Face Detection for Automatic Avatar Creation by using Deformable Template and GA

Tae-Young Park*, Ja-Yong Lee**, and Hoon Kang***

* School of Electrical and Electronics Engineering, Chung-Ang University, Seoul, Korea
(Tel : +82-2-816-8234; E-mail: tzeropak@sirius.cie.cau.ac.kr)

** School of Electrical and Electronics Engineering, Chung-Ang University, Seoul, Korea
(Tel : +82-2-816-8234; E-mail: jalnans@sirius.cie.cau.ac.kr)

*** School of Electrical and Electronics Engineering, Chung-Ang University, Seoul, Korea
(Tel : +82-2-816-8234; E-mail: hkang@cau.ac.kr)

Abstract: In this paper, we propose a method to detect contours of a face, eyes, and a mouth of a person in the color image in order to make an avatar automatically. First, we use the HSI color model to exclude the effect of various light conditions, and find skin regions in the input image by using the skin color defined on HS-plane. And then, we use deformable templates and genetic algorithm (GA) to detect contours of a face, eyes, and a mouth. Deformable templates consist of B-spline curves and control point vectors. Those represent various shapes of a face, eyes and a mouth. GA is a very useful search algorithm based on the principals of natural selection and genetics. Second, the avatar is automatically created by using GA-detected contours and Fuzzy C-Means clustering (FCM). FCM is used to reduce the number of face colors. In result, we could create avatars which look like handmade caricatures representing user's identity. Our approach differs from those generated by existing methods.

Keywords: Face detection, Deformable Template, GA, B-spline curve, HIS color model, FCM, avatar

1. INTRODUCTION

There are many methods for creating avatars in internet. Most of them are done by the method based on image templates. But, it is true that we cannot obtain the user’s feature of appearance well with the template-based method. Also, it requires so many options for users to choose, and as a result, it makes users tired. This paper proposes an effective and interactive method that automatically recognizes user’s features of appearance from a 2D Image of the user face, and creates an avatar by using these features. Therefore, the created avatars have the user appearance, interact with the user, so as to express the user’s subjective sensitivity. To detect a feature from an input image, the search for features in all regions is required. Here, reduction of search target regions as well as efficient usage of the computation time and resources are important issues.

We use the HSI color model to reduce distortion of the tone of color caused by the light conditions. It defines the color of skin in the HS plane, and then applies it to detecting a face region.

In search of face contours, it is difficult to search for such contours by using only the edge values in a complex background. So it is required that the search method may obtain an object contour even if serious variations present.

We can increase accuracy by using a deformable template consisting of B-spline snakes in the feature extraction to which we apply the genetic algorithm for efficiency improvement.

Basically, our method of creating avatars consists of two stages automatically and interactively. The first stage is detecting facial feature points from the user’s input picture. In the second one, we create avatars by referring to these points and fine-tune them in order to be enhanced during the interaction with the users.

This work was supported by grant No. 20050021 (Development of the Key Technology for Intelligent and Interactive Module) from the project of Developing SIC(Super Intelligent Chip) and its Applications under the program of Next generation technologies in 2000 of [Ministry of Commerce, Industry and Energy]

The first step in detecting user’s facial feature points from the picture is done by separating the background from the face area. The color information is used during the processing of the image. Actually, human’s skin color has a uniform color range. However, this method is not perfect due to the intensity changes and the shading. Therefore, we perform several preprocesses for the image. Once the facial region has defined, detection of facial feature points is performed on the gray scale image because the color values of the image can be distorted by many causes such as illumination and input device’s color characteristics.

The detected edges have much information on the image. We rather perform the modified edge detection than the classical methods because the traditional edge detection methods have weakness under noises.

By using these feature points, our application draws an original avatar and its variations to be shown for the user. And then, it interacts with the user to add the user’s sensitivity to the generated avatars. In evolutionary computations, the important thing is how to define a fitness function. It is proper that the user evaluate the fitness function subjectively.

2. FACE REGION DETECTION

The input image of the RGB color model is converted to the HSI color model. It is robust against variance of light changes.

2.1 RGB model VS HSI model

Generally, a stored image is represented with RGB color model. One of the problems in image recognition is that each component contains light's intensity, so the intensity values change a lot.

There are frequent occasions in such ways not only that light is distributed irregularly, but also that target of body is shaded by the light. In this case, if we use HSI color model, we may have some independence from shadow or light effects.

Intensity is used for measuring the degree of light's powerfulness in the HSI model. It is regarded as one independent component, and therefore we can design robust algorithm against variance of light if we neglect intensity. The picture below represents this concept.

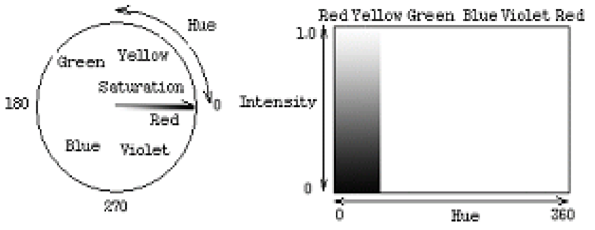


Fig. 1 HSI color model

2.2 Skin color definition in H-S plane

Detecting facial area is launched by defining the region suitable for skin candidate area. However, it is almost impossible that skin area can be defined as one region in Hue and Saturation due to changing lights and environments. So, in this paper, we determined a standard face color in H-S plane.

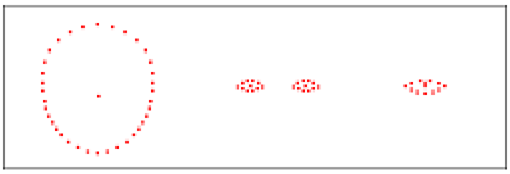
We extract skin color from various images in different conditions to define a standard skin color cluster. The center point in H-S plane is defined as a standard skin color and the skin color is defined inside the ellipsoid of the color cluster.

Using the skin color range, we could detect the facial skin.

Fig. 2 Detection of facial skin region

2.3 Definition of deformable templates

In this paper, object features are face, eyes, and lip. There are basic template control points based on B-spline snake.



(a) Face (b) Eyes (c) Lips
Fig. 3 Control points of the deformable template

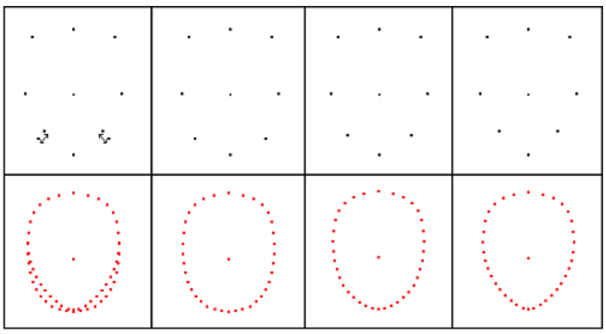


Fig. 4 Deformable of chin line with the templates

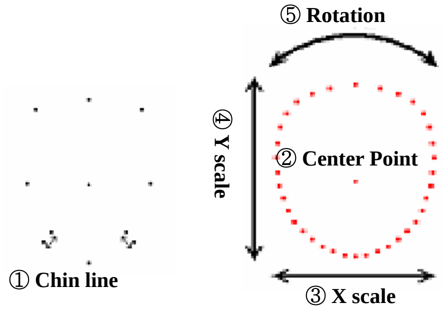
In case of the face contour, chin line control points are added in order to enhance the accuracy. To extract the face contour there 8 control point based on B-spline. There are two control point sets for chin line in them. Affine-transformed chin line control points have an influence on a face contour as shown in Fig. 4.

2.4 Deformation parameter

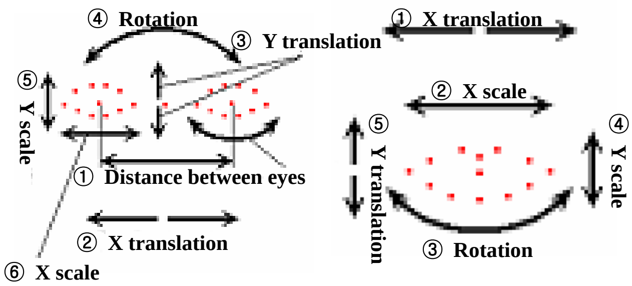
Each template is deformed independently. In the feature extraction, the processing step goes through face, eyes and lips. Definition of each parameter is as below. In case of the face contour, there are 5 parameters; X or Y axis translation, X or Y axis scale, and rotation.

In the eyes, there are 2 templates at left and right. But we change only one side because of symmetric shapes. Therefore a parameter is added for the distance between two eyes. So there are 7 parameters for the eyes; X or Y axis scale and rotation X or Y translation, rotations of the eye center points and the distance between two eyes.

In the lips, there are 5 parameters; X or Y translation, X or Y scale and rotation. Fig. 5 shows these parameters.



(a) Face



(b) Eyes (c) Lip
Fig. 5 Deformable template parameters

3. FACE DETECTION USING GA

Genetic Algorithm (GA) is one of the most powerful methods to solve an optimization problem. In this paper, we use GA to search the best solution for an input image.

3.1 Chromosomes

The chromosomes consist of parameters defined by the deformable template. Table 1~3 shows these chromosomes.

Table 1 Chromosome of face contour.

Phenotype		Genotype	
Parameter	Range	Form	Range
Chin line	-1 ~ +1 *	2Byte	0~65535
Center point	1 ~ Number of skin pixel	2Byte	0~65535
X axis scale	0.5 ~ 1.5	2Byte	0~65535
Y axis scale	0.5 ~ 1.5	2Byte	0~65535
Rotation	-10°~+10°	2Byte	0~65535

Table 2 Chromosome of eye contours.

Phenotype		Genotype	
Parameter	Range	Form	Range
Distance between eyes	-1 ~ +1 **	2Byte	0~65535
X axis translation	-1 ~ +1 **	2Byte	0~65535
Y axis translation	-1 ~ +1 **	2Byte	0~65535
Rotation from center point between eyes	-10°~+10°	2Byte	0~65535
X axis scale	0.5 ~ 1.5	2Byte	0~65535
Y axis scale	0.5 ~ 1.5	2Byte	0~65535
Rotation from center of eyes	-10°~+10°	2Byte	0~65535

Table 3 Chromosome of lip contour.

Phenotype		Genotype	
Parameter	Range	Form	Range
X axis translation	-1 ~ +1 **	2Byte	0~65535
Y axis translation	-1 ~ +1 **	2Byte	0~65535
X axis scale	0.5 ~ 1.5	2Byte	0~65535
Y axis scale	0.5 ~ 1.5	2Byte	0~65535
Rotation	-10°~+10°	2Byte	0~65535

*) Dependent on input image
**) Dependent on skin region

3.2 Fitness function

We use a tone of color and edge value for searching. We use 3 fitness functions for face contour detection.

First, we estimate a fitness by using the edge values.

$$f_{f1} = \sum_{i=1}^{57} \frac{(E_{-1i} + 2\times E_{0i} + E_{1i})}{(4\times 57\times 255)} \quad (1)$$

Second, we estimate a fitness using H, S values at the inner place of the template.

$$f_{f2} = \sum_{i=1}^{57} \frac{\sqrt{(H_i - H_{sfcolor})^2 + (S_i - S_{sfcolor})^2}}{(57\times D_{HS\max})} \quad (2)$$

Third, we estimate a fitness using H value at the outer place of the template.

$$f_{f3} = \sum_{i=1}^{40} \frac{|H_{15i} - H_{3i}|}{40\times 80} \quad (3)$$

Finally, total fitness of face contour is calculated by weights.

$$f_f = \frac{\omega_1\times f_{f1} + \omega_2\times f_{f2} + \omega_3\times f_{f3}}{\omega_1+\omega_2+\omega_3} \quad (4)$$

f_{fi} : Fitness function of face contour

E_{0i} : Edge value

E_{-1i}, E_{1i} : Edge value at side control value

H_i, S_i : H, S value

$H_{sfcolor}, S_{sfcolor}$: Standard face color

$D_{HS\max}$: Distance between standard face color and control point

H_{15i}, H_{3i} : H value

(a) f_{f1} (b) f_{f2} (c) f_{f3}

Fig. 6 Fitness measurement of face contour

The fitness measurement of the eye contour consists of 3 fitness functions.

First, we use the edge values to find an eye.

$$f_{e1} = \sum_{i=1}^{20} \frac{E_{ri} + E_{li}}{40\times 255} \quad (5)$$

Second, the gray values reflected inner place of the eyes.

$$f_{e2} = \sum_{i=1}^{N_{linepixel}} \frac{D_{rGi} + D_{lGi}}{N_{rLinePixel} + N_{lLinePixel}} \quad (6)$$

Third, we estimate the eye points of lower part to remove an eyebrow.

$$f_{e3} = \frac{(G_{ri} + G_{li})}{2} \quad (7)$$

Finally, a fitness of the eye contour is calculated by the weights.

$$f_e = \frac{\omega_1\times f_{e1} + \omega_2\times f_{e2} + \omega_3\times f_e}{\omega_1+\omega_2+\omega_3} \quad (8)$$

f_{ei} : Fitness function of eyes

$N_{lLinePixel}, N_{rLinePixel}$: Control point of side

G_i : Gray value of sample point

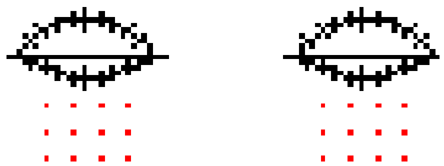


Fig. 7 Sample point for fitness measurement

In case of the lips, there are 3 fitness functions.
First, we use an edge value to measure a fitness.

$$f_{m1} = \sum_{i=1}^{24} \frac{E_{mi}}{(24 \times 255)} \tag{9}$$

Second, the tones of color near the mouth different from the skin color.

$$f_{m2} = \sum_{i=1}^{24} \frac{1 - \sqrt{(H_i - H_{sfcolor})^2 + (S_i - S_{sfcolor})^2}}{(24 \times D_{HS\max})} \tag{10}$$

Third, the lip is more red than the skin color.

$$f_{m3} = \sum_{i=1}^{24} \frac{|R_{iin} - R_{iout}|}{25 \times 255} \tag{11}$$

Finally, we get a total fitness of the lip contour.

$$f_m = \frac{\omega_1 \times f_{m1} + \omega_2 \times f_{m2} + \omega_3 \times f_m}{\omega_1 + \omega_2 + \omega_3} \tag{12}$$

f_{mi} : Fitness function of lip

E_{mi} : Edge value

R_{iin} : Red value at inner point

R_{iout} : Red value at outer point

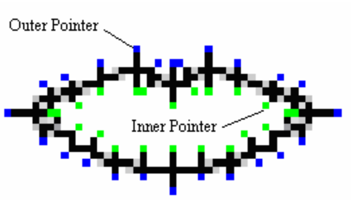


Fig. 8 Sample point for fitness measurement

4. AVATAR CREATION

4.1 Facial color simplification using FCM

It needs to simplify the color levels of the detected facial image for an avatar.

Only the skin region must be extracted in the facial region determined by the preceding procedure. It can be clustered with FCM on the basis of pixel intensity at this region. The pixels which belong to each cluster are simplified to the predefined color levels based on the degree of intensity.

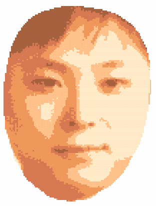


Fig. 9 Simplified image by FCM

4.2 Creation of face image

We make an image like a cartoon by using simplified image from FCM. In case of eyes, we use the contours obtained by the above procedure to make an eye image.

With the similar approach, we get a simplified lip image.

Fig. 12 represents the simplified image according to the proposed method.

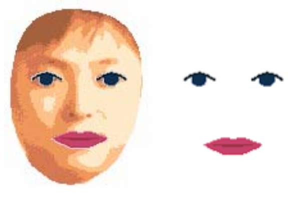


Fig. 10 Result image

5. SIMULATION

Fig .11 is flow chart of proposed method.

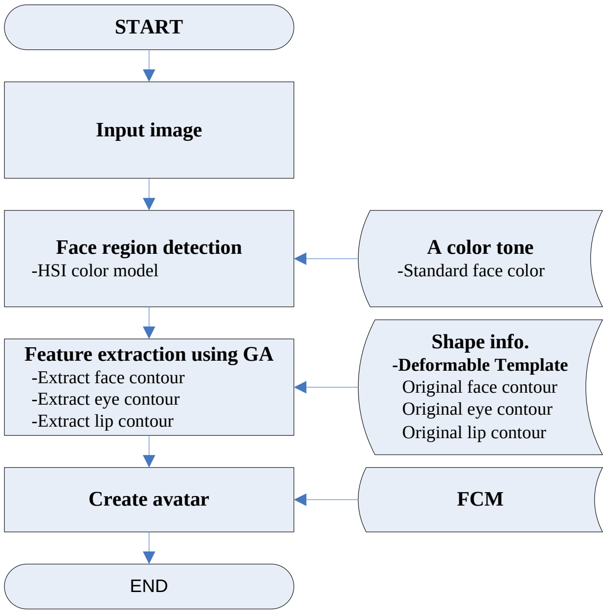


Fig. 11 Flow chart of proposed method

Fig. 12 is showing a result image made by proposed method.

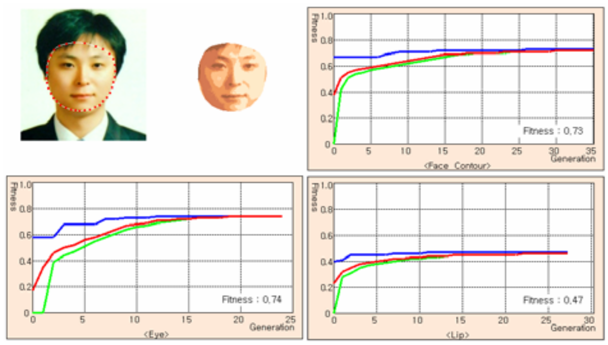


Fig. 12 Simulation result

In fig.12, the fitness and the face contour are shown. The result image shows that the face contour is refined well comparatively and the fitness graph shows that the total fitness value gets higher as the generation evolves. If the fitness value meets the condition, the searching process is terminated.

In case of the eyes and the lips, we know that the accuracy is limited because the templates are rigidly fixed in shape.

6. CONCULUTION

In the past, image recognition system used gray image and edge value but it is hard to find a good result in the uncertain environmental conditions, especially, light changes.

Therefore we are concerned about hue and saturation to find the facial feature sets. The deformable template is connected with the B-spline snakes. Then we find the facial features using genetic algorithm. As we use the deformable template to reflect various shapes of faces, the proposed genetic algorithm reduces the searching time. The genetic algorithm is the most powerful method to find the optimized feature in a global search region.

As a result, we find efficient methods of extraction in spite of the noise distortion or the lack of information. We hope to improve the proposed method and then apply it to many other uncertain or difficult situations.

REFERENCES

[1] Andrew Blake, Michalel Isard, Active Contours, Springer, 2000
[2] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, second edition, Prentice Hall, 2002
[3] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, Numerical Recipes in C++, CAMBRIDGE UNIVERSITY PRESS, 2002
[4] James D. Foley, Andries van Dam, Steven K. Feiner, John F. Hughes, Computer Graphics Principles and Practice, Addison-Wesley, 1997
[5] B.R.Moon, Genetic Algorithm, Dasung, seoul, 2001
[6] David E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley Publishing Company, Inc., 1989
[7] Hoon Kang, "Fuzzy System", Daeyoung, seoul, 2000
[8] J.S.R Jang, C.T. Sun, E. Mizutani, Neuro-Fuzzy & Soft Computing, Prentice Hall, Inc, 1997
[9] Saeed B. Niku, Introduction to Robotics Analysis, Systems, Applications, Prentice Hall, 2001
[10] Perez.P., Gidas.B. "Motion detection and tracking using deformable templates" Image Processing, 1994. Proceedings. ICIP-94., IEEE international Conference Volume 2, 12-16 Nov. 1994 pp272-276 vol.2
[11] Silsbee.P.L "Motion in deformable templates", Image Processing, 1994. Proceedings. ICIP-94., IEEE international Conference Vol.1, pp.323-327, 1994