MOTION-BASED MOVING OBJECT TRACKING USING AN ACTIVE CONTOUR

Boo Hwan Lee[1], Il Choi[2], Gi Joon Jeon[3]

[1]Agency for Defense Development, Yuseong P.O. Box 35-1, Daejeon, S. Korea, [2]Research Institute, 3B System Inc., Daegu, S. Korea, [3]School of Electrical Engineering and Computer Science, Kyungpook National University, Daegu, S. Korea, bhlee@add.re.kr, ichoi@3bsystem.co.kr, gjjeon@ee.knu.ac.kr

ABSTRACT

It is quite important to decide the local converging direction of an active contour for correctly extracting the boundary of a moving object with a deformable shape. Thus, a new energy function for the active contour is proposed based on the addition of a directional energy term using a frame difference map to the greedy snake. The frame difference map is used to obtain the motion information of an object with a faster and non-rigid motion. Plus, updating rules for the frame difference map are also developed to encourage a stable convergence of the contour points. Experimental results on real image sequences showed that the proposed method could fully track a speedy deformable object while exactly extracting the boundary of the object in every frame.

1. INTRODUCTION

The active contours (or snakes) are dynamic curves moving within an image domain to capture desired image features. The curve motion is driven by the influence of the internal energy within the curve itself and the external energy derived from the image data, which achieves a minimal state of energy when the snakes conform to an object boundary or other desired features. The active contour models are widely used in computer vision, mainly for image segmentation [1-3] and motion tracking [4-5].

The original snake introduced by Kass et al. [6] and other improved snake models all adopt a variational approach to complete their energy minimization. The search-based technique, such as a dynamic programming [7], has also been proposed for minimization. However, those approaches give rise to the problems of the numerical stability and the computational burden to obtain admissible solutions. Williams et al. [8] then proposed the greedy algorithm to solve those problems. The greedy-algorithm-based snake reasonably combines speed, flexibility and simplicity [1]. Nevertheless, the greedy snake still has several problems like the former snakes. One problem is a narrow capture range, that is, the initial contour has to be close to the true boundary, otherwise, it would be likely to converge to the wrong result. The gradient vector flow (GVF) snake [2] is an effective method to solve this problem. The GVF snake has a remarkably large capture range, making is less sensitive to the initialization. However, this method has a larger computational cost due to the involvement of vector diffusion, plus it is still sensitive to initialization despite its expanded its capture range.

Some of the approaches mentioned above have already been applied to the problem of contour tracking. In the original work by Kass et al. [6] and followed by Leymarie and Levine [4], the snake unable to track a contour if the initialization provided by the contour detected in the previous frame is not close enough to the newly deformed and displaced contour in the current frame. Other techniques using motion estimation have been developed. Kalman snake [9] is appropriate when the motion of the boundary is predictable with no large deformations between successive frames. Meanwhile, an error term is used for motion compensation by Pardas and Sayrol [5], yet their block matching-based snake is unable to cope with non-rigid objects changing their shapes drastically in consecutive image sequences.

Accordingly, this paper uses the greedy snake as the basis for a new motion-based boundary tracking method for a rigid or a non-rigid moving object in an image sequence. To cope with shape change of the moving object in tracking applications, a new directional energy function based on a frame difference (FD) map is introduced into the greedy snake to decide the local converging directions of the contour points. In addition, updating rules for the FD map are developed to encourage a stable convergence of the contour points. The introduced energy function allows the snake to increase the region of uniform convergence. The proposed snake is not only faster and more efficient, but also able to inflate or deflate its contour automatically. It has been compared with the greedy and GVF snakes using the real image sequences.

2. GREEDY SNAKE

An active contour for a discrete curve is represented as a set of contour points, snaxels, $\mathbf{v} = \{\mathbf{v}_i = (x_i, y_i) \mid i = 0,\ 1, \cdots, n-1\}$, where n is the number of the snaxels and (x_i, y_i) is the Cartesian coordinates of the snaxel i. The energy $\mathrm{E}(\mathbf{v})$ of the contour $\mathbf{v}$, which is going to be minimized, is defined as

$$\mathrm{E}(\mathbf{v}) = \sum_{i=0}^{n-1} [\mathrm{E}_{\mathrm{int}}(\mathbf{v}_i) + \mathrm{E}_{\mathrm{ext}}(\mathbf{v}_i)], \tag{1}$$

where $\mathrm{E}_{\mathrm{int}}$ and $\mathrm{E}_{\mathrm{ext}}$ represent the internal and the external energies of the snake, respectively. The internal energy is composed of the first-order continuity term E_{con} and the second-order curvature term E_{cur}. The external energy usually adopts the image term E_{image}. Description of these energies is given in [1, 8]. The local energy $\mathrm{E}(\mathbf{v}_i^j)$ is defined as

$$\mathrm{E}(\mathbf{v}_i^j) = \alpha \mathrm{E}_{con}(\mathbf{v}_i^j) + \beta \mathrm{E}_{cur}(\mathbf{v}_i^j) + \gamma \mathrm{E}_{image}(\mathbf{v}_i^j),\ j = 0,\ 1,\ \cdots, 8, \tag{2}$$

where $\mathbf{v}_i = \mathbf{v}_i^0$ and $\mathbf{v}_i^j (j \neq 0)$ represents the eight-neighborhood points of the $\mathbf{v}_i$. The α, β and γ are the weighting parameters that control the snake's elasticity, rigidity, and the attraction to the desired image feature, respectively. In the energy minimization process, every snaxel on the contour is iteratively deformed until convergence using a given termination criterion. For each snaxel $\mathbf{v}_i$, the energy function is computed for every point belonging to the eight-neighborhood of the $\mathbf{v}_i$. The point with minimum energy then replaces the current snaxel $\mathbf{v}_i$.

3. PROPOSED SNAKE MODEL

When an object moves fast and changes its shape concurrently between the two consecutive frames, its position and shape in the next frame can deviate significantly from those in the current frame shown in Fig. 1(a). Thus, since the greedy snake can not fully enclose the boundary of the object in the next frame, it loses the target object, which means the greedy snake can not be used to directly track a rigid or the non-rigid object with a large deformation between the two consecutive frames.

To cope with the shape change of a moving object, adaptive classification of the snaxels into two groups is an attractive option, as shown in Fig. 1(b), where the inflating snaxels cover the partial growth of the object, while the deflating snaxels cover the partial shrinkage of the object. In Fig. 1(b), the initial contour AB represented by the dashed line has to deflate in order to deform itself toward the contour AB represented by the solid line. Conversely, the initial contour BC represented by the dashed line must inflate in order to deform itself toward the contour BC represented by the solid line. To accommodate this strategy, a new directional energy term based on a FD map is introduced to the greedy snake to determine the local converging directions of the corresponding snaxels. In addition, updating rules for the FD map are also included to encourage a stable convergence of the snaxels.

(a) (b)

Fig. 1. (a) Two consecutive frames. Head and shoulder images marked in the white box have a large motion. (b) Basic idea of proposed snake model.

3.1. Directional Energy Based on a FD Map

3.1.1 FD Map

When an image sequence is captured by a static observer or a stabilized platform mounted on a camera, it is possible to suppose that the background is usually stationary or has a simple global motion. Thus, the background can be removed by simply differencing between the two successive frames. Moreover, it is well known that the FD is more suitable for objects with faster and non-rigid motions. Therefore, the FD is used to obtain the motion information on an object with deformation. The decision, whether a spatial position $\mathbf{x}=(x,y)$ belongs to the changed or unchanged image part of two frames $k-1$ and k, is based on the evaluation of the FD defined as

$$FD_k(\mathbf{x}) = I_k(\mathbf{x}) - \hat{I}_{k-1}(\mathbf{x}), \tag{3}$$

where $\hat{I}_k(\mathbf{x}) = (1-w)\times I_k(\mathbf{x}) - w\times \hat{I}_{k-1}(\mathbf{x})$ and w is a weighting factor ranging from 0 to 1. To distinguish between relevant changes due to object motion or brightness changes and irrelevant temporal changes due to noise, the FD is compared to the threshold I_{th}. Consequently, the moving parts of an object are classified into three different regions, and the FD map $M_k(\mathbf{x})$ is generated by

$$M_k(\mathbf{x}) = \begin{cases} 1 & \text{(leading edge),} \quad \text{if } FD_k(\mathbf{x}) < I_{th} \\ -1 & \text{(trailing edge),} \quad \text{if } FD_k(\mathbf{x}) < -I_{th}. \\ 0 & \text{(background),} \quad \text{otherwise} \end{cases} \tag{4}$$

3.1.2 Directional Energy

A new directional energy term is introduced as the external energy for the proposed snake to determine whether the snaxels are to be inflated or deflated. The directional energy term is derived as follows. First, to decompose the snaxels into inflating and deflating snaxels, the average value of the FD map $M_k(\mathbf{x})$ is computed in the eight-neighborhood of $\mathbf{v}_i$. If the average value is greater than zero, $\mathbf{v}_i$ becomes an inflating snaxel, if it is less than zero, $\mathbf{v}_i$ becomes a deflating snaxel, and if the value is zero, the snaxel acts the same as the greedy snake. Second, to determine the next updating point of $\mathbf{v}_i$, which corresponds to the point with the minimum energy, the converging direction d_i^j is derived as follows. Let $\mathbf{p}_i^j = (\mathbf{v}_i^j - \mathbf{v}_i) \Big/ \left\| \mathbf{v}_i^j - \mathbf{v}_i \right\|$ be the directional unit vector connecting the two points $\mathbf{v}_i^j$ and $\mathbf{v}_i$. Let $\mathbf{n}_i$ be the unit vector lying at right angles to the vector connecting the two adjacent snaxels of $\mathbf{v}_{i-1}$ and $\mathbf{v}_{i+1}$ at $\mathbf{v}_i$. The direction of vector $\mathbf{n}_i$ has two different directions, i.e. outward and inward. Thus, let vectors $\mathbf{n}_i^+$ and $\mathbf{n}_i^-$, as shown in Fig. 2(a), be the outward and inward directions vectors for vector $\mathbf{n}_i$ at $\mathbf{v}_i$, respectively. If $\mathbf{v}_i$ is an inflating snaxel, the converging direction is the inner product of the two vectors, $\mathbf{p}_i^j$ and $\mathbf{n}_i^+$, which is defined by $d_i^j = \mathbf{p}_i^j \cdot \mathbf{n}_i^+$. Meanwhile, If $\mathbf{v}_i$ is a deflating snaxel, the converging direction is defined as $d_i^j = \mathbf{p}_i^j \cdot \mathbf{n}_i^-$. Finally, the directional energy term can be defined as

$$\mathrm{E}_{dir}(\mathbf{v}_i^j) = 1 - \left| M_k(\mathbf{v}_i^j) \times d_i^j \right|, \tag{5}$$

where $M_k(\mathbf{v}_i^j)$ is the value of the FD map at $\mathbf{v}_i^j$. Note, if $\mathbf{v}_i$ lies in the background region, $M_k(\mathbf{v}_i^j)$ is zero for every point belonging to the eight-neighborhood of the $\mathbf{v}_i$. Thus, the directional energies have the same value, as $M_k(\mathbf{v}_i^j) = 0$, which means the role of $\mathbf{v}_i$ becomes the same as that of the greedy snake.

The validity of Eq. (5) is explained as follows. Consider that the inflating snaxel $\mathbf{v}_i$ lies on the leading edge, as shown in Fig. 2(a). For the case of $\mathbf{v}_i^1$ in the eight-neighborhood of $\mathbf{v}_i$, the second term in Eq. (5) is at its minimum value, as $M_k(\mathbf{v}_i^1) = 0$ and $\mathbf{p}_i^1 \cdot \mathbf{n}_i^+ = -1$, whereas the directional energy $\mathrm{E}_{dir}(\mathbf{v}_i^1)$ is at its maximum value, thus $\mathbf{v}_i$ does not deflate toward the $\mathbf{v}_i^1$ due to the energy minimization rule. Meanwhile, for the case of $\mathbf{v}_i^8$, the second term in Eq. (5) is at its maximum value, as $M_k(\mathbf{v}_i^8) = 1$ and $\mathbf{p}_i^8 \cdot \mathbf{n}_i^+ = 1$, whereas the directional energy $\mathrm{E}_{dir}(\mathbf{v}_i^8)$ is at its minimum value, thus $\mathbf{v}_i$ can inflate toward $\mathbf{v}_i^8$, as shown in Fig. 2(b). Similarly, $\mathrm{E}_{dir}(\mathbf{v}_l^j)$ can adaptively determine the converging direction of the deflating snaxel $\mathbf{v}_l$ lying on the trailing edge, as shown in Fig. 2(b).

By introducing Eq. (5) into Eq. (2), a new energy function can be defined as

$$\begin{aligned} \mathrm{E}_{\text{new}}(\mathbf{v}_i^j) = \alpha\, \mathrm{E}_{con}(\mathbf{v}_i^j) + \beta\, \mathrm{E}_{cur}(\mathbf{v}_i^j) + \gamma\, \mathrm{E}_{image}(\mathbf{v}_i^j) + \\ \zeta \mathrm{E}_{dir}(\mathbf{v}_i^j), \quad j = 0,1,\cdots,8, \end{aligned} \tag{6}$$

where $\alpha + \beta + \gamma + \zeta = 1$. Finally, the new updated snaxel $\mathbf{v}_i'$, as shown in Fig. 5(b), is determined as

$$\mathbf{v}_i' = \arg\min_{\mathbf{v}_i^j} \mathrm{E}_{\text{new}}(\mathbf{v}_i^j), \quad j = 0,1,\ldots,8. \tag{7}$$

Every snaxel is sequentially updated in a clockwise direction using Eq. (7), and this is repeated until all the snaxels converge at the boundary of the object. Note, since Eq. (5) provides the same value for both $\mathbf{v}_i'^1$ and $\mathbf{v}_i'^8$ for the updated snaxel $\mathbf{v}_i'$, as shown in Fig. 2(c). When the next updating snaxel $\mathbf{v}_i'$, it is impossible to

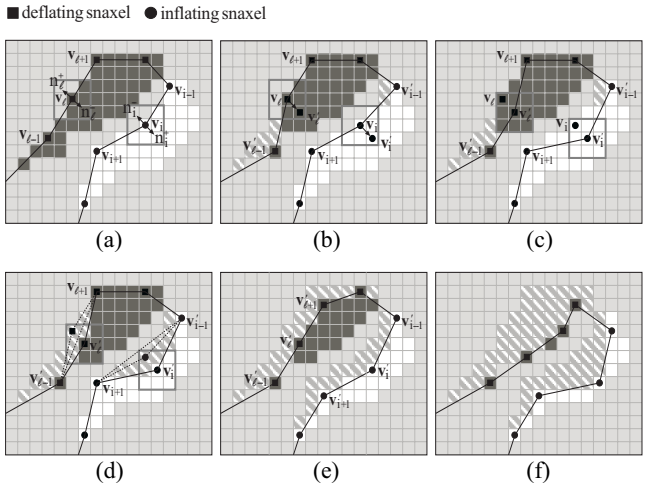


Fig. 2. Example of updating snaxels using directional energy. The white, black, and gray regions represent the leading edges, trailing edges, and background of FD map, respectively.

determine whether its location is at $\mathbf{v}_i'^1$ or $\mathbf{v}_i'^8$ in the updating procedure. Accordingly, the FD map represented by the shaded area shown in Fig. 2(d) must be updated with the value 0 during the period right after a specific snaxel has just been updated and before the next snaxel is updated. After updating the FD map represented by the shaded area shown in Fig. 2(e), the next updating point for snaxel $\mathbf{v}_i'$ can be determined using $\mathbf{v}_i'^8$. Thus, every snaxel can deform itself toward points on the boundary of the object, as shown in Fig. 2(f), through repeatedly relocating and updating the FD map.

3.1.3 Updating the FD Map

After updating the location of a snaxel using Eq. (7), the FD map must be locally updated with the value 0 around the snaxel. For this, updating rules for the FD map are proposed as follows. First, the local shape of the three adjacent snaxels $\mathbf{v}'_{i-1}$, $\mathbf{v}_i$, $\mathbf{v}_{i+1}$, including the currently considered snaxel $\mathbf{v}_i$, is categorized as convex or concave, as shown in Fig. 3. Here, θ is the counter clockwise angle from the horizontal axis to the line connecting $\mathbf{v}'_{i-1}$ and $\mathbf{v}_i$ together, while angle ϕ is the counter clockwise angle from the horizontal axis to the line connecting $\mathbf{v}'_{i-1}$ and $\mathbf{v}_{i+1}$ together. If $\theta \geq \phi$, it is categorized as convex, as shown in Fig. 3(a) and 3(b). Otherwise, if $\theta < \phi$, it is categorized as concave, as shown in Fig. 3(c) and 3(d). Second, the three adjacent snaxels including the currently considered snaxel, are located either within the right half plane (RHP) or the left half plane (LHP). As shown in Fig. 3(a) and 3(c), if the sequence of the snaxels $\mathbf{v}'_{i-1}$, $\mathbf{v}_i$, $\mathbf{v}_{i+1}$ proceeds in a clockwise direction from 12 to 6 o'clock, it belongs to the RHP, whereas if the sequence proceeds in a clockwise direction from 6 to 12 o'clock, it belongs to the LHP, as shown in Fig. 3(b) and 3(d). Third, the currently considered snaxel is either inflated or deflated based on the average value of the $M_k(\mathbf{x})$ in the eight-neighborhood of $\mathbf{v}_i$. Finally, the updatable region of the FD map at a specific snaxel is determined based on a combination of the three rules mentioned above. All the updating FD map categories are given by

C1: Convex $\cap$ RHP $\cap$ Inflation $\equiv$ Concave $\cap$ LHP $\cap$ Deflation
C2: Convex $\cap$ RHP $\cap$ Deflation $\equiv$ Concave $\cap$ LHP $\cap$ Inflation
C3: Concave $\cap$ RHP $\cap$ Inflation $\equiv$ Convex $\cap$ LHP $\cap$ Deflation
C4: Concave $\cap$ RHP $\cap$ Deflation $\equiv$ Convex $\cap$ LHP $\cap$ Inflation.

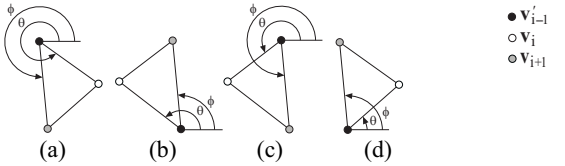


Fig. 3. (a) Convex and RHP. (b) Convex and LHP. (c) Concave and RHP. (d) Concave and LHP.

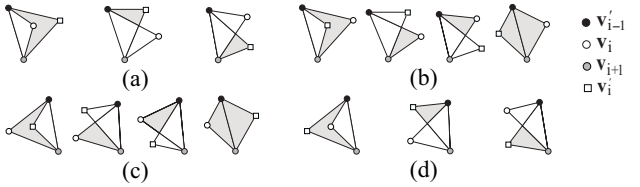


Fig. 4. Updatable regions of FD map. (a) C1 case. (b) C2 case. (c) C3 case. (d) C4 case.

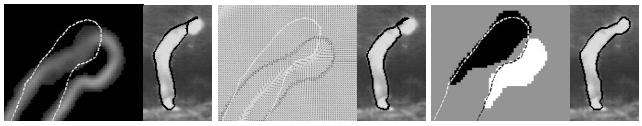
Figure 4 shows the updatable regions of the FD map according to each category. Here, the gray area represents the regions of the FD map to be updated. These regions are easily defined by one or two triangles based on the simplex coordinates [10].

4. EXPERIMENTAL RESULTS AND DICUSSION

Several experiments on the three kinds of real image sequences were conducted to examine the performance of the proposed snake model. The two image sequences were acquired by infrared imaging device with wavelength $3 \sim 5\mu m$. One is a bowing woman and the other is a walking woman with 320×240 pixels. The third sequence acquired by infrared imaging device with wavelength $8 \sim 12\mu m$ is a moving tank with 640×480 pixels. To fully explore the proposed snake, we compared it with the greedy and GVF snakes.

Figure 5 shows the boundary tracking results of the greedy, GVF, and proposed snakes for the two consecutive frames in the bowing woman image sequence as shown in Fig. 1(a). Since the initial contour does not lie within the capture range, the greedy snake with $\alpha = 0.15, \beta = 0.15, \gamma = 0.3$ fails to deform it toward the boundary of the head and shoulder as shown in Fig. 5(a). The GVF snake with $\alpha = 0.05, \beta = 0.0, \mu = 0.2$, as shown in Fig. 5(b), also does not evolve correctly its shape toward the boundary because the initial contour does not enclose the 'center of divergence' [11] of the GVF field. Whereas the proposed snake with $\alpha = 0.15$, $\beta = 0.15, \gamma = 0.3, \zeta = 0.4, w = 0.3$, as shown in Fig. 5(c), shows that it is less sensitive to the initialization and is able to track the speedy non-rigid object. In addition, Fig. 6 shows the evolving of the contour and the updating of the FD map. It can be easily seen that the initial contour, which is far away from the target object, converges correctly toward the boundary of the object while inflating and deflating its contour automatically with updating the FD map under the control of the proposed directional energy.

Figure 7 shows the results of the greedy, GVF, and proposed snakes for walking woman. In this experiment, the location of the initial contour was the same for all the snakes, while the values of the parameters were the same as those in Fig. 5. Figure 7(a) shows the four consecutive frames for walking woman image sequence, where significant motion occurs with the left leg, head, and neck among the successive frames. The greedy and GVF snakes were unable to deform toward the boundary of the walking woman due to their significant motion as shown in Fig. 7(b) and 7(c),



(a) (b) (c)

Fig. 5. Comparison of results for boundary tracking of bowing woman. The initial contours represented by the dashed line are overlaid on (a) the Gaussian weighted edge image, (b) the GVF vector field, and (c) the FD map marked in the white box shown in Fig. 1(a), respectively. The final contours, represented by the solid lines, were converged by using (a) the greedy, (b) GVF, and (c) proposed snakes, respectively.

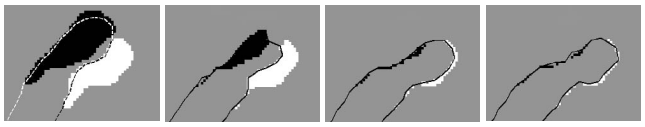
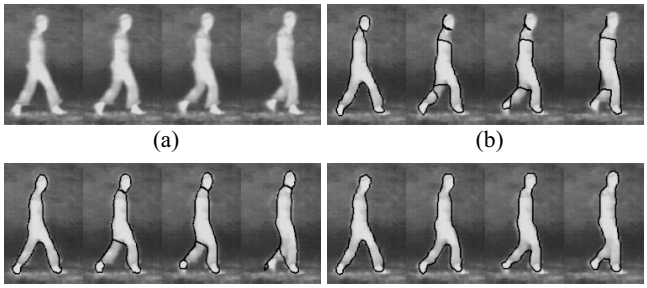


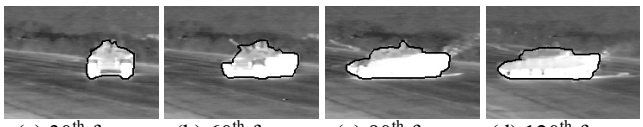
Fig. 6. Contour evolving and FD map updating sequences.



(a) (b)

(c) (d)

Fig. 7. Comparison of results for boundary tracking of walking woman. (a) Walking woman image sequence. The final contours, represented by the solid lines, were converged using (b) the greedy, (c) GVF, and (d) proposed snakes, respectively.



(a) 30th frame (b) 60th frame (c) 90th frame (d) 120th frame

Fig. 8. Tracking of moving tank within a short range using the proposed snake.

respectively. In contrast, despite the non-rigid walking motion, the proposed snake produced a satisfactory performance, as shown in Fig. 7(d).

Figure 8 shows the results for the moving tank within a short range, where the vehicle initially moved to the front then turned in opposite direction. The values of parameters in this experiment were the same as those in Fig. 5, and $w = 0.5$. Note that the proposed snake successfully converged on the final object in image sequence. For this image sequence, the results for the GVF snake were almost the same as those for the proposed snake, however, the greedy snake was unable to extract the object boundaries due to the fast motion and shape changes (These results are not shown here to the limitation of pages).

A comparison of the computation time for the greedy, GVF and proposed snakes was also carried out, and the tests were run using a Pentium IV-3GHz processor with a 512MB memory. The algorithms shown below were all written using MATLAB. Table 1 shows that the computation cost increases drastically in the GVF snake due to finding the GVF vector field. The proposed snake did require more computation time than the greedy snake, due to its additional directional energy term and the updating process for the FD map. However, when comparing the snake evolution results in Fig. 5, Fig. 7, and Fig. 8, the proposed snake clearly exhibited a better boundary localization than the greedy and GVF snakes.

Table 1. Comparison of the average computation time of the greedy, GVF, and proposed snakes.

Image sequence	Proposed snake	Greedy snake	GVF snake
moving tank	7.4sec	5.0sec	92.2sec

5. CONCLUSION

A motion-based boundary tracking algorithm was presented for use with the greedy snake to stably track a rigid or non-rigid moving object in an image sequence. To cope with the shape changes of a moving object in tracking applications, a new directional energy term based on a FD map is introduced to the greedy snake and updating rules for the FD map are developed.

The proposed snake is not only faster and more efficient, but also able to inflate or deflate its contour automatically. Moreover, the initial contour does not need to be close to the true boundary, which causes the boundary tracking problem. The proposed snake has been compared with the greedy and GVF snakes using the real image sequences. The experimental results showed that the proposed snake produced a better performance for boundary localization and quick tracking of a deformable object in every frame.

6. REFERENCES

[1] L. Ji, H. Yan, "Attractable Snakes Based on the Greedy Algorithm for Contour Extraction," Pattern Recognition, vol. 35, pp. 791-806, 2002.

[2] C. Xu, J. L. Prince, "Snakes, Shapes, and Gradient Vector Flow," IEEE Trans. IP, vol. 7, no. 3, pp. 359-369, 1998.

[3] Z. Hou, C. Han, "Force Field Analysis Snake: an Improved Parametric Active Contour Model," Pattern Recognition Letters, vol. 26, pp. 513-526, 2005.

[4] F. Leymarie, M. D. Levine, "Tracking Deformable Objects in the Plane Using an Active Contour Model," IEEE Trans. PAMI, vol. 15, no. 6, pp. 617-634, 1993.

[5] M. Pardas, E. Sayrol, "Motion Estimation Based Tracking of Active Contours," Pattern Recognition Letters, vol. 22, pp. 1447-1456, 2001.

[6] M. Kass, A. Witkin, D. Terzopoulos, "Snakes : Active Contour Models," Int. J. Comput. Vis., vol. 1, no. 4, pp. 321-331, 1998.

[7] A. A. Amini, T. Weynouth, R. C. Jain, "Using Dynamic Programming for Solving Variational Problems in Vision," IEEE Trans. PAMI, vol. 12, no. 9, pp. 855-867, 1990.

[8] D. J. Williams, M. Shah, "A Fast Algorithm for Active Contours and Curvature Estimation," CVGIP: Image Understanding, vol. 55, no. 1, pp. 14-26, 1992.

[9] N. Peterfreund, "Robust Tracking of Position and Velocity with Kalman Snakes," IEEE Trans. PAMI, vol. 21, no. 6, pp. 564-569, 1999.

[10] P. J. Schneider, D. H. Eberly, Geometric Tools for Computer Graphics, Morgan Kaufmann, 2003.

[11] G. Xingfei, T. Jie, "An Automatic Active Contour Model for Multiple Objects," ICPR, vol. 2, pp. 881-884, 2002.