OPTIMAL RADIAL CONTOUR TRACKING BY DYNAMIC PROGRAMMING

Yunqiang Chen and Thomas S. Huang

chenyq@ifp.uiuc.edu, huang@ifp.uiuc.edu
405 N. Mathews Ave., Beckman Institute
University of Illinois at Urbana-Champaign
Urbana, IL, USA 61801

Yong Rui

yongrui@microsoft.com
Microsoft Research
One Microsoft Way
Redmond, WA 98052

ABSTRACT

A common problem in most active contour methods is that the recursive searching scheme can only return a local optimal solution. Furthermore, the internal energy of the snake is not strong enough to control the shape of the contour. To overcome these difficulties, in this paper, we develop a causal internal energy term based on a radial contour representation to encode the smooth constraint of the contour, and develop a global shape priori to control contour's shape and position based on object's dynamics. The causality nature of the representation allows us to efficiently find global optimal solution using dynamic programming. To validate the efficacy and robustness of the proposed approach, we apply this approach to track people in bad illumination and cluttered environments. We report promising results in the paper.

1. INTRODUCTION

Visual tracking has become more and more important. Real-time applications such as video surveillance, video conferencing and human-computer interface in virtual environment all require the ability to track moving objects. A robust visual tracking algorithm in complex environments is a very challenging task. For example, in the virtual environment of CAVE [1], illumination and background change dramatically between frames, making color or motion based visual tracking almost impossible.

Contour-based tracking methods have been extensively studied in computer vision community during the past decade [2, 3, 4, 5, 6]. For tracking non-rigid objects in moving background and bad illumination environments, it is a promising approach. Active contour model, e.g., snake, has been proved to be a powerful tool for semi-automatic boundary delineation in still and moving images. Amini et. al. developed an iterative algorithm where at each step dynamic programming was applied to incrementally improve the current contour result[7]. However, the contour evolution can easily get stuck at local minimums and is sensitive to the initial condition. They also used internal energy terms to ensure smooth contours, but such constraint was not strong enough to ensure global shape constraints.

To track the contours in clutter, Blake and Isard [8] developed the Condensation algorithm. This sampling-based algorithm explores the prior knowledge of shape and motion by using a stochastic framework. The MAP result is achieved by propagating the conditional probability densities over time. However, to represent the density function, the required samples grow exponentially with the dimension of the state space. It also required accurate models for both shape and motion dynamics.

The most related work to our approach is the active rays method proposed by Denzler and Niemann [9]. They introduced radial representation of contours that overcomes some of the problems in snake models. This representation introduced the ordering of the contour points and prevented the snake elements from crossing each other during evolution. However, they still used the traditional iterative optimization technique rather than taking full advantage of the new representation. Furthermore, there was no prior knowledge on global shape or position dynamics integrated into the framework.

In this paper, we propose a new contour-tracking algorithm that is based on the radial representation and incorporates both shape/position priori and the local smooth constraints of the contours. The proposed approach finds global optimal solutions efficiently by dynamic programming.

The rest of the paper is organized as follows. In Section 2, we define the radial representation of contours and the new internal energy. In Section 3, we give detailed description of our proposed approach. We test our algorithm with real sequences and report promising results in Section 4. Concluding remarks and future works are in Section 5.

2. REPRESENTATION OF CONTOURS

Our objective is to design a new contour model to track people in complex environment. In this section, we will

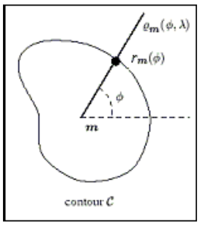

Fig. 1. Representation of a contour by active rays

describe the radial representation of object contours and re-express the traditional smooth constraint of contours in a causal way such that we can find global optimal solutions in one dynamic programming iteration.

2.1. Radial Representations

Let the image coordinate be indexed by (x, y). Let $m = (x_m, y_m)^T$ be the center of a contour. Let ϕ be the angular direction of a ray coming out from the center. Let λ be the distance from any point in the ray to the center point. An active ray $\rho_m(\phi, \lambda)$ is defined as:

$$\rho_m(\phi, \lambda) = f(x_m + \lambda \cos\phi, y_m + \lambda \sin\phi), 0 \le \lambda \le N \tag{1}$$

where N is the quantized length of the rays.

This is illustrated in Figure 1. In this radial contour representation, we assume the object contour will intersect with each ray only once. Then the contour can be represented by the center m of the object and the 1D function $r_m(\phi)$. Note that this assumption is much relaxed than the non-concave assumption. Furthermore, this assumption is valid in most human tracking situations. For example, the profile of a person's head and the top view of a human shoulder both satisfy this assumption. This new radial representation introduces ordering to all the contour points. For visual tracking, our task is to find the contour $r_m(\phi)$ that best fits the constraints and the image content, given the center point m. We could also use line segments normal to the contour as our visual feature (e.g., in [8]) if the contour is irregular. This, however, will not affect our proposed tracking scheme.

2.2. Energy Terms' Definition

Like traditional contour models, we need to impose smooth constraints on the contours. This is achieved by defining internal energy term to set penalty for rough contour points. In the traditional snake model, the roughness is defined by the first and second derivatives of the contour. This definition results in the non-causality of the snake model, because the first and second derivatives of contour depend on both the pixels before and after the current pixel on the contour. Therefore, the optimization has to be solved iteratively. To avoid this, we assume the radius on two adjacent rays will not change dramatically for smooth contours. So we can re-express the smooth constraint in a causal way:

$$E_i'(r_m(\phi_i)) = \alpha_i \cdot |r_m(\phi_i) - r_m(\phi_{i-1})|^2 \tag{2}$$

Similarly, we can also define the higher order smoothness terms:

$$\begin{aligned} E_i''(r_m(\phi_i)) = \alpha_i'' \cdot |(r_m(\phi_i) - r_m(\phi_{i-1})) \\ -(r_m(\phi_{i-1}) - r_m(\phi_{i-2})|^2 \end{aligned} \tag{3}$$

These new definitions are not the same as the first and second derivative of the contour. But the difference is small when the contour is smooth. In this paper, for computation efficiency, we only use the first order smooth constraint.

We next define the external energy to set the influence of the image on the contour. It is defined as a function of the image gradient along the direction of the ray:

$$\begin{aligned} E_e(r_m(\phi)) &= \alpha_e \cdot g\left(-\left|\frac{d}{d\lambda}\rho_m(\phi,\lambda)\right|^2\right) \\ &= \alpha_e \cdot g\left(-(\rho_m(\phi, \lambda+1) - \rho_m(\phi,\lambda))^2\right) \end{aligned} \tag{4}$$

where $g(.)$ is a non-linear monotonically increasing function.

The total energy of the contour is therefore given by:

$$E(r_m(\phi)) = \int_0^{2\pi} \left(E_i(r_m(\phi)) + E_e(r_m(\phi))\right) d\phi \tag{5}$$

The best contour $r_m(\phi)$ is the one that gives the global minimum of the total energy. Because the introduction of ordering to the contour points and the new causal definition of the smooth constraint, it is possible to find the global optimal solution efficiently by dynamic programming.

2.3. Energy Minimization

To find the best-fit contour, the algorithm starts from $\phi_0 = 0$ with $E^o(\phi_0, \lambda_k) = E_e(\lambda_k)$ and propagates the energy to $\phi_M = 2\pi$, where M is the number of active rays quantized from 0 to 2π. For every pixel along a ray ϕ_{k+1}, we can find the optimal energy for a contour ending on it by the following propagating equation:

$$\begin{aligned} E^o(\phi_{k+1}, \lambda_{k+1}) = \min_{\lambda_k \in [0,N]} \{E^o(\phi_k, \lambda_k) + E_i(\lambda_k, \lambda_{k+1})\} \\ + E_e(\lambda_{k+1}) \end{aligned} \tag{6}$$

where $E_i(\lambda_k, \lambda_{k+1}) = \alpha_i \cdot |\lambda_{k+1} - \lambda_k|^2$ is the smooth constraint.

We can find the best contour by finding the minimum energy point along the last ray $min_\lambda E^o(2\pi, \lambda)$ and then back-trace to find the whole contour that gives this optimal energy.

This method works only for open contours. For closed contours, we have an extra constraint that the beginning point and the ending point should be the same. We can still apply dynamic programming in this situation. First we fix the start point $\lambda(\phi_0)$ and find the best contour that ends at the same point. After trying all the possible start points, we choose the one giving the optimal energy. This is the global optimal solution, but it costs N times the computation over the open contours.

Unlike traditional snake model, the proposed method gives us the global optimal solution in one dynamic programming iteration.

3. TRACKING WITH GLOBAL SHAPE PRIORI

The previous section has described our modified active contour algorithm. However, with only the local smoothness constraint, it is still difficult to produce good results in cluttered environments. For example, there may exist smooth and strong edges on the background and that will severely distract the contours.

In most applications, we have prior knowledge about the shape of the foreground objects. It is common to represent the contours of human faces or hands as ellipses. Furthermore, we know the speed and position of the objects do not change dramatically between frames. In the traditional snake model, the tracking results in previous frame will only be used as initial condition for recursive search. There is no guaranty that the object's shape and position will be properly maintained.

In our new scheme, it is easy to incorporate such information into an energy term. First, we assume the object contour can be represented by some global parameters, such as ellipse. Knowing the object's shape and position in the previous frames, we can predict its new shape and position in current frame. If the prediction is accurate, the deviation of current contour from the prediction will be independent a zero-mean Gaussian as shown in following equation:

$$r_{\hat{m}}(\phi) = \hat{r}_{\hat{m}}(\phi) + N(0, \sigma_r) \tag{7}$$

where $\hat{r}_{\hat{m}}(\phi)$ is the predicted position, $r_{\hat{m}}(\phi)$ is the true position of contour points on the rays, and $\hat{m}$ is the predicted object center.

Now we can incorporate the global shape/position priors into an extra energy term:

$$E_t(r_{\hat{m}}(\phi) = \alpha_t \cdot |r_{\hat{m}}(\phi) - \hat{r}_{\hat{m}}(\phi)|^2 \tag{8}$$

Because of the parameterization of the contour, we can represent the contour with small number of parameters and it is much easier to incorporate the dynamic properties of the objects to predict where the object will be in next frame. The tracking procedure is summarized as follows:

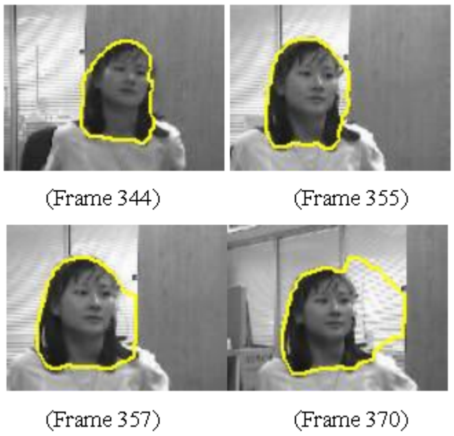

Fig. 2. Tracking with traditional contour model. The contour tracking results are severely distracted by the sharp edges on the background.

1. Predict where the object will be in current frame by object's dynamics and the results in previous frames.

2. Given the predicted contour of the object, the best contour in current frame can be found by solving the optimization of the total energy:

$$\begin{aligned} E(r_{\hat{m}}(\phi)) = \int_0^{2\pi} & (E_i(r_{\hat{m}}(\phi)) + E_e(r_{\hat{m}}(\phi)) \\ & + E_t(r_{\hat{m}}(\phi)))d\phi \end{aligned} \tag{9}$$

3. Finally, we fit an ellipse to the contour points and estimate the velocity of the object (e.g., translation and rotation). Go to step 1 for next frame.

To begin this tracking procedure, a separate initialization module is needed. This can be done either manually or by change-detection [11].

4. EXPERIMENTS

To validate the efficacy and robustness of the proposed algorithm, we test our algorithm in a cluttered office sequence [12]. There are 499 frames in this sequence, with 30 frames per second. We use gray value only. Note that the blinds and the door (sharp edges and cluster) impose a great challenge to the visual tracking algorithms. Also note that the sequence was captured by a pan/tilt/zoom camera that moves all the time. For the traditional contour tracking without global shape priori, the contour can easily stuck in the background. The tracking results of four frames are shown in Figure 2.

When the person moved close to the door, the sharp edge along the door greatly distracted the tracking. The edge is

Fig. 3. Tracking with our new contour model. The tracking results quickly recovered when the user moves away from the door.

smooth and strong – local smoothness constraints alone are not sufficient to avoid the distraction.

In our new tracking algorithm, we use 60 rays and approximate human heads with ellipses. Although the contour in frame 355 is distracted a little bit, it is quickly recovered in frame 357 because of the shape and dynamics we used in our algorithm.

5. CONCLUSION

In this paper, we introduce a new representation of contour model. The new causal definition of smoothness constraint enables us to find the global optimal solution with dynamic programming instead of iterative search methods.

With the new contour representation, we can also easily integrate the global shape/position priori with an extra energy term. Experimental results on a cluttered office environment demonstrated the robustness of our tracking algorithm.

When background noise is extremely high, this contour tracking may still fail. This is mainly because we assume the tracking result in the previous frame is correct, so the error may accumulate. We are working on integrating this model with probabilistic multiple hypothesis tracking (MHT) methods and with region-based tracking methods to further improve the robustness.

6. ACKNOWLEDGEMENTS

This work was supported in part by National Science Foundation Grant CDA 96-24396, EIA 99-75019, IIS-00-85980 and in part by a Gift Grant from Microsoft Research. Thanks Stan Birchfield for allowing us to use the test sequence.

7. REFERENCES

[1] "http://evlweb.eecs.uic.edu/pape/cave," .

[2] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[3] D. Terzopoulos and R. Szeliski, "Tracking with kalman snakes," in *Active Vision*, A. Blake and A. Yuille, Eds., pp. 3–20. MIT Press, 1992.

[4] N. Peterfreund, "Robust tracking of position and velocity with kalman snakes," *IEEE Trans. on Pattern Analysis & Machine Intell.*, vol. 21, no. 6, pp. 564–569, June 1999.

[5] C. Chesnaud, P. Refregier, and V. Boulet, "Statistical region snake-based segmentation adapted to different physical noise models," *IEEE Trans. on Pattern Analysis & Machine Intell.*, vol. 21, no. 11, pp. 1145–1157, November 1999.

[6] R. Ciampini, L. Blanc-Feraud, M. Barlaud, and E. Salerno, "Motion-based segmentation by means of active contours," in *Proc. IEEE Int'l Conf. on Image Processing*, 1998, vol. 2, pp. 667–670.

[7] A.A. Amini, T.E. Weymouth, and R.C. Jain, "Using dynamic programming for solving variational problems in vision," *IEEE Trans. on Pattern Analysis & Machine Intell.*, vol. 12, no. 9, pp. 855–867, September 1990.

[8] M. Isard and A. Blake, "Condensation – conditional density propagation for visual tracking," *Int. J. Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.

[9] J. Denzler and H. Niemann, "Active-rays: Polar-transformed active contours for real-time contour tracking," *Real-Time Imaging*, vol. 5, no. 3, pp. 203–13, June 1999.

[10] D. Geiger, A. Gupta, L.A. Costa, and J. Vlontzos, "Dynamic-programming for detecting, tracking, and matching deformable contours," *PAMI*, vol. 18, no. 5, pp. 575, May 1996.

[11] H. Tao, H. S. Sawhney, and R. Kumar, "Dynamic layer representation and its applications to tracking," in *Proc. IEEE Int'l Conf. on Comput. Vis. and Patt. Recog.*, June 2000, pp. 134–141.

[12] S.T. Birchfield, "Elliptical head tracking using intensity gradients and color histograms," in *Proc. IEEE Int'l Conf. on Comput. Vis. and Patt. Recog.*, 1998, pp. 232–237.