

A Rule-based Real-Time Face Detector

Yi-Ting Huang and Cheng-Chin Chiang¹

Email: ccchiang@mail.ndhu.edu.tw

Department of Computer Science and Information Engineering

National Dong-Hwa University

Shoufeng, Hualien 974, R. O. C.

TEL: 03-8662500 ext. 22122, FAX: 03-8662781

ABSTRACT

In this paper, a real-time face detection algorithm is proposed for locating single or multiple faces in images and videos. Not only finding the face regions, this algorithm also locates the eyes and lips for each located face. The process of the proposed algorithm starts from the skin-color pixels extraction through rules built upon a quadratic polynomial model using the normalized color coordinate (NCC). Interestingly, this polynomial model with a slight modification can also be used to extract the pixels of lips very effectively. The benefit of using the same polynomial model is the reduction of computation time. Moreover, the skin-color regions and lips can be extracted simultaneously in one scan of the image or video frame. The algorithm then verifies the extracted eyes and lips using a set of rules derived from the common spatial and geometrical relationships of facial features. Finally, the precise face regions are determined accordingly. The implemented face detection system based on the proposed algorithm exhibits satisfactory performance in terms of both accuracy and speed for faces with wide variations in size, scale, orientation, color, and expressions.

Keywords: face recognition, face detection, facial features, skin-color pixel extraction, geometrical and spatial relationships.

1. INTRODUCTION

In recent years, the fast advancement of the image processing techniques and the cost down of various image/video acquisition devices encourage the development of many computer vision applications, such as vision-based surveillance, vision-based man-machine interfaces, vision-based biometrics. Among these many applications, the face recognition is a central task and thus attracts the attentions of many researchers. Some research systems and commercial products of face recognition

¹ To whom all correspondences should be addressed.

applications have been presented in many literatures [1-6]. According to the experience of these existing systems, it shows that only a powerful recognition engine is not enough for achieving a satisfactory recognition performance. A robust face detector is another critical key component without any doubts.

Generally, the capability of face detectors puts a limit on the capability of recognition systems. The final aim of designing a face detector is to develop a robust detector that can accurately extract all faces in images regardless of their positions, scales, orientations, colors, shapes, poses, expressions and light conditions. However, for the current state of the art in image processing technologies, this is a big challenge. For this reason, the face detectors deal with only upright and frontal faces in many existing systems [14-19].

Another concern in designing a face detector is the detection speed. In many applications such as videophones and surveillance, the real-time speed is a critical requirement. Due to the requirement of real-time speed, some algorithms that can precisely extract faces but require an extensive amount of time-consuming computations have to be excluded from the design.

In this paper, we propose a novel real-time face detection method that can accurately locate not only the face regions in images, but also the eyes and lips for each located face. The detailed capability of the proposed method is described as follows:

- (1) The persons can tilt their faces left or right for about 45° .
- (2) The persons can rotate their faces to any rotations as long as their lips and eyes are not occluded.
- (3) The persons can raise or lower their head for about 30° .

- (4) The light is controlled to uniformly illuminate on the faces to be detected and the light brightness is also controlled to prevent the face skin becoming white or black.
- (5) The size of each face is constrained between 1024 (=32x32) pixels and 9216 (=96x96) pixels. This constraint is set to fit the resolution requirements for general face recognition engines. The values can be dynamically adjusted if different resolutions are demanded.

The basic concept of the proposed algorithm is to extract and verify the desired components, including skin-color regions, lips, eyes and faces through several simple rules. It is found that the defined rules can handle a large degree of variations in faces. Since the proposed method needs no pixel-based template matching to extract the desired components, the computation time can be significantly saved. In addition, due to the simplicity of the rules, the proposed algorithm can detect faces at a real-time speed.

The rest of the paper is organized as follows. Section 2 makes a brief survey on some related work. The proposed algorithm is presented in Section 3. In order to show the effectiveness of the proposed algorithm, some experimental results are provided in Section 4 to evaluate the accuracy and speed of face detection. Finally, we conclude the research in Section 5.

2. RELATED WORK

A straightforward way for detecting face is to locate the face patterns by template matching [14,15,16,17]. The template can be designed or learned through the collection of a set of face images. During the matching process, a template is convolved with the sub-images everywhere within the input image to find the possible

candidates based on a predefined similarity or distance. To handle the many possible variations in size, orientation, and shape, etc., two methods are usually adopted. The first is to convert each input image with different transformations, e.g. resizing to different dimensions, before matching. Then, the template matching is performed on images of each transformation. The other alternative method is to use multiple face templates during matching. Clearly, the matching time increases very rapidly with the numbers and the dimensions of the used templates and the converted images, no matter what approach is used.

To reduce the search time of the possible face candidates over the input images, a skin-color pixels extraction is commonly performed [7-11]. The basic idea of the skin-color pixels extraction is to build a statistic model for colors of skin-like pixels. For example, the mixture of Gaussian distributions is a frequently used probabilistic model for this purpose [7,8]. With a probabilistic model, the probability of being a skin-color pixel for each pixel can be calculated. A threshold is then set on this probability to remove those non-skin pixels. There are also some other researchers proposed the neural network approach to approximate more complex models for skin-color pixels [11,19]. However, this usually increases the cost of computations.

After extracting the skin-color regions from images, the face templates are then used to verify these regions [12,13]. Those redundant skin-color regions that do not pass the verification are then removed from the candidate list. In addition to the template matching, the neural network verifier is an alternative design for performing verification [14]. Although the neural network verifier usually exhibits higher variation tolerance than the template matching, the long period of training time and the requirement of large training sets are the major drawbacks.

The above face detection method can be referred as a top-down approach. The term “top-down” means that the face regions are determined without resorting to the identification of individual facial feature components such as eyes, noses and lips, etc. An alternative approach to locate face regions is the bottom-up approach, which means that the face regions is constructed from the identified facial feature components. In other words, the facial features have to be extracted before the face regions are located. In general, some individual facial features like eyes and lips own some visual properties that are less sensitive to the face variations mentioned previously. Thus, the extraction of these facial features would be more stable. Hence, the bottom-up approach should be intuitively more robust and accurate than the top-down one. In addition, as the individual components are much smaller than the whole face in size, the processing time during extraction is thus also much shorter.

In this paper, we adopt the bottom-up approach. The process of the proposed algorithm starts from the skin-color pixels extraction through several rules built on a quadratic polynomial model using the normalized color coordinate (NCC). Interestingly, this polynomial model also can be used to extract the pixels of lips effectively. Finally, the algorithm verifies the extracted eyes and lips based on a set of rules derived from their common spatial and geometrical relationships to remove the false facial components. The final precise face regions are determined accordingly. The detailed method is described in next section.

3. RULE-BASED FACE DETECTION ALGORITHM

3.1 Rules for Skin-Color Region Extraction

As mentioned previously, the purpose of extracting skin-color regions is to reduce the search time for possible face regions on the input image. As the goal of

our algorithm is to detect faces at a real-time speed, we exclude the approaches based on some complicated probabilistic models due to their high computation costs. Alternatively, we use a much simpler method that requires only a very few computations.

Firstly, the proposed algorithm uses the normalized color coordinate (NCC), rather than the RGB color system, for pixel colors. The conversion from RGB to NCC, denoted by (r, g) , is defined as follows:

$$r = \frac{R}{R + G + B} \text{ and } g = \frac{G}{R + G + B},$$

where R, G, B denotes the intensity of pixels in red, green and blue channels, respectively. One of the major benefits of using NCC is the brightness variation in images can be normalized according to the above transformation. The other benefit is that the dimension of NCC is 2, which is easier to visualize and analyze when building the model for skin-color pixels.

To build a model for skin-color pixels, we plot the distribution of skin-color pixels over the $r-g$ plane. Figure 1 shows the distribution. Apparently, the skin-color pixels form a compact region over the $r-g$ plane. In many existing methods, the distribution is approximated with some probabilistic models, such as Gaussian mixtures [19]. The major drawback of this kind of modeling is the high computation costs in calculating the probabilities for each pixel in the image. These computation costs are harmful to our goal of real-time detection speed. Therefore, instead of trying to build the precise probabilistic model, we modify the approach of Soriano and Martinkauppi [10], which uses two quadratic polynomials to approximate the upper and lower boundaries of the compact region formed by the skin color

distribution on the r - g plane. These two polynomials are:

$$f_{upper}(r) = -1.3767r^2 + 1.0743r + 0.1452, \quad (1)$$

$$f_{lower}(r) = -0.776r^2 + 0.5601r + 0.1766. \quad (2)$$

Therefore, the pixels falling within the area between these two polynomials can be extracted by the following rule checking on their r , g values:

$$\text{R1: } g > f_{lower}(r) \text{ and } g < f_{upper}(r).$$

In Figure 1, we see the above two inequalities define a crescent-like area on the r - g plane. Beside these two polynomials, Soriano and Martinkauppi also defined a circle on r - g plane to exclude the bright white pixels which falling around the point $(r, g)=(0.33, 0.33)$ from the region enclosed by the two polynomials. Therefore, the exclusive circle is defined by:

$$\text{R2: } (r - 0.33)^2 + (g - 0.33)^2 \leq 0.0004.$$

Apparently, the above three mathematical formulas are still not accurate enough to filter out the non-skin-color pixels, particularly for some yellow-green, blue and orange pixels that fall around the top, left end and right end of the crescent area respectively. Figure 2 illustrates the result of applying rules R1 and R2 simultaneously for skin-color pixel extraction.

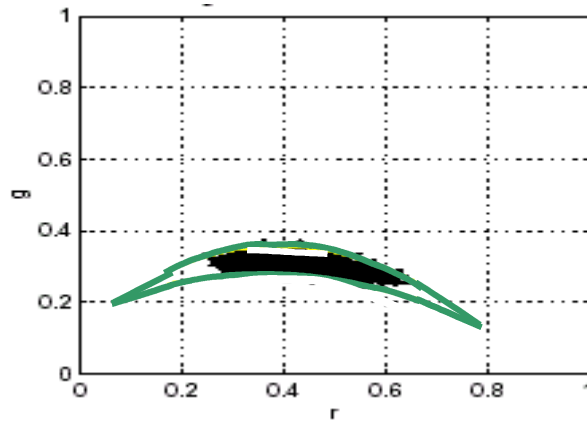


Figure 1. The distribution of skin-color pixels on r - g plane and the two

quadratic polynomials for the upper and lower boundaries. [10]

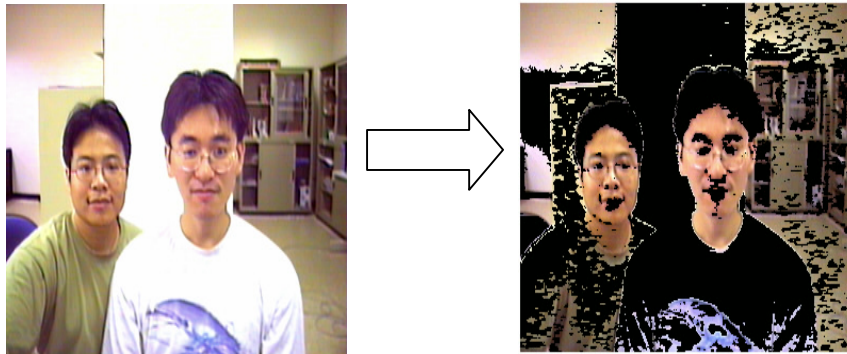


Figure 2. The skin-color pixels filtering using the method of [10]. The left image is the input image and the right image is the resultant image.

To improve the filtering results, the proposed algorithm adds two extra simple rules. These two rules are:

R3: $R > G > B$, and

R4: $R - G \geq 45$.

Rule R3 is derived from the observation that the skin-color pixels tend to be red and yellow. The blue intensity is always the smallest one among the three channels. With Rule R3, the blue pixels can be effectively removed. Rule R4 is defined in order to remove the yellow-green pixels. Figure 3 illustrates the improved result after introducing Rules R3 and R4. It is clear to see that Rules R1 through R4 involve only very simple computations. The filtering process is much faster than those probabilistic approaches.

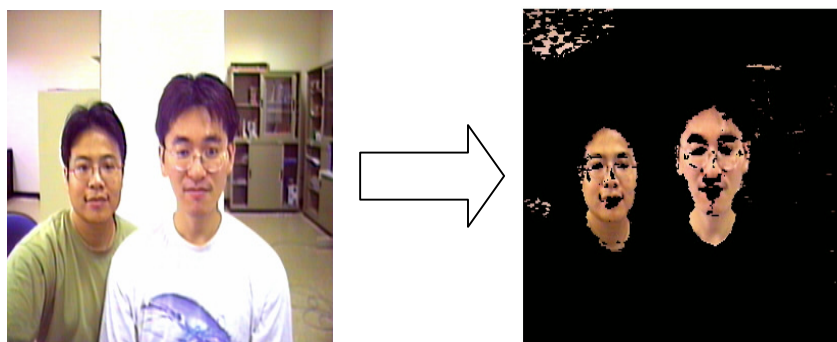


Figure 3. The improved result of skin-color pixel extraction after introducing Rules R3 and R4.

In summary, we summarize the final rule for extracting skin-color pixels as follows:

$$S = \begin{cases} 1 & \text{if } (g < f_1(r)) \& (g > f_2(r)) \text{ and } (W \geq 0.0004) \text{ and } (R > G > B) \text{ and } ((R - G) \geq 45), \\ 0 & \text{otherwise.} \end{cases}$$

where $S=1$ means the examined pixel is a skin-color pixel. After the skin-color pixel extraction, the algorithm finds the bounding boxes for the connected components of skin-color pixels. These bounding boxes become the candidate skin-color regions for further identifications of individual facial features.

3.2 Rules for Lips and Eyes Detection

Observing the pixels of lips, we find that the colors of lips range from dark red to purple under normal light condition. These colors distribute on the areas below the crescent area for skin-color region on the r - g plane. According to human's visual perception, the lips are very easy to be differentiated from the face skins for any races of persons based on their relative color contrasts. Based on this observation, the color distribution of the lips should be disjunctive to that of normal skins.

Similar to the extraction of skin-color pixels, we try to define another quadratic polynomial as the discriminant function for lip pixels. When defining this polynomial, we have the following two design goals:

- (1) The polynomial should be computationally efficient, and
- (2) The detection of lip pixels can be done in parallel with the detection of skin-color pixels.

Based on the above two goals and our observations about the lip color distribution, we find that the quadratic polynomial of the lower boundary of the $f_{lower}(r)$ defined in Eq.(2) can be reused with a slight modification. As the lip color distribute around the area below the crescent area, we just need to decrease the constant term in

$f_{lower}(r)$. Therefore, according to the statistics on lip pixels, we define the discriminant function for lip pixels as:

$$l(r) = -0.776r^2 + 0.5601r + 0.165. \quad (3)$$

Beside $l(r)$, those dark pixels, which are usually the eyes, eyebrows or nose holes, also need to be excluded. We thus define the following rule for detecting lip pixels:

$$L = \begin{cases} 1 & \text{if } g \leq l(r) \text{ and } R \geq 20 \text{ and } G \geq 20 \text{ and } B \geq 20, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where $L=1$ means the pixel is a lip pixel. Notice that this polynomial also separates the lip distribution from the skin distribution to confirm to the fact of the disjunction between these two distributions.

The benefit of using the same mathematical model in both skin detection and lips detection is that the computations for discriminant function can be significantly reduced. In addition, the detections of both skin pixels and lip pixels can be done in parallel in one scan of the image or video frame. Thus, this design completely follows up the two design principles listed above. After detecting the lip pixels on the images, the proposed algorithm also perform the connected component labeling to group the connected lip pixels into a component.

Two side effects arise in performing the detections of skin pixels and lip pixels in parallel. One is that the information of the skin color regions cannot be utilized to reduce the search space of the lip pixels. This effect does not influence the detection speed much because the extra cost in computing the discriminant function for lip pixels is very low. The other effect is that some pixels outside the skin-color regions might be wrongly detected as the lip pixels if they satisfy the discriminating

conditions. To handle this situation, we just need to perform a post check to see if the detected lip component is contained within the detected skin-color regions. This check can be easily done through the computation of intersected areas between the bounding boxes of skin-color regions and lip components. For those lip components falling outside the skin-color regions, the algorithm removes them from the candidate list of lip components.

Remind that the proposed algorithm does not deal with the faces that tilt left or right for more than 45° . Therefore, the lips are always below the positions of the two eyes for a face. With this assumption, we can exclude many combinations for lips and eyes from considerations in later geometrical and spatial relationship verifications. For those detected noise components that are neither eyes nor lips, the algorithm will also remove them in the consequent verification process that are to be described later.

Observing the eye components under normal light condition, the eyes tend to be the darkest part on the face provided that no shadows are projected. However, the pixels of eyes are not necessarily black, i.e. $r=g=0$. Therefore, the polynomial used for extracting lips cannot be used, although it can extract black pixels very effectively. For this reason, the algorithm uses another simple way to extract the dark components on the faces. The eye components are extracted through a threshold operation (e.g. $\text{threshold} = 20$) on the grayscale image converted from the original color image. The method is found to be very efficient and effective [12]. Even though some dark components are wrongly detected, later verification process will remove them according to some geometrical rules.

3.3 Rules for Geometrical Verifications and Face Region Determination

After extracting the possible lips and eyes components, the proposed algorithm removes those mismatches by verifications based on a set of rules. These rules come from the general geometrical and spatial relationships among lips and eyes under normal light conditions.

The rules used by the proposed algorithm are described in the following.

(1) Let $\mathbf{p}_{eL}(x_{eL}, y_{eL})$ and $\mathbf{p}_{eR}(x_{eR}, y_{eR})$ be the two centers of the possible left eye and right eye, respectively. Then the angle between the line $\overline{\mathbf{p}_{eL}\mathbf{p}_{eR}}$ and the horizontal line must be in the range $[-45^\circ, 45^\circ]$. This rule is used to exclude those faces tilted left or right for more than 45° .

(2) Let $\mathbf{p}_m(x_m, y_m)$ be the center of the mouth component and $\mathbf{p}_{mp}(x_{mp}, y_{mp})$ be the point of projecting $\mathbf{p}_m(x_m, y_m)$ onto the line $\overline{\mathbf{p}_{eL}\mathbf{p}_{eR}}$. Given that the slope of $\overline{\mathbf{p}_{eL}\mathbf{p}_{eR}}$ is denoted as m_e . Then the following rules must be satisfied:

$$\begin{cases} x_{eL} - 0.1 * |x_{eR} - x_{eL}| \leq x_{mp} \leq x_{eR} + 0.1 * |x_{eR} - x_{eL}|, & \text{if } |m_e| < 0.2, \\ x_{eL} \leq x_{mp} \leq x_{eR} + 0.5 * |x_{eR} - x_{eL}|, & \text{if } 0.2 \leq m_e \leq 1, \\ x_{eL} - 0.5 * |x_{eR} - x_{eL}| \leq x_{mp} \leq x_{eR}, & \text{if } -1 \leq m_e \leq -0.2. \end{cases}$$

where we assume that $x_{eL} < x_{eR}$ without loss of generality. The first rule rules out the combinations that the projected position of mouth falls far away from the areas between two eyes, provided that the face is not tilted much (about 15°). The second and the third rules exclude the combinations that the mouth position is not consistent with the tilt direction of heads. In other words, the projected point of the mouth cannot be at the left side of left eye provided that the head tilt left ($0.2 \leq m_e \leq 1$), and vice versa.

(3) Let $L(\overline{\mathbf{pq}})$ denote the length of the line segment $\overline{\mathbf{pq}}$ and $\mathbf{p}_{ec}(x_{ec}, y_{ec})$ be the center point of $\overline{\mathbf{p}_{eL}\mathbf{p}_{eR}}$. Then the following rule must be satisfied:

$$0.8 \leq \frac{L(\overline{\mathbf{p}_m\mathbf{p}_{ec}})}{L(\overline{\mathbf{p}_{eL}\mathbf{p}_{eR}})} \leq 1.5.$$

(4) Let $W(C_i)$ denote the width of the bounding box of component i . Then the following rules must be satisfied:

$$\frac{1}{1.3} \leq \frac{W(C_{Left-eye})}{W(C_{Right-eye})} \leq 1.3,$$

$$0.6 \leq \frac{W(C_{Left-eye})}{W(C_{Lip})} \leq 1.1 \text{ and } 0.6 \leq \frac{W(C_{Right-eye})}{W(C_{Lip})} \leq 1.1.$$

The second rule is useful in filtering out the noise components. Note that the eyebrows sometimes connect to the eyes in the extracted components. For such cases, the algorithm regards the connected eye and eyebrow as a component. This action usually affects only a little precision of the located positions for eyes, but not the accuracy of the extracted eye components. The detection accuracy of faces is thus not significantly affected.

According to the verification on the above four geometrical and spatial relationship rules for eyes and lips, the candidate facial triangles on the skin-color regions can be extracted for further processing. In the proposed algorithm, the clear separations of both lips components and eyes components can greatly reduce the number of extracted triangles. Besides, the assumption that requires the lips to be always below the eyes in one face is also very helpful in reducing the possible combinations, too.

Most noise components can be removed after the verification process. However, it still might be possible that some confusing combinations need to be dealt with. The confusing combinations are from those triangles that might contain the lips and eyes of different faces. We find that the remained confusing triangles are those that have the same lips/mouth components in one or two of the triangle vertices. We then proceed to the next step. A criterion, named skin color ratio (SCR), is defined for the arbitration of the remained confusing triangles. The *SCR* is defined as follows:

$$SCR(\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3) = \frac{N_{skin|\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3}}{A(\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3)} \quad (5)$$

where $A(\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3)$ denotes the area of the triangle $\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3$ and $N_{skin|\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3}$ is the number of skin color pixels within the triangle $\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3$. If *SCR* is very small, then $\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3$ is very likely to be a triangle that contains three facial components on distinct or scattered skin-color regions. For this kind of triangles, it has very low possibility to be a face. Hence, we can use *SCR* as the criterion for arbitration of tied triangles.

In order to calculate $N_{skin|\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3}$, it might need a few computations. Fortunately, the introduced cost is not very high because the number of confusing triangles is usually very small. To reduce the computation time, we use a fast method to calculate $N_{skin|\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3}$. This method is to examine each detected skin-color pixel on the image to see if it is in the interior of each triangle $\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3$. If it is, then $N_{skin|\Delta \mathbf{p}_1 \mathbf{p}_2 \mathbf{p}_3}$ is increased by one. In this way, we need not to use time-consuming algorithms for tracing out the interior pixels of each confusing triangle. We just need to perform simple triangle interior checks for each skin-color pixel. We ignore the details of triangle interior check and triangle area calculation here because they are

very trivial.

After extracting the triangles of facial components, the final face regions are simply determined according to the triangles. Figure 4 shows the dimension of the bounding box for face region. In Figure 4, L_{ee} is the distance between the two eyes. L_{me} is the distance between the center point of the line segment connecting the two eyes and the lowest point of the lips component. Note that the bounding box can be rotated according to tilted angle of the line connecting the two eyes.

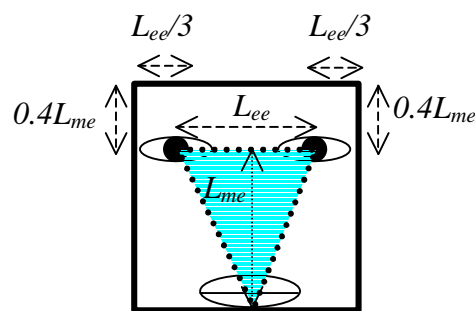


Figure 4. Determination of a face region according to the triangle of facial components.

4. PERFORMANCE EVALUATION

For the performance evaluation, we have implemented the algorithm on a PC with a Pentium 4 1.6G CPU and 128M RAM. The implemented system has two modes of operation. The first is on-line mode that is designed to detect faces of video frames captured from a PC camera in real time. The other mode is off-line mode that is designed to detect faces in still images that are prepared beforehand. To find out the detection accuracy and the detection speed, we prepare a database of 400 images. Among these 400 images, 300 images contain only single face on each image, while the other 100 images have multiple faces on each. These images contain both simple and complicate backgrounds.

Table 1 shows the detection accuracy of the proposed algorithm. The average

detection accuracy is around 92% and the false detection rate is very low. Table 2 shows the detection speed of the proposed algorithm. The dimension of each processed image is 320x240 pixels. It is easy to see that the proposed algorithm indeed achieves real-time detection of faces with a satisfactory accuracy from Table 1 and Table 2. In Figure 5, some snapshots of skin-color region detection are shown. It verifies the effectiveness of our skin-color pixel detection based on the quadratic polynomial models. In Figure 6, the results of lips component detection are illustrated. The results show that the detections are very accurate. In our experiments, the detection accuracy of lips achieves around 95% after performing the intersection of the extracted skin-color regions and the possible lips component candidates. The false detection rate of lips is below 2.5%. Therefore, this result verifies the effectiveness of our quadratic polynomial model for lip color pixels. The high accuracy of lips detection is very helpful in speeding up the verification process to be performed. In Figure 7, the extracted facial triangles composed of eyes and lips as well as the final facial regions are shown for several tested images. These images contain cases of single/multiple persons, multiple orientations, body interference, different expressions.

Table 1. The detection accuracy of the proposed algorithm.

	Single-face images (300 images)	Multiple-face images (100 images)
Total number of faces in images	300	320
Number of correct detections	282	290
Number of false detections	7	15
Detection Accuracy	94.0%	90.6%
False Detection Rate	2.33%	4.69%

Table 2. The detection speed of the proposed algorithm.

	Single-face images	Multiple-face images
Detection Speed	11 images/sec	8 images/sec

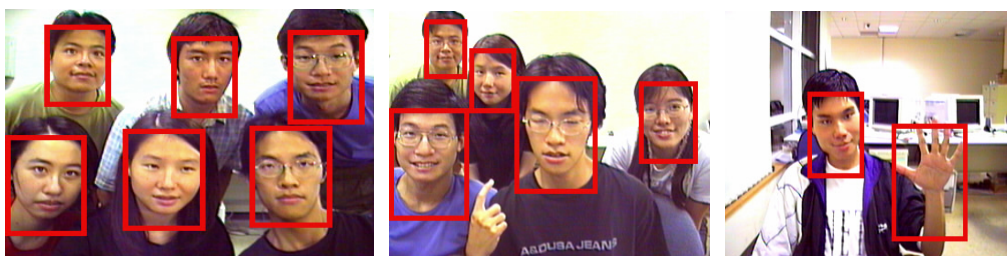


Figure 5. Results of skin-color region detection.



Figure 6. Results of lips detection.

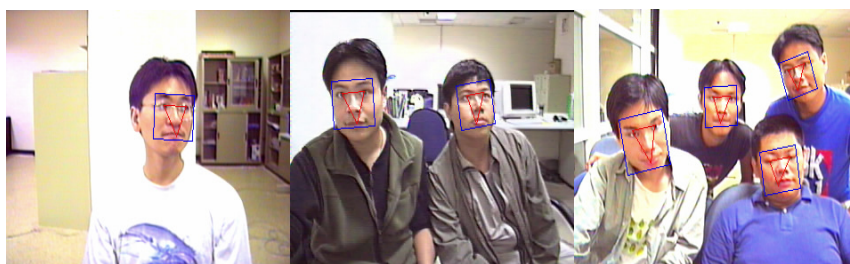


Figure 7. Results of facial triangle detection and facial region determination.

In order to test the robustness of the proposed algorithm, we also conduct the experiments of detecting faces wearing sunglasses. Figure 8 demonstrates the results. The results show that the faces still can be detected as long as the frame of the sunglasses is not very dark and thick. If the frame is dark and thick, the face is usually partitioned into two separated regions and thus usually causes the missed detection.



Figure 8. Results of detecting faces wearing sunglasses.

5. CONCLUDING REMARKS

According to the experimental results, the proposed algorithm exhibits satisfactory performances in both detection accuracy and detection speed. The cases that the proposed algorithm can handle are also very diverse. Actually, in many applications which need to handle only very simple cases, the proposed algorithm can be further improved in both speed and accuracy by introducing more verification rules for ruling out some impractical cases. However, there are a few restrictions on using the proposed algorithm:

- (1) The light condition must be normal. In other words, the faces to be detected cannot be too bright or too dark. In addition, the proposed algorithm does not allow the vast shadows on the faces for they interfere the geometry properties of facial features.
- (2) The facial features of faces must appear on the images. Thus, our

algorithm cannot detect those incomplete faces resulted from serious occlusions and large orientations.

In the future, we plan to improve the algorithms in two directions:

- (1) Increasing the robustness of light variations by developing a light compensation/correction preprocessing technique; and
- (2) Developing more precise component-based detection and verification process for facial features by designing corresponding facial component detectors.

The first is aiming at to loosen the restrictions on light condition and the second to handle the problems of occlusions and large orientations.

6. ACKNOWLEDGEMENT

This work was supported in part by the National Science Council under Grant NSC-90-2218-E-259-001.

REFERENCES

- [1] Turk M. A. and Pentland P., "Face Recognition Using Eigenfaces", Proc. of *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586-591, 1991.
- [2] Belhumeur P., Hespanha J., and Kriegman D., "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", *IEEE Trans. on PAMI*, 19(7): 711-720. 1997.
- [3] Yuela, P.C.; Dai, D.Q.; Feng, G.C., "Wavelet-Based PCA for Human Face Recognition", *Image Analysis and Interpretation*, 1998 IEEE Southwest Symposium on, pp. 223 –228, 1998.
- [4] Meng Joo Er; Shiqian Wu; Juwei Lu; Hock Lye Toh., "Face Recognition with Radial Basis Function Neural Networks", *IEEE Transactions on Neural Networks*, Volume: 13 Issue: 3 , pp. 697 –710, 2002.
- [5] R. Chellappa, C.L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces, A Survey", *Proc. of the IEEE*, Vol. 83, pp. 705-740, 1995.
- [6] "BioID: A Multimodal Biometric Identification System", *IEEE Computer*, Vol. 3,

No. 2, pp. 63-68, 2000.

- [7] Shinjiro Kawato, Jun Ohya , “Automatic Skin-color Distribution Extraction for Face Detection and Tracking”, Proc. Int. Conf. On Signal Processing, Vol.II, pp.1415-1418, 2000.
- [8] M. Störring, H. Andersen, and E. Granum , “Estimation of the Illuminant Colour from Human Skin Colour “ , In *IEEE International Conference on Face & Gesture Recognition*, pages 64-69, Grenoble, France, March 2000. IEEE Computer Society.
- [9] J.-C. Terrillon, M. N. Shirazi, H. Fukamachi, and S. Akamatsu , “ Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images “, In *Proc. of the International Conference on Face and Gesture Recognition*, pages 54 - 61, Grenoble, France, 2000
- [10] M. Soriano, S. Huovinen, B. Martinkauppi , M. Laaksonen , “ Using The Skin Locus To Cope With Changing Illumination Conditions in Color-Based Face Tracking “ , *Proc. IEEE Nordic Signal Processing Symposium (NORSIG 2000)*, Kolmarden , Sweden , P.383-386, 2000.
- [11] Ishii, H.. Fukumi, M.. Akamatsu, N. “ Face Detection Based on Skin Color Information in Visual Scenes by Neural Networks “, *IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on* , Vol.5, pp.557-563,1999.
- [12] Chiunhsiun Lin , Kuo-Chin Fan , “ Human Face Detection Using Triangle Relationship “ , *15th International Conference on Pattern Recognition*, Vol. 2, pp. 945-948, September 2000.
- [13] S.H.Jeng , H.M. Liao , Y.T.Liu , and M.Y.Chern , “ An Efficient Approach for Facial Feature Detection Using Geometrical Face Model, “ *Int , Proceedings of the ICPR 1996*,pp. 1739-1755,1993.
- [14] Henry Rowley, Shumeet Baluja, Takeo Kanade, “Neural Network-Based Face Detection”, *IEEE Tran. On PAMI*, Vol. 20, No. 1, pp. 23-38, 1998.
- [15] Kah-Kay Sung. *Learning and Example Selection for Object and Pattern Detection*. PhD thesis, MIT AI Lab, January 1996. Available as AI Technical Report 1572.
- [16] Alex Pentland, Baback Moghaddam, and Thad Starner. View-based and modular eigenspaces for face recognition. In *Computer Vision and Pattern Recognition*, pp. 84-91, 1994.
- [17] Antonio J. Colmenarez and Thomas S. Huang. Face Detection with information-based maximum discrimination. In *Computer Vision and Pattern Recognition*, pp. 782–787, 1997.
- [18] Edgar Osuna, Robert Freund, and Federico Girosi. Training Support Vector

Machines: An Application to Face Detection. In *Computer Vision and Pattern Recognition*, pp.130–136, 1997.

- [19] S. H. Lin, S. Y. Kung, and L. J. Lin. Face Recognition/Detection by Probabilistic Decision-based Neural Network. *IEEE Trans. on Neural Networks, Special Issue on Artificial Neural Networks and Pattern Recognition*, 8(1), January 1997.