15th Triennial World Congress, Barcelona, Spain

VISUAL TRACKING FOR MOVING OBJECT BY INVARIANT ACTIVE CONTOUR MODEL

Won Kim and Ju-Jang Lee

Department of Electrical Engineering and Computer Science
Korea Advance Institute of Science and Technology
373-1 Kusong-dong, Yusong-gu, Taejon, Korea, 305-701
E-mail: kimwon@odyssey.kaist.ac.kr, jjlee@ee.kaist.ac.kr

Abstract: In case where strong a priori knowledge about the object being analyzed is available, it can be embedded into the formulation of the snake model. When prior knowledge of shape is available for a specific application, information concerning the shape of the desired objects can be incorporated into the formulation of the snake model as an active contour model. This active contour model can have more robust tracking performance for the weak edge cases because the snaxels of active contour model on weak edges can estimate the likelihood from a priori knowledge about target's shape and the relations with neighbor snaxels. In this paper we show Five points algorithm can be applied to design invariant energy and the effectiveness of the snake designed on this new concept in more robust tracking for moving object in visual space than the standard snake.

Keywords: Snake; Active Contour Models; Visual Tracking; Invariant.

1. INTRODUCTION

Deformable models have been used extensively in image analysis, especially in medical or biological imaging applications. Especially active contour models have been developed as useful tools for segmenting and tracking rigid or non-rigid deformable objects. Snake, one of the active contour models, was introduced by Kass in 1987(Kass *et al.* 1988). They defined snake energies such as internal energy, image energy and external energy. Segmentation and tracking can be done by energy minimization process about these energy. They tried to solve optimization problem for energy minimization by variational approach. Amini presented dynamic programming for finding minimum points(Amini *et al.* 1990). Leymarie tried to segment and track deformable objects like amebas and proved the convergence of snake's motion(Leymarie *et al.* 1993). Yang tried to track the contours of a human face for recognition of human's intentions and emotions by computer vision technique. Xu used snake to model three dimensional objects(Xu *et al.* 1998).

Visual tracking means successive segmentations of target object's boundaries through sequential image streams. Snake's segmentation process is a kind of process related with energy minimization. Snake energies have to be defined to have the minimal values on target's boundaries. If the minima of energy surface can be found successfully through energy minimization process, snake can find the object's boundary. This is the segmentation process. After the segmentation process, next new image is captured to vision systems. When the variances of the object's location and configuration between two sequential images are small, snake can also make segmentation for this new image from the results of the former stage through process of energy minimization. This is a tracking process of snake.

By using the results of snake's tracking we can get the corresponding points between two subsequent images. The corresponding points are related with the extrinsic parameters that are the translation vector, T, and the rotation matrix, R about camera motions. So visual navigation can be implemented by using the tracking results of deformable models operating about the feature shapes in sequential images.

Snaxels, that are snake's nodes, of the standard snake can move individually based on the values of snake energy. There may be weak edges around target objects generated by the changes of illumination conditions, reflectance of surfaces and the other reasons. The changes of illumination conditions may always exist in real environment. For the standard snake the snaxels on the segments of weak edges may have poor tracking performance because the values of image energy on those points may be not low. Therefore at that instance the snaxels on weak edges may find the other local edges that are not on the boundaries of target. Moreover the tracking failure at one instance is strongly linked with overall tracking performance in the tracking based on deformable models.

In case where strong a priori knowledge about the object being analyzed is available, it can be embedded into the formulation of the snake model(Ip *et al.* 1998). When prior knowledge of shape is available for a specific application, information concerning the shape of the desired objects can be incorporated into the formulation of the snake model as an active shape model(ASM). This active contour model can have more robust tracking performance for the weak edge cases because the snaxels of active contour model on weak edges can estimate the likelihood from a priori knowledge about target's shape and the relations with neighbor snaxels. This active shape model can be applied to recover the motion of the camera given a sequence of images.

2. DESIGN OF SNAKE ENERGIES

There are several energy terms in active contour models(Kass *et al.* 1988) such as internal energy, image energy and external energy. In this paper continuity energy, smoothness energy and image energy are selected as basic energy terms explained in (Trucco *et al.* 1998).

2.1 *Continuity Energy*

Continuity Energy, E_{cont}, prevents the formulation of clusters of snaxels. Denoting $\bar{d}$ as a mean distance between two adjacent snaxles, we can express E_{cont} as follows:

$$E_{cont} = \|p_i - p_{i-1}\|^2 \tag{1}$$

where $p_i, i = 1...N$ is *i-th* snaxel of a chain of N image points on the contour.

2.2 *Smoothness Energy*

Smoothness energy has the function to make contour smooth by penalizing high contour curvatures. Because the curvature is well approximated by the second derivative of the contour, E_{curv} can be defined as

$$E_{curv} = \|p_{i-1} - 2p_i + p_{i+1}\|^2. \tag{2}$$

2.3 *Image Energy*

Gradient operation for calculating image energy can be executed by the convolution of image with a Sobel operator (Jain 1989). The operation can be expressed as follows :

$$\begin{aligned} < U, H >_{m,n} &= u(m,n) \otimes h(-m,-n) \\ &= \sum_i \sum_j h(i,j) u(i+m, j+n) \end{aligned} \tag{3}$$

where U is an image buffer, H is a $p \times p$ Sobel mask, and (m, n) is an interested location coordinate.

A Sobel mask is composed of two orthogonal operators. The equation (4) is a mask for the gradient along the x-axis, and the equation (5) is a mask for the gradient along the y-axis. The magnitude of the gradient can be obtained by applying convolving operations with these masks as follows:

$$h_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{4}$$

and

$$h_2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \tag{5}$$

3. DESIGN OF INVARIANT ENERGY

3.1 *Motivation*

There may be two kinds of problems in snake's tracking. The existence of weak edge is the first one. Some portions of target boundaries may not be strong edges in the stream of sequential images. So snake energies in weak edges may have lower values and this may guide snake to tracking failures because a failure in one image frame means total tracking failure in snake's tracking. Second one is the confusion in the case of multiple objects. There may be the other objects in images. Snake may be confused to track which object and therefore may track the other object near the target if the edges of the other's are stronger than the target's. These two problems are explained in figure 1.

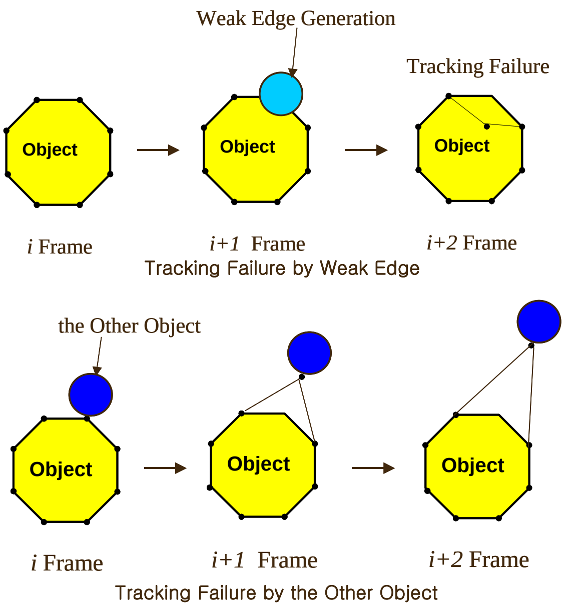

Fig. 1. The concept of two problems in Snake's tracking.

3.2 *Invariant Energy*

To overcome two tracking problems discussed earlier we introduce invariant energy term, E_{inv} which give the contour information of target to snake. So in the existence of weak edge, the other object near target and noises snake can have somewhat robust characteristics. Invariant energy, E_{inv} is defined in equation (6) as follows :

$$E_{inv}(v(s)) = \omega(s) \cdot exp(L/\sigma) \qquad (6)$$

where ω and σ are tuning parameters and L means the variation ratio of invariant values. The variation ratio of invariant values, L is defined as

$$L = (|I(k) - I'(k)| / I'(k)) * 100 \qquad (7)$$

where I(k) means invariant value of current image at node k and I'(k) means invariant value of the previous image at the same node k.

Invariant value I(k) is calculated by Five point algorithm is defined as

$$I = det(X_5X_1X_4) \cdot det(X_5X_2X_3) \qquad (8)$$
$$/(det(X_5X_1X_3) \cdot det(X_5X_2X_4))$$

where X_i is a pixel coordinate expressed as $X_i = (x_i, y_i, 1)^T$. Above equation (9) was suggested by Roh and Kweon (Roh *et al.* 1998). By using this equation we can incorporate object's contour information to snake. The concept of calculation of this invariant values on the contour is explained in figure 2.

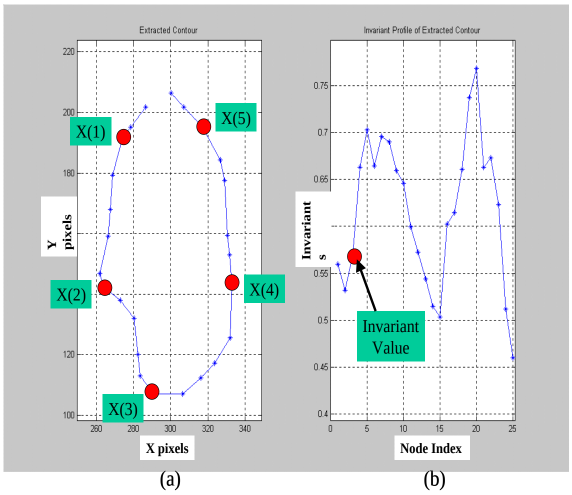

Fig. 2. The concept of calculation of invariant values on the contour : (a) A contour (b) The invariant profile for this contour.

The parameter ω should be expressed as equation (9) because the differences are big in five points simultaneously due to the characteristics of the Five points algorithm for one node which miss tracking the boundary of target as follows:

$$\omega(i) = \begin{cases} 0 & \text{if } D(i) \geq D_{TH} \\ \omega(i) & \text{otherwise} \end{cases} \qquad (9)$$

where D_{th} is a threshold value and D means a measure of the position difference of snake's node between two frames expressed as follows:

$$D(i) = \|X_i - X_i'\|_2 \,. \qquad (10)$$

This condition is needed to only consider the large snake's node variation case. Therefore invariant energy values increase exponentially along the variation of invariant values between two successive images.

4. SIMULATIONS

4.1 *Simulation Results for Synthetic Images*

Figure 3 is a simulation result of the basic snake for a case of the existence for two objects. There is the other object near target and one node fails tracking because the edge of the other object is stronger the target's one.

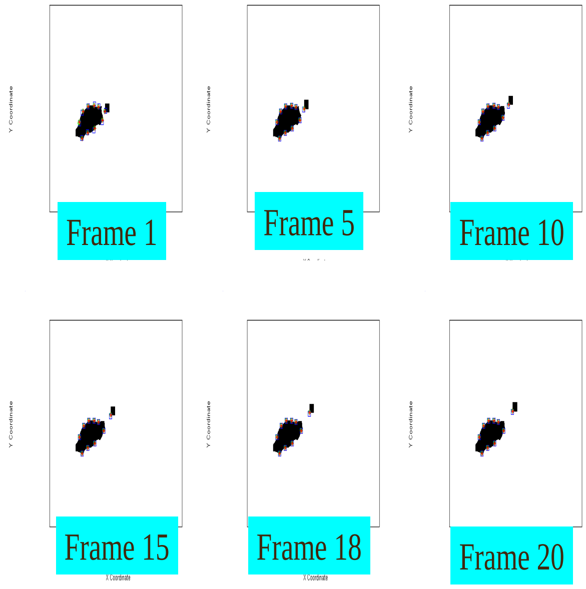

Fig. 3. The simulation result of the basic snake.

Figure 4 is a simulation result of the proposed snake for the same case. In this simulation one node tracks the other object in initial frames. When the value of D is large than D_{TH} at the 15-th frame the snake's invariant energy, E_{inv} increases abruptly and the split node can escape from the local minimum point. So robust tracking can be done by this scheme.

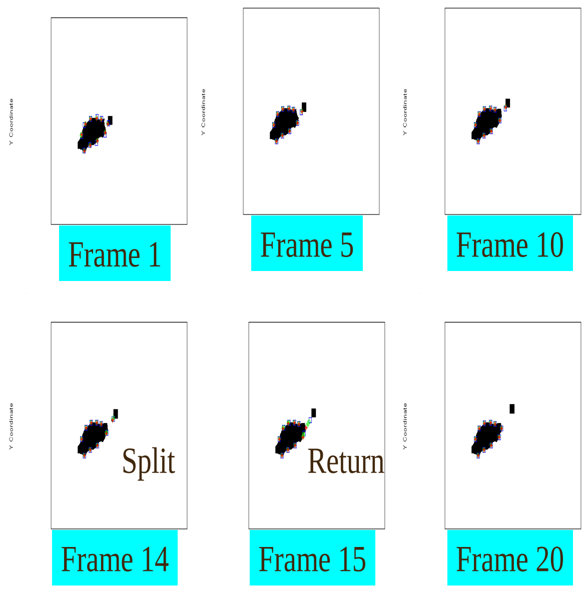

Fig. 4. The simulation result of the proposed snake.

4.2 *Simulation Results for Tracking Face-like Objects*

There are image sequences in which an face-like object is moving to the left direction in image space in figure 5.

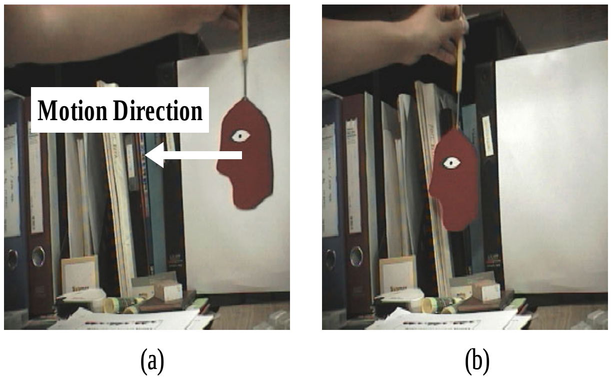

Fig. 5. Tracking of face-like object : (a) Initial position. (b) Final position.

In figure 6 the segmentation result is explained. Initially snake should be placed around target object. So in this case roughly placed snaxels contract into the face-like object and finally find the exact boundaries of the target in figure 6 (a). In figure 6 (b) the extracted snaxels on that contour are expressed in image space.

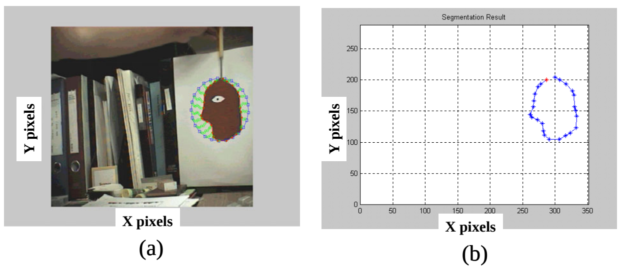

Fig. 6. Segmentation result : (a) Segmentation mode. (b) Extracted contour.

There is a comparison between the positions of the model contour and the affine transformed one to the 12-th image in figure 7 (a) in which the snaxels in a model and the affine transformed contour are labelled as squares and stars, respectively. For the comparison of invariant values between two contours an affine transform matrix between them should be found by an optimization method explained in (Ip *et al.* 1998). In figure 7 (b) the two contours in the 12-th image and the affine transformed model are expressed labelled as squares and stars, respectively. The comparison in invariant values for these two contours is explained in figure 8.

5. CONCLUSIONS

In cases where strong a priori knowledge about the object being analyzed is available, it can be em-

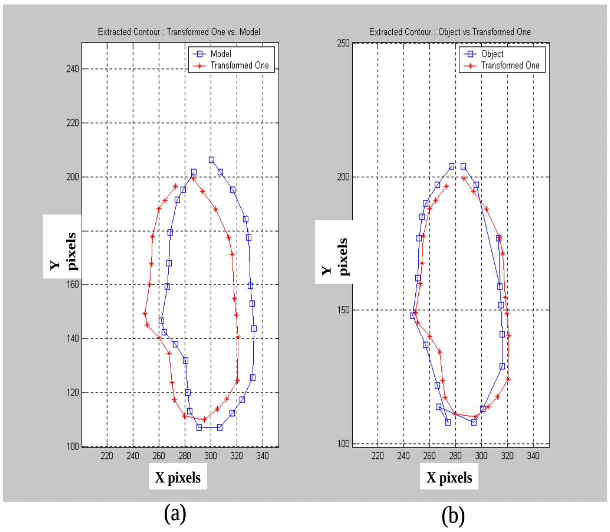

Fig. 7. Affine Transform : (a) Two contours of model and transformed contour to the 12-th one. (b)Affine transformed contour of model and the 12-th one.

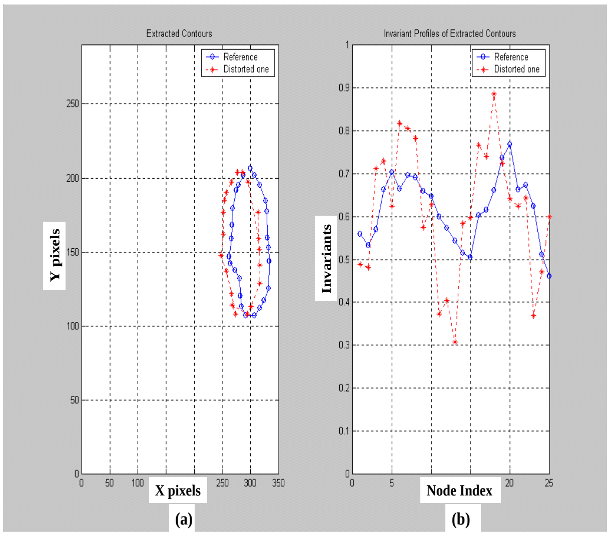

Fig. 8. Affine Transform : (a) Two contours of model and the 12-th image. (b) Comparison of invariant values for these two contours.

bedded into the formulation of the snake model. When prior knowledge of shape is available for a specific application, information concerning the shape of the desired objects can be incorporated into the formulation of the snake model as an active contour model. In this paper we show Five points algorithm can be applied to design invariant energy. This active contour model can have more robust tracking performance for the weak edge cases because the snaxels of active shape model on weak edges can estimate the likelihood from a priori knowledge about target's shape and the relations with neighbor snaxels. This active shape model can be applied to recover the motion of the camera given a sequence of images.

6. REFERENCES

Kass, M., A. Witkin, and D. Terzopoulos (1988). Snake : Active contour models. *Int. J. Computer Vision* , pp. 321-331.

Amini, A.A., T.E. Weymouth, and R.C. Jain (1990). Using dynamic programming for solving variational problems in vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **Vol.12**, pp. 855-867.

Yi, Y.J. (1993). A study on the development of the tracking algorithm using active contour model. Master's thesis, KAIST in Korea.

Leymarie, F. and M.D. Levine (1993). Tracking deformable objects in the plane using an active contour model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **Vol.15**, No. 6, pp.617-634.

Cho, Y.J. (1996). A study on the face contour tracking system using snakes model. Master's thesis, KAIST in Korea.

Xu, C. and J.L. Prince (1998). Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Processing*, **Vol.7**, No. 3, pp.359-369.

Jain, A.K. (1989). *Fundamentals of digital image processing*, NJ:Prentice-Hall.

Jain, R., R. Kasturi, and B.G. Schunck (1995). *Machine vision.* McGraw-Hill.

Giachetti, A., M. Campani, and V. Torre (1998). The use of optical flow for road navigation. *IEEE Trans. on Robotics and Automation*, **Vol.14**, No. 1, pp.34-48.

Blake, A. and A. Yuille (1992). *Active vision.* MIT-Press.

Trucco, E. and A. Verri (1998). *Introductory techniques for 3-D computer vision.* Prentice Hall.

Ip, H. and D. Shen (1998). An affine-invariant active contour model(AI-snake) for model-based segmentation. *Image and vion computing*, **Vol.16**, pp.135-146.

Arbter, K., W.E. Snyder, H. Burkhardt and G. Hirzinger (1990). Application of Affine-Invariant Fourier Descriptor to Recongnition of 3-D Object. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **Vol.12**, No. 7, pp.640-647.

Roh, K.S. and I.S. Kweon (1998). 2-D Object Recognition using Invariant Contour Descriptor and Projective Refinement. *Pattern Recognition*, **Vol.31**, No. 4.