

Contour Based Object Tracking

Chirag I. Patel and Ripal Patel

Abstract—Contour base Object tracking is an important task in computer vision. The problem involves the tracking of the boundary contour of a moving and deforming object in a sequence of images. Normally the problem is solved as follows. First, the contour of the object is obtained in the first frame. Once, a rough contour of the desired structure is available on the first image of the sequence, the system automatically outlines the contours on the subsequent images at video rate. Contour base Object tracking is useful in many areas such as motion based recognition, automated surveillance, human-computer interaction, traffic monitoring, vehicle navigation etc. Complexity in the problem may arise due to noise in the images, complex object motion, complex object shapes, occlusion etc.

Index Terms—Contour tracking, video processing, kalman filter, color histogram.

I. INTRODUCTION

Object tracking, in general, is a challenging problem. Object tracking is an important task within the field of computer vision. The proliferation of high-powered computers, the availability of high quality and inexpensive video cameras, and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms. There are three key steps in video analysis: detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior. Difficulties in tracking objects can arise due to abrupt object motion, changing appearance patterns of both the object and the scene, no rigid object structures, object-to-object and object-to-scene occlusions, and camera motion. Tracking is usually performed in the context of higher-level applications that require the location and/or shape of the object in every frame. Typically, assumptions are made to constrain the tracking problem in the context of a particular application. There exists many Contour base Object tracking algorithms in the literature. These algorithms can be divided into three categories [1]: Point Tracking (e.g. Kalman Filter [2, 3]), Kernel tracking (e.g. Mean shift [4], KLT tracker [5]), silhouette tracking (e.g. variational methods [6], condensation algorithm [7]). These algorithms mainly differ in the way they use image features and model motion, appearance and shape of the object. In development of any tracking algorithm, we need to have suitable object representation, right features for tracking and good tracking algorithm.

Manuscript received June 12, 2012; revised July 22, 2012.
The authors are with the Computer Science Department, Nirma Institute of Technology, Ahmedabad, India (e-mail: chirag453@gmail.com, ripalpatel315@gmail.com)

II. APPROACH

A. Object Representation

We mark the object in the first frame using an initial contour and represent it by estimating its color histogram. The colors having very low frequency in the color histogram are removed to remove the effect of noise. Use of color histogram to represent the object is motivated by the fact that it is a stable way to represent object that works well for varying illumination conditions [8].

B. Object Tracking

The state of the contour showing the current position of the object is defined using coordinates of its centroid, its velocity and acceleration. The next state of the contour is predicted based on the states seen so far using Kalman filter [9]. To get the correct contour in the subsequent frames in our approach, first centroid of the contour is obtained in the new frame and then the best fit contour is estimated by searching the neighborhood of the contour stochastically. Best fit contour C is a contour which has the maximum energy E , where E is defined as $nF - nB$. Here, nF is the total number of pixels lying inside the contour which belong to foreground object and nB is total number of pixels lying inside the contour which belong to background. Pixel p (which lies inside the contour) is said to be the foreground pixel if $(p_c - c)^T \Sigma^{-1} (p_c - c) < n\lambda$, where c is the centre of a bin in the object histogram and p_c is the color vector associated with pixel p . Σ is covariance matrix and only contains diagonal terms. λ is a user defined constants and n is the dimension of the feature vector. In our experiments, $\lambda = 7$ and $n = 3$.

At any time t , position of the contour is defined by specifying its centroid x_t, y_t and is predicted by using the following dynamics equations:

$$\begin{aligned} x_t &= x_{t-1} + V_{x,t-1} + 0.5a_{x,t-1} \\ y_t &= y_{t-1} + V_{y,t-1} + 0.5a_{y,t-1} \end{aligned} \quad (1)$$

where

$$\begin{aligned} V_{x,t} &= x_t - x_{t-1} \\ V_{y,t} &= y_t - y_{t-1} \end{aligned} \quad (2)$$

and

$$\begin{aligned} a_{x,t} &= V_{x,t} - V_{x,t-1} \\ a_{y,t} &= V_{y,t} - V_{y,t-1} \end{aligned} \quad (3)$$

In above equations, v and a are velocity and acceleration respectively. After estimating the new state of the contour, stochastic search is performed to get the best fit contour. Let

contour C is represented by a set of points (x_i, y_i) , $i = 1, 2, \dots, n$. To get the correct position of the i^{th} contour point (x_i, y_i) , x and y values are searched stochastically in the range $(x_i - 2\sigma_x)$ to $(x_i + 2\sigma_x)$ and $(y_i - 2\sigma_y)$ to $(y_i + 2\sigma_y)$ respectively, where σ_x and σ_y are the standard deviation of the prediction error in x and y direction respectively and are learnt as follows.

$$\sigma_{x,t} = \beta \sqrt{\left(x_{t-1} - \hat{x}_{t-1}\right)^2} + (1-\beta)\sigma_{x,t-1}$$

$$\sigma_{y,t} = \beta \sqrt{\left(y_{t-1} - \hat{y}_{t-1}\right)^2} + (1-\beta)\sigma_{y,t-1}$$

(4)

where

$(\hat{x}_{t-1} - \hat{y}_{t-1})$ and $(x_{t-1} - y_{t-1})$ are the predicted and the exact values of the contour centroid at time $t-1$, and β is the learning rate. After getting the new position is of the contour, modelling parameters of the Kalman filter's are updated based on the prediction error.

III. RESULTS



Fig. 1. One of the frame from the result-1 video.



Fig. 2. One of the frame from the result-1 video.

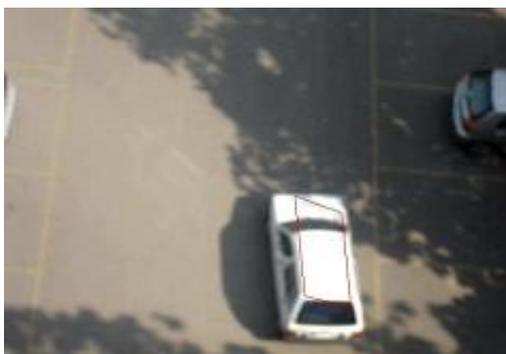


Fig. 3. One of the frame from the result-1 video.



Fig. 4. One of the frame from the result-2 video.



Fig. 5. One of the frame from the result-2 video.



Fig. 6. One of the frame from the result-2 video.



Fig. 7. One of the frame from the result-3 video.



Fig. 8. One of the frame from the result-3 video.



Fig. 9. One of the frame from the result-3 video.



Fig. 10. One of the frame from the result-4 video.



Fig. 11. One of the frame from the result-4 video.



Fig. 12. One of the frame from the result-4 video.

Fig. 1, 2 and 3 present few frames from the result-1 video. Here, we are able to track the car which moves under the shadow of trees. We are able to track the car even though the color of the car is changing.

Fig. 4, 5 and 6 shows few frames from the result-2 video. In this example, we are tracking the man who is riding the motorcycle. We are able to track even when there is some occlusion in the intermediate frames. We are not able to track the whole bike because the color of the bike is the same as the color of the background.

Few frames from the result 3 are shown in Fig. 7, 8 and 9. Here we are tracking a car on runway. The color of the

car is only slightly different from the runway. Still we can able to track the car with reasonably accuracy. Here we also show that we can track the object even when the intermediate frames are skipped. This feature can be useful when the frame rate is high. This video has been taken from [10].

In the result 4, in Fig. 10, 11 and 12, we used a video of 470 frames. In this video, we could only track the car up to 370 frames. After that we lost the track of the car. Figure 4 shows few frames of the result. We can observe that the red contour drawn in the frames does not include the object to be tracked. Reason of this failure is that when car takes turn its color distribution changes very much and becomes very close to background. We found that our method fails to track the objects correctly in the cases where background is of same color as foreground. Various parameters that are required to be tuned in the algorithm are λ , β .

IV. APPLICATION

In our approach, we are finding the velocity and acceleration of the object being tracked. This feature can be used in Traffic monitoring to detect vehicles which violate the speed limit.

We are able to track the objects even while skipping a few frames. This feature can be useful to track objects in high frame rate videos.

V. CONCLUSION

Representing objects using color histogram can be used for fast tracking of objects. This is because the method is computationally inexpensive. The method presented can track the object even when a few frames are skipped. The method presented can take care of partial occlusion of the object. The approach presented fares well only when color object is at least slightly different from that of the background

REFERENCES

- [1] A.Yilmaz, O. Javed, and M.Shah, "Object Tracking: A Survey," *ACM Comput. Surv.*, vol. 38, no. 4, pp.13,2006.
- [2] T. J. Brodia and R. Chellappa, "Estimation of object motion parameters from noisy images," *IEEE trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 1, pp. 90-99, 1986.
- [3] D. terzopoulos and R. Szeliski, "Tracking with Kalman Snakes," pp .3-20, 1993.
- [4] D. Comanicu and P. Meer, "Mean Shift Analysis and application," in *ICCV '99: Proceedings of the International Conference on Computer Vision -Volume 2.Washington, DC, USA:IEEE computer society,1999,p.1197*
- [5] C. Tomasi and J. Shi, "Good features to track," in *CVPR' 94*, 1994, pp.593-600.
- [6] M. Bertalmio, G.Sapiro, and G.Randall, "Morphing Active Contours," *IEEE trans. Pattern Anal. Mach. Intell.*, vol. 22, pp.733-737, 2000.
- [7] M. Isard and A. Blake, "Condensation -conditional density propagation for visual tracking," *Int. J. Comput. Vision*, vol.29, no. 1, pp. 5-28,1998.
- [8] M. J. Swain and D. H. Ballard, "Color Indexing," *Int. J. Comput. Vision*, vol. 7, no. 1, pp.11-32,1991.
- [9] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Prentice Hall, 2003
- [10] R. T. Collons, X. Zhous, and S. K. Teh, "The vivid tracking evaluation website, cmu," [Online]. Available: <http://www.vividevaluation.ri.cmu.edu/main.html>.



Chirag I Patelis pursuing Ph.D. in Computer Science & Engineering from Nirma Institute of Technology, Ahmedabad, Gujarat, India. He received B.E. in Information Technology Engineering from A.D.Patel Institute of Technology, New VallabhVidyanagar, Gujarat, India, in 2006, and the M.Tech. in Computer Science & Engineering from Nirma Institute of Technology, Ahmedabad, Gujarat, India in 2009. His research interests are in video processing, statistical image processing and 3D vision.



Ripal Patel obtained the B.E. degree in Electronics & Communication Engineering from A.D.Patel Institute of Technology, New VallabhVidyanagar, Gujarat, India, in 2006, and the M.E. in electronics & communication Engineering from Dharamsinh Desai University, Nadiad, Gujarat, India in 2009. Her research interests are Computer Vision, Texture Classification, Video Processing and Image Registration