21

Image Segmentation using Active Contour for Images of Complex Background

Daehee Kim and Yo-Sung Ho
Kwangju Institute of Science and Technology (K-JIST)
1 Oryong-Dong Puk-Gu, Kwangju, 500-712, Korea

Abstract

The basic idea of an active contour algorithm is to find an energy-minimizing spline, which operates under certain constraints. Conventional active contour algorithms are designed for extracting objects in the homogeneous background, but are not applicable to objects of complex background. In this paper, we employ morphological tools to define a new external energy, which guarantees to produce good segmentation results for objects of complex background.

1 Introduction

In recent years, MPEG-4 has become a very promising standard for multimedia applications. MPEG-4 enables content-based functionalities by introducing the concept of video object planes (VOP's). In order to process the image based on its content, we should segment the image into a set of meaningful objects and determine their parameters such as color, shape and motion. Knowledge about the shape of video objects in the scene helps us for better image reconstruction, especially along object boundaries.

Each frame of the input sequence is segmented into arbitrary shaped image regions such that each video object describes a semantically meaningful object or video content of interest. A video object layer (VOL) is assigned to each video object containing shape, motion and texture information.

Among the conventional approaches for segmentation, the spatio-temporal segmentation technique is particularly interesting to us because it employs the morphological tools and change detection algorithms based on a statistical model. MPEG-4 FCD includes a typical spatio-temporal algorithm in its informative annex. This algorithm extracts edge information by the morphological operation and obtains information of moving objects by the change detection mask [1, 2].

However, the spatio-temporal segmentation algorithm cannot extract each individual video object in a single frame. In addition, this algorithm is somewhat premature to obtain satisfactory segmentation results from various kinds of image sequences, because the mathematical model and the similarity measure for extracting video objects are not defined adequately. If the user can define video objects in the first frame or newly appeared video objects in a partially or completely user-assisted manner, we may obtain better segmentation results in the succeeding picture frames. Semi-automatic segmentation may be more practical in generating VOPs of moving objects.

The types of active contour algorithm can accommodate various external and constraint forces. Therefore, the active contour approach is appropriate for semi-automatic segmentation. In this paper, we propose a robust active contour algorithm for finding the shape of visual objects in a single image frame.

2 Active Contour Algorithm

Basically, the active contour algorithm tries to find an energy-minimizing spline curve from an initial curve drawn by the user. The performance of an active contour algorithm depends on how to define energy functional $E(s)$. The shape of the active contour is controlled by internal force, external force and constraint force.

Internal force is as a smoothness constraint. External force guides the active contour towards image features. Constraint force allows interactivity in manipulating the active contour. The energy functional $E(s)$ is represented as a parametric curve $v(s) = (x(s), y(s))$, where s is the arc length [3]. A functional of the snake is defined by

$$\begin{aligned} E^*_{snake} &= \int_0^1 E_{snake}(v(s))ds \\ &= \int_0^1 [E_{int}(v(s)) + E_{ext}(v(s)) \\ &\quad + E_{con}(v(s))]ds \qquad (1) \end{aligned}$$

where E_{int}, E_{ext} and E_{con}, represent the internal energy of the contour, image force, and external constraint force, respectively. The final location of the active contour corresponds to the local minimum of the energy functional.

The internal spline energy can be written as

$$E_{int} = (\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)/2 \quad (2)$$

The above equation contains a first-order term $v_s(s)$ that has a large value in a gap of the curve, and a second-order continuity term $v_{ss}(s)$ that has a large value where the curve is bending rapidly. The weights, $\alpha(s)$ and $\beta(s)$ at each point, control the extent to which the contour is allowed to stretch or bend at that point. For example, a large value of $\beta(s)$ would make the curve smoother, approaching a circle for a closed contour and a straight line for an open contour. If $\alpha(s) = 0$, discontinuity can occur at that point. If $\beta(s) = 0$, the active contour can guarantee to find the corner at that point. Therefore, $\alpha(s)$ and $\beta(s)$ should be selected appropriately according to the image content.

The second term, E_{ext}, is the energy obtained from the image, whose value is small at the point of large gradient. We can use the Laplace operator or the Sobel operator to find image edges. Since external energy has a small value at the object boundary, external energy enforces the contour to be located at image edges. When we extract image edges by a simple mask operation, we assume that the background of the image is homogeneous. Unfortunately, since most images have complex backgrounds, it is difficult to extract object features by the simple mask operation. When the minimization process of the active contour algorithm is applied on the image of complex background, the final curve could converge to an unwanted local minimum point.

If we denote $E_{EXT} = E_{ext} + E_{con}$, Eq. (1) becomes

$$\begin{aligned} E^*_{snake} &= \int_0^1 [E_{EXT}(v(s)) \\ &+ \frac{1}{2}(\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)]ds \quad (3) \end{aligned}$$

The minimum value of Eq. (3) can be obtained by techniques of variational calculus [3]. However, the minimization of $E(s)$ using variational calculus has a problem of numerical instability [4]. Since dynamic programming techniques [5] can provide more stable results, most active contour algorithms adopt dynamic programming. Unfortunately, dynamic programming methods are extremely slow with a computational complexity of $O(nm^3)$, where n is number of points in the active contour and m is the number of points in the search window. Greedy algorithms [6] retain numerical stability and improve the speed of convergence, but they are sensitive to the distance between adjacent points on the initial contour.

Conventional active contour algorithms are deigned for objects of homogeneous background, but they may not work well for objects of complex background. In this paper, we propose a new active contour algorithm that can be applied on images of complex background.

3 Morphological External Energy

In order to discretize Eq. (1), we need to represent the contour by a finite number of control points as

$$E^*_{snake} = \sum_{i=1}^{n} \lambda_i E_{int}(v_i) + (1 - \lambda_i)E_{ext}(v_i) \quad (4)$$

where n is the total number of control points on the contour.

We also discretize Eq. (2) as

$$\begin{aligned} E_{int}(v_i) &= \alpha_i|v_i - v_{i-1}|^2 \\ &+ \beta_i|v_{i+1} - 2v_i + v_{i-1}|^2 \quad (5) \end{aligned}$$

Since there is no external energy in this approximation, the contour could converge to only one point. In order to prevent the contour from being shrunk to one point, we define a new internal energy as

$$E_{int}(v_i) = \frac{1}{l(V)}|v_i - \alpha(v_{i-1} + v_{i+1})|^2 \quad (6)$$

where we choose $\alpha = 0.5$ for open contours, and $\alpha = 0.5cos^{-1}(2\pi/n)$ for closed contours. Eq. (6) resolves the contour shrinking problem when the internal energy in Eq. (2) is discretized. The internal energy can be normalized by the averaged distance $l(V)$ between two successive control points.

$$l(V) = \frac{1}{n}\sum_{i=1}^{n}|v_i - v_{i-1}|^2 \quad (7)$$

In this paper, we define two kinds of external energy functions. For the first one, we use the gradient information as

$$E_{grd}(v_i) = 1 - |n_i^T g(v_i)| \quad (8)$$

where $g(v_i)$ is a 2×1 gradient vector, and $n(v_i)$ is the unit normal vector at the control point v_i. Therefore, when the gradient vector has the same direction as the normal vector of the contour, the first external energy has the minimum value.

If we use a simple edge detector as in conventional active contour algorithms, it is difficult to obtain satisfactory segmentation results from images of complex background. Therefore, we employ morphological tools to simplify the image and we define the second external energy function by the morphological gradient value of the simplified image.

In order to derive a new morphological energy function, we need to remove the complexity of images by morphological operations. In addition, this simplification operation suppresses the number of watershed lines. Morphological open-close and close-open by reconstruction filters are used for image simplification. These filters remove regions that are smaller than the size of the structuring element, but preserve object contours in the image [7]. The size of the structuring element should be determined according to applications.

The spatial gradient of the simplified image is approximated by a morphological gradient operator [8]. The spatial gradient can be used as an input to the watershed algorithm to partition the image into homogeneous intensity regions. The resulting gradient image contains noisy gradient values, which can cause oversegmentation of the image. In order to alleviate this problem, we set a threshold for gradient values. Gradient values smaller than the threshold value are set to zero.

We then apply the watershed algorithm on the gradient values of the image to obtain the object contours. The watershed algorithm is originated from topographic works, which deals with catchment basins and their dividing lines, called as watershed lines [9]. Watershed lines partition the image by associating each catchment basin to its local minimum point.

A watershed algorithm can be performed by immersion simulation in discretized grid points [9]. We first pierce a hole in each local minimum point. We then immerse our surface into a lake slowly. Starting from the point of the lowest altitude, we fill up different catchment basins with water progressively. At each pixel where the water coming from two different minima would merge, we build a dam. At the end of this immersion procedure, each minimum point is completely surrounded by dams, which delimit its associated catchment basin. The final result of the watershed algorithm is a tessellation of the input image into its different catchment basins. These dams are considered as watershed lines.

We obtain a new binary image W_{img} by the watershed algorithm. Pixel values on the watershed lines are set to one in W_{img}. We can define a morphological external energy on the binary image W_{img} by

$$E_{mor}(v_i) = 1 - \sum_{m=-1}^{1} \sum_{n=-1}^{1} W_{img}(v_{ix}+m)(v_{iy}+n) \quad (9)$$

where v_{ix} and v_{iy} are the components of v_i. In Eq. (9), since the thickness of the watershed line is one pixel, we use the local window.

4 Minimization Process

In order to find a set of control points minimizing Eq. (4) around the initial contour, we decompose the minimization process into n independent stages. In each stage, we include only 3 neighboring points. This idea initially was proposed under the framework of dynamic programming [4].

Since it is quite difficult to select parameter values λ_i for various situations, we try to find a curve that minimizes the total energy of the active contour.

$$E_{snake}^{final} = \min_{V} \sum_{i=1}^{n} \max[E_{int}(v_i), E_{grd}(v_i), E_{mor}(v_i)] \quad (10)$$

This function satisfies the minimax criterion when $\lambda_i = 1$ or $\lambda_i = 0$ in Eq. (4). Since the energy at each point is related to internal energy, the energy at each control point can be defined from three adjacent control points. If we represent

$$\begin{aligned} E_i(v_i, v_{i+1}, v_{i+2}) &= \max[E_{int}(v_i, v_{i+1}, v_{i+2}), \\ &\quad E_{grd}(v_i), E_{mor}(v_i)] \end{aligned} \quad (11)$$

we can solve our problem by dynamic programming, as in Eq. (12).

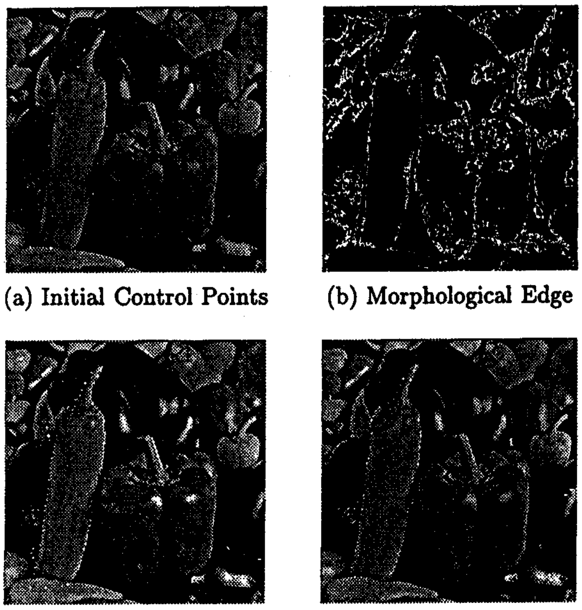
$$\begin{aligned} s_1(v_2, v_3) &= \min_{v_1} E_1(v_1, v_2, v_3) \\ s_2(v_3, v_4) &= \min_{v_2} s_1(v_2, v_3) + E_2(v_2, v_3, v_4) \\ &\vdots \\ s_{n-2}(v_{n-1}, v_n) &= \min_{v_{n-2}} s_{n-3}(v_{n-2}, v_{n-1}) \\ &\quad + E_{n-2}(v_{n-2}, v_{n-1}, v_n) \end{aligned} \quad (12)$$

5 Experimental Results

For performance evaluation of the proposed segmentation algorithm, we have chosen "Pepper" image of size 512 × 512 pixels, because this image has more complex background than any other images used in active contour algorithms. In order to avoid heavy computational requirement for the minimization process, we reduce the search range gradually as in the three-step algorithm for motion estimation [10].

Fig. 1(a) shows the initial contour provided by the user with a pointing device such as a mouse. Fig. 1(b) exhibits binary image edges obtained by morphological filers and the watershed algorithm.

Fig. 1(c) is the segmentation result obtained from the given initial contour by Lai's generalized snake algorithm [11]. If there is a local minimum point of the



(a) Initial Control Points (b) Morphological Edge

(c) G-Snake (d) Proposed Scheme

Figure 1: Experimental Results

gradient-based energy inside the area of the same object, the active contour converges to that point in Fig. 1(c). Therefore, the snake algorithm fails to extract the object boundary accurately.

Fig. 1(d) is the final result obtained by the proposed method. If there is no difference of the morphological energy in a certain region, the contour converges to the local minimum value of the gradient-based energy. If the gradient value is very small in a certain region, the contour converges to the local minimum value of the morphological external energy. Fig. 1(d) demonstrates that the morphological approach produces better results for images of complex background.

6 Conclusions

In this paper, we proposed a new active contour algorithm for intra-frame segmentation. Performance of the active contour algorithms heavily depends on definitions of energy functions. Since conventional active contour algorithms include external energies obtained from simple edge detectors, they work well for objects of homogeneous background, but they are not applicable to objects of complex background. However, our proposed algorithm performs quite well for objects of complex background owing to the introduction of the sophisticated external energy functions: the gradient-based energy and the morphological energy.

Acknowledgments

This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optic (UFON) Networks Research Center at Kwangju Institute of Science and Technology (K-JIST).

References

[1] ISO/IEC JTC1/SC29/WG11, "Information technology - generic coding of audio-visual objects Part2: visual ISO/IEC FDIS 14496-2," *MPEG98*, Oct. 1998.

[2] D. Kim, M. Kim and Y. S. Ho, "Spatio-temporal video segmentation using history information," *VLBV'98*, pp. 141-144, Oct. 1998

[3] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: active contour models," *Proc. First Int. Conf. on Compuetr Vision*, pp. 259-269, 1987.

[4] A. A. Amimi, S. Tehrani and T. E. Weymouth, "Using dynamic programming for solving variational problems in vision," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 9, pp.255-274, Sep. 1990.

[5] R. Bellman and R. Kalaba, *Dyanmic Programming and Modern Control Theory*, Academic Press, 1965.

[6] D. J. Williams and M. Shah, "A fast algorithm for active contours and curvature estimation," *CVGIP:Image Understanding*, Vol. 55, No. 1, pp.14-26, Jan. 1992.

[7] P. Salembier and M. Pardas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. Image Processing*, Vol. 3, No. 5, pp. 639-651, Sep. 1994.

[8] P. Salembier, "Morphological multiscale segmentation for image coding," *Signal Processing*, Vol. 38, No. 3, pp. 359-386, Jan. 1994.

[9] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No.5, pp. 583-598, June 1991.

[10] A. M. Tekalp, *Digital Video Processing*, Prentice-Hall, 1995.

[11] K. F. Lai and R. T. Chin, "Deformable contours: modeling, extraction," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 17, No 11, pp. 1084-1090, Nov. 1994.