

Hand Gesture Tracking Based on Improved Level Set Algorithm

N. G. Yu

Department of Electronic information and Control
Engineering
Beijing University of Technology
Beijing, China

M. Guo

Department of Electronic information and Control
Engineering
Beijing University of Technology
Beijing, China

F. F. Mo

Department of Electronic information and Control
Engineering
Beijing University of Technology
Beijing, China

X. G. Ruan

Department of Electronic information and Control
Engineering
Beijing University of Technology
Beijing, China

Abstract—The tracking and extraction the hand pose is the previous work of gesture analysis. In this paper, a mixed Gauss model based level set contour extraction algorithm is presented. The algorithm combined with the mean shift algorithm to achieve the extraction of hand contour dynamically in a video sequence. This method first get the initial contour by color information, and build Gaussian model of the inside and outside of the outline of the image features by using Gaussian mixture model. Depending on the characteristics of the Gaussian distribution and combining with active contour algorithm for image segmentation, then obtaining hand movement position by mean shift algorithm, we achieved gesture tracking. The test results show that compared to pixel-based active contour algorithm, this method has greater stability, which can get more accurate extraction region and achieve accurate contour extraction and gesture tracking.

Keywords—active contour; mean-shift algorithm; edge detection hand tracking

I. INTRODUCTION

Image is the source of most of the information about human activity. To greatly reduce the complexity of analysis or identification of the target, image analysis first to perform image segmentation and extract the target region[1]. Image sequence adds the process of moving target analysis, this includes target acquisition, target identification, target tracking, object understanding and so on.

Active contour model was presented as an effective method of image segmentation. Active contour model is computational efficiency and simplicity, especially for modelling and extracting deformation profile of arbitrary shape, so it has become one of the most active research topics in computer vision[1]. Level set algorithm raised by Sethian and Osher in 1988[2] was one method for solving geometric active contour models. The use of Level Set Methods making geometric active contour model can overcome many of the shortcomings of the parameter active contour models, such as processing topology change of curve, be insensitive to the

initial position of the active contour, only one stable numerical solution and so on[3].

Compared to traditional parametric active contour model, level set algorithm has strong topology adaptive and remains closed in the process of evolution. Although the level set method is a good motion interface tracking algorithm, there are still many shortcomings, such as the initial contour algorithm impacts the computing speed, surface evolution may not be stable, texture image segmentation is not satisfactory.

Lankton[6] proposed a model based on local regional framework to guide the evolution of the profile curve, which named Local Chan-Vese model. But this method has weakly boundary curves driving, so that will not get satisfactory segmentation. Aselles[4] and Malladi[5] proposed a level set segmentation model based on edge, such model considers the mean curvature motion of curve and target edge information as the main driving force to construct level set function so as to boot partition curve movement. The model which relied on edge information and be conducted to local minima.

To further enhance the performance of Level Set Model, this paper firstly take advantage of human skin color information, adaptively extract hand initial contour information. With level set function to achieve effective texture image segmentation, this article takes advantage of Gaussian mixture model to improve the energy function and gets the global information of the image. In the image sequence, the method combine the contour extraction algorithm with the mean shift motion tracking algorithm and automatically extract object contour information in tracking process, we achieve hand contour extraction and gesture tracking at the same time.

II. THE MODEL OF LEVEL SET

The initial level set comes with randomness in traditional methods, the quality of the initial contour direct influence the effect of the subsequent evolution of the level set. This paper mainly utilize color information to get hand initial contour,

while improving the energy function, according to the global information of the image, make it has a global optimal solution.

A. Color-based gestures initial contour extraction

For the handsegmentation form image, the main feature is the colortarget. We can use color information to obtain initial hand outer contour. Ellipsecolor model based on YCbCrcolor space showed good performance and efficiency in gesture detection and segmentation. This method calculate each point in the image so that can get the skin area. YCbCrcolor space can determine the hand region, the advantage of this model is that it can get a good skin area space through color contour region extraction.

B. Evolution equations based on the model of regional

Level set algorithm has strong topology adaptability, which can change its topology with the evolution of the level set function. Delimiting a curve C to represent zero level set, our job is to change C to get these two regions. The core idea of the algorithm is to implicitly express the two-dimensional evolution curve as the zero level set of a three-dimensional continuous function $\Phi_0(x, y, t)$, which is remaining zero to divide the the evolution region. Establish the following level set function equation through the initial outline:

$$\check{v}_0(x, y, t) = \begin{cases} -\rho(x, y) \in \wedge_0 - \partial \wedge_0 & (x, y) \in \partial \wedge_0 \\ 0 & \rho \wedge - \wedge_0 \end{cases} \quad (1)$$

RomeilSandhu[7] proposed a kind of Level Set Algorithm to split image, which defines a closed curve C to represent level set, by changing the shape and size of the curve or the level set function $\Phi_0(x, y, t)$. The best way to get the curve C is to minimize the evolution equation of level set function:

$$\frac{\partial \check{v}}{\partial t} = \nabla \check{v} E_{\text{image}} + \lambda \cdot \sigma(\check{v}) \cdot \text{div} \left(\frac{\nabla(\check{v})}{|\nabla(\check{v})|} \right) \quad (2)$$

Among them, the second item in the right side of the function (2) in order to theregion limit, the first use metric function based on the energy functional measuring distribution of gray values inside and outside the contour.

$$E_{\text{image}}(z, \check{v}) = \sqrt{\varepsilon \left\{ \left(\log \frac{p_{\text{in}}(z, \check{v})}{p_{\text{out}}(z, \check{v})} \right)^2 \right\} - \varepsilon \left\{ \left(\log \frac{p_{\text{in}}(z, \check{v})}{p_{\text{out}}(z, \check{v})} \right) \right\}^2} \quad (3)$$

In the function, p_{in} and p_{out} are probability distribution of internal and external contour of the grayvalue. Log likelihood function use to obtain the maximum difference gray valuedistribution between inside and outside the contour.

C. Gaussian mixture model

Gaussian model with a Gaussian probability density function to quantify things, expressing things as a model based on Gaussian probability density function, approximating the distribution of a variable with mixed Gaussian probability function. Gaussian mixture model is a linear combination of M numbers Gaussian density function, if divide the image into M classes by homogeneous regions of Gaussiandistribution,

themixed probability function of pixel distribution is defined as follows:

$$P(x) = \sum_{i=1}^M \pi_i N_i(x, \mu_i, \Sigma_i) \quad (4)$$

Among them, π_i represents mixing parameters, indicating a priori probability, μ_i is the mean of $N_i(x, \mu_i, \Sigma_i)$, Σ_i is the Gaussian distribution of variance, probability density function is as follows:

$$N_i(x) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left(-\frac{(1-\mu_i)^2}{2\Sigma_i} \right) \quad (5)$$

Parameter values (μ_i, Σ_i) in the Gaussian distribution process the samples X with K-means clustering algorithm, get the initial parameters of GMM, then obtain the final parameter values for each Gaussian distribution with continuous iterative EM algorithm. In image dataset, samples with different pixel values belong to the corresponding Gaussian probability distribution, the maximum Likelihood Estimation of model distribution parameters is got by EM algorithm.

D. Active contour algorithm based on Gaussian mixture model

Typically the color map is converted to gray scale in order to reduce the computational complexity of image analysis, but lost the image's color information. This article uses principal component analysis method for multi-channel color texture dimension reduction, with maximum retention image color and texture information.

The gray value of each pixel of a color image is as a three-dimensional vector, the image is regarded as image sample space composed by three-dimensional vector, as follows:

$$X(m, n) = \begin{bmatrix} X_1(m, n) \\ X_2(m, n) \\ X_3(m, n) \end{bmatrix} \quad (6)$$

To reduce the dimension of space image samples by the principal component analysis method. After subtract the mean value of the image, the eigenvectors and the i-th principal component of matrix can be present as following:

$$Y_i(m, n) = w_i^T X(m, n) \quad (7)$$

A three-dimensional color texture image will be projected onto a one-dimensional space, creating a new image, retaining the maximum original energy.

Considered previously obtained skin color area x as the data samples, then by the above Gaussian mixture model algorithm calculate the parameters of the Gaussian mixture model, construct the corresponding Gaussian mixture model. As the same, calculate the Gaussian mixture model parameters of the outer contour of the image area, z_{in} and z_{out} is respectively both inside and outside the contour corresponding to each pixel probability likelihood estimation, describing the distribution of statistical models of the gray scale image. This article established five Gaussian mixture model of the internal

and external contour, also compared the difference between internal and external contours probability distributions, for internal outline is of high similarity probability distribution, it was set higher weights. Compared with the level set method based on the pixel, this method is of greater stability and more accurate extraction region. According improved energy function by Gaussian mixture mode to get the global information about the image to have a global optimal solution. In which, two external distribution within the active contour models were:

$$p_{in}(z_{in}, \nu) = \int_{\wedge} \frac{K(z_{in}-I(x))H_{\epsilon}(-\nu)}{H_{\epsilon}(\nu)} dx \quad (8)$$

$$p_{out}(z_{out}, \nu) = \int_{\wedge} \frac{K(z_{out}-I(x))H_{\epsilon}(\nu)}{H_{\epsilon}(\nu)} dx \quad (9)$$

Among them, z is Gaussian mixture model distribution corresponding image, Heaviside function and Dirac impulse function is following:

$$H(z) = \begin{cases} 1 & z \geq 0 \\ 0 & z < 0 \end{cases} \quad (10)$$

$$\delta = \frac{d}{dz} H(z) \quad (11)$$

We define $K(z - I(x)) = \delta_{\epsilon}(z - I(x))$. When there are the requirements for taking distance distribution between the two Gaussian mixture model, commonly used similarity measure is KL divergence (Kullback-Leibler Divergence, KLD)[8][9]. Since there is no closed-form expression for solving, Kittipat Kampa proposed a Cauchy-Schwarz inequality to strike a similarity between the two Gaussian mixture model $q(x)$ and $p(x)$. The equation is as follows:

$$D_{CS}(q, p) = -\log \left(\frac{\int q(x)p(x)dx}{\sqrt{\int q(x)^2 dx \int p(x)^2 dx}} \right) \quad (12)$$

With D_{CS} we can get the similarity of both internal and external contours of the Gaussian distribution. Curve evolving according to equation (2), the curve C is as image segmentation curve when the two Gaussian distance inside and outside Contours is maximum.

III. REAL-TIME GESTURE TRACKING ALGORITHM

Mean shift algorithm [10] [11] is an effective non-parametric statistical iterative algorithm, which is to make every point drift to a local maxima of density function. In recent years, mean shift algorithm has been widely used in the field of computer vision[12].

Merging the improved level set method and mean shift algorithm in order to achieve image sequences tracking real-time gesture and achieve positioning regional hand. As follows:

1) According to the skin color information initialize the outline, and strike the centre of the contour y_0 .

2) Establish initialized search window $W(y_0, h_0)$ in the profile, y_0 is initial anchor point, h_0 is Window width,

and extract color probability distribution within this window, defined as $p(y_0) = \{p_u(y_0)\}_{u=1}^M$, and satisfy $\sum_u p_u = 1$:

$$p_u(y_0) = c_{h_0} \sum_{i=1}^{n_{h_0}} k \left(\left\| \frac{y_0 - x_i}{h_0} \right\|^2 \right) \delta[b(x_i) - u] \quad (13)$$

k is the profile function, b is the mapping function, $b(x_i)$ equals quantization characteristics of x_i , n is the number of pixels within the window W , c_{h_0} is the normalization constant.

3) Use coefficient Bhattacharyya ρ measure the degree of similarity of density of q and $p(y_0)$:

$$\rho(y_0) = \sum_{u=1}^M [q_u p_u(y_0)]^{\frac{1}{2}} \quad (14)$$

Here we analysis the function by the Taylor's development:

$$\rho(y_0) \approx c_1 + c_2 \sum_{i=1}^{n_{h_0}} w_i k \left(\left\| \frac{y_0 - x_i}{h_0} \right\|^2 \right) \quad (15)$$

$$w_i = \sum_{u=1}^M [q_u p_u(y_0)]^{\frac{1}{2}} \delta[b(x_i) - u] \quad (16)$$

c_1 And c_2 are constant, according to the mean shift algorithm:

$$y_i = \frac{\sum_{i=1}^{n_{h_0}} x_i k \left(\left\| \frac{y_0 - x_i}{h_0} \right\|^2 \right) w_i}{\sum_{i=1}^{n_{h_0}} k \left(\left\| \frac{y_0 - x_i}{h_0} \right\|^2 \right) w_i} \quad (17)$$

k' is the derivative of k .

4) Judge whether $\|y_1 - y_0\| > thld$, if yes, finish iteration, or continue iterating. Threshold value $thld$ is a small value close to zero, which new windows $W(y_1, h_0)$ and Mean Shift tracking vector can be gotten.

5) According to the offset translate the forward frame contour, according to adaptive contour models get new contour C .

6) Read the next frame, return to step 2)

Combining with the mean shift algorithm can reduce the number of iterations of active contour algorithm, thereby reducing computation time. Relying on accurate positioning of the mean shift algorithm to improve the accuracy of the hand contour extraction can track deformed objects in real-time image sequences.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In order to verify the accuracy of the improved active contour algorithm, first, use 140 * 140 pixel image verifies experiment, set the initial outline as a rectangle 55 * 75, extract hand contour of the static image. Contour iterative process shown in FIG. 1, the corresponding number of iterations were 30, 150, 200, 300, 390, and 480.

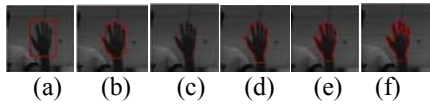


FIGURE I. IMPROVED LEVEL SET ALGORITHM TO EXTRACT STATIC HAND CONTOUR

As can be seen, the proposed algorithm, being combined with Gaussian mixture model and modeling global information can get a better contour extraction. First, set the initial outline as rectangular, the outline evolve and iterative according to energy force of inner and outer contours, gradually approaching the exact contours of the hand. The gesture contour information obtained by several iterations can describe gesture edge features more completely.

In order to verify the accuracy real-time and continuity of the tracking algorithm, we use a set of image sequence proceeding dynamic segmentation. The results shown in Figure 2. Position and shape of the hand in the image sequence are changed, the figure shows the effect of the track in frame 6,68,122,137,172, and 217,as which can be seen this method can track real-time contour accurately ,the colored rectangle is tracking window, the red area is extracted hand contour, as can be seen from the experimental results, our algorithm can track moving and changing contours of the hand well.

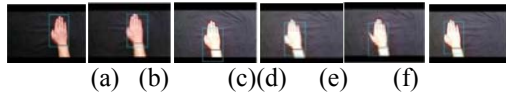


FIGURE II. GESTURE TRACKING RESULTS

V. CONCLUSIONS

In this paper propose level set algorithm based on Gaussian mixture model, and combine with the mean shift algorithm, achieving a gesture tracking process. Firstly get manpower initial contour by colorinformation, then use the mean shift algorithm positioning the hand area, move contour to the corresponding position, reduce the number of iterations of the algorithm. Finally,extract hand contour with improved level set algorithm achieving a gesture tracking. The test results show that this method has a strong relative stability, which more accurate extraction region, contour extraction and gesture tracking can be gotten.

REFERENCES

- [1] J. L. Liu, Study on Active Contour Model and Gaussian Model in Image Processing[D]. Xi'an University of Technology, 2012.
- [2] Osher S, Sethian J A. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations[J]. Journal of computational physics, 1988, 79(1): 12-49.
- [3] B. Chen, J. H. L., Active Contour Models on Image Segmentation: A Survey [J]. Chinese Journal of image and graphics, 2007, 12(1): 11-20.
- [4] Caselles V, Catté F, Coll T, et al. A geometric model for active contours in image processing[J]. Numerischemathematik, 1993, 66(1): 1-31.
- [5] Malladi R, Sethian J A, Vemuri B C. Evolutionary fronts for topology-independent shape modeling and recovery[M]//Computer Vision--ECCV'94. Springer Berlin Heidelberg, 1994: 1-13.
- [6] Lankton S, Tannenbaum A. Localizing region-based active contours[J]. Image Processing, IEEE Transactions on, 2008, 17(11): 2029-2039.
- [7] Sandhu R, Georgiou T, Tannenbaum A. A new distribution metric for image segmentation[C]//Medical Imaging. International Society for Optics and Photonics, 2008: 691404-691404-9.
- [8] Vasconcelos N. On the efficient evaluation of probabilistic similarity functions for image retrieval[J]. Information Theory, IEEE Transactions on, 2004, 50(7): 1482-1496.
- [9] Vasconcelos N, Ho P, Moreno P. The Kullback-Leibler kernel as a framework for discriminant and localized representations for visual recognition[M]//Computer Vision-ECCV 2004. Springer Berlin Heidelberg, 2004: 430-441.
- [10] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2003, 25(5): 564-577.
- [11] Cheng Y. Mean shift, mode seeking, and clustering[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1995, 17(8): 790-799.
- [12] S. M. Qi, X. W. H, J. Meng. Real Time Hand Tracking Based on Adaptive Active Contour Model[J]. Computer Science Journal, 2006, 33(11): 192-194.