Face Detection and Facial Feature Extraction Using Color Snake

Kap-Ho Seo, Won Kim, Changmok Oh and Ju-Jang Lee

Dept. of Electrical Engineering and Computer Science,
Korea Advanced Institute of Science and Technology,
373-1, Kuseong-dong, Yuseong-gu, Taejeon, 305-701, Korea
E-mail: {khseo, kimwon, cmoh}@odyssey.kaist.ac.kr jjlee@ee.kaist.ac.kr

Abstract

Face detection and facial feature extraction is a necessary first-step in face recognition systems. These tasks are one of the visual tasks, which humans can do effortlessly. However, in computer vision terms, this task is not easy.

This paper presents new active contour model using color information for extracting facial features And, we describe a new algorithm for detecting human faces and facial features, such as the location of the eyes, nose, and mouth. The algorithm is composed of three main parts: the face region estimation part, the detection part and the facial feature extraction part. In the face region estimation part, images are segmented based on human skin color. In the face detection part, the template matching method is used. And, in the facial feature extraction part, new proposed algorithm called "color snake" is applied to extract the facial feature points within the estimated face region.

1. Introduction

Automatic detection and recognition of faces and facial expressions from still images and video is an emerging application. A complete facial image processing system should be able to: 1) detect faces in a given images, 2) extract facial features, 3) recognize people, 4) describe facial expression. In this paper, we will focus on the two problems: 1 and 2.

Detecting and segmenting human faces in color images is becoming increasingly important because of developments in video coding and transmission, and the natural human-machine interaction. A general statement of the problem can be defined as follows: Given a still or video image, detect and localize an unknown number of faces. The solution to the problem involves segmentation, extraction and verification of faces and possibly facial features from an uncontrolled background. A face detection system should also be able to achieve the task regardless of illumination, orientation. To date, several approaches have been published for the detection of facial regions using color and some other information.

Use of KL transform has also been proposed in [1] where the inherent symmetry of the face is incorporated in an "eigenpicture" representation to improve the localization. However, the localization performance decreases quickly with changes in scale and orientation. Deformable template models have been utilized to locate the eyes and mouth. In addition, active contour models were employed in [2] to capture the eyebrows, nostrils and face. These techniques rely heavily on near perfect segmentation or edge detection. Furthermore, the extracted contours are highly dependent on the initialization of the snake or active contour model, and on the parameters involved in defining these models. Sung *et al.* [3] described an approach for face detection by utilizing 19*19 windows patterns and exhaustively searching a given image at all possible scales. Colmenarez *et al.* [4] introduced an algorithm for detecting faces based on a symmetry measurement that is designed to locate the center line by means of correlation. Chen et al. [5] proposed an algorithm for extracting human faces from color images. The algorithm determines the location of the face by first extracting the skin color regions, and then matching these regions, using a fuzzy pattern matching algorithm, to face models at multiple scales.

The proposed face recognition system is composed of three main parts: the face region estimation part, the detection part and the facial feature extraction part. In the face region estimation part, images are segmented based on human skin color. In the face detection part, the template matching method is used. And, in the

0-7803-7369-3/02/$17.00 © 2002 IEEE

facial feature extraction part, new proposed algorithm called "color snake" is applied to extract the facial feature points within the estimated face region.

This paper also presents a new active contour model using color information. Using color snake, the main facial feature such as eyes, mouth, eyebrows. These features are very important for analyzing the facial expression. So, we will consider the efficiency of color snake for facial expression recognition system.

The organization of the paper is as follows. Section 2 gives an difference between original active contour model and proposed contour model. Section 3 introduces the overall system of our emotional robot. Section 4 describes the proposed face recognition system. In section 5 experimental results are analyzed in detail. Conclusion and future work are addressed in section 6.

2. Proposed Color Snake Models

Kass [7] introduced snake concept in1987. Kass proposed a contour, which is deformable curve, some kinds of energies and solutions based on variational approach.

2.1 Brief Review of The Snake Models

The energy functional, which is minimized, is a weighted combination of internal and external forces. The internal forces emanate from the shape of the snake, while the external forces come from the image and/or from higher-level image understanding processes. The snake is defined parametrically as $v(s) = [x(s), y(s)]$, where $x(s)$, $y(s)$ are x, y coordinates along the contour and $s \in [0, 1]$. The energy functional to be minimized may be written as

$$E^{*}_{snake} = \int_0^1 E_{snake}(v(s))ds$$

$$= \int_0^1 \{E_{\text{int}}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))\}ds \tag{1}$$

where E_{int} represents the internal energy of the spline due to bending, E_{image} denotes image forces, and E_{con} external constraint forces. Usually, $v(s)$ is approximated as a spline to ensure desirable properties of continuity.

The minimum energy contour is determined using variational calculus. Using the snake models has proved to be a very attractive and efficient method, but it still encounters several problems. First, the contour shrinks due to the effect of internal energy when it is not subject to any external forces, and, consequently, it finds its own equilibrium at a point or a line. Secondly, control points on the contour can navigate along the contour as well as along the direction perpendicular to it, and, thereby, the points tend to be bunched together in parts of the contour where the image forces are higher. Thirdly, the results obtained by the snake model are considerably sensitive to the selection of the parameters associated with the model as well as the initial position of the contour. The other critical disadvantage of the snake models is that they do not utilize the prior knowledge about color information, which often plays a important role in improving the detection performance in the presence of some disturbing image features. The new contour model proposed in this paper is attempted to solve the fourth problem and apply to facial feature extraction problem.

2.2 Proposed Color Snake Models

Energy functional for color snake is modified from Kass' snake. As figure 1, we have designed snake's node having two components: outside and inside color. These components are compared to original image which snake's node lays. This makes facial feature extraction process faster and more accuracy.

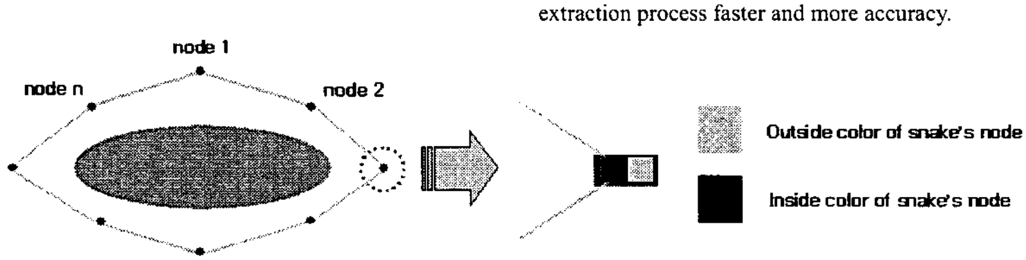


Figure 1: *Color snake models*

Newly proposed energy function is as follows:

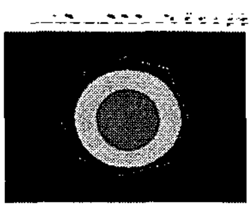
$$E^{*}_{color-snake} = \int_0^1 E_{color-snake}(v(s))ds$$
$$= \int_0^1 \{E_{\text{int}}(v(s)) + E_{image}(v(s)) + E_{node}(v(s))\}ds \quad (2)$$

$$E_{node} = \sum_{i}^{inner,\ outer} w_i \left(O_i(v(s)) - I_i(v(s))\right)^2 \quad (3)$$

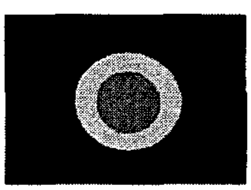
where w_i is weight for inner and outer of snake node, $O(v(s))$ is image color intensity at $v(s)$, $I(v(s))$ is snake node color intensity at $v(s)$.

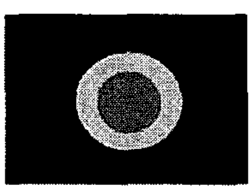
2.3 Test Simulation

Figure 2 shows the simulation results of original snake models and proposed color snake models. Fig 2.(a) is the initial snake position. Eight snake nodes are used for test simulation. Original snake model shrinks to first boundary like fig 2.(b). But newly designed color snake shrinks to second boundary like fig 2.(c). In this test simulation, each snake node has red color for inner color and orange for outer. So, snake node does not work for first boundary.



(a)





(b) (c)

Figure 2: *Test simulation (a) initial position (b) original snake model (c) proposed color snake model*

3. Overview of Robot System

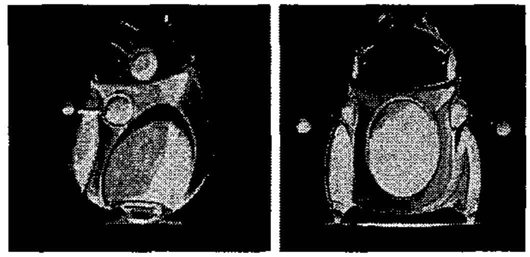


Figure 3: *Overview of emotional robot system*

Figure 3 is emotional robot "Rai". This system can function voice recognition and vision recognition for emotional interaction. So, Rai has two ears and one eye. He can hear and view his user and estimate user's emotional state and express his action with regard to user's emotion.

The detail spec of robot system cannot be described because he is under development.

DSP board was made for real time robot vision system. This board can gather 320*240 color images and process in real time. In each sampling time, robot can find face and analyze user's emotion. During the process, the new algorithm "color snake" is used for facial feature extraction.

4. Face Detection and Facial Feature Extraction

The proposed face recognition system is composed of three main parts: the face region estimation part, the detection part and the facial feature extraction part. In the face region estimation part, images are segmented based on human skin color. In the face detection part, the template matching method is used. And, in the facial feature extraction part, color snake is applied to extract the facial feature points within the estimated face region.

4.1 Part I: Face region estimation based human skin color

Face region estimation is performed pixel-by-pixel by determining if each pixel is "skin-colored" or not. In this paper, we use the normalized RGB to characterize human skin. We have chosen the following model parameter:

$$0.355 \le r \le 0.465,\ 0.27 \le g \le 0.363 \quad (4)$$

where $r = R/(R+G+B)$, $g = G/(R+G+B)$ represents the normalized Gaussian distribution of the red and green component, respectively.

(a) (b)

(a) (b)

Figure 4: *(a) Original image (b) face region estimation*

After skin color detection, a general region identification method, "component labeling" is used. This method can determine that extracted region is suitable for face or not. Regions with small pixels are excluded. Figure 4-(b) shows face region estimation result.

4.2 Part II: Face detection by template matching

Based on result of part I, in part II, we can detect face region by template matching. For template matching, face estimation region is preprocessed again. Extracted face estimation region is histogram-equalized, gray-scaled, and edge-detected. In the edge detection procedure, Sobel edge detection method is used. But, in this paper, we have chosen edge map using only vertical direction. In general, there are two main directions in our constrained face picture: horizontal and vertical.

Horizontal gradients are useful to detect the left and right boundaries of face and nose, whereas vertical gradients are useful to detect the head top, eyes, nose base, and mouth [8]. Figure 5 show that vertical edge map is more suitable for our algorithm than other edge maps. This edge map is used for template matching.

(a) (b) (c)

Figure 5: *Edge maps (a) both directions (b) horizontal direction (c) vertical direction*

Figure 6 shows template used in this paper. Template's size is 20*20 pixels. Template is composed of two parts: black region and multileveled gray region. Black region means whole colored region of template, which calculates the summation of intensity of edge in this region. And, multileveled gray region means the lower part of template, gray colored part in figure 6, which calculate the summation of intensity of lip color pixels. So we can represent the measure for template matching as follows:

$$E_{template} = \alpha \sum_{i}^{black} I_i + \beta \sum_{i}^{gray} Lip_i \qquad (5)$$

where I is intensity of edge map, Lip is intensity of lip color pixels, and α, β weights

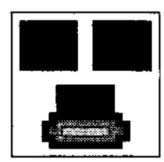


Figure 6: *Template*

Maximizing $E_{template}$ is solution of part II. For comparing with original image, template is scaled based on result of part I. Figure 7 shows the results of part II.

(a) (b)

Figure 7: *Face detection results*

4.3 Part III: Facial feature extraction by color snake

We propose color snake for facial feature extraction. Generally, the initialization of snake is very important for better performance of snake as describe in section 2. In this paper, snake initialization and facial feature extraction procedure is based on result of part II. The result of part II is initial state of color snake. As figure 7, yellow box is the initial position of color snake. Each snake point starts from boundary of yellow box. And, each node has prior color information of target feature: for example, if target is mouth, inner color is lip color and outer skin, otherwise if target is eyebrow, inner color is black, outer skin, and so on. Because we know the prior color information of human facial feature, we can expect better performance of proposed algorithm with respect to original snake models. The main result will be shown in section 5, experimental results.

5. Experimental Results

In this paper, color information is fully used to detect faces in complex color images. The robot vision system gathers 320*240 color image using CCD camera and find the face in real time.

In order to testify the efficiency of this method, experiment was carried out using facial images of 10 individuals in our lab. Each person provides five facial images with different viewpoint. Figure 8 shows experimental results: different person in complex background, different viewpoint of one person, and many people in one image.

Examples of figure 8(a) show frontal view of different person. And different face size was tested for efficiency of proposed system. Fig. 8(b) is test results of different viewpoint of one person. In this detection test, only up-down, left-right rotational images are included. The rolling face to image axis was not included. Fig. 8(c) shows the result of different number of person in one image. In this test, the system can detect all faces.

These results can be considered that if the color information of images is properly used, we can detect very easily human faces. But, if we use only color information to detect human faces, it is very sensitive to illumination condition. So, our proposed face detect system, template matching algorithm was also used. Because this algorithm is based on edge information, it is less sensitive to illumination condition.

(a)

(b)

(c)

Figure 8: *Face detection results*

(a) (b)

Figure 9: *Performance (a) original snake models (b) color snake models*

For facial feature extraction simulation, proposed color snake models was used. From the result of face detection, snake nodes are initialized.

This experiment shows an increase of performance accuracy with respect to original snake models. Figure 9 shows the comparison between original snake and proposed color snake. Original snake uses only edge information and cannot extract feature exactly. But since color snake uses both edge and color information, this algorithm can extract mouth exactly.

Figure 10 is final experimental result of color snake. In this simulation, eight snake nodes were used for each facial feature. We want to extract five parts: eyebrows, eyes, and mouth. So, totally 40 snake nodes were used. In this result, we can extract facial feature exactly.

Figure 10: *Facial feature extraction using color snake models*

But, this algorithm can be used only when we knows prior color information exactly. If illumination changes, color adaptation process is needed.

6. Conclusions and Further Works

Automatic detection and recognition of faces and facial expressions from still images and video is an emerging application. A complete facial image processing system should be able to: 1) detect faces in a given images, 2) extract facial features, 3) recognize people, 4) describe facial expression.

Detecting and segmenting human faces in color images is becoming increasingly important because of developments in video coding and transmission, and the natural human-machine interaction.

In this paper, we proposed new method "color snake" and fast face detection method. Color snake is suitable for exact facial feature extraction. Exact facial feature extraction is important for facial expression analysis. This algorithm will be used for emotional robot. Further work is the implementation of face recognition system that can detect, recognize and analysis faces.

7. References

[1] M. Kirby and L. Sirovich. "Application of the Karhunen-Loeve procedure for the characterization of human faces", IEEE Trans. on PAMI, vol. 16, pp. 689-700, 1994

[2] C. L. Huang, T. Y. Cheng, and C. C. Chen, "Color image segmentation using scale filter and Markov random field", Pattern Recognition, vol. 25, pp. 1217-1229, 1992

[3] K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection", Technical Report, 1994

[4] A. J. Colmenarez and T. S. Huang, "Frontal view face detection", SPIE, vol. 2501, pp. 90-98, 1995

[5] Q. Chen, H. Wu, and M. Yachida, "Face detection by fuzzy pattern matching", International conf. on computer vision, pp. 591-596, 1995

[6] Yanjiang Wang and Baozong Yuan, "Segmentation method for face detection in complex background", Electronics letters, vol. 36, pp. 213-214, 2000

[7] M. Kass, A. Witkin, and D. Terzopoulos, "Snake: Active contour models", Int. J. computer vision, pp. 321-331, 1988

[8] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates", IEEE trans. on PAMI, vol. 15, pp. 1042-1052, 1993