ICIC Express Letters
Part B: Applications
Volume **8**, Number **3**, March **2017**

 ISSN 2185-2766
pp. **609–614**

A MULTI-CUE INTEGRATED ADABOOST BASED TARGET TRACKING METHOD

Longkui Jiang[1] and Yuru Wang[2,3,*]

[1]School of Information Engineering
Jilin Business and Technology College
No. 1666, Kalunhu Street, Jiutai Economic Development Zone, Changchun 130507, P. R. China
longkuij@163.com

[2]School of Computer Science and Information Technology
[3]Key Laboratory of Intelligent Information Processing of Jilin Province
Northeast Normal University
No. 5268, Renmin Street, Changchun 130024, P. R. China
*Corresponding author: yuru_765@163.com

Received August 2016; accepted November 2016

Abstract. *In order to handle the difficulties of complex tracking conditions in real tracking problems, this paper proposed a novel Adaboost based tracking algorithm. In the proposed method, visual tracking is resolved as a two-class classification problem between fore-ground (the target) and its background. And the classification is resolved by designing an adaptive classifier. In detail, an on-line updated Adaboost classifier is designed to overcome the limitation of single classifier in tracking accuracy and robustness. Specially, the problem of sample's label is considered, and the confidence is assigned to each input sample as its class, so as to realize more accurate classification. In addition, multiple salient features are integrated to represent target's appearance and provided to the classifier, so as to realize the robustness. The experimental results demonstrate the robustness and accuracy of the proposed Adaboost classifier in dealing with tracking problems.*
Keywords: Computer vision, Visual tracking, Adaboost, Multi-cue

1. **Introduction.** As for target's movement in complex environments, the main difficulty comes from how to differentiate the target from its backgrounds. In the visual tracking problem, how to design a good tracking method so as to search for the target region is a key problem. In this paper, we will make major research on the tracking algorithm.

Currently, the existing visual tracking algorithms can be classified into two classes. The one is the target's pattern searching based method, which looks the tracking problem as local pattern's matching and optimizing [1]. The other is the state estimation based method, which learns the tracking conditions and evaluates various predictions in the tracking procedure [2]. Based on the theoretical analysis on these two methods, we found that the common shortcoming of them is that they consider much on the target and ignore the backgrounds when locating the target. So the existing methods cannot give a good enough solution to tracking problem, and always fail or track with large errors. Therefore, some researchers proposed to design specific classifiers to deal with tracking problem. Collins et al. [3] employed an on-line feature selection method to differentiate the color space set of the foreground and its backgrounds. Tian et al. [4] proposed an on-line ensemble SVM tracker to handle drift and whole or long-time occlusion. Penne et al. proposed a Modular Ensemble Tracker [5]. Large Margin Classifier [6] is introduced into tracking problem and embedded in an ensemble tracking framework. Besides, Zhu et al. [7] proposed a Clustering Ensemble Tracking method . In the method, the object-part predictor, the parameter clustered predictor and the feature clustered predictor are integrated together. We can see from the existing methods, that the design of classifier is an important issue. And in most of the existing classifiers, the features are fixed as HOG or HSV. Therefore, developing a classifier integrated with multiple features is supposed to show better performance. In our previous work, a particle filter guided SVM based method [8] is proposed, and the SVM classifier is demonstrated to be effective in locating target region accurately. However, the SVM should be trained ahead of time, so the tracking effectiveness cannot be guaranteed.

Above all, the two-class based classifier proposed a robust and accurate solution for the target tracking problem. However, there are still two aspects of difficulties: in real tracking applications, and on the one hand, due to the complex change of tracking environments, image noises and occlusions, it is hard for a single classifier to give an accurate classifying for the target from the backgrounds pixels. On the other hand, due to target's changes caused by environments and the target itself, a fixed classifier cannot adapt to the changes. With the consideration of the above factors, this paper proposed an ensemble classifier based tracking method. The novelty of this paper mainly lies in the two-fold. Firstly, in order to adapt to the complex changes in tracking environments, an on-line updating Adaboost classifier is designed. Especially, in order to realize more accurate classifying, the problem of samples' class label is considered and is assigned as samples' confidence values. Secondly, in order to make the method robust to target's changeable appearance, multiple features are extracted and provided to the classifier.

The rest of the paper is organized as follows. Section 2 gives the framework of the proposed semi-supervised method. Section 3 details the employed Adaboost classifier. In Section 4, experiments, comparisons and results are given. And the proposed method is concluded in Section 5.

2. **System Framework.** Focusing on the tracking difficulty of complex tracking conditions, this paper employs Adaboost classifier to realize accurate tracking. And the system framework is shown in Figure 1.

In order to realize an accurate tracking, developing effective tracking algorithm is not sufficient, and to develop target's robust model is also necessary. Therefore, this paper employs a multi-cue integration method [9]. The employed features are salient visual

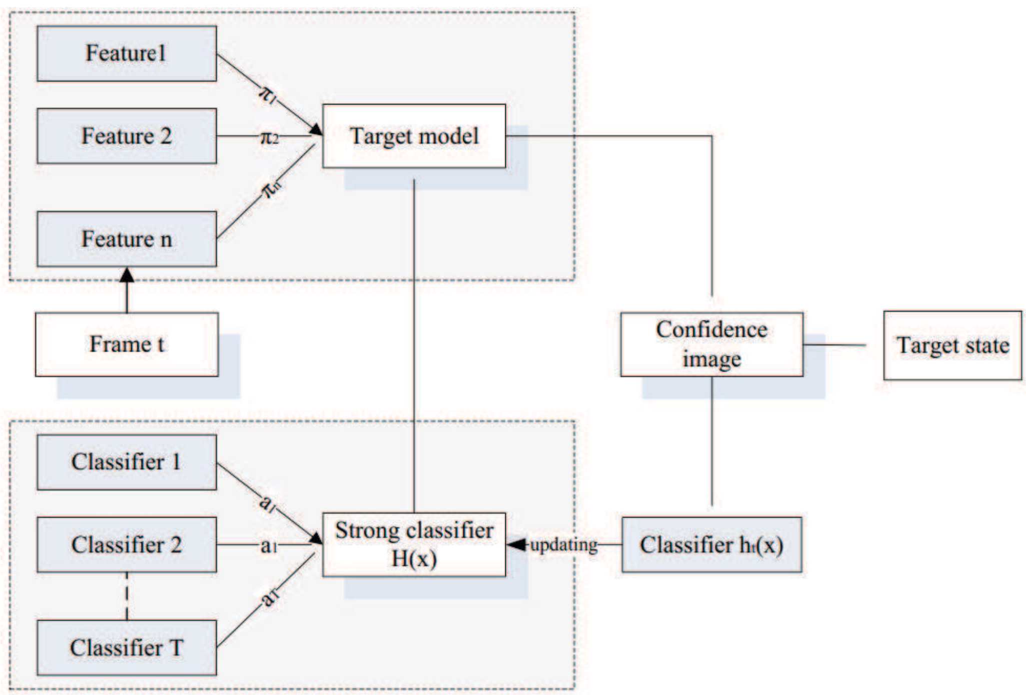

FIGURE 1. The tracking flowchart at time t in the video sequence

features as color, texture, and contour. Specially, the multi-cue integrated target model is not employed to represent the samples of classifier, but to describe the estimated target region. In such a way, the space information will be maintained, and the time continuity is guaranteed in model's online updating. The employed Adaboost classifier ensembles multiple weak ones, and the weighted minimum mean square regression method is employed.

3. **Multi-Cue Integrated Adaboost.** In order to adapt to the complex change of tracking environment, with the consideration of tracking backgrounds, this paper designed an online updating Adaboost classifier.

3.1. **Adaboost classifier.** Adaboost method constructs a strong classifier by training several weak ones, and the performance of the strong one is much better than the weak ones. The ensemble classifier realizes tracking by classifying the target and backgrounds at each frame and updating each weak classifier. And its working procedure is shown in Figure 2, where the samples in target and backgrounds regions are positive and negative, respectively. These samples are employed to train each weak classifier, and Adaboost method is employed to construct a strong classifier.

In this figure, the rectangle with solid lines is the target region, and the region between dashed and solid lines is the background. (a) At frame $t-1$, classify the estimated region with the classifier (the two solid lines represent weak classifiers). (b) At frame t, train a new weak classifier (the dashed line) and add it to the strong one.

For a new frame, the strong classifier will calculate out a confidence value to each pixel in the estimated target region, and then the target will be located based on the confidence

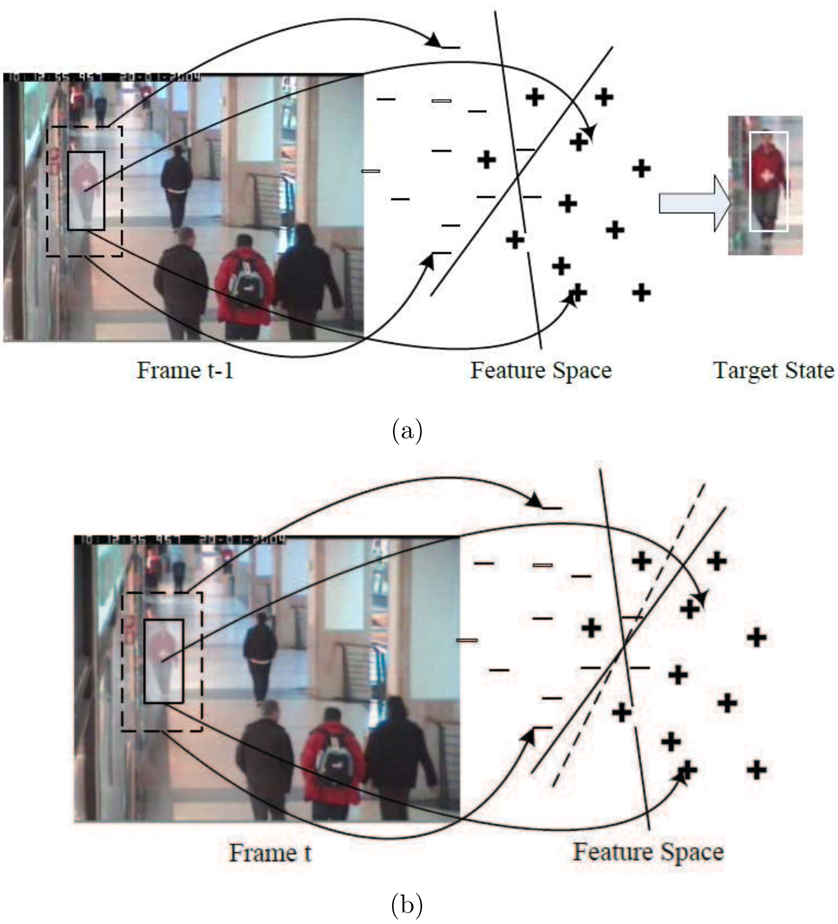

Figure 2. The updating and classifying of Adaboost

map. Once the detection is finished, a new weak classifier is trained, and added to the strong classifier.

$$H(x) = \sum\nolimits_{t=1}^{T} a_t h_t(x) \tag{1}$$

where a_t is the weight of the t-th weak classier, and is set to be the number of the wrongly classified samples. At each frame, according to this weight, we choose K classifiers and construct a new strong classifier. Over the whole video, the above procedure is repeated. In the proposed Adaboost method, at each frame, the weak classifiers can be deleted or added, so as to adapt to the change of target and backgrounds. After some frames, a strong classifier is constructed, and it guarantees the accuracy.

3.2. **Samples construction.** The classifier is not used to judge one region, but to judge some pixel by several classifiers ensemble. The classifier works in the pixel level, and each sample provided to the classifier is defined as the feature of some pixels. With each pixel in the detection region represented by multi-cue, the sample is the feature vector of its $n \times n$ window. This sample representation method defined in pixel level shows the following advantages. On the one hand, it maintains the space feature of each pixel's neighbor region. On the other hand, it is robust to the change of target appearance.

3.3. **Weak classifier.** There are many choices for weak classifier of Adaboost method. The sample set $\{x_i, y_i\}_{i=1}^{N}$ shows the N samples and their labels. And each sample is a feature vector with d dimensions, $x_i \in R^d$, $y_i \in \{-1, +1\}$. Every classifier can not only give the label of each pixel, but also calculate its distance to classifying sphere $y_i \in \{-1, +1\}$. If the confidence value is ignored and the classifier only classifies the sample to be negative or positive, then much usable information will be ignored. The weighted least square regression method is employed as the weak classifier, and the confidence value is used as the labels to train the classifier.

4. **Experimental Results.** In order to test the proposed Adaboost method, we implemented the proposed method in VC++ 6.0 and analyzed the results in MATLAB 2008. It is difficult for the existing methods to deal with the difficulty of complex tracking conditions, such as environmental changes and target's appearance changes. The proposed method is tested on two kinds of videos, target in complex environment, and pedestrian with occlusion. And the comparison experiment is implemented to illustrate the effectiveness and accuracy of the proposed algorithm.

4.1. **Test on the video with complex environment.** In order to test the effectiveness of our method in dealing with complex environments, it is tested on such videos as shown in Figure 3. In the tested video (with 449 frames, of size 352×240 pixels), a man rides on a bicycle in the street, where there exist a lot of objects with similar appearance to the target, and the environment changes a lot. It can be seen from the key video frames, that the tracking environment is so complex that it is hard even for people to identify the rider from the backgrounds. As seen from the tracking results, our method provides a good solution to such a tracking problem. Though the backgrounds show similar color distribution or textures, the proposed method can handle it. The success relies on the feature extracting method of each sample. In the proposed method, the samples are represented as each pixel's neighbor vector. However, there are also some failure examples (Frame 425), that is because of the occurrence of a similar rider. After a few frames, the tracker is corrected.

Figure 3. The tracking results of a pedestrian in a complex street

4.2. **Comparison experiment on videos with occlusion.** In order to test the performance of the proposed methods on the occlusion problem, and also to further demonstrate the effectiveness of the proposed method, they are tested on a video sequence with occlusions. As shown in Figure 4 (where the size of each frame is 320×240 pixels), a pedestrian is occluded by a tree with complex appearance. When occlusion happens, only a small part of the pedestrian can be observed. The proposed method is compared with the method proposed in our previous paper [5], where a particle filter guided SVM method is employed. As can be seen from the tracking result figures, both the two methods can handle the occlusion problem, as shown in frame 135, but the proposed method performs more accurate tracking. The winning relies on that the Adaboost method is online updated, so the classifier can adapt to the tracking conditions, even when occlusions happen. Besides, the proposed method works more effectively than the other method (3.1 seconds versus 15.5 seconds at average), because the SVM updating is realized by training all the samples. So the proposed method realizes an effective and accurate tracking.

Figure 4. The solid rectangle is the target regions located by the proposed method, and the dashed one is the ones located by the method in [5]

5. **Conclusions.** The complex environment is a major difficulty in visual tracking. In order to realize an accurate tracking, this paper proposed a novel multi-cue based Adaboost method. Specially, a multiple features model is employed to represent the target, and a pixel level classification is realized by extracting each pixel's confidence as label. Through experiments, the proposed method is demonstrated to be effective and robust against complex background changes and occlusion. In our future work, the effeciency for real-time tracking is the main direction. Specifically, we will try to adapt the design of weak classifier and the integration method, and realize a more effective and accurate tracking.

Acknowledgments. This work is supported by the National Natural Science Foundation of China under Grant No. 61300099, the China Postdoctoral Science Foundation Funded Project under Grant No. 2015M570261, the Jilin Province Postdoctoral Science Foundation funded project, the Science and Technology Development Plan of Jilin Province under Grant No. 20170101144JC, and the Open Fund of China Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education under Grant No. 93K172016K14.

REFERENCES

[1] D. Comaniciu, V. Ramesh and P. Meer, Kernel-based object tracking, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.25, no.2, pp.564-577, 2003.
[2] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*, Academic Press, 1988.
[3] T. R. Collins, Y. Liu and M. Leordeanu, Online selection of discriminative tracking features, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.27, no.10, pp.1631-1643, 2005.
[4] M. Tian, W. Zhang and F. Liu, On-line ensemble SVM for robust object tracking, *Asian Conference on Computer Vision*, vol.4843, pp.355-364, 2007.
[5] T. Penne, C. Tilmant, T. Chateau and V. Barra, Modular ensemble tracking, *International Conference on Image Processing Theory Tools and Applications*, pp.363-368, 2010.
[6] Y. Wang, Q. Liu, M. Yin and S. Wang, Large margin classifier based ensemble tracking, *Journal of Electronical Imaging*, vol.25, no.4, 2016.
[7] G. Zhu, J. Wang and H. Lu, Clustering ensemble tracking, *Lecture Notes in Computer Science*, vol.9007, pp.382-396, 2015.
[8] Y. Wang, J. Zhang and T. Gao, Particle filter guided SVM based visual tracking, *ICIC Express Letters, Part B: Applications*, vol.2, no.5, pp.1169-1174, 2011.
[9] Y. Wang, X. Tang and Q. Cui, Dynamic appearance model for particle filter based visual tracking, *Pattern Recognition*, vol.45, pp.4510-4523, 2012.