Available online at www.elixirpublishers.com (Elixir International Journal)

Electrical Engineering

Elixir Elec. Engg. 67 (2014) 21230-21233

ISSN: 2229-712X

Object tracking using Parzen window based mean shift

Nandini K. Bhandari[1], R.R. Bhambare[1], R.D. Raut[2] and M.M.Mushrif[3]

[1]Deparment of Electronics Engineering, Pravara Rural Engineering College, Loni. Tal. Rahata Dist. Ahmednagar.

[2]Department of E & TC Engineering, Ramdeobaba College of Engineering, Nagpur.

[3]Department of E& TC Engineering, Yashvantrao Chavan College Of Engineering, Nagpur.

ARTICLE INFO

Article history:

Received: 10 December 2013;
Received in revised form:
20 January 2014;
Accepted: 1February 2014;

Keywords

Mean shift,
Kernel density estimator (KDE),
Parzen window,
Gaussian kernel,
Object tracking.

ABSTRACT

Mean shift tracking is a widely used tool for robust and quick tracking of the object in video, against partial occlusion and clutter. This paper proposes a Parzen window based mean shift tracker for visual object tracking. It combines three RGB color histogram and generates its Kernel Density estimation using Gaussian Kernel. Parzen window density estimation method is used to solve the basic problems of mean shift like, divide by zero problems in the weight computation and its association to tracking interruption. The Parzen window interpolates the histogram of target candidate. The experimental results show that proposed method performs better than traditional mean shift for fast moving objects and objects having background clutter at moderate increase in computational complexity.

Introduction

Real time object tracking is gaining a lot of interest among the researchers of the computer vision. It has wide range of applications in the fields like video surveillance system, traffic monitoring, weather forecasting, human computer interaction and a lot of military applications [2-9]. An extensive survey on variety of algorithms developed for tracking is done in [1]. These algorithms can be classified in two approaches viz. top down and bottom approaches [10]. In top down approaches a set of state hypotheses based on target model are generated. These hypotheses are verified in consecutive frames in video to evaluate the performance of tracking. To have robust tracking large number of hypotheses is required but this may increase computational cost. Particle filter is one of the typical top down approach. It estimates the probability density function of state vector for nonlinear and non Gaussian distribution for robust visual object tracking [11]. Though particle filters work well in partial occlusions and cluttered background, it suffers from two major issues. Firstly computational cost increases because of large number of particles. Secondly the motion model is learned by sample images before tracking is performed. So tracking fails in real time. Bottom up approach generate target model by analyzing the image content.These methods are computationally efficient but their robustness depends on image analysis. Non parametric statistical methods fall in this category. It allows the data guide to search for the function that fits best to get the target object [12]. Colour distribution histograms are one of the non parametric approaches which are used for tracking non rigid objects. They are calculated efficiently. They are robust to partial occlusion and clutter and are rotation and scale invariant [13, 14].

Fukunaga and Hostetler introduced mean shift algorithm [15] which is a general clustering algorithm. It is a non-parametric in nature. In this histograms of grey scale intensities of object to be tracked and candidate to be matched are constructed. Similarity between the two is measured using Bhattacharya measure. For tracking weights are assigned to histogram bins, which are projected back to the image such that each pixel gets a weight equal to the weight assigned to the corresponding bin to which it belongs. The shift in location of the object being tracked is found by finding the mean of the back-projected weights in the tracking window. The tracking window is shifted to the new mean location after each iteration. This algorithm is commonly used for tracking because it is robust and easy to implement. However standard mean shift suffer from number of problems which affect tracking performance. It is not adaptable to change in object size. Its performance also depends on kernel size selected in object initialization phase. The tracking may shift or even fail when objects are occluded or object colour is same as background. Much research work has been done to reduce these problems. Cheng improved performance of mean shift using clustering, a class of kernels [5]. The idea of kernel based tracking is used by many researchers [17, 18]. The method in [17] is unfeasible as it uses repetitions of mean shift algorithm at each iteration by changing window size by ± 10%. Birchfield [19] used additional spatial information to increase robustness of tracking. Spatiogram is a type of histogram where every bin is weighted by spatial mean and covariance of pixels. Spatiogram with mean shift provide better matching performance. Mean shift with adaptive feature selection is used in [20]. CAMSHIFT [21] uses different weight for each pixel to get maximum of similarity. This can track scale and orientation but can be used only when object has single colour. Method in [22] tries to solve problem that object colour is same as background. The position of object in previous and current frame is considered to compute target model. Though it has low computational cost it fails during occlusion and merging of objects. [23] try to solve the problem of incorrect kernel scale selection in mean shift. It uses difference of Gaussians kernel to handle changes in target scale. However it is not suitable in real time tracking. Many authors combine mean shift tracking with Kalman filter [24] or particle filter [14] to get better results. Annealed Mean Shift [25] is used for global mode seeking. However these algorithms cannot

Tele:
E-mail addresses: nandinibhandari1969@gmail.com

simultaneously solve the problem of scale, orientation and position. Tracking through scale space has high computational cost and cannot reach real time requirement.

In our paper we use traditional mean shift with Parzen window. Parzen window is used to interpolate the histogram of target candidate to improve the performance of tracking. The mean shift algorithm combines three RGB color histograms into a single histogram for tracking the target [18]. Normalization is applied on image to reduce the effect of illumination variations and noise.

The remaining paper is organized as follows. Section II describes about the Parzen window and Gaussian kernel. The mean shift tracking algorithm along with a simple method of combining RGB histograms in a single histogram and Parzen window is given in section III. Section IV describes the experimental results. Section V deals with conclusion.

Parzen Window

Parzen Window technique in pattern recognition [26, section 4.3] is a popular kernel density estimation method. It defines window function as

$$\emptyset_u = \{1 \; if \; |u| \le \frac{1}{2} \qquad (1)$$

0 otherwise

This is unit hypercube centered at origin. If the function is Gaussian in d dimensions then

$$\emptyset_G(u) = \left(\frac{1}{2\pi}\right)^d e^{-u^2}/2$$

Suppose we have n data points $X_1, X_2, \ldots X_n$ in a d dimensional space R^d, then the multivariate kernel density estimator f(x) with Parzen density estimation at x is given by

$$\hat{f}(X) = \frac{1}{nh_d}\sum_{i=1}^{n} K\frac{(X-X_i)}{h} \qquad (2)$$

Or $\hat{f}(X) = \frac{1}{nh_d}\sum_{i=1}^{n} k \left\|\frac{X-X_i}{h}\right\|^2 \qquad (2)$

Where K(X) (profile of k(x)) is a kernel function which is non-negative and piecewise continuous. 'h' is the interpolating window bandwidth. So the Parzen window distribution is a true distribution and variance tend to zero when no. of samples is large. In practice we have finite number of samples and window size effectively interpolates the data.

The quality of kernel density estimator is measured by the mean squared error between the actual density and the estimate. A more commonly used kernel is a Gaussian kernel

$$K(X) = \frac{1}{\sqrt{(2\pi)^d}} exp - \frac{1}{2}\|X\|^2 \qquad (2)$$

The term $\frac{1}{\sqrt{(2\pi)^d}}$ is normalization constant.

When kernel is Gaussian $K' \alpha K$ we get a simple and elegant mean shift algorithm.

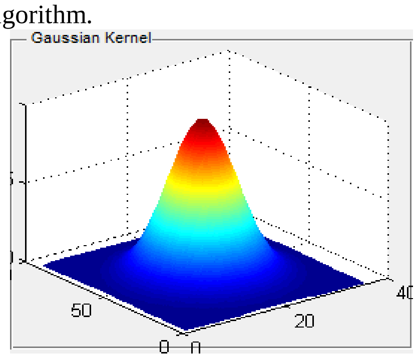

Figure 1: Gaussian kernel

Mean Shift

In Mean shift algorithm [17] an object color model is obtained by calculating its histogram in tracking rectangle. Color distributions achieve robustness against non-rigid, partial occlusions and rotations. So it tracks the object in video by matching the probability of target with that of object color model. Normalization of color is used as it divides out pixel luminance (brightness) and leaves only chromacity information so result is less sensitive to illumination variations.

Suppose Xi, i =1,2,…..n denotes the pixel locations of the object model centered at O. Color histogram [27] is used to describe the object. Each subspace of RGB color space is divided into m intervals. Each interval is known as bin. For each bin the probability of all the pixels in the object area is calculated and integrated into color histogram.

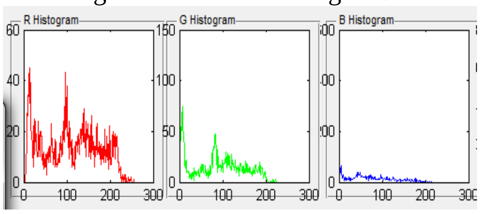

Figure 2: RGB histograms of object to be tracked

Let b(Xi) denote the color bin of the color at Xi. Choose kernel bandwidth h =1 by assuming that the size of the model is normalized.

The probability q of color u, u=1, 2… m, in the object model is

$$q_u = C\sum_{i=1}^{n} k(\|X_i\|^2)\delta(b(x_i) - u) \qquad (3)$$

where k(x) is a kernel profile, δ is the Kronecker delta function and C is a normalization constant

$$C = \left[\sum_{i=1}^{n} k(\|X_i\|^2)\right]^{-1} \qquad (4)$$

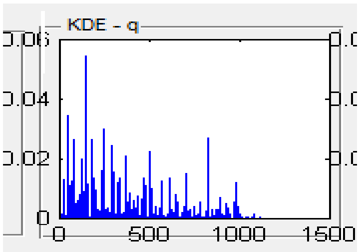

Figure 3: Kernel density Estimation of Object model

Once the frame with object is detected, every frame after that having object area is called target candidate area. Let y_i= 1, 2,……..,n_h denote the pixel locations of the target centered at y. Then the probability p of color u in the target is given by

$$p_u(y) = C_h\sum_{i=1}^{n_h} k\left(\left\|\frac{y-y_i}{h}\right\|^2\right)\delta(b(y_i) - u) \qquad (5)$$

C_h is the normalization constant.

$$C_h = \left[\sum_{i=1}^{n_h} k\left(\left\|\frac{y-y_i}{h}\right\|^2\right)\right]^{-1} \qquad (6)$$

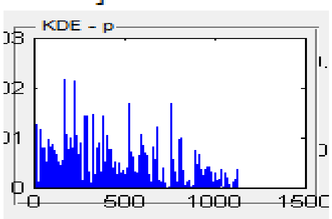

Figure 4: Kernel density estimation of Target candidate

Similarity function [28] describes the similarity between the object model q and the target candidate model p at location y. Use Bhattacharyya coefficient ρ as a simlarity function

$$\rho(y) = \rho(p(y), q) = \sum_{u=1}^{m} \sqrt{p_u(y)\, q_u} \qquad (7)$$

ρ is the cosine of vectors $(\sqrt{p_1}, \dots\dots\dots, \sqrt{p_m})^T$ and $(\sqrt{q_1}, \dots\dots\dots, \sqrt{q_m})^T$. Large vaqlue of ρ indicates a good colour match. To make $\rho(y)$ largest define target center in the current frame as target center y_0 in the previous frame. Search for a new target location z near y where the color probability density does not change drastically.

By Taylor series expansion

$$\rho(p(z), q) = \frac{1}{2}\sum_{u=1}^{m} \sqrt{p_u(y)\, q_u} + \frac{1}{2}\sum_{u=1}^{m} p_u(z) \sqrt{\frac{q_u}{p_u(y)}} \quad (8)$$

Substituting eq.(5) in eq. (8) gives

$$\rho(p(z), q) = \frac{1}{2}\sum_{u=1}^{m} \sqrt{p_u(y)\, q_u} + \frac{C_h}{2}\sum_{u=1}^{n_h} w_i\, k(\|\frac{z-y_i}{h})^2 \quad (10)$$

where weight w_i is

$$w_i = \sum_{u=1}^{m} \delta(b(y_i) - u) \sqrt{\frac{q_u}{p_u(y)}} \quad (11)$$

So to maximize $\rho(p(z), q)$, we need to maximize second term in equation (10).

Mean Shift Algorithm

Given the distribution $\{q_u\}$ of the object model and location y of the target in the previous frame

1. Initialize the target location y in the current frame.
2. Apply normalization to each frame.
3. Compute $\{p_u(y)\}$, u=1,2,…..m and find similarity function $\rho(p(y), q)$.
4. Calculate weights w_i for i=1,2,…,n_h using equation (11).
5. Applying mean shift find new location of target z as

$$z = \frac{\sum_{i=1}^{n_h} w_i\; g(\|\frac{y-y_i}{h}\|^2) y_i}{\sum_{i=1}^{n_h} w_i\; g(\|\frac{y-y_i}{h}\|^2)}$$

where g(x) = -k′(x)

6. Update $\{p_u(z)\}$ and $\rho(p(z), q)$.
7. If $\rho(p(z), q) \langle\ \rho(p(z), q)$ then

$z \leftarrow \frac{1}{2}(y + z)$

8. If$\|y-z\| \langle \in$stop otherwise set $y \leftarrow z$ and goto step1.

Experimental Results

Several video sequences are used to evaluate the proposed method of mean shift tracking. The algorithm is implemented in MATLAB 12a. In all the experiments, RGB color model was used as feature space. RGB was quantized into 16×16× 16 bins. We have normalized the image. Normalization of color divides out the pixel luminance and leaves only chromacity information, so the results are less sensitive to illumination variation. The appearance of the object is modeled by kernel density estimated (KDE) in the RGB space. For practical purpose, we have used the weighted histogram approximation of KDE using the Gaussian kernel. In all the experiments we have chosen Parzen window width as 0.9 and value of $\in$ as 0.001.Number of iterations in the mean shift is kept 10.

The first sequence is a ping-pang ball sequence. The sequence has 52 frames of spatial resolution 352×240. Experiment results shows that tracking is lost quickly as ball has fast movement.

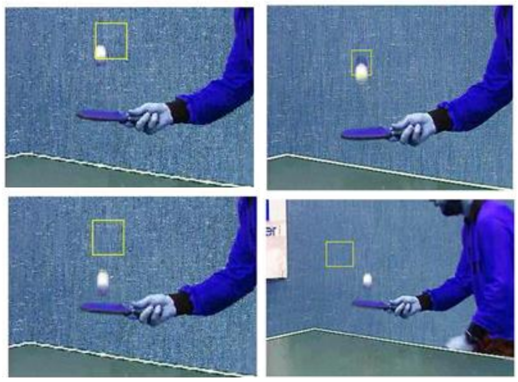

Figure 5: Mean shift tracking results on the ping pang ball sequence.

In the second experiment we have considered a sequence of table tennis player. In this we track head of the player. The target is selected such that it contained much part of background. The results show that prominent features of target model are enhanced and relevance of background is decreased significantly.

A PET 2000 sequence is also tested. The sequence has 520 frames of spatial resolution of 480×320. The sequence is tested for 10 iterations and 5 iterations and $\in$= 0.001 and tracked object path is plotted in MATLAB. It is observed that in this sequence variation in iteration have not affected the tracked path a lot. If the value of $\in$ is changed from 0.001 to 0.005 tracking stops in the midway of the sequence. As we increase the value of $\in$ from 0.001to 0.05 efficiency of tracking reduces.

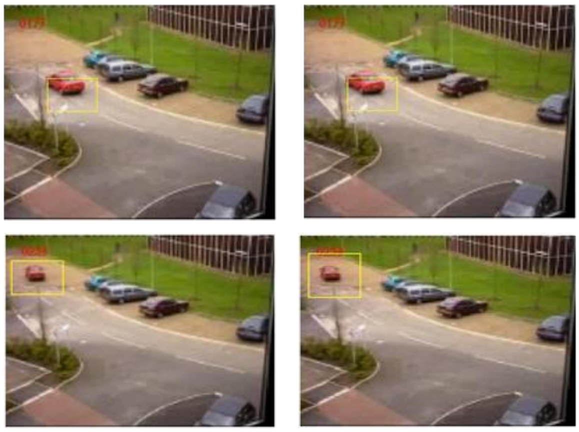

Figure 6: Mean shift tracking results on PET2000 sequence

The mean shift algorithm is also tested for black and white sequence like walk complex and walk simple. The algorithm does not track the object.

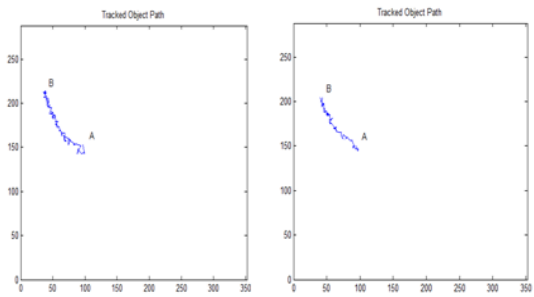

Figure 7: (a) Tracked object path with iterations 5 and $\in$=0.01 and (b) Tracked object path with iterations 10 and $\in$=0.001.

The algorithm was tested on PET2000, VIP man and Ant sequence for variations in Parzen window width. The results are tabulated in table 1. The results show that as bandwidth reduces tracking stops in earlier frames. It is also observed that if we increase window width above 1.0 mean shift algorithm fails. If we increase the number iterations tracking is lost. Results shown in table1 show that proposed Parzen window based mean shift gives better results than the traditional mean shift.

Table 1: Results with traditional mean shift and proposed Parzen window based mean shift

Sr. No	No. of iterations (10)	No. of frames after which tracking stops		
		PET2000	VIP man	Ant
1	Traditional Mean shift	6	4	4
2	Parzen window based mean shift	15	5	12

Table 2: Variation of window size

Sr. No	Size of Parzen window	No. of frames after which tracking stops		
		PET2000	VIP man	Ant
1	0.5	10	4	4
2	0.6	20	6	15
3	0.7	45	9	30
4	0.9	Continues till the end of sequence	15	54

Conclusion

In this paper a mean shift algorithm based on Parzen window is proposed. It uses Gaussian kernel to estimate kernel density. It proposed algorithm works well on both indoor and outdoor samples. The experimental results also show that proposed method performs better than traditional mean shift for fast moving objects and objects having background clutter at moderate increase in computational complexity. Though normalization of image is done to overcome illumination variations, the color analysis is not enough for tracking as color map is continuously changing. When the environment is extremely complex our system will lose the target. These problems are to be solved in future.

References

[1] Yilmaz A., Javed O., Shah M. “Object Tracking: a survey”, ACM computer survey, 38,(4),2006.

[2] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer, “Multi-camera multi-person tracking for Easy Living,” in Proc. IEEE Intl. Workshop on Visual Surveillance, Dublin, Ireland, 2000, pp. 3–10.

[3] J. Ning, Lei Zhang, David Zhang and C. Wu, “Robust Mean Shift Tracking with Corrected Background-Weighted Histogram,” to appear in IET Computer Vision.(2011)

[4] Huimin Qian, Yaobin Mao, Jason Geng, and Zhiquan Wang, “Object tracking with self-updating tracking window,” in PAISI,2007, vol. 4430, pp. 82–93.

[6] L.Li, "An Efficient Sequential Approach to Tracking Multiple Objects Through Crowds for Real-Time Intelligent CCTV Systems", IEEE transactions on system, man, and cybernetics—part B: cybernetics, vol. 38, no. 5, October 2008.

[7] P.Bai,"Person-Tracking with Occlusion Using Appearance Filters", Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems October 9 - 15, 2006, Beijing, China.

[8] Y.Wu,F.Lian, T.Chang ,"Traffic Monitoring and Vehicle Tracking using Roadside Cameras",2006 IEEE International Conference on Systems, Man, and Cybernetics, Taipei, Taiwan.

[9] P.Vadakkepat, "Multimodal Approach to Human-Face Detection and Tracking", IEEE transaction on industrial electronics, vol. 55, no. 3, March 2008.

[10] W. Ying, “Robust visual tracking by integrating multiple cues based on co-inference learning”, International Journal on Computer Vision 58, 55.71 (2004).

[11] N.J. Gordon, A. Doucet and N. D. Freitas, “On Sequential Monte Carlo Sampling Methods for Bayesian Filtering,” Statistical Computing vol.10,no.3,pp197-208,2000.

[12] J.R. Jimenez, V. Medina, and O. Yanez, “Nonparametric MRI segmentation using mean shift and edge confidence maps.” Proc. SPIE 5032, 1433.1441 (2003).

[13] D. Comaniciu, V. Ramesh, and P. Meer, “Real-time tracking of non-rigid objects using mean shift”, Computer Vision and Pattern Recognition 2, 142.149 (2000).

[14] K. Nummiaro, E. Koller-Meier, and L. Van Gool, “An adaptive colour-based particle filter.” Image and Vision Computing 21, pp99-110 ,2003

[15] Fukunaga K, and Hostetler LD, “The estimation of the gradient of a density function, with applications in pattern recognition,” IEEE Trans. Information Theory, vol. 21,pp.32-40, 1975.

[16] Cheng Y, “Mean shift, mode seeking, and clustering,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.17, no.8, pp.790-799, 1995.

[17]Comaniciu D, Ramesh V, Meer P. “Kernel-based objecttracking”.*IEEE Trans. PAMI*. Vol.25, pp564-577, 2003.

[18] Comaniciu D., Peter Meer “Mean shift: A robust approach towards feature space analysis” IEEE Trans Pattern Anal. Machine Intell. , 24, (5), pp. 603-619,2002.

[19] Birchfield S.T., Rangarajan, S.: “Spatiograms versus histograms for region-based tracking”. Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), San Diego, CA USA, June 2005, pp. 1158–1163.

[20] Vladimir Nedovic, Martijn Liem, Maarten Corzilius and Mark Smids, "Kernel-based object tracking using adaptive feature slection," project report, 2005.

[21] Bradski, G.R., "Computer Vision Face Tracking for Use in a Perceptual User Interface, "IEEE Workshop on Applications of Computer Vision, Princeton, NJ, 1998, pp. 214- 219.

[22] Boonsin, M., Wettayaprasit, W., Preechaveerakul, L.: “Improving of mean shift tracking algorithm using adaptive candidate model”. Proc. ECTI-CON, Chiang Mai, Thailand, June 2010, pp. 894–898

[23] Collins, R.T.: “Mean-shift blob tracking through scale space” Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), Madison, WI, USA, June 2003, pp. 234–240.

[24] Zhiwei Zhu, Qiang Ji, Fujimura, K. and Kuangchih Lee, “Combining Kalman filtering and mean shift for real time eye tracking under active IR illumination," Proceedings. 16th International Conference on Pattern Recognition, 2002, vol. 4, pp. 318- 321.

[25] Shen, C. Brooks, M. J. van den Hengel, A., “Fast Global Kernel Density Mode Seeking: Applications to Localization and Tracing," IEEE Transactions on Image Processing, 2007, Vol. 16, pp. 1457-1469.

[26]] R. O. Duda and P.E.Heart, “Pattern Classification and Scene Analysis,”Wiley, 1973.

[27] IdoLeichter, Michael Lindenbaum, Ehud Rivlin,” Mean Shift tracking with multiple reference color histograms” Elseiver transaction on Computer vision and image understanding., pp 400-408, 2010.

[28] Juan Wang, Jiaomin Liu, JunyingMeng, Ming Han,” Research on an improved mean shift algorithm,” Elsevier Procedia Engineering, vol.15, pp 3572-3576, 2011.