International Journal of Electronics and Electrical Engineering
ISSN : 2277-7040 Volume 2 Issue 4 (April 2012)
http://www.ijecee.com/ https://sites.google.com/site/ijeceejournal/

Kernel based Object Tracking using Color Histogram Technique

[1]Prajna Parimita Dash, [2]Dipti patra, [3]Sudhansu Kumar Mishra, [4]Jagannath Sethi,
[1,2]Dept of Electrical engineering, NIT, Rourkela, India
[3]Department of EEE, BIT, Mesra, India
[4]Department of AE&I, ITER Bhubaneswar
pdash6@gmail.com, dpatra@nitrkl.ac.in,
sudhansu.nit@gmail.com, jagannathsethi@gmail.com

Abstract. Object tracking is the process of locating moving objects in the consecutive video frames. Real time object tracking is a challenging problem in the field of computer vision, motion-based recognition, automated surveillance, traffic monitoring, augmented reality, object based video compression etc. In this paper kernel based object tracking using color histogram technique has been applied for different challenging situations. Covariance tracking algorithm has also been applied to the same problem. From the simulation studies it is clear that this two techniques effectively handle various challenges like occlusion, illumination changes etc. Experimental results reveal that the histogram based method is efficient in terms of computation time and covariance tracker is better in terms of detection rate.

Keywords: Object tracking, Color histogram, Occlusion, Covariance.

1 Introduction

The object detection and tracking are the most essential components of computer vision applications ranging from consumer electronics to smart weapons. Fast and reliable detection of moving objects is important for applications such as video surveillance, navigation, traffic management, video broadcasting, teleconferencing, human-computer interface etc. During video surveillance, it facilitates understanding of motion patterns to uncover suspicious events. In navigation systems it employs to keep the vehicles in lanes and prevent collisions. In traffic management systems it controls the flow of vehicles to keep traffic moving smoothly. Video broadcasting makes use of this to better compress data, improving download speeds for users accessing video files on the internet. Tracking of objects becomes complex due to several problems such as loss of information caused by the projection of 3D world on 2D images, noise in images, complex motion of objects, non-rigid or articulated nature of objects, partial and full object occlusion, change in scene illumination changes in background etc.

Concise survey on video tracking is given by Emanuele Trucco *et al.* and Alper *et al.* in [1,2]. Researchers have proposed many methods for object tracking. The templates and density-based appearance models being proposed by Schweitzer *et al.* in [3]. Instead of templates, color histograms or mixture models can be computed by using the appearance of pixels inside the rectangular or ellipsoidal regions. Fieguth and Terzopoulos [4] generate object models by finding the mean color of the pixels inside the rectangular object region. Comaniciu and Meer [5] use a weighted histogram computed from a circular region to represent the object. Comaniciu used a joint spatial-color histogram instead of just a color histogram [6]. In [7], Kang *et al.* used histograms of color and edges as the object models. Alper used object tracking by asymmetric kernel mean shift with automatic scale and orientation selection [8]. Zoran in [9] applied color histogram based non-rigid object tracking algorithm using mean shift kernel approach. Liu used adaptive template block for matching block of target in [10].

The rest of the paper is organized as follows. The Fundamentals of Object Tracking in Video Sequences areoutlined in Section 2. The two techniques for object tracking is presented and discussed in Section 3. Section 4 provides the simulation results of present studies. Finally the conclusion of the investigation is presented and further possible extension of the work is outlined in Section 5.

2 Fundamentals of Object Tracking in Video Sequences

The basic flow diagram of object tracking is as follows.

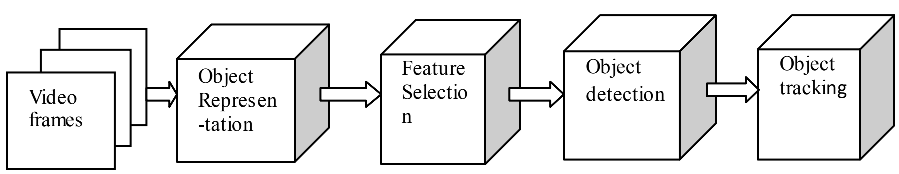

Fig.1.Block diagram of Object tracking process

The object can be represented as a single point or a set of points. It may represented as primitive geometric shapes like rectangle, ellipse or circle. It may be object silhouette, object contour, articulated shape models, skeletal model etc. The shape representation is combined with the appearance representations for tracking purpose. Some commonly used appearance representations are probability densities of object appearance, templates, active appearance model and multi view appearance model. Feature selection is closely related to the object representation. Some of the commonly used features are color, edges, texture, optical flow, gradient etc. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. Some commonly used object detection methods are: point detectors, background subtraction, segmentation and supervised learning. After detection of the object the tracker's task is to generate the trajectory of an object over time by locating the object position in every frame of the video.

3 The Algorithms for Object Tracking

3.1 Kernel-based object tracking using color histogram

In this method a feature space is chosen to characterize the target. The reference target model is represented by its pdf , q in the feature space and in the subsequent frame, a candidate model is defined at location y and is characterized by the pdf , $p(y)$ .The $pdfs$ are estimated from the m -bin histograms. A similar function $\hat{\rho}(y)$ called as the Bhattacharyya coefficient between $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ plays the role of likelihood and its local maxima in image indicate the presence of objects.

The target mode $\hat{\mathbf{q}} = \{\hat{q}_u\}_{u=1\ldots m}, \sum_{u=1}^{m} \hat{q}_u = 1$ (1)

The candidate model $\hat{\mathbf{p}}(\mathbf{y}) = \{\hat{p}_u(y)\}_{u=1\ldots m}, \sum_{u=1}^{m} \hat{p}_u = 1$ (2)

The distance between two discrete distributions is defined as

$$d(y) = \sqrt{1 - \rho[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}]} \tag{3}$$

where $\rho[\hat{\mathbf{p}}(y), \hat{\mathbf{q}}]$ is the Bhattacharyya coefficient.

$$\hat{\rho}(y) = \rho[\hat{\mathbf{p}}(y), \hat{\mathbf{q}}] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}_u} \tag{4}$$

The main advantages of this methods are :(i) this method successfully coped with complex camera motion, partial occlusion of the target, presence of significant clutter, and large variations in target scale and appearance (ii) it is a general object detection algorithm which runs very fast (ii) it is suitable for models that have multiple dominant colors (iii) when there is no object, global optimum needs much less time (iv) position, size, and orientation can be simultaneously and precisely detected and tracked. The main disadvantage of this method are: (i) in this method the spatial information of the target is lost (ii) the similarity measures like Bhattacharya coefficients and Kullback-Leibler divergence are not very discriminative (iii) representation of object histogram disregards the spatial arrangement of the features and do not scale to higher dimensions

(iv) as it depends on the color feature, it can not give good performance when an object and its background have similar colors (v) the color of an object depends on illumination, view point, and camera parameters that tend to change during a long tracking process. So fixed color features are not discriminative enough.

3.2 Object tracking using covariance tracker

For a given object region, the covariance matrix of some features is computed. It is considered as model of the object. In the current frame the region that has minimum covariance distance from the model is determined and that region is assigned as estimated location.

Let I be the observed one dimensional intensity image. F be the $W \times H \times d$ dimensional feature image extracted from I .

$$F(x,y) = \phi(I,x,y) \tag{5}$$

for a given rectangular window region $R \subset F$, $\{\mathbf{f}_k\}_{k=1\ldots n}$ is the d dimensional feature vector inside R .

$$\mathbf{f}_k = [x, y, I(x,y), I_x(x,y), I_{xx}(x,y)] \tag{6}$$

The covariance matrix for the $M \times N$ rectangular region R is calculated as follows.

$$\mathbf{C}_R = \frac{1}{MN}\sum(\mathbf{f_k} - \boldsymbol{\mu}_\mathbf{R})(\mathbf{f_k} - \boldsymbol{\mu}_\mathbf{R})^T \tag{7}$$

where $\boldsymbol{\mu}_\mathbf{R}$ is the vector of the mean of the corresponding features for the points within the region R . To get the most similar region to the target object, the distance between the covariance matrices corresponding to the target object window and the candidate region is calculated as:

$$\rho(\mathbf{C_i}, \mathbf{C_j}) = \sqrt{\sum ln^2 \lambda_k(\mathbf{C_i}, \mathbf{C_j})} \tag{8}$$

where $\lambda_k(\mathbf{C_i}, \mathbf{C_j})$ are the generalized Eigen values of $\mathbf{C_i}$ and $\mathbf{C_j}$ computed from

$$\lambda_k \mathbf{C_i}\mathbf{x_k} - \mathbf{C_j}\mathbf{x_k} = 0,\ k = 1,2,\ldots..d \tag{9}$$

where $\mathbf{x_k}$ are the generalized Eigen vectors.

At each frame we search the whole image to find the region which has the smallest distance from the current object model. The best matching region determines the location of the object in the current frame.

The main advantages of this method are: (i) the covariance matrix can take many possible features such as coordinate, color, gradient, edge, texture, motion etc. Hence it captures the spatial and statistical properties (ii) the covariance matrices are low-

International Journal of Electronics and Electrical Engineering
ISSN : 2277-7040 Volume 2 Issue 4 (April 2012)
http://www.ijecee.com/ https://sites.google.com/site/ijeceejournal/

dimensional compared to other region descriptor and due to symmetry C_R has only $(d^2 + d)/2$ distinct values (iii) it works detecting single as well as multiple rigid and non-rigid objects (iv) it works satisfactorily at object deformations and appearance change (v) noise is largely filtered out during covariance computation (vi) the covariance matrix of any region has the same size, thus it enables comparing any regions without being restricted to a constant window size (viii) covariance is invariant to the mean changes such as identical shifting of color values. This became an advantageous property when objects are tracked under varying illumination conditions.

The main disadvantages of this method are: (i) The covariance matrices lie in a Riemannian space, hence an appropriate distance metric has to be used when comparing regions (ii) computation is not very fast as tracking required global search. Hence the tracking of very fast moving object is difficult.

4 Simulation results

Fig. 2.(a)

Fig. 2.(b)

Ball sequence frame 1,10,25(left to right) Fig.1(a): Histogram based, Fig 1(b): Covariance based

International Journal of Electronics and Electrical Engineering
ISSN : 2277-7040 Volume 2 Issue 4 (April 2012)
http://www.ijecee.com/ https://sites.google.com/site/ijeceejournal/

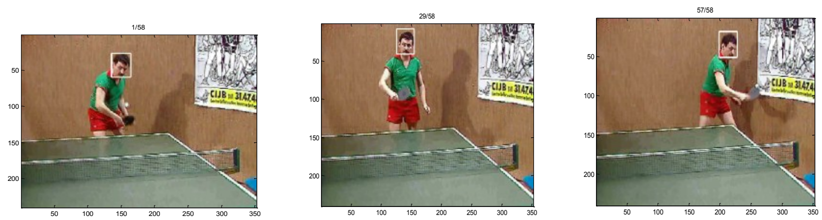

Fig.3.(a)

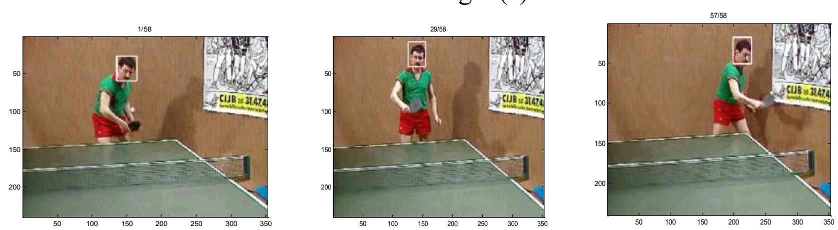

Fig. 3.(b)

Table tennis player sequence frame 1,29,57(left to right) Fig.2(a): Histogram based, fig 2(b): Covariance based method. From this results it clear that all methods are able to track the object efficiently.

Fig..4(a)

Fig.4.(b)

Figure.6.Football player sequence2 with uneven lighting condition:frame 26,49,76 (left to right) Fig.3.(a): Histogram based, Fig.3(b): Covariance based

International Journal of Electronics and Electrical Engineering
ISSN : 2277-7040 Volume 2 Issue 4 (April 2012)
http://www.ijecee.com/ https://sites.google.com/site/ijeceejournal/

Detection rate is the ratio of the number of frames the object location is accurately estimated to the total number of frames in the sequence. The detection rate of two methods applied to three different video sequences are shown in Table 1. The covariance tracking method is more efficient as compared to other method in terms of detection rate. .

Table.1 Comparison of detection rate of two methods

	Histogram Based		Covariance method		Remark
	Frame Missed/Total	Detection rate	Frame Missed/total	Detection rate	Performan ce of covariance method is best
Ball video	2/92	97.82 %	0/92	100%	
Table Tennis video	2/58	96.55 %	0/58	100%	
Football video	58/427	86.42 %	16/427	96.25%	

An experiment has been carried out to compare the two methods in terms of CPU time required in second to detect object in each frame and detection rate.

Table.2 Comparison of CPU time between two methods

	Histogram Based	Covariance method	Remark
	CPU Time/frame	CPU Time/frame	Performance of histogram based method is better in terms of computation time
Ball video	500 msec	600msec	
Tennis video	550msec	700 msec	
Football video with uneven lighting condition	580 msec	730 msec	

The experiment results are shown in Table 2 which reveals that for a histogram based method, the search takes about ~500 msec/frame for ball video which is less as compared to covariance method. The program is run in MATLAB with a p4, 3.2 GHz m/c for 320x240 images.

5 Conclusion and future work

In this paper two techniques have been applied for object tracking for different challenging situations. From the experiments results it can be conclude that the object tracking using this two methodsare a simple and efficiently handle the challenging situations like occlusion, multiple objects, faster object, complex background, etc. The kernel based object tracking using color histogram method is not very good in terms of detection rate but its computational time is less as compared to the other method. The object tracking using covariance tracker method is more efficient as compared to histogram method in terms of detection rate. Future research work in the area includes incorporation of other features like texture and motion vector to improve the robustness of tracking algorithm. Statistical models based approach may be applied to the problem. Evolutionary computing methods for object initialization may apply to the problem.

References

[1] EmanueleTrucco and KonstantinosPlakas, (2006) "Video Tracking: A Concise survey" IEEE Journal.Ocean Eng., vol. 31, no. 2, pp. 520-528.

[2] AlperYilmaz, Omar Javed and Mubarak Shah,(2006) "Object Tracking: A Survey" ACM Comput. Surv.38,4,Article 13,December.

[3] H.Schweitzer, J.W.Bell, andF.Wu. (2002), "Very fast template matching".European Conference on Computer Vision (ECCV).pp.358–372.

[4] P.Fieguth, and D.Terzopoulos.(1997)"Color-based tracking of heads and other mobile objects at videoframe rates". In IEEE Conference on Computer Vision and Pattern Recognition (CVPR).pp.21–27.

[5] D. Comaniciu,V.Ramesh, and P.Meer.(2003)"Kernel-based object tracking". IEEE Trans. Patt. Analy. Mach. Intell. 25, pp-564–575.

[6] D.Comaniciu, P. Meer, (2002) "Mean shift: A robust approach toward feature space analysis". IEEE Trans. Patt. Analy.Mach. Intell.24, 5, pp-603–619.

[7] J.Kang, I. Cohen, G. Medioni.(2004)"Object reacquisition using geometric invariant appearancemodel".International Conference on Pattern Recongnition (ICPR).pp-759–762.

[8] AlperYilmaz, (2007) "Object Tracking by Asymmetric Kernel Mean Shiftwith Automatic Scale and Orientation Selection," IEEE ComputerSociety Conference on Computer Vision and Pattern Recognition, pp.1-6.

[9] Hong Yu, Jie Wei, Jin Li, (2009) "Object Tracking by Mean Shift Based onColor Distribution and Simulated Annealing," International Seminaron Future Information Technology and Management Engineering, pp.128-131.

[10] E. Zhou ,C.Liu, Y. Sun, Z. Wang . S.Gong.(2010) "Adaptive tracking window updating algorithm based on particle filtering". IEEE International congress on Image and Signal processing.pp.303-307.