Real Time Corrected Background Weighted Histogram Based Mean Shift Tracking

Michael Daniel Nachipyangu

College of Electronics Engineering, Tianjin University of Technology and Education, Tianjin, P.R China

Abstract: ***The real time target tracking provides the capability for the user to choose object s/he want to track as it moves in real time. The Corrected Background Weighted Histogram (CBWH) based mean shift Tracking is used on designing the algorithm for tracking. The algorithm used basis on reducing the effects caused by the changes on background which in turn affect the tracking efficiency. Whenever the system starts the target model is made and the background models too, when there is a change on background of the target object the system updates it accordingly. This algorithm also use color as its feature on tracking. The system is well designed with the Graphical user interface enabling even the user with less knowledge being able to operate the system.***

Keywords: Corrected Background Weighted Histogram (CBWH), Background updating model, mean shift, tracking and MatLab

1. Introduction

The advancement of computer technology has brought so many changes on the way of doing things especially monitoring some activities or access control. In recent years Image processing has tried to make the camera as the human eye and contribute a lot on access control systems. This paper is going to discuss the use of image processing on object tracking. Object tracking has two categories, single object and multi object tracking this paper will only discuss single object tracking. The object will be tracked on real time with user given capability of choosing an object to be tracked. Mean Shift tracking algorithm is used to design the tracking algorithm. Corrected background weighted Histogram (CBWH) is used to reduce the effect of background which normally changes as the target object moves which in turn reduce the efficiency of the tracking algorithm. To encounter that problem the background is updated every time whenever necessary during the tracking process. Well designed Graphical user interface is also designed to make the proposed system be friendlier.

Figure 1: Tracking the ball as the target object

2. Related Work

Tracking has been the hot topic for so long since the emerging of image processing but it has been done in different ways. The main categories of tracking are single object tracking and multi object tracking, within those two there are real time tracking and tracking from the recorded video file. Many researches concentrate on single object tracked from a video file. Selection of features to be used for tracking an object has made a wide difference on designing different algorithms for tracking. Sometimes two features can be used to ensure good efficiency of the system like Object Tracking using Joint Color-Texture Histogram [13]. Generally the type of application and environment where the system will be used determine the selection of feature to use on tracking also whether the system to track single or multi objects.

3. Tracking

Object Tracking is one of the important fields in Image Processing as its stake in Computer Vision. The innovation of high-powered computer, high quality camera and the need of automated video analysis have brought much attention to the Object Tracking algorithms. There are three key steps in video analysis: detection of interesting moving objects, tracking of such objects from frame to frame, later depending on the application the tracks can be analyzed to recognize their behavior. In this paper tracking has been used just to track selected objects.

4. Mean Shift

Mean shift is a non parametric feature space analysis technique for locating the maxima of a density function, a so-called mode-seeking algorithm. Mean shift considers feature space as a empirical probability density function. If the input is a set of points then Mean shift considers them as sampled from the underlying probability density function. If dense regions (or clusters) are present in the feature space, then they correspond to the mode (or local maxima) of the probability density function. We can also identify clusters associated with the given mode using Mean Shift. For each data point, Mean shift associates it with the nearby peak of the dataset's probability density function.

For each data point, mean shift defines a window around it and computes the mean of the data point. Then it shifts the center of the window to the mean and repeats the algorithm till it converges After each iteration, we can consider that the window shifts to a denser region of the dataset. The main applications of Mean Shift are on Cluster Analysis for Computer Vision and Image processing.

4.1 Corrected Background-Weighted Histogram based Mean Shift Tracking

Tracking is one of the applications of the mean shift. In the mean shift tracking algorithm, the color histogram is used to represent the target because of its robustness to scaling, rotation and partial occlusion. However, the mean shift algorithm is prone to local minima when some of the target features present in the background. Thus, Comaniciu et al. further proposed the background-weighted histogram (BWH) to decrease background interference in target representation. The strategy of BWH is to derive a simple representation of the background features and use it to select the salient components from the target model and target candidate model. More specifically, BWH attempts to decrease the probability of prominent background features in the target model and candidate model and thus reduce the background's interference in target localization. Later Jifeng Ning, Lei Zhang, David Zhang and Chengke Wu realized that the BWH is just the same as normal back mean shift tracking and they proposed the Corrected Background Weighted Histogram (CBWH) which is the one used for experiments in this thesis.

An important advantage of the proposed CBWH method is that it can work robustly even if the target model contains much background information. Thus it reduces greatly the sensitivity of mean shift tracking to target initialization. In the experiments, we can see that even when the initial target is not well selected, the proposed CBWH algorithm can still correctly track the object, which is hard to achieve by the usual target representation.

4.2 Target Representation

In object tracking, a target is usually defined as a rectangle or an ellipsoidal region in the frame and the color histogram is used to represent the target. Donated by $\{x_i^*\}_{i=1\ldots n}$ the normalized pixels in the target region, which has n pixels. The probability of a feature u, which is actually one of the m color histogram bins, in the target model is computed as

$$\hat{q} = \{\hat{q}_u\}_{u=1\ldots m}\ ;\ \hat{q}_u = C\sum_{i=1}^{n} k\left(\|x_i^*\|^2\right)\delta[b(x_i^*) - u] \tag{4.1}$$

Where $\hat{q}$ is the target model, $\hat{q}_u$ is the probability of the element of $\hat{q}$, δ is the Kronecker delta function, $b(x_i^*)$ associates the pixel x_i^* to the histogram bin, k(x) is an isotropic kernel profile, and constant C is

$$C = 1/\sum_{i=1}^{n} k\left(\|x_i^*\|^2\right) \tag{4.2}$$

Similarly, the probability of the feature u=1, 2… m in the target candidate model from the target candidate region centered at position y is given by

$$\hat{q}(y) = \{\hat{q}_u(y)\}_{u=1\ldots m}\ ;\ \hat{q}_u(y) = C_h\sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\delta[b(x_i) - u] \tag{4.3}$$

Where $\hat{q}(y)$ is the target candidate model, $\hat{q}_u(y)$ is the probability of the u[th] element of $\hat{q}(y)$, $\{x_i\}_{i=1\ldots n_h}$ are pixels in the target candidate region centered at y, h is the bandwidth and C_h is the normalized constant

$$C_h = 1/\sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \tag{4.4}$$

4.3 Mean Shift Tracking Algorithm

The more important concept on Mean Shift tracking is the computation of an offset from the current location to the new location y_1 in accordance to the mean shift iteration equation. The computation of the offset is the one that determine the accuracy of tracking, the accuracy in tracking is to find the most similar region of the target object in the new frame.

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{y - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{y - x_i}{h}\right\|^2\right)} \tag{4.5}$$

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{q}_u(y)}}\,\delta[b(x_i) - u] \tag{4.6}$$

Where g(x) is the shadow of the kernel profile k(x): g (x) = −k '(x). For the convenience of expression, we denote by;

$$g_i = g\left(\left\|\frac{y - x_i}{h}\right\|^2\right)$$

Thus Eq. (4.5) can be re-written as:

$$y_1 = \sum_{i=1}^{n_h} x_i w_i g_i \Big/ \sum_{i=1}^{n_h} w_i g_i \tag{4.7}$$

With Eq. (4.7), the mean shift tracking algorithm can find the most similar region to the target object in the new frame.

4.4 The Corrected Background-Weighted Histogram

In target tracking, often the background information is included in the detected target region or the selected target region. If the correlation between target and background is high, the localization accuracy of the object will be decreased which also decrease the tracking capability of a particular system. To reduce the interference of salient background features in target localization, a representation model of background features was proposed by Comaniciu et al. [12] to select discriminative features from the target region and the target candidate region. Comaniciu et al scheme was not working well on solving the problem; there is where Jifeng Ning, Lei Zhang, David Zhang and Chengke Wu came with another scheme to solve that problem of background having the same features with the target which is called Corrected Background Weighted Histogram (CBWH). As the name itself is just the correction of the first model which was Background Weighted Histogram (BWH). The background is represented as

$\{\hat{O}_u\}_{u=1\ldots m}$ With $\sum_{i=1}^{m} \hat{O}_u = 1$

Denote by $\hat{O}^*$ the minimal non-zero value in $\{\hat{O}_u\}_{u=1\ldots m}$ the coefficients

$$\{v_u = min(\hat{O}^*/\hat{O}_u, 1)\}_{u=1\ldots m} \tag{4.8}$$

Eq. (4.8) is employed to transform only the target model but not the target candidate model. That is to say, we reduce the prominent background features only in the target model but not in the target candidate model.

From Eq. 4.6;

$$w_i' = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u'}{\hat{p}_u'(y)}}\,\delta[b(x_i) - u] \tag{4.9}$$

Let u ' be the bin index in the feature space which corresponds to point x_i in the candidate region. We have $\delta[b(x_i) - u'] = 1$. So Eq. (4.9) can be simplified as

$$w_i' = \sqrt{\hat{q}_{u'}'/\hat{p}_{u'}(y)} \tag{4.10}$$

But the new target model,

$$\hat{q}_u' = C' v_u \sum_{i=1}^{n} k(\|x_i^*\|^2)\delta[b(x_i^*) - u] \tag{4.11}$$

With the normalization Constant

$$C' = \frac{1}{\sum_{i=1}^{n} k\left(\|x_i^*\|^2\right) \sum_{u=1}^{m} v_u \delta[b(x_i^*) - u]} \tag{4.12}$$

Then the new candidate model is,

$$\hat{p}_u'(y) = C_h' v_u \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\delta[b(x_i) - u] \tag{4.13}$$

Where

$$C_h' = \frac{1}{\sum_{i=1}^{n} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \sum_{u=1}^{m} v_u \delta[b(x_i) - u]}$$

Then,

$$w_i' = \sqrt{\frac{C' v_{u'} \sum_{j=1}^{n} k\left(\|x_j^*\|^2\right)\delta[b(x_j^*) - u']}{C_h' v_{u'} \sum_{j=1}^{n_h} k\left(\left\|\frac{y - x_j}{h}\right\|^2\right)\delta[b(x_j) - u']}} \tag{4.14}$$

By removing the common factor $v_{u'}$ from the numerator and denominator and substituting the normalization factors C and C_h into the above equation, we have

$$w_i' = \sqrt{\frac{CC_h}{CC_h} \frac{C' \sum_{i=1}^{n} k\left(\|x_i^*\|^2\right)\delta[b(x_i^*) - u']}{C_h' \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\delta[b(x_i) - u']}} = \sqrt{\frac{C'C_h}{CC_h'}} \cdot \sqrt{\frac{\hat{q}_{u'}'}{\hat{p}_{u'}'}} = \sqrt{\frac{C'C_h}{CC_h'}}\, w_i$$

We define a new weight formula

$$w_i'' = \sqrt{\hat{q}_{u'}'/\hat{p}_{u'}(y)}$$

We can also obtain,

$$w_i'' = \sqrt{C'/C} \,.\, \sqrt{v_{u'}} \,.\, w_i \tag{4.15}$$

Since $\sqrt{C'/C}$ is a constant scaling factor, it has no influence on the mean shift tracking process. We can omit it and simplify Eq. (4.15) as

$$w_i^{''} = \sqrt{v_{u'}} \cdot w_i \qquad (4.16)$$

Eq. (4.16) clearly reflects the relationship between the weight calculated by using the usual target representation (i.e. w_i) and the weight calculated by exploiting the background information (i.e. $w_i^{''}$). If the color of point i in the background region is prominent, the corresponding value of $v_{u'}$ is small. Hence in Eq. (4.16) this point's weight is decreased and its relevance for target localization is reduced. This will then speed up mean shift's convergence towards the salient features of the target. Note that if we do not use the background information $v_{u'}$ will be 1 and $w_i^{''}$ will degrade to w_i with the usual target representation.

4.5 Background Updating Model

The background color model $\{\hat{O}_u\}_{u=1...m}$ is employed and initialized at the beginning of tracking. However, in the tracking process the background will often change due to the variations of illumination, viewpoint, occlusion and scene content, etc. If the original background color model is still used without updating, the tracking accuracy may be reduced because the current background may be very different from the previous background model. Therefore, it is necessary to dynamically update the background model for the good tracking performance.

The updating model of the background start with the background feature $\{\hat{O}'_u\}_{u=1...m}$ and $\{v'_u\}_{u=1...m}$ in the current frame are calculated. Bhattacharyya similarity between $\{\hat{O}'_u\}_{u=1...m}$ and the old background model $\{\hat{O}_u\}_{u=1...m}$ computed by

$$\rho = \sum_{u=1}^{m} \sqrt{\hat{O}'_u \hat{O}_u} \qquad (4.17)$$

If ρ is smaller than a threshold, this implies that there are considerable changes in the background, and then we update $\{\hat{O}_u\}_{u=1...m}$ by $\{\hat{O}'_u\}_{u=1...m}$ and update $\{v_u\}_{u=1...m}$ by $\{v'_u\}_{u=1...m}$. The transformed target model $\hat{q}'_u$ is then computed by Eq. (4.8) using $\{v'_u\}_{u=1...m}$ Otherwise, the background model is not updated.

5. Experimental Results and Discussion

The experiment was done with the use of the PC installed MatLab with Logitech webcam C170. The designed interface with MatLab GUI is used to make the system friendlier even with the people with less knowledge about Image Processing.

Figure 2: Object tracking system interface

The system has the interface as seen above which has capability of previewing the video, process which enable the user to select the object to be tracked and track it clear all video and exit the system. It also shows the status on each process on the status bar.

6. Conclusion and Recommendations

The system successfully designed, tested and work as desired as the target object can be selected and being tracked in real time. The following are the recommendations for the future work;

Using more than one feature will increase the robustness and accuracy on tracking since color bands are sensitive to illumination variations and reflectance of the target object. The use of more advanced and quality camera will give the system ability to track even distant objects compared to the capability it has now.

References

[1] C. T. Johnston, K. T. Gribbon and D. G. Bailey. "FPGA based Remote Object Tracking for Real-time Control", 1st International Conference on Sensing Technology, Palmerston North, New Zealand, 2005.

[2] C. Bibby and I. "Reid. Robust real-time visual tracking using pixel-wise posteriors" In ECCV, pages 831–844, 2008.

[3] Alper Yilmaz, Omar Javed and Mubarak Shah "Object Tracking Survey" ACM Computing Surveys, Vol. 38, No. 4, Article 13, December 2006

[4] Saad M. Khan and Mubarak Shah "A multiview approach to tracking people in crowded scenes using a planar homography constraint" European Conference on Computer Vision, 2006.

[5] Shah, H., and Morrell, D "An adaptive zoom algorithm for tracking targets using pan-tilt-zoom cameras" Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Canada (2004).

[6] E. Meijering, O. Dzyubachyk, and I. Smal “Methods for cell and particle tracking. Methods in Enzymology” Live Cell Imaging, 2012.
[7] K. Fragkiadaki and J. Shi “Detection free tracking: Exploiting motion and topology for segmenting and tracking under entanglement” CVPR, pages 2073–2080, June 2011
[8] R. Y. Wang and J. Popovic. “Real-time hand-tracking with a color glove” ACM TOG (Proc. SIGGRAPH), (3):63:163:8, July 2009.
[9] I. Leichter. Mean Shift Trackers with Cross-Bin Metrics. IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI), 34(4):695-706, 2012.
[10] Fei Qi, Xiaowei Song, and Guangming Shi, "Lda based color information fusion for visual objects tracking.," in IEEE International Conference on Image Processing, 2009, pp. 2201-2204.
[11] M. Grabner, H. Grabner, and H. Bischof, “Learning features for tracking” in In Proc. IEEE Conference on Computing Vision and Pattern Recognition, 2007.
[12] Comaniciu D., Ramesh V., and Meer P.: ‘Real-Time Tracking of Non-Rigid Objects Using Mean Shift’. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Hilton Head, SC, USA, June, 2000, pp. 142-149.
[13] J. Ning, L. Zhang, D. Zhang and C. Wu, “Robust Object Tracking using Joint Color-Texture Histogram,” International Journal of Pattern Recognition and Artificial Intelligence, vol. 23, No. 7 (2009) 1245–1263.

Author Profile

Michael D. Nachipyangu was awarded Bsc in Electronics and Communication (Hons) from the University Of Dar Es Salaam Tanzania the year 2011, He then served as RF Engineer at Startel (T) Limited (Raha) from August 2011 to September 2012 which is one best Internet Service Provider company in Tanzania, September 2012 he joined Tianjin University of Technology and Education for his Masters of Engineering in Signal and Information Processing with his research oriented in Digital Video and Image Processing.