International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Object Tracking Using Joint Color-Texture Histogram

Pallavi C. Hingane[1], S. A. Shirsat[2]

[1, 2] Department of Electronics and Telecommunication, Sinhgad College of Engineering Pune, Maharashtra, India

Abstract: ***Object tracking is active research topic in field of computer vision. It is widely used in many applications such as security, surveillance, traffic monitoring. Mean shift object tracking algorithm is presented in this paper by using the joint color-texture histogram to represent a target object. The objective of this work is to track moving object with help of mean shift tracking algorithm. Mean shift algorithm is implemented to increase tracking accuracy. Many tracking algorithms have been proposed to overcome difficulties arising from noise, changes in background environment. This algorithm is more suitable due to simplicity. Using only color histogram in mean shift tracking has some problems. First spatial information is lost. Second, when target has similar appearance to the background, color histogram will become invalid to distinguish them. To overcome this drawback local binary pattern method is used. The performance of this algorithm improves tracking accuracy with fewer mean shift iterations.***

Keywords: Object tracking; color histogram; local binary pattern; feature extraction

1. Introduction

Real time object tracking is a critical task in computer vision. Many tracking algorithms have been proposed to overcome difficulties arising from occlusion, noise, clutter and changes in the background or foreground object in the environment. Among the different tracking algorithms, mean shift object tracking algorithms have recently become more popular due to their simplicity. Mean Shift tracking is an iterative kernel based deterministic procedure. It converges to a local maximum of the measurement function with and certain assumptions on kernel behaviors. Mean shift is a low complexity based algorithm. It provides a reliable solution to object tracking. This algorithm is independent of the target representation.

Color histogram is an estimating mode of point of sample distribution. It is very robust in representing the target object appearance. Using only color histogram in mean shift tracking has some problems. First spatial information is lost. Second, when target has similar appearance to the background, color histogram will become invalid to distinguish them[1-2]. It's aim is to locate a moving object video scene. This algorithm is used for analyzing the video frame by frame. Main goal of algorithm is to determine target object in each frame and to returns in its location. Main task is to find moving object in image sequences. Tracking objects can be complex due to loss of information caused by projection of the 3D on a 2D image, noise in images, object complex motion, articulated nature of objects, partial object occlusions, complex shapes scene[1-4].Object tracking is to follow recognition step in the image processing. Better target representation, the edge or gradient features have been used in combination with color histogram. Several object representation that exploits the spatial information. This information has been developed by partitioning tracking region into fixed size fragments or the articulations of human objects. Each sub region, a color or edge feature based target model was presented. The texture pattern is to reflect the spatial structure of the object. They have features to represent and recognize targets. The texture features introduce new information. Color histogram does not convey. By using the joint color-texture histogram for target representation is more reliable than using only color histogram in tracking complex video scenes. Idea of combining edge and color for target object representation has been exploited by researchers. The local binary pattern (LBP) technique is very effective to describe image texture features.

LBP has advantages such as rotation invariance fast computation and, which facilitates wide usage in the fields of image retrieval, texture analysis, and face recognition. LBP was successfully applied to the detection of moving objects. In LBP, each pixel is assigned a texture value. It can b e naturally combined with the color value of the pixel to represent targets. LBP feature is used to construct a two dimensional histogram representation of the target for tracking monochromatic and thermo graphic video. This target representation scheme eliminates noise and smooth background in the tracking process. Compared with the traditional RGB color space based target representation, it exploits the target structural information. This methode improves greatly tracking accuracy with fewer mean shift iterations than standard mean shift object tracking. It can robustly track target object under complex scenes, such as similar background appearance and target.

2. Mean Shift Object Tracking

Mean shift object tracking is consist of two parts, first is appearance description and second one is tracking. Figure2.1 shows two parts of mean shift algorithm.

Licensed Under Creative Commons Attribution CC BY

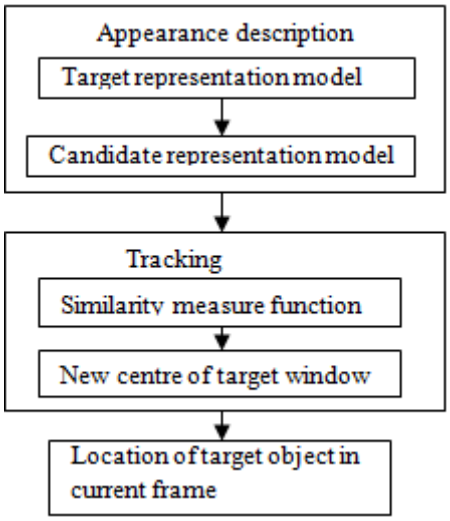


Figure 2.1: Two Parts of Mean Shift Tracking

The mean shift algorithm is used for visual tracking. A target is defined by a rectangular or an ellipsoidal region in the image frame. Most existing target objects tracking schemes are used in the color histogram. At initial frame target joint color texture histogram is prepared. This model is used for target tracking. This model estimates maximum similarity between target model and candidate model.

2.1 Target Representation

First part of mean shift tracking is target re presentation. A target is usually defined by a rectangle or an ellipsoidal region in the image. This paper is to represent a new target representation approach by using the joint color-texture histogram. Denote {xi*} i=1•••n the normalized pixel positions in the target region. This is supposed to be centered at origin point.

$$\hat{q} = \{\hat{qu}\} \quad u = 1 \ldots m \qquad \text{-------(1)}$$

$$\hat{qu} = C \sum_{i=1}^{n} k(\|xi*\|^2)\, \partial[\, b(xi*) - u] \qquad \text{-------(2)}$$

$\hat{qu}$ represents the probabilities of feature u target model. m is number of feature spaces. b(xi∗) associates the pixel xi∗ to the histogram bin. . k(x) is an isotropic kernel profile. δ is the Kronecker delta .Constant C is a normalization function defined by,

$$C = 1 \,/\, \sum_{i=1}^{n} k(\|xi*\|^2) \qquad \text{-------(3)}$$

Candidate model $\hat{pu}(y)$ corresponding to the candidate region is computed as,

$$\hat{p}(y) = \{\hat{pu}(y)\} u = 1 \ldots . m \qquad \text{-------(4)}$$

$$\hat{pu}(y) = Ch \sum_{i=1}^{nh} k(\|y - xi/h\|^2)\, \partial[\, b(xi) - u] \qquad (5)$$

$\hat{pu}(y)$ represents as the probability of feature u in the candidate model $\hat{pu}(y)$. {xi}i=1⋯nh denotes pixel positions in target candidate region. It's center at y.Ch is a constant normalization function. h is bandwidth.

$$Ch = 1 \,/\, \sum_{i=1}^{nh} k(\|y - xi/h\|^2) \qquad \text{-------(6)}$$

Bhattacharyya coefficient is defined as two normalized histograms in between $\hat{pu}(y)$ and $\hat{qu}$.Target model and candidate model two metrics are based on Bhattacharyya coefficient.

$$\rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^{m} \sqrt{[\hat{pu}(y)\ \hat{qu}]} \qquad \text{-------(7)}$$

The distance between $\hat{pu}(y)$ and $\hat{qu}$ is given by

$$d[\hat{p}(y), \hat{q}] = \sqrt{1 - \rho[\hat{p}(y), \hat{q}]} \qquad \text{-------(8)}$$

To measure similarity function between target model and candidate model are used. After estimating similarity function between target and candidate model the new center of tracking window is calculated. To minimizing the distance between target and candidate model is equivalent to maximizing bhattacharyya coefficient in shown in above equation no-5.

3. Local Binary Pattern

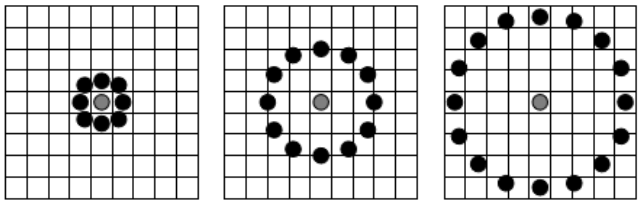
3.1 Texture Feature Extraction

The LBP operator labels pixel in an image by thresholding its neighborhood with center value and considering the result as a binary number (binary pattern).

Image textures features can be used to help in image segmentation or image classification. LBP operator is defined as:

$$LBP_{P,R}(Xc, Yc) = \sum_{p=0}^{p-1} s(gp - gc)2^p \qquad \text{------(9)}$$

gc value corresponds to gray value of centre pixel (xc , yc) of a local neighbourhood. gp is gray value of P equally spaced pixels on the circle with radius R.. Figure3.1 shows that three circularly symmetric neighbour sets for different values of P and R.



P=8,R=1.0 P=12,R=2.5 P=16,R=4.5

Figure 3.1: Three circularly symmetric neighbour sets for different values of P and R.

By varying P and R, an LBP operator goes on under the different quantization of the angular space and spatial resolution, and multiresolution. The function s(x) is defined as follows:

s(x) = 1 x>=0

= 0 x<0 ------(10)

The texture model is derived by gray-scale invariance. The grayscale and rotation invariant for LBP texture model is obtained by,

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{p-1} s(gp - gc) & \text{if } U(LBP_{P,R}) \le 2 \\ p + 1 & = \text{Otherwise} \end{cases} \qquad \text{--(11)}$$

Licensed Under Creative Commons Attribution CC BY

Where

$$U(LBP_{p,R})=U(LBP_{p,R}) = |s(g_{p-1-gc}) - s(g_0 - gc)| + \sum_{p=1}^{p-1}|s(gp - gc) - s(g_{p-1-gc})| \text{ -----(12)}$$

riu2 means that rotation invariant uniform patterns. It has U value of at most 2. P+1 is uniform binary patterns occur in a circularly symmetric neighbor set of P pixels. Equation no-11 assigns a unique label to each of them corresponding to the number of 1 bits in the pattern (0 to P). Nonuniform patterns are grouped under miscellaneous label (P +1).

LBP selects a 3*3 window and to apply threshold function over this window. The difference of pixel value of center pixel with neighbor pixels in 3*3 windows is taken pixel by pixel. The threshold function is applied on difference values. If this difference value is greater than or equal to 0 than. It takes value 1 for the pixel, 0 otherwise. Output value of each neighboring pixel is added up (s). Figure3.2 shows example of LBP texture model.

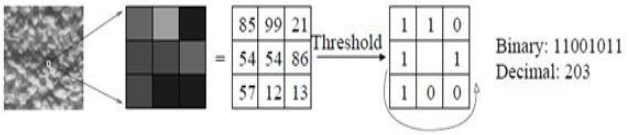


Figure 3.2: LBP texture model

3.2 Color Feature Extraction

Color feature of target object are extracted in form of R, G, B color space. Each subspace of the RGB color space R, G, B, is divided into equal k-intervals. Each interval is called a bin. The number of bins feature space is mc = K^3. Each bin the probability based on the data of all the pixels in target area is calculated and a histogram is prepared. This histogram is integrated with texture features of objet.

3.3 Joint Color-Texture Histogram

In joint color –texture histogram color information of object is extracted as well as texture information is also extracted. After extracting color and texture histogram a joint color-texture histogram is prepared. Color features of the target object are integrated with texture feature. In integration process first the texture features are estimated. Color information of useful texture feature location is used. This texture feature information and color information is integrated in the joint color-texture histogram.

4. Experimental Results

In this section, representative experiments are performed to illustrate joint color-texture model based on mean shift tracking algorithm. It is the comparison with the mean shift tracking with appearance models M1 and M2. The videos of different scenes, including the one has similar target or background colors. These are used in evaluate the performance of different algorithms. Target is represented using rectangular region in the frame. The first experiment is on a video sequence of table tennis playing with 58 frames. Spatial resolution of table tennis video is 240x352.Input of target model is table tennis video file in .avi format. The output of target model is observed on MATLAB version of R2013a software. Target object shows in terms of rectangular window format. Target object is tracked in each frame and location of object in current frame is send to next frame. Target model is calculated by given experimentation method. Output of video file is observed with minimum distance between two mean shift iterations. Tracking object is shown as in the output of target model. Figure5.1 shows tracking results of sequence table tennis playing using method M2.



Figure 5.1: Tracking of table tennis playing video

Table1 enlists the mean error and standard deviation for color histogram (M1) and joint color-texture histogram (M2) methods.

Table 1: Target localization accuracies (mean error and standard deviation) by two methods on the table tennis video

	Method	*Color histogram(M1)*	*Joint color texture Histogram(M2)*
1	Mean error	4.4	2.2
2	Standard deviation	4.2	2.0

Joint color-texture histogram tracks the moving head more reliably and more accurately than only color histogram.

The second experiment is on a video sequence of sliding ball with 121 frames. Spatial resolution of table tennis video is 480x640.The tracking target object a sliding ball. Figure5.2 shows tracking results of sequence sliding ball video using method M2.

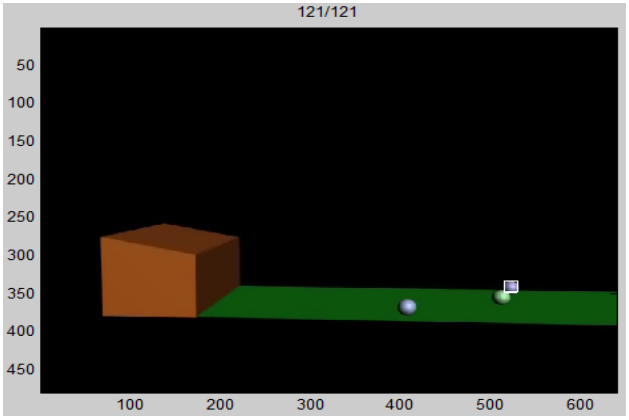


Figure 5.2: Tracking of sliding ball video

Licensed Under Creative Commons Attribution CC BY

Table2 enlists the mean error and standard deviation for color histogram (M1) and joint color-texture histogram (M2) methods.

Table 2: Target localization accuracies (mean error and standard deviation) by two methods on the sliding ball video

	Method	*Color histogram(M1)*	*Joint color texture Histogram(M2)*
1	Mean error	3.6	2.6
2	Standard deviation	5.7	5.6

Input of target model is sliding ball video file in .avi format. The output of target model is observed on MATLAB R2013a software. The target object sliding ball, which is to be tracked on the ground. Joint color-texture histogram method (M2) tracks the sliding ball more accurately than only color histogram (M1). Joint color-texture histogram is better than only color histogram. M2 has much better localization result than M1.

5. Conclusion

Object tracking is done by using features of target object. Feature extraction color and texture features of the target object are used. Combination of both color feature and texture feature is enhancing feature extraction quality. For features extraction LBP technique is used. Color aberration is caused by changing light and similar color to the background color. It affects on accurate location of target object. Both color and texture features are extracted by LBP technique. Joint color and LBP texture are used to reduce the computational complexity.

Experimental results indicate that joint color-texture methode (M2) performs much better than original color histogram methode (M1) with fewer iteration numbers. A simulation result shows that accuracy of joint color-texture histogram (M2) is better than only color histogram (M1) methode. It shows that joint color-texture histogram (M1) outperforms compared to color histogram (M1).This method is improved greatly the tracking accuracy with fewer mean shift iterations than standard mean shift tracking.

References

[1] D. Comaniciu, V. Ramesh and P. Meer, Kernel-based object tracking, *IEEE Trans. Patt. Anal. Mach. Intell.* 25(5) (2003) 564–575
[2] I. Haritaoglu and M. Flickner, Detection and tracking of shopping groups in stores, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Kauai, Hawaii, 2001, pp. 431–438.
[3] K. Nummiaro, E. Koller-Meier and L. V. Gool, An adaptive color-based particle filter, *Imag. Vis. Comput.* 21(1) (2003) 99–110.
[4] C. Yang, D. Ramani and L. Davis, Efficient mean-shift tracking via a new similiarity measure, *Proc. IEEE Conf. Computer Vision and Pattern Recognition* I (2005) 176– 183.
[5] A. Yilmaz, O. Javed and M. Shah, Object tracking: A survey, *ACM Comput. Surv.* 38(4) (2006).
[6] G. Bradski, Computer vision face tracking for use in a perceptual user interface, *Intel Technol. J.* 2(2) (1998) 12–21.
[7] Q. A. Nguyen, A. Robles-Kelly and C. Shen, Enhanced kernel-based tracking for monochromatic and thermographic video, *Proc. IEEE Conf. Video and Signal Based Surveillance* (2006), pp. 28–33.
[8] C. C. Gotlieb and H. E. Kreyszig, Texture descriptors based on co-occurrence matrices, Comput. Vis. Graph. Imag. Process. 51(1) (1990) 70–86.
[9] M. Pietik¨ainen, T. Ojala and Z. Xu, Rotation-invariant texture classification using feature distributions, Patt. Recogn. 33(1) (2000) 43–52.
[10] M. Sonka, V. Hlavac and R. Boyle, Image Processing, Analysis and Computer Vision, 3rd ed. (Thomson, 2007).
[11] T. Ojala, M. Pietik¨ainen and T. M¨aenpa¨a, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Patt. Anal. Mach. Intell. 24(7) (2002) 971–987.
[12] T. Ojala, K. Valkealahti, E. Oja and M. Pietik¨ainen, Texture discrimination with multi-dimensional distributions of signed gray level differences, Patt. Recogn. 34(3) (2001) 727–739.

Licensed Under Creative Commons Attribution CC BY