The Application of an Active Contour Method for Extracting Human Face Contour

[1]Yupeng Xia
[1]School of Mechanical and Electrical Engineering, Shenzhen Polytechnic,
Xili Lake, Shenzhen, Guangdong, 518055, China
E-mail: ypxia@szpt.edu.cn

[2]Feng Hu
[2]Shenzhen Graduate School, Harbin Institution and Technology,
Xili Lake, Shenzhen, Guangdong 518055, China
E-mail: hufeng_83@126.com

***Abstract*—The active contour is an effective method to perform image segmentation. But its key technology problem is, both the establishment of a mathematic model and the development of the calculating algorithm are excessively dependent on the initial contour position. In dynamic image analysis, the difference method and the fast template matching algorithm provide available solutions for dealing with the problem of sensitivity to initial position in active contour method. In this paper, the initial position of human body was first tested by using difference operation of dynamic image sequence and binary. Then the face position was initially located by fast template matching algorithm. And finally, the face position was located by active contour algorithm of "balloon" model to extract the face contour.**

***Index Terms*—Image processing, Active Contour, Template Matching, Dynamic Image Analysis.**

I. INTRODUCTION

Kass et al. constructed such a parametric spline curve [1] (Snake curve, also called as active contour): the curve is deformed and moved under the interaction of internal force, image force and external control force. The internal force keeps the curve smooth and continuous. While, the image force promotes significant image features of curve, such as line and edge, and so on. On the other hand, the external force pushes the curve to the most interested object contour where the local energy is in minimum value. They defined the object contour by using the three types of suitable energies. The mathematic model of total energy of object contour is defined as following equation:

$$E_{snake} = \int_0^1 [E_{\text{int}}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))]ds \qquad (1)$$

where $\bar{v}(s) = (x(s), y(s))$ denotes the parametric vector, and represents the mapping of unit parameter domain s∈ [0, 1] on image surface. E_{int}, E_{image} and E_{con} indicate the internal force, image force and external control force, respectively. The contour finally converges at the edge of interested target through minimized energy function. There are mainly two difficulties to achieve image segmentation by active contour algorithm. Firstly, the algorithm depends heavily on initial position, both the convergence of the algorithm and the convergence position are too sensitive to the choice of initial contour. In practical application, much manual intervention is usually required. Secondly, it is difficult for curve to converge at the depression part of contour, besides; the detected object contour is always not the expected one. A lot of researches have been done regarding these problems. But mainly of them were in two aspects, one of them was on the convergence method and rate of the algorithm, another was development of the suitable energy function and new mathematic models. For the former aspect, Sundaramoorthi and Yezzi reformulated the geometric active contour model by redefining the notion of gradient in accordance with Sobolev-type inner products [2]. Amini et al. discussed the dynamic programming as a novel approach to solving variation problems in vision [3]. Hu, Worrall, Sadka and Kondoz proposed a new scalable vertex-based shape coding scheme using temporal prediction [4]. Williams and Shan proposed an improved Snake model focusing on building detection from gray-level aerial image of high resolution [5]. For the latter one, Wang and Chen tried to combine Snake algorithm and DP algorithm to obtain the characteristic edge points of image [6]. Lu et al. constructed a new image segmentation algorithm which was called "balloon model". In this algorithm, the prior knowledge of image was used to establish the statistical information of image characteristics, with which the expanding force was set up and the internal balance force of the contour was proposed [7]. Farag et al. proposed an algorithm based on using four different types of deformable templates describing typical geometry and gray level distribution of lung nodules [8].

This paper focuses on the application of active contour in the contour detection and segmentation of moving

Manuscript received March 25, 2014, revised June 1, 2014; accepted July 1, 2014.
Published as submitted by the author(s).

doi:10.4304/jcp.9.11.2661-2665

face. The difference method and fast template matching method are used to deal with high dependency of Snake curve on initial contour. And the new model force is adopted to have deformable template converging to real contour approximately.

II. Methods

A. Deformation Face Template

Before the segmentation and extraction of face contour in motion image, a group of parametric curves are applied firstly to defining a popular face contour model on 2D plane. The curve can be expressed as), in which $_i=(x_i, y_i)$, i=1, 2, ..., N, denoting the coordinates of contour points in Cartesian coordinate system. Suppose that the contour is closed after segmentation, so it is obvious that $_{N+1} = \ _1$.

According to the statistical data of Chinese adult males' faces, the face is basically similar to an ellipse, with the ration of the major and minor axis between 1.15 and 1.35. So, some points on the ellipse are selected as contour points, shown in Fig. 1. The dotted portion in the figure is one ellipse model for reference. For the latter motion detection and image segmentation, several iconic points on the ellipse are fetched to form a facial basic contour which is called "prototype". With the change of parameter , the prototype also changes. And then the prototype becomes the deformable template. If the deformable template is given favorable constraints to drive the deformable template to be in control scope and finally converge at the edge of interested object or feature, the object segmentation is achieved.

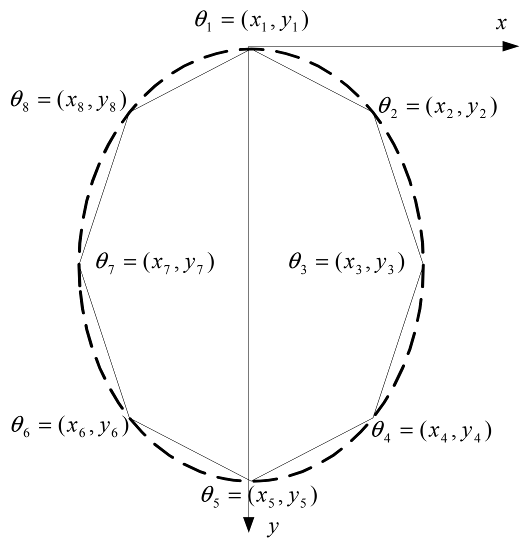


Fig. 1. Model of Human Face

In this paper, the curve parameters must be strictly controlled so that the facial deformable template can converge to real face. The various constraints and rules of facial deformable template are defined as the followings (suppose that the initial time is $|\theta_1\theta_5| = L$, $|\theta_3\theta_5| = M$):

$$-M/8 < x_1 < M/8, \quad -L/8 < y_1 < L/8;$$

$$M/4 < x_2 < 3M/4, \quad 0 < y_2 < L/5;$$

$$3M/8 < x_3 < 3M/4, \quad 3L/8 < y_3 < 5L/8;$$

$$M/4 < x_4 < 3M/4, \quad 7L/8 < y_4 < L;$$

$$-M/8 < x_5 < L/8, \quad 7L/8 < y_5 < 9L/8;$$

$$-3M/4 < x_6 < -M/4, \quad 7L/8 < y_6 < L;$$

$$-3M/4 < x_7 < -3M/8, \quad 3L/8 < y_7 < 5L/8;$$

$$-3M/4 < x_8 < -M/4, \quad 0 < y_8 < L/5.$$

B. Motion Detection

The selection of Snake initial contour should depend on motion information. In terms of the three images difference method of motion sequence image, the fore ground image is detected in this paper. If three frame images I_1, I_2, I_3 are scraped at three continuous moments $t-\Delta t$, t, $t+\Delta t$, these images contain the concerned motion information. The interval Δt of video capture is short, taking 1/25. After Gaussian low pass filter and smoothing fuzzy, the three frame images become $\tilde{I}_1, \tilde{I}_2, \tilde{I}_3$, respectively. The difference image is defined as the following:

$$D_i(i,j) = \left|\tilde{I}_1(i,j) - \tilde{I}_2(i,j)\right| \times \left|\tilde{I}_2(i,j) - \tilde{I}_3(i,j)\right| \quad (2)$$

where $D_i(i, j)$ denotes the pixel grey values of the i^{th} and j^{th} rows in the i^{th} difference image; × denotes multiply algorithm, factually it denotes the complementary operation of two frame images. The difference image provides very important position information and motion information of moving object in each frame image. Furthermore, the threshold segmentation is performed on difference image, so the difference image after threshold segmentation is got and defined as the following:

$$D_t(i,j) = \begin{cases} 1, & [D(i,j) > t_{low}, (i,j) \,\&\, t \,\&\, D(i,j) < t_{high}, (i,j)] \\ 0 \end{cases} \quad (3)$$

Therefore, the parts where the motion is not obvious or the motion is too vigorous in difference are removed.

C. Fast Template Matching

The fast Fourier template matching algorithm, proposed in literature [9] is used to reduce the redundant computational complexity of addition and subtraction in matching process through the establishment of accumulated summing value scale and accumulated square sum value scale. And the Fourier transform algorithm is applied to reducing the calculation time of multiplication. Generally, the normalized correlation coefficient (NCC) is defined as:

$$NCC(u,v) = \frac{\sum_{x,y}[I(x,y) - \bar{I}_{u,v}][t(x-u,y-v) - \bar{t}]}{\{\sum_{x,y}[I(x,y) - \bar{I}_{u,v}]^2 \sum_{x,y}[t(x-u,y-v) - \bar{t}]^2\}^{1/2}} \quad (4)$$

where I represents the image to be matched, $N\times N$; t denotes template image, $M\times M$; u and v are matching points. $\bar{I}_{u,v}$ and $\bar{t}$ are the mean value of template region, and the mean value of template image of matching points in image to be matched, respectively.

Define:

$$S(u,v)=\sum_{x=0}^{u}\sum_{y=0}^{v}f(x,y) \text{ and } \qquad S^2(u,v)=\sum_{x=0}^{u}\sum_{y=0}^{v}f^2(x,y) \tag{5}$$

where:

$$S(u,v)=I(u,v)+S(u-1,v)+S(u,v-1)-S(u-1,v-1)$$

$$S^2(u,v)=I^2(u,v)+S^2(u-1,v)+S^2(u,v-1)-S^2(u-1,v-1)$$

Let $t^{'}(x,y)=t(x,y)-\bar{t}$, and through Fourier convolution operation, it can be got:

$$NCC(u,v)=\frac{F^{-1}\{F(I)F^{*}(t^{'})\}}{\sqrt{\begin{array}{l}S^2(u+N-1,v+N-1)-S^2(u-1,v+N-1)\\-S^2(u+N-1,v-1)+S^2(u-1,v-1)\sqrt{e_t(u,v)}\end{array}}} \tag{6}$$

D. The Design of Snake Energy Function

According to the above steps, initial position of Snake curve can be initialized. Next each energy item of Snake needs to be designed, and the convergence position of Snake should be found out when the energy is the minimum. The expression of total energy is described as the follows:

$$E_{tatal}=k_1E_{\text{int}1}+k_2E_{\text{int}2}+k_3E_{image}+k_4E_{ball} \tag{7}$$

In this equation, k_1, k_2, k_3 and k_4 denote weight coefficients respectively.

a.Elastic energy E_{int1}

The elastic energy E_{int1} is defined as:

$$E_{\text{int}1}=\frac{D_i}{D_{\max}}=\frac{|\,\bar{d}-d_i\,|}{D_{\max}}=\frac{|\,\bar{d}-|\,v_i-v_{i-1}\,||}{D_{\max}} =\frac{|\,\frac{1}{n}\sum_{i=0}^{n-1}|v_i-v_{i-1}|-|v_i-v_{i-1}|\,|}{D_{\max}} \tag{8}$$

where v_i represents the i^{th} Snake point, d_i denotes the distance of adjacent Snake points, D_{max} indicates the maximum value of energy item in search window $n\times m$ of Snake point v_i . It is for balancing the weight of each force in energy function.

b. Rigidity energy E_{int2}

The rigidity energy E_{int2} is defined as:

$$E_{\text{int}2}=\frac{C_i}{C_{\max}}=\frac{|v_{i-1}-2v_i+v_{i+1}|}{C_{\max}} =\frac{(x_{i-1}-2x_i+x_{i+1})^2+(y_{i-1}-2y_i+y_{i+1})^2}{C_{\max}} \tag{9}$$

where C_{max} represents the maximum value of energy item in search window $n\times m$ of Snake point v_i.

c. Image energy E_{image}

The image energy E_{image} is:

$$E_{image}=\frac{G_{\min}-G_i}{G_{\max}-G_{\min}} \tag{10}$$

Where G_i indicates the mode of gradient vector of v_i point, G_{max} and G_{min} represent the maximum value and minimum value of energy item in search window of snake v_i. Obviously, $E_{image}\in$ [-1, 0] when G_i is the maximum one among the nine point, E_{image} is the minimum. At the same time, in order to balance the differences of image in each energy item, when the difference of G_{max} and G_{min} is small, let the difference be a suitable and fixed constant.

d. Partial balloon force E_{ball}

The partial balloon force E_{ball} is:

$$E_{ball}=\begin{cases}1 & when \quad (i,j)\in\Omega_{out}\\0 & when \quad (i,j)\in\Omega_{edge}\\-1 & when \quad (i,j)\in\Omega_{out}\end{cases} \tag{11}$$

where Ω_{out} denotes the external of motion object region; Ω_{edge} indicates the edge region of moving object; Ω_{in} denotes the internal of moving object. Because what the Snake algorithm deals with is the difference image which has been treated by binary, the pixel values in motion object region are 256. If it is 8-bit gray image, the average gray value in motion object region should be evaluated.

E. Normal Optimization

It has been talked about above that there are generally 3 methods for the minimization of Snake energy function. The greedy algorithm, the quickest one, updates each control point by iterative method. For each control point, the energy of each point in its neighbors is calculated, and the point with the minimum energy is the contour point. When the current control point is calculated, the other control points are all regarded to be relative optimal positions. The update of all control points are achieved after iteration being completed. Therefore, the iteration does not stop until all control points do not move through multi times of iterations. At this moment, the position of control points can be thought as the positions of contour points, actually the control points converge towards the direction of the real contour. Thus, with the normal optimization, it can reduce computational complexity through seeking for the minimum value of contour point in the normal direction of control point (as shown in Fig.2).

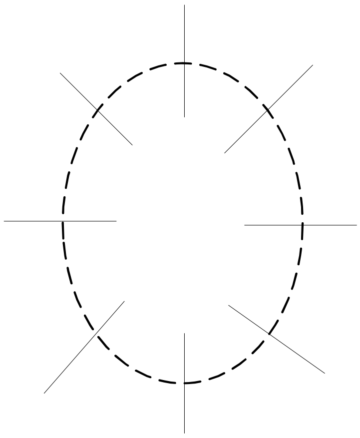


Fig. 2. The normal direction of control point.

III. ALGORITHM REALIZATION

Firstly, the moving human region in image sequence is detected in accordance with three image difference and complementary operation. To reduce computational complexity, the double thresholds operation is performed on difference image and then the high frequency part and low frequency part are removed, and consequently the good binary image can be got. Next the optimum matching position with deformable template is found out in this binary image, and then the optimum matching position is the region where the face locates most probably. The initialized contour position of Snake curve can be got by expanding slightly this region. Several control points on initial contour are selected, and then by using Greedy algorithm[6], the point with minimum energy value along the normal direction of each control point is used to replace the current control point. When the energy has been calculating, it is supposed that other points have been on the optimum contour position. The energy value of each control point is calculated cyclically and updated continuously till the iteration is completed when all points stop moving. After the completion of each time of iteration, the B-spline curve is generated and then the new normal direction of each control point should be found out for preparing the next iteration. Figure 3 shows the flow chart of the whole algorithm. Firstly, continuous three frame images $I_{k-1}(i, j)$, $I_k(i, j)$, $I_{k+1}(i, j)$ are obtained from acquisition card, and then the difference operation is performed after Gaussian smoothing. $G(i, j)$ in figure represents the Gaussian smoothing convolution operator. Through double thresholds operation, the high frequency part and low frequency part in difference image are removed, and then the binary difference image $D_t(i, j)$in figure is got. Next the complementary operation is performed on the two binary difference images, so the foreground image can be got. Consequently, the image sequences become frames of foreground images. The approximate cross position of face contour can be found out by the fast Fourier template matching algorithm discussed above, and then the Snake curve contour is initialized. Eight control points on the initialized contour are fetched and the normal vectors at the 8 points are calculated. For each control point, the one with the minimum energy in neighbors is calculated along with the point's normal direction. Then it is judged whether the move location of the point meets for the rules that defined by former deformable template. If it meets, the current control point is replaced; otherwise, the point is thought to converge. All of the control points are completed one time of calculation is called one time of iteration. After each time of iteration, the number of moving control points is recorded. If the number of control points which stop moving is more than thresholds, the Snake curve is thought to converge, stopping iteration. Otherwise, the new cubic curves are fitted on all current control points and meanwhile the positions of control points should be adjusted. The normal vector of new control point is calculated, preparing for the next cycle.

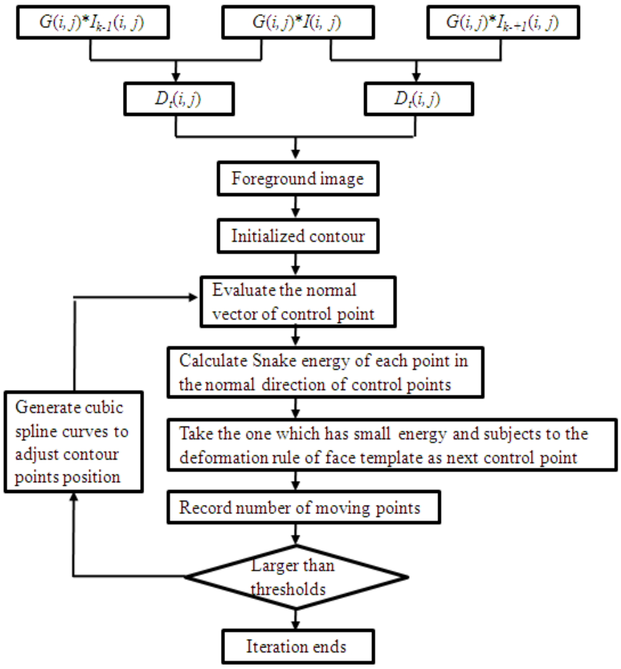


Fig. 3. Flow chart of the algorithm

IV. EXPERIMENTAL RESULT AND ANALYSIS

The new algorithm in this paper is applied to moving object detection and tracking system, and then the initial contour position of face is got through former treatment. Eight points in ellipse are fetched as control points, and then the point with minimum energy is found out along with the direction of the point to replace the current point. However, the motion of these points much meet for the rules fitted before. The original image is shown in Fig. 4 (a), and the template matching image is shown in Fig. 4 (b). The image is segmented with Snake algorithm, shown in Fig. 4 (c). The size of all images is 192×144.

In Snake algorithm, $k_1 = 1.0$, $k_2 = 0.4$, $k_3 = 4.6$, $k_4 = 3.3$. In this algorithm, the control points are less, so the weights of k_1 and k_2 in energy function are small. The algorithm does not focus on the gliding property of obtained contour curves, but the rapidity and stability. In the fast template matching method, there is no fixed template in this paper. After each time of iteration, the current target matching image region is taken as next frame template. Due to the short interval of acquisition time, the most matching target position can always be found out in the next frame image.

(a) The original image

(b) The template matching image

(c) Result based on Snake algorithm

Fig. 4. Experimental result

V. CONCLUSION

Snake algorithm has been rarely applied in dynamic image analysis. One of its key problems is that its sensitivity on the initial position limits its application. In practical operation, the manual intervention is usually used as follows: the contour initial control point-the Snake point is selected manually. Additionally, according to the characteristics of this system, the "balloon" model is improved to have the contour approaching target edge. It is found the ill-convergence of contour is likely to occur by only taking image force as external binding force, if the balloon force is not adopted.

ACKNOWLEDGMENT

This work was preformed in the school of mechanical and electrical engineering, Shenzhen Polytechnic. It was supported in part by the foundation of Shenzhen Bureau of Science Technology & Information (05KJfc050).

REFERENCES

[1] Kass M, Witkin A, Tezpoulous D. Snake, "Active contour models," *Int.J.of Comp.* Vision, 1988,1(4), pp321~331.

[2] Ganesh Sundaramoorthi, Anthony Yezzi, "Sobolev Active Contours," *Int.J. Of Comp.* Vision, 2007, pp345~365.

[3] Amini, Weymouth, Jain, "Using Dynamic Programming for Solving Variational Problems in Vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, pp855~866.

[4] Mingyou Hu, S. Worrall, A. H. Sadka, A. M. Kondoz, "An efficient Scalable object contour tracking scheme and its application for video segmentation," *2004 IEEE 6th Workshop on Multimedia Signal Processing*, 2004, pp155~158.

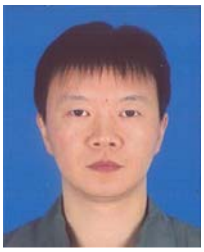
[5] Donna J. Williams, Mubarak Shan, "A Fast Algorithm for contours and curvature Estimation," *CVGIP: Image Understand,* 1992, pp14~26.

[6] Wang Heng, Chen Shaoqiang, "Image Boundary Detection Algorithm Based on Snake Model," *Friend of Science Amateurs,* 2007, No. 11, pp134~136.

[7] Lu Peizhou, Lai Shengli, Hu Rong and Chen Jiayang, "Image Segmentation Algorithm Based on Snake Model of Local Statistical Characteristic Restriction," *Journal of South Chian University of technology* (*Natural Science Edition*). 2007, Vol 35, No. 9, pp36~39.

[8] Aly A. Farag, Ayman El-Baz, Georgy G. Gimel ' farb, Robert Falk, Stephen G. Hushek, "Automatic Detection and Recognition of Lung Abnormalities in Helical CT Images Using Deformable Templates," *Medical Image Computing and Computer-Assisted Intervention,* 2004, pp560~566.

[9] Shinichiro Omachi, and Masako Omachi, "Fast Template Matching With Polynomials," *Image Processing, IEEE Transactions.* Volume: 16, 2007.



Yupeng Xia was born in Yunnan province China in May 1968, received the M. E. and Ph. D degree both from the School of mechanical engineering, Tohoku University, Sendai, Japan in 1996 and 1999 respectively.

He was a research fellow in the Virtual System Laboratory at Gifu University, Gifu, Japan during 2000-2002. Currently, he is an associate professor in the School of Mechanical and Electrical Engineering, Shenzhen Polytechnic, Shenzhen, Guangdong, China. His present research interests are the intelligent control of mechanical systems and the development of vision systems.

Feng Hu was with the Shenzhen Graduate School, Harbin Institution and Technology, Xili Lake, Shenzhen, Guangdong 518055, China (e-mail: hufeng_83@126.com)

.