International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.7, No.1 (2014), pp.181-192
http://dx.doi.org/10.14257/ijsip.2014.7.1.17

Multi-Target Adaptive On-line Tracking based on WIHM

Wenhui Li, Hongyin Ni, Ying Wang, Peixun Liu and Yuchao Zhou

College of Computer Science and Technology, Jilin University,
Changchun, 130012, China
nihongyin@163.com

Abstract

Multi-target tracking based on on-line boosting is a significant technique in computer vision. However, it is very difficult to select the optimal classifier in on-line learning process since tracking often relies on an assumption that the appearance model of target is fixed. This would directly lead to a decline in the performance of on-line boosting. In this paper, we presents a novel on-line multi-target tracking framework based on Weighted Incremental Histogram Model (called WIHM), which can be applied in some static or dynamic scenarios. First, we propose a novel method—WIHM, which is employed to obtain the optimal size of a tracked object. Second, a new update scheme is used to reach local optimum tracking appearance model with high possibility and accuracy. Third, based on the above works, a multi-target tracking framework is proposed to track multi-targets simultaneously. With the appearance model of targets changed continually, our represented approach can track these targets more powerful, especially in dealing with camera motions. Experimental results show the effectiveness and robustness of our method.

Keywords: *Target tracking, On-line boosting, WIHM, Bhattacharyya distance*

1. Introduction

Visual target tracking approaches have achieved measurable success in the domain of computer vision. Examples include video surveillance [1], assistant driving [2] or human action understanding [3]. However, on the one hand, most existing tracking methods often track each target separately, but have difficulties to deal with large number of targets in complex scenes. On the other hand, it is a tremendous challenge to design effective visual tracking algorithms, which can deal with the inevitable changes (*e.g.*, illuminations, shapes, viewpoints and occlusions) that can occur in dynamic scenes [4].

In recent years, Boosting such as AdaBoost, proposed by Y. Freund and R. Schapire [5], is an excellent off-line learning method that has been further researched by many researchers[6, 7]. In 2001, N. Oza proposes on-line boosting that updates weights of weak classifiers to adjust the target object shape [8]. Meanwhile, he has proved that on-line boosting and off-line boosting have the same performance in detection task if they are given the same training set [9]. Recent work on the on-line boosting and tracking (detection-based) seems a promising direction. Therefore, many researchers have been focusing on performing tracking tasks by means of on-line boosting. H. Grabner [10] applies an improved on-line boosting to select several weak classifiers that are trained in an off-line boosting method for target tracking. B. Wu [11] proposes an unsupervised, incremental learning approach based on on-line boosting to improve the performance of boosted object detectors, which are learned from a small amount of labeled data and have moderate accuracy. In [12], an ensemble of weak classifiers is trained on-line to distinguish between the target and the background, which is combined

ISSN: 2005-4254 IJSIP

into a strong classifier using AdaBoost. T. Parag [13] proposes a novel on-line boosting method where the form of the weak classifiers is modified to deal with scene changes. Instead of replacement [12], the parameters of the weak classifiers are altered in accordance with the new data subset presented to the on-line boosting process at each time step. Cheng-Hao Kuo [14] presents an approach for on-line learning of discriminative appearance models (OLDAMs) for robust multi-target tracking in a crowded scene from a single camera. When OLDAMs were integrated into a hierarchical association framework, the tracking accuracy has been improved significantly. Zhiquan Qi [15] proposes an improved on-line multiple instance learning based on boosting algorithm (called OMILBoost) and successfully applied it to on-line object detection problem with partial occlusions. Based on Cheng-Hao Kuo's work, a novel on-line multiple instance gradient feature selection models are established by Yuan Xie [16]. The proposed method not only allows us to achieve an efficient way of updating the discriminative feature set using gradient feature selection scheme, but also could overcome drifting problem to some extent with the help of MIL. Unlike most previous approaches which only focus on producing discriminative motion and appearance models for all targets [4, 10, 12, 14-16], Bo Yang [17,18] further consider discriminative features for distinguishing difficult pairs of targets. The tracking problem is formulated using an on-line-learned CRF model, and transformed into an energy minimization problem. Quan Miao [19] presents a robust feature-based tracking scheme by using adaptive classifiers to match the detected keypoints in consecutive frames. Experimental results show better performance on some challenging video sequences.

Although much progress has been made in developing methods for producing discriminative motion or appearance models, there has been comparatively less work on real-time changing the appearance models in on-line boosting framework, which are key factors for improving tracking performance. In this paper, to overcome these problems, firstly, we design a weighted incremental histogram model for on-line boosting framework, which can selects the real correct tracking patch with high possibility and accuracy. Then the weak classifiers produced by the on-line boosting can be updated by our dynamic update scheme effectively. At last, we evaluate our approach on several public data sets, and show significant improvements compared with some state-of-art methods.

The outline of the paper is as follows: First, in Section 2, we introduce the overview of our framework. Next, our proposed weighted incremental histogram model for on-line boosting framework is given in Section 3. The experimental results are shown in Section 4. The conclusion of this paper is given in Section 5.

2. Overview of our Approach

The framework of our multi-targets tracking is shown in Figure 1. Similar to [4, 19], the core idea of our framework is also to transform the tracking problem into a binary classification task and to achieve robustness by updating the current weak classifier of the target continuously. Given a video input, we first detect targets in each frame by a pre-trained off-line detector [4]. Then detection responses are linked into some selectors by the on-line learning approach separately. Meanwhile, corresponding gray-level histograms of tracked targets are obtained in our method. Thirdly, by constructing the confidence map and computing Bhattacharyya distance of the previous and current histograms, which belong to the same target, we can obtain the tracked target's local optimum of position and size. At last, the selectors of tracked targets are updated by re-learning for each sliding window. There are two main contributions in this paper: One is the WIHM for computing the optimal size of a tracked object in continual frames, which are necessarily capable of distinguish difficult targets, the other is the proposed on-line multi-target tracking framework based on WIHM.

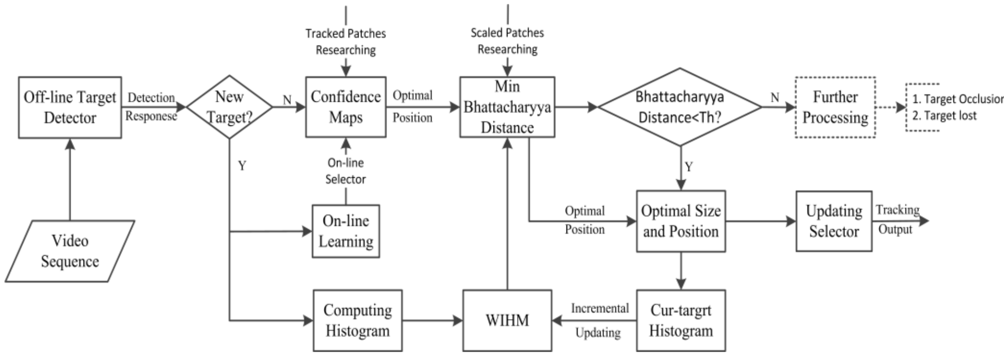

Figure 1. The Framework of our Approach

3. WIHM

Just like the multi-target tracking framework shown in Figure 1, we need to compare the degree of similarity between different scaled rectangular patches, and thus obtain the optimal patch of tracked target. Currently, the histogram model is the most commonly used to generate powerful appearance model. Such as MeanShift [20] and CamShift [21] algorithms. However, Traditional histogram model only considers target's fixed statistical information, and neglects its changing appearance constantly. Therefore, changes of illumination, appearance and partial occlusions will lead to not-stable tracking. It is necessary to real-time update histogram model for avoiding these problems.

Given a rectangular area R in frame I, the histogram model R can be denoted by:

$$H_R = \{h_k^R \mid h_k^R = \frac{1}{\chi} \sum_{(i,j)\in I} I\{L\times k/B \le V_{(i,j)} \le L\times(k+1)/B\}, k=0,...,B-1\}$$

Where k reflects k-th gray-level, L corresponds to the maximum value of selected gray space, B reflects the number of bins, $V_{(i,j)}$ is the value of I at point (i,j) and χ is the normalization factor. K. Nummiaro [22] has proposed a simple linear update method, which can be formalized by:

$$H_R^{t+1} = (1-\lambda)H_R^t + \lambda\hat{H}_R^t \tag{1}$$

Where λ corresponds to the parameter of update-rate and $\hat{H}_R^t$ is the estimate of R at time t. This update process has two problems: one is the update-rate is fixed value, which cannot be modified adaptively with the change of time or environment condition. The other is the linear update method may lose the ability to distinguish partially occluded targets when consecutive error estimates are updated to the histogram model of target.

To overcome these problems above and build a strong appearance model, the initial histogram model of the tracked target can be considered as an important factor. The update of the target model is implemented by :

$$H_R^{t+1} = (1-\alpha-\beta)H_R^t + \alpha\hat{H}_R^t + \beta H_R^0 \tag{2}$$

Where H_R^0 is the initial histogram model of R, α and β are the parameters of update rate, satisfying the following condition:

$$0 \le \alpha+\beta \le 1, 0 \le \alpha \le 1, 0 \le \beta \le 1$$

The parameters α and β are calculated in accordance with Algorithm 1.

Algorithm 1 Computing parameters α and β

Input: H_R^0, H_R^t, $\hat{H}_R^t$ and regulatory factor θ

1: for $k = 0$ to $B-1$ do

2: $S_1(k) = \left| H_R^t(k) - H_R^0(k) \right|$

3: $S_2(k) = \left| \hat{H}_R^t(k) - H_R^0(k) \right|$

4: $M_{\max} = \dfrac{\max(\{H_R^t(k), H_R^0(k), \hat{H}_R^t(k)\})}{H_R^t(k) + H_R^0(k) + \hat{H}_R^t(k)}$

5: $M_{\min} = \dfrac{\min(\{H_R^t(k), H_R^0(k), \hat{H}_R^t(k)\})}{H_R^t(k) + H_R^0(k) + \hat{H}_R^t(k)}$

6: if $S_1(k) > S_2(k)$

7: $\alpha(k) = \theta \cdot M_{\max}, \beta(k) = \theta \cdot M_{\min}$

8: else

9: $\alpha(k) = \theta \cdot M_{\min}, \beta(k) = \theta \cdot M_{\max}$

10: end if

11: end for

Output: α, β

Our proposed method is an incremental histogram update method, which will take the degree of factors among H_R^0, $\hat{H}_R^t$ and H_R^t into account. The initial histogram model H_R^0 is utilized to limit the model update rate reasonably and adapt the tracked target's partial occlusion, rotation or illumination. The parameters α and β can be considered as the weight of each bin from the target's histogram model, so we call the proposed update model as weighted incremental histogram model.

4. Adaptive on-line Tracking

Our goal is to design an adaptive on-line multi-target tracking framework base on WIHM in this section. By the WIHM information of a target, two works will be completed for tracking: calculating the optimal size of targets and determining whether the target is partially occluded by others or not. Meanwhile, an on-line learning process is adopted and modified to find the position of targets automatically in the next frame.

4.1. Optimal Size

Suppose there are two normalized histogram models P, Q associated to the tracked regions. A similarity measure should be employed to measure the difference between two histogram models. The most often used similarity measure is Bhattacharyya coefficient, which can be formularized as:

$$\rho(P,Q) = \sum_{k=0}^{B-1} \sqrt{P(k)Q(k)} \tag{3}$$

The larger $\rho(P,Q)$ is, the more similar the histogram models are. Based on the Bhattacharyya coefficient, the Bhattacharyya distance can be defined by the following equation:

$$d(P,Q) = \sqrt{1-\rho(P,Q)} \tag{4}$$

In this paper, the initial targets can be detected by an off-line classifier [4]. The position of the targets in the next frame is estimated through the confidence map in on-line boosting. We assume that the position is correct. However, the optimal size is unknown over time. How to calculate the optimal size of the target? A multi-scale grading method can be employed to obtain the optimal size of targets.

For a tracked target T, suppose we already acquired the histogram model H_T^i in the i-th frame and the position (x_i, y_i) in the (i+1)-th frame, we can calculate a set of regions $\Re_T^{i+1}$, satisfying a scaled upper and lower limits and the same regional center (x_i, y_i). The optimal size of T in (i+1)-th frame can be estimated by the minimum Bhattacharyya distance between H_T^i and the histogram models of regions in $\Re_T^{i+1}$.

$$d_{\min} = \min_d(\{d(H_T^i, R), R \in \Re_T^{i+1}\}) \tag{5}$$

$$\hat{R}_T^{i+1} = R_{index(d_{\min})}, R \in \Re_T^{i+1} \tag{6}$$

4.2. Occlusion Judgment

Occlusion judgment can be estimated using Bhattacharyya distance simply. Through the formula (6), we can obtain the optimal size $\hat{R}_T^{i+1}$ of the target T. Meanwhile, H_T^{i+1} can also be calculated by (2). The Bhattacharyya distance $d(H_T^{i+1}, H_T^i)$ between H_T^{i+1} and H_T^i can be acquired in accordance with the formula (7). A threshold η needs to be set for determining whether the target T is occluded by others or not.

$$O_T^{i+1} = \begin{cases} 1 & if\ d(H_T^{i+1}, H_T^i) > \eta \\ 0 & otherwise \end{cases} \tag{7}$$

4.3. Multi-target Tracking

In this section, we represent an adaptive on-line multi-target tracking framework to adjust to the changes of appearance model or environment. As we all know, how to select and update a series of weak classifiers for each tracked target is a crucial work in traditional on-line training framework [4]. Therefore, if tracking a target is successful, the on-line tracking framework must be able to choose appropriate weak classifiers from training samples collected in real time and take the performance of each weak classifier into account. However, in traditional on-line training process, these weak classifiers corresponding to the tracked targets may be wrong choose or update due to the occluded phenomenon of targets. This easily leads to target's tracking inaccurately or lost. In addition, we employ WHIM to estimate the optimal size of targets and determine occlusion in real-time. This means the size of tracked target is change with time. Therefore, we need to modify the traditional on-line tracking framework simultaneously.

In [4], given a set of M weak classifiers with hypothesis $\Psi = \{h_1, h_2, ..., h_M\}$. Each weak classifier $h_i \in \Psi$ corresponds to a weight w_i. A selector $h_{sel} = h_m, h_m \in \Psi$ is one of these weak classifiers, satisfying the following condition:

$$e_m = \min(\{e_1, e_2, ..., e_M\}) \tag{8}$$

Where $e_i, i = 1,2,...,M$ is the classification error of h_i in the current training samples. For each tracked pitch x, a strong classifier $H(x)$ can be selected from a set of K selectors by on-line boosting. The output of $H(x)$ is calculated by a linear combination of selectors:

$$OUT_{H(x)} = sign(\sum_{n=1}^{K} w_n h_n(x)) \tag{9}$$

In tracking process, the size of tracked target (or pitch) is not fixed, so we need to estimate new position and size of each selector before updating its weight. In order to speed up estimation, we introduce a simple yet effective way, called same-rate scaling method. The details about this method are shown in [23] and [24]. From experience, the selector responses of nearby scales can be approximated accurately enough. We can employ the scaling factor s to compensate for the difference caused by the changes of size approximately.

$$R(s) = \begin{cases} a_u s^{b_u} & if\ s > 1 \\ a_d s^{b_d} & otherwise \end{cases} \tag{10}$$

Where $R(s)$ corresponds to the ratio between a selector response at scale 1 versus scale s, and a_u, b_u, a_d, and b_d are the up-scaling and down-scaling factors. In this paper, we set $a_u = 1$, $b_u = 0$, $a_d = 0.89$ and $b_d = 1.586$ according to the empirical values from [23].

After solving the size problem above, we propose a novel on-line multi-tracking procedure in Algorithm 2.

Algorithm 2 Adaptive on-line multi-target tracking

Input: 1) Target set ψ_i in *i*th frame detected by a off-line detector (using Haar-like feature)
2) Tracked results: successful results ζ_{i-1} and failed results φ_{i-1} ($\zeta_0 = \varphi_0 = \Phi$)
3) Max tolerated frames $F = 100$, Bhattacharyya distance thresh $\eta = 0.4$

1: For each $\tau \in \varphi_{i-1}$, if tolerated frames $f_\tau < F$, $f_\tau = f_\tau + 1$; otherwise delete τ
2: For $n = 0$ to $|\Psi_i| - 1$ do
3: If R_n ($R_n \in \Psi_i$) is a new target
4: Create the detector of R_n using on-line training method[2]
5: Compute the initial gray-level histogram H_n^i of R_n
6: For each $\tau \in \varphi_{i-1}$, if $d(H_\tau, H_n^i) < \eta$, $\zeta_i = \zeta_i \cup \{\tau\}$; Otherwise delete τ
7: End if
8: End for
9: For $k = 0$ to $|\xi_{i-1}| - 1$ do

10: Estimate the position of $D_k \in \xi_{i-1}$
11: Get the optimal size $\hat{R}^i_{D_k}$ of D_k by formula (5), (6) and the histogram H
12: If $d(H, H^{i-1}_{D_k}) < \eta$
13: $\zeta_i = \zeta_i \cup \{D_k\}$, and $H^i_{D_k}$ is computed by formula (2)
14: Update the detector of D_k
15: Else
16: $\varphi_i = \varphi_i \cup \{D_k\}$
17: End if
18: End for
Output: ζ_i and φ_i

5. Experimental Results

For validate the feasibility and effectiveness of our proposed on-line multi-target tracking framework, we have carried out some experiments. All our experiments are finished on a PC with Intel(R) Core(TM) Duo E7200 CPU running at 2.53 GHz and 2GB memory. In addition, we set the empirical threshold $F = 100, \eta = 0.4$ in our experiments.

We first evaluate the effectiveness of WIHM by tracking one target from the non-occluded situation (392th frame) to completely occluded situation (685th frame). Performance of our method is compared with fixed update model, and linear update model [22]. In Figure 2, when the tracked target begins to be occluded by other targets in the 392th frame, the Bhattacharyya distance of three methods becomes large gradually. Up to the 452th frame (Only one fifth of the target is occluded), the Bhattacharyya distance of fixed update model exceeds η. The same result of our method appears in the 561th frame (About three fifth of the target is occluded). However, linear update model loss original target and track other target after the occlusion (475th frame). From the results of Figure 2, it is clear that the WIHM method is consistently better than the other two methods in most frames.

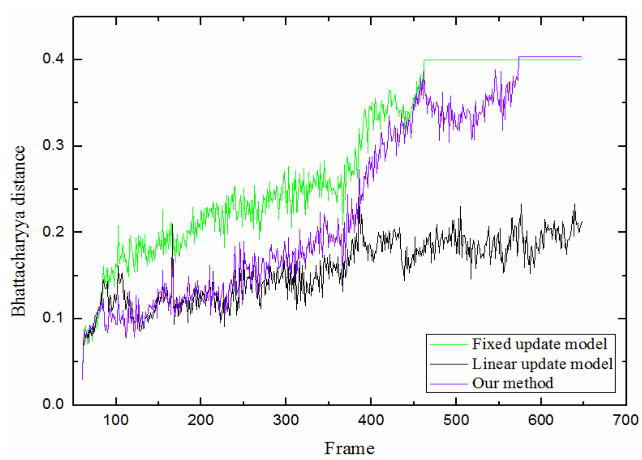

Figure 2. Performance of WIHM Approach

Second, to validate the effectiveness of our proposed multi-target tracking framework, we test our method with two tracking methods: Multiple Instance Learning (MIL) [25] and

Distribution Field (DF) [26]. MIL applies an unsupervised on-line learning over Haar features to train a discriminative tracker. DF is a representation that allows smoothing the objective function without destroying information about pixel values, which can helps in disregarding outliers during tracking without modeling.

For facilitate comparison, we use some publicly available videos, which are collected by B. Babenko[25]. The tracking accuracy is computed to compare our method with others. The correct tracking satisfies the following condition [26]:

$$T = \begin{cases} 1 & (A \cap B)/(A \cup B) > 0.5 \\ 0 & otherwise \end{cases} \quad (11)$$

Where A and B are the regions of a tracking result and ground truth respectively. The tracking accuracies by using different methods are shown in Table 1.

Table 1. Tracking Accuracies by using Different Methods. The Results of MIL and DF are Available at [26]. The First Five Videos are from [25] and the Last One is our Self-Recorded Video

Video name	Our Mehod	MIL[25]	DF[26]
Tiger1	**90.03**	43.14	88.57
sylv	**78.50**	73.88	66.79
girl	**77.95**	55.20	73.00
Coke11	71.24	17.93	**75.86**
David indoor	**100.00**	58.91	100.00
Car occlusion	**98.71**	72.65	90.26

From Table 1, the performance of our method is better than that of MIL and DF in most videos. These examples demonstrate the improved ability of our proposed method to deal with illumination changes and occlusions. However, we also found that the result in video Coke11 is slightly lower than DF. The main reason is that the color of foreground (target) and background is very close each other. This affects the optimal size of target by WIHM.

Some results are shown in Figure 3. Our approach finds the 0-target (green box, label 0) in the 747th frame. Then the tracking box constantly adapt to the size changes of 0-target. After the 975th frame, only a small part of 0-target is visible, so occlusion is output (red box, label 0). However, the 0-target is detected and tracked again with the same label at the beginning of the 1019th frame.

747 # 947 # 975

Figure 3. Examples of Tracking Results of our Approach

6. Conclusion

This paper addresses the challenging problem of tracking of multiple objects in complex scenes using a monocular, moving camera. We analyze the importance of appearance model correspond to tracked targets and put forward an adaptive on-line multi-target tracking framework base on WIHM, which can adapt to changes of illumination, appearance and partial occlusions. Experiments on some datasets show effective improvements by our proposed method. Currently, in our framework, we only use the WIHM to infer whether the occlusion occurs roughly and does not consider the position of occlusion. So the tracking accuracy may be affected to a certain extent. Therefore, determining where the occlusion occurs with our framework to improve the tracking accuracy will be the future work.

Acknowledgements

This work is supported by the grants from Jilin Planned Projects for Science Technology Development [Grant No. 20120305 and No. 20130522119JH] and Unmanned Systems based on Machine Vision Sensing Technology Project [Grant No. 450060491089].

References

[1] B. Ben and R. Ian, "Stable multi-target tracking in real-time surveillance video", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA, **(2011)** June 20-25.
[2] S. Munder C. Schnörr and D. Gavrila, "Pedestrian detection and tracking using a mixture of view-based shape-texture models", IEEE Transactions on Intelligent Transportation Systems, vol. 9, no. 2, **(2008)**.
[3] B. Zhou X. Wang and X. Tang, "Understanding Collective Crowd Behaviors: Learning a Mixture Model of Dynamic Pedestrian-Agents", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, RI, USA, **(2012)** June 16-21.
[4] H. Grabner M. Grabner and H. Bischof, "Real-Time Tracking via Online Boosting", Proceedings British Machine Vision Conference (BMVC), Edinburgh, UK, **(2006)** September 4-7.
[5] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of Computer and System Sciences, vol. 55, no. 1, **(1997)**.
[6] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, USA, **(2001)** December 8-14.
[7] S. Li and Z. Zhang, "FloatBoost learning and statistical face detection", IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 9, **(2004)**.
[8] N. Oza and S. Russell, "Online bagging and boosting", Proceedings of Artificial Intelligence and Statistics, Morgan Kaufmann, **(2001)** January 4-7.
[9] N. Oza, "Online Ensemble Learning", PhD thesis, University of California, Berkeley, **(2001)**.
[10] H. Grabner and H. Bischof, "On-line Boosting and Vision", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, **(2006)** June 17-22.
[11] B. Wu and R. Nevatia, "Improving part based object detection by unsupervised, online boosting", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota, USA, **(2007)** June 18-23.

[12] S. Avidan, "Ensemble Tracking", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 2, **(2007)**.
[13] T. Parag F. Porikli and A. Elgammal, "Boosting adaptive linear weak classifiers for online learning and tracking", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, Alaska, USA, **(2008)** June 24-26.
[14] C.-H. Kuo, C. Huang and R. Nevatia, "Multi-Target Tracking by On-Line Learned Discriminative Appearance Models", IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, **(2010)** June 13-18.
[15] Z. Qi, Y. Xu, L. Wang and Y. Song, "Online multiple instance boosting for object detection", Neurocomputing, vol. 74, no. 10, **(2011)**.
[16] Y. Xie, Y. Qu, C. Li and W. Zhang, "Online multiple instance gradient feature selection for robust visual tracking", Pattern Recognition Letters, vol. 33, no. 9, **(2012)**.
[17] B. Yang, C. Huang and R. Nevatia, "Learning affinities and dependencies for multi-target tracking using a CRF model", IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA, **(2011)** June 20-25.
[18] R. Bo Yang Nevatia, "An online learned CRF model for multi-target tracking", IEEE Conference on Computer Vision and Pattern Recognition, RI, USA, **(2012)** June 16-21.
[19] Q. Miao, G. Wang, X. Lin, Y. Wang, C. Shi and C. Liao, "Scale and rotation invariant feature-based object tracking via modified on-line boosting", 17th IEEE International Conference on Image Processing, Hong Kong, China, **(2010)** September 26-29.
[20] D. Comaniciu and P. Meer, "Mean Shift: a robust approach toward feature space analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, **(2002)**.
[21] G. R. Bradski, "Computer vision face tracking for use in a perceptual user interface", Intel Technology Journal, vol. 2, no. 2, **(1998)**.
[22] K. Nummiaro, E. Koller-Meier and L. Van Gool, "An adaptive color-based particle filter", Image and Vision Computing, vol. 21, no. 1, **(2003)**.
[23] P. Dollár, S. Belongie and P. Perona, "The fastest pedestrian detector in the west", Proceedings of British Machine Vision Conference (BMVC), Aberystwyth, UK, **(2010)** August 30-September 2.
[24] R. Mathias Benenson, M. Timofte, R. Van and L. Gool, "Pedestrian detection at 100 frames per second", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, RI, USA, **(2012)** June 16-21.
[25] B. Babenko, M. H. Yang and S. Belongie, "Visual tracking with online multiple instance learning", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, **(2009)** June 20-25.
[26] E. Learned-Miller and L. Sevilla-Lara, "Distribution Fields for Tracking", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, RI, USA, **(2012)** June 16-21.

Authors

WenHui Li, was born in 1961. He is a Professor and doctor supervisor in College of Computer Science and Technology, Jilin University, Changchun, P.R. China. His major research interests include computer vision, image processing, pattern recognition, graphics, virtual reality, CAD and geometric constraint.

HongYin Ni, was born in 1984. He received his M.S. degree majoring in Technology of Computer Application in 2009 from Computer Science and Technology, Jilin University, P. R. China. He is studying for his Ph.D. in the College of Computer Science and Technology, Jilin University. His research interests include computer image processing, pattern recognition and virtual reality technology.

Ying Wang, was born in 1983. She received her M.S. degree majoring in software engineering in 2008 from College of Software, Jilin University, P. R. China. She is now pursuing her Ph.D. at Image Processing and Virtual Reality Technology Lab in the College of Computer Science and Technology, Jilin University. Her research interests include face recognition, fire/smoke detection, pedestrian detection and object tracking.

Peixun Liu, was born in 1986. Since 2011, He has been working on the PhD degree in College of Computer Science and Technology, Jilin University, Changchun, P. R. China. His research interests include computer vision, image processing and pattern recognition.

