VOL. 10, NO. 8, MAY 2015 ISSN 1819-6608

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

MULTI TARGET TRACKING UNDER OCCLUSION USING PARTICLE FILTER AND PROJECTED GRADIENTS

Karpagavalli P. M. E., Ramprasad M. E. and Vijayadurga P.

Department of Electronics and Communication Engineering, K.L.N. College of Engineering, Madurai, India

E-Mail: karpagavallimohan@gmail.com

ABSTRACT

A Promising approach for occlusion tracking, the most complex chore in target representation and target localization is proposed. Multiple kernels provide more in-depth information than single kernels. Multiple inter related kernels have thus been exploited for tracking occlusion. Adaptive multiple kernels based tracker is attractive, because it contemporaneously combines a summary description of both the intensity values and spatial positions. The algorithm combined with projected gradients provides continuous optimization algorithm. Projected gradients furnish the best match using predefined constraints. An efficacious method is also accomplished to deal with the scale change issues. The paper provides fast convergence and lesser computation. Simulation results prove that the proposed method tracks the video object under occlusion efficaciously.

Keywords: multiple kernels, projected gradient, mean shift, particle filter, cost function.

1. INTRODUCTION

The Escalation of high powered computers, the availability of high quality and inexpensive video cameras and the increasing need for automated video analysis has generated a great deal of interest in target tracking. The objective of target tracking is to associate targets in consecutive frames. Target tracking finds application in a wide area as motion based recognition, automated surveillance, video indexing, human computer-interaction and traffic monitoring. Target tracking becomes complex due to occlusion. Innumerable methods for tracking the objects have been proposed. A brief introduction to main tracking categories is discussed below [3].

1) Point tracking: This method requires an external mechanism to track targets in successive frames. The objects are represented by points which depend on previous state, including motion and position.

2) Kernel tracking: kernel is a convolution matrix or mask representing the appearance of a target. Kernel tracking enumerates the motion of the target.

3) Silhouette tracking: A silhouette tracker bestows an accurate shape description of these objects. This uses the information encoded in the target frame. Kernel based tracking has an outstanding attention since they were established. Instinctively, kernel based trackers are attractive as they integrate intensity values and spatial positions, avoiding complex modeling of object shape, appearance and motion. The cardinal idea of this algorithm is to diminish the difference between target and candidate appearance model by masking the objects applying kernel [3]. Nevertheless, when an occlusion occurs sufficient information is not available to track the target. The imprecise judgment of occlusion situation and illegitimate target template updating has to take place during occlusion. This paper bestows on retained tracking and obtaining exactitude solution. During the occlusion single kernel is not authenticated for tracking the targets or objects and hence multiple kernels are applied. We propose a projected gradient based multiple tracking schemes that use multiple kernels to track the targets. The whole system is automated and does not need human intervention. The main benefaction of this paper is:

1) The Projected Gradient method is engrossed to procure the best match of tracking a target under Pre-defined constraints.
2) Adaptive kernels are used to effectively track the targets.
3) The gradient of the density estimator is used to efficaciously update the scale of the tracking target.

This paper proceeds with related work in section II. Then the introduction to our proposed multiple- kernel tracking is presented in section III followed by the experimental result and discussion in sections IV. The conclusion is given in section V.

2. RELATED WORK

Since it is not possible to have a detailed deliberation on all visual tracking papers, we focus on the most applicable paper regarding occlusion. Multiple Object Tracking by Kernel Based Centroids Method is one of the best methods for occlusion tracking, was pertained to find the area of a video frame that is locally most similar to a previously initialized model [4]. R. Collins [5] manipulated two approaches for tracking that intricate creating a color likelihood image with pixels having similar weight and signifying color with histogram. In object tracking [6], appearance model was smoothed temporally by the robust Kalman-filter for precise detection of severe occlusion. A new framework for efficacious tracking of objects by using constrained multiple kernels with kalman filtering continuously updates the position and tracks the occlusion [7]. The major failure in object tracking is the inability to track rapidly moving objects. A mean shift tracker using both the spatial measure and color histogram based similarity measure was procured. Boundary cues were also contemplated in kernel based tracking [8]. Though the above works, predicaments additional computations using multiple kernels are needed. In our proposed work scale change issue is elucidated effectively and has been signified using experimental results. All the above works used only single kernels for occlusion tracking which failed with severe occlusion conditions because of hidden statistics. In order to get superior tracking results multiple kernels are enrolled for these years consider scale change in [3], the color similarities and shape fitness are used for occlusion tracking. This depiction provided more pliancy while facing the problems of tracking such as occlusion. In [9] the new position was estimated using RDHOGPF. The estimated results were grey masked making the detection precise. Occlusion strategy [10] was procured based on the direct color, motion integration mechanism which tracked effectively and automatically when the lost object reappears. Multiple kernels centered on high motion areas [11] proposed by Porikli tracked targets with rapid motion and convergence property was also enhanced by two additional likelihood terms. Robust fragment based tracking [12] where objects were divided into fragments and tracked using kernels, but this method failed when the background and foreground were similar. In [13] mean shift iterations were used for tracking. Multiple distracting objects with similar features were tracked coherently. Multiple kernel learning algorithms proposed by Sonnenburg [14] elucidated large scale problems by iteratively using vector machine code, but the relationships between kernels were not appraised. Furthermore adaptive weights for kernels were also not obtruded. The interrelationships among kernels were proposed by Fan *et al*. using predefined constraints [15]. In this approach they represented target motion spatially distributed simpler motions associated with motion estimator. It works for object tracking, but under occlusion the results are deviated easily from the optimum. Adaptive weights for multiple kernels were also not considered which makes the system unstable. In order to overcome these problems completely we propose a projected gradient based multiple kernel tracking [13].

3. PROPOSED METHOD

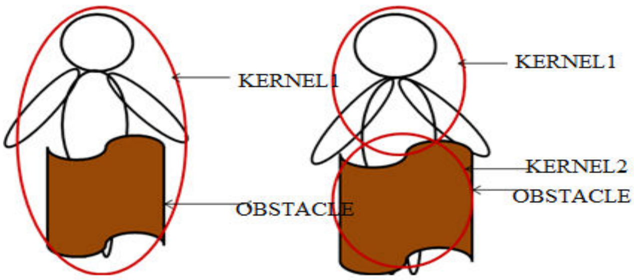
The occluded targets require to be tracked in the successive frames interlude for obtaining victorious tracking. The target model is assumed to be known and the candidate model is extracted at each location in every frame. The similarity measures between these two models have to be secured and the intention is to obtain the candidate model that has the exorbitant similarity which can be seen in [2].

A. Multiple kernels

In [7], a feature space is chosen for designating a target. The model of the target is illustrated with the pdf. For example, the color histogram of the target can be considered as the target model. The amount of benefaction of a pixel is ascertained by the difference between the kernel center and the pixel.

$$F(y) = \frac{\sum_{i=0}^{n_h} w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{\sum_{i=0}^{n_h} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} \quad (1)$$

Including the previous frame state, the total cost function needs to be efficaciously decreased and the constraints also need to be gratified during the search. This is acquired with the movement vector determined based on the projected gradient method. The movement vector iteratively solves the constrained optimization problem. The instinctive idea is to keep the values of the constraint function unchanged while projecting the gradient vector onto space. Additionally, for further handling of constraints, the second term satisfying the guarantee of the constraint is instituted. Hence, δy consists of two components δy_A and δy_B. where y is the center of the kernel; x_i is the location of pixel to be examined and n_h is the number of pixels; w_i is the weight for each pixel; k () is a monotonically decreasing function. The mean shift algorithm is implemented to obtain the optimum solution effectively. Nevertheless, in single kernel tracking the dimensionality of the histogram and the detailed information is not secure. Differently, the problem of similarity maximization can be redeveloped to minimizing the cost function. We can procure this minimization by inverse proportionate to the similarity function.



(a) (b)

Figure-1. (a) Single kernel (b) Multiple kernels.

$$J(y) \propto \frac{1}{simi(y)} \quad (2)$$

where, *simi(y)* is the similarity function of position y in the state space domain. The mean shift tracking can be embraced when the single kernel is used to track the objects. Tracking objects during occlusion it leads to error which can be avoided by using multiple kernels where the kernel is represented using the bounded box. When the occlusion occurs, the tracking accomplishment can be terribly influenced since the kernel 1 assimilates a large amount of unwanted information to the targets in Figure-1(a). When an additional kernel is applied in Figure-1 (b), though the kernel 2 is non-observant, the well observed kernel 1 is used to recompense the ill effects resulted due to occlusion after the kernels are made inter-linkable. Hence, for N multiple kernels, the total cost function J(y) is given as the sum of N individual cost J(y). Other than the cost function, the constraint functions C(y) =0, are also needed to be trusted. The constraint functions enclose the kernels depending on their interrelationships. For illustration; if the target in Figure-1 is a rigid body, then the kernel 1 will repeatedly be given as the following equation.

$$J(y) = \sum_{i=1}^{N} J_i(y) \tag{3}$$

$$\delta y = \alpha\left(-I + C_y\left(C_y^T C_y\right)^{-1} C_y^T\right) J_y + \left(-C_y \left(C_y^T C_y\right)^{-1} C(y)\right) \tag{4}$$

where y $\in R^{n;\ C}$ is the vector of constraints; C_y is the gradient of vector constraints with respect to y and consists of m constraint functions; J_y is the gradient vector of the total cost function with respect to y and $\alpha>0$ is the step size. And also, our proposed method results in less complexity because, the dimensionality $C_y^T C_y$ is equal to the number of constraints and will not increase with the number of kernels. This is also smaller than the dimensionality of the matrix which is to be inversed [15]. The iterative projected gradient in each frame for multiple kernel tracking is defined by the following algorithm where |.| takes entry-wise absolute value:

Algorithm1. Multiple kernels Optimization Algorithm

1. x is the final state of the previous frame;
a. Counter=0.
2. While counter<T
3. Counter=counter+1.
4. Evaluate C(y) and J(y).
5. ***if*** $|C(y)| < \varepsilon_C$ and $J(y) < \varepsilon_J$
6. ***then***
7. End of the iteration.
8. ***else***
9. ***if*** $|C(y)| < \varepsilon_C$ and $J(y) \geq \varepsilon_J$
10. ***then***
11. Compute δy_A
12. Apply $y = y+\delta y_A$
13. ***else***
14. ***if*** $|C(y)| \geq \varepsilon_C$ and $J(y) < \varepsilon_J$
15. ***then***
16. Compute δy_B.
17. Apply $y = y+\delta y_B$
18. ***else***
19. Compute δy_A and δy_B.
20. Apply $y = y+\delta y_A+\delta y_B$
21. ***end if***
22. ***end if***
23. ***end if***
24. ***end while***

B. Adaptive cost function

In Algorithm 1, the values of ε_C and ε_J are set in the range according to the size of the objects empirically, where the iterations will be equal to 5.As specified above, during occlusion, all the kernels are not well grounded. Each kernel is coupled with an adaptively changeable weight value w_i for calculating the total cost function;

$$J(y) = \sum_{i=1}^{N} w_i J_i(y) \tag{5}$$

Consequently, the movement vector in (4) is revised as

$$\delta y = \alpha\left(-I + C_y\left(C_y^T C_y\right)^{-1} C_y^T\right) W J_y + \left(-C_y \left(C_y^T C_y\right)^{-1} C(y)\right), \tag{6}$$

$$= \delta y_A + \delta y_B$$

and $w_i \infty$ simi(y); I is an n/N * n/N identity matrix with n being equal to the dimension of the state space, and N is the number of the kernels. w_i is i^{th} weight value which is adaptively updated and normalized depending upon the individual similarity $simi_i$. The similarity represents the degree of match between the color feature of the candidate and the target in the single kernel [6].

C. Scale issue

The size of the target or the object changes when the object keeps moving. The changes in scale are observed due to changes of the angle and distance between the target and the camera. There are many methods [5], [7] proposed for the scale change issue, but none of them can be instinctively applied.

As these methods need subsidiary processes for dealing with the scale variable, a simple efficacious solution is proposed to vanquish this issue, as affirmed by the experimental results. A similarity measure of the object delineated by the extracted histogram is highly hinging on the kernel bandwidth *h*. When the object size is changing; kernel bandwidth has to be robustly changed for acquiring the histogram that matches exactly with the original one. For example, Figure-2 displays the same object with different sizes. The two nearby pixels have the same weights for histogram calculation if *D/h = D'/h'*. Hence the object size has the high positive correlation to the kernel bandwidth. Then the change of the kernel bandwidth is employed to surmise the object size that changes. All the terms in (7), including w_i, v_i, $g(v_i)$,$k(v_i)$ and h, can straightly be devised using the tracking methods such as the mean shift vector and cost function enumeration established on these the scale change factor can be effectually determined to reflect the appropriate size.

$$\Delta f(h) = \frac{\delta f(h)}{\delta(h)}$$

$$= f(h)\left(\frac{-2\sum i(g(v_i)v_i)}{h\sum ik(v_i)} + \frac{-2\sum i(w_i g(v_i)v_i)}{h\sum i w_i k(v_i)}\right) \tag{7}$$

ARPN Journal of Engineering and Applied Sciences

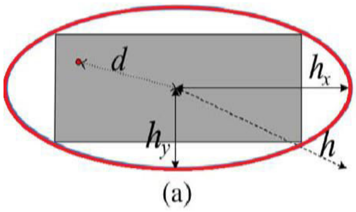


Figure-2. (a) Original size and (b) Smaller size.

The grey rectangle is the object. The Red ellipse represents the kernel with bandwidth and which is equal to and, respectively. The red dots, corresponds to the locations on the objects with two different sizes.

4. EXPERIMENTAL RESULTS

Here, we deal with the implementation steps in detailed manner. Our simulation scenarios in Figure-4 take into account of tracking multiple persons in a video who is occluded involving potential scale changes in a crowd. We compared our kernel-tracking methods with the collaborative kernels [16] using accuracy and error rate. The proposed multi-kernel tracking using projected gradients with productiveness of our approach for dealing with the scale changes will also be manifested. Thus the robust tracking of multiple persons simultaneously will be evidenced.

A. Implementation setup

Each and every object in the video is portrayed with the multiple inter-related kernels in the multiple kernel tracking section and the net state vector x is framed of the centroid of all the kernels denoted where N is the number of kernels, and (x_i,y_i) is the centroid of the i^{th} kernel. When more number of kernels is applied better results are procured as it handles occlusion efficaciously in a crowded environment. Experiments on multiple kernel using projected gradients

In order to highlight the robustness of multiple kernels tracking method, we focus on tracking multiple targets that are occluded, with the potential scale changes. The layouts for multiple kernels are given in Figure-3. Hence the number of kernels is implemented depending on the target. For the efficient procurement of the fully automatic tracking, we affiliate the multiple kernels tracking system into a particle filter based tracking system [21], where the particle prediction as the initial location for the multiple kernels tracking section is utilized and the result from the multiple kernel tracking is served as the menstruations for particle renovation. Our objective is to evolve a fully automatic tracking system that tracks multiple objects concurrently. Various video clips are utilized for the performance assessment.

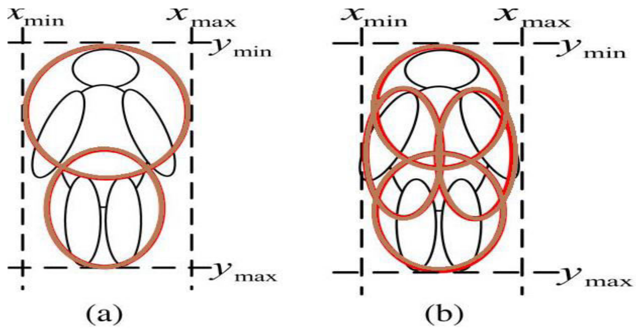


Figure-3. Layout for the multiple kernels. (a) 2 kernels. (b) 4 kernels

kernels represented as brown ellipses. Note that the object is not necessarily viewed from the front. The full process is automatic and no physical intercession is needed. Moreover, we analogize the performance in terms of the error rate, with the previous methods by using the same video clips. The error rate is secured based on all the people in all the frames in the video and the error is defined as the pixel distance between the targets' centroids of the simulation results and the ground truths which can be produced with the help of ViPER tool [22]. When the organization of the kernels is planned, the overlapping between a pair of the kernels is allowed. Moreover, the overlapping areas between the pair of the kernels are limited; when the occlusion occurs, both the Kernels will agonize from the ill-observable condition. For human tracking, we can instinctively outline two kernels, one for the upper part and one for the lower part. It is sensible to assume that the human remains upright when they are moving. Hence of upper and lower body changes the geometrical relationship.

Table-1. The cost function variations Vs iterations.

Cost function variations				
Intermediate steps	**Iteration1**	**Iteration2**	**Iteration3**	**Iteration4**
Cost function	3.448	2.255	1.484	1.329
1^{st} Constraint function value	1	1	1	1
2^{nd} Constraint function value	0	0	0	0
Update vector applied	δy_A	δy_A	δy_A	NULL

The cost function variations table indicates the changes in the constraint functions value for several iterations. The vector will be updated based on these obtained values.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

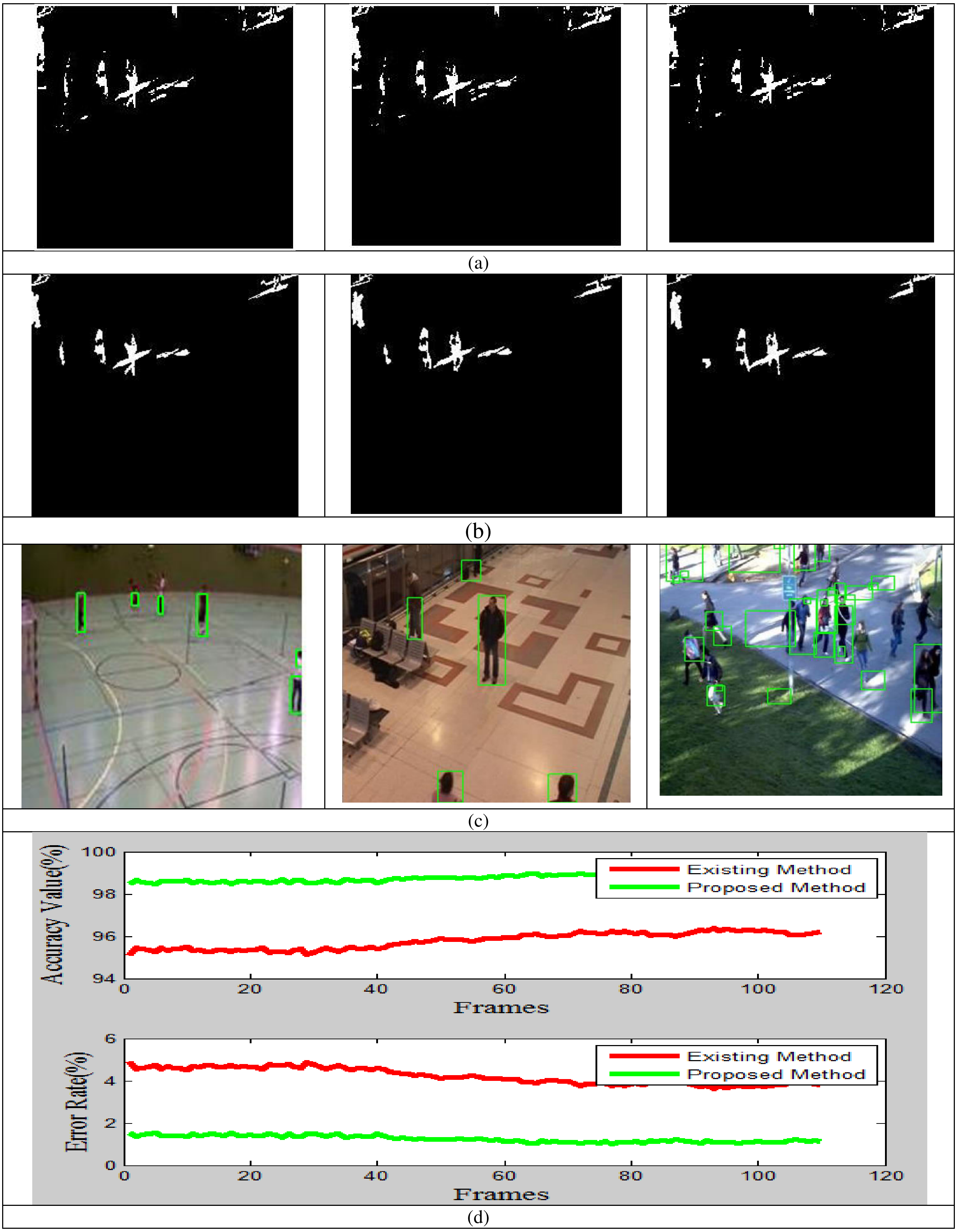


Figure-4. Experimental results of targets tracking under Occlusion of two different videos. a) Estimated Targets position Frame #25, #26, #27. b) Predicted Targets position Frame #25, #26, #27. c) Targets tracking of three different videos d) Performance measure of different videos (Error rate and accuracy).

ARPN Journal of Engineering and Applied Sciences

5. CONCLUSIONS

We suggest an ingenious method that implements multiple kernels using projected gradients for attaining the best match using predefined constraints. The adaptive weights are introduced to the kernels to track the occluded targets more effectively. Hence an efficacious method for scale change issue is proposed. The multiple kernels Tracking is combined with the particle filter to provide an automatic tracking module. From the experimental results it is evident that the proposed method successfully tracks the multiple targets under occlusion and the entire system renders that the targets are tracked concurrently.

REFERENCES

[1] Chun-Te Chu, Jenq-Neng Hwang, Hung-I Pai, and Kung-Ming Lan," Tracking Human under Occlusion Based on Adaptive Multiple Kernels with Projected Gradients," IEEE Transactions on Multimedia, Vol. 15, No. 7, November 2013.

[2] A.Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," ACM Comput. Surveys, vol. 38, no. 4, 2006.

[3] Lee, K.-H,Hwang, J.-N, Chen, S.-I, "Model-Based Vehicle Localization Based on Three-Dimensional Constrained Multiple-Kernel Tracking," IEEE Transactions on Circuits and Systems for Video Technology. vol. pp, No. 99, June 2014.

[4] Rahul Mishra1, Mahesh K. Chouhan2, Dr. Dhiiraj Nitnawwre, "Multiple Object Tracking by Kernel Based Centroid Method for Improve Localization," International Journal of Advanced Research in Computer Science and Software Engineering Volume 2, Issue 7, July 2012.

[5] R. T. Collins, "Mean-shift blob tracking through scale space," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2003, vol. 2, pp. 234-240.

[6] Hieu T. Nguyen andArnold W.M. Smeulders, Member IEEE," Fast Occluded Object Tracking by a Robust Appearance Filter," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, No. 8, August 2004.

[7] Kuan-Hui, Lee Yong-Jin Lee, Jenq-Neng Hwang, "Multiple-kernel based vehicle tracking using 3-D deformable model and license plate self-similarity," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2013.

[8] R.Venkatesh babu, Anamitra Kaur," Kernel based spatial color modeling for fast moving object tracking," IEEE International Conference on vol: 1, 2007.

[9] B. Martinez, L. Ferraz, X. Binefa, and J. Diaz-Caro, "Multiple kernel two-step tracking," in Proc. IEEE Int. Conf. Image Processing, 2006, pp. 2785-2788.

[10] Bing - FeiWu, Chih - Chung Kao, Cheng-Lung Jen, Yen - Feng Li, Ying-Han Chen, and Jhy-Hong Juang," Relative - Discriminative - Histogram-of-Oriented-Gradients-Based Particle Filter Approach to Vehicle Occlusion Handling and Tracking," IEEE Transactions On Industrial Electronics, Vol. 61, No. 8, August 2011.

[11] Zhang Jin, Sun Hongguang, Guan Weizhou, Wang Jialiang, Xie Yannan and Shang Bingan," Robust Human Tracking Algorithm Applied for Occlusion Handling," Fifth International Conference on Frontier of Computer Science and Technology (FCST), 2010.

[12] F. Porikli and O. Tuzel, "Multi-kernel object tracking," in Proc. IEEE Int. Conf. Multimedia and Expo, 2005, pp. 1234-1237.

[13] J. Fang, J. Yang, and H. Liu, "Efficient and robust fragments-based multiple kernels tracking," Int. J. Electron. Commun, vol. 65, pp. 915-923, 2011.

[14] Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "More efficiency in multiple kernel learning," in Proc. ICML, 2007.

[15] Z. Fan, Y. Wu, and M. Yang, "Multiple collaborative kernel tracking," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2005, vol. 2, pp. 502-509.

[16] Zhang, W. Huang, Z. Huang, and L. Li, "Affine object tracking with kernel-based spatial-color representation," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2005, vol. 1, pp. 293-300.

[17] C. Chu, J. Hwang, H. Pai, and K. Lan, "Robust video object tracking based on multiple kernels with projected gradients," in Proc. IEEE Conf. Acoustics, Speech and Signal Processing, May 2011, pp.1421-1424.

[18] H. Grabner and H. Bischof, "On-line boosting and vision," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2006, vol. 1, pp. 260-267.

[19] S.Wang, H. Lu, F. Yang, andM. Yang, "Superpixel tracking," in Proc. IEEE Int. Conf. Computer Vision, 2011.

[20] C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006.

[21] C. Chu, J. Hwang, S.Wang, and Y. Chen, "Human tracking by Adaptive Kalman filtering and Multiple Kernels Tracking with Projected Gradients," Proc.

ACM/IEEE Int. Conf. Distributed Smart Cameras, 2011.

[22] D. Doermann and D.Mihalcik, "Tools and techniques for video performance evaluation," in Proc. IEEE Int. Conf. Pattern Recognition, 2000, pp. 167–170. [Online]. Available: http://viper-toolkit.sourceforge.net/.

[23] M. Fashing and C. Tomasi, "Mean shift is a bound optimization," IEEE Trans. Pattern Anal. Mach. Intell. vol. 27, no. 3, pp. 471-474, March 2005.