*International Journal of Engineering Science Invention (IJESI)*
*ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726*
*www.ijesi.org ||Volume 7 Issue 7 Ver IV || July 2018 || PP 27-31*

# Multiple Object Tracking System using Discriminative Correlation Filter

V. Ramalakshmi @ Kanthimathi[1], Dr. M. Germanus Alex[2]
*[1]Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore, India*
*[1]Assistant Professor, Kamarajar Government Arts College, surandai.*
*[2]Research Guide, Department of Computer Science, Bharathiar University, Coimbatore, India*
*[2]Assistant Professor, Kamarajar Government Arts College, Surandai.*
*Corresponding Author: V. Ramalakshmi*

***Abstract -*** *Multiple Object Tracking (MOT), or Multiple Target Tracking (MTT) plays a vital role in computer vision. The objective of this paper is to identify multiple objects in various scenarios. In this paper, the objects in the video are identified and tracked using Discriminative Correlation Filter. The proposed method is tested on publicly available datasets: TUD Campus, TUD Crossing, TUDStadtmitte and PETS2009 S2-L1 datasets. The experimental results are compared with recent multiple object tracking methods. It is substantially proved that the proposed method achieves better performance when compared to other methods.*
***Keywords –*** *blob, bounding box, correlation, occlusion, tracking,*

---

Date of Submission: 29-06-2018 Date of acceptance: 14-07-2018

---

## I. Introduction

There have been various researches on the tracking of objects from video sequences over the last two decades. The interest of multiple object tracking is motivated by numerous applications such as surveillance, video conferencing, man-machine interfaces, and sports enhancement. In the conventional methods, the process of MOT can be carried out in two ways, such as data association and state estimation. The task of MOT is to identify multiple objects, yielding their individual trajectories and maintaining their identities for a given input video[1].

Objects to track can be, pedestrians on the street, vehicles in the road, and sports players on the court or groups of animals [2-4]. Multiple objects could also be viewed as different parts of single object[5]. In this paper, pedestrian tracking is mainly focused. The reason behind this research is threefold. First, compared to other common objects in our environment, pedestrians are typical non-rigid objects, which is an ideal example to study the MOT problem. Second, videos of pedestrians arise in a huge number of practical applications, which further results in great commercial potential. Third, according to all data collected, at least 70% of current MOT research efforts are devoted to pedestrians.

MOT lands high-level tasks such as pose estimation, action recognition, and behavior analysis [6-8]. Compared with Single Object Tracking (SOT), which primarily focuses on designing sophisticated appearance models and/or motion models to deal with challenging factors such as scale changes, out-of-plane rotations and illumination variations, MOT additionally requires two tasks to be solved: determining the number of objects, which typically varies over time, and maintaining their identities. Apart from the common challenges in both SOT and MOT, further key issues that complicate MOT include:1) frequent occlusions, 2) initialization and termination of tracks, 3) similar appearance, and 4) interactions among multiple objects.

Many researchers concentrated on MOT in video using different methods. Early work by J. Berclazet al. [9] has developed a method for MOT with K-shortest path optimization to track the object, which is difficult to handle intersecting trajectories. Additional work by X.Zhou et al. [10] dealt with the improved spatial color appearance with interferences using Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter, in which computational cost is high. Then, to improve the confidence of sampling and perform the iteration effectively, X. H.Xia et al.[11] have implemented the Markov chain Monte Carlo-based multi-object visual tracking method. W. Hu et al. [12] addressed the multi-feature joint sparse representation to automatically focus the visible parts of an occluded object, in which many parameters need to be set and the semantic corrections between the different features are not modeled.

I. Ali and M. N. Dailey [13] have developed the confirmation-by-classification method to detect and track multiple humans in high-density crowds in the presence of extreme occlusion. To capture the interdependence of multiple influence factors, X. Liu et al. [14] implemented the discriminative structure prediction model. To capture both the global and local spatial layout information, Log-Euclidean Riemannian

Subspace and Block-Division approach was developed by W. Hu et al. [15], in which detections was not flexible for high level environment change.

In this paper, we present an approach for tracking of multi-objects in video using discriminative correlation filter. The proposed method is tested with datasets containing multiple occlusions, multiple human with different locations. The results are analyzed and compared with recent MOT methods.

The remaining of the paper is organized as follows. Section II describes the overall System Architecture. The experimental results are analyzed in Section III followed by conclusion in Section IV.

## II. System Architecture

In this section, the proposed multiple object tracking system and its components are presented. Firstly, the system receives the video sequences from a video monitoring system such as CCTV or web camera. The next step is object detection which involves detecting the objects of interest from the first video frame. This step involves applying some background subtraction techniques after performing some preliminary preprocessing methods to remove the noise. The basic idea of background subtraction is subtracting a predefined background model frame from the current frame to identify the moving objects. Several advanced background subtraction methods have been introduced in the literature which is insensitive to external environmental conditions such as noise. These include approximate median, running Gaussian, and mixture Gaussian methods. After detecting the objects of interest, the next step in a visual object detection system is called object tracking which involves finding the location of the object in the subsequent video frames.

Over the years, tracking-by-detection methods have become great interest among the researchers for various visual object tracking tasks, because of their excellent tracking performance. In the object tracking task, the tracker needs to estimate the location of the bounding box of the object within each frame of the video sequence. These methods model the target localization problem as a classification problem.

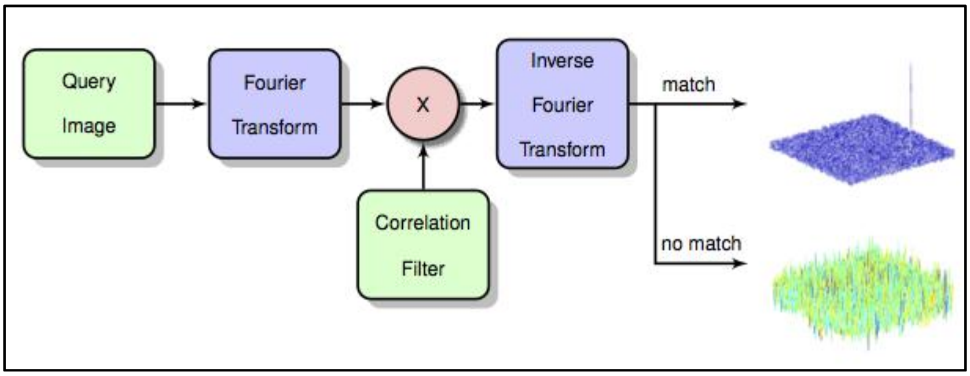


**Fig. 1** System Architecture

Initially, the target's bounding box or tracking window is selected either manually or automatically. In the manual case, a supervisor marks the rectangle boundary (left, top, right, and bottom) of the target object of interest. Whereas, in the automatic setup as in the proposed system, the prior object detection method outputs the boundary of the target. The discriminative correlation filter used in this work is based on the tracker proposed in [16]. The object tracking task is performed using these three steps: 1) after selecting object boundary, a correlation filter is trained using image patch cropped from the first frame; 2) in the subsequent frames, the target is tracked by correlating the trained correlation filter over a search window. The window location which gives maximum correlation output is marked as the new location of the target; 3) finally, based on this new location, an online update of the correlation filter is performed. Steps 2 and 3 are repeated for all the frames in a video sequence.

**Algorithm 1**: Implementation steps of the proposed visual object tracker

**Input**: Input video frames $V_i$, 0 <=i<N, where N is the total number of frames.
**Output**: Tracks of target objects in the video sequence
**Steps**:

1. Read the first video frame $V_0$ and detect the M target objects.
2. Initialize tracks Tm for M target objects such as id, bounding box, etc.
3. Start the tracker for all tracks Tm using $V_0$ and learn the correlation filter.
4. For all the video frames $V_i$ 1<=i<N
   Read the video frame $V_I$
   For all trackers in $T_j$ 1<=j<M
      Predict the new location for the tracker $T_j$
      Get the new position $P_i$ of the target from the tracker T
      Update the tracker $T_j$ with new location
5. Plot the tracks for all M objects

## III. Experimental Results

The proposed method is tested on a variety of challenging video datasets: TUD Campus, TUD Crossing, TUD Stadtmitte and PETS2009 S2-L1 [17]. These datasets are very challenging for several reasons. First, they have walking pedestrians in an outdoor environment hence lighting conditions are perfect. Second, as it covers vast area, people get very small when they are far from the camera making their tracking more challenging (PETS2009 video). Then, in TUD dataset, targets have a similar size and they walk with similar speeds. However, targets are frequently occluding each other (heavy inter-object occlusion) and are occluded by static objects. To obtain the detections, we use the detections originally provided with the videos [17]. The details of each dataset are given in Table 1.

**Table 1** Details of dataset

| **Sequence** | **# frames** | **Persons** | **Resolution** |
|---|---|---|---|
| **TUD-CAMPUS** | 71 | Up to 6 | 640x480 |
| **TUD-CROSSING** | 201 | Up to 8 | 640x480 |
| **TUD-STADTMITTE** | 179 | Up to 8 | 640x480 |
| **PETS2009-S2-L1** | 795 | Up to 10 | 768x576 |

Two performance metrics are used to evaluate the performance of the proposed MOT system as used in [18]. They are explained as follows.

(a) **Multiple Object Tracking Accuracy (MOTA)**: It is one of the widely used evaluation metrics for multiple object tracking applications. It is defined as below.

$$MOTA = 1 - \frac{\sum_t \left(FP(t) + FN(t) + ID(t)\right)}{\sum_t N_{GT}(t)}. \qquad \text{---------------- (1)}$$

(b) **Multiple Object Tracking Precision (MOTP)**: The MOTP is defined as below.

$$MOTP = \sum_{t,i} \bar{d}\left(\mathcal{GT}_i^t, \mathcal{H}_{g(i)}^t\right) \Big/ \sum_t m_t, \qquad \text{--------------- (2)}$$

The proposed method is compared with recent state-of-the-art MOT algorithms. Among the compared approaches, a first category studied MOT with the aim of improving detection responses using model-free tracker [17,19], a second category aimed to ameliorate the data association technique [21-22], and a third category aimed to improve the appearance model [23-25]. The results are obtained from the authors' papers.

The experimental results are shown in Table 2. In general, for all the performance metrics, the proposed method outperforms other object trackers by achieving up to 85% of MOTA. Our MOTA are often higher than in the previous results. On PETS2009-S2-L1, TUD-Campus and TUD-Crossing, the proposed algorithm outperforms the tracking by detection method of Breitenstein et al. [22]. On the other hand, on TUDCampus and TUD-Crossing, we perform better than Riahi et al. method [28]. The proposed method also outperforms the tracking system proposed by [29]. On TUD-Stadtmitte and PETS2009-S2-L1, we achieved better MOTA than Segal et al. [23] MOT algorithm. It is possible to observe that our MOTA is higher than Dorra et al. [30] Furthermore, the proposed method is better than Yang et al. method [26] which includes background subtraction to handle occlusion.

**Table 2** Results comparison of the proposed method with other methods

| **Dataset** | **Method** | **MOTA (%)** | **MOTP (%)** | **False Negative (%)** | **False Positive (%)** |
|---|---|---|---|---|---|
| **TUD-CAMPUS** | Proposed | 78.7 | **70** | **0** | **10** |
| | Dorra, [2014] | 78.18 | 69 | 0 | 13 |
| | Riahi, [2014] | 72 | 74 | 25 | 2 |
| | Breitenstein, [2011] | 73 | 67 | 26 | 0.1 |
| **TUD-CROSSING** | Proposed | 79 | 68 | 0 | 7 |
| | Dorra, [2014] | 78 | 66 | 1 | 8 |
| | Riahi, [2014] | 72 | 76 | 26 | 1 |
| | Breitenstein, [2011] | 84 | 71 | 14 | 1 |
| | Andriyenko, [2011] | 63 | 75.5 | - | - |
| | Pirsiavash, [2011] | 63.3 | 76.3 | - | - |
| | Tang, [2014] | 70.7 | 77.1 | - | - |
| | Segal, [2013] | 74 | 76 | - | - |
| **TUD-STADTMITTE** | Proposed | 69 | 55 | 12 | 1 |
| | Dorra, [2014] | 67 | 57.26 | 26 | 6 |
| | Andriyenko, [2011] | 60.5 | 66 | - | - |

| | | | | | |
|---|---|---|---|---|---|
| | Milan, [2013] | 56.2 | 62 | - | - |
| | Segal, [2013] | 63% | 73 | - | - |
| | Milan, [2014] | 71 | 65.5 | - | - |
| | Andriyenko, [2012] | 61.8 | 63.2 | - | - |
| **PETS2009-S2-L1** | Proposed | 85 | 69 | 11 | 3 |
| | Dorra, [2014] | 84 | 66 | 13 | 2 |
| | Yang, [2009] | 76 | 54 | - | - |
| | Breitenstein, [2011] | 80 | 56 | - | - |
| | Andriyenko, [2011] | 80 | 76 | - | - |
| | Berclaz, [2006] | 60 | 66 | - | - |
| | Fuhr, [2014] | 70 | - | - | - |
| | Milan, [2014] | 90 | 80 | - | - |
| | Sherrah, [2013] | 81.3 | 74.4 | - | - |
| | Bae, [2014] | 80.34 | 69.72 | - | - |
| | Bae, [2014] | 83 | 69.59 | - | - |

## IV. Conclusion

In this paper, we proposed a novel object tracking method for video sequences. The proposed method is capable of tracking multiple objects under different challenging scenarios such as illumination variation, occlusion, and low resolution. We evaluated the proposed method using different challenging real world benchmarking video sequences. The performance of the system was evaluated using two accuracy metrics. The experimental results show that the proposed object tracking method is able to accurately track the targets except a case when some obstacles completely hide the objects.

## References


[1]. W.Luo, J.Xing, X.Zhang, X.Zhao and T.K. Kim,Multiple Object Tracking: A Literature Review,Computer Vision and Pattern Recognition, 2015.

[2]. B. Yang, C. Huang, and R. Nevatia, "Learning affinities and dependencies for multi-target tracking using a CRF model," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2011,pp. 1233–1240.

[3]. S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking,"in Proc. IEEE Int. Conf. Comput. Vis., 2009, pp. 261–268.

[4]. D. Koller, J. Weber, and J. Malik, "Robust multiple car tracking with occlusion reasoning," in Proc. Eur. Conf. Comput. Vis., 1994,pp. 189–196.

[5]. G. Duan, H. Ai, S. Cao, and S. Lao, "Group tracking: exploring mutual relations for multiple object tracking," in Proc. Eur. Conf.Comput. Vis., 2012, pp. 129–143.

[6]. T. Pfister, J. Charles, and A. Zisserman, "Flowing convnets for human pose estimation in videos," in Proc. IEEE Int. Conf. Comput.Vis., 2015, pp. 1913–1921.

[7]. W. Choi and S. Savarese, "A unified framework for multi-target tracking and collective activity recognition," in Proc. Eur. Conf.Comput. Vis., 2012, pp. 215–230.

[8]. W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," IEEE Trans. Syst.Man Cybern. Part C-Appl. Rev., vol. 34, no. 3, pp. 334–352, Mar.2004.

[9]. Jerome Berclaz, Francois Fleuret, EnginTuretken,Multiple Object Tracking Using K-Shortest Paths Optimization, IEEE transactions on pattern analysis andmachine intelligence. **33**(9) (2011) 1806-1819.

[10]. X.Zhou, Y.Li, B.He and T.Bai, GM-PHD-Based Multi-Target Visual Tracking Using Entropy Distribution and Game Theory, IEEE transactions on industrial informatics. **10**(2) (2014) 1064-1076.

[11]. X. H.Xia, W.Y. Nan, Z. Wei, Z. Jiang, Y. X.Fang, Multi object visual tracking based on reversible jump Markov chain Monte Carlo, IET Computer Vision. **5**(5) (2011)282–290.

[12]. W. Hu, W. Li, X.Zhang and S.Maybank, Single and Multiple Object Tracking Using a Multi-Feature Joint Sparse Representation, IEEE Transactions on Pattern Analysis and Machine Intelligence. **37**(4) (2014) 816 –833.

[13]. I.Ali and M.N. Dailey, Multiple human tracking in high density crowds, Image and Vision Computing. **30** (2012)966–977.

[14]. X.Liu, D. Tao, M.Song, L.Zhang, J.Bu, and C.Chen,Learning to Track Multiple Targets, IEEE transactions on neural networks and learning systems. **26**(5) (2014)1060 – 1073.

[15]. W.Hu, X.Li, W.Luo, X.Zhang, S.Maybank, andZ.Zhang, Single and Multiple Object Tracking Using Log-Euclidean Riemannian Subspace and Block-Division Appearance Model, IEEE transactions onpattern analysis and machine intelligence. **34**(12) (2012)2420-2440.

[16]. M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg, "Accurate scale estimation for robust visual tracking," in Proceedings of the BritishMachine Vision Conference BMVC, 2014.

[17]. A. Milan, K. Schindler, S. Roth, Detection-and trajectory-level exclusion in multiple object tracking, in: CVPR, IEEE, 2013, pp. 3682–3689.

[18]. Wu, Yi, Jongwoo Lim, and Ming-Hsuan Yang. "Online object tracking: A benchmark." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2411-2418. 2013.

[19]. M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, L. Van Gool, Online multiperson tracking-by-detection from a single, uncalibratedcamera, Pattern Analysis and Machine Intelligence(PAMI), IEEE Transactions on 33 (9) (2011) 1820–1833.

[20]. A. V. Segal, I. Reid, Latent data association: Bayesian model selectionfor multi-target tracking, in: ICCV, IEEE, 2013, pp. 2904–2911.

[21]. A. Andriyenko, K. Schindler, Multi-target trackingby continuous energy minimization, in: CVPR, IEEE, 2011, pp. 1265–1272.

[22]. J. Berclaz, F. Fleuret, P. Fua, Robust people tracking with global trajectory optimization, in: CVPR, Vol. 1, IEEE, 2006, pp. 744–750.

[23]. J. Yang, P. A. Vela, Z. Shi, J. Teizer, Probabilistic multiple people tracking through complex situations, in: 11th IEEE International Workshopon Performance Evaluation of Tracking and Surveillance, 2009.
[24]. G. Fu¨hr, C. R. Jung, Combining patch matching and detection for robust pedestrian tracking in monocular calibrated cameras, Pattern Recognition Letters 39 (2014) 11–20.
[25]. D. Riahi, G.-A. Bilodeau, Multiple feature fusion in the dempster-shafer framework for multi-object tracking, in: Computer and Robot Vision (CRV), IEEE, 2014, pp. 313–320.
[26]. H. Pirsiavash, D. Ramanan, C. C. Fowlkes, Globally-optimal greedy algorithms for tracking a variable number of objects, in: CVPR, IEEE, 2011, pp. 1201–1208.
[27]. A. Andriyenko, S. Roth, K. Schindler, Continuous energy minimization for multi-target tracking, IEEE TPAMI 35 (1) (2014) 1
[28]. A. Andriyenko, K. Schindler, S. Roth, Discrete-continuous optimization for multi-target tracking, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE, 2012, pp. 1926–1933.
[29]. J. Sherrah, B. Ristic, D. Kamenetsky, A pedestrian multiple hypothesis tracker fusing head and body detections, in: Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on, IEEE, 2013, pp. 1–8.
[30]. DorraRiahia, Guillaume-AlexandreBilodeaua, "Online multi-object tracking by detection based on generative appearance models", Computer Vision and Image Understanding, 2016

V. Ramalakshmi "Multiple Object Tracking System Using Discriminative Correlation Filter "International Journal of Engineering Science Invention (IJESI), vol. 07, no. 07, 2018, pp 27-31