

# OBJECT TRACKING IN CLUTTER AND PARTIAL OCCLUSION THROUGH RULE-DRIVEN UTILIZATION OF SNAKES

*Gabriel Tsechpenakis, Kostas Rapatzikos, Nicolas Tsapatsoulis and Stefanos Kollias*

Image, Video and Multimedia Systems Laboratory,  
School of Electrical and Computer Engineering,  
National Technical University of Athens,  
9 Iroon Polytechniou Str., 15773 Athens, Greece

## ABSTRACT

Efficient moving object tracking in cases of partial occlusion is a challenging task for the researchers in the fields of computer vision and video processing. In modern coding standards, like MPEG-4 and MPEG-7, the term of video objects is used to define moving objects in a video sequence. Automatic extraction of such objects is by no means trivial, and occlusion is one of most important problems. This paper reformulates one of the most popular deformable templates for shape modeling and object tracking, the Snakes, in a probabilistic manner, in order to include providence for partial occlusion of the moving objects. Experiments of object tracking in partial occlusion, in complex natural sequences, where temporal clutter, abrupt motion and external lighting changes have been carried out, showing the efficiency of the proposed approach.

## 1. INTRODUCTION

The deformable templates known as active contours have come up and drawn special attention in the last decade, due to their efficiency in dealing with problems like image and motion segmentation [7], object detection, localization and tracking in video sequences [8, 10]. A major category of active contours, the Snakes, has been successfully applied to a variety of problems, such as edge detection and tracking [3]. Snakes are based on energy minimization along a curve; that is a curve deforms its shape so as to minimize an “internal” and an “external” energy along its boundary. However, Snakes are not problem free models: parameter tuning, shape initialization and the complexity of the energy minimization procedure are problems related to the applicability of Snakes, while the relatively poor performance in images and video sequences with complex and noisy background is mainly due to the fact that Snakes are one of the most representative examples of boundary-based active contours. Object tracking in partial occlusion is a problem which seems to have similarities with object tracking in complex back-

ground, but it is actually much harder. Partial occlusion hides some parts of the object and in this case the external (image dependent) energy term, which contributes to the total snake energy, has to be ignored. The basic issue here is to understand when partial occlusion occurs. In our approach the movement history of the snake is used a predictor of its position in the next frame. Prediction serves two purposes: (a) it provides a reasonable shape initialization and (b) it can be used for detecting partial occlusion.

The probabilistic approach of tracking contours that we use, utilizes a snake model, which is based on both boundary and region information. Region motion information about the past and the present motion of the tracked object is provided by a motion estimation scheme [1] and, based on the object’s motion history, a shape prior knowledge (boundary-based), indicated as *uncertainty region*, is extracted. The moving contour is estimated inside that region, according to an energy function. A force-based implementation, following a steepest descent-alike method is used to approximate the global minimum solution of the snake energy function. Furthermore, the obtained solution should obey several constraints in order to identify the boundaries of occluded parts of the moving objects.

## 2. THEORETICAL BACKGROUND

Active contour models were first introduced by Kass et al. [5] through the so-called snake models. In general, snakes concern model and image data analysis, through the definition of a linear energy function and a set of regularization parameters. This energy function consists of two parts, the external energy term  $E_{ext}$ , which depends on the image data according to a chosen criterion and the internal energy term  $E_{int}$ , which enforces smoothness along the snake. The goal is to minimize the total snake energy and this is achieved iteratively, after considering an initial estimate for the object shape.

The proposed total energy  $E_{snake}$  of the snake  $C_{snake}$ , is

given by  $E_{snake} = E_{int} + E_{ext}$ , where  $E_{int}$  is defined in terms of snake local curvature  $CU_{snake}$  and elasticity  $DV_{snake}$ , whereas  $E_{ext}$  is defined with the use of a modified image gradient  $G_m$ , replacing the commonly used laplacian-of-gaussian term  $|\nabla G_\sigma * I|$ , which introduces noise in the snake models. More information about the definitions of the proposed energy terms can be found in [9].

Before applying the tracking model in the current frame of a sequence, as described in the following section, we preprocess the image to eliminate noise, with the use of a morphological Alternating Sequential Filter (ASF), based on successive morphological area opening and closing operations with structure elements of increasing scale and different orientations [6]. The main advantage of such filters is that they preserve line-type image structures (edges), which is impossible to be achieved, for example, with median filtering. The modified image gradient  $G_m$  used for our purposes is actually a part of the Watershed transformation in image segmentation problems [6] and consists of the extraction of binary image markers, through a morphological geodesic erosion reconstruction of the image gradient, and successive morphological conditional erosions of these markers, so that they constitute the only local minima of the image gradient.

### 2.1. Motion Estimates Extraction

The correct extraction of moving edges in terms of position and direction is important and aids the accurate estimation of an object's position from the current to the next frame. Several existing techniques are able to adequately cope with the difficult problem of optical flow recovery given that their assumptions hold. The challenge is to achieve high robustness against strong assumption violations commonly met in real sequences. We adopt the motion estimation technique proposed by Black et al. [1] as an efficient tool for overcoming these violations. They reformulate the objective function, which consists of the optical flow equation and the spatial coherence constraint, in order to include the robust statistics tools [4] in an almost straightforward way. They simply take the standard least-squares formulation of optical flow and use a robust estimator instead of the quadratic one. This approximation is then minimized using a coarse-to-fine (multiresolution) simultaneous over-relaxation technique. The proposed reformulation results in an area-based regression technique that is robust to multiple motions due to occlusion, transparency or specular reflections and compensates for over-smoothing and noise sensitivity.

## 3. OBJECT TRACKING

The main issues that the proposed tracking approach is called upon to cope with, are: (a) non-rigid (deformable) moving

objects, (b) moving objects with a complicated (not smooth) contour, (c) object movements which are not simple translations, but also include rotations and motion towards the camera capturing the sequence, (d) sequences with strong existence of temporal clutter and external lighting changes, (e) sequences with complex background (common case in natural sequences), and (f) moving object partial occlusions. The issues (a)-(d) are covered by the formulation of probabilistic approach of snakes as it is shown in [9]. The issues (e) and (f) are the subjects of the current work and are discussed in the following sections.

### 3.1. Force-Based Approach

Given the proposed snake model presented in Section 2, the first step is to extract some regions around the curve, which are described as *uncertainty regions*. This is achieved by exploiting the motion history of the tracked contour (curve points' motion in previous time instances), estimated with the use of the motion estimation scheme proposed in subsection 2.1: the previously estimated contour is deformed according to the previously estimated point motion and the standard deviation of each point's mean motion is calculated; the uncertainty region around each point is then defined in terms of its corresponding standard deviation. The next step is to find the new position of each point of the curve, inside its corresponding uncertainty region, which corresponds to the minimum of a criterion, which is defined by the snake's energy terms described Section 2:

$$\mathbf{C}^{(I+1)} = \arg \min_{\mathbf{r} \in R} [w_1 \cdot D_{CU}^{(\mathbf{r})} + w_2 \cdot D_{DV}^{(\mathbf{r})} + w_3 \cdot E_{ext}^{(\mathbf{r})}], \quad (1)$$

where  $D_{CU}^{(\mathbf{r})}$  and  $D_{DV}^{(\mathbf{r})}$  are the differences between the curvatures and elasticities of the contour  $\mathbf{C}^{(I)}$  and the curve  $\mathbf{r} \in R$  at the  $k$ -th point, respectively. The set of all possible curves  $R$  emerge by oscillating the points of the curve inside the uncertainty region  $U$ . Finally, the weights  $w_1$  and  $w_2$  are computed according to the curvature  $CU_{snake}$  and elasticity  $DV_{snake}$  distributions' zero crossings, that denote how smooth and elastic the curve is, while  $w_3$  is computed in terms of the mean value of  $G_m$  inside the uncertainty region  $U$ , that denote how smooth (homogeneous) this region is.

The minimization of the equation (1) is a procedure of high complexity. This is a common problem when applying the minimization procedure of snakes and therefore we adopt a force-based approach, as described in the following, instead of using a dynamic programming algorithm. According to that approach, energy terms are converted in forces resulting to the minimizing of them. The initial estimation of the object's contour  $\mathbf{C}_{init}^{(I+1)} = [C_{init,k}^{(I+1)} | k = 1 \dots N]$  in the frame  $I + 1$  is computed based on contour's current location and

its instant motion, i.e.,

$$\mathbf{C}_{init}^{(I+1)} = \mathbf{C}^{(I)} + \mathbf{m}_c^{(I)}, \quad (2)$$

where  $\mathbf{m}_c^{(I)} = [m_{(c,1)}^{(I)}, \dots, m_{(c,N)}^{(I)}]$  is the estimated motion of the contour points in the frame  $I$  (obtained using the technique described in 2.1). Then, the internal snake forces are defined as,  $\mathbf{F}_c(k) = w_1(k) \cdot [CU_{(\mathbf{C}^{(I)})}(k) - CU_{(\mathbf{C}_{init}^{(I+1)})}(k)] \cdot \mathbf{n}_k^{(I+1)}$  and  $\mathbf{F}_d(k) = w_2(k) \cdot [DV_{(\mathbf{C}^{(I)})}(k) - DV_{(\mathbf{C}_{init}^{(I+1)})}(k)] \cdot \mathbf{t}_k^{(I+1)}$ , where  $\mathbf{t}_k^{(I+1)}$  and  $\mathbf{n}_k^{(I+1)}$  are the tangential unit and normal vectors of the curve  $\mathbf{C}_{init}^{(I+1)}$  at each point  $k$ , respectively. The forces  $\mathbf{F}_d(k)$  represent the stretching components (elasticity), whereas  $\mathbf{F}_c(k)$  are the deformations of the curve along its normal directions (local curvature).

Given the snake's uncertainty region  $\mathbf{U}$ , for each point  $k$  we define a function  $g_{m,k}(\mathbf{p})$ , constituting of all pixels  $\mathbf{p}$  of  $G_m$ , inside the uncertainty region  $\mathbf{U}$ , lying on the line segment that is defined by the normal direction of the curve  $\mathbf{C}_{init}^{(I+1)}$  at point  $k$ :

$$g_{m,k}(\mathbf{p}) = [G_m(\mathbf{p}) | (\mathbf{C}_{init,k}^{(I+1)} - \mathbf{p}) \cdot \mathbf{n}_k^{(I+1)} = 1, \mathbf{p} \in \mathbf{U}] \quad (3)$$

$$\tilde{\mathbf{p}}_k = \arg \max_{\mathbf{p}} [g_{m,k}(\mathbf{p})], \quad (4)$$

$\tilde{\mathbf{p}}_k$  determines the most salient edge-pixel in the line segment defined above and thus defines the direction of the external snake force

$$\mathbf{F}_e(k) = w_3(k) \cdot \text{sgn}_k \cdot e_{ext}(k) \cdot \mathbf{n}_k^{(I+1)}, \quad (5)$$

where  $w_3(k) = \|\mathbf{k} - \tilde{\mathbf{p}}_k\|$ ,  $\mathbf{k}$  is the pixel corresponding to the contour point  $k$ , and  $\text{sgn}_k$  denotes the direction of the force:  $\text{sgn} = '+'$  if  $\tilde{\mathbf{p}}$  is inside the curve, and  $\text{sgn} = '-'$  otherwise. From the definition of the external energy term it can be seen that it takes values close to zero in regions of high image gradient ( $G_m^2(k) \simeq 1$ ) and values close to unity in regions with relatively constant intensity (smooth) ( $G_m^2(k) \simeq 0$ ).

In the force-based approach, the initialization of the examined curve  $\mathbf{C}_{init}^{(I+1)}$  marches towards the object's boundaries in the next frame  $I + 1$ , according to the forces applied to it. Thus, the minimization of equation (1) can be approximated by using the internal and external snake forces defined above, in an iterative manner similar to the steepest descent approach [2]. The final contour  $\mathbf{C}^{(I+1)}$  is obtained when one of the following criteria is satisfied: (a) if the resultant force is greater than the one of the next iteration, or (b) the maximum number of iterations is reached. It must be noted that the use of the proposed steepest descent approach does not ensure that the final contour corresponds to the solution of the equation (1), but under the constraints we pose, even if  $\mathbf{C}^{(I+1)}$  corresponds to a local minimum, it is close to the desired solution (global minimum).

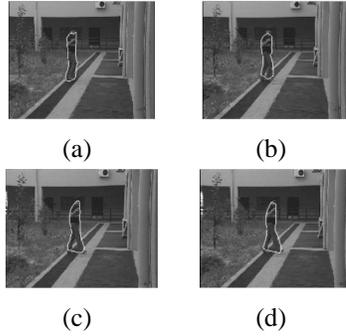
### 3.2. Rule-driven Handling of Partial Occlusion

In order to separate background and object regions, especially when the background is not homogeneous (smooth), as well as to cope with moving object's partial occlusion that may occur, we introduce more constraints that  $\tilde{\mathbf{p}}_k$  (eq. (4)) must obey, so that its estimation will be reasonable. Without loss of generality, we suppose that the background is static and possible occluding objects are also static. Let  $\bar{m}_c^{(I)} = \frac{1}{N} \sum_{i=1}^N m_{c,i}^{(I)}$  be the mean estimated motion of the contour points at frame  $I$  and  $\mathbf{k} - 1$  and  $\mathbf{k} + 1$  be the surrounding pixels of point  $\mathbf{k}$  on the line segment, along which the function  $g_{m,k}$  is calculated, then (a)  $\tilde{\mathbf{p}}_k$  must divide that line segment in two parts: an immiscibly moving and a immiscibly static one, that is  $u(\mathbf{k} - 1) \simeq \bar{m}_c^{(I)}$  and  $u(\mathbf{k} + 1) \simeq 0$ , or  $u(\mathbf{k} - 1) \simeq 0$  and  $u(\mathbf{k} + 1) \simeq \bar{m}_c^{(I)}$ , and (b)  $\tilde{\mathbf{p}}_k$  must be a moving point with velocity close to  $\bar{m}_c^{(I)}$ , where  $u(\cdot)$  denotes the instant velocity.

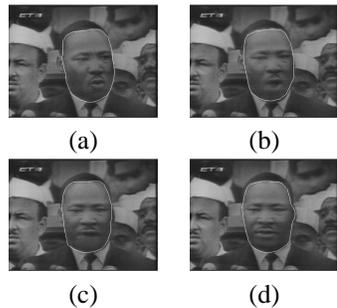
Thus, taking the above constraints into consideration, we overcome cases such as (a) when the maximum is found in background: it is not a moving one and does not separate two immiscible (according to the motion) parts of the function  $g_m$ , (b) when the maximum is found inside the moving object region: although it is a moving one, it does not divide the function  $g_m$  in such two parts, (c) when occlusion occurs and the maximum is on the occluding object boundary: the maximum is not moving, although it makes the region  $g_m$  separation and (d) when occlusion occurs and the maximum is in the occluding object region: neither the maximum is moving, nor it makes such a separation. In these cases, where these two constraints are not reached, we ignore the external force and evolve the curve according to its internal forces; in this way, we can obtain contours similar to the ones in the past frames.

## 4. EXPERIMENTAL RESULTS

In Figure 1, the proposed approach is applied to a strongly cluttered sequence, where the desired object is deforming along the time; the accuracy of the method is based on the snake's external energy definition through the image modified gradient utilization and the ASF pre-filtering. On the other hand, Figure 2 represents an indoor human head tracking: despite the complicated background, the rule-based constraints proposed in subsection 3.2 separate correctly the moving head from background. Finally, Figure 3 illustrates a case of two moving objects, where, as time goes by, the one is getting partially occluded by a static obstacle, while the other is moving in the front of the obstacle. In sub-figures 3 (a1-d1), the motion estimates are illustrated, showing that the noise is effectively eliminated on the boundaries between the static and moving regions even when the occlusion occurs, whereas the respective sub-figures 3 (a2-d2)



**Fig. 1.** Example of a man walking in a cluttered sequence.



**Fig. 2.** Head tracking in with complicated background.

show that both objects' contours are estimated with sufficient accuracy, due to the additional constraints in which the maximum of equation (4) is imposed.

## 5. CONCLUSIONS

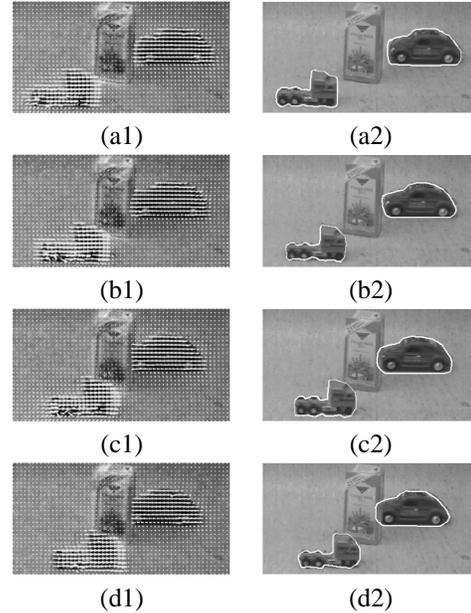
In this work we have presented a rule-driven application of Snakes for object tracking in partial occlusion and complex backgrounds. Rules represent constraints on the continuity of the contour motion in successive frames as well as differences in the mean motion of the object and the background or the occluding object. Obtained results are encouraging, while the authors are working so as to include mathematical representation of the rules into the snake's energy minimization function.

## 6. REFERENCES

[1] M.J. Black and P. Anandan, "The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields," *CVIU*, 63(1), pp. 75-104, 1996.

[2] S. Haykin, "Neural Networks" Macmillan College Publishing Company, *Section 5.3: The Steepest Descent Method*, pp. 124-126, 1994.

[3] M. Hoch and P. Litwinowicz, "A Practical Solution for



**Fig. 3.** Occlusion case: motion vectors and obtained object contours.

Tracking Edges in Image Sequences with Snakes," *The Visual Computer*, 12(2), pp. 75-83, 1996.

- [4] P. Huber, *Robust Statistics*, Wiley eds., NY, 1981.
- [5] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active Contour Models," *Int. Journal of Comp. Vis.*, 1(4), pp. 321-331, 1988.
- [6] V.K Madisetti and D.B Williams Eds., *The Digital Signal Processing Handbook*, CRC Press, 1998, Chap. 74, pp. 20-26.
- [7] A. Mansouri, T. Chomaud and J. Konrad, "A comparative evaluation of algorithms for fast computation of level set PDEs with applications to motion segmentation," *In Proc. of ICIP'01*, pp. 636 -639, vol.3, Thessaloniki, Greece, 2001.
- [8] N. Paragios and R. Deriche, "Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects," *IEEE Trans. on PAMI* 22(3), pp. 266-280, 2000.
- [9] G. Tsechpenakis, N. Tsapatsoulis and S. Kollias, "Probabilistic Boundary-Based Contour Tracking with Snakes in Natural Cluttered Video Sequences," *submitted to Int. Journal of Image and Graphics (IJIG)*. <http://www.image.ece.ntua.gr/~gtsech/>
- [10] C. Vieren, F. Cabestaing and J.G. Postaire, "Catching Moving Objects with Snakes for Motion Tracking" *Pattern Recognition Letters*, 16, pp. 679-685, 1995.