An adaptive speed function of level set method for moving object extraction

An Adaptive Speed Function of Level Set Method for Moving Object Extraction

Kousuke Imamura*, Hideo Hashimoto†
Graduate School of Natural Science and Technology, Kanazawa University
Kakuma-machi, Kanazawa, Ishikawa 920-1192 Japan
{*imamura,†hasimoto}@ec.t.kanazawa-u.ac.jp

Abstract— **The convergence and stability of the level set method depend on the speed function. Therefore, it is important to define the speed function in a manner that is suitable for the individual application. In the present paper, we propose a novel speed function of the level set method for moving object extraction from a video sequence with a stationary background. The speed function focuses on the precision of moving object extraction. In the proposed extraction method, the outline between moving object regions and the background is estimated in advance based on the Gaussian noise distribution in the frame. The speed function is changed adaptively using the obtained outline in order to improve the convergence and precision of extraction. In addition, a new energy term in the direction of the contour of an object is incorporated into the speed function. The precision of the moving object extraction method using the proposed speed function is evaluated through computer simulations.**

I. Introduction

The extraction of moving objects from a video sequence is important for new content-based applications. As such a number of object extraction techniques, chromakey, texture analysis, contour extraction, and contour tracking, have been proposed. However, a high-precision, general purpose technique for moving object extraction has not yet been established.

The level set method was proposed by S. Osher *et al.* as a topology-free active contour model[1]. The level set method has attracted a great deal of attention, and techniques for improving have been proposed[2][3][4]. The level set method was applied to various applications, such as three-dimensional (3D) geometric modeling[6], simulation of the pattern formation in crystal growth[7], and surface editing[8]. Kurazume *et al.* applied the level set method to real-time detection and tracking of moving objects for surveillance[9]. However, few studies have focused on the precision of moving object extraction from a video sequence by the level set method. Moreover, since the convergence and stability of the level set method depend on the speed function, it is important to define the speed function in a manner that is suitabe for the individual application.

In the present paper, we propose a novel speed function of the level set method for moving object extraction from a video sequence with a stationary background. The speed function focuses on the precision of moving object extraction. In the moving object extraction method using the proposed speed function, the outline between moving object regions and the background is estimated in advance based on the Gaussian noise distribution in the frame. The speed function is changed adaptively using the obtained outline in order to improve convergence. In addition, a new energy term in the direction of the contour of an object is incorporated into the speed function for precise extraction. The energy term causes the contour obtained by the level set method to correspond to the Laplacian zero-crossing of the pixel intensity.

II. Level Set Method

In this section, we describe the basic principles of the level set method. In the level set method, the implicit function ϕ is defined in a space that is one dimension higher than that in which contour C exists at time t. The contour is represented as the cells for which the implicit function has a value of zero. The implicit function ϕ is updated according to a renewal function. We obtain the contour at time $t + \Delta t$ as the cells for which $\phi = 0$ by iteration of the updated process until time $t + \Delta t$. Figure 1 shows the concept map of the level set method.

In the implementation of the level set method for an image, let $C(\mathbf{p}, t)$ be a contour at time t, which denotes the set of points for which $\phi = 0$, where $\mathbf{p}$ is (p_x, p_y).

Let $\mathbf{p}(t)$ be a point on the contour $C(\mathbf{p}, t)$ such that

$$\phi(\mathbf{p}(t), t) = 0. \tag{1}$$

By the chain rule, we have

$$\phi_t + \nabla\phi(\mathbf{p}(t), t)\mathbf{p}_t = 0. \tag{2}$$

The speed function $F(\kappa)$ in the direction $\mathbf{N}$ normal to the contour $C(\mathbf{p}, t)$ is given by

$$F(\kappa) = \mathbf{p}_t \cdot \mathbf{N}, \tag{3}$$

where the normal vector $\mathbf{N}$ on the curve is given by

$$\mathbf{N} = \frac{\nabla\phi}{|\nabla\phi|}, \tag{4}$$

and κ is the local curvature of ϕ. Namely, the evolution equation for ϕ is defined as

$$\begin{aligned} \phi_t + F(\kappa)|\nabla\phi| &= 0, & (5)\\ \phi(C_0(\mathbf{p}), 0) &= 0 \quad \text{(initial condition)}, & (6) \end{aligned}$$

where C_0 is the initial contour.

978-1-4244-5309-2/10/$26.00 ©2010 IEEE

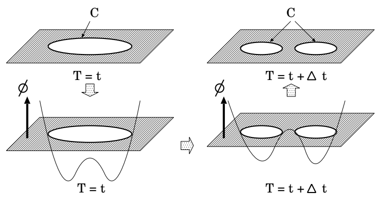

Fig. 1. Concept Map of the Level Set Method.

The implicit function ϕ is updated iteratively by Eq. (5), and we obtain a new contour satisfying the condition $\phi(x, y, t) = 0$. Thus, the level set method uses an implicit representation to express the movement of the contour and is able to handle the topological change of the contour.

The general form of the speed function F in Eq. (5) is

$$F_{i,j} = K_{i,j}(-a - b\kappa_{i,j}), \tag{7}$$

where a and b are positive constants, $\kappa_{i,j}$ is the local curvature of the implicit function ϕ, and the term $K_{i,j}$ is based on image features at the observed position (i, j). The image-based term $K_{i,j}$ is given by

$$K_{i,j} = \frac{1}{1 + |M|^n}, \tag{8}$$

where n is a positive constant.

In Eq. (8), M has a strong influence on the convergence and stability of the extraction. When the value of M is small, the evolution speed is fast and the contour moves inward, and when the value of M is sufficiently large, the evolution speed approaches zero. Generally, $\nabla^2 G \otimes I_{i,j}$ is used as M for contour extraction from an image. Here, $\nabla^2 G \otimes I_{i,j}$ represents the filtered intensity $I_{i,j}$ obtained by the Laplacian of Gaussian (LoG).

Figure 2 shows the extraction results obtained using $\nabla^2 G \otimes D_{i,j}$ as M for moving object extraction, where $D_{i,j}$ is the frame difference and the Gaussian kernel size is 15×15 pixels. The test sequence included Gaussian noise of $N(0, 4)$. In Figure 2, the shapes of the extracted contours are inaccurate due to the influence of the smoothing operator. In addition, the extracted contours did not include the shapes of the feet because of the insufficient frame difference.

Next, Figure 3 shows the locus of the extracted contours in the case of using $D_{i,j}$ as M. The locus of the contour is unsmooth because of the adverse effect of noise. However, the extraction results of Figure 3 provide the relatively precise shapes of the moving objects.

III. Pre-processing of the Proposed Method

The moving object extraction method using the proposed speed function estimates the outline between moving object regions and the background in advance. The speed function is changed using the obtained outline in order to improve the convergence and precision of extraction.

We assume that a frame of a video sequence includes Gaussian noise with a normal distribution $N(0, \sigma^2)$. From

Fig. 2. Extraction from the Frame Difference using the LoG Filter.

Fig. 3. Locus of Contour Obtained by the Speed Function using the Frame Difference.

the additivity of normal distributions, the frame difference includes Gaussian noise with $N(0, 2\sigma^2)$. The outline of the regions is estimated based on this assumption.

A. Estimation of the Gaussian Noise Distribution

We roughly divide a frame into moving object regions and the background.

First, the frame difference image was partitioned into 16×16 pixel blocks, and the mean value m_i of the absolute frame difference for each block is calculated. A histogram is constructed for m_i. The threshold TH_m for detecting a block as part of a moving object is determined as the value around the upper tail of the histogram. We test the connectivity of the detected blocks as the moving object part ($m_i \geq TH_m$), and the spatial isolation block is deleted.

Next, a dilation operation with a 3×3 block window is applied to the region of the moving object block. The variance of frame difference in the background block is then calculated, and the obtained variance is used as the estimated variance of the Gaussian noise distribution in the frame difference.

B. Pixel Detection in the Noise-only Region

The observed pixel (x, y) is located in either the noise-only region or the other regions by

$$\mathrm{Map}_{(x,y)} = \begin{cases} 0 & \text{if } \; S^2 \leq 2\sigma^2(1+\alpha)\text{: noise-only region} \\ 1 & \text{if } \; S^2 > 2\sigma^2(1+\alpha)\text{: other regions,} \end{cases} \tag{9}$$

where S^2 is defined as

$$S^2 = \frac{1}{N} \sum_{(i,j)\in\Omega} \left\{ D_{i,j} - \frac{1}{N} \sum_{(i,j)\in\Omega} D_{i,j} \right\}^2, \tag{10}$$

S^2 denotes the variance of frame difference in a $N = (2n+1) \times (2n+1)$ square block window centered at the observed pixel, and α is determined by the significance level as follows:

$$P_e = \Pr\left[S^2 > 2\sigma^2(1+\alpha)|\text{noise-only region}\right], \qquad (11)$$

because $N\frac{S^2}{\sigma^2}$ denotes a χ^2 distribution with $N-1$ degrees of freedom.

C. Estimation of Moving Object Regions

The obtained noise-only region (Map = 0) in the frame difference represents the background, and we assume that the pixels in the other regions (Map = 1) belong to moving object regions. However, we regard small moving object regions as unsemantic objects. The morphological closing operation with a 5×5 pixel window is first applied to the moving object regions in Map in order to remove the unsemantic objects. Next, for the case in which the obtained moving object region is less than 0.1% of the entire image, the moving object region is assigned to the background.

D. Exception of Uncovered-Background

Generally, the contour extracted from the frame difference is not precise because the frame difference includes the uncovered background. Hence, the object map is constructed by Eq. (12) so as to exclude the uncovered background as follows:

$$\text{Rmap}_t = \text{Map}[t] \wedge \text{Map}[t-1], \qquad (12)$$

where $\wedge$ denotes logical conjunction, and Map$[t]$ is the Map constructed from the frame difference between frame $I(t+1)$ and frame $I(t)$. Therefore, the Rmap$_t$ is constructed from the three consective frames.

The obtained Rmap$_t$ represents the outline of the moving object regions and background at time t, and the proposed speed function uses the outline for adaptive processing.

IV. Speed Function

A. Energy Term for the Improvement of Extraction Precision

The proposed speed function includes a new energy term in the direction of the contour of an object in order to improve the precision of extraction. The energy term causes the contour obtained by the level set method to correspond to the Laplacian zero-crossing of the pixel intensity.

The energy is calculated as follows:

1. A Gaussian filter is applied to the frame.
2. A Sobel filter is applied to the obtained smoothed frame. This process provides the gradient of the smoothed frame g.
3. The potential vector v_p at a point in one of the moving object regions (Rmap = 1) is calculated as the distance from the nearest background region (Rmap = 0), and v_p of the background region is the zero vector.
4. The energy E in the direction of the contour of object is given by

$$E = \nabla g \cdot \nabla v_p. \qquad (13)$$

B. Definition of the Novel Speed Function

The proposed speed function is changed adaptively using the obtained outline. The novel speed function of the proposed method is defined as follows:
Background (Rmap$_t$ = 0):

$$F_{i,j} = K_{I,i,j}(-a - b\kappa_{i,j}), \qquad (14)$$

$$K_{I,i,j} = 1, \qquad (15)$$

Moving object regions (Rmap$_t$ = 1):

$$F_{i,j} = \begin{cases} K_{I,i,j}(-a' - b'\kappa_{i,j}) & \text{if } K_{I,i,j} \geq 0 \\ K_{I,i,j}(-a' + b'\kappa_{i,j}) & \text{if } K_{I,i,j} < 0, \end{cases} \qquad (16)$$

$$K_{I,i,j} = K_d + (1 - K_d)K_v, \qquad (17)$$

where

$$K_d = \frac{1}{1 + \left(\frac{|D_{i,j}|}{\sigma_D}\right)^n}, \qquad (18)$$

$$K_v = max\left(-1.0, min\left(1.0, H_{i,j,\alpha}\right)\right), \qquad (19)$$

$$H_{i,j,\alpha} = -0.5 + 0.5\ \frac{(E_{i,j} + \alpha)}{\alpha} + \frac{1}{\pi}\sin\left(\frac{\pi(E_{i,j} + \alpha)}{\alpha}\right), \qquad (20)$$

and a, a', b, b', n, and α are positive constants.

The image-based term of the proposed speed function is constructed from the term K_d based on the frame difference and the term K_v based on the new energy E.

In other words, the speed function causes the contour to move based on the local curvature term κ in the background, so that the convergence is immune to the influence of noise. In addition, the contour is made to correspond to the zero-crossing point of the Laplacian of the pixel intensity as a result of the energy in the direction of the contour of object in the neighboring boundary of the moving object region. Consequently, the proposed speed function provides precise extraction of the shape of the moving object.

V. Simulation and Results

The proposed moving object extraction was examined by computer simulations. "Japanese Room" and "Table Tennis" (grayscale) were used as test sequences. The parameters for the speed function were set as $a = 0.5$, $a' = 0.1$, $b = 2.0$, $b' = 0.3$, $n = 1$, and $\alpha = 30$.

First, we verified the estimated outline between the moving object regions and the background. Figure 4 shows the estimated moving object outline for the "Japanese Room" sequence for changing of speed function.

Next, we verified the improvement in stability provided by the proposed method. Figure 5 shows the locus of the contour ($\phi = 0$) obtained by the proposed method. From Figure 5, since the locus of the contour was smooth, the proposed method improved the stability of extraction by avoiding the influence of noise in the background.

Finally, we evaluated the precision of the moving object extraction by the proposed method. For the purpose of comparison, Figure 6 shows the extraction results obtained by

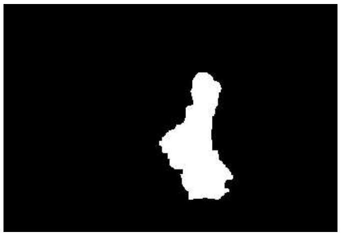

Fig. 4. Outline between the Moving Object Region and the Background.

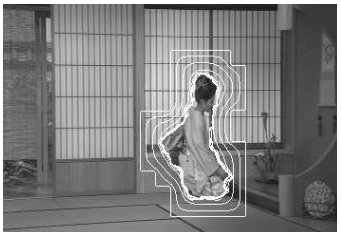

Fig. 5. Locus of the Contour Obtained using the Proposed Method.

(a) Previous Level Set Method.

(b) Proposed Method

Fig. 6. Extraction Results (Japanese Room).

(a) Previous Level Set Method.

(b) Proposed Method

Fig. 7. Extraction Results (Table Tennis).

the level set method using the LoG filtered frame difference, and Figure 7 shows the extraction results for the sequence of "Table Tennis". Comparison of these results indicates that the proposed method extracts the contour more accurately than the level set method using the previous speed function.

VI. CONCLUSIONS

In the present paper, we proposed a novel speed function of the level set method for moving object extraction from a video sequence with a stationary background. The speed function is changed adaptively using the obtained outline in order to improve convergence. In addition, a new energy term in the direction of the contour of an object is incorporated into the speed function for precise extraction. The simulation results revealed that the proposed method improved the convergence and precision of moving object extraction by the level set method.

In the future, we intend to improve the method used to estimate the Gaussian noise distribution σ^2. In addition, we intend to examine the possibility of adapting the proposed method to the extraction of moving objects from a moving background.

REFERENCES

[1] S. Osher and J. A. Sethian, "Fronts Propagating with Curvature-dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations," *Journal of Computational Physics*, Vol.79 No.1, pp.12-49, 1988.

[2] Adalsteinsson, D. and J. A. Sethian, "A Fast Level Set Method for Propagating Interfaces," *Journal of Computational Physics*, Vol.118, pp.269-277, 1995.

[3] J. A. Sethian, "Level Set Methods and Fast Marching Methods," *Volume 3 of Cambridge Monographs on Applied and Computational Mathematics*, Cambridge University Press, Cambridge, Second Edition, 1999.

[4] M. Rousson and N. Paragious, "Prior Knowledge, Level Set Representations & Visual Grouping" *International Journal of Computer Vision*, Vol.76, No. 3, p.231-243, 2008.

[5] J. A. Sethian, "Level Set Methods" Cambridge University Press, Cambridge, 1996.

[6] C. Baillard, P. Hellier, and C. Barillot, "Cooperation between Level Set Techniques and Dense 3D Registration for the Segmentation of Brain Structure," *International Conference on Pattern Recognition*, Vol.1, pp.991-994, 2000.

[7] Kei, K. and Kokichi, S., "Approximation of Multiplicatively Weighted Crystal Growth Voronoi Diagram and Its Application," *The Transactions of the Institute of Electronics, Information and Communication Engineers. A*, Vol.J83-A, No.12, pp.1495-1504, 2000.

[8] K. Museth, D. Breen, R. T. Whitaker, and A. H. Barr, "Level Set Surface Editing Operators," *ACM Transactions on Graphics*, Vol.21, No.3, ACM SIGGRAPH '02, pp.330-338, 2002.

[9] R. Kurazume, S. Yui, T. Tsuji, Y. Iwashita, K. Hara, and B. Hasegawa, "Fast Level Set Method and Realtime Tracking of Moving Objects in a Sequence of Images," *Transactions of Information Processing Society of Japan*, Vol.44, No.8, pp.2244-2254, 2003.

[10] K. Haris, G. Tziritas, and S. Orphanoudakis, "Smoothing 2-D or 3-D Images Using Local Classification," *Proc. EUSIPCO*, Edinburgh, U.K., Sept. 1994.