AN ACTIVE CONTOUR MODEL FOR VIDEO SEGMENTATION

Daniele D. Giusto, Francesco Massidda, Cristian Perra

Department of Electrical and Electronic Engineering
University of Cagliari, Cagliari, ITALY

ABSTRACT

This paper presents an active contour model for object segmentation in video sequences. The proposed model is completely automatic if there is a single moving object. It requires user intervention if there are more objects in the scene. At first the video content is analyzed and preprocessed by means of a watershed algorithm. Then the active contour curvature is dynamically evaluated and controlled in order to obtain a good representation of the object contour.

1. INTRODUCTION

Segmenting objects in video sequences is a very challenging field. The most successful methods for such tasks are probably those based on deformable models. In particular the active contours or snakes proposed in [1,2] and their improvements [3,5,6,9] are often used as reference in the present literature.

Active contours are deformable curves confined to the plane. They move within the image plane guided by forces depending from the curve configuration and from the image data. The main forces that govern this process are called external forces and internal forces. The former usually depend on image boundaries whilst the latter depend on the curve configuration.

The objective is to identify an active contour configuration that is solution of an energy minimization problem. An appropriate choice of the energy functions should provide a contour solution corresponding to the object boundary. The process starts from an initial active contour configuration that moves until stability is reached.

The image noise can cause the snake to be trapped before the desired solution is obtained [2]. The starting contour must be very close to the object contour otherwise it is difficult to make the active contour move towards the object. Another drawback [2] is that the active contour cannot move into boundary concavities or indentation.

External forces defined as the negative gradient of a Euclidean distance map are often used [5,6] in active contour models. In particular gradient vector flow (GVF) [6] reduces the drawbacks of closeness between initial contour and object contour and allows that active contour moves inside concavities and indentation. The GVF-based model proposed in [10] is effective in contour extraction of interframe moving masks that is used as a starting point for successive object segmentation in video sequences.

This paper presents an active contour evolution method for object segmentation in video sequences, based on a modified GVF field and curvature analysis. In order to improve the effectiveness of the field, a preprocessing scene classification step is applied to the image. This provides a better external GVF field. Several experiments on standard video sequences have been performed in order to assess the goodness and the computational capacities of the proposed approach.

The paper is organized as follows. An overview of active contours is presented in Section 2. The proposed framework is presented in Section 3. The experimental results are presented in Section 4. Section 5 concludes this paper.

2. ACTIVE CONTOUR

An active contour is a parametric curve $V(s) = (x(s), y(s))$, $s \in [0,1]$ confined to the plane. The related discrete deformable model changes in subsequent iterative steps. The deformation of the active contour is performed in order to find a configuration that minimizes the following energy functional:

$$E_{snake} = \int_0^1 (E_{\text{int}} + E_{ext})ds \,. \tag{1}$$

Eq. (1) defines a sum of energy terms. The first term represents an internal energy which is a function of the $V(s)$ contour position. The second term represents an external energy which is a function that takes into account characteristics of the processed images such as edges and luminance peaks. The former has the objective of preserving, for the active contour, characteristics like steadiness, smoothness, tension and stiffness. The latter drives the active contour towards the desired points or boundaries within the image plane.

The parametric curve representing the active contour is dynamically deformed guided by a process that tends to minimize Eq. (1). The internal and external energy functionals should be chosen in order to have their minima in the boundaries of the object of interest.

Internal energy is often defined as follows:

$$E_{\text{int}}(V) = \frac{1}{2}\alpha\left|V_s\right|^2 + \frac{1}{2}\beta\left|V_{ss}\right|^2 \qquad (2)$$

where the first component of the sum represents the local tension of the active contour whilst the second represents the active contour curvature or stiffness.

The external energy used in this work is based on the Gradient Vector Flow (GVF) [6]. This is a vector field similar to the image gradient that realizes a spread field over the image plane. This field allows the starting contour to lay theoretically in every position inside the image plane or being captured and guided by a significant intensity of forces. If the external energy is simply the image gradient the starting active contour should lay next to the boundaries of the target object. In the GVF approach, Eq. (1) is minimized using variational methods and solving the Euler equation

$$\begin{aligned} &\mu \cdot \nabla^2 u - (u - f_x)(f_x^{\,2} + f_y^{\,2}) = 0 \\ &\mu \cdot \nabla^2 v - (v - f_y)(f_x^{\,2} + f_y^{\,2}) = 0 \end{aligned} \qquad (3)$$

where f is an edge map of the image, μ is a constant parameter and *(u,v)* are the GVF vector field components. In order to increase the GVF capture capabilities, more complex preprocessing step could be used in connection with the GVF field iterative approach.

A simple measure of the local active contour curvature normally is used within the snake minimization process in order to obtain curves that preserve properties of smoothness and regularity. Often a common solution is the estimation based on a functional of three contiguous control points of the curve [2,3,6], where such control points are evenly spaced along the curve. Several approaches have been presented in literature [4,7,8] providing a more effective estimation of the active contour curvature. In [7] several numerical approximations are presented, and it is shown how the curvature estimation performed using the following expression is the better compromise between speed and accuracy:

$$\left|\frac{\mathbf{U}_i}{\left|\mathbf{U}_i\right|} - \frac{\mathbf{U}_{i+1}}{\left|\mathbf{U}_{i+1}\right|}\right| = \left(\frac{\Delta x_i}{\Delta s_i} - \frac{\Delta x_{i+1}}{\Delta s_{i+1}}\right) + \left(\frac{\Delta y_i}{\Delta s_i} - \frac{\Delta y_{i+1}}{\Delta s_{i+1}}\right) \qquad (4)$$

where $\mathbf{U}_i = (x_i - x_{i-1}, y_i - y_{i-1})$ is a control point of the active contour $V(s)$.

Fig. 1. Watershed segmentation of two frames from video sequences Akiyo and Toyclock

3. VIDEO SEGMENTATION FRAMEWORK

Many segmentation algorithms use a preprocessing step in order to decrease the complexity of the image content and the amount of texture and noise. Simple gradient of the image has been used in the original Xu and Prince approach. More complex edge extracting operators have been used in the recent literature [10]. The approach proposed in this paper uses the fast Watershed algorithm proposed in [11] (Fig. 1). The GVF field is then computed using the classified and simplified image.

The main problem is that a complex preprocessing step increases the computational time needed to compute the external field. For several years, many researchers have focused on the problem of computational complexity reduction for active contours. External field computation and iterative active contour functional minimization are complex processes that require a high and variable number of iteration in order to converge. One of the simplest methods for controlling the trade off between speed and accuracy is based on a regularization process performed, for each iteration step, on the number of control points within the parametric curve. Standard snake implementations (e.g. [2,6]) allows a simple regularization of the control point positions, in term of an almost equi-spacing performed with a minimum $d_{\min}$ and maximum $d_{\max}$ distance constraint between consecutive control points for each iterative step. With this approach, a simple method for controlling the process speed and the final contour accuracy is achieved.

Then, the computational complexity of the active contour convergence process is related to the number of control points and, therefore, to the implementation capacity in minimizing the total number of control points or optimizing the location with respect to the characteristics and positions of the desired boundaries.

The main problem of the classical approach is that simple methods of regularization do not take into account

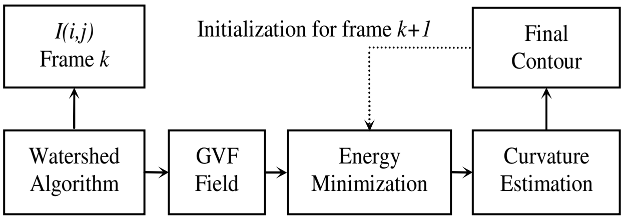


Fig. 2. Proposed active contour segmentation framework

the topology of the active contour that depends on the external field intensity and direction. In other terms, it does not depend on the boundaries of the objects within the image plane.

The purpose of the proposed algorithm is to obtain a compromise between the number of control points sufficient to an effective and efficient description of the desired contour and the minimum number of points necessary for allowing a fast snake convergence.

This procedure have to be done taking into account the topology of the active contour for each time step of its evolution until the final contour is reached. When the regularization is performed using the $d_{\min}$ and $d_{\max}$ approach, where the curvature of the desired edges is big, the contour approximation is done in a less accurate manner, resulting in smooth boundaries because of the reduced number of control points.

The proposed method is based on the connection between the active contour shape and the local curvature of the desired object boundaries. The main idea is that if the point elimination could be done according to the local curvature of the active contour configuration, the same description of the final object can be achieved using a minimum number of control points for each iteration step. When the initial contour starts to approach the desired boundaries, the snake shape starts to look like the object contour. Then, in each iteration step, the point elimination process can be done obtaining the best possible description of high curvature zones of the final object boundary. The algorithm should delete only the points having local curvature smaller then a given local threshold (for each j*th* iteration). Control points insertion/deletion is performed keeping into account the local value of curvature and the local and global distance between reference and optimized configuration.

A configuration energy term E_{conf} is introduced in the classical definition (Eq. (1)) and defined as:

$$E_{conf} = \left| d\left(P_{N(t)}(t), P_{M(t)}(t)\right) - C \right| \quad (5)$$

where *N(t)* and *M(t)* are the number of control points for the reference and optimized configurations $P_{N(t)}(t)$ and $P_{M(t)}(t)$ in each t*th* iterative step; *d(.)* is a measure of distance between the two contours and *C* is a constant value. The new energy term gives a measure of the actual state and position of the active contour for the t*th* iteration step. An appropriate threshold value *C* should be defined for an acceptable maximum distance between the reference and the optimized configurations, which can easily take into account different needs and tolerance thresholds of several application fields. The optimized configuration $P_{M(t)}(t)$ is iteratively changed in each time step, starting from the reference one, according to a point elimination phase performed using local curvature estimation. In order to evaluate the local curvature, the curvature estimation is done according to Eq. (4). The procedure ends when the *d(.)* have a value as close as possible to the chosen threshold *C*. Fig. 2 shows the framework of the proposed video segmentation method.

4. RESULTS

In order to assess the effectiveness of the curvature based Snake model, several QCIF and CIF standard test sequences have been used. The proposed algorithm, implemented in C++ language on an AMD Athlon 2500+ with 512MBytes of DDR-RAM, has been applied only to the luminance component.

In order to prove the tracking capability of this approach, when used in videoconference test sequences, the final active contour extracted for the previous frame is used to initialize the active contour evolution procedure in the following frame.

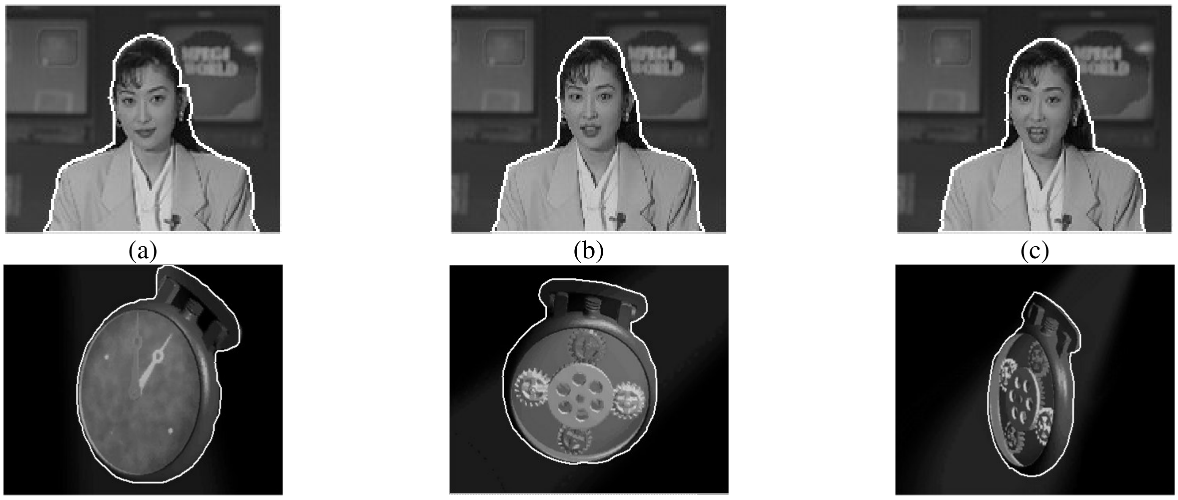
In Fig. 3(a-c) three final contours extracted from Akiyo test sequence are shown. When used to extract single or simple objects within the image, a static initialization can be used for every frame.

In Fig. 3(d-e) three final contours are shown, obtained for the Toyclock sequence, using a constant four-control-point initialization snake located in the corners of the image plane. Table 1 presents the computational times and iteration numbers obtained using the proposed approach with or without the Watershed algorithm.

The control point optimization process decreases the snake control point number, without losing accuracy especially in high curvature object edges. Considering the addition of the Watershed algorithm, the global process of snake convergence results faster with respect to the classical GVF method.

5. CONCLUSIONS

An active contour method for video segmentation and tracking was developed and presented. The proposed algorithm starts with an image preprocessing (fast watershed algorithm). The external field is then computed



(a) (b) (c)

(d) (e) (f)

Fig. 3. Video segmentation for Akiyo (a-c) and Toyclock (d-e). (a,d) Frame 1, (b,e) Frame 25, (c,f) Frame 75

	Sequence Format			Processed Frames per Seconds		
	Width	Height	fps	Field	Snake	Tot
Akiyo	176	144	15	0.17	0.30	6.38
Claire	176	144	15	0.18	0.18	8.33
Miss_Am	176	144	15	0.63	0.20	1.20
Toyclock	352	288	10	0.22	0.33	1.81
Ball	352	288	10	0.91	0.23	0.88

Table 1. Performance achieved by the proposed method. Field column reports the mean computing time for the external field while Snake column reports the mean time for the snake evolution. The watershed algorithm is used only for Miss America and Ball sequences.

using a modified GVF approach and a regularization of the control points based on more effective curvature estimation. The presented results reveal the accuracy of the proposed method (comparable with the original GVF approach). Furthermore, the minimization of the control points brings an acceptable computational speed even though a watershed algorithm is introduced in the GVF field construction.

6. REFERENCES

[1] D. Terzopoulos, and K. Fleischer, "Deformable models," The Visual Computer, 4 (6) (1988) 306–331.

[2] M. Kass, A. Witkin, D. Terzopoulos, "Snakes: Active Contour Models," International Journal of Computer Vision, (1) (1987) 321-331.

[3] A. A. Amini, T. E. Weymouth, R. C. Jain, "Using Dynamic Programming for Solving Variational Problems in Vision," IEEE Transaction On PAMI, 12 (9) (September 1990).

[4] D. J. Williams, and M. Shah, "A Fast Algorithm for Active Contours and Curvature Estimation," CVGIP (Image Understanding), 55 (1) (1992) 14-26.

[5] L. D. Cohen, I. Cohen, "Finite-Element Methods for Active Contour Models and Balloons for 2-D and 3-D Images," IEEE Trans. Pattern Analysis and Machine Intelligence, 15 (11) (1993) 1131-1147.

[6] C. Xu e J. L. Prince, "Snakes, Shapes, and Gradient Vector Flow," IEEE Transactions on Image Processing, 7 (3) (March 1998).

[7] F. Mohanna, and F. Mokhtarian, "Improved Curvature Estimation for Accurate Localization of Active Contours," Proc. International Conference on Image Processing, vol. II, Thessaloniki, Greece, 2001, pp. 781-784.

[8] D. Perrin and C. Smith, "Rethinking classical internal forces for active contour models," In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2001.

[9] L. D. Cohen, "On Active Contour Models and Balloons," Computer Vision Graphics and Image Processing: Image Understanding 53 (2) (March 1991) 211-218.

[10] C. Perra, F. Massidda, D.D. Giusto, "Active contour for automatic segmentation of video sequences," Workshop on Image Analysis for Multimedia Interactive Services, April 9-11, 2003, London, UK.

[11] S. L. Stoev, "RafSi - a Fast Watershed Algorithm Based on Rainfalling Simulation," In: Proceedings of 8-th International Conference on Computer Graphics, Visualization, and Interactive Digital Media (WSCG'2000), (2000) 100-107.