Design and Implement of Image Segmentation Algorithms

Wang Li 119034910222

***Abstract*—Location detection and segmentation of human is an important process for face identification. Many things influencing the success of detection and segmentation of human, between it, complex color background, styles of the clothes, and the luminate etc. For precise segmentation, this paper proposes a color image face segmentation algorithm to segment the face, and proposes another similar algorithm to segment the clothes. Image face segmentation algorithm and clothes segmentation algorithm, combines skin color model, human face structure features, morphological algorithms, edge information, of which the performance is well. Firstly, construct an elliptical skin model in the YCbCr color space to segment the skin color pixels from the background image. While ginkgo is similar to skin in color, we use edge information, then separate face region from ginkgo. To get the further steps, we use mathematical morphological operators and topological algorithm to filter the image. And, candidate face regions will be found., finally, we try to use face template to do matching for suspected face areas. In addition, we imply this method to some different images, the results show the target region successfully.**

***Index Terms*—Face segmentation, morphological operation, topological description, edge detection**

I. Introduction

This paper does the segmentation of human face and body separately. And the algorithms are similar but a little different due to the different features of them. For face segmentation, up to now, much work has been done on detecting and locating faces in color images and the methods like template-based [2], neural network-based [3], feature-based [4], machine learning-based [5] have been well studied by many researchers. Among many face detection algorithms, the method based on skin color model has been widely used for its convenient use, simple performance and high detection speed [1]. It is unreliable to make face detection only using skin color features when there are a large number of objects similar to skin color. So, we need to utilize the other features of human face to further verify. We use edge detection to get the edge information, then combine the skin color segmentation, separate the leave from face region to make preparation for morphological algorithm such as open operator and close operator. We also try topological algorithm to retain face while filtering those areas containing too many holes. For body segmentation, due to the difference between clothes and the background, the combination of image edge segmentation, morphological operation and image threshold segmentation is used to solve this problem. Edge is a collection of gray-level abrupt pixels in the image, which are usually detected by differentiation. The basic edge detection algorithms are: Roberts operator, Prewitt operator, Sobel operator. The more advanced algorithms are: Marr-Hilderth edge detector, Canny edge detector. In this paper, we try Canny operators to do the edge detection. On this basis, dilate and erode operations are used to eliminate fine leaves and keep body parts.

The paper is submitted in the course of Digital Image Processing on November23, 2019. It is completed with the support of Prof. Rui Zhang and my partners. My heartfelt thanks to them. Wang Li majors in Electronic and Communication Engineering in School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University

II. The Basic Theory of Segmentation

A. Skin Color Segmentaion

- From RGB to YCbCr color spaces [6]:
  Processing color information can be performed in different color spaces, but in order to improve the performance of skin color clustering, YCbCr is used [6] to build a skin color model, as the chrominance components are almost independent of luminance component in this space. The equation of transformation between RGB to YCbCr is given as the following:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & -37.797 & 112.000 \\ 128.553 & -74.203 & -93.786 \\ 24.966 & 112.000 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

- Skin-color Modelling [1]:
  As the statistics show, different races' color is mainly affected by the brightness information. So we directly need consider CbCr in the YCbCr space. Skin color show good clustering properties, diagram is as follows:

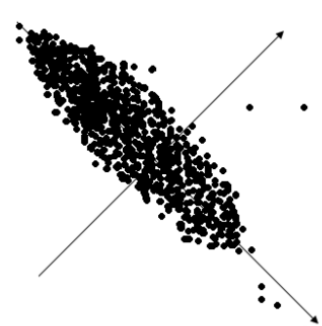


Fig. 1. The color clustering results in the YCbCr.

There are many ways of skin-color modelling, including Gaussian model, mixture of Gaussian model, and elliptical boundary model, etc. Gaussian skin color model algorithm is relatively complex and long-running, so elliptical model is used in this paper. The expression of skin color elliptical model function is:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} Cb - c_x \\ Cr - c_y \end{bmatrix} \tag{2}$$

$$\frac{(\mathrm{x} - \mathrm{ec_x})^2}{\mathrm{a}^2} + \frac{(y - ec_y)^2}{b^2} \tag{3}$$

if the pixel information satisfies the following condition, then mark it as skin color.

$$\begin{aligned} &c_x = 109.38 \quad c_y = 152.02 \quad ec_x = 1.60 \quad ec_y = 2.41 \\ &\theta = 2.53 \quad a = 25.39 \quad b = 14.03 \end{aligned} \tag{4}$$

the original image to be detected is shown in Fig.2, it was taken under a big ginkgo tree, to get better detection, we did not do downsampling, the size of the image is 1483×1483. From fig.2 we can find that the color of ginkgo tree is similar to skin, as expected, after skin color segmentation, ginkgo and skin region are retained together, as fig.3 shows.

B. Edge-Based Detection

[3] The so-called edge refers to the collection of consecutive pixel points on the boundary line of two different regions in the image, which is a reflection of the local feature discontinuity of the image, and reflects the mutation of image characteristics such as gray, color and texture. In general, the edge-based segmentation method refers to edge detection based on gray values, which is based on the observation that the edge gray value will exhibit step or roof type changes. Commonly used edge detection operators include Laplacian operators, Sobel operators, Canny operators, etc., which will be briefly introduced below:

1) Sobel Operators
Sobel operators are usually divided into two directions, one for detecting horizontal edges and the other for detecting vertical edges. The principle of the Sobel operator can be explained by the following formula:

$$\nabla \mathrm{f} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \tag{5}$$

$$\frac{\partial f}{\partial x} = f(x+1, y) - f(x, y) \tag{6}$$

$$\frac{\partial f}{\partial y} = f(x, y+1) - f(x, y) \tag{7}$$

Where $f(x, y)$ represents the pixel value at point (x, y). From the above formula, we can deduce the final Sobel formula as follows:

$$\begin{aligned} \nabla \mathrm{f} =& [f(x+1, y) - f(x, y)] \\ &+ [f(x, y+1) - f(x, y)] \end{aligned} \tag{8}$$

The Sobel operator can also be represented in the form of a template, and the values in the template represent the weighting factors corresponding to the corresponding pixels.

$$\mathrm{M}_v = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \mathrm{M}_H = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{9}$$

Equations (9) represent template functions in the vertical and horizontal directions, respectively.

C. Morphological Filtering

[2] Morphological filters are an important class of nonlinear filters that have emerged in recent years. They evolved from early binary filters to later multivalued (grayscale) morphological filters for shape recognition, edge detection, and texture analysis. Wide range of applications in areas such as image restoration and segmentation. Common mathematical morphology methods include erosion, dilation, opening operation and closing operation.

1) Dilation
The so-called dilation is to expand the neighborhood of the highlighted part of the image, so that the highlighted part of the expanded effect image becomes larger. The expression is:

$$A \oplus B = \{x, y | (B)_{xy} \cap A \neq \varnothing\} \tag{10}$$

Where, first, we design the structural element B, then make a Convolution of the structural element B and the original image, so the maximum value of the pixel of the area covered by the kernel B is calculated, and this maximum value is assigned to the pixel specified by the reference point. This will cause the highlights in the image to grow gradually, as shown in the following image:

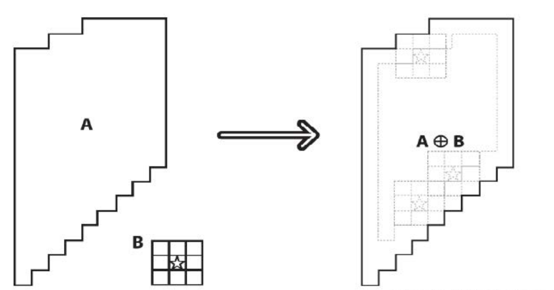


Fig. 2. The dilation operation.

2) Erosion
The process of erosion and expansion is just the opposite. Erosion is a way to eliminate boundary points and eliminate small and meaningless objects in the image. Erosion is the process of shrinking the image boundary to the inside, so erosion is the operation of finding the local minimum. The local minimum is obtained by convolving the structural element B with the original image. The expression is:

$$A \ominus B = \{x, y | (B)_{xy} \subseteq A\} \tag{11}$$

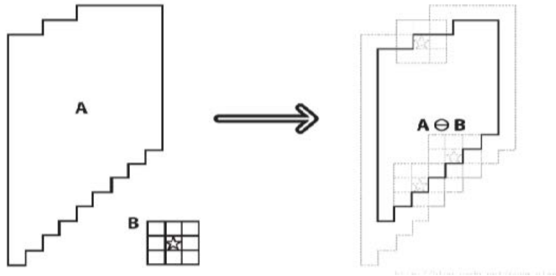


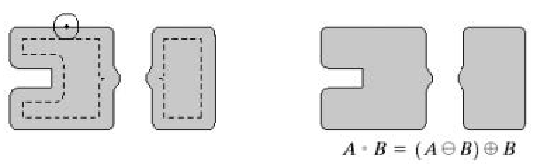
Fig. 3. The erosion operation.

The process is as follows: 3) Opening and Closing The image opening operation can be regarded as the first erosion operation, and then the dilation operation. Conversely, the image closure operation can be regarded as the first dilation operation, and then the erosion operation. The opening and closing operations of the image can be expressed by equations (10) and (11) respectively.

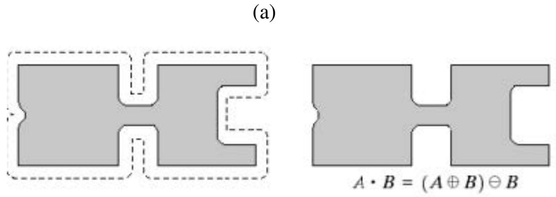
$$A_B = A \circ B = (A\Theta B) \oplus \mathrm{B} \tag{12}$$

$$A^B = A \cdot B = (A \oplus B)\Theta B \tag{13}$$

Where A is the original image, B is the structural element, symbol Θ represents erosion, and symbol $\oplus$ represents dilation. The effect of the opening and closing operation can be represented by Fig.4. As shown in Fig.6, Opening generally



(a)



(b)

Fig. 4. (a) Opening operation, (b) Closing operation.

smoothes the contour of an object, breaks narrow isthmuses, and eliminates thin protrusions. Closing tends to smooth sections of contours, generally fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour.

D. Topological Algorithm

[2] Topological properties are useful for global descriptions of regions in the image plane. Topology is the study of properties of a figure that are unaffected by any deformation, as long as there is no tearing or joining of the figure. The expression is:

$$E = C - H \tag{14}$$

where, C is the number of connected components, and H is the number of holes , E is the Euler number in a region. As is shown in Fig.5, where, C is 1, H is 2, and the Euler number is $E = C - H = -1$.

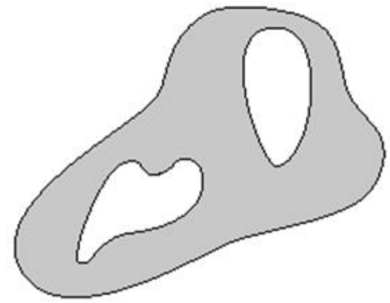


Fig. 5. A region with two holes.

III. Our Image Segmentation Algorithm

In this section, we will introduce two algorithms based on the relevant theories presented in the second section, of which one is used to face segmentation, the other is used to body segmentation.

A. Image Face Segmentation

In this paper, image face segmentation is based on Skin Color Modelling, but the background of the image we should take is ginkgo tree, which makes the algorithm a little difficult while the color of skin and ginkgo leave is similar. To segment the face region, we use morphological and topological algorithms, the process of this algorithm is shown as Fig.6.

- Step 1: Do the pre-processing:
  The input image is shown in Fig.7(a), first, we use Gaussian Filter as the pre-processing, after filtering, the image is shown in Fig.7(b). Then, transfer the image from RGB color space to YCbCr color space.
- Step 2: In YCbCr color space, we do Skin-color Modelling, the result is shown in Fig.7(c).
- Step 3: Topological description:
  Due to the similar color between ginkgo and skin, but, the structure of them is different, skin is smooth with obvious hole while the region of ginkgo is Loose and porous, to separate the area of skin and the ginkgo, we discard those regions whose Euler number less than 1 or more than 15, the result is shown in fig.7(d), this method performs well, most background is removed, the face is retained as expected.
- Step 4: Morphological operation
  As fig.7(d) shows, after topological operation, there are still some small dots with narrow slit and thin protrusions retained,we apply morphological open operation with a structuring element 'disk' (as fig.7(e) shows). Now, we use topological operation to eliminate those noises, the retained background noises are almost filtered, at the

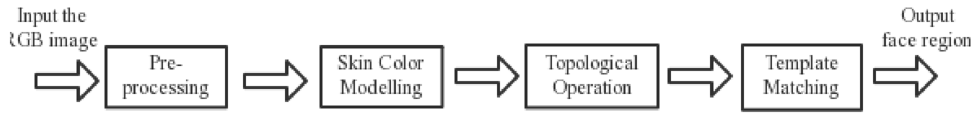


Fig. 6. Process of Image Face Segmentation.

same time, the shape of the human face is kept well. The result is shown in fig.7(f).

- Step 5: After skin color segmentation, we get a series of connected candidate face regions as fig.7(f) shows. Furthermore, we have to select suitable regions from all skin regions which could be potential human faces. Thus, to narrow down our search for human faces, we define a number of criteria:
  (1) In general, we set the height to width ratio in the range between 0.7 and 2. (2)To detect the face precisely, a template model is use to match the candidate areas. SSIM [7] is used to measure the similarity. The template is shown as fig.7(g). The value of SSIM are shown as TABEL I. We can find that the third of the candidate areas shows the maximum value, obviously, it is the face region and shows in fig.7(h).

TABLE I
THE SSIM OF CANDIDATE AREAS AND TEMPLATE

the ith of areas	$i=1$	$i=2$	$i=3$
ssim	0.2319	0.1769	0.2521

B. Image Human Body Segmentation

In the last section, image human face segmentation is very successful, in this section, we shall introduce our image human body segmentation in details. The process mainly includes edge detection, morphological operation, Image threshold segmentation. From the original image, we can find that the jacket is divided into two colors, red and white. The red part is an overall color block, which can be detected easily by edge operation. But the white part, there are many folds on the sleeves which cannot be divided from the background leaves, so we use image threshold segmentation since the white color has large value in a digital image.

- Step 1: Firstly, do the edge detection as fig.8(a) shows, from the image, we can find that the body does not have so much edge, but the boundary of the body is full of leaves who have fine edges. So, it is a good way to use dilate operator to "smooth" the leaves, the result is shown in fig.8(b).
- Step 2: To separate the edge from the background, we use erode operation as fig.8(c) shows.
- Step 3: Invert the binary image, we get fig.8(d), then keep the Euler number between -10 to 0, as fig.8(e) shows. The problem is, the arms are missing because there are many fine edges in the sleeves which are eliminated by dilate operation.
- Step 4: To solve the problem in Step 3, we use image threshold segmentation to segment the sleeves separately. Since the Sleeves are white, so turn the original image to Gray level and set the thresh to get the white sleeves. Then, add the white sleeves to fig.8(e), we get fig.8(f).
- Step 5: The outline of the human body has been roughly shown, now, we put the face segmented in the last section and do Close operation, the result is shown in fig.8(g).
- Step 6: Combine the original image with fig.8(g), we get the final image which segment the human body successfully.

In general, this algorithm is not that complicated, we apply this method to other images, the process is shown in fig.9. We can see that this algorithm is useful for the above specific pictures, but, there are many restrictions on the pictures, such as when the face has too many expressions or the hair covers the face, the face cannot be segmented by morphology algorithm, for the clothing, when the clothes are simple with sigle color, the segmentation is more easy, but for clothes that have difficult structures, there should be other segmentation method adding to this base algorithm.

IV. CONCLUSION AND PROSPECTION

In this paper, an image segmentation algorithm is introduced with simulation on computer work using the software MATLAB. From the performances, the human face and human body are segmented well separately. Different image is simulated and the result is good since the algorithm is a little different. In short, there is no image segmentation method that is applicable to any scene. We can only roughly summarize which method is relatively better for a certain type of image, and this is also relative. In practice, we need to combine actual Specific analysis of the use of theoretical knowledge.

REFERENCES

[1] B. Menser and M. Brunig, "Segmentation of human faces in color images using connected operators," Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348), Kobe, 1999, pp. 632-636 vol.3.
[2] X. Li, N. Roeder, "Face contour extraction from front-view images", Pattern Recognition, vol. 28, no. 8, pp. 1167-1179.
[3] S.H. Jeng, H.M. Liao, Y.T. Liu, M.Y. Chern, Jeng et al., "An efficient approach for facial feature detection using geometrical face model", Proceedings of the ICPR 1996, pp. 426-430, 1996.
[4] H.A. Rowley, S. Baluja, T. Kanade, "Neural Network-based Face Detection", IEEE Trans. Pattern Anal. Machine Intell, vol. 20, no. 1.
[5] Chiunhsiun Lin, Kuo-Chin Fan, "Triangle-based Approach to the Detection of Human Face", Pattern Recognition, vol. 34, no. 6, pp. 1271-1283, 2001.
[6] R L. Hsu, "Face Detection in Color Images [J]", Pattern Analysis and Machine Intelligence, no. 5, pp. 696-706, 2002.
[7] Z Wang, A.C Bovik, H.R Sheikh. Image Quality Assessment: From Error Visibility to Structural Similarity[J]. IEEE Trans Image Process, 2004, 13(4).

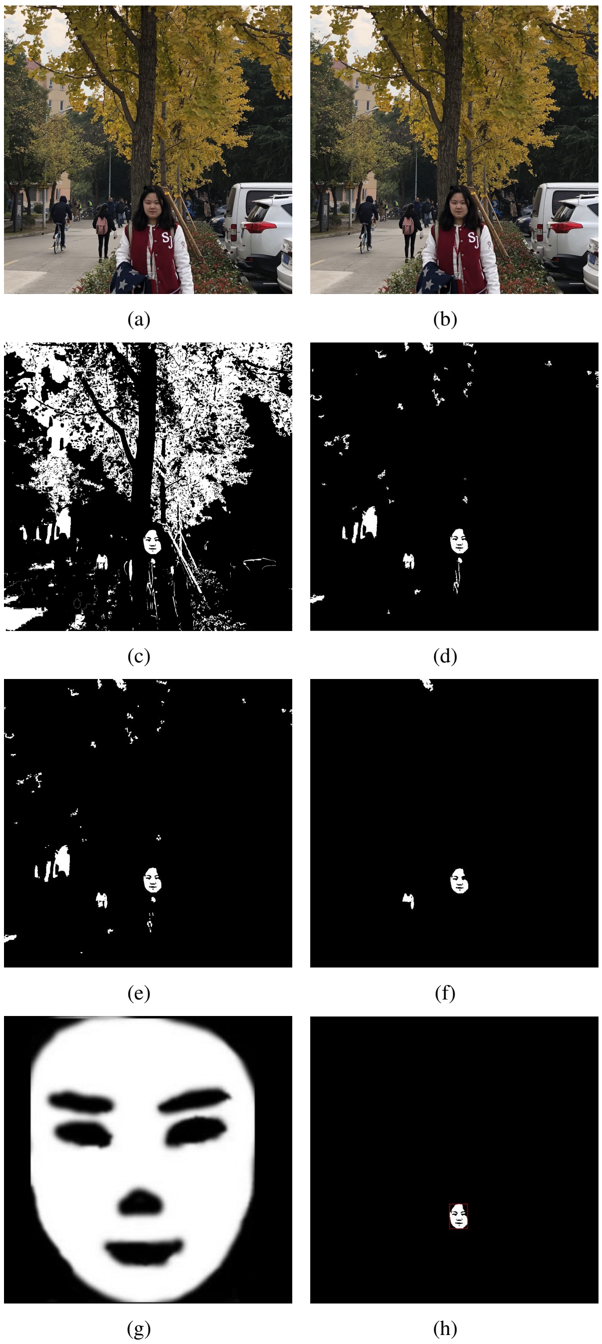


Fig. 7. Face Segmentation. (a)Original image,(b)Image after gaussian filter,(c) image after skin-color modelling of(b), (d)retain the Euler number between -10 to 0 of (c),(e)Do the open operation,(f)The candidate areas after finding the Euler number between -10 to 0 again,(g)Template,(h)Face region after doing template matching.

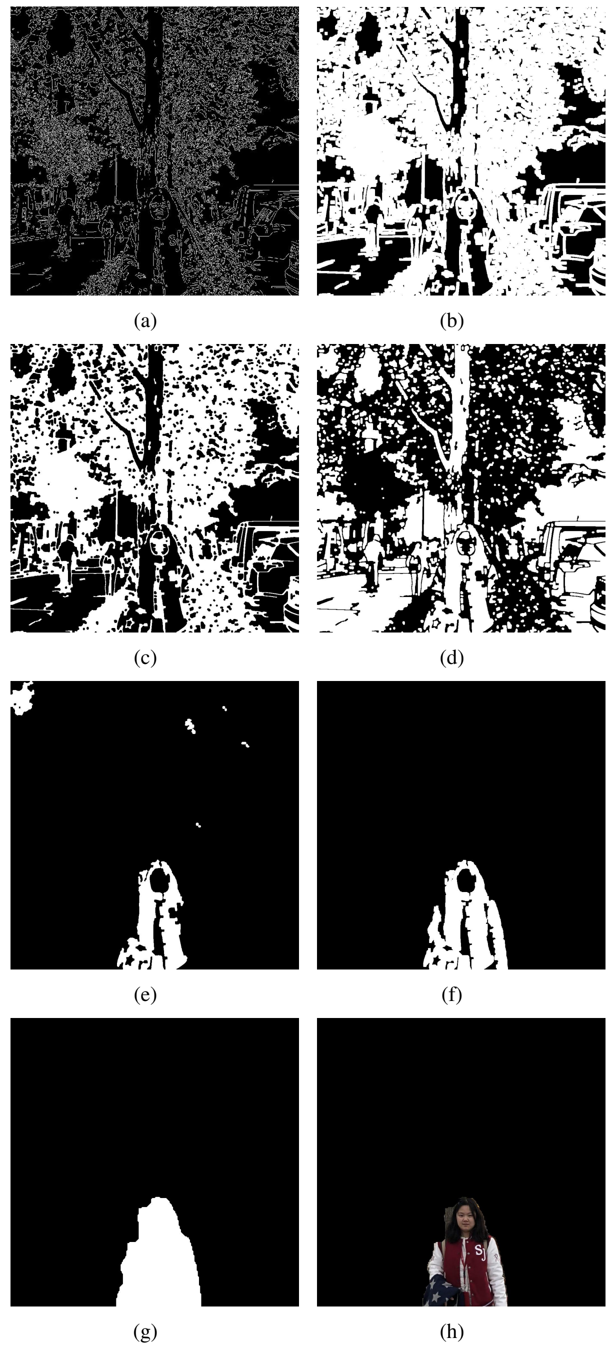


Fig. 8. Body Segmentation and the final results. (a)Edge discription,(b)Dilate operation with a structuring element 'disk', whose size is 8, (c) Erode operation with a structuring element 'disk', whose size is 4, (d)Do the opposite operation,(e)Combine with the arms, (f)Do the close operation. (g)The final image after segmentation.

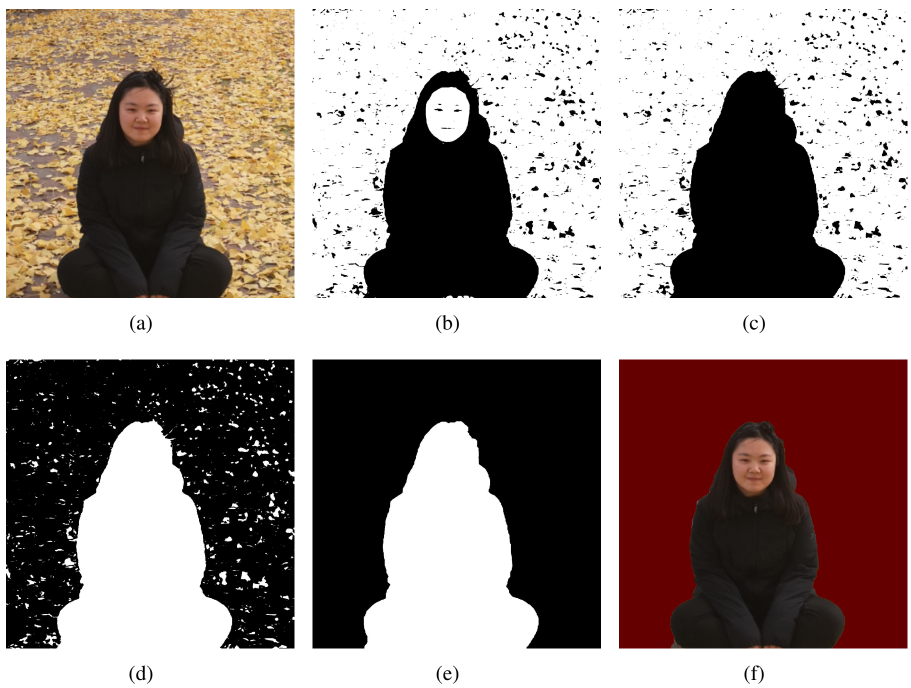


Fig. 9. Simulation for another image. (a)The original image, (b) Do the Skin-color modelling, (c)Topologial operation, (d)invert the binary image, (e) Morphological open operation, (f) The segmented image.