Tracking Moving Contours Using Energy-Minimizing Elastic Contour Models

Naonori UEDA[1] *and Kenji MASE*[2]

[1] NTT Communication Science Laboratories, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan
[2] NTT Human Interface Laboratories, Kanagawa, 238-03, Japan

Abstract. This paper proposes a method for tracking an arbitrary object contour in a sequence of images. In the contour tracking, energy-minimizing elastic contour models are utilized, which is newly presented in this paper. The proposed method makes it possible to establish object tracking even when complex texture and occluding edges exist in or near the target object. We also newly present an algorithm which efficiently solves energy minimization problems within dynamic programming framework. The algorithm enables us to obtain optimal solution even when the variables to be optimized are not ordered.

1 Introduction

Detecting and tracking moving objects is one of the most fundamental and important problems in motion analysis. When the actual shapes of moving objects are important, higher level features like object contours, instead of points, should be used for the tracking. Furthermore, since these higher level features make it possible to reduce ambiguity in feature correspondences, the correspondence problem is simplified.

However, in general, the higher the level of the features, the more difficult the extraction of the features becomes. This results in a tradeoff, which is essentially insolvable as long as a two-stage processing is employed. Therefore, in order to establish high level tracking, object models which embody a *priori* knowledge about the object shapes are utilized[1][2].

On the other hand, Kass *et al.*[3] have recently proposed *active contour models*(Snakes) for the contour extraction. Once the snake is interactively initialized on an object contour in the first frame, it will automatically track the contour from frame to frame. That is, contour tracking by snakes can be achieved. It is a very elegant and attractive approach because it makes it possible to simultaneously solve both the extraction and tracking problems. That is, the above tradeoff is completely eliminated.

However, this approach is restricted to the case that the movement and deformation of an object are very small between frames. As also pointed out in Ref.[2], this is mainly due to the excessive flexibility of the spline composing the snake model.

In this paper, we propose a robust contour tracking method which can solve the above problem while preserving the advantages of snakes. In the proposed method, since the contour model itself is defined by elastics with moderate "stiffness" which does not permit local major deformations, the influence of texture and occluding edges in or near the target contour is minimal. Hence, the proposed method becomes more robust than the original snake models in that it is applicable to more general tracking problems.

In this paper, we also present a new algorithm for solving energy minimization problems using dynamic programming technique. Amini *et al.*[4] have already proposed a

Lecture Notes in Computer Science, Vol. 588
G. Sandini (Ed.)
Computer Vision – ECCV '92

dynamic programming(DP) algorithm which is superior to *variational approach* with regard to optimality and numerical stability. In order to use DP, however, the original decision process should be *Markovian*. From this point of view, with Amini's formulation, optimality of the solution is ensured only in the case of open contours. That is, for closed contours, reformulation is necessary. In this paper, we clarify the problem of Amini's formulation, and furthermore, within the same DP framework, we present a new formulation which guarantees global optimality even for closed contours.

2 Formulating the contour tracking problem

2.1 Elastic contour models

A model contour is defined as a polygon with n discrete vertices. That is, the polygonally approximated contour model is represented by an ordered list of its vertices: $\mathbf{C} = \{v_i = (x_i, y_i)\}, 1 \le i \le n$. A contour model is constrained by two kinds of "springs" so that it has a moderate "stiffness" which preserves the form of the tracked object contour in the previous frame as much as possible. That is, each side of the polygon is composed of a spring with a restoring force proportional to its expansion and contraction, while the adjacent sides are constrained by another spring with a restoring force proportional to the change of the interior angle. Assume that these springs are *original length* when the contour model is at the initial contour position $\{v_i^0\}_{i=1}^n$ in the current frame. Therefore, at that time, for the springs no force is at work. Clearly, the initial position in the current frame corresponds to the tracking result in the previous frame.

2.2 Energy minimization framework

Let $\{v_i^0\}_{i=1}^n$ denote a tracked contour in the preceding frame. Then our goal is to move and deform the contour model from $\{v_i^0\}_{i=1}^n$ to the best position $\{v_i^*\}_{i=1}^n$ in the current frame such that the following total energy functional is minimized:

$$E_{total}(v) = \sum_{i=1}^{n} \Big(E_{elastic}(v_i) + E_{field}(v_i) \Big). \tag{1}$$

Here, $E_{elastic}$ is elastic energy functional derived from the deformation of the contour model and can be defined as:

$$\begin{aligned} E_{elastic}(v_i) = \frac{1}{2}\Big(&\mu_1(|v_{i+1} - v_i| - |v_{i+1}^0 - v_i^0|)^2 \\ &+\mu_2(ang(v_i, v_{i+1}, v_{i+2}) - ang(v_i^0, v_{i+1}^0, v_{i+2}^0))^2\Big) \end{aligned} \tag{2}$$

where $ang(v_i, v_{i+1}, v_{i+2})$ means the angle made by sides $v_{i+1}v_i$ and $v_{i+1}v_{i+2}$. (see Fig.1b.) μ_1 and μ_2 are non-negative constants. In Eq.(2), the first energy term corresponds to the deformation due to the expansion and contraction of each side of the polygonally approximated contour model, while the second energy term corresponds to the deformation due to the change of interior angle between the two adjacent sides.

E_{field} is the potential energy functional which gives rise to edge potential field forces newly defined in this paper. The potential field is derived from the edges in the current frame including target contour. The potential field used here, since it is obtained with distance transformation[5], unlike that used in the original snakes, smoothly extends over a long distance. Therefore, it can influence the contour model even if the contour model is remote from the target contour.

Assuming that $z(\boldsymbol{v}_i)$ denotes the height or potential value at $\boldsymbol{v}_i$ on the potential field, then the potential energy, E_{field}, can easily be defined by the classical gravitational potential energy equation. That is,

$$E_{field}(\boldsymbol{v}_i) = m\mathcal{G}z(\boldsymbol{v}_i) = \mu_3 z(\boldsymbol{v}_i). \tag{3}$$

where m is the constant mass of the $\boldsymbol{v}_i$, $\mathcal{G}$ is the magnitude of the gravitational acceleration. μ_3 is a negative constant.

It can be intuitively interpreted that Eq.(1) becomes minimum when the contour model is localized on the contour whose shape most nearly resemples the contour tracked in the previous frame. Accordingly, even if the contour model is not remote from the target contour, the model can move to the target contour while preserving its shape as much as possible. As a result, a tracking desired contour can be achieved.

3 Optimization algorithm

From Eqs.(2) and (3), the total energy functional shown in Eq.(1) can be formally brought to the general form:

$$E_{total}(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n) = \sum_{i=1}^{n}\{f_i(\boldsymbol{v}_i) + g_i(\boldsymbol{v}_i, \boldsymbol{v}_{i+1}) + h_i(\boldsymbol{v}_i, \boldsymbol{v}_{i+1}, \boldsymbol{v}_{i+2})\}, \tag{4}$$

Note that the general form of Eq.(4) is the same as that of snakes.

The minimization of Eq.(4), like snakes, returns us to the problem of finding the optimum values $\{\boldsymbol{v}_i^*\}_{i=1}^n$ which give the local minimum, starting from the initial values $\{\boldsymbol{v}_i\}_{i=1}^n$. One way to find the minimum is by employing *exhaustive enumeration*. However, with this approach, combinatorial explosion is inevitable. Therefore, we must devise a more efficient algorithm.

Recently, Amini *et al.*[4] proposed a dynamic programming approach to energy minimization of the snakes. In the dynamic programming approach, the minimization of Eq.(4) is viewed as a discrete multistage decision process, with $\boldsymbol{v}_i$ corresponding to the *state variable* in the i-th decision stage. However, this DP formulation is for open contours which preserve the *ordering* of the variables $\{\boldsymbol{v}_i\}_{i=1}^n$. In other words, $\boldsymbol{v}_1$ and $\boldsymbol{v}_n$ are not connected and constrained. Consequently, reformulation of DP equation for the closed contours is necessary.

Let V be a set of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$. Being focused on $\boldsymbol{v}_1$ in Eq.(4), $\boldsymbol{v}_1$ is included by $f_1(\boldsymbol{v}_1), g_1(\boldsymbol{v}_1, \boldsymbol{v}_2), h_1(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3), h_{n-1}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_1), g_n(\boldsymbol{v}_n, \boldsymbol{v}_1)$, and $h_n(\boldsymbol{v}_n, \boldsymbol{v}_1, \boldsymbol{v}_2)$. Thus, for convenience, we here use S for the sum of these functions. That is,

$$S = f_1(\boldsymbol{v}_1) + g_1(\boldsymbol{v}_1, \boldsymbol{v}_2) + h_1(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) + h_{n-1}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_1) + g_n(\boldsymbol{v}_n, \boldsymbol{v}_1) + h_n(\boldsymbol{v}_n, \boldsymbol{v}_1, \boldsymbol{v}_2). \tag{5}$$

Then, the minimization of E_{total} can be written as:

$$\begin{aligned} \min_{V} E_{total} &= \min_{V-\{\boldsymbol{v}_1\}} \min_{\boldsymbol{v}_1} E_{total} \\ &= \min_{V-\{\boldsymbol{v}_1\}} \Big\{(E_{total} - S) + \min_{\boldsymbol{v}_1}\{S\}\Big\}. \end{aligned} \tag{6}$$

Hence, the first step of the optimization procedure is to perform the minimization with respect to $\boldsymbol{v}_1$ in Eq.(6). Clearly, from Eq.(5), one can see that the minimization is a function of v_2, v_3, v_{n-1}, and v_n. Therefore, this minimization is made and stored for all possible assignments of v_2, v_3, v_{n-1}, and v_n. Formally, the minimization can be written as:

$$\psi_1(v_2, v_3, v_{n-1}, v_n) = \min_{v_1}\{S\}. \tag{7}$$

Note that in the minimization in Eq.(7), exhaustive enumeration is employed.

Then, the problem remaining after the minimization with respect to v_1,

$$\min_{V} E_{total} = \min_{V-\{v_1\}}\Big\{\big(E_{total} - S\big) + \psi_1(v_2, v_3, v_{n-1}, v_n)\Big\}, \tag{8}$$

is of the same *form* as the original problem, and the function $\psi_1(v_2, v_3, v_{n-1}, v_n)$ can be regarded as a component of the new objective function.

Applying the same minimization procedure for the rest of the variables, $v_2, v_3, \ldots$ in this order, we can derive the following DP equations. That is, for $2 \le i \le n-4$,

$$\psi_i(v_{i+1}, v_{i+2}, v_{n-1}, v_n) = \min_{v_i}\Big\{\psi_{i-1}(v_i, v_{i+1}, v_{n-1}, v_n) + f_i(v_i) + g_i(v_i, v_{i+1}) + h_i(v_i, v_{i+1}, v_{i+2})\Big\}, \tag{9}$$

where, for $i = n-3, n-2, n-1$, the corresponding DP equations can be obtained respectively.

The time complexity of the proposed DP algorithm then becomes $\mathcal{O}(nm^5)$ because, in Eq.(9), the optimum decision is done over m^4 combinations. However, since, in general, each optimum decision stage in DP can be independently achieved, computation time can be drastically reduced with parallel processing.

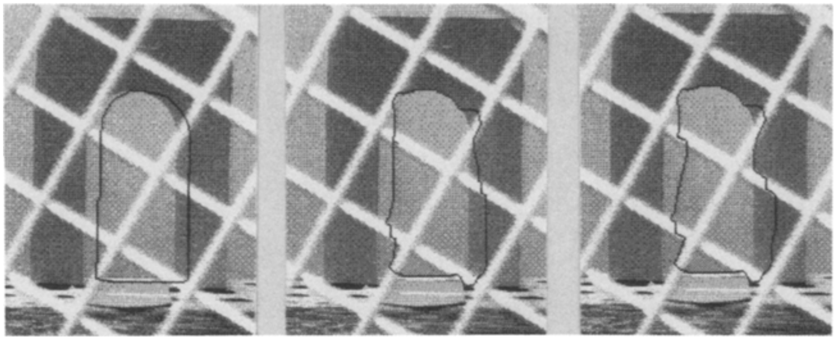
4 Experiments

The proposed contour tracking method has been tested experimentally on several synthetic and real scenes. Figure 1 compares the snake model(Fig.1a) with our model(Fig.1b) when occluding edges exist. The scene in Fig.1 is an actual indoor scene and corresponds to one frame from a sequence of a moving bookend on a turntable over a static grid. Since the snake model is influenced by occluding edges, the model was not able to track the target contour. On the other hand, the proposed model successfully tracked it without being influenced by the occluding edges. We also obtained successful results for the trackings of moving car, deforming ball, and so on.

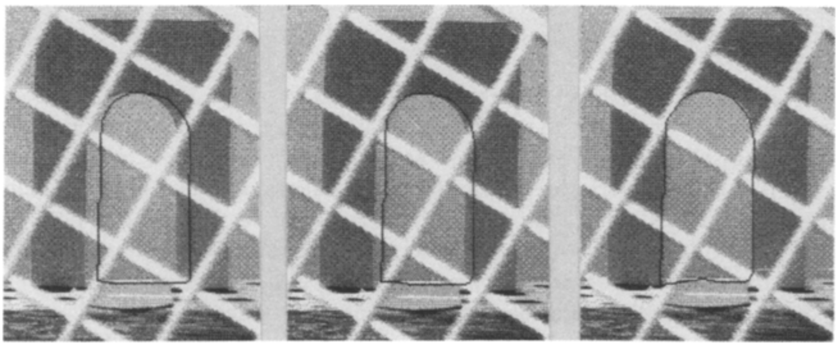
In this approach, since the contour model itself moves toward the target contour, point correspondences are established between frames. That is, correspondence based optical flows are also obtained. Therefore, feature point trajectories over several frames can easily be obtained by the proposed method.

5 Conclusions

We have presented here an energy-minimizing elastic contour model as a new approach to moving contour tracking in a sequence of images. Compared to the original snake model, the proposed method is more robust and general because it is applicable even when movements and deformation of the object between frames are large and there exist occluding edges. Moreover, we have newly devised an optimization algorithm with a dynamic programming framework, which is efficient and mathematically complete.



$I = 0$ $I = 4$ $I = 12$(Result)
(a)Tracking by the snake model



$I = 0$ $I = 2$ $I = 7$(Result)
(b)Tracking by the proposed model

Fig.1. Comparison of the results of tracking contour with occluding edges. I denotes the number of iterations.

References

1. Dreschler and Nagel H. H.: "Volumetric model and 3D trajectory of a moving car derived from monocular tv-frame sequence of a street scene", in *Proc.* **IJCAI81**, 1981.
2. Yuille A. L., Cohen D. S. and Hallinan P. W.: "Feature extraction from faces using deformable templates", in *Proc.*, **CVPR89**, pp. 104–109, 1989.
3. Kass A., Witkin A. and Terzopoulos D.: "Snakes: Active contour modes", *Int. J. Comput. Vision*, 1, 3, pp. 321-331, 1988.
4. Amini A. A., Weymouth T. E. and Jain R. C.: "Using dynamic programming for solving variational problems in vision", *IEEE Trans. Pattern Anal. Machine Intell.*, **PAMI-12**, 9, pp. 855–867, 1990.
5. Rosenfeld A. and Pfaltz J. L.:"Distance functions on digital pictures", *Pattern Recognition*, 1, pp.33-61, 1968.

This article was processed using the LaTeX macro package with ECCV92 style