devices, such as a DSP platform. This system-level implementing ability brings great potential to customized and reusable applications as well as miniature systems.**

I. Introduction

Applications based on human face detection have been significantly developed recently—surveillance systems, digital monitoring, intelligent robots, notebook, PC cameras, digital cameras, 3G cell phones, and the like. These applications consequently play an important role in our daily life. Nevertheless, the algorithms of the applications are quite complicated and hard to meet real-time requirements of specific frame-rate.

Over the past decade, many approaches for improving the performance of human face detection have been proposed, which are categorized into to main types (1) Knowledge-based method: This method is aimed at finding invariant features of a face within a complex environment, thereby localizing the position of the face. Relationships among the features helpfully determine whether a human face appears in an image or not [1]. (2)Feature invariant approaches: Invariant features, unresponsive to different positions, brightness, and viewpoints, are utilized in this approach to detect human faces. A statistical model is usually built up for describing the relations among face features and the presence of the detected faces. Such face features, for instance, are Facial Features [2], Texture [3], and Skin Color [4]. (3)Template matching method: A template cohering with human face features is used to perform a pattern-matching operation based on the template and an input image. Shape template [5] and Active Shape Model [6] are common examples of this method. (4)Appearance-based method: This method, such as Eigen face [7], Neural Network [8], and Hidden Markov Model [9], employs a series of face images to train and establish a face model for the face detection. In general, method (2), (3), and (4) are more complex than method (1); yet the more features are used in method (1), the more complicated it is.

In this paper, a complexity-reduced algorithm for detecting human faces in real-time is proposed. This algorithm, on the strength of Knowledge-based method with minimum features, adopts the geometric characteristics of skin and hair color to detect human faces. In addition, this algorithm is able to be expectedly transplanted to an embedded system, like the developing pet robot so as to perform dynamic face detection and tracking.

II. Theory

Color information helps extract features of skin color and hair from a complex background. The flowchart of the proposed algorithm is shown as Figure 1.

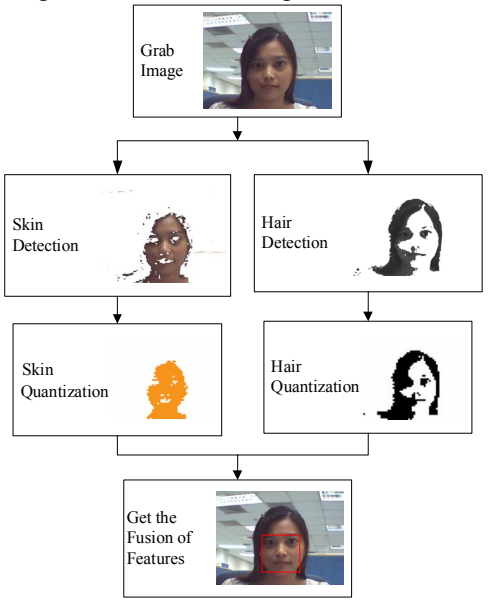

Fig. 1 An Overview of Face Detection Algorithm

Figure 1 includes five main modules--(A)Skin Detection: Using color information to detect possible skin color in an image (B)Hair Detection: Utilizing brightness information to find out where hair probably is (C)Skin Quantization: Quantizing skin color pixels and identifying blocks of the skin (D)Hair Quantization: Quantizing hair color pixels and identifying blocks of the hair (E) Get the Fusion of Features: Determining whether the detected skin area is a part of a human face according to the relative positions between skin and hair regions.

Detailed descriptions of each module are as follows:

A. Skin Detection

Many studies have been involved in defining the range of skin color in an image. Extracting skin color from the Normalized RGB color model [10] was found to be effective since the RGB model without normalization was sensitive to variations of light. The RGB model was therefore transformed to the Normalized RGB model. The formulas for the transformation are listed as equation 1 and 2:

$$r = \frac{R}{R+G+B} \tag{1}$$

$$g = \frac{G}{R+G+B} \tag{2}$$

Equation 1 represents the normalization of a red pixel while equation 2 stands for that of a green pixel. The distribution of skin color is observably concentrated in the correlation chart of r with respect to g, as indicated in Figure 2[2].

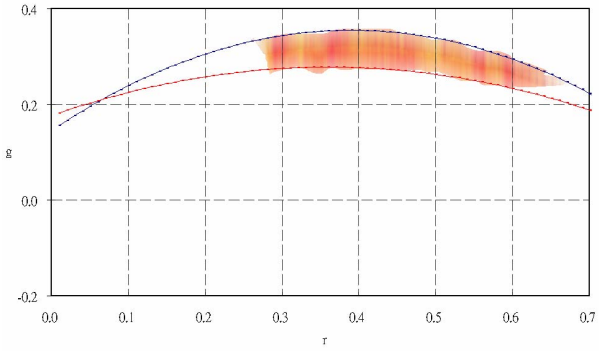

Fig. 2 Statistical chart of skin color

The two equations stated above accordingly specify the upper limit $F_1(r)$ and lower limit $F_2(r)$ of the skin color (through analyses of experiments) [11]:

$$F_1(r) = -1.376r^2 + 1.0743r + 0.2 \tag{3}$$

$$F_2(r) = -0.776r^2 + 0.5601r + 0.18 \tag{4}$$

White color ($r = 0.33$ and $g = 0.33$) is also in the defined range, so the following condition is added for excluding the white color:

$$w = (r - 0.33)^2 + (g - 0.33)^2 > 0.001 \tag{5}$$

Taken together, the skin color range is specified as follows.

$$Skin = \begin{cases} 1 & if\left(g < F_1(r) \cap g > F_2(r) \cap w > 0.001\right) \\ 0 & otherwise \end{cases} \tag{6}$$

Figure 3 shows the original image and Figure 4 demonstrates the result of skin color detection processed using equation 6. Most skin color pixels are successfully obtained while non-skin color pixels are replaced by black ones. Nonetheless, pixels of blue, yellow, and few of gray are extracted as well.

Fig. 3 Original image

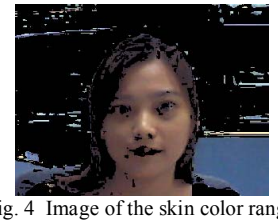

Fig. 4 Image of the skin color range

Another hues element of the robust HSI color model [12] [13] is adopted as a new criterion for estimating skin color pixels so as to improve the above deficiency. The following illustrates the relation between the RGB color model and the HSI color model.

$$\theta = \cos^{-1}\left\{\frac{0.5\left[(R-G)+(R-B)\right]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}$$

$$\begin{cases} H = \theta & if \quad B \le G \\ H = 360^\circ - \theta & if \quad B > G \end{cases} \tag{7}$$

We therefore modified the conditions for skin detection after experiments, as equation 8 shows.

$$Skin = \begin{cases} 1 & if\left(\begin{array}{l} g < F_1(r) \cap g > F_2(r) \cap w > 0.001 \cap \\ (H > 240 \cup H \le 20) \end{array}\right) \\ 0 & otherwise \end{cases} \tag{8}$$

Figure 5 illustrates the image dealt with by equation 8. The result in Figure 5 is apparently improved, as observed from a comparison with the one in Figure 4.

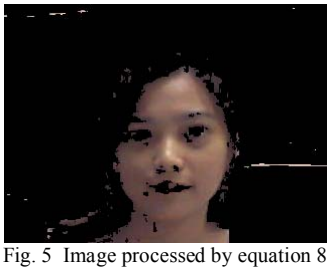

Fig. 5 Image processed by equation 8

B. Hair Detection

The intensity element of HSI color model is employed for detecting the hair color. The relation between the intensity element and RGB elements is as follows:

$$I = \frac{1}{3}(R + G + B) \tag{9}$$

This intensity element is available for evaluating the brightness of an image; moreover, the hair color is specified to be in the range of dark. Equation 10 defines the intensity range of the hair color based on experiments.

$$Hair = \begin{cases} & (I < 80 \cap (B - G < 15 \cup B - R < 15)) \\ 1 \quad if & \cup \\ & (20 < H \le 40) \\ 0 \quad otherwise & \end{cases} \tag{10}$$

Figure 6 shows the image processed by the equation 10. Non-hair color pixels are substituted by white pixels.

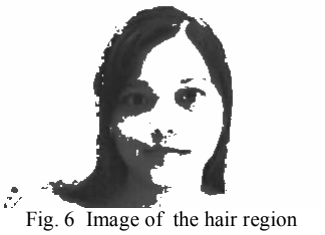

Fig. 6 Image of the hair region

The condition $(B - G < 15 \cup B - R < 15)$ precludes the pixels which are easily categorized as deep blue while hair color pixels commonly and statistically match the condition $(B - G < 15 \cup B - R < 15)$. The condition $(I < 80)$ includes the pixels which are dark. The condition $(20 < H \le 40)$ includes the pixels which are brown. Figure 6 shows that many non-hair pixels are misjudged, yet the Hair Quantization module perform the follow-up operations of the whole algorithm so as to minimize the impact from the misjudging.

C. Skin Quantization

This module quantizes skin color pixels and uses a 5x5 pixel square to express whether or not a region included pixels of skin color. The quantization lowers the image resolution; nonetheless, the following calculation of the geometric locations of pixels is speeded up. The number of black pixels within 25 pixels is counted and these 25 pixels regarded as non-skin color blocks if the number is beyond a threshold value 12. The rest are viewed as skin color blocks. Figure 7 demonstrates the image of Figure 5 after the Skin Quantization.

Fig. 7 Image of Fig. 5 after Skin Quantization

D. Hair Quantization

This module, similar to the Skin Quantization, quantizes hair color pixels and uses a 5x5 pixel square to express whether or not a region included pixels of hair color. The number of white pixels within 25 pixels is counted and these 25 pixels considered as non-hair color blocks if the number is beyond a threshold value 12. The rest are treated as hair color blocks. Figure 8 displays the image of Figure 6 processed by the Hair Quantization.

Fig. 8 Image of Fig. 6 after Hair Quantization

E. Get the Fusion of Features

This module, referring to the results from Hair Quantization, analyzes the result from Skin Quantization and determines the existence of a human face as well as its relation. The module also provides two options for human face detection; one is for single human face detection and the other is for multi-faces. Figure 9 illustrates the flowchart of the single human face detection:

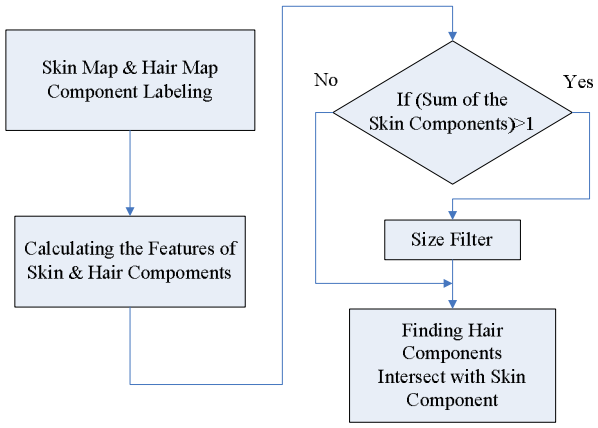

Fig.9 Flowchart of single human face detection

Component Labeling [14]:

The result from Skin Quantization and Hair Quantization is further calculated using Component Labeling. Adjacent pixels are defined as one component; two components have no adjoining pixels on the boundary. Component Labeling, applying 8-connectivity method, is able to search out all components in an image and get each of them labeled. Figure 10 shows the outcome from Component Labeling the Skin Quantization result showed in Figure 7.

Fig.10 Component Labeling

Only one component is specified, as Figure 10 indicates. However, as Figure 11 illustrates, several components are probably found out in some circumstances, such as many faces or objects with skin colors appears in an image.

Fig. 11 Image with two human faces

After the picture in Figure 11 goes through Skin Detection, Skin Quantization and Component Labeling, we can derive seven components with different labels in Figure 12.

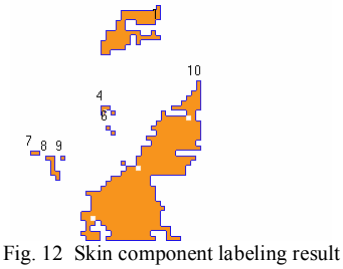

Fig. 12 Skin component labeling result

Calculating the Features of Skin & Hair Components:

The detection algorithm in Figure 9 calculates the features for all currently labeled components. The features to be calculated are area, center of gravity (COG), and the coordinates of the extreme left, top, right, and bottom pixels, respectively, of a component.

Size Filter [14]:

Size Filter is used to filter out noises and to preserve main skin features. The filtering conditions can be changed according to requirements such as the area of a component. To detect a single human face, only one component is required. Therefore among the components in Figure 12, only the largest component is preserved while the other small components are eliminated.

Finding Hair Components Intersect with Skin Component：

In this step we find the hair components that intersect with the skin component. The presence of human face requires that hair components and skin component intersect. Therefore, when there is no hair component intersecting with skin component, there is no human face present in the picture.

From practical observations, the most common twelve hair styles that we would like to detect are shown in Figure 13.

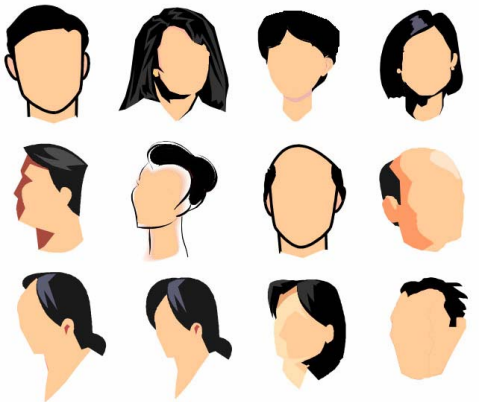

Fig. 13 Twelve hair styles we would like to detect

To determine whether the skin and hair components are extracted from a human face, we judge by the joint geometric relationship. Figure 13 shows the geometric relation between a skin component and a hair component. In order to simplify the

twelve complex geometric relations into some simple mathematical operations, we judge the joint geometric relationship using the "minimum bounding box" approach. The minimum bounding box is calculated for each skin component as well as each hair component. If the minimum bounding box of a skin component intersects with the minimum bounding box of a hair component, they are "related". Notice that the coordinates of each minimum bounding box has already been obtained in the step "Calculating the Features of Skin & Hair Components". The intersection of skin bounding box and hair bounding box is shown in Figure 14.

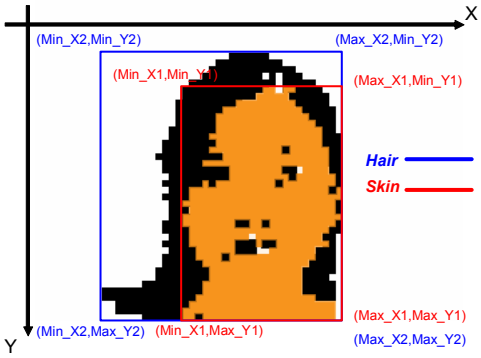

Fig. 14 the intersection of minimum bounding boxes

Figure 15 summarizes the several intersection relations between skin and hair bounding boxes form the picture in Figure 13.

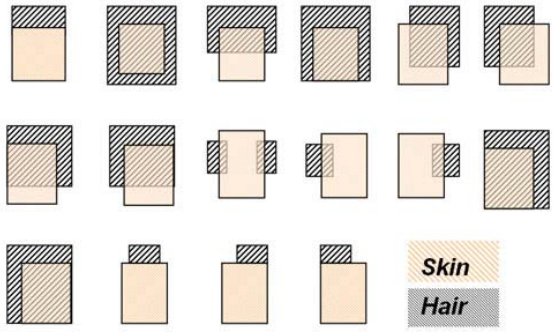

Fig. 15 Intersection relations between skin and hair bounding boxes

Note that not all intersection relationships are possible from a human face. We restrict the intersections to be the ones shown in Figure 15. That is, a hair component is associated with a skin component only if the hair bounding box intersects with skin bounding box according to one of the relations shown in Figure 15. If one hair component is associated with a skin component, the operation for finding hair component is thus finished. Figure 16 shows the consequence of the face detection.

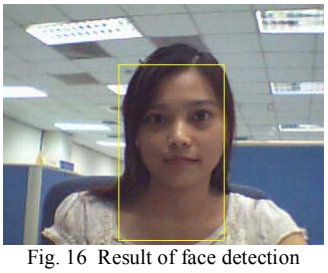

Fig. 16 Result of face detection

For faces with less hair, using the intersection of hair and skin bounding boxes still successfully detects a human face. Figure 17 shows some examples that our algorithm detects.

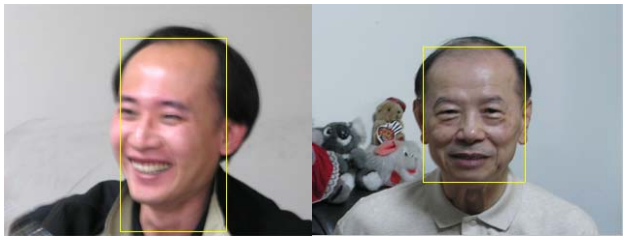

Fig. 17 Detection of the faces with less hair

Figure 18 demonstrates the flowchart of the multi-face dete ction.

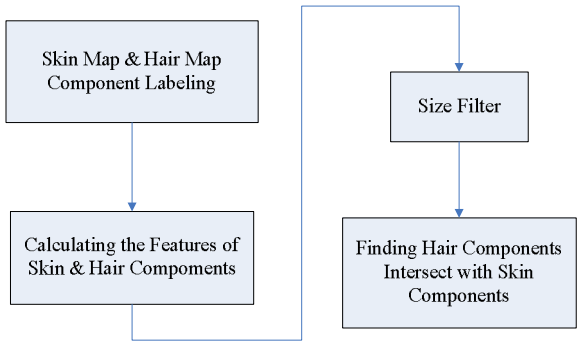

Fig. 18 Flowchart of the multi-face detection

The proposed algorithm can be extended to perform multi-face detections. The difference between single-face and multi-face detections is in the operation of the Size Filter. For single-face detection we only preserve the largest skin component. For multi-face detection, we preserve any component larger than 10 pixels. For each such skin bounding box we find the hair bounding box intersecting with it, to determine if a corresponding human face exists. A result of the multi-face detection is illustrated in Figure 19.

.

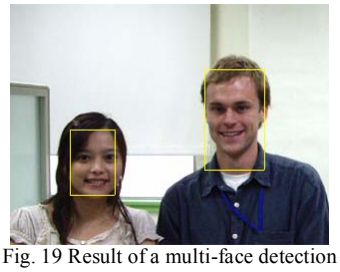

Fig. 19 Result of a multi-face detection

III. DISCUSSION

This algorithm had been applied to the developing pet robots for tracking human faces in real-time and cooperated with motor controllers in the pet robots. Figure 20 provides the architecture of a real-time face-tracking module on a pet robot:

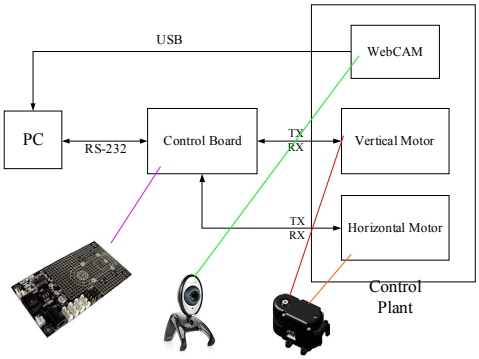

Fig. 20 Architecture of a real-time face-tracking module

This module comprises two servo motors for generating the vertical and horizontal motions of a web camera. Images are captured by the web camera and then transmit to a personal computer via a USB interface. The image data are then dealt with by the computer for detecting human faces and calculating the face positions as well. The execution of the real-time face-tracking is shown in Figure 21.

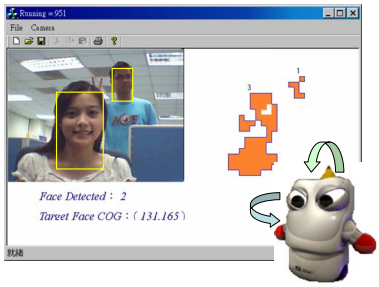

Fig. 21 Execution of the real-time face-tracking

IV. CONCLUSION

An algorithm for real-time human face tracking is realized. The algorithm takes the advantage not only of geometric relations between a human face and hair, but also of a precise skin color extraction. The frame-rate is up to 10 fps on a computer with a 1.4GHz Pentium 4 CPU. The accuracy of the single-face detection is better than 92% with a simple background and sufficient light source. From experiments, the multi-face detection results in slightly higher rate of misjudgment. The face detection is accomplished regardless of the viewpoints no matter it is a front view or a side view. However, the bottleneck of this algorithm consists in the interference during hair detection. For instance, detection errors might occur if anything other than human faces but with skin color stayed in the image. Another feature, in addition to complexion and hair, should be put in as a third basis for further betterment. The color model should be simultaneously simplified for reducing the perplexity which the third feature might contribute.

REFERENCES

[1] Jerome M. Shapiro, "Embedded Image Coding Using Zerotress of Wavelet Coefficients", IEEE Transaction on Signal Processing Vol.41 No.12 Decemrer 1993.

[2] T.K.Leung, M.C.Burl,and P.Perona, "Finding Face in Cluttered Scenes Using Random Labeled Graph Matching",Proc. Fifth IEEE Int'l Conf. Computer Vision, 1995, pp637-644.

[3] .Y. Dai and Y.Nakano, "Face-Texture Model Based on SGLD and Its Application in Face Detection in a Color Scence" ,Pattern Recognition,vol. 29,no. 6, 1996, pp.1007-1017.

[4] J. Yang and A.Waibel, "A Real-Time Face Tracker" ,Proc. Third Workshop Applications of Computer Vision, 1996,pp. 142-147.

[5] I. Craw, D. Tock, and A. Bennett, "Finding Face Features," Proc. Second European Conf. Computer Vision, 1992, pp. 92-96.

[6] A. Lanitis,C.J. Taylor,and T.F. Cootes, "An Automatic Face Identification System Using Flexible Appearance Models" ,Image and Vision Computing,vol. 13,no. 5, 1995, pp. 393-401.

[7] M. Turk and A.Pentland,"Eigenface for Recognition, "J.Cognitive Neuroscience,vol.3, 1991, pp.71-86.

[8] H. Rowley,S. Baluja,and T. Kanade, "Neural Network-Based Face Detection,"IEEE Trans. Pattern Analysis and Matchine Intelligence,vol.20,no.1,Jan. 1998.

[9] A. Rajagopalan,K. Kumar,J.Karlekar,R. Manivasakan,M. Patil,U. Desai, P. Poonacha, and S. Chaudhuri, "Finding Faces in Photographs" ,Proc. Sixth IEEE Int, l Conf. Computer Vision,1998.

[10] Linda G. Shapiro, George C. Stockman , "Computer Vision" , Prentice Hall, 2001, pp.192-193.

[11] M. Soriano, S. Huovinen, B. Martinkauppi, and M. Laaksonen, "Using the Skin Locus to Cope with Changing Illumination Conditions in Color-Based Face Tracking," Proc. of IEEE Nordic Signal Processing Symposium, pp. 383-386, 2000.

[12] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", 2nd Edition, Prentice Hall, 2002, pp.299-300.

[13] William K. Pratt, "Digital Image Processing", 3rd Edition, John Willey & Sons INC. 2001, pp.63-87.

[14] Ramesh Jain, Rangachar Kasturi, Brian G. Schunck, "Machine Vision", McGraw-Hill, 1995 ,pp.44-48.